A Practical Introduction to Differential Forms

Alexia E. Schulz

and

William C. Schulz

October 3, 2016

Transgalactic Publishing Company
Flagstaff, Vienna, Cosmopolis

Every creator painfully experiences the chasm between his inner vision and its ultimate expression. Isaac Bashevis Singer

Dedicated to our parents, children, and cats

Contents

Chapter 1

Applications to Differential Geometry

The Great Way is not difficult
If only you do not pick and choose.
Neither love nor hate
And you will clearly understand
Be off by a hair
And you are as far apart as heaven and earth
Seng Can(c. 590 CE)

1.1 Introduction

In this Chapter we wish to give an introduction to the ways differential forms can be used in Differential Geometry. There has been considereable neglect of this usage, and we discuss this below. There are many fine books on differential geometry and we do not want to plow this field again, so here we will concentrate on the important ideas that lie behind the use of differential forms. We will need to introduce the concepts of vector bundle and frame bundles but these are natural concepts, at least for the bundle of natural frames that comes from the coordinate system. We will mostly stick with frame bundles, but the theory extends with almost no change to any sort of bundle.

In classical differential geometry (which has certainly *not* been replaced by more modern treatments, although progress is being made), the principal compuational tool was tensor analysis. Tensor analysis is fine as a computational tool but poor as an aid to understanding what is really going on. For example, the concept *vector* is defined by saying an array of numbers (of the proper size) is a vector if it transforms in a certain way under coordinate change. This is not too illuminating. Then, when objects are differentiated, certain surprising quantities, the affine connections, are introduced so that the derivatives are not coordinate dependent. All of this can be made quite clear and straightforward with a little help from modern mathematical concepts.

Our basic goal is to explain the sentence

A connection is a Lie Algebra valued differential form

Once you understand what this really means, the whole of connections and curvature become accessible. This is important because curvature is the basic notion of differential geometry and one which has a fairly clear emotional meaning. When you control curvature, you control it all.

We do not want to suggest that tensors and their algebra and calculus are obsolete; we want to complement them by the use of differential forms, which often drastically simplify calculations. Tensors retain their importance as tools in places where differential forms are not appropriate, and even when differential forms are used tensors can be used a practical calculational adjuncts. The occasional tendency to regard tensors as obsolete is just picking and choosing. One needs both.

We will begin by developing some classical results for an n-manifold embedded in $(n+1)$-dimensional Euclidean space. From this we get a feeling of how the theory should go and we easily introduce connection coefficients and the Riemann curvature tensor, as well as the differential form equivalents. Then we move over to abstract manifolds and show how to construct analogs of the connections and curvature there, which is relatively easy. In fact, one of the reasons for doing this chapter is to show just how easy it is.

Next we will apply the previous material to 2-manifolds in 3-space and abstract 2-manifolds. Here we will introduce the most important invariant of a 2-manifold, the Gaussian Curvature and develop some formulas for computing it.

We will be staying almost entirely with *local* differential geometry but after the above material we will have a short chapter on the Gauss-Bonnet theorem which is one of the greatest theorems in mathematics, and the jumping off point for a vast amount of modern mathematics. The generalizations to n-manifolds, begun by Chern[1], uses tools which too advanced for us to discuss in a book of this nature, unfortunately, but the 2-manifold case is extremely impressive on its own.

We will develop the theory in a naive way first, so that you get to know the actors. At the end we will introduce the Lie Groups and Algebras to put the development in a modern context. These sections are a little more difficult and may be omitted by less enthusiastic students.

1.2 A Little History

When C. F. Gauss got the job of Professor of Astronomy at Göttingen U. he found out that it came with certain extra duties, including being responsible for surveying the Kingdom of Hannover. Gauss was not by nature a hiker and camper, and thus found it useful to think about the theory of surfaces partially to minimize the amount of actual measurement in the great outdoors that would be needed. He wrote a classic book on the subject, and this essentially began the subject as an independent discipline. Euler and others had already made some progress in applying Calculus to surfaces but Gauss pointed out the important concepts and proved many important theorems, thus creating a systematic body of knowledge from which new researchers could move forward sytematically. Another reason for Gauss's interest was non-Euclidean geometry, which could be done in a not totally satifying way on a curve called the pseudosphere, and differential geometry was helpful in these investigations.

One of Gauss's greatest theorems was that the Gaussian curvature depended only on the first fundamental form, or as we would say now the coefficiants g_{ij} of the inner product that the surface inherited from the surrounding, or ambient, 3-dimensional Euclidean space. Gauss realized that the g_{ij} could be determined by measurements *on the surface itself* and thus did not depend on how the surfaces was embedded in the Euclidean 3-space, and thus was an invariant of

[1]Professor Chern had the misfortune of transliterating his name in a Romanization of Chinese which failed to catch on. The r in Chern is a tone mark, indicating a rising tone; it is not pronounced at all. Chern thus rhymes with gun.

the surface itself. Riemann, with his usual surprising insight, reinterpreted this result to mean that you could have "surfaces" of any dimension, or as we would say, manifolds, and you could have curvature provided you gave the manifold an inner product at every point (continuously varying of course). For Gauss, the use of the surrounding three dimensional space of the surface was to provide this inner product, but for Riemann there was no need of any surrounding space since each point had a Riemann-given inner product. Riemann showed in his famous paper that this was enough to define curvature[2]. We also mention the probability that Riemann discovered his curvature tensor as an outgrowth of his researches on the integrability of first order partial differential equations, in which he was an expert.

Riemann's work opened up the field of Manifold theory, which was then carried on by the German Christoffel (1829-1900) and by Italians, who met Riemann on his trips south to a climate that was better for his health. Levi-Civita and others developed the tensor methods that allowed ordinary mathematicians, physicists, and engineers into this new world, in which eventually Cartan and Einstein became enthusiastic explorers.

Riemann presented his ideas on what became Riemannian Geometry in a lecture in 1854. Gauss was in the audience and, a very rare occurance, he was very impressed by Riemann's ideas. For unclear reasons, perhaps because he wanted to add more detail to the manuscript, Riemann had not published the manuscript when he died in 1866. Riemann was a very diffident man and did not make friends easily, but he was fortunate that the one friend he did make was Richard Dedekind, who arranged for the publication of the lecture in 1868, the most important paper in geometry in a thousand years. By 1869 Chistoffel had already introduced the Christoffel symbols and showed how to get the Riemann Curvature Tensor from them. An important book about Riemann's lecture and its influence has recently been published by Jürgen Jost.

However, the tensor methods introduced by the Italians, while natural for some people were very difficult for others, and there was a desire to have a more immediate approach to the material that was less off putting than the "forests of Indices" typical of tensor analysis. There are various methods, but differential forms, along with a bit of Lie Groups and Lie Algebras, are one entree into this world which is a little less offputting. This is what we are going to introduce you to. It is very easy in differential geometry to lose ones way and become distracted by all the beautiful objects to either side of the path. We are not going to do that (very much). We are going to march straight along the path as suggested by Seng Can's poem, and then it is not difficult. The method was pioneered by Elie Cartan in the first half of the 20th Century. However, enthusiasm has been less than one might anticipate, and one reason for this chapter has been to show how it can be done easily.

[2]It is possible that Gauss was actually aware of much of this, but Gauss disliked straining his contemporaries with radical new ideas. Riemann had no such reluctance.

1.3 Embedded n-manifolds in Euclidean $(n+1)$-space

1.3.1 Connection and Curvature Forms

We begin with an n-manifold embedded in an $n-1$-dimensional Euclindian space. We are interested in a theory which will describe the local properties of the manifold at a point. Specifically, we would like to be able to take derivatives of vector fields and to obtain expressions for the curvature. The manifold will be described locally by the coordinates $u^1, \dots u^n$ and the Euclidean space will have global coordinates $x^1, \dots x^n, x^{n+1}$. The x-coordinates of points on the manifold will thus be functions of the u coordinates

$$x^1 = x^1(u^1, \dots u^n), \ \ \dots, \ \ x^{n+1} = x^{n+1}(u^1, \dots u^n)$$

We can put these x-coordinates together into a position vector $\mathbf{R}$ for the surface,

$$\mathbf{R} = \langle x^1(u^1, \dots u^n), \ \ \dots, \ \ x^{n+1}(u^1, \dots u^n)\rangle$$

We assume the coordinates $u^1, \dots u^n$ are independent which means that the $(n+1) \times n$ matrix

$$\begin{pmatrix} \frac{\partial x^1}{\partial u^1} & \dots & \frac{\partial x^1}{\partial u^n} \\ \dots & \dots & \dots \\ \frac{\partial x^{n+1}}{\partial u^1} & \dots & \frac{\partial x^{n+1}}{\partial u^n} \end{pmatrix} \quad \text{has rank } n.$$

We can now form the vectors

$$\mathbf{e}_i = \frac{\partial \mathbf{R}}{\partial u^i}$$

These vectors are linearly independent in view of the requirement on the rank of the matrix above. We can then form the vector $\mathbf{n_1}$ by using the base vectors $\mathbf{i}_1, \dots, \mathbf{i}_{n+1}$ of the Euclidean $(n+1)$-space:

$$\mathbf{n}_1 = \begin{vmatrix} \frac{\partial x^1}{\partial u^1} & \dots & \frac{\partial x^1}{\partial u^n} & \mathbf{i}_1 \\ \dots & \dots & \dots & \dots \\ \frac{\partial x^{n+1}}{\partial u^1} & \dots & \frac{\partial x^{n+1}}{\partial u^n} & \mathbf{i}_{n+1} \end{vmatrix}$$

Now $\mathbf{n}_1$ is perpedicular to each of the $\mathbf{e}_i$ because the inner product of $\mathbf{e}_i$ can be written as a determinant

$$\begin{vmatrix} \frac{\partial x^1}{\partial u^1} & \dots & \frac{\partial x^1}{\partial u^n} & \frac{\partial x^1}{\partial u^i} \\ \dots & \dots & \dots & \dots \\ \frac{\partial x^{n+1}}{\partial u^1} & \dots & \frac{\partial x^{n+1}}{\partial u^n} & \frac{\partial x^{n+1}}{\partial u^i} \end{vmatrix}$$

which, since it has a repeated column, is equal to 0. We now set

$$\mathbf{n} = \frac{\mathbf{n}_1}{||\mathbf{n}_1||}$$

and this is the unit normal vector for the manifold. Because of the position of the basis vectors in the definition of $\mathbf{n}_1$ the vectors

$$\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{n}$$

form a basis for $\mathbb{R}^{n+1}$ which is whose orientation is the same as that of $\mathbf{i}_1, \ldots, \mathbf{i}_{n+1}$, as is easily seen (check for $n = 1$ or $n = 2$). The $n = 1$ case will convince you that the column of base vectors needs to go at the *end* rather than at the beginning, where we put it for ordinary vector algebra.

If we don't like the way the vector $\mathbf{n}$ is pointing, for example if we want the *exterior* normal for a closed surface, it is only necessary to renumber the parameters u^i.

The vectors $\mathbf{e}_i$ are tangent vectors to the coordinate lines for the coordinates u_i. This is quite possibly familiar to you from advanced calculus courses. The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ form a basis of the tangent space $T_p(M)$ at each $p \in M$.

Although it is not of much use to us in our circumstances, we note that if M happens to be (even locally) the zero set of some function f, then we can get a normal vector $\mathbf{n}_1$ by using the gradient

$$\mathbf{n}_1 = \nabla f = \Big(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^{n+1}}\Big)$$

Try it with the 2-sphere in 3-space with $f(x, y, z) = x^2 + y^2 + z^2$.

We wish to take the derivative of tangent vectors $v = \mathbf{e}_i v^i$ and to do this we need first to take the derivative of $\mathbf{e}_i$. Since $\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{n}$ is a basis of $\mathbb{R}^{n+1}$ we have

$$\frac{\partial \mathbf{e}_i}{\partial u^j} = \mathbf{e}_k \Gamma^k_{ij} + \mathbf{n} b_{ij}$$

for some coefficients Γ^k_{ij} and b_{ij}. For reasons to be explained later the Γ^k_{ij} are called *connection coefficients*. There is a symmetry here of the indices

$$\begin{aligned} \frac{\partial \mathbf{e}_i}{\partial u^j} &= \frac{\partial}{\partial u^j}\frac{\partial \mathbf{R}}{\partial u^i} \\ &= \frac{\partial}{\partial u^i}\frac{\partial \mathbf{R}}{\partial u^j} \\ &= \frac{\partial \mathbf{e}_j}{\partial u^i} \end{aligned}$$

so

$$\mathbf{e}_k \Gamma^k_{ij} + \mathbf{n} b_{ij} = \mathbf{e}_k \Gamma^k_{ji} + \mathbf{n} b_{ji}$$

which gives the symmetry conditions

$$\Gamma^k_{ij} = \Gamma^k_{ji} \qquad\qquad b_{ij} = b_{ji}$$

At this point we must introduce an idea of great importance. The basis $\mathbf{e}_k$ used to prove the last equation is derived from the position vector $\mathbf{R}$ of the manifold via $\mathbf{e}_k = \frac{\partial \mathbf{R}}{\partial u^k}$ and this particular choice of a basis of the tangent space $T(M)$ is called the *natural basis*. Naturally there is a natural basis for each coordinate system. However, for many purposes limitation to the use of only natural bases is too restrictive. We will now widen the applicability of our theory by removing this restriction. Henceforth, when it is not explicitly stated that the basis is natural, we will assume it is general. This means that at each point $p \in M$ the set of vectors $\{\mathbf{e}_1(p), \mathbf{e}_2(p), \ldots, \mathbf{e}_n(p)\}$ will be assumed to be a basis for the tangent space $T_p(M)$ at p. Naturally we assume that the $\mathbf{e}_k(p)$ are smooth vector functions of the coordinates u^i. If $\tilde{\mathbf{e}}_l$ is the natural basis and $\mathbf{e}_k$ is our general basis there will be functions α_l^k on the manifold so that $\mathbf{e}_k = \tilde{\mathbf{e}}_l \, \alpha_l^k$ but with a little luck we will not need to use this. The matrix (α_l^k) is of course invertible.

Notice, and this is important, that the symmetry $\Gamma_{ij}^k = \Gamma_{ji}^k$ is true for the natural basis but there is no reason to think it is true for a general basis. Hence when using this symmetry be sure you are using a natural basis or that you have proved the symmetry in whatever situation you are in by some other means. Remember *symmetry of the Γ_{ij}^k is a very special circumstance and not something to be expected in general*[3]. However, the equation

$$\boxed{d\mathbf{e}_i = \frac{\partial \mathbf{e}_i}{\partial u^j}\, du^j = \mathbf{e}_k \Gamma_{ij}^k \, du^j + \mathbf{n} b_{ij} \, du^j}$$

remains just as true for general bases as for natural bases.

Now some philosophy. Quantities in differential geometry are of two types. *intrinsic quantities* are quantities that depend only on the surface itself and not on how it happens to lie in $\mathbb{R}^{n+1}$ and non-intrinsic quantities. We consider the g_{ij} to be intrinsic, which sounds odd since as we have set things up it looks like M inherits the g_{ij} from the surrounding $\mathbb{R}^{n+1}$. The way to understand this is to think of the surface as having an inner product at each point and then being *mapped* into $\mathbb{R}^{n+1}$ in such a way that the inner product it came with and the inner product it inherits from the surrounding $\mathbb{R}^{n+1}$ coincide. After you read the next section this may seem clearer. For a non-intrinsic quantity, the unit normal vector $\mathbf{n}$ is very representative, and the b_{ij} are also naturally non-intrinsic. We will eventually show that the connection coefficients Γ_{ij}^k are intrinsic but this is by no means obvious. The standard way to show something is intrinsic is to show it depends on the ur-intrinsic quantities g_{ij}, the coefficients of the inner product, which, I emphasize again, are thought of as provided with the surface.

We now want to change the game in an essential way. We are not interested very much in the non-intrinsic aspects of the situation, so we would like the derivative of a tangent vector $\mathbf{v}$ to be again a tangent vector, and the differential of a tangent vector to be an element of $T^*(M)$. These things are not true as things stand, and we must modify the definitions to make them true. We do this by creating an alternative to $d\mathbf{v}$ by orthogonally projecting $d\mathbf{v}$ onto $T^*(M)$, which means along $\mathbf{n}$. Because we saw this coming we set things up so this is a

[3]Einstein played with the idea of finding a place for the electromagnetic field in an asymmetry of the Γ_{ij}^k and thus having a unified field theory for gravity and electromagnetics. It didn't work, and we would not expect it to now because of the existence of strong force fields and weak force fields

trivial task; the covariant differential and covariant derivative are

$$\boxed{D\mathbf{e}_i = \tfrac{D\mathbf{e}_i}{\partial u^j}\, du^j \;=\; \mathbf{e}_k \Gamma^k_{ij}\, du^j}$$

$$\frac{D\mathbf{e}_i}{\partial u^j} = \mathbf{e}_k \Gamma^k_{ij}$$

We are not too interested in the second equation, but the first is critical for all that follows.

We need the operator D in a couple of other circumstances, and we define it

$$\begin{aligned} Df &= d\,f \qquad \text{for functions f} \\ D(\mathbf{e}_i \omega) &= \mathbf{e}_i\, d\omega + (D\mathbf{e}_i) \wedge \omega \quad \text{for differential forms } \omega \end{aligned}$$

(We are gliding over some technicalities here with tensor products $\otimes$ but it would just complicate the notation for very small payoff.) We emphasize that keeping the order consistent is very important here. We can now, using the above formula, find the formula for the differential (in the new sense,) of a vector

$$\begin{aligned} D\mathbf{v} = D(\mathbf{e}_i v^i) &= \mathbf{e}_i\, dv^i + (D\mathbf{e}_i)\, v^i \\ &= \mathbf{e}_i \frac{\partial v^i}{\partial u^j} du^j + \mathbf{e}_k \Gamma^k_{ij}\, du^j\, v^i \\ &= \mathbf{e}_k \Big(\frac{\partial v^k}{\partial u^j} + \Gamma^k_{ij}\, v^i \Big)\, du^j \\ &= \mathbf{e}_k\, v^k{}_{|j}\, du^j \end{aligned}$$

where

$$v^k{}_{|j} = \frac{\partial v^k}{\partial u^j} + \Gamma^k_{ij}\, v^i$$

is the classical notation for the covariant derivative in coordinate form. This will help orient you if you have previous experience with differential geometry or general relativity.

Now that we have $D\mathbf{v}$ what could be more natural than to take another coveriant differential. It is cosmically significant that while $d^2 = 0$, this is not generally true for D and it turns out that D^2 churns up one of the most important things in differential geometry.

$$\begin{aligned} D^2\mathbf{e}_i &= D(D\mathbf{e}_i) \\ &= D\big(\mathbf{e}_k \Gamma^k_{ij} du^j\big) \\ &= \mathbf{e}_k\, d(\Gamma^k_{ij}\, du^j) + (D\mathbf{e}_k)\Gamma^k_{ij}\, du^j \\ &= \mathbf{e}_\ell \frac{\partial \Gamma^\ell_{ij}}{\partial u^m}\, du^m \wedge du^j + \mathbf{e}_\ell \Gamma^\ell_{km} du^m \wedge \Gamma^k_{ij}\, du^j \\ &= \mathbf{e}_\ell \Big(\frac{\partial \Gamma^\ell_{ij}}{\partial u^m} + \Gamma^\ell_{km} \Gamma^k_{ij} \Big) du^m \wedge du^j \end{aligned}$$

Now I rewrite the previous expression interchanging the summation indices m and j to get

$$\begin{aligned} D^2\mathbf{e}_i &= \mathbf{e}_\ell\Big(\frac{\partial\Gamma^\ell_{im}}{\partial u^j} + \Gamma^\ell_{kj}\Gamma^k_{im}\Big)du^j\wedge du^m \\ &= -\mathbf{e}_\ell\Big(\frac{\partial\Gamma^\ell_{im}}{\partial u^j} + \Gamma^\ell_{kj}\Gamma^k_{im}\Big)du^m\wedge du^j \end{aligned}$$

The reason for this sleight of hand is that since $D^2\mathbf{e}_i$ is a vector valued 2-form and we want it's coefficients to be skew symmetric in m and j. Now we have two different expressions for $D^2\mathbf{e}_i$ and we can get still another expression by adding the two and dividing by 2. The result is then

$$\begin{aligned} D^2\mathbf{e}_i &= \frac{1}{2}\,\mathbf{e}_\ell\Big(\frac{\partial\Gamma^\ell_{ij}}{\partial u^m} - \frac{\partial\Gamma^\ell_{im}}{\partial u^j} + \Gamma^\ell_{km}\Gamma^k_{ij} - \Gamma^\ell_{kj}\Gamma^k_{im}\Big)\,du^m\wedge du^j \\ &= \mathbf{e}_\ell\frac{1}{2}\,\Omega_i{}^\ell \end{aligned}$$

where

$$\Omega_i{}^\ell = R_i{}^\ell{}_{mj}\,du^m\wedge du^j$$

is the *curvature 2-form* and

$$R_i{}^\ell{}_{mj} = \frac{\partial\Gamma^\ell_{ij}}{\partial u^m} - \frac{\partial\Gamma^\ell_{im}}{\partial u^j} + \Gamma^\ell_{km}\Gamma^k_{ij} - \Gamma^\ell_{kj}\Gamma^k_{im}$$

is the famous Riemann Curvature Tensor. Note that it is skew symmetric in m and j. Here is is again with more standard letters[4].

$$R_j{}^i{}_{kl} = \frac{\partial\Gamma^i_{jl}}{\partial u^k} - \frac{\partial\Gamma^i_{jk}}{\partial u^l} + \Gamma^i_{mk}\Gamma^m_{jl} - \Gamma^i_{ml}\Gamma^m_{jk}$$

Now there is nothing wrong with this derivation and we have presented it in this way to make sure you see how it works. However, it is relatively inefficient, with far more writing than necessary. We now want to present it in a more convenient form which takes far less writing. Recall that

$$\begin{aligned} D\mathbf{e}_i &= \mathbf{e}_k\Gamma^k_{ij}\,du^j \\ &= \mathbf{e}_k\,\omega_i{}^k \end{aligned}$$

where $(\omega_i{}^k)$ is the $n\times n$ matrix of 1-forms

$$(\omega_i{}^k) = (\Gamma^k_{ij}\,du^j)$$

Writing things this way, and remembering the old formula

$$d(\omega\wedge\eta) = (d\omega)\wedge\eta + (-1)^{\deg(\eta)}\omega\wedge d\eta$$

[4]The sign of the Riemann Curature Tensor is not a matter of universal agreement, and some books define it to be the negative of the definition given here. It is wise always to check how the reference you are using defines it. The position of the upper index also varies. Somewhat oddly, Riemann does not seem to have written it down anywhere.

which now becomes critically important, we have

$$\begin{aligned} D^2\mathbf{e}_i &= D(\mathbf{e}_k\omega_i{}^k) \\ &= \mathbf{e}_k\,d\omega_i{}^k + D(\mathbf{e}_k)\wedge\omega_i{}^k \\ &= \mathbf{e}_l\,d\omega_i{}^l + \mathbf{e}_l\,\omega_k{}^l\wedge\omega_i{}^k \\ &= \mathbf{e}_l\left(d\omega_i{}^l + \omega_k{}^l\wedge\omega_i{}^k\right) \end{aligned}$$

Pretending that we haven't already defined Ω it would be natural at this point to define

$$\Omega_i{}^l = d\omega_i{}^l + \omega_k{}^l\wedge\omega_i{}^k$$

Let's check that this is the same as our previous result.

$$\begin{aligned} \Omega_i{}^l &= d(\Gamma^l_{ij}\,du^j) + (\Gamma^l_{km}\,du^m)\wedge(\Gamma^k_{ij}\,du^j) \\ &= \left(\frac{\partial\Gamma^l_{ij}}{\partial u^m} + \Gamma^l_{km}\Gamma^k_{ij}\right)du^m\wedge du^j \end{aligned}$$

which is exactly the previous result. Notice however how much more elegant and efficient the new method is.

But if you REALLY want elegance, watch this. We set

$$\begin{aligned} \vec{\mathbf{e}} &= (\mathbf{e}_1,\ldots,\mathbf{e}_n) && \text{row vector} \\ \omega &= (\omega_i{}^l) && n\times n \text{ matrix of 1-forms} \\ \Omega &= (\Omega_i{}^l) && n\times n \text{ matrix of 2-forms} \end{aligned}$$

Then

$$\begin{aligned} D\vec{\mathbf{e}} &= \vec{\mathbf{e}}\,\omega \\ D^2\vec{\mathbf{e}} &= D(\vec{\mathbf{e}}\,\omega) \\ &= \vec{\mathbf{e}}\,d\omega + \vec{\mathbf{e}}\,\omega\wedge\omega \\ &= \vec{\mathbf{e}}\,(d\omega + \omega\wedge\omega) \\ &= \vec{\mathbf{e}}\,\Omega \end{aligned}$$

and we see, with implied matrix multiplication

$$\Omega = d\omega + \omega\wedge\omega$$

Note that $\omega\wedge\omega \neq 0$ because it is not multiplication of 1-forms but multiplication of *matrices* of 1-forms, and so need not be 0.

Here's another example; set

$$\vec{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

and then

$$\begin{aligned} Dv &= D(\mathbf{e}\vec{v}) \\ &= \mathbf{e}\, d\vec{v} + (\mathbf{e}\omega)\vec{v}) \\ &= \mathbf{e}\left(d\vec{v} + \omega\,\vec{v}\right) \end{aligned}$$

If we decode $d\vec{v} + \omega\,\vec{v}$ we get for the i^{th} component

$$\begin{aligned} dv^i + \Gamma^i_{jk} v^j\, du^k &= \left(\frac{\partial v^i}{\partial u^k} + \Gamma^i_{jk} v^j\right) du^k \\ &= v^i{}_{|k}, du^k \end{aligned}$$

which is as it should be.

For the fun of it we should find $d\Omega$. We use $d\omega = \Omega - \omega \wedge \omega$.

$$\begin{aligned} d\Omega &= d(d\omega + \omega \wedge \omega) \\ &= 0 + (d\omega) \wedge \omega + (-1)^{\deg(\omega)} \omega \wedge d\omega \\ &= (\Omega - \omega \wedge \omega) \wedge \omega - \omega \wedge (\Omega - \omega \wedge \omega) \\ &= \Omega \wedge \omega - \omega \wedge \Omega \\ &= [\Omega, \omega] \end{aligned}$$

where we adopt the usual notation $[A, B] = AB - BA$ and we keep in mind that the AB and BA are matrix multiplications.

Mixed Covariant Derivatives

We want to find a connection between mixed covariant derivatives and the Riemann Curvature Tensor which is a little surprising and very important. Recall that for ordinary derivatives we have $\partial^2 f/(\partial u^i \partial u^j) = \partial^2 f/(\partial u^j \partial u^i)$ but this is not true for covariant derivatives and we would like to know just how much it fails. This is a simple calculation. Recall the covariant differential

$$D\mathbf{v} = \mathbf{e}_i v^i{}_{|k} du^k \quad \text{where} \quad v^i{}_{|k} = \frac{\partial v^i}{\partial u^k} + \Gamma^i_{jk} v^j$$

and we have the covariant partial derivative notation

$$\frac{D\mathbf{v}}{\partial u^k} = \mathbf{e}_i\, v^i{}_{|k} = \mathbf{e}_i \left(\frac{\partial v^i}{\partial u^k} + \Gamma^i_{jk} v^j\right)$$

Now we take another covariant derivative

$$\begin{aligned} \frac{D^2\mathbf{v}}{\partial u^l \partial u^k} &= \frac{D}{\partial u^l}(\mathbf{e}_i v^i{}_{|k}) \;=\; \mathbf{e}_i \frac{\partial v^i{}_{|k}}{\partial u^l} + \left(\frac{D}{\partial u^l}\,\mathbf{e_i}\right) v^i{}_{|k} \\ &= \mathbf{e}_m \frac{\partial v^m_{|k}}{\partial u^l} + (\mathbf{e}_m \Gamma^m_{il})\; v^i{}_{|k} \end{aligned}$$

$$
\begin{aligned}
&= \mathbf{e}_m \left(\frac{\partial v^m{}_{|k}}{\partial u^l} + \Gamma^m_{il}\, v^i{}_{|k} \right) \\
&= \mathbf{e}_m \left(\frac{\partial^2 v^m}{\partial u^l \partial u^k} + \frac{\partial \Gamma^m_{ik}}{\partial u^l}\, v^i + \Gamma^m_{ik} \frac{\partial v^i}{\partial u^l} + \Gamma^m_{il} \left(\frac{\partial v^i}{\partial u^k} + \Gamma^i_{jk} v^j \right) \right) \\
&= \mathbf{e}_m \left(\frac{\partial^2 v^m}{\partial u^l \partial u^k} + \left(\frac{\partial \Gamma^m_{jk}}{\partial u^l} + \Gamma^m_{il} \Gamma^i_{jk} \right) v^j + \left(\Gamma^m_{ik} \frac{\partial v^i}{\partial u^l} + \Gamma^m_{il} \frac{\partial v^i}{\partial u^k} \right) \right)
\end{aligned}
$$

From this we easily get, leaving out the terms symmetric in k and l,

$$
\begin{aligned}
\frac{D^2 \mathbf{v}}{\partial u^l \partial u^k} - \frac{D^2 \mathbf{v}}{\partial u^k \partial u^l} &= \mathbf{e}_m \left(\frac{\partial \Gamma^m_{jk}}{\partial u^l} - \frac{\partial \Gamma^m_{jl}}{\partial u^k} + \Gamma^m_{il} \Gamma^i_{jk} - \Gamma^m_{ik} \Gamma^i_{jl} \right) v^j \\
&= \mathbf{e}_m R_j{}^m{}_{lk}\, v^j
\end{aligned}
$$

which can also be written

$$
v^m{}_{|k|l} - v^m{}_{|l|k} = R_j{}^m{}_{lk}\, v^j
$$

It is interesting and important that the final formula has no derivatives of the v^i in it. This is related to the fact that $R_j{}^m{}_{lk}$ is a tensor.

Warning: The placement of indices in the Riemann Curvature Tensor $R_j{}^m{}_{lk}$ varies from book to book. Hence you should not expect, on consulting a different text, to find the indices lined up as we have done it here. This is annoying but unavoidable. One can usually see a book's convention by checking the definition of $R_j{}^m{}_{lk}$ in terms of the Christoffel symbols.

Computation of the Γ^i_{jk}

In this section it is critical that $\Gamma^k_{ij} = \Gamma^k_{ji}$. This can be arranged by using a natural basis. The formulas given here are *not valid for a general basis* unless it happens that the Christoffel symbols are symmetric: $\Gamma^k_{ij} = \Gamma^k_{ji}$.

Recall that

$$
\frac{\partial \mathbf{e}_i}{\partial u^j} = \mathbf{e}_k\, \Gamma^k_{ij} + n\, b_{ij}
$$

You may have wondered just how in a particular case you compute Γ^k_{ij} and we are going to work that out now. It is easy but intricate. The first thing is to take the inner product with e_l to get

$$
\begin{aligned}
\left(\frac{\partial \mathbf{e}_i}{\partial u^j}, \mathbf{e}_l \right) &= (\mathbf{e}_k\, \Gamma^k_{ij}, \mathbf{e}_l) + b_{ij}(\mathbf{n}, \mathbf{e}_l) \\
&= g_{kl}\, \Gamma^k_{ij} + 0 \\
&= \Gamma_{ij;\, l}
\end{aligned}
$$

where for convenience we have introduced the abbreviation

$$
\Gamma_{ij;\, l} = g_{kl}\, \Gamma^k_{ij}
$$

We use this formula to coupute $\partial g_{ij}/\partial u^k$ as follows. The second and third formulas are derived by applying the cyclic permuation $i \to j$, $j \to k$, $k \to i$ twice. From

$$g_{ij} = (\mathbf{e}_i,\, \mathbf{e}_j)$$

we get

$$\begin{aligned}
\frac{\partial g_{ij}}{\partial u^k} &= \Big(\frac{\partial \mathbf{e}_i}{\partial u^k},\, \mathbf{e}_j\Big) + \Big(\mathbf{e}_i,\, \frac{\partial \mathbf{e}_j}{\partial u^k}\Big) \;=\; \Gamma_{ik;\,j} + \Gamma_{jk;\,i} \\
\frac{\partial g_{jk}}{\partial u^i} &= \Gamma_{ji;\,k} + \Gamma_{ki;\,j} \\
\frac{\partial g_{ki}}{\partial u^j} &= \Gamma_{kj;\,i} + \Gamma_{ij;\,k}
\end{aligned}$$

Now we add the last two equations and subtract the one before them. Up to this point all we have done is valid for a general basis. However, at this point we require the symmetry of the Christoffel symbols: $\Gamma^k_{ij} = \Gamma^k_{ij}$ which we would have, for example, for a natural basis. This transfers immediately to $\Gamma_{jk;i} = \Gamma_{jk;i}$ so we get

$$\begin{aligned}
\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k} &= \Gamma_{ji;\,k} + \Gamma_{ki;\,j} + \Gamma_{kj;\,i} + \Gamma_{ij;\,k} - \Gamma_{ik;\,j} - \Gamma_{jk;\,i} \\
&= \Gamma_{ji;\,k} + \Gamma_{ij;\,k} + \Gamma_{ki;\,j} - \Gamma_{ik;\,j} + \Gamma_{kj;\,i} - \Gamma_{jk;\,i} \\
&= 2\Gamma_{ij;\,k}
\end{aligned}$$

Then we have

$$\begin{aligned}
g^{km}\Gamma_{ij;\,k} &= \frac{1}{2} g^{km}\Big(\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k}\Big) \\
g^{km} g_{kl}\Gamma^l_{ij} &= \frac{g^{km}}{2}\Big(\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k}\Big) \\
\delta^m_l \Gamma^l_{ij} &= \frac{g^{km}}{2}\Big(\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k}\Big)
\end{aligned}$$

$$*** \qquad \Gamma^m_{ij} = \frac{g^{km}}{2}\Big(\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k}\Big) \qquad \text{Natural Basis} \qquad ***$$

which is the required formula for the Γ^i_{jk} in terms of the g_{ij}. This formula was derived from the $\mathbf{e}_i = \frac{\partial \mathbf{R}}{\partial u^i}$ using $g_{ij} = (\mathbf{e}_i, \mathbf{e}_j)$. It is critical that the coordinates u^i in the formula and the g_{ij} are connected in this way, so don't go using this formula with g_{ij} that come from some other basis, or from somewhere else.

There is another formula for the Γ^I_{ij} which can be useful when many of the g_{ij} are 0, for example when the coordinates are orthogonal. Recalling the first formula, we derive

$$\frac{\partial \mathbf{e}_i}{\partial u^j} = \mathbf{e}_k \Gamma^k_{ij} + \mathbf{n} b_{ij}$$

$$
\begin{aligned}
(\mathbf{e}_l,\, \frac{\partial \mathbf{e}_i}{\partial u^j}) &= g_{kl}\Gamma^k_{ij} \\
g^{lm}(\mathbf{e}_l,\, \frac{\partial \mathbf{e}_i}{\partial u^j}) &= \Gamma^k_{ij} \;=\; \delta^m_k \Gamma^k_{ig^{lm}g_{kl}j} = \Gamma^m_{ij}
\end{aligned}
$$

so

$$\boxed{\Gamma^k_{ij} = g^{kl}(\mathbf{e}_l,\, \tfrac{\partial \mathbf{e}_i}{\partial u^j})}$$

Now we wish to derive an interesting formula which will be important when we get to Riemannian Geometry. We return at this point to a general basis; the material used in the derivation does not depend on the symmetry of the Christoffel symbols. We want a formula for $d(\mathbf{v}, \mathbf{w})$ for $\mathbf{v}, \mathbf{w} \in T_p(M)$. Our approach is very direct.

$$
\begin{aligned}
\frac{\partial}{\partial u^k}(\mathbf{v}, \mathbf{w}) &= \frac{\partial}{\partial u^k}\big(g_{ij}v^i w^j\big) \\
&= \Big(\frac{\partial g_{ij}}{\partial u^k}\Big)v^i w^j + g_{ij}\Big(\frac{\partial v^i}{\partial u^k}\Big)w^j + g_{ij}v^i\Big(\frac{\partial w^j}{\partial u^k}\Big) \\
&= (\Gamma_{ik;j} + \Gamma_{jk;i})\, v^i w^j + g_{ij}\Big(\frac{\partial v^i}{\partial u^k}\Big)w^j + g_{ij}v^i\Big(\frac{\partial w^j}{\partial u^k}\Big) \\
&= g_{lj}\Big(\Gamma^l_{ik}\, v^i w^j + \frac{\partial v^l}{\partial u^k}\Big)w^j + g_{il}v^i\Big(\Gamma^l_{jk}\, v^i w^j + \frac{\partial w^l}{\partial u^k}\Big) \\
&= \Big(\frac{D\mathbf{v}}{\partial u^k},\, \mathbf{w}\Big) + \Big(\mathbf{v},\, \frac{D\mathbf{w}}{\partial u^k}\Big) \\
d(\mathbf{v}, \mathbf{w}) &= \frac{\partial}{\partial u^k}\big(g_{ij}v^i w^j\big)\, du^k \\
&= \Big(\frac{D\mathbf{v}}{\partial u^k},\, \mathbf{w}\Big)\, du^k + \Big(\mathbf{v},\, \frac{D\mathbf{w}}{\partial u^k}\Big)\, du^k \\
&= \Big(\frac{D\mathbf{v}}{\partial u^k}\, du^k,\, \mathbf{w}\Big) + \Big(\mathbf{v},\, \frac{D\mathbf{w}}{\partial u^k}\, du^k\Big) \\
&= (D\mathbf{v},\, \mathbf{w}) + (\mathbf{v},\, D\mathbf{w})
\end{aligned}
$$

Some additional formulas

This subsection deals with the derivation of some formulas which we need for proving the Theorema Egregium of Gauss. *Once again we are deriving formulas for general bases.* These formulas have various uses but are not as important as the previous material. Essentially they show that the Christoffel symbols Γ^i_{jk} are not completely independent from the b_{ij} and the theorema egregium comes out of the dependence. We are going to prove these formulas for n-manifolds embedded in $\mathbb{R}^{n+1}$ although we will get the most use of them for the case of 2-manifolds embedded in $\mathbb{R}^3$.

There are a few preliminaries to the proof. First we consider what we can deduce from $A_{ij}\, du^i \wedge du^j = 0$. Most emphatically this does not mean that $A_{ij} = 0$, because the objects $du^i \wedge du^j$ are not linearly independent. We can get a linearly independent set by taking only the objects $du^i \wedge du^j$ with $i < j$. Let us look at how this works.

$$\begin{aligned} \omega &= A_{ij}\, du^i \wedge du^j \\ &= A_{ji}\, du^j \wedge du^i \qquad \text{Interchanging } i \text{ and } j \\ &= -A_{ji}\, du^i \wedge du^j \end{aligned}$$

Adding the two expressions for ω we have

$$2\omega = (A_{ij} - A_{ji})\, du^i \wedge du^j$$

Notice now that $A_{ij} - A_{ji}$ is antisymmetric in i and j, as is $du^i \wedge du^j$. Hence, switching i and j, we have

$$(A_{ji} - A_{ji})\, du^j \wedge du^i = (A_{ij} - A_{ji})\, du^i \wedge du^j$$

since the two minus signs cancel. Thus if we take only $i < j$ terms we will have only half as many terms as the full i and j and thus

$$\begin{aligned} 2\omega &= 2\sum_{i<j}(A_{ij} - A_{ji})\, du^i \wedge du^j \\ \omega &= \sum_{i<j}(A_{ij} - A_{ji})\, du^i \wedge du^j \end{aligned}$$

This is often called antisymmetrizing the coefficients. Now because the $du^i \wedge du^j$ with $i < j$ form a linearly independent set, we can say that

$$\text{If} \quad \omega = A_{ij}\, du^i \wedge du^j = 0 \quad \text{then} \quad A_{ij} - A_{ji} = 0$$

We will use this several times in what follows. Clearly this could be generalized to higher order forms but it would look quite complicated.

Next we need formulas for $d\mathbf{n} = \frac{\partial\, \mathbf{n}}{\partial u^j}\, du^j$. Since $(\mathbf{n}, \mathbf{n}) = 1$, we have

$$\Big(\frac{\partial \mathbf{n}}{\partial u^j}, \mathbf{n}\Big) = 0$$

so $\frac{\partial\, \mathbf{n}}{\partial u^i}$ is a linear combination of the $\mathbf{e}_k$. Since $(\mathbf{n}, \mathbf{e_j}) = 0$ we have

$$\Big(\frac{\partial\, \mathbf{n}}{\partial u^j}, \mathbf{e}_i\Big) = -\Big(\mathbf{n}, \frac{\partial \mathbf{e}_i}{\partial u^j}\Big) = -b_{ij}$$

If $\frac{\partial\, \mathbf{n}}{\partial u^j} = \mathbf{e}_k c^k_j$ then we have

$$\begin{aligned} (\mathbf{e}_k c^k_j, \mathbf{e}_i) &= -b_{ij} \\ g_{ki} c^k_j &= -b_{ij} \\ g^{li} g_{ki} c^k_j &= -g^{li}\, b_{ij} \overset{\text{def}}{=} -b^l_j \\ c^l_j = \delta^l_k c^k_j &= -b^l_j \end{aligned}$$

so we have

$$\frac{\partial\,\mathbf{n}}{\partial u^j} = -\mathbf{e}_k b^k_j = -\mathbf{e}_k g^{ki} b_{ij}$$

Recall that when we calculated $D^2\mathbf{e}_i$ it churned up the curvature tensor. We are now going do the same thing with $d^2\mathbf{e}_i$ which will naturally be 0. We recall that

$$d\mathbf{e}_i = \mathbf{e}_k \Gamma^k_{ij}\, du^j + \mathbf{n} b_{ij}\, du^j$$

We recall that $d(\mathbf{e}_i\omega) = \mathbf{e}_i \wedge d\omega + (d\mathbf{e}_i)\wedge\omega$, so the calculation runs

$$\begin{aligned}
0 &= d^2\mathbf{e}_i \;=\; d(\mathbf{e}_k\Gamma^k_{ij}\, du^j + \mathbf{n}\, b_{ij})\, du^j \\
&= \mathbf{e}_k d(\Gamma^k_{ij}\, du^j) + (d\mathbf{e}_k)\Gamma^k_{ij}\, du^j + \mathbf{n}\, d(b_{ij}\, du^j) + (d\mathbf{n})\wedge b_{ij}\, du^j \\
&= \mathbf{e}_k\Big(\frac{\partial\Gamma^k_{ij}}{\partial u^m}\Big)\, du^m\wedge du^j + \Big(\mathbf{e}_l\Gamma^l_{km} du^m + \mathbf{n} b_{km} du^m\Big)\wedge\Gamma^k_{ij}\, du^j \\
&\quad + \mathbf{n}\,\frac{\partial b_{ij}}{\partial u^m}\, du^m\wedge du^j + \frac{\partial\mathbf{n}}{\partial u^m}\, du^m\wedge b_{ij}\, du^j \\
&= \mathbf{e}_l\Big(\frac{\partial\Gamma^l_{ij}}{\partial u^m} + \Gamma^l_{km}\Gamma^k_{ij}\Big)\, du^m\wedge du^j + \mathbf{n}\Big(b_{km}\Gamma^k_{ij} + \frac{\partial b_{ij}}{\partial u^m}\Big)\, du^m\wedge du^j \\
&\quad - \mathbf{e}_l\, b^l_m b_{ij}\, du^m\wedge du^j \\
&= \mathbf{e}_l\Big(\frac{\partial\Gamma^l_{ij}}{\partial u^m} + \Gamma^l_{km}\Gamma^k_{ij} - b^l_m b_{ij}\Big)\, du^m\wedge du^j + \mathbf{n}\Big(b_{km}\Gamma^k_{ij} + \frac{\partial b_{ij}}{\partial u^m}\Big)\, du^m\wedge du^j
\end{aligned}$$

From this we see

$$\begin{aligned}
0 &= \Big(\frac{\partial\Gamma^l_{ij}}{\partial u^m} + \Gamma^l_{km}\Gamma^k_{ij} - b^l_m b_{ij}\Big)\, du^m\wedge du^j \\
0 &= \Big(b_{km}\Gamma^k_{ij} + \frac{\partial b_{ij}}{\partial u^m}\Big)\, du^m\wedge du^j
\end{aligned}$$

Remembering the need to antisymmetrize the coefficients which we discussted earlier we have

$$\begin{aligned}
\frac{\partial\Gamma^l_{ij}}{\partial u^m} - \frac{\partial\Gamma^l_{im}}{\partial u^j} + \Gamma^l_{kj}\Gamma^k_{im} - \Gamma^l_{kj}\Gamma^k_{im} - b^l_m b_{ij} + b^l_j b_{im} &= 0 \\
\frac{\partial b_{ij}}{\partial u^m} - \frac{\partial b_{im}}{\partial u^j} + b_{km}\Gamma^k_{ij} - b_{kj}\Gamma^k_{im} &= 0
\end{aligned}$$

The second equations are the equations of Codazzi-Mainardi. The first equations, when rewritten as

$$\boxed{R_i{}^l{}_{mj} + b^l_m b_{ij} - b^l_j b_{im} = 0} \qquad \text{Equations of Gauss}$$

are called the equations of Gauss. Both equations show that the connection coefficients Γ^i_{jk} and the b_{ij} are not independent. We will use the equations of Gauss to prove Gauss's theorema egregium in the section on 2-manifolds in $\mathbb{R}^3$.

1.3.2 Curves and Geodesics

It has been a while since we used the fact that the n-manifold M was embedded in $\mathbb{R}^{n+1}$. For this section we will again make use of this.

The idea is to investigate some of the manifold's local properties by means of curves going through the point of interest. For this to be effective we need to know a bit about curves. Sadly, we must limit ourselves to only the most essential material.

Let $C(s) : (a, b) \to \mathbb{R}^n$ be a C^∞ function from an interval $(a, b) \subseteq \mathbb{R}$ to $\mathbb{R}^n$. We assume that the parameter s is arc length as this makes things simpler without restricting generality very much, although it precludes kinematic interpretations. Also for simplicity we will assume that C is one to one, so the curve does not have double points, and we further assume that $C'(s)$ is never 0, so the curve does not have cusps and does not reverse direction. If we were interested in curves themselves these would be very restrictive conditions but we are interested only in curves on manifolds and for our purposes the conditions cause no problems.

We define the *unit tangent vector* $\mathbf{T}$ as

$$\mathbf{T} = \frac{dC}{ds}$$

We know $\mathbf{T}$ is a unit vector because the parameter is arc length. We now take another derivative to form $\mathbf{N} = \frac{d\mathbf{T}}{ds}$. Because $(\mathbf{T}, \mathbf{T}) = 1$ we have

$$\begin{aligned} (\frac{d\mathbf{T}}{ds}, \mathbf{T}) + (\mathbf{T}, \frac{d\mathbf{T}}{ds}) &= 0 \\ (\frac{d\mathbf{T}}{ds}, \mathbf{T}) &= 0 \\ \frac{d\mathbf{T}}{ds} &\perp \mathbf{T} \end{aligned}$$

We now make another assumption and this one will not always be satisfied, but under normal conditions it will be. We assume

$$\frac{d\mathbf{T}}{ds} \neq 0$$

The only way this can fail over an interval is if the unit tangent vector is constant and the curve is a straight line for that interval. However, it is certainly *possible* that $\frac{d\mathbf{T}}{ds} = 0$ at the occasional isolated point, and our analysis will fail at that point. The assumption that the curve $C(s)$ is C^∞ precludes some of the bad things that can happen. We will talk a bit more about this later.

Since, as we have assumed, $\frac{d\mathbf{T}}{ds} \neq 0$, we can form the *unit principal normal* vector $\mathbf{N}$ by

$$\mathbf{N} = \frac{\frac{d\mathbf{T}}{ds}}{||\frac{d\mathbf{T}}{ds}||}$$

and we will also define the *curvature of the curve* κ at a point as

$$\kappa = \left|\left|\frac{d\mathbf{T}}{ds}\right|\right| > 0$$

(There is a technical carp here; this works only for curves an ambient space of dimension ≥ 3; for curves in $\mathbb{R}^2$ we have a slightly different situation and κ may be negative. However, this case does not occur for our work.) Trivially we have

$$\frac{d\mathbf{T}}{ds} = \kappa\mathbf{N}$$

Now we suppose that the curve lies in the manifold M, in which case the unit tangent vector $\mathbf{T}$ will lie in the tangent space to the manifold. However, there is no reason to think that $\kappa\mathbf{N}$ will be in the tangent space. We now resolve $\kappa\mathbf{N}$ into a component parallel to n, the normal to the manifold, and a component *in* the tangent space. This idea turns out to be of staggering import.

Here is the explanation. The curvature of a curve on a surface comes from two imputs. There is a component that is due the curvature of the manifold itself and there is a component caused by the curve curving *in the manifold*. To illustrate the second idea, imagine that you are lost in a barren countryside with many hills and valleys and you wish to travel as straight as possible in some direction. You have three sticks of length, say, two meters. Stick one in the ground. You walk in the direction you wish to go for a way, and, with the first stick still visible you stick a second stick in the ground. You then walk a while further, keeping both sticks in sight, and sighting along the third stick you adjust your position so that the first two sticks are in line with your present position, and then stick the third stick in the ground. Now go get the first stick, and repeat the process above keeping in line with the second and third sticks, and put the stick in the ground. Repeating this process over and over, you will trace a path that is *as straight as possible* in a hilly world[5]. Such as path is called a *geodesic*. When you travel on a geodesic, you are *not* curving in the surface, and the normal $\mathbf{N}$ to your path points the same direction as the normal to the surface (or opposite to it). If you deviate from the path indicated by the sticks, then you are curving *in the surface* and the principle normal to your path will deviate from the normal to the surface. An example on the earth of a geodesic (no curving in the surface) is a great circle. Now take a smaller circle tangent to a great circle on the earth and you see that the second circle is curved *on the earth's surface*. Think about the normal vectors to the curves in each case.

Now we would like to make all this mathematical and we have the equipment to do this; we discussed it in the paragraph above about resolving $\kappa\mathbf{N}$. Resolving $\kappa\mathbf{N}$ into its components perpendicular and parallel to the tangent space, we have

$$\frac{d\mathbf{T}}{ds} = \kappa\,\mathbf{N} = \kappa_g\,\mathbf{s} + \kappa_n\,\mathbf{n}$$

(Recall n is the unit normal vector to the surface.) Here $\mathbf{s}$ is a unit vector in the tangent space, κ_g is the *geodesic curvature* and κ_n is the *normal curvature.*

Def A curve in the manifold is a *geodesic* $\Longleftrightarrow \kappa_g = 0$.

[5] And will travel three times as far.

To clarify a bit, it is true that a geodesic minimizes the distance between nearby points on it; any other path between the two points is longer than the geodesic path, but that is NOT the definition. It is a theorem. We will prove it in the section on Riemannia Geometry. Note that the definition is quite close to the intuitive discussion involving the sticks.

Now we must find formulas for κ_g and κ_n. This is fairly easy. The points of the curve, being on the surface, have surface coordinates $u^1(s), \ldots, u^n(s)$. Thus we have

$$\begin{aligned}
C(s) &= R(u^i(s)) \\
\mathbf{T} = \frac{dC}{ds} &= \frac{\partial R}{\partial u^i}\frac{du^i}{ds} = e_i \frac{du^i}{ds} \\
\kappa\,\mathbf{N} = \frac{d\mathbf{T}}{ds} &= e_i \frac{d^2u^i}{ds} + \frac{\partial e_i}{\partial u^j}\frac{du^j}{ds}\frac{du^i}{ds} \\
&= e_k \frac{d^2u^k}{ds} + (e_k\Gamma^k_{ij} + \mathbf{n}b_{ij})\frac{du^j}{ds}\frac{du^i}{ds} \\
\mathbf{s}\,\kappa_g + \mathbf{n}\,\kappa_n = \kappa\,\mathbf{N} &= e_k\,\Big(\frac{d^2u^k}{ds^2} + \Gamma^k_{ij}\frac{du^i}{ds}\frac{du^j}{ds}\Big) + \mathbf{n}\,b_{ij}\frac{du^i}{ds}\frac{du^k}{ds}
\end{aligned}$$

from which we get immediately

$$\begin{aligned}
\mathbf{s}\,\kappa_g &= e_k\,\Big(\frac{d^2u^k}{ds^2} + \Gamma^k_{ij}\frac{du^i}{ds}\frac{du^j}{ds}\Big) \\
\kappa_n &= b_{ij}\frac{du^i}{ds}\frac{du^k}{ds}
\end{aligned}$$

Since $\mathbf{s}$ is a unit vector we have

$$\begin{aligned}
(k_g\,\mathbf{s}, k_g\,\mathbf{s}) &= \Big(e_k\,\Big(\frac{d^2u^k}{ds^2} + \Gamma^k_{ij}\frac{du^i}{ds}\frac{du^j}{ds}\Big),\, e_l\,\Big(\frac{d^2u^l}{ds^2} + \Gamma^l_{rs}\frac{du^r}{ds}\frac{du^s}{ds}\Big)\Big) \\
\kappa_g^2 &= g_{kl}\Big(\frac{d^2u^k}{ds^2} + \Gamma^k_{ij}\frac{du^i}{ds}\frac{du^j}{ds}\Big)\Big(\frac{d^2u^l}{ds^2} + \Gamma^l_{rs}\frac{du^r}{ds}\frac{du^s}{ds}\Big)
\end{aligned}$$

and we have the formula for κ_g. Notice that κ_g depends on the curve and the Γ^i_{jk} which means that κ_g is intrinsic and does not depend on the embedding. It is more difficult to get $\mathbf{s}$ but we have no real need for it so we won't bother.

Now look at κ_n. Note that the formula for κ_n uses only the b_{ij} which have nothing to do with the curve C, and the components of the vector $\mathbf{T}$, that is $\mathbf{T} = e_i \frac{du^i}{ds}$. Hence any two curves $C_1(s)$ and $C_2(s)$ which go through $p \in M$ in the same direction have the same normal curvature. This means that the normal curvature κ_n is a function only of the manifold and the direction in the tangent space, and it is thus a property of the *manifold* itself and does not depend on the curve.

With regard to κ_g we have the following important theorem. Since the e_i are linearly independent,

Theorem A curve $C(s) = \mathbf{R}(u^i(s))$ in the manifold is a geodesic if and only if it satifies the (nonlinear) system of ordinary differential equations

$$\frac{d^2u^k}{ds^2} + \Gamma^k_{ij}\frac{du^i}{ds}\frac{du^j}{ds} = 0$$

(Here the parameter s is arc length.)

Note that by the usual existence theorems for differential equations there is a unique geodesic stating at a point $p \in M$ and going in a chosen direction.

Geodesics of of great importance and there are many books devoted to them. An example of a geodesic question is, if you pick a direction at a point on a torus and follow it, will it eventually come back on itself or will it wind around forever. Mostly it winds forever, but if you choose the direction carefully it closes up. On a general manifold, the existence of closed geodesics is a matter of great interest.

From the physics point of view, geodesics are extremely important because in four dimensional space time light travels in geodesics.

But we must move on.

1.3.3 Special Case; Surfaces in $\mathbb{R}^3$

It is worth looking at the special case $n = 2$ because it has interesting features, one of the most important of which is the Gaussian Curvature. Gaussian Curvature is not only extremely interesting in itself but also is the jumping off place to a great deal of modern work on manifolds. This was pioneered by S. S. Chern.

Gaussian Curvature

In this section we will take a classical approach to the Gaussian Curvature. We begin by recalling that the normal curvature of a surface depends only on the surface. If we sit at a point $p \in M$ and rotate a unit tangent vector v around the point the value of the normal curvature will change and because we are dealing with a continuous function κ_n will have a maximum $\kappa_{\max}$ and a minimum $\kappa_{\min}$.

To be explicit, we wish to maximize and minimize $\kappa_n(v)$ under the constraint that $||v|| = g_{ij}v^iv^j = 1$. This is a problem tailored to the method of Lagrange Multipliers (refer to your old Calculus textbook) whereby we form the auxiliary function

$$F(v^i) = \kappa_n(v^i) - \lambda(g_{ij}v^iv^j - 1)$$

and seek to maximize or minimize F. (Physicists and applied mathematicians take note; the value of λ in problems with Lagrange Multipliers is often of interest, for no obvious reason.) We recall that

$$\kappa_n(v^i) = b_{ij}v^iv^j \qquad i, j = 1, 2$$

We now take the partial derivatives of F with respect to the variables v^i and λ and set them equal to 0.

$$\frac{\partial F}{\partial u^i} \;=\; b_{i,j}v^j - \lambda g_{i,j}v^j \;=\; 0 \qquad i, j = 1, 2$$

$$\frac{\partial F}{\partial \lambda} \;=\; -(g_{ij}v^iv^j - 1) \;=\; 0$$

Note that the last equation is simply the constraint, as is always the case. The first two equations are a 2×2 system of homogeneous linear equations, and thus have a solution if and only if the coefficient determinant is 0.

$$\det\begin{pmatrix} b_{11}-\lambda g_{11} & b_{12}-\lambda g_{12} \\ b_{21}-\lambda g_{21} & b_{22}-\lambda g_{22} \end{pmatrix} = 0$$

Suppose now that we find a solution λ to this equation. Then F will be a max or a min (by the nature of the problem we should have one of each) and moreover the v^i corresponding to that λ will satisfy $||v|| = g_{ij}v^iv^j = 1$. Since there is a max and a min, we have found them and moreover the max and min of F are the max and min of

$$F(v^i) = \kappa_n(v^i) - \lambda(g_{ij}v^iv^j - 1) = \kappa_n(v^i) - \lambda\cdot 0$$

and are thus $\kappa_{\max}$ and $\kappa_{\min}$. Thus it comes down to solving $\det(b_{ij} - \lambda g_{ij}) = 0$.

This is not particularly challenging. We have

$$\begin{aligned} (b_{11}-\lambda g_{11})(b_{22}-\lambda g_{22}) - (b_{12}-\lambda g_{12})(b_{21}-\lambda g_{21}) &= 0 \\ (g_{11}g_{22}-g_{12}g_{21})\lambda^2 - (g_{11}b_{22}+g_{22}b_{11}-g_{12}b_{21}-g_{21}b_{12})\lambda + b_{11}b_{22}-b_{12}b_{21} &= 0 \end{aligned}$$

Recalling that $g = \det(g_{ij})$ and $b = \det(b_{ij})$ and that $g_{ij} = g_{ji}$ and $b_{ij} = b_{ji}$ we can rewrite this is

$$g\lambda^2 - (g_{11}b_{22} + g_{22}b_{11} - 2g_{12}b_{21})\lambda + b = 0$$

and the two solutions of this equation are $\kappa_{\max}$ and $\kappa_{\min}$. As it happens we are not so interested in these individual values; we want their average and their product. Recall that if α_1 and α_2 are the two roots of a quadratic equation $a(x-\alpha_1)(x-\alpha_1) = ax^2 + bx + c = 0$ then $a(\alpha_1+\alpha_2) = -b$ and $a(\alpha_1\alpha_2) = c$ by comparing the coefficients of powers of x. Hence we have the equations

$$\begin{aligned} H = \frac{1}{2}(\kappa_{\max}+\kappa_{\min}) &= \frac{g_{11}b_{22}+g_{22}b_{11}-2g_{12}b_{21}}{2g} && \text{Mean Curvature} \\ K = \kappa_{\max}\cdot\kappa_{\min} &= \frac{b}{g} && \text{Gaussian Curvature} \end{aligned}$$

The Mean Curvature has its enthusiasts. For example, if we have a closed curve in 3-space, we might wonder what is the surface of minimum area bounded by the closed curve. You can find a model by making the closed curve out of wire and dipping it in soapy water. The soap film will (usually) be the surface of minimum area. This minimal suface has 0 mean curvature, so this is a necessary condition for a minimal surface. Study of the mean curvature was initiated by Sophie Germain in her work on elasticity theory.

The Gaussian Curvature is much more important for the following reason. The formula given, $K = b/g$, certainlty makes it look like K depends on the embedding of the surface in 3-space. However, this is an illusion. In fact,

Theorem (Gauss's Theorma Egregium[6]) The Gaussian Curvature depend only on the (g_{ij}) and is thus an intrinsic invariant of the surface.

This means, to restate, that the Gaussian curvature does not depend on how the surface is embedded in 3-space.

To make it clearer what this means, let Φ be a one to one map between two surfaces M_1 and M_2. Let u^1 and u^2 be local coordinates around $p \in M_1$. Then u^1 and u^2 can also be used as coordinates around $\Phi(p)$ on M_2, and a metric can be put on M_2 by using the (g_{ij}) at $\Phi(p) \in M_2$. To be explicit, we take (v, w) at $\Phi(p)$ (for $v, w \in T_{\Phi(p)}$) to be $(\Phi^*(v), \Phi^*(w))$ on M_1. Now suppose that M_2 has a metric already defined on it. Then Φ is an *isometry* if and only if, when using the u^i as coordinates on both surfaces, the inner products coincide, which means that they have the same (g_{ij}) and thus, by the Theorma Egregium, the same K. Hence it is not possible to have an accurate flat map of the earth (which would preserve the (g_{ij})) because the earth has $K = 1/R^2$ and the flat map has $K = 0$ where R is the radius of the earth.

However, we *can* map a portion of a plane onto a cylinder of any radius as suggested by the fact that plane and cylinder both have $K = 0$.

The proof of the Theorema Egregium is not at all difficult because we have previously set everything up for it. Recall that $K = \det(b_{ij})/\det(g_{ij})$. Recall also the formula of Gauss

$$R_i{}^l{}_{mj} - b^l_m b_{ij} + b^l_j b_{im} = 0$$

We manipulate this equation a bit

$$\begin{aligned} R_i{}^l{}_{mj} &= b^l_m b_{ij} - b^l_j b_{im} \\ g_{kl} R_i{}^l{}_{mj} &= g_{kl} b^l_m b_{ij} - g_{kl} b^l_j b_{im} \;=\; b_{ij} b_{km} - b_{im} b_{kj} \end{aligned}$$

Now we cleverly substitute particular values for the i and j; we let $i = j = 1$ and $k = m = 2$ and we get

$$\begin{aligned} g_{2\ell} R_1{}^\ell{}_{21} &= b_{11} b_{22} - b_{12} b_{21} \;=\; \det(b_{ij}) \\ K &= \frac{\det(b_{ij})}{\det(g_{ij})} = \frac{g_{2\ell}}{\det(g_{ij})} R_1{}^\ell{}_{21} \end{aligned}$$

Since $R_1{}^\ell{}_{21}$ depends only on the Γ^i_{jk} and the Γ^i_{jk} depend only on the g_{ij}, K ultimately depends only on the g_{ij} and thus is intrinsic; the Theorema Egregium is proved.

Example: The Sphere

We think it might be worthwhile to give an example of all of this, so we will present the results for the sphere of radius a. The conscientious student will

[6] Egregium is Latin for distinguished, excellent, admirable

work out the details and compare her results with the summary we give. A couple of things will be worked out in more detail.

The standard parametrization of the sphere of radiua a is to use colatitude $\phi = \pi/2-$ latitude and longitude θ (in that order). Then the parametrization is

$$\mathbf{R} = \langle a\sin\phi\cos\theta, a\sin\phi\sin\theta, a\cos\phi\rangle$$

Then

$$\begin{aligned} e_1 &= \frac{\partial \mathbf{R}}{\partial u^1} = \frac{\partial \mathbf{R}}{\partial \phi} = \langle a\cos\phi\cos\theta, a\cos\phi\sin\theta, -a\sin\phi\rangle \\ e_2 &= \frac{\partial \mathbf{R}}{\partial u^2} = \frac{\partial \mathbf{R}}{\partial \theta} = \langle -a\sin\phi\sin\theta, a\sin\phi\cos\theta, 0\rangle \end{aligned}$$

Using the standard inner product on $\mathbb{R}^3$ we have

$$(g_{ij}) = \begin{pmatrix} a^2 & 0 \\ 0 & a^2\sin^2\phi \end{pmatrix}$$

Using the cross product we have, dividing out the length,

$$\mathbf{n} = \langle \sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi\rangle$$

To find the Christoffel symbols we use the formula

$$\Gamma^k_{ij} = g^{kl}(\mathbf{e}_l, \frac{\partial \mathbf{e}_i}{\partial u^j})$$

We will need

$$\frac{\partial \mathbf{e}_2}{\partial u^1} = \langle -a\cos\phi\sin\theta,\ a\sin\phi\sin\theta,\ 0\rangle$$

so for example

$$\begin{aligned} \Gamma^2_{21} &= g^{2\ell}(\mathbf{e}_l, \frac{\partial \mathbf{e}_2}{\partial u^1})) = g^{22}(\mathbf{e}_2, \frac{\partial \mathbf{e}_2}{\partial u^1})) \\ &= \frac{1}{a^2\sin^2\phi}\big(a^2\sin\phi\cos\phi\big) = \frac{\cos\phi}{\sin\phi} = \cot\phi \end{aligned}$$

After a page or two of calculation, we find

$$\omega = \begin{pmatrix} 0 & -\sin\phi\cos\phi\, d\theta \\ \cot\phi\, d\theta & \cot\phi\, d\phi \end{pmatrix}$$

Now to find the curvature form we just use

$$\Omega = d\omega + \omega\wedge\omega$$

We have

$$d\omega = \begin{pmatrix} 0 & (-\cos^2\phi + \sin^2\phi)\, d\phi d\theta \\ -\csc^2\phi\, d\phi d\theta & 0 \end{pmatrix}$$

and

$$\begin{aligned}\omega \wedge \omega &= \begin{pmatrix} 0 & -\sin\phi\cos\phi\, d\theta \\ \cot\phi\, d\theta & \cot\phi\, d\phi \end{pmatrix} \wedge \begin{pmatrix} 0 & -\sin\phi\cos\phi\, d\theta \\ \cot\phi\, d\theta & \cot\phi\, d\phi \end{pmatrix} \\ &= \begin{pmatrix} 0 & \cos^2\phi\, d\phi d\theta \\ \cot^2\phi\, d\phi d\theta & 0 \end{pmatrix}\end{aligned}$$

so (remembering that $-\csc^2\phi = -1 - \cot^2\phi$)

$$\Omega = d\omega + \omega \wedge \omega = \begin{pmatrix} 0 & \sin^2\phi \\ -1 & 0 \end{pmatrix} d\phi d\theta$$

from which we read off

$$\begin{aligned} R_1{}^1{}_{12} &= 0 & R_2{}^1{}_{12} &= \sin^2\phi \\ R_1{}^2{}_{12} &= -1 & R_2{}^2{}_{12} &= 0 \end{aligned}$$

Now we can compute the Gaussian Curvature

$$K = \frac{g_{2j}}{g} R_1{}^j{}_{21} = \frac{g_{22}}{g} R_1{}^2{}_{21} = -\frac{g_{22}}{g} R_1{}^2{}_{12} = -\frac{a^2\sin^2\phi}{a^4\sin^2\phi}(-1) = \frac{1}{a^2}$$

You can also get K by noting that $\kappa_{\max} = \kappa_{\min} = 1/a$ but our method was more enlightening.

Lastly we will show that a great circle on the sphere is a geodesic. Recall that a curve is a geodesic if and only if it satisfies the geodesic equation

$$\frac{d^2u^k}{ds^2} + \Gamma^k_{ij}\frac{du^i}{ds}\frac{du^j}{ds} = 0$$

Since all great circles are pretty much the same, we can take a particular great circle, say the equator, and check it for being a geodesic. The equator has the coordinates $u^1 = \phi = \pi/2$ and $u^2 = \theta$ which runs from 0 to 2π. For $k = 1$ in the geodesic equation the only possible nonzero term in the equation is when $i = j = 2$. But $\Gamma^2_{22} = 0$ since $\omega_2{}^2 = \cot\phi\, d\phi$ and has no $d\theta$ term. Thus we need only worry about $k = 2$. Note $ds = a d\theta$ so pulling out the $1/a^2$ the equation becomes, remembering $u^1 =$ constant,

$$\frac{d^2u^2}{d\theta^2} + \Gamma^1_{22}\frac{du^2}{d\theta}\frac{du^2}{d\theta} = 0$$

But $\omega_2{}^1 = -\sin\phi\cos\phi\, d\theta$ so $\Gamma^1_{22} = -\sin\phi\cos\phi = -\sin\pi/2\cos\pi/2 = 0$ Hence the equator, and thus all great circles, satisfies the geodesic equation and thus all great circles are geodesics.

It would have been easy to reason geometrically, since the the unit normal to a great circle is just $-\mathbf{n}$ (where $\mathbf{n}$ is the normal to the sphere) and thus $\kappa_g = 0$ but we wanted to see how this could be done analytically.

The Gauss-Bonnet theorem

The Gauss-Bonnet theorem is one of the most important theorems in mathematics, partly because it unifies a lot of geometry by proving many elementary formulas in a coherant way and partly because it was the jumping off point for such advanced topics as Chern cohomology classes and following this road has lead to vast generalizations of our good old surface theory. Here differential geometry intersects the algebraic topology of manifolds in highly non-trivial ways.

The Gauss-Bonnet theorem comes in several forms and we will state different looking forms of it. The first might be called the Polygonal form.

Theorem (Gauuss-Bonnet I) Let M be a 2-manifold and $S \subseteq M$ be a simply connected subset of M whose boundary is a smooth curve except for a finite number of corners with exterior angles α_i. Then

$$\int_S K\, dS + \int_{\partial S} \kappa_g\, ds + \sum_i \alpha_i = 2\pi$$

Here K is the Gaussian curvature of S, κ_g is the geodesic curvature of ∂S, dS is the element of surface area, and ds is the element of arc length.

The second form might be called the Topological form. We need the following ideas. The genus g of a closed (no boundary) 2-manifold M is the number of holes. For example a sphere has genus $g = 0$, a 2-torus has $g = 1$ and a pretzel (skin only) has $g = 3$. If a manifold is cut into regions by "polygons"", and there are V vertices, E edges and F faces then the Euler-Poincare characteristic of M is

$$\chi(M) = V - E + F$$

and it is known that

$$\chi(M) = V - E + F = 2(1 - g)$$

The topological form of the Gauss-Bonnet theorem is then

Theorem (Gauss-Bonnet II) Let M be an oriented 2-manifold of genus g. Then

$$\int_M K\, dS = 4\pi(1 - g) = 2\pi\chi(M)$$

Since the proof of this theorem is a little harder than most results in this book and we don't want you to miss the applications since they are so beautiful, we have decided that we will first present the applications and then do the proof.

App 1. Surface area of a sphere. Use GBII and the fact that the Gaussian curvature is $K = 1/a^2$ to get

$$\begin{aligned} \int_M K\, dS &= 4\pi(1 - g) \\ \int_M \frac{1}{a^2}\, dS &= 4\pi(1 - 0) \end{aligned}$$

$$\begin{aligned} \frac{1}{a^2}\int_M dS &= 4\pi \\ \text{surface area of sphere} &= 4\pi a^2 \end{aligned}$$

App 2 Surface area of Lune, with opening angle ψ. The lune is bounded by great circles that have $\kappa_g = 0$ and the exterior angles are $\pi - \psi$ so from GBI we have

$$\begin{aligned} \int_S K\,dS + \int_{\partial S} \kappa_g\,ds + \sum_i \alpha_i &= 2\pi \\ \int_S \frac{1}{a^2}\,dS + \int_{\partial S} 0\,ds + \pi - \psi + \pi - \psi = 2\pi & \\ \text{surface area of lune } = 2a^2\psi & \end{aligned}$$

App 3 Area of a Geodesic Triangle on a Sphere. The "straight lines" on a sphere are the geodesics which are great circles. Hence the closest thing to a triangle on a sphere is a the area bounded by three great circles. Three great circles actually bound a number of areas, but the area is selected by choosing the angles. An angle on a sphere between two curves are defined as the angle between the tangent lines to the curves where the curves meet. Suppose the interior angles of the triangle are α, β and γ. Then the exterior angles for the triangle are $\pi - \alpha$, $\pi - \beta$, and $\pi - \gamma$. Once again, for a sphere of radius a we have the Gaussian Curvature $K = 1/a^2$ and $\kappa_g = 0$ since we are dealing with geodesics, so

$$\begin{aligned} \int_S K\,dS + \int_{\partial S} \kappa_g\,ds + \sum_i \alpha_i &= 2\pi \\ \int_S \frac{1}{a^2}\,dS + \int_{\partial S} 0\,ds + \pi - \alpha + \pi - \beta + \pi - \gamma &= 2\pi \\ \frac{1}{a^2}S &= 2\pi + \alpha + \beta + \gamma - 3\pi \\ \text{surface area of tiangle} &= a^2(\alpha + \beta + \gamma - \pi) \end{aligned}$$

Triangles on Spheres always have an angle sum in excess of π in contrast to triangles on the plane, where the angle sum is exactly π. The amount by which the angle sum exceeds π is called the *excess* and what we have shown is the that the area $= a^2\times$ (the excess).

App 4 Geodesic Polygon on a sphere. Let the interior angles be $\beta_i;\ i = 1, \ldots, n$ so the exterior angles are $\pi - \beta_i;\ i = 1, \ldots, n$. Then, as above,

$$\begin{aligned} \int_S \frac{1}{a^2}\,dS + \int_{\partial S} 0\,ds + \sum_{i=1}^{n}(\pi - \beta_i) &= 2\pi \\ \frac{1}{a^2}\cdot \text{Area} + n\pi - \sum_{i=1}^{n}\beta_i &= 2\pi \end{aligned}$$

$$\text{Area} \quad = \quad a^2(\sum_{i=1}^{n} \beta_i - (n-2)\pi)$$

App 5 You can only understand this one if you know something about Lobachevsky geometry, also called hyperbolic geometry. The Lobachevsky plane has a constant a analogous to the radius for a sphere and now the Gaussian curvature is $K = -1/a^2$. Negative Gaussian Curvature is characteristic of saddle-like surfaces and one model of part of the Lobachevski plane is the pseudosphere which indeed is saddle like at every point. In Lobachevsky geometry the angle sum is always *less* than π and there is thus a *deficit* which is $\pi - (\alpha + \beta + \gamma)$. Almost precisely the same calculation as in App 3 tells us that the area of a geodesic triangle is $a^2\times$ (the deficit). Geodesics in Lobachevsky geometry can meet one another with angle 0. A geodesic triangle in which this happens will then have the maximum possible area for a geodesic triangle which will be πa^2.

App 6 Geodesic Curvature of a parallel. A parallel on a sphere is a curve defined by $\phi =$ constant. (This is a ludicrously bad name because these curves are not great circles. Since any two great circles clearly meet each other, there are no "parallel lines" on a sphere. However the name parallel is traditional and we are stuck with it.) The equator $\phi = \pi/2$ is a geodesic, but the parallels with $\phi < \pi/2$ are *not* geodesics, and thus have positive geodesic curvature κ_g which we are now going to compute analytically with the help of the Gauss Bonnet theorem. Let ϕ_0 determine the parallel. The area of the cap above the parallel is

$$\int_{\text{cap}} a^2 \sin\phi \, d\phi d\theta = \int_0^{2\pi}\int_0^{\phi_0} a^2 \sin\phi \, d\phi d\theta = 2\pi a^2(1-\cos\phi_0)$$

We then have, making use of the obvious $\kappa_g =$ constant,

$$\begin{aligned}
\int_S K\,dS + \int_{\partial S} \kappa_g\,ds + \sum_i \alpha_i &= 2\pi \\
\frac{1}{a^2}2\pi a^2(1-\cos\phi_0) + 2\pi a \sin\phi\,\kappa_g + 0 &= 2\pi \\
-\cos\phi_0 + a\sin\phi_0\,\kappa_g &= 0 \\
\kappa_g &= \frac{1}{a}\cot\phi_0
\end{aligned}$$

Proof of the Gauss Bonnet Theorem

We first prove the theorem for a 2-manifold (surface) and a region $S \subseteq M$ whose boundary ∂S is a smooth curve (we are putting off the corners till later) and which is contained in a single coordinate patch. It will be important to us that the general basis we use is orthonormal. To arrange this let u^1, u^2 be an oriented coordinate system for M and set $\tilde{\mathbf{e}}_1 = \frac{\partial \mathbf{R}}{\partial u^1}$, where $\mathbf{R}$ is the position vector of M in $\mathbf{R}^3$. Let $\tilde{g}_{11} = (\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_1)$ and set

$$\mathbf{e}_1 = \frac{1}{\sqrt{\tilde{g}_{11}}}\tilde{\mathbf{e}}_1$$

Thus $\mathbf{e}_1$ is a unit vector. Let $\mathbf{e}_2$ be a unit vector in $T_p(M)$ which at each point is orthogonal to $\mathbf{e}_1$ and in a direction so that $\{\mathbf{e}_1, \mathbf{e}_2\}$ has the same orientation as M. We have now created a orthonormal general basis of $T_p(M)$. The orthogonality will have an effect on the curvature form ${\omega_i}^k$. We have

$$\begin{aligned}
(\mathbf{e}_i, \mathbf{e}_j) &= \delta_{ij} \\
(D\mathbf{e}_i, \mathbf{e}_j) + (\mathbf{e}_i, D\mathbf{e}_j) &= D\delta_{ij} = 0 \\
(\mathbf{e}_k {\omega_i}^k, \mathbf{e}_j) + (\mathbf{e}_i, \mathbf{e}_l {\omega_j}^l) &= 0 \\
{\omega_i}^k \delta_{kj} + {\omega_j}^l \delta_{il} &= 0 \\
{\omega_i}^j + {\omega_j}^i &= 0 \\
{\omega_i}^j &= -{\omega_i}^j
\end{aligned}$$

Now we are now ready to launch into the proof, which consists of 3 parts. In the first, we effect a cosmetic change on $\int_{\partial S} \kappa_g \, ds$ to get the integral of a certain form η so the integral becomes $\int_{\partial S} \eta$. The second part is to apply stokes theorem, which is very quick. The third part is more cosmetic shifting on $d\eta$ so that it comes out what we need for the Gauss Bonnet theorem. None of this is deep; it's just a bit tricky. Also, the only substantive step is the use of Stokes theorem; the rest is cosmetic shifting.

For the points along the curve ∂S we must use another basis. The direction of integration for ds is the one determined on ∂S by the orientation on M and inherited by S. The first vector is the unit tangent vector $\mathbf{T}$ to the curve directed in the direction of the orientation and the second is the unit vector we called $\mathbf{s}$ in the section on curves but which we are now going to call $\mathbf{b}$ to avoid variable clash. We arrange $\mathbf{b}$ so that it is in the tangent space $T_p(M)$ for $p \in \partial S$ and is perpendicular to $\mathbf{T}$. The direction of $\mathbf{b}$ is determined so that when we go around the curve ∂S according to its orientation, $\mathbf{T}, \mathbf{b}$ has the same orientation as M. Then we have

$$\begin{aligned}
\frac{D\mathbf{T}}{ds} &= \mathbf{b}\kappa_g \\
\frac{D\mathbf{b}}{ds} &= -\mathbf{T}\kappa_g
\end{aligned}$$

Note the coefficients are skew symmetric as we showed they had to be for an orthonormal basis. On the other hand we have

$$\begin{aligned}
\mathbf{e}_1 &= \mathbf{T}\cos\theta - \mathbf{b}\sin\theta \\
\mathbf{e}_2 &= \mathbf{T}\sin\theta + \mathbf{b}\cos\theta
\end{aligned}$$

so that

$$\begin{aligned}
\frac{D\mathbf{e}_1}{ds} &= (-\mathbf{T}\sin\theta - \mathbf{b}\cos\theta)\frac{d\theta}{ds} + \frac{D\mathbf{T}}{ds}\cos\theta - \frac{D\mathbf{b}}{ds}\sin\theta \\
&= -\mathbf{e}_2\frac{d\theta}{ds} + \mathbf{b}\kappa_g\cos\theta - (-\mathbf{T}\,\kappa_g)\sin\theta
\end{aligned}$$

$$\begin{aligned} &= -\mathbf{e}_2\frac{d\theta}{ds} + \mathbf{e}_2\kappa_g \\ &= \mathbf{e}_2\Big(\kappa_g - \frac{d\theta}{ds}\Big) \end{aligned}$$

The reader may wish to practise by calculating $\frac{D\mathbf{e}_2}{ds}$ and she should come out with

$$\frac{D\mathbf{e}_2}{ds} = \mathbf{e}_1\Big(- \kappa_g + \frac{d\theta}{ds}\Big)$$

Note skew symmetry. (We do not actually need this result.) Now we have

$$D\mathbf{e}_1 = \mathbf{e}_i{\omega_1}^i = \mathbf{e}_1{\omega_1}^1 + \mathbf{e}_2{\omega_1}^2 = \mathbf{e}_2{\omega_1}^2$$

since ${\omega_1}^1 = 0$ due to skew symmetry and we have

$$\begin{aligned} {\omega_1}^2 &= \Big(\kappa_g - \frac{d\theta}{ds}\Big)ds \\ &= \kappa_g ds - d\theta \end{aligned}$$

We have finished part 1; we have derived the differential form η to which we will apply Stokes theorem and it is ${\omega_1}^2 = \kappa_g ds - d\theta$. This is of course not obvious.

The next step is to apply Stokes theorem. Before we do this let us observe the following.

$${\omega_k}^2 \wedge {\omega_1}^k = {\omega_1}^2 \wedge {\omega_1}^1 + {\omega_2}^2 \wedge {\omega_2}^k = 0$$

because, due to skew symmetry, ${\omega_1}^1 = {\omega_2}^2 = 0$. Next note that

$$\begin{aligned} {\Omega_1}^2 &= d{\omega_1}^2 + {\omega_k}^2 \wedge {\omega_1}^k \\ d{\omega_1}^2 &= {\Omega_1}^2 - {\omega_k}^2 \wedge {\omega_1}^k \end{aligned}$$

and so we have

$$\begin{aligned} \int_{\partial S}(\kappa_g - d\theta) &= \int_{\partial S} {\omega_1}^2 \\ &= \int_S d{\omega_1}^2 \qquad \text{the substantive step} \\ &= \int_S {\Omega_1}^2 - {\omega_k}^2 \wedge {\omega_1}^k \\ &= \int_S {\Omega_1}^2 \end{aligned}$$

since, as we showed, the second term in the integrand is 0.

The general formula for the Gaussian Curvature K is

$$K = \frac{g_{2\ell}}{\det(g_{ij})}{R_1}^{\ell}{}_{21}$$

1.4 Some Tensors and the Proof of the Gauss-Bonnet theorem

1.4.1 Tensors and their algebra

The reader may have wondered how tensors fit into the structure that we have been building, and this is a convenient place to discuss this. We will discuss tensors in a very superficial way, and in addition we will discuss them in a non-traditional way. Lore on Tensors is vast and for readers who want to pursue them in greater depth we recommend [9] for the pure theory and [**?**] and [**?**] for more applications oriented treatments. For those who wish to sail their rowboats through index storms across the tensor sea, there is [**?**].

The most difficult thing to learn about tensors is that there is almost nothing to learn. Let V be a vector space with dual V^*. For our purposes these are $V = T(M)$ and $V^* = T^*(M)$ but the concepts have applications elsewhere.Let $v_1, \dots, v_k \in V$ and $\lambda^1, \dots, \lambda^\ell \in V^*$. A tensor product is a formal product (where α is a scaler)

$$\alpha\, v_1 \otimes \dots \otimes v_k \otimes \lambda^1 \otimes \dots \otimes \lambda^\ell$$

This product is multilinear and the scalers may move around freely. That is

$$\begin{aligned}&\alpha\, v_1 \otimes v_2 \otimes \dots \otimes v_k \otimes \lambda^1 \otimes \dots \otimes \lambda^\ell + \beta\, v_1 \otimes \tilde{v}_2 \otimes \dots \otimes v_k \otimes \lambda^1 \otimes \dots \otimes \lambda^\ell \\ &= v_1 \otimes (\alpha\, v_2 + \beta\, \tilde{v}_2) \otimes \dots \otimes v_k \otimes \lambda^1 \otimes \dots \otimes \lambda^\ell\end{aligned}$$

and similarly with any other slot. For convenience we put λs from V^* at the end so, for example

$$\begin{aligned}&(v_1 \otimes v_2 \otimes \lambda^1 \otimes \lambda^2 \otimes \lambda^3) \quad \otimes \quad (v_3 \otimes v_4 \otimes v_5 \otimes \lambda^4 \otimes \lambda^5) \\ &= v_1 \otimes v_2 \otimes v_3 \otimes v_4 \otimes v_5 \quad \otimes \quad \lambda^1 \otimes \lambda^2 \otimes \lambda^3 \otimes \lambda^4 \otimes \lambda^5\end{aligned}$$

We reiterate, because it seems hard to learn, that $v_1 \otimes v_2$ is not some vector or some other previously known object; it is a formal product which can mutate under bilinerity and is associativity but that is all. It is remarkable that such a construction is so useful, although the use often requires additional structure. Tensor analysis itself does not have powerful theorems like Stokes' theorem for differential forms.

Naturally if $e_1, \dots e_n$ is a basis for V with dual basis $e^1, \dots e^n$ for V^* we can, using bilinearity, express all tensors as sums of products of the form

$$\alpha\, e_1 \otimes \dots \otimes e_k \otimes e^1 \otimes \dots \otimes e^\ell$$

Thus we could write such a sum as

$$t^{i_1 \dots i_k}_{j_1 \dots j_\ell}\, e_{i_1} \otimes \dots \otimes e_{i_k} \otimes e^{j_1} \otimes \dots \otimes e^{j_\ell}$$

with the Einstein summation convention in force for the indices. In practice k and ℓ are stable in a calcultion.

This finishes our introduction to the algebra of tensors. Next we will bring on tensor calculus, which takes place on manifolds.

1.4.2 Tensor Calculus

You cannot actually do any tensor calculus unless you have an affine connection. With an affine connection you can do quite a lot, but the full power of tensor calculus only comes when you have an inner product on the tangent space of the manifold which means you have a (pseudo-)Riemannian metric. We will deal here with only the basic stuff which involves the affine connection. Essentially this means that we have coefficients Γ^i_{jk} so that we have

$$\begin{aligned} De_j &= e_i\,\Gamma^i_{jk} du^k \\ \frac{De_i}{\partial u^k} &= e_j\,\Gamma^i_{jk} \end{aligned}$$

Affine connections are the subject of the next section but if you can just believe the formulas temporarily we can deal with tensors and a few other minor matters.

It is *very* important that you realize that

$$\Gamma^i_{jk} \quad \text{is} \;\; *NOT* \;\; \text{a tensor}$$

It does not have the correct transformation formula for a tensor when you change the coordinates. We discuss this later in this section and much more extensively in the next section.

Next we want a formula for $\frac{De_i}{\partial u^k}$. Let us regard the formula $e^i(e_j) = \delta^i_j$, or using a more symmetric notation $\langle e^i,\, e_j\rangle = \delta^i_j$, as a kind of *product*, and then let us insist that Leibniz' formula for products works. Then we have, setting $\frac{De^i}{\partial u^k} = \gamma^i_{jk} e^j$ for some constants γ^i_{jk}

$$\begin{aligned} \frac{D\delta^i_j}{\partial u^k} &= \left\langle \frac{De^i}{\partial u^k}, e_j \right\rangle + \left\langle e^i,\, \frac{De_j}{\partial u^k} \right\rangle \\ 0 &= \langle \gamma^i_{mk} e^m,\, e_j\rangle + \langle e^i,\, e_\ell \Gamma^\ell_{jk}\rangle \\ &= \gamma^i_{mk}\delta^m_j + \Gamma^\ell_{jk}\delta^i_\ell \\ &= \gamma^i_{jk} + \Gamma^i_{jk} \\ \gamma^i_{jk} &= -\Gamma^i_{jk} \\ \frac{De^i}{\partial u^k} &= -\Gamma^i_{jk}\, e^j \end{aligned}$$

Now let us take the derivative of some random tensor

$$t^i_j e_i \otimes e^j$$

using, as usual, Leibniz' rule. We get

$$\begin{aligned} \frac{(D\, t^i_j\, e_i \otimes e^j)}{\partial u^k} &= \frac{\partial t^i_j}{\partial u^k}\, e_i \otimes e^j + t^i_j\, \frac{De_i}{\partial u^k} \otimes e^j + t^i_j\, e_i \otimes \frac{De^j}{\partial u^k} \\ &= \frac{\partial t^i_j}{\partial u^k}\, e_i \otimes e^j + t^i_j\, \Gamma^\ell_{ik} e_\ell \otimes e^j - t^i_j\, \Gamma^j_{\ell k} e_i \otimes e^\ell \end{aligned}$$

$$
\begin{aligned}
&= \frac{\partial t^i_j}{\partial u^k}\, e_i \otimes e^j + t^\ell_j\, \Gamma^i_{\ell k}\, e_i \otimes e^j - t^i_\ell \Gamma^\ell_{jk}\, e_i \otimes e^j \\
&= \Big(\frac{\partial t^i_j}{\partial u^k} + \Gamma^i_{\ell k}\, t^\ell_j - \Gamma^\ell_{jk}\, t^i_\ell \Big)\, e_i \otimes e^j
\end{aligned}
$$

The object in parentheses is the *covariant derivative* of the tensor. In tensor notation we call t^i_j the tensor, i a contravariant index and j a covariant index. (We recall contravariant means changes like the basis vectors and contravariant means changes like the coefficients of a vector[7].) Then to take the covariant derivative of the tensor you have to add terms involving the connection coefficients for each contravariant index (upstairs index) and subtract terms involving the connection coefficients for each coveriant index (downstairs index.) Common notations are

$$\nabla_k\, t^i_j = v^i_{j|k} = \frac{\partial t^i_j}{\partial u^k} + \Gamma^i_{\ell k}\, t^\ell_j - \Gamma^\ell_{jk}\, t^i_\ell$$

Note

$$\nabla_k\, v^i = v^i_{|k} = \frac{\partial v^i}{\partial u^k} + \Gamma^i_{\ell k}\, v^\ell$$

which is our old familiar coveriant derivative so that our new stuff is a generalization of the material we got from differential forms to a wider class of objects. We note that the formulas we have derived are very general and do not depend on $\Gamma^i_{jk} = \Gamma^i_{kj}$.

Another possible way to look at it is

$$\frac{(D\, t^i_j\, e_i \otimes e^j)}{\partial u^k} = (\nabla_k\, t^i_j)\, e_i \otimes e^j = v^i_{j|k}\, e_i \otimes e^j$$

As a useful example of this, let's find the coveriant derivative to metric tensor g_{ij}

$$
\begin{aligned}
\nabla_k\, g_{ij} &= \frac{\partial g_{ij}}{\partial u^k} - \Gamma^m_{ik}\, g_{mj} - \Gamma^m_{jk}\, g_{im} \\
&= \frac{\partial g_{ij}}{\partial u^k} - \Gamma_{ik;j} - \Gamma_{jk;i} \\
&= 0
\end{aligned}
$$

as we found when we were finding a formula for the Γ^i_{jk}. Note also

$$\nabla_k g = \frac{\partial g}{\partial g_{ij}}\, \nabla_k g_{ij} = 0$$

Now it is time to figure out why some indices are upstairs (contravariant) and some are downstairs (covariant) and what happens to a tensor's coefficients when the basis is changed. We will handle this for natural bases as this is traditional. We will also handle it for a simple example and leave it to the

[7]Some people feel that the names should be reversed, which would add more to a subject already awash in confusion. Anyway a good case can be made the the current choice is correct

reader to guess the general form, which is a trivial genralization. For your convenience we will rederive some things right here.

Recall that e_i is just a convenient notation for the tangent vector $\partial/\partial u^i$ and e^i is just convenient for du^i, which forms the dual basis to $\partial/\partial u^i$. Hence we see

$$\begin{aligned} \tilde{e}_i &= \frac{\partial}{\partial \tilde{u}^i} &= \frac{\partial u^j}{\partial \tilde{u}^i}\frac{\partial}{\partial u^j} &= \frac{\partial u^j}{\partial \tilde{u}^i} e_j \qquad \text{contravariant change} \\ \tilde{e}^i &= d\tilde{u}^i &= \frac{\partial \tilde{u}^i}{\partial u^j} du^j &= \frac{\partial \tilde{u}^i}{\partial u^j} e^j \qquad \text{covariant change} \end{aligned}$$

Now how does the tensor t^i_{jk} change. Well it is short for $T = t^i_{jk} e_i \otimes e^j \otimes e^k$ so it must change like

$$\begin{aligned} T = t^i_{jk} e_i \otimes e^j \otimes e^k &= t^i_{jk} \frac{\partial \tilde{u}^r}{\partial u^i}\tilde{e}_r \otimes \frac{\partial u^j}{\partial \tilde{u}^s}\tilde{e}^s \otimes \frac{\partial u^k}{\partial \tilde{u}^t}\tilde{e}^t \\ &= t^i_{jk} \frac{\partial \tilde{u}^r}{\partial u^i}\frac{\partial u^j}{\partial \tilde{u}^s}\frac{\partial u^k}{\partial \tilde{u}^t}\, \tilde{e}_r \otimes \tilde{e}^s \otimes \tilde{e}^t \end{aligned}$$

Since

$$T = \tilde{t}^r_{st}\, \tilde{e}_r \otimes \tilde{e}^s \otimes \tilde{e}^t$$

we have

$$\tilde{t}^r_{st}(\tilde{u}^m) = t^i_{jk}(u^p(\tilde{u}^m))\, \frac{\partial \tilde{u}^r}{\partial u^i}\frac{\partial u^j}{\partial \tilde{u}^s}\frac{\partial u^k}{\partial \tilde{u}^t}$$

where we have slipped in the change of variables in the arguments of of t^i_{jk} as well.

This is the tensor change law. In olden times a tensor was defined as an indexed quantity that changed in this manner; the mantra was *a tensor is an indexed quantity that changes like a tensor*, that is according to the above model. Mathematicians (in contrast to physicists) are no longer completely comfortable with this approach and so we have taken a different approach, which may clarify things for some people.

Now the big question is, is ${v^j}_{|k}$ a tensor; does the"deriviative index" k change properly? We will do this example that shows that this is indeed the case. You can use the same technioques to show that it works for any tensor. For convenience, it helps to modify the change of basis law for the connection coefficients a bit. We will derive this modification from scratch and the formulas may come in handy for other things.

$$\begin{aligned} e_k\Gamma^k_{ij} = \frac{De_i}{\partial u^j} = \frac{D}{\partial u^j}\Big(\tilde{e}_n\frac{\partial \tilde{u}^n}{\partial u^i}\Big) &= \frac{D}{\partial \tilde{u}^m}(\tilde{e}_n)\frac{\partial \tilde{u}^m}{\partial u^j}\frac{\partial \tilde{u}^n}{\partial u^i} + \tilde{e}_n\frac{\partial^2 \tilde{u}^n}{\partial u^j \partial u^i} \\ e_\ell\Gamma^\ell_{ij} - \tilde{e}_n\frac{\partial^2 \tilde{u}^n}{\partial u^j \partial u^i} &= \tilde{e}_p\tilde{\Gamma}^p_{mn}\frac{\partial \tilde{u}^m}{\partial u^j}\frac{\partial \tilde{u}^n}{\partial u^i} \\ e_\ell\Gamma^\ell_{ij} - e_\ell\frac{\partial u^\ell}{\partial \tilde{u}^n}\frac{\partial^2 \tilde{u}^n}{\partial u^j \partial u^i} &= e_\ell\frac{\partial u^\ell}{\partial \tilde{u}^p}\tilde{\Gamma}^p_{mn}\frac{\partial \tilde{u}^m}{\partial u^j}\frac{\partial \tilde{u}^n}{\partial u^i} \\ \Gamma^\ell_{ij} - \frac{\partial u^\ell}{\partial \tilde{u}^n}\frac{\partial^2 \tilde{u}^n}{\partial u^j \partial u^i} &= \frac{\partial u^\ell}{\partial \tilde{u}^p}\tilde{\Gamma}^p_{mn}\frac{\partial \tilde{u}^m}{\partial u^j}\frac{\partial \tilde{u}^n}{\partial u^i} \\ \Big(\Gamma^\ell_{ij} - \Big(\frac{\partial u^\ell}{\partial \tilde{u}^n}\frac{\partial^2 \tilde{u}^n}{\partial u^j \partial u^i}\Big)\Big)\frac{\partial \tilde{u}^p}{\partial u^\ell}\frac{\partial u^j}{\partial \tilde{u}^m}\frac{\partial u^i}{\partial \tilde{u}^n} &= \tilde{\Gamma}^p_{mn} \end{aligned}$$

Note that the first term on the left is a standard tensorial change; it is the second term we are interested in. It compensates for things that appear because the basis frame is moving. Keep your eye on how it works. We now show the ${v^j}_{|k}$ changes in a tensorial manner. This is straightforward but tedius, as is common with tensor calculations.

$$
\begin{aligned}
{\tilde{v}^\ell}_{|m} &= \frac{\partial \tilde{v}^\ell}{\partial \tilde{v}^m} + {\tilde{\Gamma}^\ell}_{nm}\tilde{v}^n \\
&= \frac{\partial}{\partial u^i}\Big(v^j \frac{\partial \tilde{u}^\ell}{\partial u^j}\Big)\frac{\partial u^i}{\partial \tilde{u}^m} + \bigg(\Gamma^k_{ij} - \Big(\frac{\partial u^k}{\partial \tilde{u}^p}\frac{\partial^2 \tilde{u}^p}{\partial u^j \partial u^i}\Big)\bigg)\frac{\partial \tilde{u}^\ell}{\partial u^k}\frac{\partial u^j}{\partial \tilde{u}^m}\frac{\partial u^i}{\partial \tilde{u}^n}\tilde{v}^n \\
&= \frac{\partial v^j}{\partial u^i}\frac{\partial \tilde{u}^\ell}{\partial u^j}\frac{\partial u^i}{\partial \tilde{u}^m} + v^j\frac{\partial^2 \tilde{u}^\ell}{\partial u^i \partial u^j}\frac{\partial u^i}{\partial \tilde{u}^m} + \Gamma^k_{ij}\frac{\partial \tilde{u}^\ell}{\partial u^k}\frac{\partial u^j}{\partial \tilde{u}^m}v^i - \frac{\partial \tilde{u}^\ell}{\partial u^k}\frac{\partial u^k}{\partial \tilde{u}^p}\frac{\partial^2 \tilde{u}^p}{\partial u^i \partial u^j}\frac{\partial u^j}{\partial \tilde{u}^m}v^i \\
&= \Big(\frac{\partial v^j}{\partial u^i} + \Gamma^j_{ki}v^k\Big)\frac{\partial \tilde{u}^\ell}{\partial u^j}\frac{\partial u^i}{\partial \tilde{u}^m} + v^j\frac{\partial^2 \tilde{u}^\ell}{\partial u^i \partial u^j}\frac{\partial u^i}{\partial \tilde{u}^m} - \frac{\partial^2 \tilde{u}^\ell}{\partial u^i \partial u^j}\frac{\partial u^j}{\partial \tilde{u}^m}v^i \\
&= {v^j}_{|i}\frac{\partial \tilde{u}^\ell}{\partial u^j}\frac{\partial u^i}{\partial \tilde{u}^m}
\end{aligned}
$$

which is the proper tensor change law. All such covariant derivatives of tensors can be proved to be tensors with exactly the same computation. We have presented this so that the reader can observe how the second derivative term in the change of variables formula just cancels the second derivative term which shows up in the derivative $\partial/\partial u^i(v^j\, \partial\tilde{u}^\ell/\partial u^j)\partial u^i/\partial \tilde{u}^m$.

1.4.3 Raising and Lowering Indices

In the presence of an inner product it is possible to shift covariant and contravariant indices in a tensor. It is usually necessary to be quite careful about keeping the indices in the same vertical strip as we move them up and down, and for clarity a dot may be used as a placeholder. For example, the Riemann Curvature Tensor may be written with the dot as

$$R_j{}^i{}_{k\ell} = R_j{}^{i}_{\cdot}{}_{k\ell}$$

where the dot makes it clear that i is above the second slot.

The raising and lowering is done with the g_{ij} and g^{ij} as follows.

$$R_{j\,i\,kl} = g_{im}\, R_j{}^{m}_{\cdot}{}_{k\ell}$$

The reason for doing this is that the modified tensor may have more congenial symmetries, or be easier to calculate, or a number of other reasons. Remember that this is just convenience; there is no real content here. Also note that we can always undo this

$$R_j{}^{m}_{\cdot}{}_{k\ell} = g^{mi}R_{j\,i\,kl}$$

Another example is the lowering of an index on a contrvariant vector to get a covariant vector.

$$v_j = g_{ij}v^i$$

Remember that g is symmetric, e.g. $g_{ij} = g_{ji}$, so the order of the indices doesn't matter. When this is done we then have the inner product written as

$$(v, w) = g_{ij}\, v^i w^j = v_j\, w^j$$

and this can be convenient.

1.4.4 Epsilon tensors

That the Kronecker delta is a tensor is a trivial matter and we leave it to the reader. The next tensor of somewhat similar type is the epsilon tensor, which we will discuss after some introductory remarks. Epsilon tensors were used to form dual tensors before the knowledge of the $*$-operator became common [8]. They are a rather clumsy tool to do the $*$-operator with, and the techniques used earlier in the book are much more convenient. However, in low dimensions they do have their uses.

Let $(i_1, i_2, \ldots, i_n)$ be a permutation of $(1, 2, \ldots, n)$ which we think ofa as a function $f(1) = i_1, f(2) = i_2, \ldots, f(n) = i_n$ and symbolize by

$$\sigma = \begin{pmatrix} 1 & 2 & ,\ldots, & n \\ i_1 & i_2 & ,\ldots, & i_n \end{pmatrix}$$

We symbolize the sign of this permutation by $\operatorname{sgn}(i_1, i_2, \ldots, i_n)$. Now recall that the determinant of an $n \times n$ matrix a_{ij} is a sum over all permutations σ of 1 to n

$$\begin{aligned} \det(a_{ij}) &= \sum_\sigma \operatorname{sgn}(\sigma)\, a_{1i_1} \cdots a_{ni_n} \qquad \text{or} \\ \det(a_{ij}) &= \sum_\sigma \operatorname{sgn}(\sigma)\, a_{i_1 1} \cdots a_{i_n n} \end{aligned}$$

and similarly for (a^i_j) and (a_{ij}).

We remind the reader that we assume that on each Tangent space $T_p(M)$ we have an inner product $(v, w)_p)$ which is smooth in p. For an embedded manifold this comes from the embedding space, but, as we will learn later, it can be supplied to the manifold by fiat, and this is called a (pseudo-)Riemannian manifold. It is pseudo-Riemannian if the inner product is non-degenerate but not positive definite, and we have consideredsuch cases previously, but will not consider them at present for simplicity. There is not much change from Riemannian to pseudo-Riemannian in the formalism. Of course there is a big change in the geometry.

For any basis $e_1, \ldots, e_n$ we have a matrix (g_{ij}) for the inner product given by $g_{ij} = (e_i,\, e_j)$ with inverse matrix $(g^{ij}) = (g_{ij})^{-1}$. Now recall Grassmann's theorem

$$(e_1 \wedge \ldots \wedge e_n,\, e_1 \wedge \ldots \wedge e_n) = \det\big((e_i,\, e_j)\big) = \det(g_{ij}) = g$$

[8]Although the work had already been done by Grassmann in the AUSDEHNUNGSLEHRE of 1861.

Now we will look at the situation when we change from one *natural* basis to another. We have

$$e_i = \frac{\partial}{\partial u^i} \qquad \tilde{e}_i = \frac{\partial}{\partial \tilde{u}^i}$$

$$\begin{aligned} e_i &= \frac{\partial}{\partial u^i} &= \frac{\partial \tilde{u}^j}{\partial u^i}\frac{\partial}{\partial \tilde{u}^j} &= \frac{\partial \tilde{u}^j}{\partial u^i}\,\tilde{e}_j \\ e_1 \wedge \ldots \wedge e_n &= \det\left(\frac{\partial \tilde{u}^j}{\partial u^i}\right)\tilde{e}_1 \wedge \ldots \wedge \tilde{e}_n \end{aligned}$$

Taking the inner product of each side with itself we get

$$g = \det\left(\frac{\partial \tilde{u}^j}{\partial u^i}\right)^2 \tilde{g}$$

$$\sqrt{\frac{g}{\tilde{g}}} = \det\left(\frac{\partial \tilde{u}^j}{\partial u^i}\right) \qquad \sqrt{\frac{\tilde{g}}{g}} = \det\left(\frac{\partial u^i}{\partial \tilde{u}^j}\right)$$

Now we want a tensor that acts rather like sgn but sgn itself does not quite transform like a tensor. It must be multiplied by a function as follows

$$\varepsilon_{i_1\ldots i_n} = \sqrt{g}\,\mathrm{sgn}(i_1\ldots i_n) \qquad \varepsilon^{i_1\ldots i_n} = \frac{\mathrm{sgn}(i_1\ldots i_n)}{\sqrt{g}}$$

Now we will check the tensor character of $\epsilon^{i_1\ldots i_n}$.

$$\begin{aligned} \tilde{\varepsilon}^{j_1,\ldots j_n}\frac{\partial u^{i_1}}{\partial \tilde{u}^{j_1}}\cdots\frac{\partial u^{i_n}}{\partial \tilde{u}^{j_n}} &= \frac{1}{\sqrt{\tilde{g}}}\det\begin{pmatrix} \frac{\partial u^{i_1}}{\partial \tilde{u}^1} & \cdots & \frac{\partial u^{i_1}}{\partial \tilde{u}^n} \\ \cdots & \cdots & \cdots \\ \frac{\partial u^{i_n}}{\partial \tilde{u}^1} & \cdots & \frac{\partial u^{i_n}}{\partial \tilde{u}^n} \end{pmatrix} \\ &= \frac{1}{\sqrt{\tilde{g}}}\mathrm{sgn}(i_1\cdots i_n)\det\begin{pmatrix} \frac{\partial u^1}{\partial \tilde{u}^1} & \cdots & \frac{\partial u^1}{\partial \tilde{u}^n} \\ \cdots & \cdots & \cdots \\ \frac{\partial u^n}{\partial \tilde{u}^1} & \cdots & \frac{\partial u^n}{\partial \tilde{u}^n} \end{pmatrix} \\ &= \frac{1}{\sqrt{\tilde{g}}}\mathrm{sgn}(i_1\cdots i_n)\sqrt{\frac{\tilde{g}}{g}} = \frac{1}{\sqrt{g}}\mathrm{sgn}(i_1\ldots i_n) \\ &= \varepsilon^{i_1\ldots i_n} \end{aligned}$$

1.4.5 Epsilon tensors and Dual tensors in Two Dimensions

We can use the epsilon tensors to construct dual tensors although the process is clumsy. However, in two dimensions it works rather well. Let $v = e_i\,v^i$ Then $*v$ is defined by the usual relation

$$w \wedge *v = (w,\, v)\Omega_0$$

for all $w \in T(M)$ where $\Omega_0 = e_1 \wedge e_2/||e_1 \wedge e_2|| = e_1 \wedge e_2/\sqrt{g}$. We claim

$$*w = e_j\,(\varepsilon_k^{\cdot j}\,v^k) = e_j\,(g_{ki}\,\varepsilon^{ij}\,v^k)$$

This can be done by a calculation but since the dimension is 2 it is interesting to look at what happens in detail; to see the gears whirr and clank. To do this we first determine $\varepsilon_k^{\cdot j} = g_{ik}\,\varepsilon^{kj}$.

$$\varepsilon_1^{\cdot 1} = g_{1k}\varepsilon^{k1} = -g_{12}\frac{1}{\sqrt{g}} \qquad \varepsilon_2^{\cdot 1} = g_{2k}\varepsilon^{k1} = -g_{22}\frac{1}{\sqrt{g}}$$
$$\varepsilon_1^{\cdot 2} = g_{1k}\varepsilon^{k2} = g_{11}\frac{1}{\sqrt{g}} \qquad \varepsilon_2^{\cdot 2} = g_{2k}\varepsilon^{k2} = g_{21}\frac{1}{\sqrt{g}}$$

Now we compute $*v$

$$\begin{aligned} *v &= e_j\,\varepsilon_i^{\cdot j}\,v^k \\ &= e_1\,(\varepsilon_1^{\cdot 1}\,v^1 + \varepsilon_2^{\cdot 1}\,v^2) + e_1\,(\varepsilon_1^{\cdot 2}\,v^1 + \varepsilon_2^{\cdot 2}\,v^2) \\ &= \frac{1}{\sqrt{g}}\Big[e_1\,(-g_{12}\,v^1 - g_{22}\,v^2) + e_2\,(g_{11}\,v^1 + g_{21}\,v^2)\Big] \end{aligned}$$

Next we calculate

$$\begin{aligned} w \wedge *v &= (e_1 w^1 + e_2 w^2) \wedge \frac{1}{\sqrt{g}}\Big[e_1\,(-g_{12}\,v^1 - g_{22}\,v^2) + e_2\,(g_{11}\,v^1 + g_{21}\,v^2)\Big] \\ &= \Big[w^1(g_{11}\,v^1 + g_{21}\,v^2) - w^2(-g_{12}\,v^1 - g_{22}\,v^2)\Big]\frac{e_1 \wedge e_2}{\sqrt{g}} \\ &= \big(g_{ij}\,v^i w^j\big)\,\Omega_0 \;=\; (v,w)\,\Omega_0 \end{aligned}$$

which proves our assertion $*v = e_j\,(g_{ki}\,\varepsilon^{ij}\,v^k)$.

Now recall that

$$**\,v = (-1)^{\deg(v)(\dim(V)-\deg(v))} = (-1)^{1(2-1)}v = -v$$

This relation is also easily proved by using the above formula for $*v$ twice. We suggest you do this so you can see how the factor $1/\sqrt{g}$ once again saves the day. Using this we can show $||*v|| = ||v||$ and that $v, *v$ form an oriented orthonormal basis for $V = T_p(M)$. Indeed we have

$$(*v, *v)\,\Omega_0 = *v \wedge **\,v = *v \wedge (-v) = v \wedge *v = (v,v)\,\Omega_0$$

So, $\Lambda^2(V)$ being one-dimensional, we have $(*v, *v) = (v,v)$. Also note the formula $v \wedge *v = (v,v)\,\Omega_0$ which, since $(v,v) > 0$, tells us that v, $*v$ is oriented like e_1, e_2.

Finally,

$$(v, *v)\,\Omega_0 = v \wedge **\,v = v \wedge (-v) = -v \wedge v = 0$$

so v is perpedicular to $*v$.

1.4.6 The Riemann Curvature Tensor in Two Dimenions

We will now look at some aspects of the Riemann curvature tensor for a two dimensional manifold embedded in 3-space. We will see that much information can be extracted from the symmetries of the coefficients. We recall some previous equations to remind the reader who the actors are. First, recall that

$$\frac{\partial e_i}{\partial u^j} = e_k \Gamma^k_{ij} + b_{ij}\, n$$

where e_1, e_2 is a basis of $T_p(M)$ and n is the unit normal vector. The Riemann curvature tensor we remember is

$$R_i{}^\ell{}_{mj} = \frac{\partial \Gamma^\ell_{ij}}{\partial u^m} - \frac{\partial \Gamma^\ell_{im}}{\partial u^j} + \Gamma^\ell_{km}\Gamma^k_{ij} - \Gamma^\ell_{kj}\Gamma^k_{im}$$

The first equation following is the equation of Gauss and it will be critical for us. In the second equation we lower the contravariant index.

$$\begin{aligned} R_i{}^l{}_{mj} &= b^l_m b_{ij} - b^l_j b_{im} \\ R_{ikmj} = g_{kl} R_i{}^l{}_{mj} &= g_{kl} b^l_m b_{ij} - g_{kl} b^l_j b_{im} = b_{ij} b_{km} - b_{im} b_{kj} \end{aligned}$$

The second equation allows us to easily determine some symmetries of the Riemann curvature tensor. It is not available in general Riemannian Geometry so there the symmetries are a little harder to prove. First note, interchangeing the final two indices j and m, that

$$R_{ikjm} = b_{im} b_{kj} - b_{ij} b_{km} = -(b_{ij} b_{km} - b_{im} b_{kj}) = -R_{ikmj}$$

so R_{ikmj} is antisymmetric in the last two indices. (We already knew this from earlier theory.) Next we swap the first two indices i and k to get

$$R_{kimj} = b_{kj} b_{im} - b_{km} b_{ij} = -(b_{ij} b_{km} - b_{im} b_{kj}) = -R_{ikmj}$$

so R_{ikmj} is also antisymmetric in the first two indices. Then of course

$$R_{kimj} = -R_{ikmj} = R_{ikjm}$$

Now we are going to swap the first *pair* of indices ik with the second *pair* of indices jm

$$R_{mjik} = b_{mk} b_{ji} - b_{mi} b_{jk} = b_{ij} b_{km} - b_{im} b_{kj} = R_{ikmj}$$

So swapping the first pair and the second pair leaves R_{ikmj} invariant. Now we know the value of one set of indices from the the theorem egregium of Gauss. We have

$$R_{1221} = b_{11} b_{22} - b_{12} b_{21} = \det(b_{ij})$$

But we know that

$$K = \frac{\det(b_{ij})}{\det(g_{ij})} = \frac{R_{1221}}{g}$$

so

$$R_{1221} = gK$$

Using the antibymmetries found above we can now write down the complete set of values of R_{ikmj}. They are

$$\begin{array}{llll} R_{1111}=0 & R_{1112}=0 & R_{1121}=0 & R_{1122}=0\\ R_{1211}=0 & R_{1212}=-gK & R_{1221}=gK & R_{1222}=0\\ R_{2111}=0 & R_{2112}=gK & R_{2121}-gK & R_{2122}=0\\ R_{2211}=0 & R_{2212}=0 & R_{2221}=0 & R_{2222}=0 \end{array}$$

Because of the antisymmetries it is typical of the Riemann curvature tensor to have a large number of zero components. The fully covariant Riemann curvature tensor is simpler than the original tensor. For example

$$R_1{}^2{}_{21} = g^{2\ell}R_{1\ell\,21} = g^{21}R_{11\,21} + g^{22}R_{1221} = g^{22}gk$$

We will need the formula $(\nabla_k\nabla_j - \nabla_j\nabla_k)\,v^i = R_m{}^i{}_{kj}\,v^m$. Indeed this is how the Gaussian Curvature K enters into the Gauss-Bonnet theorem. To get a little tensor practise we will derive the formula with tensor methods.

$$\begin{aligned}
\nabla_j v^i &= \frac{\partial v^i}{\partial u^j} + \Gamma^i_{mj}v^m = v^i{}_{|j}\\
\nabla_k\, v^i{}_{|j} &= \frac{\partial}{\partial u^k}\, v^i{}_{|j} + \Gamma^i_{nk}v^n{}_{|j} - \Gamma^\ell_{jk}v^i{}_{|\ell}\\
&= \frac{\partial}{\partial u^k}\Big(\frac{\partial v^i}{\partial u^j} + \Gamma^i_{mj}v^m\Big) + \Gamma^i_{nk}\Big(\frac{\partial v^n}{\partial u^j} + \Gamma^n_{mj}v^m\Big) - \Gamma^\ell_{jk}\Big(\frac{\partial v^i}{\partial u^\ell} + \Gamma^i_{m\ell}v^m\Big)\\
&= \frac{\partial^2 v^i}{\partial u^k\partial u^j} + \frac{\partial\Gamma^i_{mj}}{\partial u^k}\,v^m + \Gamma^i_{mj}\frac{\partial v^m}{\partial u^k} + \Gamma^i_{nk}\frac{\partial v^n}{\partial u^j} - \Gamma^\ell_{jk}\frac{\partial v^i}{\partial u^\ell}\\
&\quad + \Big(\Gamma^i_{nk}\Gamma^n_{mj} - \Gamma^\ell_{jk}\Gamma^i_{m\ell}\Big)\,v^m
\end{aligned}$$

Now we swap j and k and hope a lot of stuff cancels out.

$$\begin{aligned}
\nabla_j\, v^i{}_{|k} &= \frac{\partial^2 v^i}{\partial u^j\partial u^k} + \frac{\partial\Gamma^i_{mk}}{\partial u^j}\,v^m + \Gamma^i_{mk}\frac{\partial v^m}{\partial u^j} + \Gamma^i_{nj}\frac{\partial v^n}{\partial u^k} - \Gamma^\ell_{kj}\frac{\partial v^i}{\partial u^\ell}\\
&\quad + \Big(\Gamma^i_{nj}\Gamma^n_{mk} - \Gamma^\ell_{kj}\Gamma^i_{m\ell}\Big)\,v^m
\end{aligned}$$

The terms involving derivatives of v^i cancel out in pairs, as do two of the terms multipling v^m leaving us with

$$\begin{aligned}
(\nabla_k\nabla_j - \nabla_j\nabla_k)\,v^i &= \Big(\frac{\partial\Gamma^i_{mj}}{\partial u^k} - \frac{\partial\Gamma^i_{mk}}{\partial u^j} + \Gamma^i_{nk}\Gamma^n_{mj} - \Gamma^i_{nj}\Gamma^n_{mk}\Big)\,v^m\\
&= R_m{}^i{}_{kj}\,v^m = R_m{}^i{}_{kj}\,v^m
\end{aligned}$$

This can also be written, where we abbreviate $v^i{}_{|j|k}$ as $v^i{}_{|jk}$, (and note the order)

$$v^i{}_{|jk} - v^i{}_{|kj} = R_m{}^i{}_{kj}\,v^m$$

The advantage of the differential form method for deriving these things is that often terms that cancel out in the tensor derivation never appear at all in the differential form derivation.

1.5 General Manifolds and Connections

In this short section we introduce the idea of Connections on the tangent bundle of an arbitrary differentiable manifold. This is easy. A connection on the tangent bundle is often called an affine connection. It is trivial to put connections on any vector bundle on a manifold, and we will eventually talk about this.

A manifold is covered by a set of coordinate patches. We choose one of these patches with coordinates $u^1, \dots, u^n$. Recall that in this case the basis vectors $\mathbf{e}_i$ are defined to be the differential operators $\frac{\partial}{\partial u^i}$ but we will continue to call them $\mathbf{e}_i$ so things will look familiar. In a few places we have to dredge up $\mathbf{e}_i = \frac{\partial}{\partial u^i}$ but not often.

We now introduce *on this coordinate patch* a Covariant Differential D by setting

$$De_i = \mathbf{e}_j\,\omega_i{}^j = \mathbf{e}_j\,\Gamma^j_{ik}\,du^k$$

Now you ask, what are the conditions on the Γ^j_{ik} for this to be a connection and the answer is none whatever. You pick any Γ^j_{ik} *on this patch* and you have a connection *on this patch*. You may choose the connection according to some mathematical or physical situation or by pure whimsey. **However**, once you have chosen it for this patch it will migrate to the overlaps of other patches with this patch, and your choice of connection on other patches is constrained by this. This is very important. Let us now look at the overlap of our patch with a nearby patch which has coordinates $\tilde{u}_1, \dots, \tilde{u}_n$. This is one of those moments when we need to remember what e_i actually is. We have

$$\begin{aligned}
\tilde{\mathbf{e}}_j &= \frac{\partial}{\partial \tilde{u}^j} = \frac{\partial u^i}{\partial \tilde{u}^j}\frac{\partial}{\partial u^i} = \mathbf{e}_i \frac{\partial u^i}{\partial \tilde{u}^j} \\
(\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n) &= (\mathbf{e}_1, \dots, \mathbf{e}_n)\Big(\frac{\partial u^i}{\partial \tilde{u}^j}\Big) \\
\vec{\tilde{\mathbf{e}}} &= \vec{\mathbf{e}}\,C
\end{aligned}$$

where C is the $n \times n$ matrix

$$C = \Big(\frac{\partial u^i}{\partial \tilde{u}^j}\Big)$$

Quantities that change in this way are called coveriant vectors, have low indices and are written as rows.

Consider now a vector $\mathbf{v} = \mathbf{e}_i v^i$. We can write this as

$$\mathbf{v} = \mathbf{e}_i v^i = (\mathbf{e}_1, \dots, \mathbf{e}_n)\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} = \vec{\mathbf{e}}\,\vec{v}$$

To change bases, we compute as follows

$$\begin{aligned}
\vec{\tilde{\mathbf{e}}}\,\vec{\tilde{v}} = \mathbf{v} &= \vec{\mathbf{e}}\,\vec{v} \\
\vec{\mathbf{e}}\,C\,\vec{\tilde{v}} &= \vec{\mathbf{e}}\vec{v}
\end{aligned}$$

$$\begin{aligned} C\,\vec{\tilde{v}} &= \vec{v} \\ \frac{\partial u^i}{\partial \tilde{u}^j}\,\tilde{v}^j &= v^i \end{aligned}$$

Quantities that change in this way are called contravariant vectors, have high indices, and are written as columns.

Now we return to the change in the connection coefficients due to the change of coordinates. We are so focused on change of coordinates for the reasons discussed above and we also want our equations not to depend on particular coordinate choices, but to be valid whatever coordinates we use. The system is set up to take care of all this by itself. We have

$$D\vec{\tilde{\mathbf{e}}} = \vec{\tilde{\mathbf{e}}}\,\tilde{\omega} \qquad D\vec{\mathbf{e}} = \vec{\mathbf{e}}\,\omega \qquad \vec{\tilde{\mathbf{e}}} = \vec{\mathbf{e}}\,C$$

Thus we have

$$\begin{aligned} \vec{\tilde{\mathbf{e}}} &= \vec{\mathbf{e}}\,C \\ D\vec{\tilde{\mathbf{e}}} &= D(\vec{\mathbf{e}}\,C) \\ \vec{\tilde{\mathbf{e}}}\,\tilde{\omega} &= \vec{\mathbf{e}}\,dC + D(\vec{\mathbf{e}})\,C \\ \vec{\mathbf{e}}\,C\,\tilde{\omega} &= \vec{\mathbf{e}}\,dC + \vec{\mathbf{e}}\,\omega\,C \\ C\,\tilde{\omega} &= dC + \omega\,C \end{aligned}$$

so we have

$$\boxed{\tilde{\omega} = C^{-1}\,dC + C^{-1}\,\omega\,C}$$

This is an important formula and the way this change of variable is often described in advanced books. Make sure you understand it. If you have had experience with tensors, which we have not emphasized in the book, note that the formula is made up of two parts, the second of which is a standard tensorial formula. The first part, is not tensorial, and is what keeps the Christoffel symbols from being tensors.

Now we just have to decode this, noting that $C^{-1} = (\frac{\partial \tilde{u}^j}{\partial u^i})$ and that

$$dC = d\Big(\frac{\partial u^i i}{\partial \tilde{u}^j}\Big) = \Big(\frac{\partial^2 u^i}{\partial \tilde{u}^j \partial u^k}\,du^k\Big)$$

so the above formula decodes as

$$\tilde{\omega}_s{}^r = \frac{\partial \tilde{u}^r}{\partial u^i}\frac{\partial^2 u^i}{\partial u^j \partial u^k}\,du^k + \frac{\partial \tilde{u}^r}{\partial u^i}\,\omega_j{}^i\,\frac{\partial u^j}{\partial \tilde{u}^s}$$

which is nice to know. However, we would also like to go down another level and see how the formula is expressed in terms of the Christoffel symbols Γ^i_{jk}. This is intricate but easy. We replace the $\omega_s{}^r$ by their formulas in terms of the Christoffel symbols to get

$$\begin{aligned} \tilde{\Gamma}^r_{sm} d\tilde{u}^m &= \frac{\partial \tilde{u}^r}{\partial u^i}\frac{\partial^2 u^i}{\partial u^j \partial u^k}\,du^k + \frac{\partial \tilde{u}^r}{\partial u^i}\,\Gamma^i_{jk} du^k\,\frac{\partial u^j}{\partial \tilde{u}^s} \\ &= \frac{\partial \tilde{u}^r}{\partial u^i}\frac{\partial^2 u^i}{\partial u^j \partial u^k}\,\frac{\partial u^k}{\partial \tilde{u}^m}\,d\tilde{u}^m + \frac{\partial \tilde{u}^r}{\partial u^i}\,\Gamma^i_{jk}\,\frac{\partial u^k}{\partial \tilde{u}^m}\,d\tilde{u}^m\,\frac{\partial u^j}{\partial \tilde{u}^s} \\ \tilde{\Gamma}^r_{sm} &= \frac{\partial \tilde{u}^r}{\partial u^i}\frac{\partial^2 u^i}{\partial u^j \partial u^k}\,\frac{\partial u^k}{\partial \tilde{u}^m} + \frac{\partial \tilde{u}^r}{\partial u^i}\,\Gamma^i_{jk}\,\frac{\partial u^j}{\partial \tilde{u}^s}\,\frac{\partial u^k}{\partial \tilde{u}^m} \end{aligned}$$

In differential geometry books this is often presented in the form

$$\tilde{\Gamma}^r_{sm} = \frac{\partial^2 u^i}{\partial u^j \partial u^k}\frac{\partial \tilde{u}^r}{\partial u^i}\frac{\partial u^k}{\partial \tilde{u}^m} + \Gamma^i_{jk}\frac{\partial \tilde{u}^r}{\partial u^i}\frac{\partial u^j}{\partial \tilde{u}^s}\frac{\partial u^k}{\partial \tilde{u}^m}$$

but the first form is preferable since it reflects the matrix multiplications from which the formula comes.

We have now got the coefficients of the connection on the Tangent Bundle on the overlap of the original patch with each neighbouring patch. Next, on each such patch, we construct Γ^i_{jk} arbitrarily except for the critical constraint that on the overlaps it must coincide with the transform of the original Γ^i_{jk} to the new coordinates. Care must of course be taken where more than two coordinate patches overlap, but it turns out that because of the form of the transformation this is automatically taken care of. However annoying in practise there is no problem in principle. However, this is a very crude method of introducing a global connection on the manifold and we will soon see a much better method. However, the better method requires a digression through a bit of Lie Groups and Lie Algebras.

The basic idea of a connection is to allow us to take the derivative of sections of a vector bundle. In our case the bundle is the tangent bundle. Thus we need a machine into which we insert a section, and then a tangent vector, and the outcome is the directional derivative of the section in the direction of the tangent vector. This amounts to having an operator ∇ so that for a section $s : U \to T_(M)$ (think of U as a coordinate patch and for $p \in U$, $s(p) \in T_p(M)$ is a tangent vector at p.) Then ∇s is a function which eats tangent vectors $X \in T_p(M)$ and spits out a real number; the directional derivative of s in the direction X. Thus $\nabla s(X)$ is a real number giving the change of s in the direction X. This is often written $\nabla_X s$. Now the only question is how to *compute* $\nabla s(X)$,

Let us now discuss why what has occurred to you won't work. The natural thing to want to do is to set

$$\nabla s(X) = \lim_{t\to 0} \frac{s(p+tX) - s(p)}{t} \qquad \text{WRONG!!}$$

This has two defects. First, $p + tX$ makes no sense on a general manifold. One might attempt to get around this by using a flow, but that won't work either because $s(p + tX)$ and $s(p)$ are in different tangent spaces so how are we going to subtract them? Nothing like this is going to work directly in this context. The notion of parallel transfer was invented to overcome this problem, but this is just another way of introducing a connection.

We remind the reader that there is a function π from $T(M)$ to M which sends a tangent vector $Y \in T_p(M)$ to p ; we have $\pi(Y) = p$. Then a *section* s can be defined as a function on an open set $U \subseteq M$ to $T(M)$ for which

$$\pi(s(p)) = p \qquad\qquad \pi \circ s = I$$

The set of sections of U is denoted by $\Gamma(T(M))(U)$ [9]. We will usually omit the

[9]This has no relation to the use of Γ in Γ^i_{jk} as connection coefficients, but hopefully (the standard wish) no confusion will arise.

(U) since it is always there. Keep this in mind.

With this equipment we can now say that ∇ is a function

$$\nabla : \Gamma(T(M)) \to \Gamma(T(M)) \otimes \Gamma(T^*(M))$$

See if you can decode this based on what was said above.

A manifold with a connection on the Tangent Bundle (referred to as an affine connection) is called an Affine Manifold or (older term) an Affinely related manifold. As we said before, given a differential manifold we may put an affine connection on it patch by patch being careful that the coefficients transform properly on the overlaps. This having been done, almost all the remaining work coincides with what we found in the embedded case, which I will now recall.

On a coordinate patch we have coordinates $u^1, \ldots, u^n$ and a local basis of tangent vectors of the Tangent Bundle

$$e_1 = \frac{\partial}{\partial u^1} \quad \ldots \quad e_n = \frac{\partial}{\partial u^n}$$

Because we have been given an affine connection we can then write

$$De_i = e_k \Gamma^k_{ij}\, du^j$$

This was the basic formula from which we derived all the other formulas. We set

$$\omega_i^{\ k} = \Gamma^k_{ij}\, du^j \qquad \omega = (\omega_i^{\ k})$$

and then we have

$$\Omega = d\omega + \omega \wedge \omega$$

with

$$\Omega = (R_i^{\ j}{}_{k\ell}\, du^k \wedge du^\ell)$$

with the components $R_i^{\ j}{}_{k\ell}$ of the Riemann Curvature tensor. The same calculations as in the embedded case give us just as before

$$R_j^{\ i}{}_{k\ell} = \frac{\partial \Gamma^i_{j\ell}}{\partial u^k} - \frac{\partial \Gamma^i_{jk}}{\partial u^\ell} + \Gamma^i_{mk}\Gamma^m_{j\ell} - \Gamma^i_{m\ell}\Gamma^m_{jk}$$

We can then calculate, for s a local section over an open subset of the coordinate patch

$$\begin{aligned} Ds = D(e_i s^i) &= e_i ds^i + De_i s^i \\ &= e_k ds^k + e_k \Gamma^k_{ij} s^i\, du^j \\ &= e_k (ds^k + \Gamma^k_{ij} s^i\, du^j) \\ &= e_k \Big(\frac{\partial s^k}{\partial u^j} + \Gamma^k_{ij} s^i \Big) du^j \\ &= e_k s^k{}_{|j}\, du^j \end{aligned}$$

where

$$s^k{}_{|j} = \frac{\partial s^k}{\partial u^j} + \Gamma^k_{ij} s^i$$

and for a tangent vector $X \in T(M)$ the directional derivative of s in the X direction is

$$D_X s = Ds(X) = e_k s^k{}_{|j}\, du^j \left(e_\ell X^\ell\right) = e_k s^k{}_{|j} \delta^j_\ell X^\ell = e_k s^k{}_{|j} X^j$$

and we have now solved the problem of taking derivatives of sections on an Affine Manifold. Thus our D may serve for the ∇ discussed above, and for sections s we can use D_s for ∇_s. We introduced ∇ because in many books it is used instead of D and we thought it might be useful for the reader to have seen it in this context should she meet it elsewhere.

To complete our study of affine connections we will need to get a deeper insight into what a connection is globally and this requires a look at parallel displacement along a curve and a small excursion into Lie Groups and Lie Algerbras.

1.6 Parallel Displacement Along Curves

Although we have not previously mentioned parallel displacement, it is an important idea and indeed the whole theory can be based on it. This is a matter of personal preference. We will need a little of the theory to put connections into a more abstract setting. Fortunately there is not much difficulty with this subject and it is rather interesting.

The idea of parallel displacement of a vector along a curve C is that, to the extent possible, the vector (in the tangent bundle) does not change as we move along the curve. This is of course an impossible requirement because the various vectors lie in different tangent spaces; if $u^i(t)$ describes the curve and $w(t)$ the vectors along the curve then $w(t) \in T_{p(t)}M$ and so vectors at different points cannot be directly compared. However, we already have equipment which will allow us to express this idea conveniently with the covariant derivative.

Let $U \subseteq M$ be a coordinate patch and C a curve in the patch given by $u^1(t), \ldots, u^n(t)$. Let $p(t)$ be the point on the manifold given by $u^i(t)$ and $w(t) \in T_{p(t)}M$ be a vector at each point of the curve. Let $e_i = \frac{\partial}{\partial u^i}$ be the natural basis. Since we do not want $w(t)$ to change, the obvious requirement is that

$$\frac{Dw}{dt} = 0$$

We may write this out in coordinates as

$$\begin{aligned}
\frac{Dw}{dt} = \frac{D}{dt}(e_i w^i) &= e_i \frac{dw^i}{dt} + \frac{De_i}{\partial u^j}\frac{du^j}{dt} w^i \\
&= e_k \frac{dw^k}{dt} + e_k \Gamma^k_{ij} w^i \frac{du^j}{dt} \\
&= e_k \left(\frac{dw^k}{dt} + \Gamma^k_{ij} w^i \frac{du^j}{dt} \right)
\end{aligned}$$

Thus for the vector field w along C we have

Def The vector field $w(t)$ is parallel along $C \Longleftrightarrow$

$$\frac{dw^k}{dt} + \Gamma^k_{ij} w^i \frac{du^j}{dt} = 0$$

An amusing question is to seek the condition that the tangent vector to the curve is itself parallel along its own curve. In this case $w^i = \frac{du^i}{dt}$ and the condition becomes

$$\frac{d^2u^i}{dt^2} + \Gamma^k_{ij} \frac{du^i}{dt} \frac{du^j}{dt} = 0$$

which you will recall is the equation which says that C is a geodesic. That is, a geodesic C is characterized by its tangent vector being parallel along itself. This is good evidence that the notion of parallel displacement is an important idea, though we will not make a lot of use for it in this book.

Now let us look at this in a slightly different way. Consider a tangent vector $w \in T_pM$ to the manifold at the point $p = p(0)$ of C. We would like to manufacture a vector field $w(t) \in T_{p(t)}M$ along the curve C which is parallel along C and for which $w(0) = w$. This is called *parallely transporting the vector w along C* or *parallely translating the vector w along C* and each $w(t)$ is called a *parallel transport of w along C*. Now we ask, is it possible always to do this. For it to happen, we must satisfy the differential equation an initial condition

$$\frac{dw^k}{dt} + \Gamma^k_{ij} w^i \frac{du^j}{dt} = 0 \qquad\qquad w^i(0) = w^i$$

Since this is a linear system of differential equations with C^∞ coefficients there is always a unique solution. Thus parallel transport is always possible for any curve C, and it is easy to shift coordinates as we go across coordinate patch boundaries.

Now here is a critical fact. Suppose we have two curves from point p to point q on the manifold and we transport a tangent vector w at p along both curves. Will they coincide? No, in general *they will not coincide.* If they do coincide for any two curves it indicates a flatness in the manifold in an area containing p and q. It is then no big jump to guess that the failure of the transports to coincide is related to the Riemann Curvature Tensor. Sadly we will not go into this in any further detail.

Notice that parallel transport from p to q sets up an isomorphism (because the transport equation is linear) between T_pM and T_qM. This seems to be the origin of the term "connection" since the connection connects the two tangent spaces. However, because the isomorphism depends on the choice of curve this is not as exciting as it looks at first glance.

Looked at another way, if we take two curves C_1 and C_2 from p to q we can also think of w as translated back to p by going around the combined curve $C_1 - C_2$. This will then give an automorphism of T_pM into itself, and thus a mapping from curves C to $\mathrm{AUT}(T_p(M))$, the space of automorphisms of $T_p(M)$. This is another object of study.

In some treatments of differential geometry the concept of parallel displacement is given center stage and other concepts are then derived from it. One reason is because it can be easily visualized, and this is helpful. For example, on a two dimensional manifold embedded in three space and with a curve parametrized by arc length, parallel translation is easily visualized; starting with a vector w at p and a curve C with points $p(t)$ find the angle between w and the tangent vector to the curve C at $p(0)$ and maintain this same angle as you move along C. The result will be the parallel transport of w along C. We will discuss this further when we get to Riemannian geometry where the notion of angle is again available. (The need for the angle is why I had to return to the embedded case.) Our need for parallel transport is to produce certain Lie Groups from which we will get certain Lie Algebras which will then be used to justify the mantra *a connection is a Lie Algebra valued differential form.*

1.7 A little about Lie Groups and Lie Algebras

We will need a short introduction to Lie Groups and Lie Algebras in order to put connections into a more abstract setting. This is a vast and beautiful subject but we can only give the shortest possible introduction here. Fortunately only some very basic concepts are necessary for our purposes. At the end of the section we will give a short bibliography for those who wish to learn more about this central subject of modern mathematics.

Lie Groups were among the first crossover abstract structures which means structures combining two fundamental structures, in this case Groups and Differentiable Manifolds. Crossover structures have been a great source of interesting mathematics.

Lie Groups are defined as follows

Def G is a Lie Group $\Longleftrightarrow$

1. G is a group

2. G is a (separable) differentiable manifold

3. The operations $(a, b) \to ab$ and $a \to a^{-1}$ are C^∞ functions.

This marriage of group and manifold theory leads to a theory of great subtlety and beauty and occasional surprising difficulty for some of its important theorems.

While there is a beautiful theory of abstract Lie Groups, it is a little more difficult than the special case of Lie Groups whose elements are matrices. A Lie Group whose elements are $n \times n$ matrices is a subgroup of the group $\mathrm{GL}(n, \mathbb{R})$ of *invertible* $n \times n$ matrices. There are also complex Lie Groups like $\mathrm{GL}(n, \mathbb{C})$ whose entries are complex rather than real numbers but we will not be using them. There are also infinite dimensional Lie Groups and this is an area of active research which we will not even glance at.

Our favorite two example of Lie Groups are $\mathrm{GL}(n,\mathbb{R})$ and $\mathrm{SO}(n,\mathbb{R})$. We define $\mathrm{SO}(n,\mathbb{R})$ by

Def $A \in \mathrm{SO}(n,\mathbb{R}) \Longleftrightarrow$

1. $AA^\top = I$

2. $\det A = +1$

$\mathrm{SO}(n,\mathbb{R})$ is called the *special orthogonal group* and it is the second condition that makes it *special*. The *orthogonal group* $\mathrm{O}(n,\mathbb{R})$ is defined by only the first condition $AA^\top = I$.

A lie Group is connected if and only if it is connected as a topological space. $\mathrm{O}(n,\mathbb{R})$ is not connected; it has two componenents, one with $\det A = 1$ and the other, a coset, with $\det A = -1$. (It is not completely obvious that the former is connected.) The component of the identity is always itself a Lie Group and a subgroup of the original group with discrete factor group. In the present case the component of the identity is $\mathrm{SO}(n,\mathbb{R})$. A Lie subgroup H of a Lie Group G is an algebraic subgroup whose inclusion map $\iota : H \to G$ is an immersion, which means $d\iota$ is onto. We do NOT want to require that that the topology of H is the same as the relative topology of H as a subset of G, as this would exclude many interesting and useful examples. This is one of the subtleties of Lie theory.

Now we are going to introduce the Lie Algebra of G. This turns out to be $T_I(G)$ but we will be more explicit about it. Let $A(t)$ be a C^∞ path in G where $t \in [a,b]$ where $a < 0 < b$ and $A(0) = I$. Then $X = A'(0)$ will be a tangent vector at I to G and the set of all such tangent vectors will be the Lie Algebra $\mathfrak{g}$ of G. (It is customary to use the lower case fraktur letters for Lie Algebras.)

It is very important that the tangent space $T_I(G)$ is not just a vector space; it has a product structure on it. For matrix Lie Algerbras this is given by

$$[X,Y] = XY - YX$$

It is called the *Lie Bracket* and naturally we must show that if $X, Y \in \mathfrak{g}$ then $[X,Y] \in \mathfrak{g}$ which is not obvious. It is easy to prove the following two identities for Matrix Lie Algebras by direct computation.

$$\begin{aligned} [X,Y] &= -[Y,X] && \text{antisymmetry} \\ [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X.Y]] &= 0 && \text{the Jacobi Identity} \end{aligned}$$

We now prove that, for matrix Lie Groups G, $[X,Y] \in \mathfrak{g}$. To do this (and for purposes of general education) we note that G acts on $\mathfrak{g}$ as we now explain. for $h, g \in G$, we have $ghg^{-1} \in G$. Now let $Y \in \mathfrak{g}$ and $Y = \frac{dh(t)}{dt}\Big|_{t=0}$ for some path $h(t)$ in G where $h(0) = I$. Then $gh(t)g^{-1}$ is a path in G and thus

$$gYg^{-1} = \frac{d}{dt}(gh(t)g^{-1})\Big|_{t=0} \in \mathfrak{g}$$

This is important enough to get a special name

Def $\mathrm{Ad}(g)(Y) = gYg^{-1}$

$\mathrm{Ad}(g)$ is an endomorphism[10] of $\mathfrak{g}$ which is vector space. Thus

$$\mathrm{Ad} : G \to \mathrm{End}(\mathfrak{g})$$

where $\mathrm{End}(V)$ is the ring of all endomorphism of V. A homomorphism of a group G into the ring of endomorphisms of a vector space is called a *representation*. By choosing a basis for V we may take the matrix representations of the endomorphisms and we get *matrix representation*. Since $\mathfrak{g}$ is a vector space we may choose a basis of $\mathfrak{g}$ and then, in the usual way, $\mathrm{Ad}(G)$ will be represented in this basis as a matrix and so we have a homomorphism of G into a set of matrices. Such a homomorphism is referred to as a *matrix representation of* G. Ad: $G \to \mathrm{End}(\mathfrak{g})$ is called the *Left Regular Represetation* of G and the same name is used for the corresponding matrix representation.

Now we want to show that for $X, Y \in \mathfrak{g}$ we have $[X, Y] = XY - YX \in \mathfrak{g}$. To do this, we take $g(t)$ a path in G with $X = \frac{d}{dt}g(t)\big|_{t=0}$ and $g(0) = I$. Then we have

$$\begin{aligned}
\mathrm{Ad}(g(t))Y &= g(t)Yg(t)^{-1} \\
\frac{d}{dt}\big(\mathrm{Ad}(g(t))Y\big) &= \frac{d}{dt}(g(t)Yg(t)^{-1}) \\
&= \frac{dg(t)}{dt}Yg(t)^{-1}) + g(t)Yg(t)\frac{dg(t)^{-1}}{dt} \\
&= \frac{dg(t)}{dt}Yg(t)^{-1}) + g(t)Yg(t)(-1)g(t)^{-1}\frac{dg(t)}{dt}g(t)^{-1}
\end{aligned}$$

Taking the value for $t = 0$ we have

$$\frac{d}{dt}\big(\mathrm{Ad}(g(t))Y\big)\Big|_{t=0} = XY - YX$$

Finally, we showed before that $\mathrm{Ad}(g(t))Y = g(t)Yg(t)^{-1}$ is in $\mathfrak{g}$ and therefore

$$\frac{d}{dt}\big(\mathrm{Ad}(g(t))Y\big) = \lim_{u\to 0}\frac{1}{u}\Big(\mathrm{Ad}(g(t+u))Y - \mathrm{Ad}(g(t))Y\Big)$$

will be in $\mathfrak{g}$ since a vector space is closed under limits, and thus its value at $t = 0$ which is $XY - YX = \frac{d}{dt}\big(\mathrm{Ad}(g(t))Y\big)\big|_{t=0}$ is in $\mathfrak{g}$. This proof works for matrix Lie Groups; it requires some slight modification for abstract Lie Groups.

For clarity, and because it is really all we need, let us specialize to $\mathrm{GL}(n, \mathbb{R})$. Then $\mathrm{GL}(n, \mathbb{R})$ is inside the linear space $\mathrm{M}(n, \mathbb{R})$ of $n \times n$ matrices and the tangent vector $A'(0)$ will be in in $\mathrm{M}(n, \mathbb{R})$ also, which is handy. We will show that $\mathfrak{g}l(n, \mathbb{R}) = \mathrm{M}(n, \mathbb{R})$. We may now imitate the construction for $\mathrm{SO}(n, \mathbb{R})$ (or

[10] An endomorphism is a homomorphism of a vector space into itself.

any other real matrix Lie Group) to get its Lie Algebra $\mathfrak{so}(n, \mathbb{R})$ but we will do this later.

First, it is obvious that $\mathfrak{gl}(n, \mathbb{R}) \subseteq \mathrm{M}(n, \mathbb{R})$. We must show the opposite. To do this we must introduce the exponential map exp: $\mathfrak{g} \to G$. This can be done for abstract groups too but is easier for matrix groups. We may put a "natural" inner product, and thus a norm, on $\mathrm{M}(n, \mathbb{R})$ quite easily by, for $X, Y \in \mathrm{M}(n, \mathbb{R})$

$$\begin{aligned} (X, Y) &= \sum_{ij} X_{ij} Y_{ij} \\ \|X\|^2 = (X, X) &= \sum_{ij} X_{ij}^2 \end{aligned}$$

Then we can define the exponential map $\exp : \mathrm{M}(n, \mathbb{R}) \to \mathrm{G}(n, \mathbb{R})$ by, for $X \in \mathrm{M}(n, \mathbb{R})$,

$$\exp(X) = \sum_{j=0}^{\infty} \frac{1}{j!} X^j$$

Then one proves the convergence of this series much as one proves it in real or complex variables using the Weierstrass M test which can easily be shown to work in these circumstance. It is a little trickier to prove

$$\det(\exp(X)) = \exp(\mathrm{Tr}(X))$$

where $\mathrm{Tr}(X) = \sum_i X_{ii}$ is the *trace* of X. One proves it for diagonal matrices, which is easy, and then proves it for all matrices similar to diagonal matrices, which is also easy, and then since diagonable matrices are dense in $\mathrm{M}(n, \mathbb{R})$ the result follows by continuity. You would think this would be easier but it is often the case with Lie theory that things that should be easy are harder than expected. Note that

$$\det(\exp(X)) \neq 0 \quad \text{so} \quad \exp(X) \in \mathrm{GL}(n, \mathbb{R})$$

Now given $X \in \mathrm{M}(n, \mathbb{R})$ we may form a path in $\mathrm{M}(n, \mathbb{R})$ by

$$A(t) = \exp(tX) = \sum_{j=0}^{\infty} \frac{1}{j!} (tX)^j = \sum_{j=1}^{\infty} \frac{t^j}{j!} X^j$$

Note $A(0) = I$. Then, imitating the standard trickery for differentiating power series, we have

$$\left.\frac{dA(t)}{dt}\right|_{t=0} = \exp(tX) \cdot X|_{t=0} = IX = X$$

and since X is the tangent vector to a curve at the origin we have $X \in \mathfrak{gl}(n, \mathbb{R})$ and thus

$$\mathfrak{gl}(n, \mathbb{R}) = \mathrm{M}(n, \mathbb{R})$$

as we promised.

We can try to imitate this construction for any matrix Lie Group but often there are easier ways. Now note that, just as with power series in complex variables, we may prove that

$$\exp(tX)\cdot\exp(uX) = \exp((t+u)X)$$

which tells us that $A(t) = \exp(tX)$ is a lot better than a path; it is a one-dimensional subgroup of $\mathrm{GL}(n,\mathbb{R})$. Each tangent vector X at I has one of these one dimensional subgroups to which it is tangent, so that an alternate definition of the Lie Algebra would be the set of one dimensional subgroups of G. Note also that $A : \mathbb{R} \to GL(n,\mathbb{R})$ is a homomorphism, so that X is not just a tangent vector to a curve but a tangent vector to a curve which is a homomorphism.

Caution In general $\exp(X)\cdot\exp(Y) \neq \exp(X+Y)$. However, just as in complex variables it is not difficult to prove that if X and Y commute then indeed $\exp(X)\cdot\exp(Y) = \exp(X+Y)$.

These results or analogous ones can be proved for abstract Lie Groups (Lie Groups which are *not* groups of matrices) but it takes a little more effort.

We now want to find the Lie Algebra $\mathfrak{so}(n,\mathbb{R})$ of $\mathrm{SO}(n,\mathbb{R})$, which is characterized by $AA^\top = I$ and $\det(A) = +1$. Let $X \in T_I(\mathrm{SO}(n,\mathbb{R}))$ and select a curve $A(t)$ with with $A(0) = I$ and

$$X = \frac{dA(t)}{dt}\Big|_{t=0}$$

Then

$$\frac{d}{dt}\big(A(t)A^\top(t)\big) = \frac{d}{dt}I = 0$$

so

$$\begin{aligned}
\frac{dA(t)}{dt}A^\top(t) + A(t)\frac{dA^\top(t)}{dt} &= 0\\
\frac{dA(t)}{dt}\Big|_{t=0}A^\top(0) + A(0)\frac{dA^\top(t)}{dt}\Big|_{t=0} &= 0\\
XI + IX^\top &= 0\\
X + X^\top &= 0
\end{aligned}$$

Thus $\mathfrak{so}(n,\mathbb{R}) \subseteq \{\text{skew symmetric } n\times n \text{ matrices}\}$.

On the other hand, let X be a skew symmetric $n\times n$ matrix. Then

$$\begin{aligned}
\exp(tX)(\exp(tX))^\top &= \exp(tX)\exp(tX^\top) = \exp(tX)\exp(-tX)\\
&= \exp(tX)\exp(tX)^{-1} = I
\end{aligned}$$

This shows that $\exp(tX) \in \mathrm{SO}(n,\mathbb{R})$ so

$$X = \frac{d}{dt}\exp(tX)\Big|_{t=0} \in \mathfrak{so}(n,\mathbb{R})$$

and we have shown

$$\boxed{\mathfrak{so}(n,\mathbb{R}) = \{\text{skew symmetric } n\times n \text{ matrices}\}}$$

Now we wish to do an example so we can see this all in action. The group here is $\mathrm{SO}(2,\mathbb{R})$ and we will take $X \in \mathfrak{so}(2,\mathbb{R})$ where X is the skew symmetric matrix

$$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Then we have

$$X^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \qquad X^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
$$X^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad X^5 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
$$\text{etc}$$

Then we set

$$\begin{aligned} A(t) = \exp(tX) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + t\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \frac{t^2}{2!}\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \\ &\quad + \frac{t^3}{3!}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \frac{t^4}{4!}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{t^5}{5!}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 - \frac{t^2}{2!} + \frac{t^4}{4!} + \dots & -t + \frac{t^3}{3!} - \frac{t^5}{5!} + \dots \\ t - \frac{t^3}{3!} + \frac{t^5}{5!} + \dots & 1 - \frac{t^2}{2!} + \frac{t^4}{4!} + \dots \end{pmatrix} \\ &= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \end{aligned}$$

Thus $A(t) \in \mathrm{SO}(2,\mathbb{R})$ is the one parameter group of rotations in the (x,y)-plane. Note

$$\left.\frac{dA}{dt}\right|_{t=0} = \left.\begin{pmatrix} -\sin t & -\cos t \\ \cos t & -\sin t \end{pmatrix}\right|_{t=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = X$$

which is just as it should be. In physics X is called the infinitesimal generator of the Lie group of rotations. It should be clear that in 3-spacethe infinitesimal generator of the group of the one parameter group of rotations around the z axis is

$$X_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

From this you should be able to figure out the rotation matrix $T_z(t)$ for angle t around the z axis and then by symmetries the rotations $T_x(t)$ and $T_y(t)$ by yourself. Then you can work out the commutators $[X_z, X_x],\ [X_x, X_y],\ [X_y, X_z]$. These are of great interest in Quantum Mechanics.

The Lie Algebra has a great deal of information about its Lie Group, but if two Lie Groups are alike around I their Lie Algebras will be the same. For example, the unit circle is a Lie Group with elements $\{e^{it}|t \in (-\pi, \pi]\}$ and $\mathbb{R}$ is a Lie Group under addition. Both have Lie Algebras isomorphic to $\mathbb{R}$ with bracket identically 0, so the Lie Algebra cannot determine the Lie Group. However, if we restrict ourselves to *simply connected* Lie Groups, then it is the case that every finite dimensional Lie Algebra does determine a unique simply connected Lie Group. This theorem is very difficult for people outside Lie studies.

For persons wishing to pursue this most beautiful and important part of modern mathmatics we give now a short annotated bibliography of the books that we found helpful. They are arranged more or less in order of difficulty (in our opinion).

SHORT BIBLIOGRAPHY OF LIE GROUPS AND LIE ALGEBRAS

These books cover only finite dimensional Lie theory.

1. Wu, Loring Tu. AN INTRODUCTION TO MANIFOLDS. Springer, New York, 2011. Chapter 4 of this book has a very gentle introduction to Lie Groups and Lie Algebras much more complete than that in the present book.

2. Baker, Andrew, MATRIX GROUPS, AN INTRODUCTION TO LIE GROUP THEORY, Springer, London 2002. An introduction to matrix Lie Groups with many examples and lots of motivation. A fine introduction to the subject including many subjects important for applications.

3. Knopp, Anthony W. LIE GROUPS BEYOND AN INTRODUCTION 2nd editon. Birkhäuser, Boston 2002. Despite the name, the 2nd edition has a initial chapter which is an introduction to Matrix Lie Groups. The book goes on from there to develop the abstract theory of Lie Groups in a way suitable for mathematical studies. Knopp explains things well and the book covers the classical theory. If you read this book you will be a beginning professional in the subject.

4. Bump, Daniel LIE GROUPS, Springer, New York, 2004. Covers the classical material in a somewhat shorter treatment than 3.

5. Duistermaat, J.J. & Kolk, J.A.C. LIE GROUPS, Springer, New York, 2000. This is a modern treatment of Lie Groups and Lie Algebras meant for the mathematically mature reader. It is not for the timid or untrained, but we found it rewarding to look at their treatment when we already had some familiarity with the material.

6. Chevelley, Claude, THEORY OF LIE GROUPS, Princeton 1946. An important classic in the field by a master. The reader must accustom herself to some terminology that is no longer current, but this is still a wonderful source. We highly recommend this book for readers who want to get

the flavor of Lie Groups quickly. Very self contained book which develops some functional analysis for use in the Peter Weyl theorem. Cognoscenti will recognize material here in the definition of manifolds which later was incorporated into sheaf theory, which is equivalent to but looks different from the treatment of manifolds in this book.

1.8 Frame Bundles and Principle Bundles

This is a natural place to introduce the *Frame Bundle* of a Differentiable Manifold. The frame bundle is the fundamental example of a *Principle Bundle* and so, with a little extra work, we can include this concept in our repertoire. Principle bundles have some importance in advanced areas of quantum mechanics so it is worth the slight extra effort to understand how frame bundles are examples of principle bundles. However, since this topic is pretty advanced we will not be using it much in the material that follows, so if you prefer you can just skim the abstract part. But the easy stuff about frame bundles it would be good to understand completely.

Recall the concept of Tangent Bundle of a manifold. At each point of the manifold $T_p(M)$ is an n-dimensional vector space and and we can, in a C^∞ way, choose a local basis of sections of $T_p(M)$. This means, for some $U \subseteq M$ (which may or may not be a coordinate patch) we have at each point $p \in U$ a basis

$$\sigma(p) = (e_1(p), \ldots, e_n(p))$$

of $T_p(M)$, where the $e_i(p)$ are C^∞ sections. The basis $\sigma(p)$ of $T_p(M)$ is called a *frame*. We can always come up with a *natural basis* $\sigma = (\partial/\partial u^1, \ldots, \partial/\partial u^n)$ but we do not want to emphasize the natural basis at all in the section. It is not possible in general to come up with a σ for the entire manifold; in fact it may not be possible to come up with a single non-zero section of the manifold, but this is a local book and such things are global results covered in books on Differential Topology.

The frame bundle resembles the tangent bundle in certain respects, but instead of a vector space over each $p \in U$ we have the set of frames $\sigma(p)$ for $T_p(M)$. This is just a set, but by requiring $\sigma(p)$ to be C^∞ we have already put some additional structure on the frame bundle. We want more.

Recall that, given a frame $\sigma(p)$ at each point we get an isomorphism of $T_p(M)$ onto $\mathbb{R}^n$ which takes $v = e_i v^i \in T_p(M)$ to $(v^1, \ldots, v^n)$. Recall that $\pi : T(U) \to M$ is given by $\pi(v) = p$ were $v \in T_p(U)$. We thus have a commutative triangle

$$\begin{array}{ccccc} v & & \longleftrightarrow & & ((v^1, \ldots, v^n), p) \\ & \pi \searrow & & \swarrow \pi_2 & \\ & & p & & \end{array}$$

We wish to do a similar thing for the frame bundle, giving coordinates of a sort, but this is rendered tricky by certain rules that principle bundles have. We will look at this soon.

1.8.1 Group Actions

An *Automorphism* of a set S is just a one to one onto mapping (a bijection) of S to S. The set of all automorphisms of S is denoted by $\mathrm{Aut}(S)$.

A *right action* of a group G on a set S is a mapping of G into $\mathrm{Aut}(S)$ denoted generally by $(s, g) \to sg$ subject to the following law:

$$(sg)h = s(gh) \quad \text{for } s \in S \text{ and } g.h \in G$$

Naturally if the set S and the group G have topologies it is assumed that $(s, g) \to sg$ is C^∞ and will will not talk much explicitly about these topological matters, assuming the reader can just insert them as necessary.

The action is *transitive* if and only if for any two elements $s_1, s_2 \in S$ there exists an element $g \in G$ for which $s_1 g = s_2$. The action is *free* if and only if, if for any $s \in S$ we have $sg = s$ then $g = e$, e being the identity of G. It is easy to prove the if the auction is *both* transitive and free then the g taking s_1 to s_2 is unique.

For transitive free actions an interesting (and possibly slightly confusing) thing happens. Select $s_0 \in S$ arbitrarily. Then there is a one to one correspoindence between S and G; given $s \in S$ there is a unique g_s for which $s = s_0 g_s$. Hence S is bijective to the group G but since the choice of s_0 (which corresponds to e in the bijection) is arbitrary S lacks an identifiable identity, not to mention the binary operation of a group. We express this whole situation, S, G and the transitive and free action of G on S by saying that

$$S \text{ is a } G\text{-torsor}$$

Our important and nearly unique example is the frame bundle. Let $U \subseteq M$ be a subset of M over which a local basis of sections

$$\sigma_0 = (e_1, \ldots, e_n)$$

of $T_p(U)$ can be defined. Then $G = \mathrm{GL}(n, \mathbb{R})$ acts on the frame bundle by

$$\sigma = \sigma_0 g = (e_1, \ldots, e_n)(g^i_j) \quad \text{where } g \in G$$

where if $\sigma = (f_1, \ldots, f_n)$ then $f_j = e_i g^i_j$. This action is easily seen to be transitive and free. Hence the frame bundle is (locally) a $\mathrm{GL}(n, \mathbb{R})$-torsor, but there is no distinguished section σ_0 which is the typical torsor situation.

1.8.2 Principal Bundles

We now return to the Frame bundle although we will set up the notation so that it looks very like the notation in an abstract treatment. The principal bundle here will be the bundle of frames where we are working over $U \subseteq M$ where we can set up a basis of local sections $\sigma(p) = (e_1(p)), \ldots, e_n(p))$ where $p \in U$ and $(e_1(p)), \ldots, e_n(p))$ is a basis of $T_p(U)$. The definition of Priniple Bundle requires us to come up with a mapping $\Phi : \pi^{-1}[U] \to G \times U$ which is one to one and onto (a bijection), C^∞, and which satisfies the following condition which is often called equivariance. That is

$$\begin{array}{rrcl} \text{if} & \Phi(\sigma) & = & (g,p) \\ \text{then} & \Phi(\sigma h) & = & (gh,p) \end{array}$$

(The σh were defined in the last section for the frame bundle but it works the same for any bundle with a transitive free action.) The mappings Φ are correlated with distinguished sections σ in the follwing manner.

$\rightarrow$ Given a Φ, define a section $\sigma_0(p)$ by

$$\sigma_0(p) = \Phi^{-1}(e,p)$$

$\leftarrow$ Given a section $\sigma_0(p)$, for any σ we can find a $g(p)$ in $\mathrm{GL}(n,\mathbb{R})$ for which $\sigma(p) = \sigma_0(p)g(p)$. Then we set

$$\Phi(\sigma(p)) = (g,p)$$

1.9 Affine Connections

To make my point more clearly in this section I want to begin with a little review. Let us start with two bases

$$\sigma = (e_1,\dots,e_n) \qquad \tilde{\sigma} = (\tilde{e}_1,\dots,\tilde{e}_n)$$

and let them be connected by the matrix $C = (h^i_j)$ so that

$$(\tilde{e}_1,\dots,\tilde{e}_n) = (e_1,\dots,e_n)(h^i_j) \qquad \tilde{\sigma} = \sigma C$$

Then we know that if $v = e_1v^1 + \cdots + e_nv^n = \tilde{e}_1\tilde{v}^1 + \cdots + \tilde{e}_n\tilde{v}^n$ that we will have

$$\begin{pmatrix} \tilde{v}^1 \\ \vdots \\ \tilde{v}^n \end{pmatrix} = C^{-1} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

Now it is possible to define a vector as a column whose entries change under change $\tilde{\sigma} = \sigma C$ in exactly this way. This analogous to defining a duck by
Def If it quacks like a duck it's a duck[11].
Note that having feathers or having n entries is not enough; it must have the proper quack or transformation behaviour under basis change.

Now this kind of definition, singling out some characteristic property and defining the concept as something having it, is not in itself bad, and is used in physics productively and often. Vectors are things that transform like vectors, tensors are things that transform like tensors, etc. However, in mathematics we

[11] I am indepted to Anthony Zee for this example

like to have things defined by having them belong to sets. This may sometimes be just pyschological quibbling, but that's the way we like to do it, and if the sets have structures this can be genuinely illuminating. So that's what we will attempt to do for Affine Connections.

In a somewhat similar way we can define an affine connection ω on a manifold like this. In each coordinate patch U with coordinates $(u^1, \dots, u^n)$ a affine connection is represented by an $n \times n$ matrix of first order differential forms, and thus looks like

$$\omega \text{ is represented locally by } \left(\Gamma^i_{jk}\, du^k\right)$$

where the Γ^i_{jk} are called the connection coefficients. Now for the tricky part. On the overlap of coordinate patches, the ω representatives must change in a specified way. Using the letter ω itself for the representative on the first patch, let us derive the transition rule more time. As usual we have $e_i = \partial/\partial u^j$ and $\tilde{e}^i = \partial/\partial \tilde{u}^i$. They are related by

$$\begin{aligned} (\frac{\partial}{\partial \tilde{u}^1}, \dots, \frac{\partial}{\partial \tilde{u}^n}) &= (\frac{\partial}{\partial u^1}, \dots, \frac{\partial}{\partial u^n})\left(\frac{\partial u^i}{\partial \tilde{u}^j}\right) \\ (\tilde{e}_1, \dots, \tilde{e}_n) &= (e_1, \dots, e_n)C \end{aligned}$$

with $C = (\partial u^i/\partial \tilde{u}^j)$. Then, using the abbreviation $e = (e_1, \dots, e^n)$ etc we have from our old formulas

$$\begin{aligned} D\tilde{e} &= \tilde{e}\tilde{\omega} \\ D(eC) &= eC\tilde{\omega} \\ e\, dC + (De)C &= eC\tilde{\omega} \\ e\, dC + e\omega\, C &= eC\tilde{\omega} \\ dC + \omega\, C &= C\tilde{\omega} \\ C^{-1}\, dC + C^{-1}\,\omega\, C &= \tilde{\omega} \end{aligned}$$

This is the formula which must be satisfied between the representatives ω and $\tilde{\omega}$ on the overlap of the patches $U \cap \tilde{U}$. We leave it to the reader to show that on the overlap of *three or more* coordinate patches the requirement is consistent.

Another point of importance is that in deriving the change of basis formula for ω I never used the fact that the bases involved originally were natural bases. This formula is good for any change of basis.

Now we must recall Euler's method for solving a first order differential equation. The equation with initial condition is

$$\frac{dy}{dt} = f(y, t) \qquad\qquad y(t_0) = y_0$$

The important point for us is that at $t = t_0$ the value of $\frac{dy}{dt}$ tells us which way to travel from the initial point. We now choose a small Δt and compute $\Delta y = \frac{dy}{dt}|_{t=t_0}\Delta t$. We will get an approximation $y_1 = y_0 + \Delta y$. Iterating the process we will get a sequence of values $y_j(t_j) = y_{j-1}(t_{j-1}) + \frac{dy}{dt}\big|_{t=t_{j-1}} \Delta t$. Connecting the dots in the sequence of points $(t_j, y_j(t_j))$ we get an approximation of the solution with points $(t, y(t))$. The smaller the Δt the better the approximation. Again, the value of $\frac{dy}{dt}\big|_{t=t_{j-1}}$ tells us which way to go to get to the next point in the approximation.

The domain (t_1, t_2) of the solution with $t_1 < t_0 < t_2)$ is not very predictable for general $f(y,t)$ but for *linear* funcitons $f(y,t) = Ay$ the solution is defined for all t. The situation with systems where y is a column vector and $A \in \mathrm{GL}(n, \mathbb{R})$ is similar and the solution exists for all t.

We now apply this to the following situation. Let C be a curve in the n-dimensional manifold M given by $(u^1(t), \ldots, u^n(t))$. Let $v(t)$ be the tangent vector to $C(t)$ with coordinates $(u'^1(t), \ldots, u'^n(t))$. Let $\sigma_0(t)$ be a curve in the frame bundle of M lying over C so that $\pi(\sigma_0(t)) = C(t)$.

We assume that we have an affine connection given in our coordinate patch by ω. Now consider the system of equations on the curve C

$$\frac{dg_i^j(t)}{dt} = \omega(v(t)) = -\Gamma^j_{ik} du^k(v(t)) = -\Gamma^j_{ik}\frac{du^k}{dt} \qquad\qquad g_i^j(0) = \delta_i^j$$

This system will have a solution and for small t continuity guarantees that $\det(g_i^j(t)) \neq 0$. (Actually, a little extra work would give us this for all t but we don't need it. A local solution will do.) Our solution can thus be regarded a curve in the Lie Group $\mathrm{GL}(n, \mathbb{R})$ and thus it's derivative at $t = 0$, that is $-\big(\Gamma^j_{ik}\frac{du^k}{dt}\big)\big|_{t=0}$, is an element of the Lie algebra of $\mathrm{GL}(n, \mathbb{R})$, which is $\mathfrak{g}l(n, \mathbb{R})$. Thus we can say that ω is a $\mathfrak{g}l(n, \mathbb{R})-$valued one-form. The entire process can be generalized to other groups than $\mathrm{GL}(n, \mathbb{R})$ and other principle bundles. Later we will show that for Riemannian geometry the natural group is $\mathrm{SL}(n, \mathbb{R})$ and the connection is thus an $\mathfrak{s}l(n, \mathbb{R})-$valued one-form which is to say that ω is skew symmetric: $\omega^i_j = \omega^j_i$.

The curious reader may wonder what the geometric significance of $g_i^j(t)$ is and it is very interesting.

1.10 Riemannian Geometry

In a previous section we studied manifolds which are embedded in a Eucliden space of one higher dimension. This gave us, among other things, a normal vector and an inner product. Riemann realized from his knowledge of Gauss's work on intrinsic geometry (for example, the Theorema Egregium), that the embedding was unnecessary. Whether Riemann foresaw Einstein's use of non-embedded manifolds (perhaps the Universe as a spacial 3-manifold) is not known; various people have suspected this was one of his motivations, but he left no solid clue behind. Riemann realized that the basic tool one needed to develop a geometry familiar enough to work with but much more general was the inner product.

Def A Riemannian Manifold is a finite dimensional differentiable manifold M (C^∞ for us) for which each point $p \in M$ has an inner product on $T_p(M)$ associated with it. We also require that if v, w are C^∞ vector fields over some open set of M then (v, w) is a C^∞ function on the open set.

It is easy to see that this just means the coefficients (g_{ij}) are C^∞ functions on M. The inner product lets us measure lengths and angles for vectors in $T_p(M)$.

We will introduce here an idea which has a lot of uses later. For any open set U a *section* of $T(M)$ over U is a function $v : U \to T(M)$ for which $v(p) \in T_p(M)$. Thus $v(p)$ gives a vector in the tangent space at p for each $p \in U$. This is often called a tangent vector field on U, which is the older terminology. Section of the tangent bundle and tangent vector field are thus synonyms.

Now it should be clear that locally (that is, in a neighborhood U of any point $p \in M$) we can choose n vector fields which are linearly independent at each point of U. It probably won't be possible to do this globally, but no matter; this is mostly a local book. We call the n vector fields $e_1, \ldots, e_n$. We could even use the Gram-Schmidt orthogonalization procedure to get (locally) an orthonormal basis which might, or might not, be a good idea. Note we can do this because the inner product lets us measure lengths and angles.

Our tour of embedded manifolds has taught us some things. We are going to need a covariant derivative which we assume works like it did before; we need one forms on the manifold so that for a local basis of sections of the Tangent bundle $e_1, \ldots, e_n$ w have a matrix $({\omega_i}^j)$ so that

$$De_i = e_j\, {\omega_i}^j$$

As before the ${\omega_i}^j = \Gamma^j_{ik}\, du^k$ are called the *connection 1-forms*. Since all we have to begin with are the (g_{ij}), we must somehow get the connection 1-forms from the (g_{ij}). Once we have these we have curvature and geodesics.

We have to start somewhere, and one of quite a large number of ways to start is to require that a formula we previously derived for embedded manifolds continues to hold. It is quite natural to require the following analogue of Leibniz' rule:

$$d(v, w) = (Dv, w) + (v, Dw)$$

We would certainly *like* this to be true, and perhaps it will narrow down the choice of ${\omega_i}^j$. Applying this to e_i and e_j, we have

$$\begin{aligned}
De_i &= e_k {\omega_i}^k \;=\; e_k \Gamma^k_{im}\, du^m \\
De_j &= e_l {\omega_j}^k \;=\; e_l \Gamma^l_{jn}\, du^n \\
d(e_i, e_j) &= (De_i, e_j) + D(e_i, De_j) \\
dg_{ij} &= (e_k \Gamma^k_{im}\, du^m,\, e_j) + (e_i,\, e_l \Gamma^l_{jn}\, du^n) \\
&= g_{lj} \Gamma^l_{im}\, du^m + g_{il} \Gamma^l_{jn}\, du^n \\
\frac{\partial g_{ij}}{\partial u^k}\, du^k &= \left(g_{lj} \Gamma^l_{ik} + g_{il} \Gamma^l_{jk}\right) du^k
\end{aligned}$$

Using the linear independence of the du^m and cyclically permuting the $i\,j\,k$ we have

$$\begin{aligned}\frac{\partial g_{ij}}{\partial u^k} &= g_{lj}\Gamma^l_{ik} + g_{il}\Gamma^l_{jk}\\ \frac{\partial g_{jk}}{\partial u^i} &= g_{lk}\Gamma^l_{ji} + g_{jl}\Gamma^l_{ki}\\ \frac{\partial g_{ki}}{\partial u^j} &= g_{li}\Gamma^l_{kj} + g_{kl}\Gamma^l_{ij}\end{aligned}$$

This almost does it. If we add the second and third equations and subtract the first we have

$$\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k} = g_{lk}\Gamma^l_{ji} + g_{jl}\Gamma^l_{ki} + g_{li}\Gamma^l_{kj} + g_{kl}\Gamma^l_{ij} - g_{lj}\Gamma^l_{ik} - g_{il}\Gamma^l_{jk}$$

The second and fifth terms on the right, and the third and sixth terms, would cancel if $\Gamma^l_{ki} = \Gamma^l_{ik}$. Hence, besides the analogue of Leibniz rule we started with it is also necessary to assume the connection coefficients Γ^l_{ik} are symmetric in the lower two indices. We recall this was true in the embedded case and so it seems fairly natural to assume it here also. With this assumption the terms cancel and the surviving first and fourth terms double up and we are left with

$$\begin{aligned}\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k} &= 2g_{kl}\Gamma^l_{ij}\\ \frac{1}{2}g^{mk}\left(\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k}\right) &= g^{mk}g_{kl}\Gamma^l_{ij} \;=\; \delta^m_l\Gamma^l_{ij}\end{aligned}$$

giving

$$\Gamma^m_{ij} \;=\; \frac{1}{2}g^{mk}\left(\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k}\right)$$

as the only possible choice of connection coefficients which are symmetric and satisfies the analogue of Leibniz rule.

Now that we have ${\omega_j}^i$ we may define geodesics as curves parametrized by arc length which satisfy the geodesic equation

$$\frac{d^2u^k}{ds^2} + \Gamma^k_{ij}\frac{du^i}{ds}\frac{du^j}{ds} = 0$$

We may study curvature by defining the Curvature Form Ω as

$$\Omega = d\omega + \omega \wedge \omega$$

where $\omega = ({\omega_j}^i)$ is the matrix of curvature forms. This cherns up the Riemann Curvature Tensor exactly as in the embedded case.

Bibliography

[1] Akhiezer, N. I. and Glazman, I. M. THEORY OF LINEAR OPERATORS IN HILBERT SPACE, Dover, New York, 1993.

[2] Bott, Raoul and Tu, Loring W., DIFFERENTIAL FORMS IN ALGEBRAIC TOPOLOGY, Springer, New York, 1982

[3] Frankel, Theodore THE GEOMETRY OF PHYSICS, 3^{rd} edition, Cambridge 2012 Cambridge University Press, Cambridge,etc, 2012

[4] Gilkey, Peter B. INVARIANCE THEORY, THE HEAT EQUATION AND THE ATIZYA-SINGER INDEX THEOREM, 2nd Edition Boca Raton 1995

[5] Jost, Jürgen, ON THE HYPOTHESES WHICH LIE AT THE BASES OF GEOMETRY, Brikhäuser, Switzerland 2016

[6] Jost, Jürgen, GEOMETRIC ANALYSIS 4th edition, Springer, New York, 2005

[7] Laugwitz, Detlef, DIFFERENTIAL AND RIEMANNIAN GEOMETRY, Academic Press, New York 1965

[8] Liusternik L. A. and Sobolev V. J. ELEMENTS OF FUNCTIONAL ANALYSIS New York, 1961

[9] Lovelock, David and Rund, Hanno TENSORS, DIFFERENTIAL FORMS AND VARIATIONAL PRINCIPLES, Wiley Interscience

[10] Misner, Thorne and Wheeler, GRAVITATION, W.H. Freeman, San Francisco, 1973

[11] Nakahara, Mikio GEOMETRY, TOPOLOGY, AND PHYSICS, Institute of Physics Publishing, London, 2003

[12] Pederson, Peter, RIEMANNIAN GEOMETRY, Springer, New York 1998

[13] Roe, John, ELLIPTIC OPERATORS, TOPOLOGY AND ASYMPTOTIC METHODS, 2nd Edition, Boca Raton, 1998

[14] Rosenberg, S. THE LAPLACIAN ON A RIEMANNIAN MANIFOLD, Cambridge, 1997

[15] Schmüdgen, Konrad, UNBOUNDED SELF ADJOINT OPERATORS ON HILBERT SPACE, Springer, Dordrecht, 2012

[16] Schulz, William, THEORY AND APPLICATION OF GRASSMANN ALGEBRA, http://www.cefns.nau.edu/ schulz/grassmann.pdf

[17] Yosida, Kôsaku FUNCTIONAL ANALYSIS, 4^{th} Edition, Springer, New York, 1974.

[18] Zee, Anthony, EINSTEIN GRAVITY IN A NUTSHELL, Princeton U.P, Princeton New Jersey, 2013