Math 456/501 Course Notes

Spring 2014

These are daily notes for the Spring 2014 Introduction to Geometry course at Penn. The notes are a combination of notes of my own presentation, and notes derived from sources. Where notes are derived from books of other authors, I shall attempt to give references.

Preliminaries

0.1 Non-Calculus Basics

0.1.1 Distance

Given two points $\vec{x}, \vec{y} \in \mathbb{R}^n$ where $\vec{x} = (x^1, \ldots, x^n)^T$ and $\vec{y} = (y^1, \ldots, y^n)^T$, we have the dot product

$$\vec{x} \cdot \vec{y} \triangleq x^1 y^1 + x^2 y^2 + \ldots + x^n y^n. \tag{1}$$

Given a point $\vec{x} \in \mathbb{R}^n$ where $\vec{x} = (x^1, \ldots, x^n)^T$, the *norm* of $\vec{x}$, denoted $\|\vec{x}\|$ is

$$\|\vec{x}\| = \sqrt{(x^1)^2 + (x^2)^2 + \ldots + (x^n)^2} \tag{2}$$

One can prove the *triangle inequality*

$$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \tag{3}$$

and the *Cauchy-Schwarz* inequality

$$|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \, \|\vec{y}\| . \tag{4}$$

Given two points $\vec{x} = (x^1, \ldots, x^n)$ and $\vec{y} = (y^1, \ldots, y^n)$ in $\mathbb{R}^n$, the distance between them is

$$\mathrm{dist}(\vec{x}, \vec{y}) \triangleq \|\vec{x} - \vec{y}\| . \tag{5}$$

Given any three points $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$, we have the triangle inequality

$$\mathrm{dist}(\vec{x}, \vec{z}) \leq \mathrm{dist}(\vec{x}, \vec{y}) + \mathrm{dist}(\vec{y}, \vec{z}). \tag{6}$$

0.1.2 Functions

Given a function f, it is often convenient to specify its domain and range, even without specifying the exact form of the function itself. For example the function

$$f(x) = x^2 + cos(x) \quad for \quad x \in [a, b] \tag{7}$$

could be written

$$f : [a, b] \;\longrightarrow\; \mathbb{R} \tag{8}$$

if we don't know (or don't care about) exactly what the function is. Of course we may have multi-variable functions; for instance

$$\vec{f}(x, y) \;=\; \big(x^2,\, y^2 + xy,\, xy + \log(1 + x^2)\big) \tag{9}$$

could be written

$$f : \mathbb{R}^2 \;\longrightarrow\; \mathbb{R}^3. \tag{10}$$

It is normally more important to be careful with the range than the domain. For instance, since log(0) does not exist, the function

$$\vec{f}(x, y) \;=\; \big(x^2,\, y^2 + xy,\, xy + \log(x^2 + y^2)\big) \tag{11}$$

should be written

$$f : \mathbb{R}^2 \setminus \{0\} \;\longrightarrow\; \mathbb{R}^3. \tag{12}$$

0.2 Calculus: Integration

The process of integration is breaking a region into very small (indeed infinitesimal) parts, and summing. In the case of a continuous 1-variable function $f : [a, b] \to \mathbb{R}$, we can approximate the (signed) area under the curve by selecting N many points $\{x_1 = a,\, x_2,\, \ldots,\, x_N\}$ and setting the (discrete) length of the i^{th} interval by $\triangle x_i \triangleq x_{i-1} - x_i$ and computing

$$\sum_{i=1}^{N} f(x_i) \triangle x_i. \tag{13}$$

As the partition becomes finer and finer, so $N \to \infty$ and $\triangle x_i \to 0$ for each i, the discrete sum $\sum$ becomes an "infintesimal" sum $\int$, and we have the exact signed area under the curve

$$\int_a^b f(x)\, dx \;=\; \lim_{N \to \infty} \sum_{i=1}^{N} f(x_i) \triangle x_i. \tag{14}$$

0.3 Calculus: Differentiation

Single variable functions

The process of differentiation is conceptually more difficult; perhaps the best we to think about it is as a way to find the best linear approximation to a function. The definition

$$f'(x_0) \;=\; \lim_{\triangle x \to 0} \frac{\triangle f\big|_{x_0}}{\triangle x} \tag{15}$$

where $\triangle f\big|_{x_0} = f(x_0 + \triangle x) - f(x_0)$. Then the best linear approximation to $f(x)$ at x_0 is the linear expression

$$f(x_0) \; + \; f'(x_0) \cdot (x \; - \; x_0) \, . \tag{16}$$

However there are further interpretations.

Real-valued functions of several variables

In the case of a function

$$f : \mathbb{R}^n \longrightarrow \mathbb{R} \tag{17}$$

the linear approximation (equation of a hyperplane) to a $f(\vec{x})$ at $\vec{x}_0$ is

$$f(\vec{x}_0) \; + \; \vec{\nabla} f(x_0) \; \cdot \; (\vec{x} \; - \; \vec{x}_0) \tag{18}$$

where the $\cdot$ is matrix multiplication and the gradient is the row vector

$$\vec{\nabla} f \;\; = \;\; \left(\frac{\partial f}{\partial x^1}, \; \frac{\partial f}{\partial x^2}, \; \ldots, \; \frac{\partial f}{\partial x^n} \right) \tag{19}$$

0.3.1 Vector-valued functions of several variables

In the case of a function

$$\vec{F} : \mathbb{R}^n \longrightarrow \mathbb{R}^m \tag{20}$$

the linear approximation (equation of a hyperplane) to a $\vec{F}(\vec{x})$ at $\vec{x}_0$ is

$$\vec{F}(\vec{x}_0) \; + \; \vec{\nabla} \vec{F}(x_0) \; \cdot \; (\vec{x} \; - \; \vec{x}_0) \tag{21}$$

where the $\cdot$ is matrix multiplication and $\vec{\nabla}\vec{f}$ is the *Jacobian matrix* of $\vec{f}$:

$$\vec{\nabla}\vec{F} \;\; = \;\; \begin{pmatrix} \frac{\partial F^1}{\partial x^1} & \frac{\partial F^1}{\partial x^2} & \cdots & \frac{\partial F^1}{\partial x^n} \\ \frac{\partial F^2}{\partial x^1} & \frac{\partial F^2}{\partial x^2} & \cdots & \frac{\partial F^2}{\partial x^n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial F^m}{\partial x^1} & \frac{\partial F^m}{\partial x^2} & \cdots & \frac{\partial F^m}{\partial x^n} \end{pmatrix} \tag{22}$$

Lecture 1 - Curves in $\mathbb{R}^n$

Delivered Thursday Jan. 16

1.1 A word on notation

In the past, coordinates on $\mathbb{R}^n$ would be denumerated with subcripts, that is as $x_1, x_2, \ldots, x_n$, and coordinates given as a row vector, for instance $(x_1, \ldots, x_n)$. As will be increasingly clear throughout the course, it will be convenient to use superscripts instead, and also to consider points in $\mathbb{R}^n$ to be column vectors, not row vectors.

To be precise, let

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \qquad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \tag{1.1}$$

be the basis vectors of $\mathbb{R}^n$. Then an arbitrary vector $\vec{v} \in \mathbb{R}^n$ has components $v^1, \ldots, v^n$, and $\vec{v}$ is the column vector we write

$$\vec{v} \;=\; \sum_{i=1}^{n} v^i \vec{e}_i \;=\; \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} \tag{1.2}$$

We may abbreviate this by (v^i). At times we want to deal with row vectors; these are indicated by lower indices. A row vector would be denoted $(v_1, v_2, \ldots, v_m)$. For instance if $v_1 = 1$, $v_2 = 0$, $v_3 = -2$, the row vector (v_i) would be $(1, 0, -2)$.

This gives a convenient way to denote matrices as well. The matrix (a^i_j) would be

$$(a^i_j) \;=\; \begin{pmatrix} a^1_1 & a^1_2 & \ldots & a^1_n \\ a^2_1 & a^2_2 & & a^2_n \\ \vdots & & \ddots & \vdots \\ a^m_1 & a^m_2 & \ldots & a^m_n \end{pmatrix} \tag{1.3}$$

so the notation is

$$a^{row}_{column} \tag{1.4}$$

1.2 Curves

Given a 1-dimensional object in some space, there is a difference between the geometric object itself and the parametrization used to describe it.

A parametrized curve in $\mathbb{R}^n$ is a continuous function

$$\vec{\gamma} : I \longrightarrow \mathbb{R}^n \tag{1.5}$$

where $I \subseteq \mathbb{R}$ is any interval. Following with our convention of considering points in $\mathbb{R}^n$ to be column vectors, the path has components $\gamma^i(t)$ and we have

$$\vec{\gamma}(t) = \begin{pmatrix} \gamma^1(t) \\ \gamma^2(t) \\ \vdots \\ \gamma^n(t) \end{pmatrix} \tag{1.6}$$

If $\vec{\gamma}$ is differentiable, the tangent vector at $x_0 \in I$ is just the column vector

$$\frac{d\vec{\gamma}}{dt} = \begin{pmatrix} \frac{d\gamma^1}{dt} \\ \frac{d\gamma^2}{dt} \\ \vdots \\ \frac{d\gamma^n}{dt} \end{pmatrix} \tag{1.7}$$

Reparametrizations

If $\vec{\gamma}(t)$ is any parametrized curve, the length of the curve between points $p_0 = \vec{\gamma}(t_0)$ and $p_1 = \vec{\gamma}(t_1)$ is *geometric*, meaning independent of parametrization. It is given by

$$\int_{p_0}^{p_1} |d\vec{\gamma}| = \int_{t_0}^{t_1} \left| \frac{d\vec{\gamma}}{dt} \right| dt = \int_{t_0}^{t_1} \sqrt{\left(\frac{d\gamma^1}{dt}\right)^2 + \left(\frac{d\gamma^2}{dt}\right)^2 + \ldots + \left(\frac{d\gamma^n}{dt}\right)^2}\, dt \tag{1.8}$$

Given a fixed point $p_0 = \vec{\gamma}(t_0)$ on the curve, this allows us to define the arclength as a function of t:

$$s = s(t) = \int_{t_0}^{t} \left| \frac{d\vec{\gamma}}{d\tau} \right| d\tau \tag{1.9}$$

Example. Consider the curve in $\mathbb{R}^2$ given by

$$\vec{\gamma}(t) \;=\; \begin{pmatrix} 2t \\ t^2 \end{pmatrix}. \tag{1.10}$$

From the initial point $(0,0)^T = \vec{\gamma}(0)$, determine arclength as a function of time.

Solution. We compute

$$\frac{d\vec{\gamma}}{dt} \;=\; \begin{pmatrix} 2 \\ 2t \end{pmatrix} \tag{1.11}$$

and $|d\vec{\gamma}/dt| = 2\sqrt{1+t^2}$. From (1.9) we then compute

$$s \;=\; \int_0^t \sqrt{1+\tau^2}\, d\tau \;=\; t\sqrt{1+t^2} \,+\, \ln\left(|t| + \sqrt{1+t^2}\right) \tag{1.12}$$

1.2.1 Calculus of Curves

The *unit tangent* to the curve, denoted $\vec{T}$ (or $\vec{T}_{\vec{\gamma}}$ if the function $\vec{\gamma}$ needs to be specified) is the normalized tangent vector

$$\vec{T} \;=\; \frac{d\vec{\gamma}/dt}{\|d\vec{\gamma}/dt\|} \;=\; \frac{d\vec{\gamma}/ds}{\|d\vec{\gamma}/ds\|}. \tag{1.13}$$

The *principal normal* is the direction (not the magnitude) in which the unit tangent is bending in:

$$\vec{N} \;=\; \frac{d\vec{T}/dt}{\left\|d\vec{T}/dt\right\|} \;=\; \frac{d\vec{T}/ds}{\left\|d\vec{T}/ds\right\|}. \tag{1.14}$$

The magnitude of the change in the tangent vector is called the *path curvature* or *geodesic curvature* of the curve, and is denoted κ:

$$\kappa \;=\; \left\|\frac{d\vec{T}}{ds}\right\| \;=\; \frac{\left\|\frac{d\vec{T}}{dt}\right\|}{\left|\frac{ds}{dt}\right|} \tag{1.15}$$

1.2.2 Computation of κ

For twice-differentiable curves, from (1.14) we have the formula

$$\frac{d\vec{T}}{dt} \;=\; \kappa\,\vec{N} \tag{1.16}$$

which is really only useful if we have a shortcut for computing κ. We compute

$$\begin{aligned}
\frac{d\vec{T}}{dt} &= \frac{d}{dt}\left(\frac{d\vec{\gamma}}{dt}\left\|\frac{d\vec{\gamma}}{dt}\right\|^{-1}\right) \\
&= \frac{d^2\vec{\gamma}}{dt^2}\left\|\frac{d\vec{\gamma}}{dt}\right\|^{-1} + \frac{d\vec{\gamma}}{dt}\frac{d}{dt}\left\|\frac{d\vec{\gamma}}{dt}\right\|^{-1} \\
&= \frac{d^2\vec{\gamma}}{dt^2}\left\|\frac{d\vec{\gamma}}{dt}\right\|^{-1} + \frac{d\vec{\gamma}}{dt}\frac{d}{dt}\left(\left\|\frac{d\vec{\gamma}}{dt}\right\|^{2}\right)^{-\frac{1}{2}} \\
&= \frac{d^2\vec{\gamma}}{dt^2}\left\|\frac{d\vec{\gamma}}{dt}\right\|^{-1} - \frac{1}{2}\frac{d\vec{\gamma}}{dt}\frac{d}{dt}\left\|\frac{d\vec{\gamma}}{dt}\right\|^{2}\left(\left\|\frac{d\vec{\gamma}}{dt}\right\|^{2}\right)^{-\frac{3}{2}} \qquad (1.17) \\
&= \frac{\frac{d^2\vec{\gamma}}{dt^2}\left\|\frac{d\vec{\gamma}}{dt}\right\|^{2} - \frac{1}{2}\frac{d\vec{\gamma}}{dt}\frac{d}{dt}\left\|\frac{d\vec{\gamma}}{dt}\right\|^{2}}{\left\|\frac{d\vec{\gamma}}{dt}\right\|^{3}} \\
&= \frac{\frac{d^2\vec{\gamma}}{dt^2}\frac{d\vec{\gamma}}{dt}\cdot\frac{d\vec{\gamma}}{dt} - \frac{d\vec{\gamma}}{dt}\frac{d^2\vec{\gamma}}{dt^2}\cdot\frac{d\vec{\gamma}}{dt}}{\left\|\frac{d\vec{\gamma}}{dt}\right\|^{3}}
\end{aligned}$$

Then

$$\begin{aligned}
\left\|\frac{d\vec{T}}{dt}\right\|^{2} &= \left\|\frac{\frac{d^2\vec{\gamma}}{dt^2}\frac{d\vec{\gamma}}{dt}\cdot\frac{d\vec{\gamma}}{dt} - \frac{d\vec{\gamma}}{dt}\frac{d^2\vec{\gamma}}{dt^2}\cdot\frac{d\vec{\gamma}}{dt}}{\left\|\frac{d\vec{\gamma}}{dt}\right\|^{3}}\right\|^{2} \\
&= \frac{\frac{d^2\vec{\gamma}}{dt^2}\cdot\frac{d^2\vec{\gamma}}{dt^2}\left(\frac{d\vec{\gamma}}{dt}\cdot\frac{d\vec{\gamma}}{dt}\right)^{2} - 2\left(\frac{d^2\vec{\gamma}}{dt}\cdot\frac{d\vec{\gamma}}{dt}\right)^{2}\left(\frac{d\gamma}{dt}\cdot\frac{d\gamma}{dt}\right) + \frac{d\vec{\gamma}}{dt}\cdot\frac{d\vec{\gamma}}{dt}\left(\frac{d^2\vec{\gamma}}{dt^2}\cdot\frac{d\vec{\gamma}}{dt}\right)^{2}}{\left\|\frac{d\vec{\gamma}}{dt}\right\|^{6}} \qquad (1.18) \\
&= \frac{\left(\frac{d^2\vec{\gamma}}{dt^2}\cdot\frac{d^2\vec{\gamma}}{dt^2}\right)\left(\frac{d\vec{\gamma}}{dt}\cdot\frac{d\vec{\gamma}}{dt}\right) - \left(\frac{d^2\vec{\gamma}}{dt}\cdot\frac{d\vec{\gamma}}{dt}\right)^{2}}{\left\|\frac{d\vec{\gamma}}{dt}\right\|^{4}} = \frac{\left\|\ddot{\vec{\gamma}}\right\|^{2}\left\|\dot{\vec{\gamma}}\right\|^{2} - \left(\ddot{\vec{\gamma}}\cdot\dot{\vec{\gamma}}\right)^{2}}{\left\|\dot{\vec{\gamma}}\right\|^{4}} \\
\kappa^2 &= \left\|\frac{d\vec{T}}{ds}\right\|^{2} = \frac{\left\|\frac{d\vec{T}}{dt}\right\|^{2}}{\left|\frac{ds}{dt}\right|^{2}} = \frac{\left\|\ddot{\vec{\gamma}}\right\|^{2}\left\|\dot{\vec{\gamma}}\right\|^{2} - \left(\ddot{\vec{\gamma}}\cdot\dot{\vec{\gamma}}\right)^{2}}{\left\|\dot{\vec{\gamma}}\right\|^{6}}
\end{aligned}$$

$$\kappa = \frac{\sqrt{\left\|\ddot{\vec{\gamma}}\right\|^{2}\left\|\dot{\vec{\gamma}}\right\|^{2} - \left(\ddot{\vec{\gamma}}\cdot\dot{\vec{\gamma}}\right)^{2}}}{\left\|\dot{\vec{\gamma}}\right\|^{3}} \qquad (1.19)$$

1.2.3 Interpretation of κ

Consider a parametrized circle of radius $r > 0$ in the plane:

$$\vec{\gamma}(t) \;=\; \begin{pmatrix} r\cos(t) \\ r\sin(t) \end{pmatrix} \tag{1.20}$$

We compute

$$\begin{aligned} \frac{d\vec{\gamma}}{dt} \;&=\; \begin{pmatrix} -r\sin(t) \\ r\cos(t) \end{pmatrix}, \qquad \frac{d^2\vec{\gamma}}{dt} \;=\; \begin{pmatrix} -r\cos(t) \\ -r\sin(t) \end{pmatrix} \\ \left\| \frac{d\vec{\gamma}}{dt} \right\| \;&=\; \left\| \frac{d^2\vec{\gamma}}{dt^2} \right\| \;=\; r, \qquad \frac{d\vec{\gamma}}{dt}\cdot\frac{d^2\vec{\gamma}}{dt^2} \;=\; 0 \end{aligned} \tag{1.21}$$

so that by (1.19)

$$\kappa \;=\; \frac{\sqrt{r^2\cdot r^2 \;-\; 0}}{r^3} \;=\; r^{-1}. \tag{1.22}$$

The curvature of a circle, therefore, is the inverse of its radius.

The *osculating circle* to a curve $\vec{\gamma} : I \to \mathbb{R}^n$ at a point $\vec{\gamma}(t_0)$ is the best approximating circle: that is, the circle that is incident on $\vec{\gamma}(t_0)$, has the same tangent $\vec{T}$, principal normal $\vec{N}$, and curvature κ as $\vec{\gamma}$ at t_0.

Example: Circles

Circles in the plane are easy to parametrize, mainly due to the fact that there is no question as to which 2-plane the circle lies in. If a circle in $\mathbb{R}^2$ has center $\vec{p} = \begin{pmatrix} p^1 \\ p^2 \end{pmatrix}$ and radius r, a good parametrization is

$$\vec{\eta}(\tau) \;=\; \begin{pmatrix} p^1 + r\cos\tau \\ p^2 + r\sin\tau \end{pmatrix}. \tag{1.23}$$

In higher dimensional settings, it can be a little tougher to figure out due to the fact that the circle's 2-plane has to be specified. Assuming a circle in $\mathbb{R}^n$ passes through point $\vec{p}$ and has radius r, and, at $\vec{p}$, has unit tangent vector $\vec{T}$ and unit normal vector $\vec{N}$, we should have enough information to find an equation for this circle. Indeed

$$\vec{\eta}(\tau) \;=\; \vec{p} \,+\, r\left(\cos(\tau) \,-\, 1\right)\vec{N} \,+\, r\sin(\tau)\vec{T} \tag{1.24}$$

fits the bill. To see this, note that clearly this path lies in the plane through $\vec{p}$ that is spanned by $\vec{T}$ and $\vec{N}$, that it passes through $\vec{p}$ at time $t = 0$, and that its curvature κ is the constant r^{-1}.

1.3 Exercises

1) Using only the triangle inequality (3), prove the Cauchy-Schwarz inequality (4).

2) Using only the triangle inequality for norms (3), prove the triangle inequality for distances (6).

3) Consider the path $\vec{\gamma} : \mathbb{R} \to \mathbb{R}^3$ given by $\gamma^1(t) = \frac{2}{3}t^{\frac{3}{2}}$, $\gamma^2(t) = \sin(t)$, $\gamma^3(t) = \cos(t)$. Determine arclength s as a function of time, where $s(0) = 0$.

4) Reparametrize the path $\vec{\gamma}$ from problem (3) in terms of arclength.

5) Determine the best linear approximation to the given functions at the given points.

   a) $f : \mathbb{R} \to \mathbb{R},\ f(x) \ = \ x^2 e^x$ at $x_0 = 1$

   b) $f : \mathbb{R}^2 \to \mathbb{R},\ f(\vec{x}) \ = \ 2 + x^1 x^2 \ + \ \left(x^1\right)^2$ at $\vec{x}_0 = (1, -1)^T$

   c) $\vec{f} : \mathbb{R}^2 \to \mathbb{R}^3,\ \vec{f}(\vec{x}) \ = \ \begin{pmatrix} \left(x^1\right)^2 x^2 + 1 \\ x^1 x^2 \\ x^1 + x^2 \end{pmatrix}$ at $\vec{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

6) Determine the unit tangent $\vec{T}$ and the principal normal $\vec{N}$ for the following curves:

   a) $\vec{\gamma}(t) \ = \ \begin{pmatrix} \frac{1}{2}t \\ t^2 \end{pmatrix}$

   b) $\vec{\eta}(t) \ = \ (\cos(t),\, \sin(t),\, t)^T$

7) For curves (*a*) and (*b*) in problem (6), graph a portion of each curve, and graph $\vec{T}(t_0)$ and $\vec{N}(t_0)$ at $t_0 = 1$

8) Consider the plane curve $\vec{\gamma}(t) = (\cos(t),\, 2\sin(t))^T$.

   a) Compute κ as a function of t.

   b) Determine the equations of the osculating circles for $t = 0$ and $t = \frac{\pi}{2}$.

   c) Sketch the graph of $\vec{\gamma}$ along with the osculating circles for $t = 0$ and $t = \frac{\pi}{2}$.

9) Consider the plane curve $\vec{\gamma}(t) = (t, t^3)^T$; its graph is the standard cubic. Determine the osculating circle to the graph when $t = 0$.

10) If time is measured in seconds (s) and space is measured in meters (m), then what are the units of geodesic curvature, κ?

No additional problems for 501 students this time.

Lecture 2 - Euler Curvature

Lecture given Tuesday Jan 21 — class shortened by snow.

2.1 Old calculus and new notation

2.1.1 Directional derivatives

In the new notational conventions, rows are parametrized with upper indices and columns with lower indices. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real-valued function. Its partial derivatives are given the notation

$$f_{,i} \triangleq \frac{\partial f}{\partial x^i}. \tag{2.1}$$

so that it is natural to consider its gradient to be a row vector:

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \ldots, \frac{\partial f}{\partial x^n} \right) = (f_{,1}, f_{,2}, \ldots, f_{,n}) \tag{2.2}$$

Letting

$$\vec{v} = \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} \tag{2.3}$$

be a vector, the rate of change of f in the $\vec{v}$-direction is just the matrix product

$$\left(\vec{\nabla} f\right)(\vec{v}) = (f_{,1}, f_{,2}, \ldots, f_{,n}) \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} = \sum_{i=1}^{n} f_{,i} v^i. \tag{2.4}$$

Consider another calculus problem. You are given a vector field

$$\vec{F}(x^1, \ldots, x^n) \;=\; \begin{pmatrix} F^1(x^1, \ldots, x^n) \\ \vdots \\ F^n(x^1, \ldots, x^n) \end{pmatrix} \;=\; \begin{pmatrix} F^1 \\ \vdots \\ F^n \end{pmatrix} \tag{2.5}$$

and you wish to determine how the field is changing, at the point $\underline{\vec{x}}$ in the direction $\vec{v}$. The answer, of course, is to take the Jacobian of $\vec{F}$

$$\begin{aligned} \vec{\nabla}\vec{F} \;&=\; \left(F^i_{,j}\right) \\ &=\; \begin{pmatrix} F^1_{,1} & F^1_{,2} & \cdots & F^1_{,n} \\ F^2_{,1} & F^2_{,2} & & F^1_{,n} \\ \vdots & & \ddots & \vdots \\ F^n_{,1} & F^n_{,2} & \cdots & F^n_{,n} \end{pmatrix} \end{aligned} \tag{2.6}$$

where we have defined $F^i_{,j} \triangleq \frac{\partial F^i}{\partial x^j}$, and to apply $\vec{v} = (v^1, \ldots, v^n)^T$:

$$\frac{d\vec{F}}{d\vec{v}} \;=\; \vec{\nabla}\vec{F}\cdot\vec{v} \;=\; \begin{pmatrix} F^1_{,1} & F^1_{,2} & \cdots & F^1_{,n} \\ F^2_{,1} & F^2_{,2} & & F^1_{,n} \\ \vdots & & \ddots & \vdots \\ F^n_{,1} & F^n_{,2} & \cdots & F^n_{,n} \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} \;=\; \begin{pmatrix} \sum F^1_{,j} v^j \\ \sum F^2_{,j} v^j \\ \vdots \\ \sum F^n_{,j} v^j \end{pmatrix} \tag{2.7}$$

A related question is the following. Given a function $F : \mathbb{R}^n \to \mathbb{R}$ and straight line $\gamma(t) = \vec{p} + \vec{v}t$ in $\mathbb{R}^2$, what are the first and second derivatives of $F \circ \gamma$? We have

$$\frac{dF}{d\vec{v}} \;=\; \vec{\nabla}F\;\vec{v} \;=\; (F_{,1},\, F_{,2},\, \ldots,\, F_{,n}) \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} \;=\; \sum_i F_{,i} v^i \tag{2.8}$$

The Hessian of F, namely

$$\vec{\nabla}^2 F \;=\; (F_{,ij}) \tag{2.9}$$

is a 2×2 array whose indexing indicates it is a column-column matrix (instead of the usual row-column matrix). This obviously can't be drawn, but consider the usefulness of our indexing scheme:

$$\frac{d}{d\vec{v}}\frac{dF}{d\vec{v}} \;=\; \sum_{i,j=1}^{n} F_{,ij} v^i v^j \;=\; \nabla^2 F \cdot \vec{v} \cdot \vec{v} \tag{2.10}$$

This is often denoted $\nabla^2 F(\vec{v}, \vec{v})$. In particular, any function $F : \mathbb{R} \to \mathbb{R}$ provides is with a bilinear form:

$$\nabla^2 F(\vec{v}, \vec{w}) \;=\; \sum_{i,j=1}^{n} F_{,ij} v^i w^j. \tag{2.11}$$

2.1.2 The Einstein sum convention

The Einstein sum convention consists of simply leaving off the sum sign, and whenever repeated upper and lower indices appear as a product, one knows to sum over them. So for instance

$$\begin{aligned} \frac{dF}{d\vec{v}} &= F_{,i}v^i \\ \nabla^2 F\left(\vec{v},\,\vec{w}\right) &= F_{,ij}v^iw^j. \end{aligned} \tag{2.12}$$

and we can also write the inner product

$$\langle \vec{v},\,\vec{w}\rangle \;=\; \vec{v}^i\vec{w}^j\delta_{ij} \tag{2.13}$$

Indeed the transpose of the column vector $\vec{v} = (v^i)$ is the row vector $(\delta_{ij}v^j)$.

2.2 Curves on a surface

First consider the issue of unit-parametrized, straight-line paths in $\mathbb{R}^2$—these are determined by a point $\vec{p} \in \mathbb{R}^2$ and a direction, encoded by an angle θ. Then a unit-parametrized path in $\mathbb{R}^2$ through $\vec{p} = \begin{pmatrix} p^1 \\ p^2 \end{pmatrix}$ is given by

$$\vec{\mu}(t) \;=\; \begin{pmatrix} p^1 + t\cos\theta \\ p^2 + t\sin\theta \end{pmatrix}. \tag{2.14}$$

If further clarity is needed, we can explicitly index $\vec{\mu}$ by its starting point $\vec{p}$ and direction θ:

$$\vec{\mu}_{\vec{p},\theta}(t) \;=\; \begin{pmatrix} p^1 + t\cos\theta \\ p^2 + t\sin\theta \end{pmatrix}. \tag{2.15}$$

Now consider the graph of a function $f(x^1, x^2)$. A path $\vec{\mu}_{\vec{p},\theta}(t)$ lifts to a graph on the surface, given by

$$\vec{\gamma}_{\vec{p},\theta}(t) \;=\; \begin{pmatrix} p^1 \;+\; t\cos\theta \\ p^2 \;+\; t\sin\theta \\ f(p^1 + t\cos\theta,\, p^2 + t\sin\theta) \end{pmatrix} \tag{2.16}$$

2.3 Euler's notions of surface curvature

Let $f(x^1, x^2)$ be some function, and let $(p^1, p^2, f(p^1, p^2))^T$ be a point on the graph, and consider the family of paths through this point:

$$\vec{\gamma}_\theta(t) \;=\; \begin{pmatrix} p^1 \;+\; t\cos\theta \\ p^2 \;+\; t\sin\theta \\ f\left(p^1 + t\cos\theta,\, p^2 + t\sin\theta\right) \end{pmatrix} \tag{2.17}$$

(where we have left $\vec{p}$ implicit).

Now for each θ the curve γ_θ lies on the surface and has its own geodesic curvature. Each path passes through the point $(p^1, p^2, f(p^1, p^2))^T$ at time zero, so we can record the curvature κ at that point. We define

$$\kappa(\vec{p},\, \theta) \;=\; \text{curvature of the path } \vec{\gamma}_\theta \text{ at time 0.} \tag{2.18}$$

Obviously κ will vary with θ. Fixing the point $\vec{p}$ and allowing θ to vary, we will find a largest and smallest curvature.

To arrive at Euler's notion of surface curvature, we must also require that the principal normal of these curves is also normal to the surface. When this is the case, the largest and smallest curvatures are called the *Eulerian principal curvatures* $\kappa_1(\vec{p})$ and $\kappa_2(\vec{p})$. Specifically

$$\begin{aligned} \kappa_1(\vec{p}) &= \sup_{\theta\in[0,\pi)} \kappa(\vec{p},\, \theta) \\ \kappa_2(\vec{p}) &= \inf_{\theta\in[0,\pi)} \kappa(\vec{p},\, \theta). \end{aligned} \tag{2.19}$$

From these, Euler determined two measures of the curvature of a surface at a point: the *Eulerian mean curvature* and the *Eulerian curvature*, defined to be, respectively,

$$\begin{aligned} M &= \frac{1}{2}\left(\kappa_1 \,+\, \kappa_2\right) \\ K &= \kappa_1\kappa_2. \end{aligned} \tag{2.20}$$

2.4 Exercises

1) (Practice with notation.) Suppose the 3×3 matrix (A^i_j) is given by $A^i_j = -1 + i + j$. Also let

$$\vec{v} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \vec{w} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \tag{2.21}$$

   a) Find the vector $(A^i_j v^j)$.

   b) Express $(\delta_{ij} v^j)$ as a vector (is it a row or a column vector?).

   c) Find the scalar $A^i_j \delta_{ik} v^j w^k$.

2) Consider the vector-valued function $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$\vec{F}(x^1,\, x^2,\, x^3) \;=\; \begin{pmatrix} x^1x^2 \\ \frac{1}{2}\left(\left(x^1\right)^2 + \left(x^2\right)^2 + x^2x^3\right) \\ x^1 - x^2 \end{pmatrix} \tag{2.22}$$

a) Determine $F^2_{,2}$ as a function of x^1, x^2 and x^3.

b) At the point $\underline{\vec{x}} = (1, 1, 1)^T$, how is the vector field changing in the direction $\vec{v} = (\sqrt{3}, \sqrt{3}, \sqrt{3})^T$?

3) Consider the surface given by the graph of $f(x^1, x^2) = \frac{1}{2}\left(\left(x^1\right)^2 + \frac{\left(x^2\right)^2}{10}\right)$. Simply set $\vec{p} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and determine the curvature $\kappa(\vec{p}, \theta)$ as a function of θ. Determine the principal curvatures at this point.

Problems due Thursday Jan 30.

Lecture 3 - Euler Curvature, continued

Lecture given Thursday Jan 23.

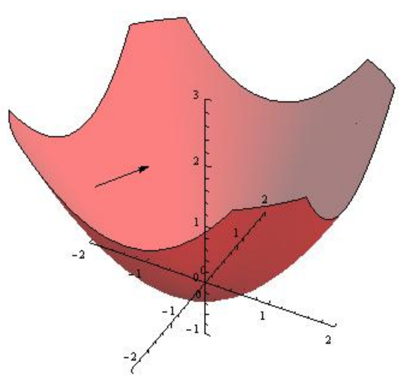
3.1 A conceptual look at Euler curvature

Formal definition of principal curvature

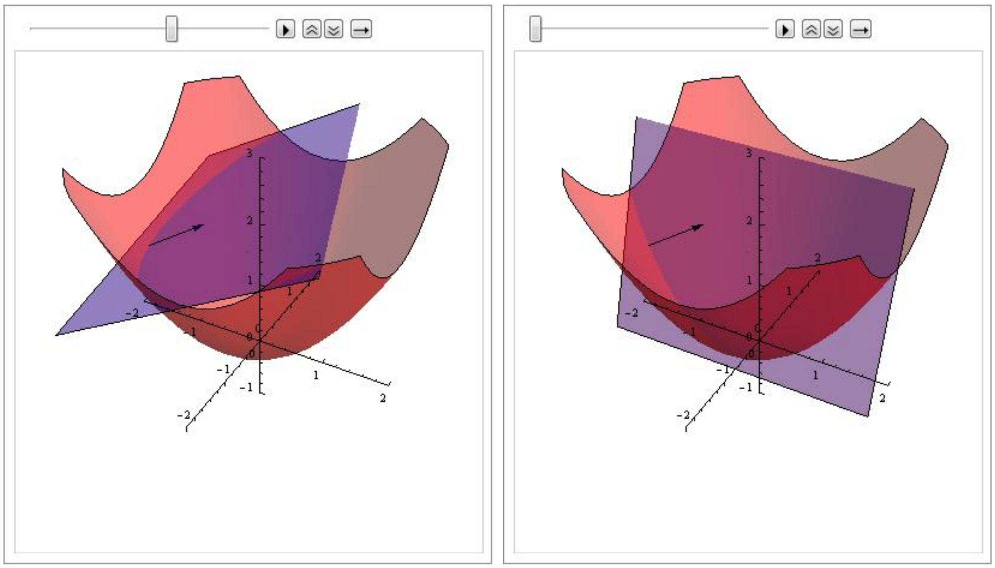
Euler's work was in surfaces in $\mathbb{R}^3$, so let's spend a moment looking into the geometry here.

Given a point $\vec{p} \in \mathbb{R}^3$ and a vector $\vec{v}$ based at $\vec{p}$, how many planes exist that both pass through $\vec{p}$ and contain the vector $\vec{v}$? The answer, of course, is half a circle's worth.

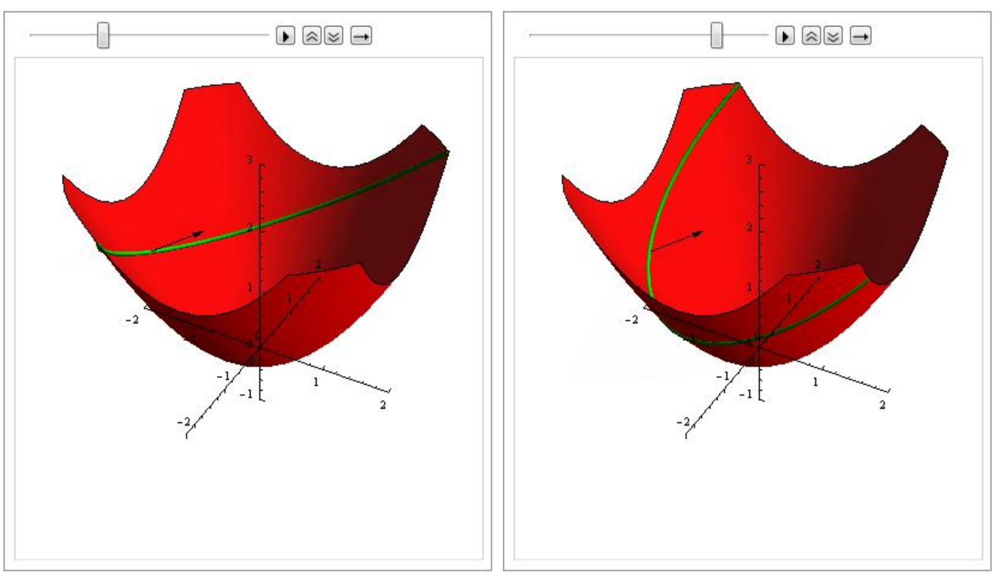
Now consider any (twice differentiable) surface Σ in $\mathbb{R}^3$, and let $\vec{p}$ be a point on Σ. At $\vec{p}$, the surface itself has a normal vector $\vec{N}$, so follows:



Now consider the family of planes that passes through $\vec{p}$ and that have a direction in common with $\vec{N}$. Two examples of such planes are as follows:



As you can see, each plane cuts out a curve that is both a plane-curve, and also a curve on the surface. At $\vec{p}$ itself, the tangent to each curve lies tangent to the surface, and because it is a plane curve, the normal to the curve must lie both along the plane and perpendicular to the unit tangent tangent, so each curve has as its normal the vector $\vec{N}$, the normal to the surface. Two examples of such curves are as follows:



Having fixed the point $\vec{p}$ and the normal vector $\vec{N}$, we obtain a family of curves passing through $\vec{p}$, and so that the principal normal of each curve at $\vec{p}$ is parallel to $\vec{N}$. At p, each curve has a curvature κ. We make one more note: we define κ to be positive if the curve's principal normal is $\vec{N}$, and define κ to be negative if the curve's principal normal is $-\vec{N}$.

Now we can precisely define principal curvatures. Among all such curves passing through $\vec{p}$, there will be a largest and a smallest curvature (possibly both negative!) . These are called the *principal curvatures*, and we label them

$$\kappa_1(\vec{p}) \quad \text{and} \quad \kappa_2(\vec{p}). \tag{3.1}$$

This geometrical formulation has two main points. First, it is completely independent of any coordinate system. Thus there is no question of ambiguity arising from choosing different origins, coordinate axes, or choosing spherical over rectangular coordinates, say.

The second point is even more significant. The reason for insisting that the principal normals of the paths themselves are parallel to the surface's normal is to unsure that, as much as possible, the paths are bending the way the surface is bending, and that the paths are not bending in any way *within* the surface itself.

Procedure for computing principal curvatures

Somehow we have to parametrize the curves we obtained from the slicings described above. Since everything is coordinate-independent, the first step is to choose a propitious coordinate system: choose a system so that the point under consideration, $\vec{p}$, is at the origin, and choose the x^1-x^2 coordinate plane to be tangent to the surface at $\vec{p}$.

In this case, the normal to the graph is simply the vector that points straight up along the z-axis: $\vec{N} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. The planes through the origin $\mathcal{O}$ that contain $\vec{N}$ can be parametrized by the angle they make with the x^1-axis, which we may call θ:

$$\begin{aligned} P(\theta) &= \text{the 2-plane through } \mathcal{O} \text{ spanned by } \vec{N} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \\ &= \text{all linear combinations of } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \end{aligned} \tag{3.2}$$

We can assume that the surface is given by the graph of a function $f(x^1, x^2)$. The next question is, how can we parametrize the path produced by the intersection of the graph $x^3 = f(x^1, x^2)$ and the plane $P(\theta)$? If we call this path $\vec{\gamma}_\theta$, then

$$\vec{\gamma}_\theta(t) = \begin{pmatrix} t\cos\theta \\ t\sin\theta \\ f\left(t\sin\theta,\, t\cos\theta\right) \end{pmatrix}. \tag{3.3}$$

3.2 Computation of curvatures, assuming $\vec{\nabla} f = 0$

3.2.1 Geodesic curvatures of curves on a surface

Now let Σ be a surface in $\mathbb{R}^3$ given by the graph of a function, meaning

$$\Sigma \;=\; \left\{ \begin{pmatrix} x^1 \\ x^2 \\ f(x^1,\, x^2) \end{pmatrix} \;\middle|\; \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \in \Omega \right\} \tag{3.4}$$

where Ω is some domain in $\mathbb{R}^2$. We have assumed the coordinate system is chosen so that $f(0,0) = 0$ and $\vec{\nabla} f = \vec{0}$ The paths $\vec{\gamma}_\theta(t)$, described in the previous section, are given by

$$\vec{\gamma}_\theta(t) \;=\; \begin{pmatrix} t\cos\theta \\ t\sin\theta \\ f\left(t\cos\theta,\; t\sin\theta\right) \end{pmatrix} \tag{3.5}$$

The velocity and acceleration vectors are

$$\begin{aligned} \frac{d\vec{\gamma}}{dt} \;&=\; \begin{pmatrix} \cos\theta \\ \sin\theta \\ \cos\theta\, f_{,1} \;+\; \sin\theta\, f_{,2} \end{pmatrix} \\ \frac{d^2\vec{\gamma}}{dt^2} \;&=\; \begin{pmatrix} 0 \\ 0 \\ \cos^2\theta\, f_{,11} \;+\; 2\cos\theta\sin\theta\, f_{,12} \;+\; \sin^2\theta\, f_{,22} \end{pmatrix} \end{aligned} \tag{3.6}$$

and we compute

$$\begin{aligned} \left\| \frac{d\vec{\gamma}}{dt} \right\|^2 \;&=\; 1 \;+\; \left(\cos\theta\, f_{,1} \;+\; \sin\theta\, f_{,2}\right)^2 \\ \left\| \frac{d^2\vec{\gamma}}{dt^2} \right\|^2 \;&=\; \left(\cos^2\theta\, f_{,11} \;+\; 2\cos\theta\sin\theta\, f_{,12} \;+\; \sin^2\theta\, f_{,22}\right)^2 \\ \left(\frac{d^2\vec{\gamma}}{dt^2} \cdot \frac{d\vec{\gamma}}{dt} \right)^2 \;&=\; \left(\cos\theta\, f_{,1} \;+\; \sin\theta\, f_{,2}\right)^2 \left(\cos^2\theta\, f_{,11} \;+\; 2\cos\theta\sin\theta\, f_{,12} \;+\; \sin^2\theta\, f_{,22}\right)^2 \end{aligned} \tag{3.7}$$

so that

$$\kappa \;=\; \frac{\cos^2\theta\, f_{,11} \;+\; 2\cos\theta\sin\theta\, f_{,12} \;+\; \sin^2\theta\, f_{,22}}{\left(1 \;+\; \left(\cos\theta\, f_{,1} \;+\; \sin\theta\, f_{,2}\right)^2\right)^{\frac{3}{2}}} \tag{3.8}$$

Notice that we dropped the absolute value. This is to account for the sign of κ, as discussed in the previous section. Finally, our assumption that $\vec{\nabla} f = \vec{0}$ forces $f_{,1} = f_{,2} = 0$, so we obtain

$$\kappa(\theta) \;=\; \cos^2\theta\, f_{,11} \;+\; 2\cos\theta\sin\theta\, f_{,12} \;+\; \sin^2\theta\, f_{,22} \tag{3.9}$$

3.2.2 Principal curvatures

At our chosen point (the origin), we now have κ as a function of θ being

$$\kappa(\theta) \;=\; \cos^2\theta\, f_{11} \;+\; 2\cos\theta\sin\theta\, f_{12} \;+\; \sin^2\theta\, f_{22}. \tag{3.10}$$

We want to extremize κ, so we take a derivative and set to zero:

$$\begin{aligned} 0 \;&=\; \frac{d\kappa}{d\theta} \\ &=\; -2\cos\theta\sin\theta\, f_{11} \;+\; 2\left(\cos^2\theta \;-\; \sin^2\theta\right)\; f_{12} \;+\; 2\cos\theta\sin\theta\, f_{22} \\ &=\; -\sin(2\theta)\,(f_{11} \;-\; f_{22}) \;+\; 2\cos(2\theta) f_{12} \end{aligned} \tag{3.11}$$

We obtain

$$\cot(2\theta) \;=\; \frac{f_{11} \;-\; f_{22}}{2 f_{12}}. \tag{3.12}$$

Noting that we must accept all solutions for $\theta \in [0, \pi)$ so that 2θ may be in $[0, 2\pi)$, we therefore obtain two solutions for θ, characterized by

$$\begin{aligned} \sin(2\theta) \;&=\; \pm\frac{2 f_{12}}{\sqrt{(f_{11} - f_{22})^2 \;+\; 4\,(f_{12})^2}} \\ \cos(2\theta) \;&=\; \pm\frac{f_{11} \;-\; f_{22}}{\sqrt{(f_{11} - f_{22})^2 \;+\; 4\,(f_{12})^2}}. \end{aligned} \tag{3.13}$$

From the half-angle formulas we get

$$\begin{aligned} \sin^2\theta \;&=\; \frac{1 - \cos(2\theta)}{2} \;=\; \frac{\mp(f_{11} - f_{22}) + \sqrt{(f_{11} - f_{22})^2 + 4\,(f_{12})^2}}{2\sqrt{(f_{11} - f_{22})^2 + 4\,(f_{12})^2}} \\ \cos^2\theta \;&=\; \frac{1 + \cos(2\theta)}{2} \;=\; \frac{\pm(f_{11} - f_{22}) + \sqrt{(f_{11} - f_{22})^2 + 4\,(f_{12})^2}}{2\sqrt{(f_{11} - f_{22})^2 + 4\,(f_{12})^2}} \end{aligned} \tag{3.14}$$

Putting this in to (3.10) we find that the two principal curvatures are

$$\begin{aligned} \kappa_1 \;&=\; \frac{f_{11} \;+\; f_{22} \;+\; \sqrt{(f_{11} - f_{22})^2 \;+\; 4\,(f_{12})^2}}{2} \\ \kappa_2 \;&=\; \frac{f_{11} \;+\; f_{22} \;-\; \sqrt{(f_{11} - f_{22})^2 \;+\; 4\,(f_{12})^2}}{2}. \end{aligned} \tag{3.15}$$

3.2.3 Euler's first theorem and Gauss' second derivative test

Theorem 3.2.1 (Euler, 1760) *The two paths that represent the principal curvatures of a surface Σ at any point $\vec{p}$ intersect at right angles.*

Proof. The principal curvatures at $\vec{p}$, and the paths that through $\vec{p}$ that carry those curvatures, are independent of the coordinate system that is chosen. Thus we can select a special coordinate system, where $\vec{p}$ is the origin, and the x^1-x^2 plane is tangent to the surface at $\vec{p}$, and using these coordinates we can extremize $\kappa = \kappa(\theta)$, where, recall, $\kappa(\theta)$ was defined to be the curvature of the path $\vec{\gamma}_\theta(t)$ at $t = 0$.

We found that the two angles θ_1 and θ_2 that extremize $\kappa(\theta)$ are characterized by the fact that each is in $[0, \pi)$, and that

$$\begin{aligned}
\sin(2\theta_1) &= +\frac{2f_{,12}}{\sqrt{(f_{,11} - f_{,22})^2 - 4(f_{,12})^2}}, \quad \cos(2\theta_1) = +\frac{f_{,11} - f_{,22}}{\sqrt{(f_{,11} - f_{,22})^2 - 4(f_{,12})^2}} \\
\sin(2\theta_2) &= -\frac{2f_{,12}}{\sqrt{(f_{,11} - f_{,22})^2 - 4(f_{,12})^2}}, \quad \cos(2\theta_2) = -\frac{f_{,11} - f_{,22}}{\sqrt{(f_{,11} - f_{,22})^2 - 4(f_{,12})^2}}
\end{aligned} \tag{3.16}$$

The angles $2\theta_1$ and $2\theta_2$ are therefore π different from each other, so θ_1 and θ_2 are $\pi/2$ different. Therefore the two paths are orthogonal at $\vec{p}$, and remain so regardless of whatever coordinate system is chosen. □

We have defined *mean curvature* to be the average of the principal curvatures: $M = \frac{1}{2}(\kappa_1 + \kappa_2)$, and *Euler curvature* to be the product of the principal curvatures: $K = \kappa_1\kappa_2$. Working, as above, under the assumption that $\vec{\nabla} f = 0$, we compute

$$\begin{aligned}
M &= \frac{1}{2}(f_{11} + f_{22}) \\
K &= f_{11}f_{22} - (f_{12})^2.
\end{aligned} \tag{3.17}$$

Notice that these expressions can be seen in terms of the Hessian matrix: $M = \frac{1}{2}Tr(\vec{\nabla}^2 f)$ and $K = \det(\vec{\nabla}^2 f)$.

Gauss noticed that, at any extreme point of the graph of a function $f(x^1, x^2)$ (meaning that $f_{,1} = f_{,2} = 0$), K could be used to determine if the extreme point was a saddle, a maximum, or a minimum. Specifically, at a max or a min, the quantity K is either positive or zero. At a saddle, the quantity K is negative or zero. Gauss arrived at the following rule:

$$\begin{array}{llll}
\text{if } K > 0 & \text{and} & M > 0 & \text{the function has a minimum} \\
\text{if } K > 0 & \text{and} & M < 0 & \text{the function has a maximum} \\
\text{if } K < 0 & \multicolumn{3}{l}{\text{the function has a saddle}} \\
\text{if } K = 0 & \multicolumn{3}{l}{\text{the test gives no information.}}
\end{array} \tag{3.18}$$

3.3 Exercises

1) This exercise is to illustrate the necessity of rechoosing the coordinate system before computing the principal curvatures. Consider the unit 2-sphere, given by $(x^1)^2 +$

$(x^2)^2 + (x^3)^2 = 1$ in x^1, x^2, x^3 coordinates. Let $\vec{p} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^T$. We can consider the upper-half sphere to be the graph of $f(x^1, x^2) = \sqrt{1 - (x^1)^2 - (x^2)^2}$. Let $\vec{\gamma}_\theta$ be the path, passing through $\vec{p}$ at time $t = 0$ whose osculating plane is parallel to the x^3-axis and makes the angle θ with the x^1-axis. This path is given by

$$\vec{\gamma}_\theta(t) \quad = \quad \begin{pmatrix} \frac{1}{\sqrt{2}} + t\cos\theta \\ t\sin\theta \\ f\left(\frac{1}{\sqrt{2}} + t\cos\theta,\ t\sin\theta\right) \end{pmatrix}. \tag{3.19}$$

a) Find the curvature κ of the path $\vec{\gamma}_\theta$, as a function of θ.

b) What is the maximum and the minimum of the function $\kappa(\theta)$ from (*a*)?

c) For the unit sphere, we know that $K = 1$ and $M = 1$ (see problem (2)). However if you were to find κ_1 and κ_2 from (*b*), you would compute larger values for both. Explain why; especially explain why the spurious values from this problem are larger and not smaller than the true values.

2) Consider the sphere of problem (1), along with the point $\vec{p} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^T$. Go through the whole process of determining the mean and Eulerian curvatures at $\vec{p}$, including re-choosing the coordinate system, compute κ as a function of θ, and so on. You must describe the method step by step, but if you use some common sense and your basic knowledge of spheres, this won't be a difficult problem. Was there anything special about the point $\vec{p}$?

3) Consider the surface obtained by graphing $f(x^1, x^2) = x^1x^2$. Find K and M at the point $(0, 0, 0)^T$ on the surface. Explain why κ_1 and κ_2 have opposite signs.

4) Given our computations (3.10) and (3.13), verify formula (3.15).

5)* Given any $n \times n$ symmetric matrix $A = (a^i_j)$, prove that its eigenvalues are real and non-degenerate, and it has orthogonal eigenvectors.

6)* Euclidean space $\mathbb{R}^n$ is the n-dimensional vector space with distance measured by the Pythagorean theorem, infinitesimally expressed by

$$ds \;=\; \sqrt{(dx^1)^2 + \ldots + (dx^n)^2}. \tag{3.20}$$

Now define $\mathbb{R}^{n,k}$ to be the $(n+k)$-dimensional vector space with distance measured (infinitesimally) by

$$ds \;=\; \sqrt{(dx^1)^2 + \ldots + (dx^n)^2 - (dx^{n+1})^2 - \ldots - (dx^{n+k})^2}. \tag{3.21}$$

In particular, the spaces $\mathbb{R}^{1,k}$ (or sometimes $\mathbb{R}^{n,1}$) are the Lorentzian vector spaces of special relativity. Sketch the following paths, and compute their lengths as a function of t

$a)\ \vec{\gamma}(t) = \begin{pmatrix} t \\ 1 \end{pmatrix}$ in $\mathbb{R}^{1,1}$

$b)\ \vec{\gamma}(t) = \begin{pmatrix} t \\ t \end{pmatrix}$ in $\mathbb{R}^{1,1}$

$c)\ \vec{\gamma}_\epsilon(t) = \begin{pmatrix} \epsilon t \\ t \end{pmatrix}$ (as a function of ϵ and t) in $\mathbb{R}^{1,1}$

$d)\ \vec{\gamma}(t) = \begin{pmatrix} t \\ \cos(t) \\ \sin(t) \end{pmatrix}$, $t \in [0, t_0]$, any t_0, in $\mathbb{R}^{1,2}$

Exercises marked with a "∗" are for 501 students.

Problems due Thursday Jan 30.

Lecture 4 - Euler's Theorem. The Gauss Map.

4.1 Aside: Symmetric 2×2 matrices

Let A be a matrix of the form

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix}. \tag{4.1}$$

Any symmetric matrix has real, orthogonal eigenvalues. Further, computation reveals that the two eigenvectors and corresponding eigenvectors are

$$\begin{aligned} \lambda_1 &= \frac{a+b+\sqrt{(a-b)^2\ +\ 4c^2}}{2}, \qquad \vec{v}_1 = \begin{pmatrix} 2c \\ \\ -a+b+\sqrt{(a-b)^2+4c^2} \end{pmatrix} \\ \lambda_2 &= \frac{a+b-\sqrt{(a-b)^2\ +\ 4c^2}}{2}, \qquad \vec{v}_2 = \begin{pmatrix} 2c \\ \\ -a+b-\sqrt{(a-b)^2+4c^2} \end{pmatrix} \end{aligned} \tag{4.2}$$

4.2 Euler's Second Theorem

Consider again the surface Σ, with a point $\vec{p} \in \Sigma$, and a choice of coordinate systems that makes $\vec{p}$ the origin with the x^3-axis parallel to the normal $\vec{N}$ at $\vec{p}$.

At $\vec{p}$ denote the principal directions by $\vec{V}_1$ and $\vec{V}_2$. These are the initial velocities of the paths that carry the principal curvatures. From (3.14) we easily compute

$$\begin{aligned} \vec{V}_1 &= \begin{pmatrix} 2f_{12} \\ \\ -(f_{11}-f_{22})+\sqrt{(f_{11}-f_{22})^2+4(f_{12})^2} \end{pmatrix} \\ \vec{V}_2 &= \begin{pmatrix} 2f_{12} \\ \\ -(f_{11}-f_{22})-\sqrt{(f_{11}-f_{22})^2+4(f_{12})^2} \end{pmatrix} \end{aligned} \tag{4.3}$$

which of course are the eigenvectors of the Hessian $\vec{\nabla}^2 f$. We make a final computation before wrapping things up. Consider a path along the graph through $\vec{p}$ that makes an angle of φ with the $\vec{V}_1$-direction. Then using (3.15)

$$\cos^2\varphi\,\kappa_1 \;+\; \sin^2\varphi\,\kappa_2 \;=\; \frac{f_{11} \;+\; f_{22} \;+\; \cos(2\varphi)\sqrt{(f_{11}-f_{22})^2 \;+\; 4\,(f_{12})^2}}{2} \tag{4.4}$$

Recall that we used θ to denote the angle with respect to the x^1-axis in our special coordinate system. Now φ and θ are related by an offset that we shall call θ_1, which is just the angle $\vec{V}_1$ makes with the x^1-axis, whose sine and cosine are given by (3.14). We compute

$$\begin{aligned}
\cos(2\varphi) \;&=\; \cos(2\theta - 2\theta_1) \\
&=\; \cos(2\theta)\cos(2\theta_1) \;+\; \sin(2\theta)\sin(2\theta_1) \\
&=\; \cos(2\theta)\frac{f_{11}-f_{22}}{\sqrt{(f_{11}-f_{22})^2 \;+\; 2\,(f_{12})^2}} \;+\; \sin(2\theta)\frac{2f_{12}}{\sqrt{(f_{11}-f_{22})^2 \;+\; 2\,(f_{12})^2}}
\end{aligned} \tag{4.5}$$

so that

$$\begin{aligned}
\cos^2\varphi\,\kappa_1 \;+\; \sin^2\varphi\,\kappa_2 \;&=\; \frac{f_{11} \;+\; f_{22} \;+\; \cos(2\theta)\,(f_{11}-f_{22}) \;-\; 2\sin(2\theta)\,f_{12}}{2} \\
&=\; \frac{1+\cos(2\theta)}{2}f_{11} \;+\; \frac{1-\cos(2\theta)}{2}f_{22} \;+\; \sin(2\theta)\,f_{12} \\
&=\; \cos^2\theta\,f_{11} \;+\; \sin(2\theta)\,f_{12} \;+\; \sin^2\theta\,f_{22} \;=\; \kappa
\end{aligned} \tag{4.6}$$

This is the *Euler curvature formula*, which is Euler's second theorem.

Theorem 4.2.1 (Euler, 1760) *Let κ_1 and κ_2 be the principal curvatures at a point $\vec{p}$ on a surface, and let $\vec{V}_1$, $\vec{V}_2$ be the corresponding principal directions. Then $\vec{V}_1$ and $\vec{V}_2$ are orthogonal, and the direction along the surface that makes an angle of φ with $\vec{V}_1$ has curvature*

$$\kappa(\varphi) \;=\; \cos^2(\varphi)\;\kappa_1 \;+\; \sin^2(\varphi)\;\kappa_2. \tag{4.7}$$

4.3 Some geometry of level-sets

4.3.1 Gauss' idea

Recall the Frenet formulas

$$\frac{d}{ds}\begin{pmatrix}\vec{T}\\ \vec{N}\\ \vec{B}\end{pmatrix} \;=\; \begin{pmatrix}0 & \kappa & 0\\ -\kappa & 0 & \tau\\ 0 & -\tau & 0\end{pmatrix}\begin{pmatrix}\vec{T}\\ \vec{N}\\ \vec{B}\end{pmatrix} \tag{4.8}$$

In Euler's methodology, the curves used to study the surface are also plane curves, so $\tau = 0$. This means that

$$\begin{aligned} \frac{d\vec{T}}{ds} &= \kappa \vec{N} \\ \frac{d\vec{N}}{ds} &= -\kappa \vec{T}. \end{aligned} \tag{4.9}$$

for each of the curves. So instead of studying the tangent vectors of these paths, we could study the way the normal vectors change. However these normal vectors are chosen to coincide with the normal to the surface itself.

Gauss' idea was to study the normal vector to the surface itself, without the complication of having to choose specific curves on the surface. Gauss' methods, which we will spend the next few lectures on, have the added advantage of being easily applicable in higher dimension.

4.3.2 Level-sets and normals

We have been considering surfaces given by the graph of a function, but in the study of geometry it is often better to study *level sets* of functions. This allows us to study objects that cannot be expressed as the graph of anything. For instance, consider the objects

$$\begin{aligned} F(x^1,x^2,x^3) &= \left(\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2\right), \quad F(x^1,x^2,x^3) \;=\; c \\ F(x^1,x^2,x^3) &= \left(\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2 - 1\right)\left(x^1 + x^2 - x^3\right), \quad F(x^1,x^2,x^3) \;=\; 0. \end{aligned} \tag{4.10}$$

The first is the sphere of radius $\sqrt{c}$, and the second is the union of sphere of radius 1 with the plane $x^3 = x^1 + x^2$.

A hypersurface in $\mathbb{R}^n$ will, from now on, be considered to be a level-set of the form $F = 0$ for some function $F : \mathbb{R}^n \to \mathbb{R}$. The gradient of a function F is always orthogonal to its level-sets (as long as $|\vec{\nabla}F|$ is non-zero), so we have the two unit normals:

$$\hat{n} \;=\; \pm\frac{\vec{\nabla}F}{\left|\vec{\nabla}F\right|}. \tag{4.11}$$

A surface $F = 0$ is called *non-singular* at $\vec{p} = (x^1, \ldots, x^n)^T$ if (in addition to $F(\vec{p}) = 0$) we have $\vec{\nabla}F(\vec{p}) \neq \vec{0}$. A point on the surface $F = 0$ is singular if also $\vec{\nabla}F = \vec{0}$. Note that the system of equations $F = 0$ and $\vec{\nabla}F = \vec{0}$ is overdetermined, meaning you usually can't find any critical points. This is a primitive version of Sard's theorem.

4.4 The Gauss Map

Let M^{n-1} be the $(n-1)$-dimensional surface in $\mathbb{R}^n$ given by a non-singular level set $F = 0$, where $F : \mathbb{R}^n \to \mathbb{R}$. The Gauss map, $\mathcal{G}$, is simply the map that sends each point $p \in M^{n-1}$ to the normal vector at that point. Since any normal vector lies on the sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$, this map can be thought of as a map

$$\mathcal{G} : M^{n-1} \longrightarrow \mathbb{S}^{n-1}. \tag{4.12}$$

We have a formula for the unit normal to a surface, provided the surface is given by the level-set $F = 0$. Let's look at a few examples of the Gauss map for surfaces $\Sigma \in \mathbb{R}^3$.

Example: The Gauss map for the flat 2-plane

Choosing $\Sigma \subset \mathbb{R}^3$ to be the x^1-x^2 plane, the normal vector field $\vec{N}$ is the constant

$$\vec{N} = (0,\, 0,\, 1)\,. \tag{4.13}$$

The image of the Gauss map is therefore simply the north pole on the 2-sphere.

Example: The Gauss map for the cylinder.

Consider the function $F(x^1, x^2, x^3) = \left(x^1\right)^2 + \left(x^2\right)^2 - 1$. The level-set

$$\Sigma = \{(x^1, x^2, x^3)^T \mid F(x^1, x^2, x^3) = 0\} \tag{4.14}$$

is simply the cylinder of radius 1 that is translated up and down the x^3-axis. If $\vec{p} = (p^1, p^2, p^3)^T \in \Sigma$, then the unit normal is

$$\frac{\nabla F}{|\nabla F|}(p^1, p^2, p^3) \;=\; \frac{\left(2p^1,\, 2p^2,\, 0\right)}{2\sqrt{\left(p^1\right)^2 + \left(p^2\right)^2}} \;=\; \left(p^1,\, p^2\right). \tag{4.15}$$

The image of the Gauss map is therefore the equator of the unit sphere.

Example: The Gauss map for the sphere of radius r.

The point $\vec{p} = (p^1, p^2, p^3)^T$ is on the sphere of radius r if $\left(p^1\right)^2 + \left(p^2\right)^2 + \left(p^3\right)^2 = r^2$. The unit normal at $\vec{p}$ is quite simply $\vec{p}/|\vec{p}|$. The image of the Gauss map is therefore the entire sphere.

Indeed, the Gauss map on the unit 2-sphere can be considered the identity map.

Example: The Gauss map for a paraboloid.

Let $F(x^1, x^2, x^3) = x^3 - \frac{1}{2}\left(\left(x^1\right)^2 + \left(x^2\right)^2\right)$, and set $\Sigma = \{F = 0\}$. The unit normal is

$$\frac{\nabla F}{|\nabla F|} \;=\; \frac{(-x^1, -x^2, 1)}{\sqrt{1 + (x^1)^2 + (x^2)^2}} \;=\; \frac{(-x^1, -x^2, 1)}{\sqrt{1 + 2x^3}} \tag{4.16}$$

Notice that the inner product $\left\langle (0,0,1), \frac{\nabla F}{|\nabla F|} \right\rangle > 0$ at all times. This means that the angle between the unit normal and the vertical vector is always acute, so the image of the Gauss map is within the upper hemisphere. It is easy to see that the image in indeed the entire (open) upper hemisphere.

4.5 Exercises

1) Verify the formulas in (4.2). Examine the cases that $a = b$ and that $c = 0$.

2) Consider the case of "hypersurfaces" in $\mathbb{R}^2$ (that is, curves). Specifically, consider the level sets $F_c(x^1, x^2) = 0$, where F_c is the cubic, parametrized by c, given by:

$$F_c(x^1,\, x^2) \;=\; (x^2)^2 \;-\; \frac{1}{3}(x^1)^3 \;+\; x^1 \;-\; c. \tag{4.17}$$

For generic c, determine all singular points of the curve. Graph several representative curves, and explicitly label all singular points.

3) Consider the family of surfaces in $\mathbb{R}^3$, parametrized by c, given by the quadratic polynomials $F_c(x^1, x^2, x^3) = -\left(x^1\right)^2 - \left(x^2\right)^2 + \left(x^3\right)^2 + c$. Sketch several representative surfaces for c in the range, say, $[-2, 2]$. For each c, determine all singular points and label them on your graph.

*4) Prove that mean curvature is not just the mean of the principal curvatures, but the mean of all curvatures as θ varies through the entire circle of directions.

**5) Can Euler curvature can be expressed as a total integral for $\theta \in [0, 2\pi)$, as you expressed mean curvature in (3)?

6) Consider the parabola in $\mathbb{R}^2$; this is given by the zero-set of $F = -x^2 + \left(x^1\right)^2$. Sketch the image of the Gauss map, and label, on $\mathbb{S}^1$, the images of the points $(0,0)^T$, $(1,1)^T$ and $(-1,1)^T$.

7) Compute, explicitly, the Gauss map $\mathcal{G}$ of $\Sigma = \{F = 0\}$, where $F(x^1, x^2, x^3) = x^3 - \frac{1}{2}\left(\left(x^1\right)^2 - \left(x^2\right)^2\right)$. Label the images of the points $(0,0,0)^T$, $(1,1,0)^T$, and $(\sqrt{3}, 1, 1)^T$?. As a set, what is the image $\mathcal{G}(\Sigma)$?

*8) Let Σ be a surface in $\mathbb{R}^3$ that is the graph of a (twice continuously differentiable) function. Prove that the image of the Gauss map is contained in a hemisphere of $\mathbb{S}^2$.

Exercises marked with a "$*$" are for 501 students. Exercises marked with a "$**$" are optional.

Problems due Thursday Feb 6.

Lecture 5 - The First and Second Fundamental Forms

5.1 Gauss' definition of curvature

If Σ^2 is a surface in $\mathbb{R}^3$, the Gauss map $\mathcal{G} : \Sigma^2 \to \mathbb{R}^3$ distorts the shape of Σ rather severely. In particular, consider a domain $\mathcal{R} \subset \Sigma$, and consider how it maps to $\mathbb{S}^2$. Its image is $\mathcal{G}(\mathcal{R})$. Now consider a point $p \in \Sigma$, and consider all domains $\mathcal{R}_p$ in Σ that contain p. The *Gaussian curvature* of Σ at p, denoted K_G, was defined by Gauss to be

$$K_G(p) \;=\; \lim_{\mathcal{R}_p \to \{p\}} \frac{Area(\mathcal{G}(\mathcal{R}_p))}{Area(\mathcal{R}_p)} \tag{5.1}$$

Now this definition is not rigorous, because exactly what is meant by $\lim_{\mathcal{R}_p \to \{p\}}$ is problematic. Possibly a clear notion of this limit can be obtained, but it would require an involved argument to prove that the limit doesn't depend on the way that the domains $\mathcal{R}_p$ shrink down to the set $\{p\}$. Further below, we take a more modern approach to all this.

5.2 Tangent Spaces and the First Fundamental Form

Let Σ be the hyper-surface in $\mathbb{R}^n$ given by $F(x^1, \ldots, x^n) = 0$. A point $p \in \Sigma$ is said to be a singular point provided $\vec{\nabla} F = \vec{0}$. If the point p is non-singular, then it is possible to determine a tangent space to Σ at p. This is denoted $T_p\Sigma$, and defined to be the following vector space:

$$T_p\Sigma \;\triangleq\; \left\{ \text{vectors } \vec{v} \text{ based at } p \text{ such that } \quad \left\langle \vec{V},\, \vec{N} \right\rangle \;=\; 0 \right\}. \tag{5.2}$$

Gauss defined a bilinear form, *the first fundamental form* to be simply the restriction of the inner product to the tangent space:

$$I\left(\vec{v},\, \vec{w}\right) \;=\; \left\langle \vec{v},\, \vec{w} \right\rangle, \quad \text{provided} \quad \vec{v}, \vec{w} \in T_p\Sigma. \tag{5.3}$$

Often the form I is denoted by g.

5.3 The Second Fundamental Form

The Gauss map is a map that sends any non-singular surface $\Sigma \in \mathbb{R}^n$ to the unit sphere $\mathbb{S}^{n-1}$. Gauss wants to look at how regions in Σ map to regions on $\mathbb{S}^{n-1}$, but his method is not rigorously defined. Still, if we want to look at the infinitesimal way that areas are distorted, we look at the Jacobian of the Gauss map. This is the matrix

$$\vec{\nabla}\left(\frac{\vec{\nabla}F}{\left|\vec{\nabla}F\right|}\right) = \left(\left(\frac{F_i}{|\nabla F|}\right)_j\right)_{i,j=1}^{n} \tag{5.4}$$

This Jacobian is useful in a number of ways. For instance if you want to know how the normal vector $\frac{\vec{\nabla}F}{|\vec{\nabla}F|}$ is changing in the direction $\vec{v}$, you simply compute

$$\vec{\nabla}\left(\frac{\vec{\nabla}F}{\left|\vec{\nabla}F\right|}\right)(\vec{v})\,. \tag{5.5}$$

We define the second fundamental form to be this Jacobian, restricted to the tangent space at $p \in \Sigma$:

$$\begin{aligned} II &\triangleq \vec{\nabla}\left(\frac{\vec{\nabla}F}{\left|\vec{\nabla}F\right|}\right)\Bigg|_{T_p\Sigma} \\ II\,(\vec{v},\,\vec{w}) &= \nabla\left(\frac{\nabla F}{|\nabla F|}\right)(\vec{v},\,\vec{w})\,, \qquad \text{for} \quad \vec{v},\vec{w} \in T_pM \end{aligned} \tag{5.6}$$

In coordinates, after letting $\vec{v} = (v^i)$ and $\vec{w} = (w^i)$ be vectors in $T_p\Sigma$, we have

$$\begin{aligned} II\,(\vec{v},\,\vec{w}) &= \nabla\left(\frac{\nabla F}{|\nabla F|}\right)(\vec{v},\,\vec{w}) \\ &= \left(\frac{F_{,i}}{\left|\vec{\nabla}F\right|}\right)_{,j} v^i\,w^j \\ &= \frac{\partial}{\partial x^i}\left(\frac{1}{\left|\vec{\nabla}F\right|}\frac{\partial F}{\partial x^j}\right) v^i\,w^j \end{aligned} \tag{5.7}$$

5.4 Examples

Consider the surface $\Sigma \subset \mathbb{R}^3$ given by $F = 0$ where $F(x^1, x^2, x^3) = \left(x^1\right)^2 + \left(x^2\right)^2 - \left(x^3\right)^2 - 1$. Assuming $p \in \Sigma$, we will determine $\vec{N}(p)$, a basis for $T_p\Sigma$, and the two fundamental forms expressed in this basis.

The normal vector is

$$\begin{aligned}\vec{N} &= \frac{\left(2x^1,\, 2x^2,\, -2x^3\right)}{\sqrt{4\left(x^1\right)^2 + 4\left(x^2\right)^2 + 4\left(x^3\right)^2}} \\ &= \frac{\left(x^1,\, x^2,\, -x^3\right)}{\sqrt{\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2}}.\end{aligned} \tag{5.8}$$

At the point $p = (p^1, p^2, p^3)^T$, we can simplify this (slightly) to

$$\vec{N}(p) = \frac{\left(x^1,\, x^2,\, -x^3\right)}{\sqrt{1 + 2\left(x^3\right)^2}} \tag{5.9}$$

Let $\vec{Z} = (0, 0, 1)$ be the unit vector in the x^3-direction, and let $\vec{R} = (-x^2, x^1, 0)$ be the radial vector field. At the moment, there is no reason to believe that $\vec{Z}, \vec{R} \in T_p\Sigma$. But these vectors are independent of $\vec{N}$, so we can apply a normalization process. Set

$$\begin{aligned}\vec{v}_1(p) &= \vec{Z} - \frac{\left\langle \vec{Z},\, \vec{N} \right\rangle}{\left\langle \vec{N},\, \vec{N} \right\rangle} \vec{N} \\ &= (0,0,1) - \frac{\left\langle (0,0,1),\, (p^1, p^2, -p^3) \right\rangle}{\sqrt{1 + 2\left(p^3\right)^2}} \cdot \frac{\left(p^1,\, p^2,\, -p^3\right)}{\sqrt{1 + 2\left(p^3\right)^2}} \\ &= \frac{\left(p^1p^3,\, p^2p^3,\, 1 + \left(p^3\right)^2\right)}{1 + 2\left(p^3\right)^2}\end{aligned} \tag{5.10}$$

and

$$\begin{aligned}\vec{v}_2(p) &= \vec{R} - \frac{\left\langle \vec{R},\, \vec{N} \right\rangle}{\left\langle \vec{N},\, \vec{N} \right\rangle} \vec{N} \\ &= \left(-p^2,\, p^1,\, 0\right) - \frac{\left\langle (-p^2, p^1, 0),\, (p^1, p^2, -p^3) \right\rangle}{\sqrt{1 + 2\left(p^3\right)^2}} \cdot \frac{\left(p^1,\, p^2,\, -p^3\right)}{\sqrt{1 + 2\left(p^3\right)^2}} \\ &= \left(-p^2,\, p^1,\, 0\right)\end{aligned} \tag{5.11}$$

(so it turns out that $\vec{v}^2$ was orthogonal to $\vec{N}$ after all). Now the two vectors $\vec{v}_1$ and $\vec{v}^2$ are orthogonal to $\vec{N}$ (and indeed are orthogonal to each other), so that

$$T_p\Sigma = span\ \{\vec{v}_1,\, \vec{v}_2\}. \tag{5.12}$$

In this basis, we compute the first fundamental form:

$$I(p) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} = \begin{pmatrix} \langle \vec{v}_1, \vec{v}_1 \rangle & \langle \vec{v}_1, \vec{v}_2 \rangle \\ \langle \vec{v}_2, \vec{v}_1 \rangle & \langle \vec{v}_2, \vec{v}_2 \rangle \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1+\left(p^3\right)^2}{1+2(p^3)^2} & 0 \\ 0 & 1+\left(p^3\right)^2 \end{pmatrix} \tag{5.13}$$

For the second fundamental form, the easiest thing to do is to compute $\nabla(\nabla F/|\nabla F|)$ as a 3×3 matrix in the variables x^1, x^2, and x^3, and then evaluate at $\vec{v}_1$ and $\vec{V}_2$. We compute

$$\nabla\left(\frac{\nabla F}{|\nabla F|}\right) = \begin{pmatrix} \frac{\partial}{\partial x^1}\left(\frac{1}{|\nabla F|}\frac{\partial F}{\partial x^1}\right) & \frac{\partial}{\partial x^1}\left(\frac{1}{|\nabla F|}\frac{\partial F}{\partial x^2}\right) & \frac{\partial}{\partial x^1}\left(\frac{1}{|\nabla F|}\frac{\partial F}{\partial x^3}\right) \\ \frac{\partial}{\partial x^2}\left(\frac{1}{|\nabla F|}\frac{\partial F}{\partial x^1}\right) & \frac{\partial}{\partial x^2}\left(\frac{1}{|\nabla F|}\frac{\partial F}{\partial x^2}\right) & \frac{\partial}{\partial x^2}\left(\frac{1}{|\nabla F|}\frac{\partial F}{\partial x^3}\right) \\ \frac{\partial}{\partial x^3}\left(\frac{1}{|\nabla F|}\frac{\partial F}{\partial x^1}\right) & \frac{\partial}{\partial x^3}\left(\frac{1}{|\nabla F|}\frac{\partial F}{\partial x^2}\right) & \frac{\partial}{\partial x^3}\left(\frac{1}{|\nabla F|}\frac{\partial F}{\partial x^3}\right) \end{pmatrix} \tag{5.14}$$

Using (5.8) (NOT (5.9)), we compute

$$\nabla\left(\frac{\nabla F}{|\nabla F|}\right) = \frac{1}{\left((x^1)^2+(x^2)^2+(x^3)^2\right)^{\frac{3}{2}}} \begin{pmatrix} \left(x^2\right)^2+\left(x^3\right)^2 & -x^1\,x^2 & x^1\,x^3 \\ -x^1\,x^2 & \left(x^1\right)^2+\left(x^3\right)^2 & x^2\,x^3 \\ -x^1\,x^3 & -x^2\,x^3 & -\left(x^1\right)^2-\left(x^2\right)^2 \end{pmatrix} \tag{5.15}$$

As an aside, notice that this matrix is not symmetric.

Now we can compute II. We have

$$II_p = \begin{pmatrix} II_p\left(\vec{v}_1, \vec{v}_1\right) & II_p\left(\vec{v}_1, \vec{v}_2\right) \\ II_p\left(\vec{v}_2, \vec{v}_1\right) & II_p\left(\vec{v}_2, \vec{v}_2\right) \end{pmatrix}$$
$$= \begin{pmatrix} -\frac{1+\left(p^3\right)^2}{\left(1+2(p^3)^2\right)^{5/2}} & 0 \\ 0 & \frac{1+(p^3)^2}{\left(1+2(p^3)^2\right)^{1/2}} \end{pmatrix} \tag{5.16}$$

5.5 Exercises

1) (First fundamental form) Consider the surface given by $F=0$ and $x^3>0$ where

$$F(x^1,x^2,x^3) = -\left(x^1\right)^2-\left(x^2\right)^2+\left(x^3\right)^2-1. \tag{5.17}$$

Determine a basis $\vec{v}_1$, $\vec{v}_2$ for $T_p\Sigma$, where $p = (p^1, p^2, p^3)^T$. In your basis, determine the matrix I. (Hint: Let $\vec{X} = (1,0,0)$ and $\vec{Y} = (0,1,0)$, and apply a normalization process.)

2) (Second fundamental form) Consider the surface from problem (1). In your chosen basis, and at the point $p = (0,0,1)^T$, compute II as a 2×2 matrix.

3) Let Σ be a hypersurface given as the graph of some function $x^3 = f(x^1, x^2)$. Compute $\vec{N}$, and determine a basis for $T_p\Sigma$ at $p = (p^1, p^2, p^3)^T$.

*4) In the situation of problem (3), compute $\nabla \left(\frac{\nabla F}{|\nabla F|} \right)$ in components.

5) (Parametriations) A third way to describe surfaces is via parametrizations. A *parametrized surface* in $\mathbb{R}^n$ is a surface $\Sigma \subset \mathbb{R}^n$ along with a map $P : \Omega \to \Sigma$ where Ω is a domain in $\mathbb{R}^2$ that is differntiable, one-to-one and onto.

   a) Graph the surface

$$\begin{aligned} P : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ P(s,t) &= \begin{pmatrix} s^2 + t^2 \\ s \\ t \end{pmatrix}. \end{aligned} \tag{5.18}$$

   **b*) What kind of parametrized surface is this?

$$\begin{aligned} P : [0, 2\pi) \times [0, 2\pi) &\longrightarrow \Sigma \subset \mathbb{R}^4 \\ P(s,t) &= \begin{pmatrix} \frac{1}{\sqrt{2}} \cos(t) \\ \frac{1}{\sqrt{2}} \sin(t) \\ \frac{1}{\sqrt{2}} \cos(s) \\ \frac{1}{\sqrt{2}} \sin(s) \end{pmatrix}. \end{aligned} \tag{5.19}$$

   Also, verify that Σ is contained within the unit sphere $\mathbb{S}^3 \subset \mathbb{R}^4$.

6) (Coordinate Charts) If Σ is some surface in $\mathbb{R}^n$ and Ω is some domain in Σ, then a *coordinate chart* is a map $S : \Omega \to \mathbb{R}^2$ that is differentiable and one-to-one (but needn't be onto). If $p \in \Sigma$, then the *coordinates* of p are the two coordinates of $S(p)$. Let Σ be the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$, given by $\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2 - 1 = 0$. For a point $p = (p^1, p^2, p^3) \in \mathbb{S}^3$, define

$$S(p) = \begin{pmatrix} \frac{x^1}{\sqrt{1-x^3}} \\ \\ \frac{x^2}{\sqrt{1-x^3}} \end{pmatrix} \tag{5.20}$$

   What are the coordinates of the following points?

$$p = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \quad q = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \quad r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad s = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \tag{5.21}$$

*7) Under what conditions is it true that $II = \frac{\nabla^2 F}{|\nabla F|}$? That $II = \nabla^2 F$? In the case of a graph of a function: $F(x^1, x^2, x^3) = x^3 - f(x^1, x^2)$, what does $II = \frac{\nabla^2 F}{|\nabla F|}$ at a point $(x^1, x^2, f(x^1, x^2))^T$ imply about f at the point (x^1, x^2)?

Exercises marked with a "∗" are for 501 students.

Problems due Thursday Feb 6.

Lecture 6 - Gaussian Curvature

Lecture from Thursday Feb 6, 2012.

6.1 Identity and transpose operators

Symbolically, all of δ_i^j, δ_{ij} and δ^{ij} are the same. However, consider a vector $\vec{v} = (v^i)$, which is a column vector. The vector $(\delta_j^i v^j)$ is simply $\vec{v}$ again. But $(\delta_{ij} v^j)$ is not $\vec{v}$ but rather $\vec{v}^T$. Similarly $(f_{,j}\delta^{ij})$ is a column vector, instead of a row vector.

If $A = (A_j^i)$ is a matrix, we can transpose either of its indices if we wish: we could define

$$B_{ij} = A_j^k \delta_{ik} \tag{6.1}$$

or

$$C^{ij} = A_k^i \delta^{kj}. \tag{6.2}$$

We briefly consider the case of symmetric matrices; these are matrices so that $A = A^T$. Symbolically, it is true that $A_j^i = A_i^j$. But this is weird: we are equating symbols with mixed-up indices. It is notationally correct to write

$$A_j^i \ = \ \delta^{ik}\delta_{lj}A_k^l \tag{6.3}$$

which means the same thing.

6.2 Coordinate systems

Let $\{x^i\}_{i=1}^n$ be a coordiate system on $\mathbb{E}^n$. A basic fact is that each of the x^i is a *function*. This means that, if $\{y^a\}_{a=1}^n$ is a second coordinate system, then we can take the derivative of any of the x^i with respect to any of the y^a.

Example (Polar and rectangular coordinates on $\mathbb{E}^2$.)

Let x^1, x^2 be orthonormal coordinates on 2-dimensional Euclidean space. Let

$$y^1 \;=\; \sqrt{\left(x^1\right)^2 \,+\, \left(x^2\right)^2} \qquad y^2 \;=\; \tan^{-1}\left(x^1/x^2\right). \tag{6.4}$$

However we can also define the original coordinates in terms of the new coordinates:

$$x^1 \;=\; y^1\,\cos\left(y^2\right) \qquad x^2 \;=\; y^1\,\sin\left(y^2\right). \tag{6.5}$$

We have the transition matrices

$$\begin{aligned} \left(\frac{\partial y^a}{\partial x^i}\right) &= \begin{pmatrix} \frac{\partial y^1}{\partial x^1} & \frac{\partial y^1}{\partial x^2} \\ \frac{\partial y^2}{\partial x^1} & \frac{\partial y^2}{\partial x^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{x^1}{\sqrt{(x^1)^2+(x^2)^2}} & \frac{x^2}{\sqrt{(x^2)^2+(x^2)^2}} \\ \frac{-x^2}{(x^1)^2+(x^2)^2} & \frac{x^1}{(x^1)^2+(x^2)^2} \end{pmatrix} \end{aligned} \tag{6.6}$$

and

$$\begin{aligned} \left(\frac{\partial x^i}{\partial y^a}\right) &= \begin{pmatrix} \frac{\partial x^1}{\partial y^1} & \frac{\partial x^1}{\partial y^2} \\ \frac{\partial x^2}{\partial y^1} & \frac{\partial x^2}{\partial y^2} \end{pmatrix} \\ &= \begin{pmatrix} \cos(y^2) & -y^1\sin(y^2) \\ \sin\left(y^2\right) & y^1\cos(y^2) \end{pmatrix} \end{aligned} \tag{6.7}$$

Now consider the product:

$$\begin{aligned} \left(\frac{\partial y^a}{\partial x^i}\right)\cdot\left(\frac{\partial x^i}{\partial y^a}\right) &= \begin{pmatrix} \frac{\partial y^1}{\partial x^1} & \frac{\partial y^1}{\partial x^2} \\ \frac{\partial y^2}{\partial x^1} & \frac{\partial y^2}{\partial x^2} \end{pmatrix}\begin{pmatrix} \frac{\partial x^1}{\partial y^1} & \frac{\partial x^1}{\partial y^2} \\ \frac{\partial x^2}{\partial y^1} & \frac{\partial x^2}{\partial y^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{x^1\cos(y^2)+x^2\sin(y^2)}{\sqrt{(x^1)^2+(x^2)^2}} & \frac{-x^1y^1\sin(y^2)+x^2y^1\cos(y^2)}{\sqrt{(x^1)^2+(x^2)^2}} \\ \frac{-x^2\cos(y^2)+x^1\sin(y^2)}{(x^1)^2+(x^2)^2} & \frac{x^1y^1\cos(y^2)+x^2y^1\sin(y^2)}{(x^1)^2+(x^2)^2} \end{pmatrix} \end{aligned} \tag{6.8}$$

Now since $\cos(y^2) = x^1/\sqrt{(x^1)^2+(x^2)^2}$ and $\sin(y^2) = x^2/\sqrt{(x^1)^2+(x^2)^2}$, we compute

$$\left(\frac{\partial y^a}{\partial x^i}\right)\cdot\left(\frac{\partial x^i}{\partial y^a}\right) \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{6.9}$$

Is this an accident? No.

Theorem 6.2.1 *Let $\{x^i\}_{i=1}^n$ and $\{y^a\}_{a=1}^n$ be two coordinate systems. Then*

$$\begin{aligned} \delta^i_j &= \frac{\partial x^i}{\partial y^a}\frac{\partial y^a}{\partial x^j} \qquad \textit{and} \\ \delta^a_b &= \frac{\partial y^a}{\partial x^i}\frac{\partial x^i}{\partial y^b}. \end{aligned} \tag{6.10}$$

Proof. By the chain rule we have

$$\frac{\partial}{\partial x^j} = \frac{\partial y^a}{\partial x^j}\frac{\partial}{\partial y^a} \tag{6.11}$$

Noting that $\frac{\partial x^i}{\partial x^j} = \delta^i_j$, we have

$$\delta^i_j = \frac{\partial x^i}{\partial x^j} = \frac{\partial y^a}{\partial x^j}\frac{\partial x^i}{\partial y^a} \tag{6.12}$$

which proves the first assertion. The second assertion can be shown likewise. □

6.3 Gaussian Curvature

Recall Gauss' definition of curvature

$$K_G(p) \text{ “ = ” } \lim_{\mathcal{R}\to\{p\}} \frac{Area\,(\mathcal{G}(\mathcal{R}))}{Area\,(\mathcal{R})} \tag{6.13}$$

where the limit is, in principal, taken as the open regions $\mathcal{R}$ get smaller and smaller, converging on the point p. Of course, this is not rigorous. We formally define

$$K_G(p) = \frac{\det II_p}{\det I_p}. \tag{6.14}$$

The determinants are taken after a basis $\vec{v}_1$, $\vec{v}_2$ is chosen at $p \in T_p\Sigma$. To be specific, the matrices are

$$\begin{aligned} I &= \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} = \begin{pmatrix} \langle \vec{v}_1, \vec{v}_1\rangle & \langle \vec{v}_1, \vec{v}_2\rangle \\ \langle \vec{v}_2, \vec{v}_1\rangle & \langle \vec{v}_2, \vec{v}_2\rangle \end{pmatrix} \\ II &= \begin{pmatrix} II_{11} & II_{12} \\ II_{21} & II_{22} \end{pmatrix} = \begin{pmatrix} \nabla\left(\frac{\nabla F}{|\nabla F|}\right)(\vec{v}_1, \vec{v}_1) & \nabla\left(\frac{\nabla F}{|\nabla F|}\right)(\vec{v}_1, \vec{v}_2) \\ \nabla\left(\frac{\nabla F}{|\nabla F|}\right)(\vec{v}_2, \vec{v}_1) & \nabla\left(\frac{\nabla F}{|\nabla F|}\right)(\vec{v}_2, \vec{v}_2) \end{pmatrix} \end{aligned} \tag{6.15}$$

6.4 Exercises

We didn't get through a lot this time, so we have only a few workbook-style exercises.

1) (Coordinates) Let x^1, x^2 be the standard coordinates on $\mathbb{R}^2$. Define new coordinates

$$\begin{aligned} y^1 &= \left(x^1\right)^2 - \left(x^2\right)^2 \\ y^2 &= x^1 - x^2. \end{aligned} \tag{6.16}$$

a) Determine x^1 and x^2 as functions of y^1 and y^2.

b) Make a sketch of the coordinate system $\{y^1, y^2\}$. Label the points $(0,0)^T)$, $(1,1)^T$ and $(-1,-2)^T$ (given in y-coordinates). Sketch the graph of the function

$$y^2 = 3y^1 + 1.$$

c) Compute the matrices

$$\left(\frac{\partial x^i}{\partial y^a}\right) \quad \textit{and} \quad \left(\frac{\partial y^a}{\partial x^i}\right) \tag{6.17}$$

and prove that they are inverses of each other.

2) Assume that the coordinate systems $\{x^i\}$ and $\{y^i\}$ are mutually orthogonal; this means that the matrices

$$\left(\frac{\partial x^i}{\partial y^a}\right)_{i,a=1}^{n} \quad \textit{and} \quad \left(\frac{\partial y^a}{\partial x^i}\right)_{i,a=1}^{n} \tag{6.18}$$

are not just inverses (which we proved is always the case), but that they are transposes. Prove that

$$\delta^{ab}\frac{\partial x^i}{\partial y^a}\frac{\partial x^j}{\partial y^b} = \delta^{ij}. \tag{6.19}$$

To get full credit, you must use the correct symbolic calculus as outlined in Section 1.

3) Consider $F : \mathbb{R}^3 \to \mathbb{R}$ given by $F(x^1,x^2,x^3) = \left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2 - 1$, and let Σ be the surface $\{F = 0\}$. Let $\vec{X} = (1,0,0)^T$ and $\vec{Y} = (0,1,0)^T$ be vector fields on $\mathbb{R}^3$.

a) Given any point $(p^1,p^2,p^3)^T \in \Sigma$, find $\vec{N}$.

b) At each point of Σ, project $\vec{X}$ and $\vec{Y}$ onto the tangent space of Σ (there is a basic formula for this). Call the new vectors $\vec{v}_1$ and $\vec{v}_2$.

c) Show that $\vec{v}_1$ and $\vec{v}_2$ span $T_p\Sigma$, except for when p is on the "equator" (the intersection of Σ and the x^1-x^2 plane).

d) Let $p \in \Sigma$ be any point that's not on the equator. In the $\vec{v}_1$, $\vec{v}_2$ basis, show that

$$I_p = \begin{pmatrix} 1-(x^1)^2 & -x^1x^2 \\ -x^1x^2 & 1-(x^2)^2 \end{pmatrix}, \tag{6.20}$$

and show that, as expected, I_p is singular when p is on the equator.

4) Let F and Σ be as in (3).

a) Compute $\nabla\left(|\nabla F|^{-1}\nabla F\right)$ as a 3×3 matrix. Show that it is singular.

b) In the basis from (3b), compute II as a matrix.

c) Prove that $K_G(p)=1$.

*5) On tangent spaces in $\mathbb{R}^{n,k}$, we use the inner product

$$\langle \vec{v},\,\vec{w}\rangle \;=\; \vec{v}^T I_{n,k}\vec{w} \tag{6.21}$$

where

$$I_{n,k} \;=\; \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ 0 & & & & & -1 \end{pmatrix} \tag{6.22}$$

where there are 0's in every entry except along the main diagonal, where there are n many 1's followed by k many -1's. Likewise, recall the $\mathbb{R}^{n,k}$ Pythagorean theorem: in standard coordinates, if $p=(p^1,\ldots,p^{n+k})^T$ is a point, then the distance from the origin to p is

$$\mathrm{dist}(\mathcal{O},p) \;=\; \sqrt{(p^1)^2+\cdots+(p^n)^2\;-\;(p^{n+1})^2-\cdots-(p^{n+k})^2}. \tag{6.23}$$

The set of all points $p\in\mathbb{R}^{1,2}$ that have distance 1 to the origin is a surface of two components; the component with $x^1>0$ is known as the *unit pseudosphere*. Make a sketch of this pseudosphere. Prove that its first fundamental form is in fact positive definite.

Problems due Thursday 2/13

Lecture 7 - Gaussian and Eulerian Curvature

Lecture from Tuesday Feb 11, 2012

7.1 Bases and Change of Basis

7.1.1 Bases

Given an n-dimensional vector space $\mathcal{V}^n$, a basis is normally indexed with lower indices:

$$\vec{e}_1, \ldots, \vec{e}_n. \tag{7.1}$$

For instance, we could have the standard basis when $\mathcal{V}^n = \mathbb{R}^n$

$$\vec{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \tag{7.2}$$

where the 1 is in the i^{th} position. Then a vector $\vec{v}$ can be expressed as a linear combination of the basis vectors:

$$\vec{v} = v^i e_i = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \tag{7.3}$$

7.1.2 Change of basis

Assume $\{\vec{e}_i\}_{i=1}^n$ and $\{\vec{f}_i\}_{i=1}^n$ are two bases for $\mathcal{V}^n$. We have change of basis symbols

$$\begin{aligned} \vec{e}_j &= A^i_j \vec{f}_i \\ \vec{f}_j &= B^i_j \vec{v}_i. \end{aligned} \tag{7.4}$$

We have that matrices (A^i_j) and (B^i_j) are inverses of one another: $A^i_j B^j_k = \delta^i_k$. This is simple to prove:

$$\vec{e}_i = A^j_i \vec{f}_j = A^j_i B^k_j \vec{e}_k. \tag{7.5}$$

But, because $\{\vec{e}_j\}$ is a basis, the *only* linear combination of the $\{\vec{e}_j\}$ that gives $\vec{e}_i$ is given by $\vec{e}_i = \delta^k_i \vec{e}_k$. Therefore $A^j_i B^k_j = \delta^k_i$.

7.2 Gaussian curvature and change of basis

Recall the first and second fundamental forms:

$$\begin{aligned} I &= \langle \cdot, \cdot \rangle \big|_{T_p\Sigma} \\ II &= \nabla \left(\frac{\nabla F}{|\nabla F|} \right) \Big|_{T_p\Sigma} \end{aligned} \tag{7.6}$$

After a basis $\vec{v}_1, \vec{v}_2$ is chosen for $T_p\Sigma$, we can express these as matrices:

$$\begin{aligned} I &= \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} = \begin{pmatrix} \langle \vec{v}_1, \vec{v}_1 \rangle & \langle \vec{v}_1, \vec{v}_2 \rangle \\ \langle \vec{v}_2, \vec{v}_1 \rangle & \langle \vec{v}_2, \vec{v}_2 \rangle \end{pmatrix} \\ II &= \begin{pmatrix} II_{11} & II_{12} \\ II_{21} & II_{22} \end{pmatrix} = \begin{pmatrix} \nabla \left(\frac{\nabla F}{|\nabla F|} \right) (\vec{v}_1, \vec{v}_1) & \nabla \left(\frac{\nabla F}{|\nabla F|} \right) (\vec{v}_1, \vec{v}_2) \\ \nabla \left(\frac{\nabla F}{|\nabla F|} \right) (\vec{v}_2, \vec{v}_1) & \nabla \left(\frac{\nabla F}{|\nabla F|} \right) (\vec{v}_2, \vec{v}_2) \end{pmatrix}. \end{aligned} \tag{7.7}$$

After doing all this, the Gaussian curvature K_G and the Gaussiam mean curvature M_G are defined as follows:

$$\begin{aligned} K_G &= \det \left(I^{-1} II \right) = \frac{\det(II)}{\det(I)} \\ M_G &= \frac{1}{2} Tr \left(I^{-1} II \right). \end{aligned} \tag{7.8}$$

A glaring issue is in the choice of basis: if we choose a different basis for $T_p\Sigma$, the matrix representations of II and I are obviously not the same. So aren't the values of K_G and M_G affected? The answer is no.

Theorem 7.2.1 *The values of K_G and M_G are independent of the basis that is chosen.*

Proof. Let $\{\vec{v}_1, \vec{v}_2\}$ and $\{\vec{w}_1, \vec{w}_2\}$ be two different bases for $T_p\Sigma$, with transitions

$$\begin{aligned} \vec{v}_i &= A_i^j \vec{w}_j \\ \vec{w}_i &= B_i^j \vec{v}_j \end{aligned} \tag{7.9}$$

Let

$$I = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} \quad II = \begin{pmatrix} II_{11} & II_{12} \\ II_{21} & II_{22} \end{pmatrix} \tag{7.10}$$

$$I' = \begin{pmatrix} I'_{11} & I'_{12} \\ I'_{21} & I'_{22} \end{pmatrix} \quad II' = \begin{pmatrix} II'_{11} & II'_{12} \\ II'_{21} & II'_{22} \end{pmatrix} \tag{7.11}$$

be the first and second fundamental forms expressed in the two bases. Explicitly, let

$$\begin{aligned} I_{ij} = \langle \vec{v}_i, \vec{v}_j \rangle \quad II_{ij} = \nabla \left(\frac{\nabla F}{|\nabla F|} \right) (\vec{v}_i, \vec{v}_j) \\ I'_{ij} = \langle \vec{w}_i, \vec{w}_j \rangle \quad II'_{ij} = \nabla \left(\frac{\nabla F}{|\nabla F|} \right) (\vec{w}_i, \vec{w}_j). \end{aligned} \tag{7.12}$$

Now both of these forms are bilinear, so we compute

$$\begin{aligned} I_{ij} &= \langle \vec{v}_i, \, \vec{v}_j \rangle = \left\langle A_i^k \vec{w}_k, \, A_j^l \vec{w}_l \right\rangle \\ &= A_i^k A_l^j \, \langle \vec{w}_k, \, \vec{w}_l \rangle = A_i^k A_l^j I'_{kl} \end{aligned} \tag{7.13}$$

and

$$\begin{aligned} II_{ij} &= \nabla \left(\frac{\nabla F}{|\nabla F|} \right) (\vec{v}_i, \, \vec{v}_j) = \nabla \left(\frac{\nabla F}{|\nabla F|} \right) \left(A_i^k \vec{w}_k, \, A_j^l \vec{w}_l \right) \\ &= A_i^k A_j^l \, \nabla \left(\frac{\nabla F}{|\nabla F|} \right) (\vec{w}_k, \, \vec{w}_l) = A_i^k A_j^l II'_{kl} \end{aligned} \tag{7.14}$$

With this, we compute

$$\begin{aligned} \frac{\det(II_{ij})}{\det(I_{ij})} &= \frac{\det(A_i^k A_j^l II'_{kl})}{\det(A_i^k A_j^l I'_{kl})} = \frac{(\det(A))^2 \det(II'_{kl})}{(\det(A))^2 \det(I'_{kl})} = \frac{det(II'_{kl})}{\det(I'_{kl})} \\ Tr(I^{-1} II) &= Tr((AI'A)^{-1}(AII'A)) \\ &= Tr(A^{-1} I'^{-1} II' A) = Tr(I'^{-1} II' A A^{-1}) = Tr(I'^{-1} II'). \end{aligned} \tag{7.15}$$

In the last line, we used the cyclic property of traces: $Tr(A_1 A_2 \ldots A_k) = Tr(A_2 \ldots A_k A_1)$. □

7.3 Gaussian and Eulerian Curvature

In this section we prove the first significant theorem from Gauss' 1827 paper.

Theorem 7.3.1 *Let Σ be a surface in 3-space $\mathbb{E}^3$, given as the zero-set of some function: $\Sigma = \{F = 0\}$. If $K(p)$, $M(p)$ are the Eulerian curvature and Eulerian mean curvature at $p \in \Sigma$, and if $K_G(p)$, $M_G(p)$ are the Gaussian curvature and Gaussian mean curvature at $p \in \Sigma$, then*

$$K_G(p) \;=\; K(p) \quad \textit{and} \quad M_G(p) \;=\; M(p). \tag{7.16}$$

Proof. We have already known that K and M are invariant under choosing a different coordinate system. We now know that this is true for K_G and M_G as well: changing the coordinate systems amounts to a change of basis, and we just proved that K_G and M_G are independent of change of basis.

Letting p be an arbitrary point of Σ, this fact allows us to pick a coordinate system $\{x^1, x^2, x^3\}$ so that p is the origin, and $T_p\Sigma$ is just the x^1-x^2 plane. Now in this coordinate system we can express the surface Σ as a graph: $x^3 = f(x^1, x^2)$ for some function f. Then $F(x^1, x^2, x^3) = f(x^1, x^2) - x^3$. Because of the way we chose the coordinates, we have that

$$\frac{\partial f}{\partial x^1}(p) \;=\; \frac{\partial f}{\partial x^2}(p) \;=\; 0. \tag{7.17}$$

Now letting $\vec{v}_1, \vec{v}_2 \in T_p\Sigma$ be the vectors

$$\vec{v}_1 \;=\; \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_2 \;=\; \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{7.18}$$

we compute

$$\begin{aligned} II(\vec{v}_i, \vec{v}_j) \;&=\; \nabla\left(\frac{\nabla F}{|\nabla F|}\right)(\vec{v}_i, \vec{v}_j) \\ &=\; \frac{\partial}{\partial x^j}\left(\frac{\frac{\partial F}{\partial x^i}}{|\nabla F|}\right) \\ &=\; \frac{\frac{\partial^2 f}{\partial x^i \partial x^j}}{\sqrt{1 + |\nabla f|^2}} \;-\; \frac{\frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^j}\sqrt{1 + |\nabla f|^2}}{1 + |\nabla f|^2}. \end{aligned} \tag{7.19}$$

Then using the fact that $f_{,i} = 0$ and $f_{,j} = 0$ at p we have

$$II_p(\vec{v}_i, \vec{v}_j) \;=\; \frac{\partial^2 f}{\partial x^i \partial x^j}. \tag{7.20}$$

Therefore, in our special coordinate system, we have

$$\begin{aligned} I_p \;&=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ II_p \;&=\; \begin{pmatrix} f_{,11} & f_{,12} \\ f_{,21} & f_{,22} \end{pmatrix} \end{aligned} \tag{7.21}$$

We now compute

$$\begin{aligned} K_G(p) \;&=\; \frac{\det\, II_p}{\det\, I_p} \;=\; f_{,11}f_{,22} \,-\, (f_{,12})^2 \;=\; K(p) \\ M_G(p) \;&=\; \frac{1}{2}Tr(I_p^{-1}II_p) \;=\; \frac{1}{2}\,(f_{,11} \,-\, f_{,22}) \;=\; M(p). \end{aligned} \tag{7.22}$$

Since p was an arbitrary point on Σ, we have that $K_G = K$ and $M_G = M$ on Σ. □

7.4 Exercises

1) Assume that Σ is the graph of a function: $x^3 = f(x^1, x^2)$, where $f : \mathbb{R}^2 \to \mathbb{R}$ is continuously twice differentiable on $\mathbb{R}^2$.

   a) Prove that it is always possible to choose two complete, continuously differentiable vector fields $\vec{v}_1$, $\vec{v}_2$ on Σ, so that at every point $p \in \Sigma$ they span $T_p\Sigma$.

   b) Show that it is possible to choose $\vec{v}_1$, $\vec{v}_2$ to be orthonormal at each point $p \in \Sigma$.

   c) Given $p \in \Sigma$, what is I_p in the $\vec{v}_1$, $\vec{v}_2$ basis?

   **d*) If Σ is not the graph of a surface, show that it is not necessarily the case that Σ has two complete, continuous vector fields that span $T_p\Sigma$ at every point.

2) (Generalization of Euler's formula) Let Σ be the graph of a function $x^3 = f(x^1, x^2)$. If $p = (p^1, p^2, p^2)^T$ is a point on Σ where $(p^1, p^2)^T$ is a critical point of f, we know that $K(p) = f_{,11}f_{,22} - (f_{,12})^2$ and $M(p) = \frac{1}{2}\,(f_{,11} + f_{,22})$.

   a) Letting F be the defining function for Σ, compute $\nabla\left(\frac{\nabla F}{|\nabla F|}\right)$ in terms of the $f_{,i}$, $f_{,ij}$, etc. Do not use the symbol ∇f in your final formula.

   **b*) If f is a quadratic function, what is the falloff rate of M and K at infinity? If f is any polynomial in two variables, what can you say about the curvature falloff at infinity? (Hint: You'll probably have to use the frame $\vec{v}_1$, $\vec{v}_2$ from Problem (1) to demonstrate your answer rigorously. But it would be a bad idea to attempt to evaluate M and K explicitly.)

Lecture 8 - Examples. Minimal Surfaces

Lecture given Tuesday Feb 18, 2012.

8.1 Definition

A surface is called a *minimal surface* if its mean curvature is zero.

8.2 Example

We look at the example $f(x, y) = xy$, where $F(x, y, z) = f(x, y) - z$. We compute

$$\nabla F = (y,\, x,\, -1)$$

$$\vec{N} \;=\; \left(\frac{y}{\sqrt{1+x^2+y^2}},\; \frac{x}{\sqrt{1+x^2+y^2}},\; \frac{-1}{\sqrt{1+x^2+y^2}} \right) \tag{8.1}$$

$$\nabla \left(\frac{\nabla F}{|\nabla F|} \right) = \left(1+x^2+y^2\right)^{-\frac{3}{2}} \begin{pmatrix} -xy & 1+x^2 & 0 \\ 1+y^2 & -xy & 0 \\ -x & -y & 0 \end{pmatrix}$$

Now we pick a pair of vector fields. Let $\vec{X} = (1, 0, 0)$ and $\vec{Y} = (0, 1, 0)$ be vector fields on $\mathbb{R}^3$, and let

$$\vec{V}_1 = \vec{X} - \left\langle \vec{X}, \vec{N} \right\rangle \vec{N} \quad \vec{V}_2 = \vec{Y} - \left\langle \vec{Y}, \vec{N} \right\rangle \vec{N} \tag{8.2}$$

From this we compute the first fundamental form and its inverse:

$$\begin{aligned}
I &= \begin{pmatrix} 1 - \left\langle \vec{X}, \vec{N} \right\rangle^2 & -\left\langle \vec{X}, \vec{N} \right\rangle \left\langle \vec{Y}, \vec{N} \right\rangle \\ -\left\langle \vec{X}, \vec{N} \right\rangle \left\langle \vec{Y}, \vec{N} \right\rangle & 1 - \left\langle \vec{Y}, \vec{N} \right\rangle^2 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1+x^2}{1+x^2+y^2} & \frac{-xy}{1+x^2+y^2} \\ \frac{-xy}{1+x^2+y^2} & \frac{1+y^2}{1+x^2+y^2} \end{pmatrix} \\
I^{-1} &= \begin{pmatrix} 1+y^2 & xy \\ xy & 1+x^2 \end{pmatrix}
\end{aligned} \tag{8.3}$$

For the second fundamental form, we obtain

$$II = \left(1+x^2+y^2\right)^{-\frac{5}{2}} \begin{pmatrix} -xy(1+x^2) & 1+x^2+y^2+2x^2y^2 \\ 1+x^2+y^2+2x^2y^2 & -xy(1+y^2) \end{pmatrix} \tag{8.4}$$

Then we have

$$I^{-1}II = (1+x^2+y^2)^{-\frac{5}{2}} \begin{pmatrix} x^3y^3 & (1+x^2)(1+y^2)^2 \\ (1+x^2)^2(1+y^2) & x^3y^3 \end{pmatrix} \tag{8.5}$$

Therefore

$$\begin{aligned}
K &= \frac{-1}{1+x^2+y^2} \\
M &= 0
\end{aligned} \tag{8.6}$$

We see that this is a minimal surface. Note that this surface becomes flatter and flatter

More to come!

Lecture 9 - The Theorema Egregium

Lecture given on Thursday Feb 20, 2012

Section 9.4 of today's notes is based on portions of Spivak's Volume II, Chapter 3

9.1 Directional Derivatives

9.1.1 Basic computations

If $\vec{v}$ is a vector field, we denote the derivative in the direction of $\vec{v}$ by

$$\frac{\partial}{\partial \vec{v}}. \tag{9.1}$$

For instance if f is a function then

$$\frac{\partial f}{\partial \vec{v}} = \langle \nabla f, \vec{v} \rangle . \tag{9.2}$$

But we can also take the directional derivative of a vector field. If $\vec{A}$ is any (continuously differentiable) vector field on $\mathbb{E}^n$, we denote

$$\frac{\partial \vec{A}}{\partial \vec{v}} \tag{9.3}$$

to be the derivative of $\vec{A}$ in the direction $\vec{v}$. Letting $\vec{e}_1, \dots, \vec{e}_n$ be the standards orthonormal fields, then we can write $\vec{A} = A^i \vec{e}_i$ and $\vec{v} = v^i \vec{e}_i$. We have

$$\begin{aligned} \frac{\partial \vec{A}}{\partial \vec{v}} &= \frac{\partial (A^i \vec{e}_i)}{\partial \vec{v}} = \frac{\partial A^i}{\partial \vec{v}} \vec{e}_i \\ &= v^j \frac{\partial A^i}{\partial \vec{e}_j} \vec{e}_i \\ &= \left\langle \vec{v}, \nabla A^i \right\rangle \vec{e}_i. \end{aligned} \tag{9.4}$$

9.1.2 The Product Rule

The product rule holds for directional derivatives:

$$\frac{\partial}{\partial \vec{v}} \langle \vec{x}, \vec{y} \rangle \;=\; \left\langle \frac{\partial \vec{x}}{\partial v}, \vec{y} \right\rangle + \left\langle \vec{x}, \frac{\partial \vec{y}}{\partial \vec{v}} \right\rangle. \tag{9.5}$$

9.2 The Second Fundamental Form

The main purpose of this sectin is to prove that II is symmetric: $II(\vec{v}, \vec{w}) = II(\vec{w}, \vec{v})$ for all $\vec{v}, \vec{w} \in T_p\Sigma$. Let's take a closer look.

Lemma 9.2.1 *If $\vec{v}$, $\vec{w}$ are vectors in $T_p\Sigma$, then*

$$II(\vec{v}, \vec{w}) \;=\; \left\langle \frac{\partial \vec{N}}{\partial \vec{v}}, \vec{w} \right\rangle. \tag{9.6}$$

Proof. We compute

$$\begin{aligned} II(\vec{v}, \vec{w}) \;&=\; \vec{\nabla}\vec{N}(\vec{v}, \vec{w}) \;=\; \left(\frac{F_{,j}}{|\nabla F|} \right)_{,i} v^i w^j \\ &=\; v^i \frac{\partial}{\partial x^i} \left(\frac{F_{,j}}{|\nabla F|} \right) w^j \;=\; v^i \frac{\partial}{\partial x^i} \left(\frac{F_{,k}}{|\nabla F|} \right) w^l \left\langle \vec{e}^k, \vec{e}_l \right\rangle \\ &=\; \left\langle v^i \frac{\partial}{\partial x^i} \left(\frac{F_{,k}}{|\nabla F|} \right) \vec{e}^k, w^l \vec{e}_l \right\rangle \;=\; \left\langle \frac{\partial}{\partial \vec{v}} \left(\frac{F_{,k}}{|\nabla F|} \right) \vec{e}^k, \vec{w} \right\rangle \\ &=\; \left\langle \frac{\partial}{\partial \vec{v}} \left(\frac{\nabla F}{|\nabla F|} \right), \vec{w} \right\rangle \end{aligned} \tag{9.7}$$

□

Theorem 9.2.2 *If $\vec{v}, \vec{w} \in T_p\Sigma$, then $II(\vec{v}, \vec{w}) \;=\; II(\vec{w}, \vec{v})$.*

Proof. We begin with one of the stages of the previous computation. Using the product rule and then a bit of simplification, we have

$$\begin{aligned} II(\vec{v}, \vec{w}) \;&=\; \left\langle v^i \frac{\partial}{\partial x^i} \left(\frac{F_{,k}}{|\nabla F|} \right) e^k, \vec{w} \right\rangle \\ &=\; \left\langle v^i \frac{F_{,ki}}{|\nabla F|} \vec{e}^k, \vec{w} \right\rangle - \left\langle v^i \frac{F_{,k} |\nabla F|_{,i}}{|\nabla F|^2} \vec{e}^k, \vec{w} \right\rangle \\ &=\; \left\langle v^i \frac{F_{,ki}}{|\nabla F|} \vec{e}^k, \vec{w} \right\rangle - \frac{v^i |\nabla F|_{,i}}{|\nabla F|} \left\langle \frac{F_{,k}}{|\nabla F|} \vec{e}^k, \vec{w} \right\rangle \\ &=\; \left\langle v^i \frac{F_{,ki}}{|\nabla F|} \vec{e}^k, w^l \vec{e}_l \right\rangle - \frac{v^i |\nabla F|_{,i}}{|\nabla F|} \left\langle \vec{N}, \vec{w} \right\rangle \end{aligned} \tag{9.8}$$

But because $\vec{N}$ and $\vec{w}$ are perpendicular, we arrive at

$$II(\vec{v},\, \vec{w}) \;=\; v^i w^j \frac{F_{,ji}}{|\nabla F|} \tag{9.9}$$

Switching the roles of $\vec{v}$ and $\vec{w}$ we get

$$II(\vec{w},\, \vec{v}) \;=\; w^i v^j \frac{F_{,ji}}{|\nabla F|} \;=\; v^i w^j \frac{F_{,ij}}{|\nabla F|} \tag{9.10}$$

Therefore

$$II(\vec{w},\, \vec{v}) \;-\; II(\vec{w},\, \vec{v}) \;=\; v^i w^j \left(\frac{F_{,ji}}{|\nabla F|} \;-\; \frac{F_{,ij}}{|\nabla F|} \right) \tag{9.11}$$

which is zero, by the commuatitvity of second partial derivatives. □

9.3 Vector Fields and Directional Derivatives on Σ

9.3.1 Extending vector fields

Often we are given a vector located at a point $p \in \Sigma$, but it will sometimes be necessary (or useful) to use such a vector as though it were an entire vector field.

Let $\vec{V} \in T_p\Sigma$. The problem is to extend $\vec{V}$ to a vector field on Σ, in such a way that it remains tangent to Σ. To do this, let $\vec{A}$ be the constant vector field on $\mathbb{R}^n$ that agrees with $\vec{V}$ at p. Obviously this is not going to be tangent to Σ. So we define

$$\vec{v} \;=\; \vec{A} - \left\langle \vec{A},\, \vec{N} \right\rangle \vec{N}. \tag{9.12}$$

9.3.2 Comparing Directional Derivatives

Let $\vec{V}, \vec{W} \in T_p\Sigma$ be vectors based at $p \in \Sigma$. Letting $\vec{A}$, $\vec{B}$ be the extension of these fields to $\mathbb{E}^3$, we define

$$\begin{aligned} \vec{v} \;&=\; \vec{A} - \left\langle \vec{A},\, \vec{N} \right\rangle \vec{N} \\ \vec{w} \;&=\; \vec{B} - \left\langle \vec{B},\, \vec{N} \right\rangle \vec{N} \end{aligned} \tag{9.13}$$

We wish to compare the directional derivatives $\frac{\partial \vec{v}}{\partial \vec{w}}$ and $\frac{\partial \vec{w}}{\partial \vec{v}}$. We compute, at the point $\vec{p}$, that

$$
\begin{aligned}
\left.\frac{\partial \vec{v}}{\partial w}\right|_p &= -\frac{\partial}{\partial w}\left(\left\langle \vec{A},\,\vec{N}\right\rangle \vec{N}\right) && \text{because } \vec{A} \text{ is constant} \\
&= -\frac{\partial}{\partial w}\left(\left\langle \vec{A},\,\vec{N}\right\rangle\right) \vec{N} - \left\langle \vec{A},\,\vec{N}\right\rangle \frac{\partial \vec{N}}{\partial w} && \text{product rule} \\
&= -\frac{\partial}{\partial w}\left(\left\langle \vec{A},\,\vec{N}\right\rangle\right) \vec{N} && \text{because at } \vec{p},\ \left\langle \vec{A}, \vec{N}\right\rangle = 0 \\
&= -\left\langle \vec{A},\,\frac{\partial \vec{N}}{\partial w}\right\rangle \vec{N} && \text{by the product rule, and because } \vec{A} \text{ is constant} \\
&= -\left\langle \vec{v},\,\frac{\partial \vec{N}}{\partial w}\right\rangle \vec{N} && \text{at } \vec{p}, \text{ we have } \vec{A} = \vec{v}
\end{aligned}
\tag{9.14}
$$

Now we are in a good position, for by Lemma 9.2.1 the inner product is the second fundamental form. We have

$$\frac{\partial \vec{v}}{\partial w} \;=\; -II\left(\vec{v},\,\vec{w}\right)\,\vec{N}. \tag{9.15}$$

Therefore

$$\frac{\partial \vec{v}}{\partial w} - \frac{\partial \vec{w}}{\partial v} \;=\; -\left(II\left(\vec{v},\,\vec{w}\right) \,-\, II\left(\vec{v},\,\vec{w}\right)\right)\,\vec{N}. \tag{9.16}$$

which is zero, by Theorem 9.2.2. We summarize this in the following Lemma.

Lemma 9.3.1 *Let $\vec{V}, \vec{W} \in T_p\Sigma$ be vectors at $p \in \Sigma$. Let $\vec{A}, \vec{W}$ be the constant vector fields on $\mathbb{E}^3$ that equal $\vec{V}, \vec{W}$ at p (respctively). Then define*

$$
\begin{aligned}
\vec{v} \;&=\; \vec{A} - \left\langle \vec{A},\,\vec{N}\right\rangle \vec{N} \\
\vec{w} \;&=\; \vec{B} - \left\langle \vec{B},\,\vec{N}\right\rangle \vec{N}
\end{aligned}
\tag{9.17}
$$

at points of Σ. Then $\vec{v}, \vec{w}$ are vector fields that are tangent to Σ, and at $\vec{p}$, we have

$$\left.\frac{\partial \vec{v}}{\partial \vec{w}}\right|_p \;=\; \left.\frac{\partial \vec{w}}{\partial v}\right|_p. \tag{9.18}$$

One must be careful here: at other points of Σ, it is usually **not** the case that $\frac{\partial \vec{v}}{\partial \vec{w}} = \frac{\partial \vec{w}}{\partial v}$.

9.3.3 A final useful expression of II

In the calculations in the following section, it will be necessary to express the second fundamental form *without* taking a derivative of $\vec{N}$. But II is a first order operator, so some derivatives will be necessary: in $II(\vec{v}, \vec{w})$ we will have to take derivatives of $\vec{v}$ and/or $\vec{w}$, and to do so, we will have to extend $\vec{v}$ and $\vec{w}$ to fields.

Lemma 9.3.2 *Let $p \in \Sigma$ and let $\vec{v}$ and $\vec{w}$ be the vector fields of Lemma 9.3.1. Specifically, $\vec{v}(p), \vec{w}(p) \in T_p\Sigma$ and $\frac{\partial \vec{v}}{\partial \vec{w}}\big|_p = \frac{\partial \vec{w}}{\partial v}\big|_p$. Then*

$$II(\vec{v}, \vec{w}) \;=\; -\left\langle \vec{N}, \frac{\partial \vec{v}}{\partial \vec{w}} \right\rangle. \tag{9.19}$$

Proof. Exercise. □

9.4 The Theorema Egregium

Gauss' great Theorema Egregium states that the second fundamental form can be determined entirely from the first fundamental form, along with some first and second derivatives of its entries. What this means is that if two surfaces have, in some sense, the same first fundamental forms, then the two surfaces have the same Gaussian curvature.

Theorem 9.4.1 (Gauss) *The Gaussian curvature $K = det(I^{-1}II)$ can be expressed entirely in terms of entries of I along with first and second derivatives of its entries.*

Corollary 9.4.2 (The Theorema Egregium) *If one surface can be mapped to another is such a way that intrinsic distances are preserved, then the two surfaces have the same Gaussian curvature.*

Proof of the Theorem.

Going back to Gauss, it is traditional to label the components of I and II as follows:

$$\begin{aligned} I \;&=\; \begin{pmatrix} E & F \\ F & G \end{pmatrix} \\ II \;&=\; \begin{pmatrix} l & m \\ m & n \end{pmatrix} \end{aligned} \tag{9.20}$$

so that

$$K \;=\; \frac{ln \,-\, m^2}{EG \,-\, F^2}. \tag{9.21}$$

Now we use the fact that the cross product $\vec{v} \times \vec{w}$ is proportional to the unit normal $\vec{N}$. Indeed

$$\vec{N} \;=\; \frac{\vec{v} \times \vec{w}}{\sqrt{det(I)}} \;=\; \frac{\vec{v} \times \vec{w}}{\sqrt{EG \,-\, F^2}} \tag{9.22}$$

Combining this with Lemma 9.3.2 we have

$$\begin{aligned} l &= \left\langle \vec{N},\, \frac{\partial \vec{v}}{\partial \vec{v}} \right\rangle = \frac{1}{\sqrt{EG-F^2}} \left\langle \vec{v}\times\vec{w},\, \frac{\partial \vec{v}}{\partial \vec{v}} \right\rangle \\ m &= \left\langle \vec{N},\, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle = \frac{1}{\sqrt{EG-F^2}} \left\langle \vec{v}\times\vec{w},\, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle \\ n &= \left\langle \vec{N},\, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle = \frac{1}{\sqrt{EG-F^2}} \left\langle \vec{v}\times\vec{w},\, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle \end{aligned} \tag{9.23}$$

The inner products on the right are triple products, which can be expressed as determinants:

$$\begin{aligned} l &= \frac{1}{\sqrt{EG-F^2}} Det\left(\frac{\partial \vec{v}}{\partial \vec{v}}, \vec{v}, \vec{w}\right) \\ m &= \frac{1}{\sqrt{EG-F^2}} Det\left(\frac{\partial \vec{w}}{\partial \vec{v}}, \vec{v}, \vec{w}\right) \\ n &= \frac{1}{\sqrt{EG-F^2}} Det\left(\frac{\partial \vec{w}}{\partial \vec{w}}, \vec{v}, \vec{w}\right). \end{aligned} \tag{9.24}$$

Therefore

$$\begin{aligned} Det(I)^2 \cdot K &= Det\left(\frac{\partial \vec{v}}{\partial \vec{v}}, \vec{v}, \vec{w}\right) \cdot Det\left(\frac{\partial \vec{w}}{\partial \vec{w}}, \vec{v}, \vec{w}\right) - Det\left(\frac{\partial \vec{w}}{\partial \vec{v}}, \vec{v}, \vec{w}\right)^2 \\ &= Det\left(\frac{\partial \vec{v}}{\partial \vec{v}}, \vec{v}, \vec{w}\right) \cdot Det\left(\frac{\partial \vec{w}}{\partial \vec{w}}^T, \vec{v}^T, \vec{w}^T\right) - Det\left(\frac{\partial \vec{w}}{\partial \vec{v}}, \vec{v}, \vec{w}\right) Det\left(\frac{\partial \vec{w}}{\partial \vec{v}}^T, \vec{v}^T, \vec{w}^T\right) \\ &= Det\left(\left(\frac{\partial \vec{v}}{\partial \vec{v}}, \vec{v}, \vec{w}\right)\left(\frac{\partial \vec{w}}{\partial \vec{w}}^T, \vec{v}^T, \vec{w}^T\right)\right) - Det\left(\left(\frac{\partial \vec{w}}{\partial \vec{v}}, \vec{v}, \vec{w}\right)\left(\frac{\partial \vec{w}}{\partial \vec{v}}^T, \vec{v}^T, \vec{w}^T\right)\right). \end{aligned} \tag{9.25}$$

We can multiply out these matrices to get

$$Det(I)^2 \cdot K = Det\begin{pmatrix} \langle \frac{\partial \vec{v}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{w}} \rangle & \langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{w}} \rangle & \langle \vec{w}, \frac{\partial \vec{w}}{\partial \vec{w}} \rangle \\ \langle \frac{\partial \vec{v}}{\partial \vec{v}}, \vec{v} \rangle & \langle \vec{v}, \vec{v} \rangle & \langle \vec{w}, \vec{v} \rangle \\ \langle \frac{\partial \vec{v}}{\partial \vec{v}}, \vec{w} \rangle & \langle \vec{v}, \vec{w} \rangle & \langle \vec{w}, \vec{w} \rangle \end{pmatrix} - Det\begin{pmatrix} \langle \frac{\partial \vec{w}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{v}} \rangle & \langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{v}} \rangle & \langle \vec{w}, \frac{\partial \vec{w}}{\partial \vec{v}} \rangle \\ \langle \frac{\partial \vec{w}}{\partial \vec{v}}, \vec{v} \rangle & \langle \vec{v}, \vec{v} \rangle & \langle \vec{w}, \vec{v} \rangle \\ \langle \frac{\partial \vec{w}}{\partial \vec{v}}, \vec{w} \rangle & \langle \vec{v}, \vec{w} \rangle & \langle \vec{w}, \vec{w} \rangle \end{pmatrix} \tag{9.26}$$

We will continue the proof in class next time.

9.5 Exercises

1) Prove Lemma 9.3.2.

2) Let $\vec{A} = A^i \vec{e}_i$ be a vector field in $\mathbb{E}^3$, where the $\vec{e}_i$ are the standard basis vectors, and $A^1 = 1$, $A^2 = x^1$, $A^3 = x^1x^2 - (x^2)^2$. Compute $\frac{\partial \vec{A}}{\partial \vec{v}}$ when

   a) $\vec{v} = e_1$

b) $\vec{v} = e_1 + \frac{1}{2}e_2$

c) $\vec{v} = v^1 e_1 + v^2 e_2$

3) With $\vec{N} = \frac{\nabla F}{|\nabla F|}$, we have defined

$$II(\vec{v},\, \vec{w}) \;=\; \vec{\nabla}\vec{N}\,(\vec{v},\, \vec{w}) \;=\; v^i \left(\frac{F_{,j}}{|\nabla F|} \right)_i w^j \tag{9.27}$$

when $\vec{v}, \vec{w} \in T_p\Sigma$. This interprets $\vec{\nabla}\vec{N}$ as an operator

$$\vec{\nabla}\vec{N} : T_p\Sigma \times T_p\Sigma \;\longrightarrow\; \mathbb{R}. \tag{9.28}$$

If we insert just one vector instead of two, we can interpret

$$\vec{\nabla}\vec{N} : T_p\Sigma \;\longrightarrow\; T_p\Sigma \tag{9.29}$$

To be specific,

$$\vec{\nabla}\vec{N}\,(\vec{v}) \;=\; v^i \left(\frac{F_{,k}}{|\nabla F|} \right)_i \delta^{kj} \vec{e}_j. \tag{9.30}$$

Interpretted this way, $\vec{\nabla}\vec{N}$ is known as the *Weingarten map*, or the *shape operator*.

a) Justify the claim that $\vec{\nabla}\vec{N} : T_p\Sigma \to T_p\Sigma$ (that is, show that the target space is what is claimed).

b) A surface Σ is called *umbilic* at $p \in \Sigma$ if the shape operator is a multiple of the identity operator. At an umbilic point, show that II is proportional to I.

**c)* If $p \in \Sigma$ is an umbilic point, what can you say about curvatures at p? If Σ is without boundary, is non-singular, and is umbilic at *all* points, what can you say about Σ?

4) When $\vec{v}$ and $\vec{w}$ are the special vector fields constructed above, we showed that $\frac{\partial \vec{w}}{\partial \vec{v}}\big|_p = \frac{\partial \vec{v}}{\partial \vec{w}}\big|_p$. However $\frac{\partial \vec{w}}{\partial \vec{v}} = \frac{\partial \vec{v}}{\partial \vec{w}}$ is almost always false if $\vec{v}$ and $\vec{w}$ are not these special fields.

a) Show by example that $\frac{\partial \vec{w}}{\partial \vec{v}} = \frac{\partial \vec{v}}{\partial \vec{w}}$ is not always true.

**b)* Let $\vec{\varphi} : \Omega \to \Sigma$ be a parametrization, where Ω is a domain in $\mathbb{R}^2$. We can express $\vec{\varphi}(s,t)$ as

$$\vec{\varphi}(s,t) \;=\; \begin{pmatrix} \varphi^1(s,t) \\ \varphi^2(s,t) \\ \varphi^3(s,t) \end{pmatrix} \tag{9.31}$$

Let $\vec{X}_1 = (1,0)^T$ and $\vec{X}_2 = (0,1)^T$ be the standard coordinate fields in $\mathbb{R}^2$. These fields "push forward" under $\vec{\varphi}$ to become fields

$$\begin{aligned} \vec{v}_1 &= \frac{\partial \varphi^i}{\partial s}\vec{e}_i \\ \vec{v}_2 &= \frac{\partial \varphi^i}{\partial t}\vec{e}_i \end{aligned} \tag{9.32}$$

Prove that $\frac{\partial \vec{v}_1}{\partial \vec{v}_2} = \frac{\partial \vec{v}_2}{\partial \vec{v}_1}$. This is a weakened version of what is known as the *Frobenius integrability theorem.*

Problems due Thursday 2/27

Lecture 10 - The Theorema Egregium

Lecture given on Tuesday Feb 25, 2012.

Section 10.2 of today's notes is based on portions of Spivak's Volume II, Chapter 3

10.1 Two lemmas. Commuting vector fields.

The first lemma is a fact about determinants of matrices

Lemma 10.1.1 *Let*

$$A = \begin{pmatrix} a & b & c \\ d & I & J \\ e & K & L \end{pmatrix}, \quad B = \begin{pmatrix} \alpha & \beta & \gamma \\ \delta & I & J \\ \gamma & K & L \end{pmatrix} \tag{10.1}$$

be two matrices with a common same lower-right 2×2 *minor. Then*

$$\begin{aligned} & Det \begin{pmatrix} a & b & c \\ d & I & J \\ e & K & L \end{pmatrix} \pm Det \begin{pmatrix} \alpha & \beta & \gamma \\ \delta & I & J \\ \gamma & K & L \end{pmatrix} \\ = \quad & Det \begin{pmatrix} a \pm \alpha & b & c \\ d & I & J \\ e & K & L \end{pmatrix} \pm Det \begin{pmatrix} 0 & \beta & \gamma \\ \delta & I & J \\ \gamma & K & L \end{pmatrix} \end{aligned} \tag{10.2}$$

Proof. Let $\mathcal{A}_{ij}$ (respectively, $\mathcal{B}_{ij}$) be the ij minor of A (repsectively, B). Then

$$\begin{aligned} Det(A) &= a\mathcal{A}_{11} - b\mathcal{A}_{12} + c\mathcal{A}_{13} \\ Det(B) &= \alpha\mathcal{B}_{11} - \beta\mathcal{B}_{12} + \gamma\mathcal{B}_{13}. \end{aligned} \tag{10.3}$$

But since $\mathcal{A}_{11} = \mathcal{B}_{11}$ we have

$$\begin{aligned} Det(A) \pm Det(B) &= (a \pm \alpha)\mathcal{A}_{11} - b\mathcal{A}_{12} - c\mathcal{A}_{13} \pm (0 \cdot \mathcal{B}_{11} - \beta\mathcal{B}_{12} + \gamma\mathcal{B}_{13}) \\ &= Det \begin{pmatrix} a \pm \alpha & b & c \\ d & I & J \\ e & K & L \end{pmatrix} \pm Det \begin{pmatrix} 0 & \beta & \gamma \\ \delta & I & J \\ \gamma & K & L \end{pmatrix} \end{aligned} \tag{10.4}$$

□

This lets us establish some new terminology. Two vector fields are said to *commute* if and only if

$$\frac{\partial \vec{v}}{\partial \vec{w}} = \frac{\partial \vec{w}}{\partial \vec{v}}. \tag{10.5}$$

We define the bracket of two vector fields to be

$$[\vec{v}, \vec{w}] \triangleq \frac{\partial \vec{w}}{\partial \vec{v}} - \frac{\partial \vec{v}}{\partial \vec{w}} \tag{10.6}$$

The second lemma should be familiar-sounding: it asserts the commutativity of mixed partials in a new situation. But we learn that mixed partial commute only when $\frac{\partial \vec{v}}{\partial \vec{w}} = \frac{\partial \vec{w}}{\partial \vec{v}}$.

Lemma 10.1.2 *Let $\vec{v}$, $\vec{w}$, and $\vec{z}$ be any three vector fields in $\mathbb{E}^n$. Then*

$$\frac{\partial^2 \vec{z}}{\partial \vec{v} \partial \vec{w}} - \frac{\partial^2 \vec{z}}{\partial \vec{w} \partial \vec{v}} = \frac{\partial \vec{z}}{\partial \, [\vec{v}, \vec{w}]}. \tag{10.7}$$

Therefore, if, at some point p, we have

$$\left.\frac{\partial \vec{v}}{\partial \vec{w}}\right|_p = \left.\frac{\partial \vec{w}}{\partial \vec{v}}\right|_p \tag{10.8}$$

then

$$\left.\frac{\partial^2 \vec{z}}{\partial \vec{v} \partial \vec{w}}\right|_p = \left.\frac{\partial^2 \vec{z}}{\partial \vec{w} \partial \vec{v}}\right|_p. \tag{10.9}$$

<u>*Proof*</u>.

First we express these fields in component form

$$\vec{v} = v^i \vec{e}_i, \quad \vec{w} = w^i \vec{e}_i, \quad \textit{and} \quad \vec{z} = z^i \vec{e}_i, \tag{10.10}$$

and then we compute:

$$\begin{aligned} \frac{\partial}{\partial \vec{w}} \frac{\partial \vec{z}}{\partial \vec{v}} &= \frac{\partial}{\partial \vec{w}} \left(\frac{\partial z^i}{\partial \vec{v}} \right) e^i \\ &= w^k \frac{\partial}{\partial x^k} \left(v^j \frac{\partial z^i}{\partial x^j} \right) e^i \end{aligned} \tag{10.11}$$

and likewise

$$\frac{\partial}{\partial \vec{v}} \frac{\partial \vec{z}}{\partial \vec{w}} = v^k \frac{\partial}{\partial x^k} \left(w^j \frac{\partial z^i}{\partial x^j} \right) e^i \tag{10.12}$$

Therefore

$$\begin{aligned} \frac{\partial}{\partial \vec{w}} \frac{\partial \vec{z}}{\partial \vec{v}} - \frac{\partial}{\partial \vec{v}} \frac{\partial \vec{z}}{\partial \vec{w}} &= w^k \frac{\partial}{\partial x^k} \left(v^j \frac{\partial z^i}{\partial x^j} \right) e^i - v^k \frac{\partial}{\partial x^k} \left(w^j \frac{\partial z^i}{\partial x^j} \right) e^i \\ &= \left(w^k \frac{\partial v^j}{\partial x^k} - v^k \frac{\partial w^j}{\partial x^k} \right) \frac{\partial z^i}{\partial x^j} e^i + \left(w^k v^j \frac{\partial^2 z^i}{\partial x^k \partial x^j} - v^k w^j \frac{\partial^2 z^i}{\partial x^k \partial x^j} \right) e^i \end{aligned} \tag{10.13}$$

The second term vanishes simply by commutativity of partial derivatives. For the first term, consider the computation

$$\frac{\partial \vec{v}}{\partial \vec{w}} - \frac{\partial \vec{w}}{\partial \vec{v}} = \left(w^k \frac{\partial v^j}{\partial x^k} - v^k \frac{\partial w^j}{\partial x^k} \right) \vec{e}_j \tag{10.14}$$

This is precisely the coefficient of the first term in (10.13). Therefore (10.13) is zero at p provided $\frac{\partial \vec{v}}{\partial \vec{w}} = \frac{\partial \vec{w}}{\partial \vec{v}}$ at p. □

10.2 Continuation of the Proof of the Theorema Egregium

10.2.1 Set-up from last time

Recall we had labelled the components of I and II as follows:

$$\begin{aligned} I &= \begin{pmatrix} E & F \\ F & G \end{pmatrix} \\ II &= \begin{pmatrix} l & m \\ m & n \end{pmatrix}. \end{aligned} \tag{10.15}$$

We have

$$\begin{aligned} l &= \frac{1}{\sqrt{EG - F^2}} \left\langle \vec{v} \times \vec{w}, \frac{\partial \vec{v}}{\partial \vec{v}} \right\rangle \\ m &= \frac{1}{\sqrt{EG - F^2}} \left\langle \vec{v} \times \vec{w}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle \\ n &= \frac{1}{\sqrt{EG - F^2}} \left\langle \vec{v} \times \vec{w}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle \end{aligned} \tag{10.16}$$

and we computed that

$$
\begin{aligned}
Det(I)^2 \cdot K \;=\; Det & \begin{pmatrix} \left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle & \left\langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle & \left\langle \vec{w}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle \\ \left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \vec{v} \right\rangle & \langle \vec{v}, \vec{v} \rangle & \langle \vec{w}, \vec{v} \rangle \\ \left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \vec{w} \right\rangle & \langle \vec{v}, \vec{w} \rangle & \langle \vec{w}, \vec{w} \rangle \end{pmatrix} \\
- \; Det & \begin{pmatrix} \left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle & \left\langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle & \left\langle \vec{w}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle \\ \left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \vec{v} \right\rangle & \langle \vec{v}, \vec{v} \rangle & \langle \vec{w}, \vec{v} \rangle \\ \left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \vec{w} \right\rangle & \langle \vec{v}, \vec{w} \rangle & \langle \vec{w}, \vec{w} \rangle \end{pmatrix}
\end{aligned}
\tag{10.17}
$$

10.2.2 Continuation of the Proof

The next step is to use Lemma 10.1.1 to rewrite (10.17). We have

$$
\begin{aligned}
(Det\, I)^2 K \;=\; Det & \begin{pmatrix} \left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle - \left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle & \left\langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle & \left\langle \vec{w}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle \\ \left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \vec{v} \right\rangle & \langle \vec{v}, \vec{v} \rangle & \langle \vec{w}, \vec{v} \rangle \\ \left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \vec{w} \right\rangle & \langle \vec{v}, \vec{w} \rangle & \langle \vec{w}, \vec{w} \rangle \end{pmatrix} \\
- \; Det & \begin{pmatrix} 0 & \left\langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle & \left\langle \vec{w}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle \\ \left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \vec{v} \right\rangle & \langle \vec{v}, \vec{v} \rangle & \langle \vec{w}, \vec{v} \rangle \\ \left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \vec{w} \right\rangle & \langle \vec{v}, \vec{w} \rangle & \langle \vec{w}, \vec{w} \rangle \end{pmatrix}
\end{aligned}
\tag{10.18}
$$

Then we make the simple observations that

$$
\begin{aligned}
\left\langle \frac{\partial \vec{w}}{\partial \vec{w}}, \vec{v} \right\rangle &= \frac{\partial}{\partial \vec{w}} \langle \vec{w}, \vec{v} \rangle - \frac{1}{2}\frac{\partial}{\partial \vec{v}} \langle \vec{w}, \vec{w} \rangle = \frac{\partial F}{\partial \vec{w}} - \frac{1}{2}\frac{\partial G}{\partial \vec{v}} \\
\left\langle \frac{\partial \vec{w}}{\partial \vec{w}}, \vec{w} \right\rangle &= \frac{1}{2}\frac{\partial}{\partial \vec{w}} \langle \vec{w}, \vec{w} \rangle = \frac{1}{2}\frac{\partial G}{\partial \vec{w}} \\
\left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \vec{v} \right\rangle &= \frac{1}{2}\frac{\partial}{\partial \vec{v}} \langle \vec{v}, \vec{v} \rangle = \frac{1}{2}\frac{\partial E}{\partial \vec{v}} \\
\left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \vec{w} \right\rangle &= \frac{\partial}{\partial \vec{v}} \langle \vec{v}, \vec{w} \rangle - \frac{1}{2}\frac{\partial}{\partial \vec{w}} \langle \vec{v}, \vec{v} \rangle = \frac{\partial F}{\partial \vec{v}} - \frac{1}{2}\frac{\partial E}{\partial \vec{w}} \\
\left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \vec{v} \right\rangle &= \frac{1}{2}\frac{\partial E}{\partial \vec{w}} \\
\left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \vec{w} \right\rangle &= \frac{1}{2}\frac{\partial G}{\partial \vec{v}}
\end{aligned}
\tag{10.19}
$$

and rewrite the expression as

$$
\begin{aligned}
Det(I)^2 \cdot K \;=\; Det & \begin{pmatrix} \left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle - \left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle & \frac{\partial F}{\partial \vec{w}} - \frac{1}{2}\frac{\partial G}{\partial \vec{v}} & \frac{1}{2}\frac{\partial G}{\partial \vec{w}} \\ \frac{1}{2}\frac{\partial E}{\partial \vec{v}} & E & F \\ \frac{\partial F}{\partial \vec{v}} - \frac{1}{2}\frac{\partial E}{\partial \vec{w}} & F & G \end{pmatrix} \\
- \; Det & \begin{pmatrix} 0 & \frac{1}{2}\frac{\partial E}{\partial \vec{w}} & \frac{1}{2}\frac{\partial G}{\partial \vec{v}} \\ \frac{1}{2}\frac{\partial E}{\partial \vec{w}} & E & F \\ \frac{1}{2}\frac{\partial G}{\partial \vec{v}} & F & G \end{pmatrix}
\end{aligned}
\tag{10.20}
$$

Lastly we deal with the second derivative term. We compute

$$\begin{aligned}&\left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle - \left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle \\ &= \frac{\partial}{\partial \vec{v}} \left\langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle - \left\langle \vec{v}, \frac{\partial^2 \vec{w}}{\partial \vec{v} \partial \vec{w}} \right\rangle - \frac{\partial}{\partial \vec{w}} \left\langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle - \left\langle \vec{v}, \frac{\partial^2 \vec{w}}{\partial \vec{w} \partial \vec{v}} \right\rangle\end{aligned} \tag{10.21}$$

By the commutativity of partial derivatives, we have

$$\begin{aligned}\left\langle \frac{\partial \vec{v}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle - \left\langle \frac{\partial \vec{w}}{\partial \vec{v}}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle &= \frac{\partial}{\partial \vec{v}} \left\langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{w}} \right\rangle - \frac{\partial}{\partial \vec{w}} \left\langle \vec{v}, \frac{\partial \vec{w}}{\partial \vec{v}} \right\rangle \\ &= \frac{\partial}{\partial \vec{v}} \left(\frac{\partial}{\partial \vec{w}} \langle \vec{v}, \vec{w} \rangle - \frac{1}{2} \frac{\partial}{\partial \vec{v}} \langle \vec{w}, \vec{w} \rangle \right) - \frac{1}{2} \frac{\partial}{\partial \vec{w}} \left(\frac{\partial}{\partial \vec{w}} \langle \vec{v}, \vec{v} \rangle \right) \\ &= \frac{\partial^2 F}{\partial \vec{w} \partial \vec{v}} - \frac{1}{2} \frac{\partial^2 G}{\partial \vec{v} \partial \vec{v}} - \frac{1}{2} \frac{\partial^2 E}{\partial \vec{w} \partial \vec{w}}\end{aligned} \tag{10.22}$$

At last we arrive at

$$K = \frac{Det \begin{pmatrix} \frac{\partial^2 F}{\partial \vec{w} \partial \vec{v}} - \frac{1}{2} \frac{\partial^2 G}{\partial \vec{v} \partial \vec{v}} - \frac{1}{2} \frac{\partial^2 E}{\partial \vec{w} \partial \vec{w}} & \frac{\partial F}{\partial \vec{w}} - \frac{1}{2} \frac{\partial G}{\partial \vec{v}} & \frac{1}{2} \frac{\partial G}{\partial \vec{w}} \\ \frac{1}{2} \frac{\partial E}{\partial \vec{v}} & E & F \\ \frac{\partial F}{\partial \vec{v}} - \frac{1}{2} \frac{\partial E}{\partial \vec{w}} & F & G \end{pmatrix} - Det \begin{pmatrix} 0 & \frac{1}{2} \frac{\partial E}{\partial \vec{w}} & \frac{1}{2} \frac{\partial G}{\partial \vec{v}} \\ \frac{1}{2} \frac{\partial E}{\partial \vec{w}} & E & F \\ \frac{1}{2} \frac{\partial G}{\partial \vec{v}} & F & G \end{pmatrix}}{Det(I)^2} \tag{10.23}$$

10.3 Exercises

1) Formally prove that, for any function f, we have

$$\frac{\partial^2 f}{\partial \vec{v} \partial \vec{w}} - \frac{\partial^2 f}{\partial \vec{w} \partial \vec{v}} = \frac{\partial f}{\partial [\vec{v}, \vec{w}]}. \tag{10.24}$$

2) Prove the *Jacobi identity*: given three arbitrary (differentiable) fields $\vec{v}$, $\vec{w}$, and $\vec{z}$, show

$$[\vec{v}, [\vec{w}, \vec{z}]] + [\vec{w}, [\vec{z}, \vec{v}]] + [\vec{z}, [\vec{v}, \vec{w}]] = 0. \tag{10.25}$$

Hint: do not try to evaluate directly. Rather, use (10.7) to write

$$\begin{aligned}\frac{\partial}{\partial [\vec{v}, [\vec{w}, \vec{z}]]} &= \frac{\partial}{\partial \vec{v}} \frac{\partial}{\partial [\vec{w}, \vec{z}]} - \frac{\partial}{\partial [\vec{w}, \vec{z}]} \frac{\partial}{\partial \vec{v}} \\ &= \frac{\partial^3}{\partial \vec{v} \partial \vec{w} \partial \vec{z}} - \frac{\partial^3}{\partial \vec{v} \partial \vec{z} \partial \vec{w}} - \frac{\partial^3}{\partial \vec{w} \partial \vec{z} \partial \vec{v}} + \frac{\partial^3}{\partial \vec{z} \partial \vec{w} \partial \vec{v}}\end{aligned} \tag{10.26}$$

and so forth. Finally, conclude that when a directional derivative is zero on all functions, the direction itself must be the "zero" direction.

Lecture 11 - The end of Gauss' paper

Tuesday March 20, 2014

11.1 Review

If $\vec{v}$, $\vec{w}$ are vector fields n a surface Σ and $[\vec{v},\vec{w}]=0$, and if

$$I = \begin{pmatrix} E & F \\ F & G \end{pmatrix} = \begin{pmatrix} |\vec{v}|^2 & \langle \vec{v},\vec{w}\rangle \\ \langle \vec{v},\vec{w}\rangle & |\vec{w}|^2 \end{pmatrix} \tag{11.1}$$

then we determined that

$$K = \frac{Det\begin{pmatrix} \frac{\partial^2 F}{\partial\vec{w}\partial\vec{v}} - \frac{1}{2}\frac{\partial^2 G}{\partial\vec{v}\partial\vec{v}} - \frac{\partial^2 E}{\partial\vec{w}\partial\vec{w}} & \frac{\partial F}{\partial\vec{w}} - \frac{1}{2}\frac{\partial G}{\partial\vec{v}} & \frac{1}{2}\frac{\partial G}{\partial\vec{w}} \\ \frac{1}{2}\frac{\partial E}{\partial\vec{v}} & E & F \\ \frac{\partial F}{\partial\vec{v}} - \frac{1}{2}\frac{\partial E}{\partial\vec{w}} & F & G \end{pmatrix} - Det\begin{pmatrix} 0 & \frac{1}{2}\frac{\partial E}{\partial\vec{w}} & \frac{1}{2}\frac{\partial G}{\partial\vec{v}} \\ \frac{1}{2}\frac{\partial E}{\partial\vec{w}} & E & F \\ \frac{1}{2}\frac{\partial G}{\partial\vec{v}} & F & G \end{pmatrix}}{Det(I)^2}. \tag{11.2}$$

This was Gauss' computation of K solely in terms of the surface's intrinsic geometry (as encoded by its first fundamental form).

11.2 Results from the In-Class Worksheet

Here we will outline the intended results from our in-class worksheet, culminating in the formula ????.

11.2.1 The intrinsic distance function

Let Σ be a non-singular surface in $\mathbb{R}^3$, and choose a point $p \in \Sigma$. Define the function

$$\begin{aligned} r : \Sigma \to \mathbb{R} \\ r(q) \;=\; \mathrm{dist}(p, q) \end{aligned} \tag{11.3}$$

where $dist(p, q)$ indicate the length of the shortest path from p to q, where the paths under consideration are required to lie on the surface Σ itself. This is the *intrinsic distance function*.

11.2.2 Geodesic polar coordinates

We shall take (without proof) that if $p, q \in \Sigma$ are points, there is at least one shortest path form p to q. Pick an arbitrary vector $\vec{v}_0 \in T_p\Sigma$; we shall call this the reference vector. Given a point $q \in \Sigma$, define the value $\theta(q)$ to be

$$\theta(q) \;=\; \text{angle that the initial vector of the shortest path from } p \text{ to } q \text{ makes with } \vec{v} \tag{11.4}$$

If there is more than one shortest vector from p to q, then leave $\theta(q)$ undefined. Also, $\theta(p)$ is undefined. We shall take for granted that θ is uniquely defined on a at least some region around p.

We can take r, θ to a coordinate system for a region near p. That is, the point q has the coordinates $(r(p), \theta(q))$. These are called *geodesic polar coordinates* around p. In the can of $\mathbb{R}^2$, if we choose $\vec{v}$ to be the unit vector in the x^1-direction, then the geodesic normal coordinates at the origin are simply the standard polar coordinates on $\mathbb{R}^2$.

11.2.3 The coordinate fields

Given the geodesic polar coordinates (r, θ), we have partial derivatives:

$$\begin{aligned} \frac{\partial f}{\partial r} &= \text{rate of change in } f \text{ in the direction of unit change in } r \text{ and no change in } \theta \\ \frac{\partial f}{\partial \theta} &= \text{rate of change in } f \text{ in the direction of unit change in } \theta \text{ and no change in } r \end{aligned} \tag{11.5}$$

We can then define the coordinate fields $\vec{v}_r$ and $\vec{v}_\theta$ implicitly by

$$\begin{aligned} \frac{\partial}{\partial \vec{v}_r} &= \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \vec{v}_\theta} &= \frac{\partial}{\partial \theta} \end{aligned} \tag{11.6}$$

Note that these two fields actually commute. To see this, note that

$$\frac{\partial}{\partial[\vec{v}_r, \vec{v}_\theta]} \;=\; \frac{\partial^2}{\partial\vec{v}_r\partial\vec{v}_\theta} - \frac{\partial^2}{\partial\vec{v}_\theta\partial\vec{v}_r} \;=\; \frac{\partial^2}{\partial r\partial\theta} - \frac{\partial^2}{\partial\theta\partial r} \;=\; 0 \tag{11.7}$$

The final equality is simply the usual commutativity of partial derivatives. Because $\frac{\partial}{\partial[\vec{v}_r,\vec{v}_\theta]}$ is the zero directional derivative, then $[\vec{v}_r, \vec{v}_\theta]$ must be the "zero" direction, meaning the zero vector.

Finally, it is easy to see that $|\vec{v}_r|^2 = 1$, and it is reasonable to conclude that $\langle\vec{v}_r, \vec{v}_\theta\rangle = 0$ (this is a bit difficult to prove formally; it is a version of what is called "Gauss' Lemma").

11.2.4 Curvature computation

Defining the function $g = g(r, \theta)$ by $g = |\vec{v}_\theta|^2$, we have the first fundamental form

$$I \;=\; \begin{pmatrix} 1 & 0 \\ 0 & g \end{pmatrix} \tag{11.8}$$

Because $[\vec{v}_r, \vec{v}_\theta] = 0$, we can use (11.2) to compute

$$K \;=\; -\frac{1}{\sqrt{g}}\frac{\partial^2\sqrt{g}}{\partial r^2} \tag{11.9}$$

This formula is remarkable for its simplicity and compactness. The hard work we did in carefully defining our coordinate system lead to a far, far simpler expression for K.

11.2.5 Taylor series for K

It is clear that $\lim_{r\to 0}\sqrt{g}(r,\theta) \;=\; 0$. Although somewhat less clear, it is reasonable to guess that $\lim_{r\to 0}\frac{\partial\sqrt{g}}{\partial r}(r,\theta) \;=\; 1$. From this and (11.9) we find $\lim_{r\to 0}\frac{\partial^2\sqrt{g}}{\partial r^2}(r,\theta) = 0$ and $\lim_{r\to 0}\frac{\partial^3\sqrt{g}}{\partial r^3}(r,\theta) = -K$ Therefore the Taylor series is

$$\begin{aligned} \sqrt{g}(r,\theta) \;&=\; r \;-\; \frac{1}{6}K(0,\theta)r^3 \;+\; \text{terms in } r \text{ and } \theta \text{ of order 4 or more in } r \\ \sqrt{g}(r,\theta) \;&=\; r \;-\; \frac{1}{6}K(p)r^3 \;+\; r^4\,h(r,\theta). \end{aligned} \tag{11.10}$$

11.2.6 Circumference of geodesic circles

For a fixed number ρ, consider the locus $r = \rho$. This will usually not actually be a circle, but, by abusing terminology, we may call it a *geodesic circle*. The circumference of the geodesic circle is then

$$C(\rho) \;=\; \int_0^{2\pi} |\vec{v}_\theta|\, d\theta \;=\; \int_0^{2\pi} \sqrt{g}(\rho,\theta)\, d\theta. \tag{11.11}$$

11.2.7 Conclusion: A new curvature formula

Combining (11.10) and (11.11), we have

$$\begin{aligned} C(r) \;&=\; \int_0^{2\pi} \left(r \,-\, \frac{1}{6}K(p)r^3 \,+\, r^4 h(r,\theta) \right)\, d\theta \\ &=\; 2\pi r \,-\, \frac{1}{6} r^3 K(p) \,+\, r^4 \int_0^{2\pi} h(r,\theta)\, d\theta \end{aligned} \tag{11.12}$$

Re-arranging this, we have

$$\frac{3}{\pi}\frac{2\pi r \,-\, C(r)}{r^3} \;=\; K \,-\, \frac{3}{\pi} r \int_0^{2\pi} h(r,\theta)\, d\theta. \tag{11.13}$$

Therefore

$$K \;=\; \frac{3}{\pi} \lim_{r\to 0} \frac{C_E(r) \,-\, C(r)}{r^3} \tag{11.14}$$

where $C_E(r) = 2\pi r$ is the circumference of the Euclidean circle of radius r.

This is a remarkable formula. It gives a purely intrinsic way of computing the Gaussian curvature K at a point p. Namely, if geodesic circles are too small compared to Euclidean circles, the curvature is positive. If geodesic circles are too big compared to Euclidean circles, the curvature is positive.

11.3 Exercises

1) Let p be a point on a sphere of radius ρ.

   a) Letting $r(q) = \mathrm{dist}(p,q)$ be the intrinsic distance on the sphere. If $r_0 \in [0, 2\pi\rho]$ is a constant, show that the arclength of the locus $r = r_0$ is $2\pi\rho\sin(r_0/\rho)$.

   b) Using only part (*a*) and (11.14), determine the Gaussian curvature of the sphere of radius ρ.

2) Do problem C.1.2.

3) Do problem C.1.3.

*4) Consider the vector space $\mathcal{V}$ of functions $f : \mathbb{R} \to \mathbb{R}$ that are infinitely differentiable, and that have compact support (for those who know the terminology, $\mathcal{V}$ is the vector space $C_c^\infty(\mathbb{R})$). This is called the space of *test functions.* The vector space $\mathcal{V}^*$ is called the space of *distributions.*

a) Show that the function

$$f(x) \;=\; \begin{cases} 0 & x \in (-\infty, -1) \cup (1, \infty) \\ e^{-\frac{1}{1-x^2}} & x \in [-1, 1] \end{cases} \tag{11.15}$$

is indeed an element of $\mathcal{V}$.

b) Show that if $x \in \mathbb{R}$, namely x is a point in the domain of the functions $f \in \mathcal{V}$, then we can consider $x \in \mathcal{V}^*$.

c) If $g : \mathcal{R} \to \mathcal{R}$ is any function that is locally integrable, give a way of interpreting $g \in \mathcal{V}^*$ (so functions are themselves also distributions). In particular, exhibit a natural embedding $\mathcal{V} \subset \mathcal{V}^*$.

d) (Rigorous definition of the Dirac delta.) Let $\delta : \mathcal{V} \to \mathbb{R}$ be the linear functional given by

$$\delta(f) \;=\; f(0). \tag{11.16}$$

Explain why δ should be considered the Dirac delta "function."

Lecture 12 - New Vector Spaces from Old

Thursday March 20, 2014

12.1 Vector Spaces and Dual Spaces

12.1.1 Definition of the dual of a vector space

To every vector space V we can canonically associate a second vector space, V^*, called the *dual* of V. The space V^* is defined to be the *vector space of linear functionals on* V. Specifically

$$V^* \;=\; \left\{\, F : V \to \mathbb{R} \;\middle|\; F \text{ is linear} \,\right\}. \tag{12.1}$$

Obviously there must be an addition operation and a scalar multiplication:

$$\begin{aligned} &cF : V \to \mathbb{R} \quad \text{is the operator} \quad (cF)(x) = c\,F(x) \\ &(F+G) : V \to \mathbb{R} \quad \text{is the operator} \quad (F+G)(x) \;=\; F(x) + G(x). \end{aligned} \tag{12.2}$$

An element of the vector space V^* is usually called a *covector*.

It is typical to represent vectors (elements of V) with Latin letters and lower indices: $\vec{v}_1, \vec{v}_2$ and so forth. Likewise it is typical to represent covectors (elements of V^*) with Greek letters and upper indices: η^1, η^2 and so forth.

12.1.2 Equality of dimension

Obviously a basic tool in the study of vector spaces is a choice of basis. Let

$$\vec{e}_1,\, \dots,\, \vec{e}_n \tag{12.3}$$

be a basis of V. Given such a basis for V, we can find a corresponding set of basis covectors in V^*, namely linear functionals $\{\eta^1,\, \dots,\, \eta^n\}$ defined (implicitly) by

$$\eta^i(\vec{e}_j) \;=\; \delta^i_j. \tag{12.4}$$

It is easy to see that the functionals $\eta^1, \ldots, \eta^n$ are independent: assuming $a_1, \ldots, a_n$ are constants, then consider the operator $a_1\eta^1 + \ldots + a_n\eta^n$ and its action on the vector $\vec{e}_i$:

$$\begin{aligned}\left(a_1\eta^1 + \ldots + a_n\eta^n\right)(\vec{e}_i) &= a_1\eta^1(\vec{e}_i) + \ldots + a_n\eta^n(\vec{e}_i) \\ &= a_1\delta_i^1 + \ldots + a_n\delta_i^n = a_i.\end{aligned} \tag{12.5}$$

Therefore $a_1\eta^1 + \ldots + a_n\eta^n$ is the zero operator if and only if each a_i is zero. Thus the $\eta^1, \ldots, \eta^n$ are linearly independent.

Second, we can show that the independent set $\{\eta^i\}$ actually spans V^*. So assume $F \in V^*$ is a linear functional. Define numbers F_i by $F_i = F(\vec{v}_i)$, and then define a second linear operator $\overline{F} = F_1\eta^1 + \ldots + F_n\eta^n$. We shall show that $F = \overline{F}$, verifying that F is indeed a linear combination of the η^i.

So let $\vec{v} = v^1\vec{e}_1 + \ldots + v^n\vec{e}_n$ be an arbitrary vector in V. We compute the action of both F and $\overline{F}$ on $\vec{v}$:

$$\begin{aligned}F(\vec{v}) &= F\left(v^1\vec{e}_1 + \ldots + v^n\vec{e}_n\right) \\ &= v^1F(\vec{e}_1) + \ldots + v^nF(\vec{e}_n) \\ &= v^1F_1 + \ldots + v^nF_n \\ \overline{F}(\vec{v}) &= \left(F_1\eta^1 + \ldots + F_n\eta^n\right)\left(v^1\vec{e}_1 + \ldots + v^n\vec{e}_n\right) \\ &= F_1\eta^1\left(v^1\vec{e}_1 + \ldots + v^n\vec{e}_n\right) + \ldots + F_n\eta^n\left(v^1\vec{e}_1 + \ldots + v^n\vec{e}_n\right) \\ &= F_1v^1\eta^1(\vec{e}_1) + \ldots + F_1v^n\eta^1(\vec{e}_n) \\ &\quad + \ldots \\ &\quad + F_nv^1\eta^n(\vec{e}_1) + \ldots + F_nv^n\eta^n(\vec{e}_n) \\ &= F_1v^1\delta_1^1 + \ldots + F_1v^n\delta_n^1 \\ &\quad + \ldots \\ &\quad + F_nv^1\delta_1^n + \ldots + F_nv^n\delta_n^n \\ &= F_1v^1 + \ldots + F_nv^n.\end{aligned} \tag{12.6}$$

Therefore, indeed, $F(\vec{v}) = \overline{F}(\vec{v})$ for any vector $\vec{v}$, so $F = \overline{F}$ as operators.

12.1.3 The double-dual V^{**}

An obvious question is, what happens when we take the dual of the dual? In fact, the double dual V^{**} is canonically isomorphic to V.

To see why this is, consider that $\vec{v}$ acts on a linear function F via *transposition*:

$$\vec{v}(F) = F(\vec{v}) \in \mathbb{R}. \tag{12.7}$$

The Kernel of this map is trivial: if $\vec{v}$ is the zero operator, then indeed $\vec{v}$ must be zero. To see this, again choose a basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ (so that we automatically also have a dual basis

$\{\eta^1, \ldots, \eta^n\}$) and express $\vec{v}$ in this basis:

$$\vec{v} \;=\; v^1\vec{e}_1 \,+\, \ldots \,+\, v^n\vec{e}_n, \qquad F = F_1\eta^1 \,+\, \ldots \,+\, F_n\eta^n. \tag{12.8}$$

Then consider the evaluation of $\vec{v}$ on each of the basis elements of the dual space:

$$\begin{aligned} \vec{v}(\eta^i) \;&=\; \left(v^1\vec{e}_1 \,+\, \ldots \,+\, v^n\vec{e}_n\right)\left(\eta^i\right) \\ &=\; v^1\vec{e}_1\left(\eta^i\right) \,+\, \ldots \,+\, v^n\vec{e}_n\left(\eta^i\right) \\ &=\; v^1\delta^i_1 \,+\, \ldots \,+\, v^n\delta^i_n \;=\; v^i \end{aligned} \tag{12.9}$$

Therefore $\vec{v}$ is the zero operator if and only if v^i is zero for every i, which means $\vec{v}$ is exactly the zero vector.

12.1.4 The Einstein Convention

The Einstein summation convention is a notational convenience that simplifies working in a basis.

By convention, a vector basis is indexed with lower indices, and the coefficients are indexed with upper indices:

$$\vec{v} \;=\; v^1\vec{e}_1 \,+\, \ldots \,+\, v^n\vec{e}_n \;=\; \sum_{i=1}^{n} v^i\vec{e}_i \tag{12.10}$$

and a covector basis is indexed with upper indices, with coefficients indexed with lower indices:

$$F \;=\; F_1\eta^1 \,+\, \ldots \,+\, F_n\eta^n \;=\; \sum_{i=1}^{n} F_i\eta^i. \tag{12.11}$$

The Einstein convention is simply to leave off the sum symbol, with an implicit understanding that repeated upper- and lower-index pairs are summed:

$$v^i\vec{e}_i \quad \text{means precisely} \quad \sum_{i=1}^{n} v^i\vec{e}_i. \tag{12.12}$$

12.2 Exercises

1) Do C.3.1. Specifically, prove that $Hom(V, W)$ is a vector space.

2) Consider the vector space V of quadratic polynomials in one variable. Let $\mathbf{v}_1 = 1$, $\mathbf{v}_2 = x$, and $\mathbf{v}_3 = x^2$ be a basis, and let $\{\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3\}$. Let $A : V \to \mathbb{R}$ be the linear functional

$$A(f) \;=\; \int_{\mathbb{R}} e^{-x^2} f(x)\, dx \tag{12.13}$$

What are A_1, A_2, A_3? Express A in terms of the given dual basis.

3) Do C.4.1.

Lecture 13 - The Tensor Algebra

Tuesday March 25, 2014

13.1 Material

We discussed more material from Appendix C of the book. Some aspects of class differed from the book's presentation, so we'll review our class material in the notes.

13.1.1 Review of Basics

Let V be a vector space. There is a second vector space, its *dual*, which is the vecotr space of linear operators:

$$V^* \;=\; \left\{ A : V \to \mathbb{R} \;\middle|\; A \text{ is linear} \right\}. \tag{13.1}$$

If it is finite dimensional, we may choose a basis

$$\mathcal{B} \;=\; \{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \tag{13.2}$$

after which we get a *dual basis*

$$\mathcal{B}^* \;=\; \left\{\eta^1, \ldots, \eta^n\right\}. \tag{13.3}$$

where the linear operators η_i are each defined, implicitly, by

$$\eta^i(\mathbf{e}_j) \;=\; \delta^i_j. \tag{13.4}$$

Given vector spaces V, W, we define the Hom-space:

$$Hom(V, W) \;=\; \left\{ F : V \to W \;\middle|\; F \text{ is linear}\right\}. \tag{13.5}$$

and also the tensor product of two vector spaces:

$$V \otimes W \;=\; \text{finite linear combinations of elements } \mathbf{v} \otimes \mathbf{w} \text{ where } \mathbf{v} \in V,\, \mathbf{w} \in W \tag{13.6}$$

13.1.2 The Tensor Algebra over V

The graded tensor algebra

Let V be a vector space. Then we have a variety of new vector spaces:

$$V^{\otimes k} \triangleq \bigotimes^k V \triangleq V \otimes V \otimes \cdots \otimes V \qquad (k\ many\ factors) \tag{13.7}$$

and by convention we define

$$V^{\otimes 0} \triangleq \bigotimes^0 V \triangleq \mathbb{R}. \tag{13.8}$$

We define the *graded Tensor algebra* to be the space

$$\begin{aligned} \bigotimes^* V &\triangleq \mathbb{R} \oplus V \oplus V^{\otimes 2} \oplus V^{\otimes 3} \oplus \dots \\ \bigotimes^* V &= \bigoplus_{k=0}^{\infty} \bigotimes^k V \end{aligned} \tag{13.9}$$

An element of $V^{\otimes k}$ is said to be homogeneous of degree k, or to be a tensor of degree k. The product is the obvious product.

The bigraded tensor algebra

We define the *bi-graded* tensor algebra over V to be

$$\bigotimes^{*,*} V = \left(\bigoplus_{k=0}^{\infty} \bigotimes^k V^* \right) \otimes \left(\bigoplus_{l=0}^{\infty} \bigotimes^l V \right) \tag{13.10}$$

Elements of this huge vector space have forms like

$$\begin{array}{ll} \mathbf{v}^1 \otimes \mathbf{v}_2 & \text{(homogeneous of bidegree (1,1))} \\ \mathbf{v}^2 \otimes \mathbf{v}^3 \otimes \mathbf{v}_2 & \text{(homogeneous of bidegree (2,1))} \\ \mathbf{v}_9 & \text{(homogeneous of bidegree (0,1))} \\ \mathbf{v}_9 + \mathbf{v}^1 \otimes \mathbf{v}_2 & \text{(not homogeneous)} \end{array} \tag{13.11}$$

Definition: An element of the subspace $\bigotimes^i V^* \otimes \bigotimes^j V \subset \bigotimes^{*,*} V$ is called a homogeneous tensor of bidegree (i, j). We define the algebra product as follows:

$$\begin{aligned} \text{if} \quad & \mathbf{v}^1 \otimes \dots \mathbf{v}^i \otimes \mathbf{v}_1 \otimes \dots \mathbf{v}_j \in \bigotimes^{i,j} V \\ \text{and} \quad & \mathbf{w}^1 \otimes \dots \mathbf{w}^k \otimes \mathbf{w}_1 \otimes \dots \mathbf{w}_l \in \bigotimes^{k,l} V \end{aligned} \tag{13.12}$$

then we define

$$\left(\mathbf{v}^1 \otimes \ldots \mathbf{v}^i \otimes \mathbf{v}_1 \otimes \ldots \mathbf{v}_j\right) \otimes \left(\mathbf{w}^1 \otimes \ldots \mathbf{w}^k \otimes \mathbf{w}_1 \otimes \ldots \mathbf{w}_l\right) \tag{13.13}$$

to be

$$\mathbf{v}^1 \otimes \ldots \mathbf{v}^i \otimes \mathbf{w}^1 \otimes \ldots \mathbf{w}^k \otimes \mathbf{v}_1 \otimes \ldots \mathbf{v}_j \otimes \mathbf{w}_1 \otimes \ldots \mathbf{w}_l \;\in\; \bigotimes^{i+k,j+l} V \tag{13.14}$$

To be specific, we allow elements of V^* and V to commute, but elements of V^* do not commute with each other, and elements of V do not commute with each other.

The tensor algebra

Lastly there is the full tensor algebra. This is similar to the bi-graded case, except we do not allow elements of V^* to commute with elements of V. For instance

$$\mathbf{v}^3 \otimes \mathbf{v}_4 \otimes \mathbf{v}^2 \tag{13.15}$$

is not the same as

$$\mathbf{v}^3 \otimes \mathbf{v}^2 \otimes \mathbf{v}_4 \tag{13.16}$$

13.1.3 An example

So what good are tensors? As an example, we'll see how all the theory of $n \times n$ matrices is merely a subset of the theory of tensors.

Consider a finite dimensional vector space V with basis $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. Then a generic element $v \in V$ can be expressed

$$v \;=\; v^i \mathbf{e}_i \tag{13.17}$$

where the v^i are scalars (that is, elements of $\mathbb{R}$). We can also express v as a column vector:

$$v \;=\; \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}_{[\mathcal{B}]} \;\triangleq\; v^1\mathbf{e}_1 \,+\, \ldots \,+\, v^n\mathbf{e_n} \tag{13.18}$$

If $\eta \in V^*$ is a generic element of the dual space, then $\eta = \eta_i \mathbf{e}^i$, and we can express η as a row vector:

$$\eta \;=\; (\eta_1,\, \ldots,\, \eta_n) \;\triangleq\; \eta_i \mathbf{e}^i \tag{13.19}$$

Now a generic element of $V^* \otimes V$ is a linear combination of basis vectors $\mathbf{e}^j \otimes \mathbf{e}_i$ can be expressed as

$$\begin{aligned} A \;&\in\; V^* \otimes V \\ A \;&=\; A^i_j\, \mathbf{e}^j \otimes \mathbf{e}_i \end{aligned} \tag{13.20}$$

From above, we know that $V^* \otimes V \in Hom(V, V)$. But we also know that $Hom(V, V)$, the space of endomorphisms on V, is, after choice of basis, the space of $n \times n$ matrices.

Considering an element $A \in Hom(V, V) = V^* \otimes V$ and element $v \in V$, we have that $A(v)$ is precisely

$$\begin{aligned} A(v) &= A(v^i \mathbf{e_i}) \\ &= v^i A(\mathbf{e_i}) \\ &= v^i A_i^j \mathbf{e}_j. \end{aligned} \tag{13.21}$$

and therefore, in column vector notation, we have

$$\begin{aligned} A(v) &= \begin{pmatrix} A_i^1 v^i \\ A_i^2 v^i \\ \vdots \\ A_i^n v^i \end{pmatrix}_{[\mathcal{B}]} \\ &= \begin{pmatrix} A_1^1 v^1 + A_2^1 v^2 + \ldots + A_n^1 v^n \\ A_1^2 v^1 + A_2^2 v^2 + \ldots + A_n^2 v^n \\ \vdots \\ A_1^n v^1 + A_2^n v^2 + \ldots + A_n^n v^n \end{pmatrix}_{[\mathcal{B}]} \\ &= \begin{pmatrix} A_1^1 & A_2^1 & \ldots & A_n^1 \\ A_1^2 & A_2^2 & & \\ \vdots & & \ddots & \vdots \\ A_1^n & A_2^n & \ldots & A_n^n \end{pmatrix}_{[\mathcal{B}]} \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix}_{[\mathcal{B}]} \end{aligned} \tag{13.22}$$

where

$$A = A_j^i \mathbf{e}^j \otimes \mathbf{e}_i \qquad \text{and} \qquad v = v^i \mathbf{e}_i. \tag{13.23}$$

Thus the action of A on v, when expressed in matrix notation, is nothing more than matrix multiplication.

13.2 Exercises

1) Let V be the vector space of polynomials of order up to 5 in the variable x. Let $A : V \to V$ be the operator $A = x\frac{d}{dx}$. Pick a basis, and express A as a matrix.

2) (Change of basis) Let $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\mathcal{C} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ be two different bases for the vector space V. Let A_j^i be the collection of numbers with the property that

$$\mathbf{f}_i = A_i^j \mathbf{e}_j. \tag{13.24}$$

Obviously A_j^i can be considered a matrix. Let B_j^i be the inverse matrix.

a) We took the inverse of the matrix A^i_j, but is this matrix even invertible? Explain how you know. (Hint: what is its rank space?)

b) Specifically what does it mean that A^i_j and B^i_j are inverse matrices? Express your answer with correct notation.

c) Express $\mathbf{e}_i$ in terms of the $\mathbf{f}_i$.

d) The two given bases each have their own set of dual bases: $\mathcal{B}^* = \{\mathbf{e}^1, \ldots, \mathbf{e}^n\}$ and $\mathcal{C}^* = \{\mathbf{f}^1, \ldots, \mathbf{f}^n\}$. Compute $\mathbf{e}^i(\mathbf{f}_j)$.

e) Using your computation from (*d*), determine the transition matrices for the dual bases.

3) If V is a finite dimensional vector space, then consider the vector space $V^* \otimes V$; we already know that this is isomorphic to $Hom(V, V)$. Show that it's also isomorphic to $Hom(V^*, V^*)$.

Lecture 14 - Inner and Outer Products

Thursday March 27, 2014

14.1 Contraction

If $r, s \geq 1$ then there are maps

$$\bigotimes^{r,s} V \longrightarrow \bigotimes^{r-1,s-1} V \tag{14.1}$$

given by

$$\mathbf{v}^{i_1} \otimes \cdots \otimes \mathbf{v}^{i_{r-1}} \otimes \mathbf{v}^{i_r} \otimes \mathbf{v}_{j_1} \otimes \mathbf{v}_{j_2} \otimes \cdots \otimes \mathbf{v}_{s_1} \mapsto \mathbf{v}^{i_r}(\mathbf{v}_{j_1}) \cdot \mathbf{v}^{i_1} \otimes \cdots \otimes \mathbf{v}^{i_{r-1}} \otimes \mathbf{v}_{j_2} \otimes \cdots \otimes \mathbf{v}_{s_1} \tag{14.2}$$

For instance consider an element of $V^* \otimes V$. The contraction of

$$A^i_j \mathbf{e}^j \otimes \mathbf{e}_i \tag{14.3}$$

is

$$A^i_j \mathbf{e}^j (\mathbf{e}_i) = A^i_j \delta^j_i = A^i_i \tag{14.4}$$

which is the trace!

14.2 The Symmetric Algebra

14.2.1 The Symmetric Product

Given a pure k-tensor $\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k}$, we define its symmetrization to be

$$Sym(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k}) = \frac{1}{k!} \sum_{\pi \in S_k} \mathbf{e}_{i_{\pi 1}} \otimes \cdots \otimes \mathbf{e}_{i_{\pi n}} \tag{14.5}$$

This has the porperty that

$$Sym\,(Sym\,(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k})) \;=\; Sym\,(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k}) \tag{14.6}$$

It also has the property that if two pure tensors $\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k}$ and $\mathbf{e}_{j_1}\otimes\cdots\otimes\mathbf{e}_{j_k}$ are re-arrangements of each other, their symmetrizations are identical.

We can extend *Sym* to the entire tensor algebra by linearity.

14.2.2 Example

Let's compute an example. Let $V = span\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and let $T = \mathbf{e}_1\otimes\mathbf{e}_2$. Then, using $i_1 = 1$ and $i_2 = 2$, we have $T = \mathbf{e}_{i_1}\otimes\mathbf{e}_{i_2}$. There are exactly two elements in the symmetric group S_2. Thus we have

$$\begin{aligned} Sym\,(\mathbf{e}_1\otimes\mathbf{e}_2) \;&=\; \frac{1}{2!}\sum_{\pi\in S_2}\mathbf{e}_{i_{\pi 1}}\otimes\mathbf{e}_{i_{\pi 2}} \\ &=\; \frac{1}{2}\,(\mathbf{e}_{i_1}\otimes\mathbf{e}_{i_2} \,+\, \mathbf{e}_{i_2}\otimes\mathbf{e}_{i_1}) \\ &=\; \frac{1}{2}\,(\mathbf{e}_1\otimes\mathbf{e}_2 \,+\, \mathbf{e}_2\otimes\mathbf{e}_1) \end{aligned} \tag{14.7}$$

Let's try another. Consider $T = \mathbf{e}_1\otimes\mathbf{e}_1$ so that $i_1 = i_2 = 1$. We have

$$\begin{aligned} Sym\,(\mathbf{e}_1\otimes\mathbf{e}_1) \;&=\; \frac{1}{2!}\sum_{\pi\in S_2}\mathbf{e}_{i_{\pi 1}}\otimes\mathbf{e}_{i_{\pi 2}} \\ &=\; \frac{1}{2}\,(\mathbf{e}_1\otimes\mathbf{e}_1 \,+\, \mathbf{e}_1\otimes\mathbf{e}_1) \\ &=\; \mathbf{e}_1\otimes\mathbf{e}_1 \end{aligned} \tag{14.8}$$

Let's try one more. This times set $T = \mathbf{e}_1\otimes\mathbf{e}_2\otimes\mathbf{e}_3$. Now we are dealing with the symmetry group S_3 which has 6 elements. Setting $i_1 = 1$, $i_2 = 2$, $i_3 = 3$ we get

$$\begin{aligned} Sym\,(\mathbf{e}_1\otimes\mathbf{e}_2\otimes\mathbf{e}_3) \;&=\; \frac{1}{3!}\sum_{\pi\in S_3}\mathbf{e}_{i_{\pi 1}}\otimes\mathbf{e}_{i_{\pi 2}}\otimes\mathbf{e}_{i_{\pi 3}} \\ &=\; \frac{1}{6}\,(\mathbf{e}_1\otimes\mathbf{e}_2\otimes\mathbf{e}_3 + \mathbf{e}_1\otimes\mathbf{e}_3\otimes\mathbf{e}_2 + \mathbf{e}_2\otimes\mathbf{e}_1\otimes\mathbf{e}_3 \\ &\quad +\mathbf{e}_2\otimes\mathbf{e}_3\otimes\mathbf{e}_1 + \mathbf{e}_3\otimes\mathbf{e}_1\otimes\mathbf{e}_2 + \mathbf{e}_3\otimes\mathbf{e}_2\otimes\mathbf{e}_1) \end{aligned} \tag{14.9}$$

14.2.3 The Symmetric Product

If T and S are aymmetric tensors of degree k and l, respectively, then we define the *symmetric product* to be

$$T \odot S \triangleq \frac{(k+l)!}{k!l!} Sym\,(T \otimes S) \tag{14.10}$$

This has the property that

$$Sym(T \odot S) = T \odot S. \tag{14.11}$$

We can iterate the symmetric product, for instance defining $\mathbf{v} \odot \mathbf{w} \odot \mathbf{z} = (\mathbf{v} \odot \mathbf{w}) \odot \mathbf{z}$. This product is associative.

This leads to the Symmetric Algebra:

$$\begin{aligned} \bigodot^* V &= \mathbb{R} \oplus V \oplus V^{\odot 2} \oplus V^{\odot 3} \odot \ldots \\ &= \bigoplus_{k=1}^{\infty} \bigodot^k V \end{aligned} \tag{14.12}$$

14.2.4 Examples

Let's compute some examples.

$$\begin{aligned} \mathbf{e}_1 \odot \mathbf{e}_1 &= \frac{(1+1)!}{1!1!} Sym\,(\mathbf{e}_1 \otimes \mathbf{e}_1) \\ &= 2! \frac{1}{2!} \sum_{\pi \in S_2} \mathbf{e}_1 \otimes \mathbf{e}_1 \\ &= \mathbf{e}_1 \otimes \mathbf{e}_1. \end{aligned} \tag{14.13}$$

Another: with $i_1 = 1$, $i_2 = 1$, $i_3 = 2$ we have

$$\begin{aligned} \mathbf{e}_1 \odot \mathbf{e}_1 \odot \mathbf{e}_2 &= (\mathbf{e}_1 \odot \mathbf{e}_1) \odot \mathbf{e}_2 \\ &= \frac{3!}{2!1!} Sym\,((\mathbf{e}_1 \otimes \mathbf{e}_1) \otimes \mathbf{e}_2) \\ &= \frac{3!}{2!1!} \frac{1}{3!} \sum_{\pi \in S_3} \mathbf{e}_{i_{\pi 1}} \otimes \mathbf{e}_{i_{\pi 2}} \otimes \mathbf{e}_{i_{\pi 3}} \\ &= \frac{1}{2} (\mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2 \\ &\quad + \mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1) \\ &= \mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1 \end{aligned} \tag{14.14}$$

14.3 The Exterior Product

Given a pure k-tensor $\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k}$, we define its antisymmetrization to be

$$Alt\left(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k}\right) \;=\; \frac{1}{k!}\sum_{\pi \in S_k} (-1)^{\pi} \mathbf{e}_{i_{\pi 1}} \otimes \cdots \otimes \mathbf{e}_{i_{\pi n}} \tag{14.15}$$

This has the porperty that

$$Alt\left(Alt\left(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k}\right)\right) \;=\; Alt\left(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k}\right) \tag{14.16}$$

We can extend Alt to the entire tensor algebra by linearity.

This has the property that if $\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_n}$ has any repeated index, then its alternation is zero.

Assuming V is an n-dimensional vector space, then this forces $Alt(T) = 0$ whenever T is a tensor of order greater than n.

14.3.1 The exterior product and the exterior algebra

Assuming T and S are alternating tensors of degrees k and l, respectively, the define the *exterior product* or the *wedge product* of T and S to be

$$T \wedge S \;=\; \frac{(k+l)!}{k!l!} Alt(T \otimes S). \tag{14.17}$$

This leads to the exterior algebra:

$$\begin{aligned} \bigwedge^{*} V \;&=\; \mathbb{R} \oplus V \oplus \bigwedge^{2} V \oplus \bigwedge^{3} V \oplus \\ &=\; \bigoplus_{k=0}^{\infty} \bigwedge^{k} V \end{aligned} \tag{14.18}$$

14.3.2 Examples

We compute

$$\begin{aligned} \mathbf{e}_1 \wedge \mathbf{e}_2 \;&=\; \frac{(1+1)!}{1!1!} Alt\left(\mathbf{e}_1 \otimes \mathbf{e}_2\right) \\ &=\; 2!\frac{1}{2!}\left(\mathbf{e}_1 \otimes \mathbf{e}_2 \,-\, \mathbf{e}_2 \otimes \mathbf{e}_1\right) \\ &=\; \mathbf{e}_1 \otimes \mathbf{e}_2 \,-\, \mathbf{e}_2 \otimes \mathbf{e}_1. \end{aligned} \tag{14.19}$$

We compute

$$
\begin{aligned}
\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3 &= (\mathbf{e}_1 \wedge \mathbf{e}_2) \wedge \mathbf{e}_3 \\
&= \frac{(2+1)!}{2!1!} Alt\left((\mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1) \otimes \mathbf{e}_3\right) \\
&= \frac{3!}{2!1!}\frac{1}{3!}\left(Alt\left(\mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_3\right) - Alt\left(\mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_3\right)\right) \\
&= \frac{3!}{2!1!}\left(Alt\left(\mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_3\right) - Alt\left(\mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_3\right)\right) \\
&= (\mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_3 - \mathbf{e}_1 \otimes \mathbf{e}_3 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_3 \\
&\quad + \mathbf{e}_2 \otimes \mathbf{e}_3 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_3 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1)
\end{aligned} \tag{14.20}
$$

14.4 Inner Product Spaces

14.4.1 Inner Products

A vector space with a bilinear map $g : V \otimes V \to \mathbb{R}$ is called an *inner product space* provided g satisfies

i) (Positivity) $g(\mathbf{v}, \mathbf{v}) \geq 0$ with equality if and only if $\mathbf{v} = o$.

ii) (Symmetry) $g(\mathbf{v}, \mathbf{w}) = g(\mathbf{w}, \mathbf{v})$

iii) (Bilinearity) $g(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2, \mathbf{w}) = \alpha g(\mathbf{v}_1, \mathbf{w}) + \beta g(\mathbf{v}_2, \mathbf{w})$ and $g(\mathbf{v}, \alpha\mathbf{w}_1 + \beta\mathbf{w}_2) = \alpha g(\mathbf{v}, \mathbf{w}_1) + \beta g(\mathbf{v}, \mathbf{w}_2)$

Then $g(\cdot, \cdot)$ is called an inner product. By (*iii*) we know that $g \in V^* \otimes V^*$. By (*ii*) we know that, in addition, $g \in V^* \odot V^*$.

14.4.2 The Musical Isomorphisms

If V is an inner product space, then there is a canonical isomorphism $V \to V^*$ which is denoted with a $\flat$. If $\mathbf{v} \in V$ we define

$$\mathbf{v} \longmapsto g(\mathbf{v}, \cdot) = \mathbf{v}_\flat \in V^* \tag{14.21}$$

This is a linear funcional on V, which we denote by $\mathbf{v}_\flat$.

Conversely, the $\sharp$ map is the inverse $V^* \to V$. If $\eta \in V^*$, this is defined implicitly by

$$
\begin{aligned}
&\eta \longmapsto \eta^\sharp \in V \\
&\eta(\mathbf{w}) = \left\langle \eta^\sharp, \mathbf{w} \right\rangle .
\end{aligned} \tag{14.22}
$$

14.4.3 The interior product

The interior product is a kind of converse of the exterior product (aka the wedge product). Consiider the case that $T \in \bigwedge^k V^*$ and $v \in V$. We have the exterior product:

$$\begin{aligned} v \,:\, \bigwedge^k V^* &\longrightarrow \bigwedge^{k+1} V^* \\ T &\longmapsto v_\flat \wedge T. \end{aligned} \tag{14.23}$$

We also have the interior product:

$$\begin{aligned} v \,:\, \bigwedge^k V^* &\longrightarrow \bigwedge^{k-1} V^* \\ T(\cdot, \cdot, \ldots, \cdot) &\longmapsto T(v, \cdot, \ldots, \cdot). \end{aligned} \tag{14.24}$$

Let's do an example. Consider the 3-dimensional vector space V with the inner product g. We can choose and orthonormal basis

$$\mathcal{B} \;=\; \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \tag{14.25}$$

Let T be the alternating tensor

$$T \;=\; \mathbf{e}^1 \wedge \mathbf{e}^2 \;\in\; \bigwedge^2 V^*. \tag{14.26}$$

We make three computations: $i_{\mathbf{e}_1}T$, $i_{\mathbf{e}_2}T$, and $i_{\mathbf{e}_3}T$. We get

$$\begin{aligned} i_{\mathbf{e}_1}T &= T(\mathbf{e}_1, \cdot) \;=\; (\mathbf{e}^1 \wedge \mathbf{e}^2)(\mathbf{e}_1) \\ &= (\mathbf{e}^1 \otimes \mathbf{e}^2 - \mathbf{e}^2 \otimes \mathbf{e}^1)(\mathbf{e}_1, \cdot) \\ &= \mathbf{e}^1(\mathbf{e}_1)\mathbf{e}^2 - \mathbf{e}^2(\mathbf{e}_1)\mathbf{e}^1 \\ &= \mathbf{e}^2 \end{aligned} \tag{14.27}$$

$$\begin{aligned} i_{\mathbf{e}_2}T &= T(\mathbf{e}_2, \cdot) \;=\; (\mathbf{e}^1 \wedge \mathbf{e}^2)(\mathbf{e}_2) \\ &= (\mathbf{e}^1 \otimes \mathbf{e}^2 - \mathbf{e}^2 \otimes \mathbf{e}^1)(\mathbf{e}_2, \cdot) \\ &= \mathbf{e}^1(\mathbf{e}_2)\mathbf{e}^2 - \mathbf{e}^2(\mathbf{e}_2)\mathbf{e}^1 \\ &= -\mathbf{e}^1 \end{aligned} \tag{14.28}$$

$$\begin{aligned} i_{\mathbf{e}_3}T &= T(\mathbf{e}_3, \cdot) \;=\; (\mathbf{e}^1 \wedge \mathbf{e}^2)(\mathbf{e}_3) \\ &= (\mathbf{e}^1 \otimes \mathbf{e}^2 - \mathbf{e}^2 \otimes \mathbf{e}^1)(\mathbf{e}_3, \cdot) \\ &= \mathbf{e}^1(\mathbf{e}_3)\mathbf{e}^2 - \mathbf{e}^2(\mathbf{e}_3)\mathbf{e}^1 \\ &= 0 \end{aligned} \tag{14.29}$$

14.5 Exercises

1) If V is a 3-dimensional space, show that every homogeneous form is pure. In 4-dimensions, show that there are some 2-forms that are not pure.

*2) We consider the case of a Lorentzian inner product on 3-space. To be specific, given a basis $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of V, let g be the symmetric bilinear operator given by $g(\mathbf{e}_1, \mathbf{e}_1) = 1$, $g(\mathbf{e}_2, \mathbf{e}_2) = -1$, $g(\mathbf{e}_3, \mathbf{e}_3) = -1$, and $g(\mathbf{e}_i, \mathbf{e}_j) = 0$ when $i \neq j$.

a) We know that $g \in V^* \otimes V^*$. Express g in tensor form.

b) Show that some vectors are self-orthogonal.

c) Show that the inner product g still indices a $\flat$-map, $\flat : V \to V^*$, that is invertible.

d) If g is *any* bilinear form (that is $g \in V^* \odot V^*$), show that $g \in Hom(V, V^*)$. Under what conditions, specifically, is g not an isomporphism?

3) Let $T = T_i{}^j \mathbf{e}^i \otimes \mathbf{e}_j$ be a $(1,1)$-tensor.

a) If $\mathcal{C} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ is another basis with transitions $\mathbf{e}_i = A_i^j \mathbf{f}_j$, express the tensor T in the new basis.

To be specific, we already have numbers T_i^j where $T = T_i^j \mathbf{e}^i \otimes \mathbf{e}_j$, but the same tensor T can be expressed with a different array of numbers $\widetilde{T}_i^j$ where $T = \widetilde{T}_i^j \mathbf{f}^i \otimes \mathbf{f}_j$. The challenge here is to express the numbers $\widetilde{T}_i^j$ in terms of the numbers T_i^j and the various transition matrices.

b) Show that the contraction of T is invariant with respect to change of basis. That is, whichever basis you compute the contraction in, you get the same number.

Lecture 15 - Raising and Lowering Indices

Tuesday April 1, 2014

15.1 Contraction

Recall the contraction: this is a way of taking a (r, s) tensor and producing a $(r-1, s-1)$ tensor (as long as $r, s > 1$). For example if $T = T_{ij}{}^{k}\mathbf{e}^i \otimes \mathbf{e}^j \otimes \mathbf{e}_k$ is a $(2,1)$-tensor, there are two distinct ways to contract:

$$\begin{aligned}
\mathtt{C}T &= T_{ij}{}^{k}\mathbf{e}^i(\mathbf{e}_k) \otimes \mathbf{e}^j \\
&= T_{ij}{}^{k}\delta^i_k \mathbf{e}^j \\
&= T_{ij}{}^{i}\,\mathbf{e}^j \\
\mathtt{C}T &= T_{ij}{}^{k}\mathbf{e}^i\mathbf{e}^j(\mathbf{e}_k) \\
&= T_{ij}{}^{k}\delta^i_k \mathbf{e}^i \delta^j_k \\
&= T_{ij}{}^{j}\,\mathbf{e}^i
\end{aligned} \tag{15.1}$$

15.2 The metric

If V is a vector space of n dimensions and $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis, there is an automatic dual basis $\mathcal{B}^* = \{\mathbf{e}^1, \ldots, \mathbf{e}^n\}$. A *metric* is an element $g \in \bigodot^2 V^*$ that satisfies the property that $g(\mathbf{v}, \mathbf{v}) > 0$ unless $\mathbf{v} = o$. A metric is the same thing as an inner product. A metric can be expressed $g \;=\; g_{ij}\,\mathbf{e}^i \odot \mathbf{e}^j$.

Proposition 15.2.1 *If $g \in \bigodot^2 V^*$, then it can be expressed as*

$$g \;=\; g_{ij}\,\mathbf{e}^i \otimes \mathbf{e}^j \tag{15.2}$$

where $g_{ij} = g_{ji}$.

Proof. Any such g can of course be expressed

$$g \;=\; g_{ij}\,\mathbf{e}^i \odot \mathbf{e}^j \tag{15.3}$$

Then by the definition of the symmetric product, we have

$$\begin{aligned} g &= g_{ij}\,\mathbf{e}^i \odot \mathbf{e}^j \\ &= g_{ij}\,\left(\mathbf{e}^i \otimes \mathbf{e}^j + \mathbf{e}^j \otimes \mathbf{e}^i\right) \\ &= g_{ij}\,\mathbf{e}^i \otimes \mathbf{e}^j + g_{ij}\,\mathbf{e}^j \otimes \mathbf{e}^i \\ &= g_{kl}\,\mathbf{e}^k \otimes \mathbf{e}^l + g_{lk}\,\mathbf{e}^k \otimes \mathbf{e}^l \\ &= (g_{kl} + g_{lk})\,\mathbf{e}^k \otimes \mathbf{e}^l \end{aligned} \tag{15.4}$$

Therefore

$$g \;=\; \tilde{g}_{kl}\,\mathbf{e}^k \otimes \mathbf{e}^l \tag{15.5}$$

where $\tilde{g}_{kl} = g_{ij} + g_{ji}$ obviously satisfies $\tilde{g}_{kl} = \tilde{g}_{lk}$. □

Recall that if V has a metric g, then there is a canonical map $\flat : V \to V^*$ defined by

$$\mathbf{v}_\flat \;=\; g(\mathbf{v},\,\cdot). \tag{15.6}$$

This map is invertible, and its inverse is denoted $\sharp : V^* \to V$. Recall that if g is a metric on V, then there is an automatic metric on V^*, defined by

$$g(\eta,\,\gamma) \;\triangleq\; g(\eta^\sharp,\,\gamma^\sharp) \qquad \text{where} \qquad \eta, \gamma \in V^*. \tag{15.7}$$

Theorem 15.2.2 (Characterization of the metric on the dual space) *If $g = g_{ij}\mathbf{e}^i \otimes \mathbf{e}^j \in V^* \odot V^*$ is a metric on v, then the corresponding metric on V^* is an element*

$$g \in V \odot V \tag{15.8}$$

where $g = g^{ij}\mathbf{e}_i \otimes \mathbf{e}_j$ has the property that g^{ij} is the inverse matrix of g_{ij}. Specifically, $g_{ik}g^{kj} = \delta_i^j$.

Proof. We *define* the symbols g^{ij} so that the matrix (g^{ij}) is the inverse of the matrix (g_{ij}). Given a basis covector $\mathbf{e}^i$, let's find out how to express $(\mathbf{e}^i)^\sharp$ in terms of the basis vectors $\mathbf{e}_i$. Our job is to determine the coefficients A^{ij} of

$$(\mathbf{e}^i)^\sharp \;=\; A^{ij}\mathbf{e}_j. \tag{15.9}$$

The implicit definition of $(\mathbf{e}^i)^\sharp$ is that

$$g((\mathbf{e}^i)^\sharp,\,\cdot\,) \;=\; g\left(A^{ik}\mathbf{e}_k,\,\cdot\right) \;=\; A^{ik}g(\mathbf{e}_k,\,\cdot) \tag{15.10}$$

Then using $g = g_{ij}\mathbf{e}^i \otimes \mathbf{e}^j$ we get

$$\begin{aligned}
\mathbf{e}^i &= g((\mathbf{e}^i)^\sharp, \cdot) \\
&= A^{ik}\left(g_{st}\mathbf{e}^s \otimes \mathbf{e}^t\right)(\mathbf{e}_k, \cdot) \\
&= A^{ik}\, g_{st}\, \mathbf{e}^s(\mathbf{e}_k)\mathbf{e}^t \\
&= A^{ik}\, g_{st}\, \delta^s_k\, \mathbf{e}^t \\
&= A^{ik}\, g_{kt}\, \mathbf{e}^t
\end{aligned} \tag{15.11}$$

This forces $A^{ik}g_{kt} = \delta^i_t$. Therefore A^{ik} in fact equals g^{ik}.

So we have proved that $(\mathbf{e}^i)^\sharp = g^{ij}\mathbf{e}_j$. Then using the definition of g on V^* we have

$$\begin{aligned}
g(\mathbf{e}^i, \mathbf{e}^j) &\triangleq g\left((\mathbf{e}^i)^\sharp, (\mathbf{e}^j)^\sharp\right) \\
&= g\left(g^{ik}\mathbf{e}_k, g^{jl}\mathbf{e}_l\right) \\
&= g^{ik}g^{jl}g\left(\mathbf{e}_k, \mathbf{e}_l\right) \\
&= g^{ik}g^{jl}g_{kl} \\
&= g^{ik}\delta^j_k \\
&= g^{ij}
\end{aligned} \tag{15.12}$$

15.3 The Musical isomorphisms: raising and lowering indices

We now have the following:

$$\begin{aligned}
(\mathbf{e}_i)_\flat &= g_{ij}\mathbf{e}^j \\
(\mathbf{e}^i)^\sharp &= g^{ij}\mathbf{e}_j.
\end{aligned} \tag{15.13}$$

Now consider a vector $\mathbf{v} = v^i\mathbf{e}_j$. We have that

$$\mathbf{v}_\flat = v^i(\mathbf{e}_i)_\flat = v^i g_{ij}\mathbf{e}^j. \tag{15.14}$$

So if the coefficients of $\mathbf{v}$ are v^i then the coefficients of $\mathbf{v}_\flat$ are $v^i g_{ij}$ This conversion of a vector to a covector is called the lowering of indices.

Conversely if $\eta = \eta_i\mathbf{e}^i$ is a covector, we have that

$$\eta^\sharp = (\eta_i\mathbf{e}^i)^\sharp = \eta_i(\mathbf{e}^i)^\sharp = \eta_i g^{ij}\mathbf{e}_j. \tag{15.15}$$

So if the coefficients of η are η_i then the coefficients of $\eta^\sharp$ are $\eta_i g^{ij}$ This conversion of a covector to a vector is called the raising of indices.

The raising and lowering of indices extends to tensors of any order. Consider, for instance, the tensor $(2,1)$-tensor

$$T \ = \ T_{ij}{}^{k}\, \mathbf{e}^i \otimes \mathbf{e}^j \otimes \mathbf{e}_k \tag{15.16}$$

This can be converted to a $(3,0)$ tensor by lowering of an index:

$$T \ = \ T_{ijk} \mathbf{e}^i \otimes \mathbf{e}^j \otimes \mathbf{e}^k \tag{15.17}$$

where the symbols T^{ijk} are defined as such:

$$T_{ijk} \ = \ T_{ij}{}^{s} g_{sk}. \tag{15.18}$$

There are two ways to convert this to a $(1,2)$ tensor

$$\begin{aligned} T \ &= \ T_i{}^{jk}\, \mathbf{e}^i \otimes \mathbf{e}_j \otimes \mathbf{e}^k \\ T \ &= \ T^i{}_j{}^k\, \mathbf{e}_i \otimes \mathbf{e}^j \otimes \mathbf{e}_k. \end{aligned} \tag{15.19}$$

where

$$T_i{}^{jk} \ \triangleq \ T_{is}{}^{k} g^{sj} \quad \text{and} \quad T^i{}_j{}^k \ \triangleq \ T_{sj}{}^{k} g^{si}. \tag{15.20}$$

15.4 Tracing

The trace is the combination of raising (or lowering) followed by contraction. For example, if $T = T_{ij}\mathbf{e}^i \otimes \mathbf{e}^j$ is some $(2,0)$-tensor, then to take its trace we raise an index

$$T_{ij}\mathbf{e}^i \otimes \mathbf{e}^j \quad \longrightarrow \quad T_i{}^j \mathbf{e}^i \otimes \mathbf{e}_j \tag{15.21}$$

followed by a contraction:

$$T_{ij}\mathbf{e}^i \otimes \mathbf{e}^j \quad \longrightarrow \quad T_i{}^j \mathbf{e}^i \otimes \mathbf{e}_j \quad \longrightarrow \quad T_i{}^j \mathbf{e}^i(\mathbf{e}_j) \ = \ T_i{}^i \tag{15.22}$$

Because $T_i{}^j = T_{is}g^{sj}$ we have that the trace is

$$T_{ij}g^{ij}. \tag{15.23}$$

Likewise if

$$T \ = \ T^{ij}\mathbf{e}_i \otimes \mathbf{e}_j \tag{15.24}$$

is a $(0,2)$-tensor then its trace is

$$T^{ij}g_{ij}. \tag{15.25}$$

15.5 The interior product

15.5.1 First Definition

If η is a k-form and $v \in V$, meaning

$$\eta \in \bigwedge^k V^*, \qquad v \in V \tag{15.26}$$

then we define the interior product by setting

$$i_v \eta \;=\; \eta(v, \cdot, \ldots, \cdot) \;\in\; \bigwedge^{k-1} V^* \tag{15.27}$$

For example if $\eta = \mathbf{e}^1 \bigwedge \mathbf{e}^2$ then we have

$$\begin{aligned} i_{\mathbf{v}}(\mathbf{e}^1 \wedge \mathbf{e}^2) \;&=\; (\mathbf{e}^1 \wedge \mathbf{e}^2)(\mathbf{v}, \cdot) \\ &=\; (\mathbf{e}^1 \otimes \mathbf{e}^2 \;-\; \mathbf{e}^2 \otimes \mathbf{e}^1)(\mathbf{v}, \cdot) \\ &=\; \mathbf{e}^1(\mathbf{v})\,\mathbf{e}^2 \;-\; \mathbf{e}^2(\mathbf{v})\mathbf{e}^1 \end{aligned} \tag{15.28}$$

We have the following:

Theorem 15.5.1 *If $\eta \in \bigwedge^k V^*$ and $\gamma \in \bigwedge^l V^*$ and $\mathbf{v} \in V$ then we have the following Leibniz rule*

$$i_{\mathbf{v}}(\eta \wedge \gamma) \;=\; (i_{\mathbf{v}}\eta) \wedge \gamma \;+\; (-1)^k \eta \wedge (i_{\mathbf{v}}\gamma)\,. \tag{15.29}$$

15.5.2 Second definition

There is an axiomatic way to define the interior product

(Constant Rule) If $a \in \bigwedge^0 V^* = \mathbb{R}$ is a constant, then $i_{\mathbf{v}} a = 0$.

(Sum and Difference Rule) If $\eta, \gamma \in \bigwedge^* V^*$ then $i_{\mathbf{v}}\,(\eta \pm \gamma) = i_{\mathbf{v}}\eta \pm i_{\mathbf{v}}\gamma$

(Action on Covectors) If $\eta \in \bigwedge^1 V^* = V^*$ is a covector, then $i_{\mathbf{v}}\eta = \eta(\mathbf{v})$

(Leibniz Rule) If $\eta \in \bigwedge^k V^*$ and $\gamma \in \bigwedge^k V^*$ and $i_{\mathbf{v}}(\eta \wedge \gamma) = (i_{\mathbf{v}}\eta) \wedge \gamma + (-1)^k \eta \wedge (i_{\mathbf{v}}\gamma)$

Using these rules, we can compute the action of $i_{\mathbf{v}}$ on any form. For instance

$$\begin{aligned} i_{\mathbf{e}_1}\left(\mathbf{e}^2 \wedge \mathbf{e}^1 \wedge \mathbf{e}^3\right) \;&=\; (i_{\mathbf{e}_1}\mathbf{e}^2) \wedge \mathbf{e}^1 \wedge \mathbf{e}^3 \;-\; \mathbf{e}^2 \wedge (i_{\mathbf{e}_1}\left(\mathbf{e}^1 \wedge \mathbf{e}^3\right) && \text{Leibniz Rule} \\ &=\; (i_{\mathbf{e}_1}\mathbf{e}^2) \wedge \mathbf{e}^1 \wedge \mathbf{e}^3 \;-\; \mathbf{e}^2 \wedge (i_{\mathbf{e}_1}\mathbf{e}^1) \wedge \mathbf{e}^3 \;+\; \mathbf{e}^2 \wedge \mathbf{e}^1 \wedge (i_{\mathbf{e}_1}\mathbf{e}^3) && \text{Leibniz Rule again} \\ &=\; \mathbf{e}^2(\mathbf{e}_1)\mathbf{e}^1 \wedge \mathbf{e}^3 \;-\; \mathbf{e}^1(\mathbf{e}_1)\mathbf{e}^2 \wedge \mathbf{e}^3 \;+\; \mathbf{e}^3(\mathbf{e}_1)\mathbf{e}^2 \wedge \mathbf{e}^1 && \text{Evaluation on covectors} \\ &=\; -\mathbf{e}^2 \wedge \mathbf{e}^3 \end{aligned} \tag{15.30}$$

15.6 Exercises

1) Consider $\mathbb{R}^2$ with its standard metric; we know that an orthonormal basis is $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ with corresponding dual basis is $\{(1,0),(0,1)\} \subset (\mathbb{R}^2)^*$. Now consider a new basis $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2\}$ where

$$\mathbf{e}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \qquad \mathbf{e}_2 = \begin{pmatrix} -2 \\ 2 \end{pmatrix} \tag{15.31}$$

a) Determine the corresponding dual basis $\mathcal{B}^* = \{\mathbf{e}^1, \mathbf{e}^2\}$

b) Express g in terms of the basis $\mathcal{B}^*$. Verify explicitly that $g_{ij} = g_{ji}$.

c) Express the metric on $(\mathbb{R}^2)^*$ in terms of the basis $\mathcal{B}$.

d) Compute the conjugate basis. Graph both the basis and the conjugate basis for $\mathbb{R}^2$ on the same set of coordinate axes. Be sure to clearly label your graph.

2) The moral of this problem is that all linear operations are tensor products and traces.

a) Let A be a (1,1)-tensor

$$A \;=\; A_j{}^i \, \mathbf{e}^j \otimes \mathbf{e}_i \tag{15.32}$$

If $\mathbf{v}$ is a vector, we know that $A(\mathbf{v})$ is also a vector. Show that $A(\mathbf{v})$ is equal to the tensor product $A \otimes \mathbf{v}$ followed by a contraction.

b) Let $A = A_j{}^i \, \mathbf{e}^j \otimes \mathbf{e}_i$ and $B = B_j{}^i \, \mathbf{e}^j \otimes \mathbf{e}_i$ be (1,1)-tensors. Show that the matrix product between A and B is the tensor product $A \otimes B$ (a (2,2)-tensor) followed by a contraction (resulting in a (1,1)-tensor).

3) Let V be a 2-dimensional vector space, and let $T_{ij}{}^k$ be the $(2,1)$ tensor given by

$$T_{1j}{}^k \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T_{2j}{}^k \;=\; \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}. \tag{15.33}$$

a) Express T explicitly in terms of the $\mathbf{e}_i$ and $\mathbf{e}^i$.

b) Determine, explicitly, the two traces of T. In particular show that the two covectors you obtain are not equal.

4) Let V be an n-dimensional vector space. Show that the trace of $g = g_{ij}\mathbf{e}^i \otimes \mathbf{e}^j$ is n.

5) Show that, if η is any k-form then $i_{\mathbf{v}} i_{\mathbf{v}} \eta = 0$. Do this with an induction argument: show it is true when $\eta \in \bigwedge^1$, and then assuming this is true for all $\eta' \in \bigwedge^k V^*$, show it is true whenever $\eta \in \bigwedge^{k+1} V^*$ (you will have to use the Leibniz rule).

*6) If η is any k-form and $\mathbf{v} \in V$, then show that

$$i_{\mathbf{v}}(\mathbf{v}_\flat \wedge \eta) \;+\; \mathbf{v}_\flat \wedge (i_{\mathbf{v}}\eta) \;=\; |\mathbf{v}|^2 \eta. \tag{15.34}$$

(Hint: rely on the properties of $i_{\mathbf{v}}$, and this is a one-line proof!)

Lecture 16 - The Hodge Duality Operator

Thursday April 3, 2014

16.1 Inner products on tensor spaces

16.1.1 Heuristics

If T and S are tensors of the same degree, we can define an inner product. For instance if

$$T \;=\; T_{ij}{}^{k}\mathbf{e}^i \otimes \mathbf{e}^j \otimes \mathbf{e}_k \qquad S \;=\; S_{ij}{}^{k}\mathbf{e}^i \otimes \mathbf{e}^j \otimes \mathbf{e}_k \tag{16.1}$$

then

$$\langle T,\, S\rangle \;=\; T_{ij}{}^{k} S_{lm}{}^{n}\, g^{il} g^{jm} g_{kn} \tag{16.2}$$

Let's see this in an example. Let $T = \mathbf{e}^1 \otimes \mathbf{e}^2$ and $S = \mathbf{e}^1 \otimes \mathbf{e}^1$. Then $T_{12} = 1$ and $S_{11} = 0$ and all other symbols are zero. Therefore

$$\begin{aligned}
\langle T,\, S\rangle \;&=\; T_{ij} S_{kl} g^{ik} g^{jl} \\
&=\; T_{12} S_{kl} g^{1k} g^{2l} \qquad \text{because } T_{ij} = 0 \text{ unless } i = 1, j = 2 \\
&=\; T_{12} S_{11} g^{11} g^{21} \qquad \text{because } S_{ij} = 0 \text{ unless } i = 1, j = 1 \\
&=\; g^{11} g^{21} \qquad \text{because } T_{12} = 1 \text{ and } S_{11} = 0 \\
&=\; \left\langle \mathbf{e}^1, \mathbf{e}^1 \right\rangle \left\langle \mathbf{e}^2, \mathbf{e}^1 \right\rangle .
\end{aligned} \tag{16.3}$$

Thus we have found that

$$\left\langle \mathbf{e}^1 \otimes \mathbf{e}^2,\, \mathbf{e}^1 \otimes \mathbf{e}^1 \right\rangle \;=\; \left\langle \mathbf{e}^1, \mathbf{e}^1 \right\rangle \left\langle \mathbf{e}^2, \mathbf{e}^1 \right\rangle . \tag{16.4}$$

The same sort of computation would give, for arbitrary fixed i, j, k, l, the expression

$$\left\langle \mathbf{e}^i \otimes \mathbf{e}^j,\, \mathbf{e}^k \otimes \mathbf{e}^l \right\rangle \;=\; \left\langle \mathbf{e}^i,\, \mathbf{e}^l \right\rangle \left\langle \mathbf{e}^j,\, \mathbf{e}^l \right\rangle . \tag{16.5}$$

16.1.2 The inner product on the bigraded tensor algebra

If

$$\begin{aligned} T &= T^{i_1\dots i_l}_{j_1\dots j_k}\mathbf{e}^{j_1}\otimes\dots\mathbf{e}^{j_k}\otimes\mathbf{e}_{i_1}\otimes\mathbf{e}_{i_l} \\ S &= S^{i_1\dots i_l}_{j_1\dots j_k}\mathbf{e}^{j_1}\otimes\dots\mathbf{e}^{j_k}\otimes\mathbf{e}_{i_1}\otimes\mathbf{e}_{i_l} \end{aligned} \tag{16.6}$$

are (k,l)-tensors, then their inner product is defined to be

$$\langle T,\, S\rangle \;=\; T^{i_1\dots i_l}_{j_1\dots j_k}S^{m_1\dots m_l}_{n_1\dots n_k}\, g^{j_1 n_1}\dots g^{j_l n_l} g_{i_1 m_1}\dots g_{i_l m_l}. \tag{16.7}$$

16.1.3 Second Definition of the inner product on the bigraded tensor algebra

The equation (16.5) leads us to the second

Definition. Let $\{\mathbf{v}_1,\dots,\mathbf{v}_l\},\{\mathbf{w}_1,\dots,\mathbf{w}_l\}$ be vectors and let $\{\eta^1,\dots,\eta^k\},\{\gamma^1,\dots,\gamma^k\}$ be covectors (no claim is made that these are bases). Then the inner product of two monomials of bidegree (k,l) is defined by

$$\begin{aligned} &g\left(\eta^1\otimes\dots\otimes\eta^k\otimes\mathbf{v}_1\otimes\dots\otimes\mathbf{v}_l,\,\gamma^1\otimes\dots\otimes\gamma^k\otimes\mathbf{w}^1\otimes\dots\otimes\mathbf{w}^l\right) \\ &\quad= g\left(\eta^1,\gamma^1\right)\dots g\left(\eta^k,\gamma^k\right) g\left(\mathbf{v}_1,\mathbf{w}_1\right)\dots g\left(\mathbf{v}_l,\mathbf{w}_l\right). \end{aligned} \tag{16.8}$$

We extend this inner product, by bilinearity, to the entire tensor algebra.

16.2 Inner Products on the Exterior Algebra

16.2.1 Diret extension to the exterior algebra

Forms are just examples of tensors, so the inner product explained above carries over. For example, if $\mathbf{v}^i,\mathbf{w}^i$ are various covectors (no claim is made that this is a basis), then using the definitions above we get

$$\begin{aligned} &g\left(\mathbf{v}^i\wedge\mathbf{v}^j,\,\mathbf{w}^k\wedge\mathbf{w}^l\right) \\ &\quad= g\left(\mathbf{v}^i\otimes\mathbf{v}^j-\mathbf{v}^j\otimes\mathbf{v}^i,\,\mathbf{w}^k\otimes\mathbf{w}^l-\mathbf{w}^l\otimes\mathbf{w}^k\right) \\ &\quad= g\left(\mathbf{v}^i\otimes\mathbf{v}^j,\,\mathbf{w}^k\otimes\mathbf{w}^l\right)-g\left(\mathbf{v}^i\otimes\mathbf{v}^j,\,\mathbf{w}^l\otimes\mathbf{w}^k\right)-g\left(\mathbf{v}^j\otimes\mathbf{v}^i,\,\mathbf{w}^k\otimes\mathbf{w}^l\right)+g\left(\mathbf{v}^j\otimes\mathbf{v}^i,\,\mathbf{w}^l\otimes\mathbf{w}^k\right) \\ &\quad= g(\mathbf{v}^i,\mathbf{w}^k)g(\mathbf{v}^j,\mathbf{w}^l)-g(\mathbf{v}^i,\mathbf{w}^l)g(\mathbf{v}^j,\mathbf{w}^k)-g(\mathbf{v}^j,\mathbf{w}^k)g(\mathbf{v}^i,\mathbf{w}^l)+g(\mathbf{v}^j,\mathbf{w}^l)g(\mathbf{v}^i,\mathbf{w}^k) \\ &\quad= 2\left(\left\langle\mathbf{v}^i,\mathbf{w}^k\right\rangle\left\langle\mathbf{v}^j,\mathbf{w}^l\right\rangle-\left\langle\mathbf{v}^i,\mathbf{w}^l\right\rangle\left\langle\mathbf{v}^j,\mathbf{w}^k\right\rangle\right) \\ &\quad= 2\det\begin{pmatrix}\left\langle\mathbf{v}^i,\mathbf{w}^k\right\rangle & \left\langle\mathbf{v}^j,\mathbf{w}^k\right\rangle \\ \left\langle\mathbf{v}^i,\mathbf{w}^l\right\rangle & \left\langle\mathbf{v}^j,\mathbf{w}^l\right\rangle\end{pmatrix}. \end{aligned} \tag{16.9}$$

Similarly

$$g\left(\mathbf{v}^i \wedge \mathbf{v}^j \wedge \mathbf{v}^k,\, \mathbf{w}^l \wedge \mathbf{w}^m \wedge \mathbf{w}^n\right) \;=\; 6\det\begin{pmatrix} \left\langle \mathbf{v}^i,\, \mathbf{w}^l \right\rangle & \left\langle \mathbf{v}^j,\, \mathbf{w}^l \right\rangle & \left\langle \mathbf{v}^k,\, \mathbf{w}^l \right\rangle \\ \left\langle \mathbf{v}^i,\, \mathbf{w}^m \right\rangle & \left\langle \mathbf{v}^j,\, \mathbf{w}^m \right\rangle & \left\langle \mathbf{v}^k,\, \mathbf{w}^m \right\rangle \\ \left\langle \mathbf{v}^i,\, \mathbf{w}^n \right\rangle & \left\langle \mathbf{v}^j,\, \mathbf{w}^n \right\rangle & \left\langle \mathbf{v}^k,\, \mathbf{w}^n \right\rangle \end{pmatrix} \tag{16.10}$$

In general we have

$$g\left(\mathbf{v}^1 \wedge \cdots \wedge \mathbf{v}^k,\, \mathbf{w}^1 \wedge \cdots \wedge \mathbf{w}^k\right) \;=\; k!\,\det\begin{pmatrix} \left\langle \mathbf{v}^1,\, \mathbf{w}^1 \right\rangle & \dots & \left\langle \mathbf{v}^k,\, \mathbf{w}^1 \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle \mathbf{v}^1,\, \mathbf{w}^k \right\rangle & \dots & \left\langle \mathbf{v}^k,\, \mathbf{w}^k \right\rangle \end{pmatrix} \tag{16.11}$$

16.2.2 Modified Definition

The factors of $k!$ are annoying, so it is conventional to simply define inner products by the following
Definition.

$$\left\langle \mathbf{v}^1 \wedge \cdots \wedge \mathbf{v}^k,\, \mathbf{w}^1 \wedge \cdots \wedge \mathbf{w}^k \right\rangle \;=\; \det\begin{pmatrix} \left\langle \mathbf{v}^1,\, \mathbf{w}^1 \right\rangle & \dots & \left\langle \mathbf{v}^k,\, \mathbf{w}^1 \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle \mathbf{v}^1,\, \mathbf{w}^k \right\rangle & \dots & \left\langle \mathbf{v}^k,\, \mathbf{w}^k \right\rangle \end{pmatrix} \tag{16.12}$$

Unless otherwise stated, this is always the inner product that isused on exterior forms.

16.3 The Volume Form

16.3.1 Definition of the volume form

If V is a vector space with $\dim(V) = n$, then we know that

$$\dim\left(\bigwedge^n V^*\right) \;=\; 1. \tag{16.13}$$

Assume V has an inner product and that $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is an *orthonormal* basis, then $\mathcal{B}^* = \{\mathbf{e}^1, \ldots, \mathbf{e}^n\}$ is also an orthonormal basis, and we define the *volume form* to be

$$dVol \;=\; \mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^n \;\in\; \bigwedge^n V^*. \tag{16.14}$$

Notice that

$$|dVol|^2 \;=\; \det\left(Id_{n\times n}\right) \;=\; 1. \tag{16.15}$$

However this is not quite canonical. The choice of a volume form is determined only up to a sign. A choice of an element of $\bigwedge^n V^*$ of length 1 (of which there are two possible choices) is called a *choice of orientation*; a volume form is also sometimes called an *orientation*.

16.3.2 Why is it called a "volume form"?

Let $dVol = \mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^n$ be a volume form, where $\mathcal{B}^* = \{\mathbf{e}^1, \ldots, \mathbf{e}^n\}$ is an orthonormal basis of covectors. Consider a collection $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of n-many vectors (no claim is made that this is a basis). Then the action of $dVol$ on these n many vectors gives the signed n-volume of the parallelotope spanned by these vectors:

$$\begin{aligned} dVol\,(\mathbf{v}_1, \ldots, \mathbf{v}_n) \;&=\; \begin{pmatrix} \mathbf{e}^1(\mathbf{v}_1) & \mathbf{e}^1(\mathbf{v}_2) & \ldots & \mathbf{e}^1(\mathbf{v}_n) \\ \mathbf{e}^2(\mathbf{v}_1) & \mathbf{e}^2(\mathbf{v}_2) & & \mathbf{e}^2(\mathbf{v}_n) \\ \vdots & & \ddots & \vdots \\ \mathbf{e}^n(\mathbf{v}_1) & \mathbf{e}^n(\mathbf{v}_2) & \ldots & \mathbf{e}^n(\mathbf{v}_n) \end{pmatrix} \\ &=\; \text{signed volume of the parallotope spanned by } \{\mathbf{v}_1, \mathbf{v}_2\}. \end{aligned} \tag{16.16}$$

Once an *ordered* orthonormal basis is chosen, an orientation is automatic. If any two elements of the ordered basis are switched, then the orientation reverses sign.

16.3.3 The right-hand rule

In dimensions 2 and 3, the notion of orientation is already quite familiar. In particular, a choice of orientation is equivalent to a choice of using either the right-hand rule or the left-hand rule.

For instance, let $\mathbf{e}^1, \mathbf{e}^2$ be the standard orthonormal covectors, with volume form $dVol = \mathbf{e}^1 \wedge \mathbf{e}^2$. Let's pick two vectors, $\mathbf{v} = 2\mathbf{e}_1 + \mathbf{e}_2$ and $\mathbf{w} = \mathbf{e}_1 + 2\mathbf{e}_2$, for instance. In accordance with the right-hand rule, the ordering $(\mathbf{v}, \mathbf{w})$ should be the positive ordering and the ordering $(\mathbf{w}, \mathbf{v})$ should be the negative ordering. So let's compute:

$$\begin{aligned} dVol(\mathbf{v}, \mathbf{w}) \;&=\; \mathbf{e}^1(\mathbf{v})\mathbf{e}^2(\mathbf{w}) \;-\; \mathbf{e}^2(\mathbf{v})\mathbf{e}^1(\mathbf{w}) \\ &=\; 2 \cdot 2 \;-\; 1 \cdot 1 \;=\; 3 \\ dVol(\mathbf{w}, \mathbf{v}) \;&=\; \mathbf{e}^1(\mathbf{w})\mathbf{e}^2(\mathbf{v}) \;-\; \mathbf{e}^2(\mathbf{w})\mathbf{e}^1(\mathbf{v}) \\ &=\; 1 \cdot 1 \;-\; 2 \cdot 2 \;=\; -3 \end{aligned} \tag{16.17}$$

which is what we expected.

16.4 The Hodge Duality Operator

16.4.1 The volume form as a tensor

Recall the *Levi-Civita* symbol:

$$\varepsilon_{i_1 i_2 \ldots i_n} \;=\; \begin{cases} 0 & \text{if any two of the } i_j \text{ are equal to each other} \\ sgn(\pi) & \text{if all the } i_j \text{ are unequal, where } \pi \text{ is the permutation} \\ & \quad\quad \text{of } n \text{ letters that takes } (i_1, i_2, \ldots i_n) \text{ to } (1, 2, \ldots, n). \end{cases} \tag{16.18}$$

This is also called the *totally antisymmetric symbol.* We can use the Levi-Civita symbol ε to express $dVol$ in tensor form:

$$dVol \;=\; \varepsilon_{i_1 \ldots i_n} \mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_n}. \tag{16.19}$$

16.4.2 First definition of the Hodge star

The *Hodge star operator* is simply any of the tensoral forms of the volume form. For instance, if we raise one index, we get

$$\begin{aligned} * : \bigwedge^1 V \otimes \bigwedge^{n-1} V^* \;&\approx\; Hom\left(\bigwedge^1 V^*, \bigwedge^{n-1} V^* \right) \\ * \;&=\; g^{i_1 j_i} \varepsilon_{i_1 \ldots i_n} \, \mathbf{e}_{i_1} \otimes \mathbf{e}^{i_2} \otimes \cdots \otimes \mathbf{e}^{i_n}. \end{aligned} \tag{16.20}$$

If we raise two indices, we get

$$\begin{aligned} * : \bigwedge^2 V \otimes \bigwedge^{n-2} V^* \;&\approx\; Hom\left(\bigwedge^2 V^*, \bigwedge^{n-1} V^* \right) \\ * \;&=\; g^{i_1 j_i} g^{i_2 j_2} \varepsilon_{i_1 \ldots i_n} \, \mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \mathbf{e}^{i_2} \otimes \cdots \otimes \mathbf{e}^{i_n}. \end{aligned} \tag{16.21}$$

Raising $k \leq n$ many indices we obtain an operator

$$* \; : \; \bigwedge^k \; \longrightarrow \; \bigwedge^{n-k}. \tag{16.22}$$

16.4.3 Second definition of the Hodge star

An equivalent definition is as follows. If η is a pure k-form, then we define

$$*\eta \;=\; \text{the } (n-k)\text{-form so that } \eta \wedge *\eta = |\eta|^2 dVol \tag{16.23}$$

and extend by linearlity.

For instance consider $V \approx R^3$ with $\mathcal{B}^* = \{\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3\}$ an orthonormal basis and with volume form $dVol = \mathbf{e}^1 \wedge \mathbf{e}^2 \wedge \mathbf{e}^3$. Then we compute

$$\begin{aligned} *\mathbf{e}^1 &= \text{the 2-form so that } \mathbf{e}^1 \wedge *\mathbf{e}^1 = 1 \cdot dVol \\ &= \mathbf{e}^2 \wedge \mathbf{e}^3 \\ *\mathbf{e}^2 &= \text{the 2-form so that } \mathbf{e}^3 \wedge *\mathbf{e}^3 = 1 \cdot dVol \\ &= -\mathbf{e}^1 \wedge \mathbf{e}^3 \\ *\mathbf{e}^3 &= \text{the 2-form so that } \mathbf{e}^3 \wedge *\mathbf{e}^3 = 1 \cdot dVol \\ &= \mathbf{e}^1 \wedge \mathbf{e}^3 \end{aligned} \tag{16.24}$$

Similarly

$$\begin{aligned} *(\mathbf{e}^1 \wedge \mathbf{e}^2) &= \mathbf{e}^3 \\ *(\mathbf{e}^1 \wedge \mathbf{e}^3) &= -\mathbf{e}^2 \\ *(\mathbf{e}^2 \wedge \mathbf{e}^3) &= \mathbf{e}^1 \end{aligned} \tag{16.25}$$

Theorem 16.4.1 *The Hodge star* $* : \bigwedge^k \to \bigwedge^{n-k}$ *is an isomorphism. Further,* $** : \bigwedge^k \to \bigwedge^k$ *is given by* $** = (-1)^{k(n-k)} Id$.

Proof. The second assertion clearly follows from the first, as the inverse of $*$ is then $\pm *$. To compute $**$, we work with a basis.

Any pure element of length 1 in $\bigwedge^k$ can be expressed in the form

$$\mathbf{e}^1 \wedge \ldots \mathbf{e}^k \tag{16.26}$$

where $\{\mathbf{e}^1, \ldots, \mathbf{e}^k\}$ is an orthonormal set. We can complete this to an ordered orthonormal basis $\{\mathbf{e}^1, \ldots, \mathbf{e}^n\}$ of V^*. Then we compute

$$*\left(*\mathbf{e}^1 \wedge \ldots \mathbf{e}^k\right) = *\mathbf{e}^{k+1} \wedge \cdots \wedge \mathbf{e}^n \tag{16.27}$$

Second we compute

$$\begin{aligned} &\left(\mathbf{e}^{k+1} \wedge \cdots \wedge \mathbf{e}^n\right) \wedge \left(\mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^k\right) \\ &\quad = (-1^{n-k}) \mathbf{e}^1 \wedge \mathbf{e}^{k+1} \wedge \cdots \wedge \mathbf{e}^n \wedge \mathbf{e}^2 \wedge \cdots \wedge \mathbf{e}^k \\ &\quad = (-1^{2(n-k)}) \mathbf{e}^1 \wedge \mathbf{e}^2 \wedge \mathbf{e}^{k+1} \wedge \cdots \wedge \mathbf{e}^n \wedge \mathbf{e}^3 \wedge \cdots \wedge \mathbf{e}^k \\ &\qquad \vdots \\ &\quad = (-1^{k(n-k)}) \mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^k \wedge \mathbf{e}^{k+1} \wedge \cdots \wedge \mathbf{e}^n \\ &\quad = (-1)^{k(n-k)}\, dVol. \end{aligned} \tag{16.28}$$

So on the one hand

$$\mathbf{e}^{k+1} \wedge \cdots \wedge \mathbf{e}^n \wedge * \left(\mathbf{e}^{k+1} \wedge \cdots \wedge \mathbf{e}^n\right) = dVol \tag{16.29}$$

and on the other hand

$$\left(\mathbf{e}^{k+1} \wedge \cdots \wedge \mathbf{e}^n\right) \wedge \left(\mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^k\right) \;=\; (-1)^{k(n-k)} dVol \tag{16.30}$$

which means that we must have $** = (-1)^{k(n-k)}$ as promised. $\square$

16.5 Exercises

1) Compute $*1$ and $*dVol$. Show that $*\,(\eta \wedge *\eta) = |\eta|^2$.

2) Let V be an inner product space of dimension 4, and let $\mathcal{B}^* = \{\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3, \mathbf{e}^4\}$ be an orthonormal basis of V^*. Let $\eta = 2\mathbf{e}^1 - \mathbf{e}^3$ and let $\gamma = 3\mathbf{e}^2$. Compute $*(\eta \wedge \gamma)$. What is $|\eta \wedge \gamma|$? What is $|*(\eta \wedge \gamma)|$?

3) If $n \equiv 0\,(mod\,2)$, then show that $* : \bigwedge^{\frac{n}{2}} V^* \to \bigwedge^{\frac{n}{2}} V^*$. Considering the action of $*$ on this middle dimension, if $n \equiv 0\,(mod\,4)$ then show that $** = Id$, and if $n \equiv 2\,(mod\,4)$ then $** = -Id$.

4) (Complex numbers) Consider $V = \mathbb{R}^2$ and give V^* the standard basis $\mathbf{e}^1 = (1, 0)$, $\mathbf{e}^2 = (0, 1)$ and orientation $dVol = \eta^1 \wedge \eta^2$.

   a) Show that $* : \bigwedge^1 V^* \to \bigwedge^1 V^*$ and that $** = -Id$.

   b) Show that under the standard identification $\mathbb{R}^2 \approx \mathbb{C}$, then $* \approx i$ (where $i = \sqrt{-1}$).

*5) (Spinor spaces.) We know that $** : \bigwedge^k \to \bigwedge^k$ is the operator $** = (-1)^{k(n-k)}$.

   a) If $n = 4$, then show that $** : \bigwedge^2 \to \bigwedge^2$ is the identity. Using this, show that the eigenvalues of the operator $* : \bigwedge^2 \to \bigwedge^2$ are ± 1.

   b) The eigenspaces of $\bigwedge^2$ are denoted $\bigwedge^+$ and $\bigwedge^-$, where $* : \bigwedge^+ \to \bigwedge^+$ acts by multiplication by 1 and $* : \bigwedge^- \to \bigwedge^-$ acts by multiplication by -1. If $\eta \in \bigwedge^2$, show that $\eta + *\eta \in \bigwedge^+$ and that $\eta - *\eta \in \bigwedge^-$.

   c) We will construct bases for $\bigwedge^+$ and for $\bigwedge^-$. Let $\mathcal{B}^* = \{\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3, \mathbf{e}^4\} \subset V^*$ be an orthonormal basis. Let η be one of $\mathbf{e}^1 \wedge \mathbf{e}^2$, $\mathbf{e}^1 \wedge \mathbf{e}^3$, $\mathbf{e}^1 \wedge \mathbf{e}^4$. Using the method of part (b), find three independent elements of $\bigwedge^+$, $\bigwedge^-$, respectively.

   d) $\bigwedge^+$ and $\bigwedge^-$ are sometimes called the *right-* and *left-handed spinor spaces*, respectively (technically, they are Pauli spinor spaces, not Dirac spinor spaces). Show that your bases for these spinor spaces are in fact orthogonal, and of length $\sqrt{2}$.

Lecture 17 - Manifolds

Tuesday April 8, 2014

17.1 The Definition of a Manifold

We begin our migration from geometry as a study of surfaces within an ambient space to fully intrinsic geometry.

Definition. A *manifold* M is a second countable Hausdorff topological space M such that every $p \in M$ has a neighborhood that is homeomorphic to an open set in $\mathbb{R}^n$.

(Note: in class I mistakenly failed to include the "second countable" criterion.)

To unwrap this definition, we remind ourselves of a few other definitions.

Definition. A set X with a collection $\mathcal{T}$ of subsets is called a *topological space* provided

i) $\varnothing \in \mathcal{T}$ and $X \in \mathcal{T}$

ii) (Closure under finite intersections) If $U_1, \ldots, U_k \in \mathcal{T}$ then $U_1 \cap \cdots \cap U_k \in \mathcal{T}$

iii) (Closure under arbitrary unions) If $\{U_\alpha\}_{\alpha \in \Lambda}$ is any collection of subsets in $\mathcal{T}$, then $\bigcup_{\alpha \in \Lambda} U_\alpha \in \mathcal{T}$.

If X with $\mathcal{T}$ is a topological space, then the sets in $\mathcal{T}$ are called the *open sets*. As a warning, the notion of "open set" does not necessarily have anything to do with the usual open sets of $\mathbb{R}^n$ that we're used to.

Example. (The discrete topology). If X is any set, then the *discrete topology* on X is the topology where *every* subset of X is an open set.

Example. (The standard topology on $\mathbb{R}$). Here $X = \mathbb{R}$. and the open sets are the standard open sets that we're used to: $U \in \mathcal{T}$ means that U is any union of a collection of open intervals.

Example. (The Zariski topology on $\mathbb{R}$). The open sets are arbitrary unions of sets of the form $\mathbb{R} \setminus \{a_1, \ldots, a_k\}$ where a_k is any finite collection (possibly empty) of points in $\mathbb{R}$. This is also called the *topology of finite compliments*: open sets are those sets U whose compliments $\mathbb{R} \setminus U$ consist of finitely many points. (The one exception is the empty set.)

Definition. A topological space is called *Hausdroff* if the open sets separate points.

To be more precise, a topological space is Hausdorff provided whenever $x, y \in X$, there are open sets $U, V \in \mathcal{T}$ so that $x \in U$, $y \in V$, $x \neq y$, and $U \cap V = \varnothing$.

Example. The discrete topology on any set X is Hausdorff. Specifically, if $x, y \in X$ and $x \neq y$, then the open sets $U = \{x\}$ and $V = \{y\}$ contain x and y, respectively, and clearly $U \cap V = \varnothing$.

Example. The Zariski topology on $\mathbb{R}$ is not Hausdorff. If $x \in U$ and $y \in V$ where U, V are open, then $U \cap V$ is also open and non-empty, and therefore the compliment of $U \cap V$ consists of only finitely many points. Therefore $U \cap V$ is not the empty set.

Example. (A set with a non-Hausdorff double point.) Let X be the space formed as follows: start with the open interval $(-1, 1)$, remove 0, and add in two points p and q. We will let the open set be arbitrary unions of sets of the following four types:

i) (a, b) where $0 \leq a < b \leq 1$

ii) $(-b, -a)$ where $0 \leq a < b \leq 1$

iii) $(-a, 0) \cup \{p\} \cup (0, b)$ where $0 < a \leq 1$ and $0 < b \leq 1$

iii) $(-a, 0) \cup \{q\} \cup (0, b)$ where $0 < a \leq 1$ and $0 < b \leq 1$

In this case, any open set that contains p has a subset of type (*iii*) and any open set that contains q has a subset of type (*iv*). Yet sets of types (*iii*) and (*iv*) always intersect. Thus the points p and q cannot be separated by open sets.

This last example is particularly interesting, as every point *does* have a neighborhood that is homeomorphic to an open subset of $\mathbb{R}$. Thus with the "Hausdorff" criterion, pathological spaces like this would have to be considered manifolds.

17.2 Atlases

Theorem 17.2.1 *If M is a manifold, then every point p has a neighborhood U that is homeomorphic to a neighborhood in some $\mathbb{R}^n$. If M is connected, then the dimension of the target space $\mathbb{R}^n$ does not vary from point to point.*

If the manifold M is connected, then it has a well-defined dimension. To indicate the dimension by a superscript: $M = M^n$.

Now consider the case of a *compact* manifold.

Theorem 17.2.2 *If M^n is a compact manifold, then there are finitely many open subsets $\{U_1, \ldots, U_N\}$ of M so that each U_i is homeomorphic to a neighborhood in $\mathbb{R}^n$ and so that*

$$M = \bigcup_{i=1}^{N} U_i. \tag{17.1}$$

Proof. This is a simple application of the definition of compactness. □

A collection of open sets $\{U_\alpha\}$ of M is called an *atlas* for M provided

$$M = \bigcup_{\alpha} U_\alpha. \tag{17.2}$$

What the theorem says is that every compact manifold has an atlas that consists of finitely many open sets.

Theorem 17.2.3 *If M^n is any manifold (not necessarily compact), then there is a finite collection $\{U_i\}_{i=1}^{\infty}$ of open sets such that*

i) Each U_i is homeomorphic to an open set in $\mathbb{R}^n$

ii) We have $M^n = \bigcup_{i=1}^{\infty} U_i$

iii) The collection U_i is locally finite. That is, each $p \in M$ is in only finitely many of the sets U_i.

Example. (An atlas on $\mathbb{S}^1$). The manifold $\mathbb{S}^1$ can be considered to be a subset of $\mathbb{R}^2$, given by $x^2 + y^2 = 1$, where $\begin{pmatrix} x \\ y \end{pmatrix}$ are the standard $\mathbb{R}^2$ coordinates. Let $\mathbb{R}^{\subset}\mathbb{R}$ be the x-axis, with coordinate u. Let $N = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ be the *north pole* of $\mathbb{S}^1$ and let $S = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ be the *south pole.*

Then we have two maps, northern polar stereographic projection

$$u = \psi_N(x, y) = \frac{x}{1 - y} \tag{17.3}$$

and southern polar stereographic projection

$$u = \psi_S(x, y) = \frac{x}{1 + y}. \tag{17.4}$$

The atlas consists of the two open sets $U_N = \mathbb{S}^1 \setminus \{N\}$ and $U_S = \mathbb{S}^1 \setminus \{S\}$. Both charts are homeomorphic to $\mathbb{R}^1$ (which is, of course, an open set in $\mathbb{R}^1$) via the maps we have just defined:

$$\begin{aligned} \psi_N &: \mathbb{S}^1 \setminus \{N\} \longrightarrow \mathbb{R}^1 \\ \psi_S &: \mathbb{S}^1 \setminus \{S\} \longrightarrow \mathbb{R}^1 \end{aligned} \tag{17.5}$$

17.3 Exercises

1) Construct some other atlas for the manifold $\mathbb{S}^1$, in such a way that does not involve a construction like stereographic projection.

2) Find or construct 3 examples of non-manifolds other than the ones described in the notes or discussed in class.

Lecture 18 - Charts and Vector Fields

Thursday April 10, 2014

18.1 Atlases

Recall that a manifold M^n can be covered in at most countably many open sets $\{U_i\}_{i=1}^{\infty}$ where for each U_i there is a coordinate chart

$$\varphi_i : U_i \longrightarrow \widetilde{U}_i, \tag{18.1}$$

where $\widetilde{U}_i$ is some domain in $\mathbb{R}^n$, and φ_i is a homeomorphism.

18.2 Transitions

Suppose a point $p \in M^n$ lies in two charts:

$$\begin{aligned} &\varphi_U : U \to \widetilde{U} \subset \mathbb{R}^n \\ &\varphi_V : V \to \widetilde{V} \subset \mathbb{R}^n \\ &p \in U \cap V. \end{aligned} \tag{18.2}$$

For convenience, we shall denote the coordinates of $\widetilde{U}$ by $x^1, \ldots, x^n$ and the coordinates of $\widetilde{V}$ be $y^1, \ldots, y^n$. Thus we can express

$$\begin{aligned} \varphi_U(p) &= \begin{pmatrix} x^1(p) \\ x^2(p) \\ \vdots \\ x^n(p) \end{pmatrix} \\ \textit{and} & \\ \varphi_V(p) &= \begin{pmatrix} y^1(p) \\ y^2(p) \\ \vdots \\ y^n(p) \end{pmatrix} \end{aligned} \tag{18.3}$$

Now, if the x-coordinates of p are $(x^1, \ldots, x^n)^T$, then what are the y-coordinates of p? The *transition function* (also called the *change of coordinates map*) is simply

$$\varphi_{VU} \triangleq \varphi_V \circ \varphi_U^{-1} : \varphi_U(U \cap V) \longrightarrow \varphi_V(U \cap V). \tag{18.4}$$

Notice that this is a map from a region in $\mathbb{R}^n$ to another region of $\mathbb{R}^n$.

18.3 Example: Stereographic Projection

Let $\mathbb{S}^2$ be the locus $(x^1)^2 + (x^2)^2 + (x^3)^2 = 1$, and let $\varphi_U : \mathbb{S}^2 \setminus \{N\} \to \mathbb{R}^2$, $\varphi_V : \mathbb{S}^2 \setminus \{S\} \to \mathbb{R}^2$ be the usual stereographic projections. We have the projections and their inverses:

$$\begin{aligned} \varphi_U(p) &= \begin{pmatrix} u^1(p) \\ u^2(p) \end{pmatrix} = \begin{pmatrix} \frac{x}{1-z} \\ \frac{y}{1-z} \end{pmatrix}, \qquad \varphi_U^{-1}(u^1, u^2) = \begin{pmatrix} \frac{2u^1}{1+(u^1)^2+(u^2)^2} \\ \frac{2u^2}{1+(u^1)^2+(u^2)^2} \\ \frac{-1+(u^1)^2+(u^2)^2}{1+(u^1)^2+(u^2)^2} \end{pmatrix} \\ \varphi_V(p) &= \begin{pmatrix} v^1(p) \\ v^2(p) \end{pmatrix} = \begin{pmatrix} \frac{x}{1+z} \\ \frac{y}{1+z} \end{pmatrix}, \qquad \varphi_V^{-1}(v^1, v^2) = \begin{pmatrix} \frac{2v^1}{1+(v^1)^2+(v^2)^2} \\ \frac{2v^2}{1+(v^1)^2+(v^2)^2} \\ \frac{1-(v^1)^2-(v^2)^2}{1+(v^1)^2-(v^2)^2} \end{pmatrix} \end{aligned} \tag{18.5}$$

So suppose p has u-coordiantes $(u^1, u^2)^T$. Its v-coordiates must be

$$\begin{aligned}
\varphi_{VU}(u^1, u^2) &= \varphi_V \circ \varphi_U^{-1}(u^1, u^2) \\
&= \varphi_V \begin{pmatrix} \frac{2u^1}{1+(u^1)^2+(u^2)^2} \\ \\ \frac{2u^2}{1+(u^1)^2+(u^2)^2} \\ \\ \frac{-1+(u^1)^2+(u^2)^2}{1+(u^1)^2+(u^2)^2} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\frac{2u^1}{1+(u^1)^2+(u^2)^2}}{1+\frac{-1+(u^1)^2+(u^2)^2}{1+(u^1)^2+(u^2)^2}} \\ \\ \frac{\frac{2u^2}{1+(u^1)^2+(u^2)^2}}{1+\frac{-1+(u^1)^2+(u^2)^2}{1+(u^1)^2+(u^2)^2}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{2u^1}{1+(u^1)^2+(u^2)^2-1+(u^1)^2+(u^2)^2} \\ \\ \frac{2u^2}{1+(u^1)^2+(u^2)^2-1+(u^1)^2+(u^2)^2} \end{pmatrix} \\
&= \begin{pmatrix} \frac{u^1}{(u^1)^2+(u^2)^2} \\ \\ \frac{u^2}{(u^1)^2+(u^2)^2} \end{pmatrix}
\end{aligned} \tag{18.6}$$

You may have seen this operation elsewhere. It is *inversion in the unit circle.* We have derived the change-of-coordinate maps

$$\begin{aligned}
v^1 &= \frac{u^1}{(u^1)^2+(u^2)^2}, \\
v^2 &= \frac{u^2}{(u^1)^2+(u^2)^2}.
\end{aligned} \tag{18.7}$$

18.4 Vector Fields

There are several approaches to the notion of a vector field on a manifold. Today we shall explore the first of them

A vector field *is* a directional derivative. This is probably the most important notion in intrinsic geometry.

Assume $\varphi_U : U \to \widetilde{U} \subset \mathbb{R}^n$ is a coordinate chart, where we consider $\mathbb{R}^n$ to have

$u^1, \ldots, u^n$ coordinates. At a point $p \in M^n$, any directional derivative can be expressed in the following form:

$$a^1 \left.\frac{\partial}{\partial u^1}\right|_p + \ldots + a^n \left.\frac{\partial}{\partial u^n}\right|_p \tag{18.8}$$

Now let $\mathbf{w}$ be a vector

$$\mathbf{w} = w^1 \frac{\partial}{\partial u^1} + \ldots + w^n \frac{\partial}{\partial u^n} \tag{18.9}$$

expressed in the u^i-coordinate system. If p is in a second chart: $p \in V$ where $\varphi_V : V \to \widetilde{V} \subset \mathbb{R}^n$ and we take this $\mathbb{R}^n$ to have coordinates $v^1, \ldots, v^n$, then recall we have the transition map

$$\begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix} = \varphi_{VU} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \tag{18.10}$$

where $\varphi_{VU} = \varphi_V \circ \varphi_U^{-1}\big|_{\varphi_U(U \cap V)}$. If we denote the i^{th}-component of the map φ_{VU} by $(\varphi_{VU})^i$, then of course this is merely an expression of the y-coordinates as functions of the x-coordinates:

$$\begin{aligned} y^1 &= (\varphi_{VU})^i \,(x^1, \ldots, x^n) \\ &\;\;\vdots \\ y^n &= (\varphi_{VU})^n \,(x^1, \ldots, x^n). \end{aligned} \tag{18.11}$$

which is just y^i in terms of the x^j. This lets us use the chain rule to express

$$\begin{aligned} \mathbf{w} &= w^1 \frac{\partial}{\partial u^1} + \ldots + w^n \frac{\partial}{\partial u^n} \\ &= w^1 \frac{\partial v^j}{\partial u^1} \frac{\partial}{\partial v^j} + \ldots + w^n \frac{\partial v^j}{\partial u^n} \frac{\partial}{\partial v^j} \\ &= w^1 \frac{(\varphi_{VU})^j}{\partial u^1} \frac{\partial}{\partial v^j} + \ldots + w^n \frac{(\varphi_{VU})^j}{\partial u^n} \frac{\partial}{\partial v^j}. \end{aligned} \tag{18.12}$$

18.4.1 Example

We will compute the change-of-basis formulas for the two charts on $\mathbb{S}^2$ given by the two stereographic projections. So consider the coordinate maps from above:

$$\begin{aligned} \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} &= \varphi_U(x, y, z) = \begin{pmatrix} \frac{x}{1-z} \\ \frac{y}{1-z} \end{pmatrix} \\ \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} &= \varphi_V(x, y, z) = \begin{pmatrix} \frac{x}{1+z} \\ \frac{y}{1+z} \end{pmatrix}. \end{aligned} \tag{18.13}$$

which gives u to v transition

$$\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \varphi_{VU} \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} = \begin{pmatrix} \frac{u^1}{(u^1)^2+(u^2)^2} \\ \frac{u^2}{(u^1)^2+(u^2)^2} \end{pmatrix}. \tag{18.14}$$

We shall compute the $\frac{\partial}{\partial u^i}$-to-$\frac{\partial}{\partial v^i}$ transitions. We have

$$\begin{aligned}
\frac{\partial}{\partial u^1} &= \frac{\partial v^i}{\partial u^1}\frac{\partial}{\partial v^i} \\
&= \frac{\partial v^1}{\partial u^1}\frac{\partial}{\partial v^1} + \frac{\partial v^2}{\partial u^1}\frac{\partial}{\partial v^2} \\
&= \frac{\partial}{\partial u^1}\left(\frac{u^1}{(u^1)^2+(u^2)^2}\right)\frac{\partial}{\partial v^1} + \frac{\partial}{\partial u^1}\left(\frac{u^2}{(u^1)^2+(u^2)^2}\right)\frac{\partial}{\partial v^2} \\
&= \left(\frac{-(u^1)^2+(u^2)^2}{((u^1)^2+(u^2)^2)^2}\right)\frac{\partial}{\partial v^1} - \left(\frac{2u^1u^2}{((u^1)^2+(u^2)^2)^2}\right)\frac{\partial}{\partial v^2} \\
&= \left(\frac{-(v^1)^2+(v^2)^2}{((v^1)^2+(v^2)^2)^2}\right)\left(\frac{((v^1)^2+(v^2)^2)^4}{((v^1)^2+(v^2)^2)^2}\right)\frac{\partial}{\partial v^1} - \left(\frac{2v^1v^2}{((v^1)^2+(v^2)^2)^2}\right)\left(\frac{((v^1)^2+(v^2)^2)^4}{((v^1)^2+(v^2)^2)^2}\right)\frac{\partial}{\partial v^2} \\
&= \left(-(v^1)^2+(v^2)^2\right)\frac{\partial}{\partial v^1} - 2v^1v^2\frac{\partial}{\partial v^2}
\end{aligned} \tag{18.15}$$

Similarly we compute

$$\frac{\partial}{\partial u^2} = -2v^1v^2\frac{\partial}{\partial v^1} + \left((v^1)^2 - (v^2)^2\right)\frac{\partial}{\partial v^2}. \tag{18.16}$$

Therefore given an arbitrary vector field $\mathbf{w}$ expressed in the u^i-coordinate system

$$\mathbf{w} = w^1\frac{\partial}{\partial u^1} + w^2\frac{\partial}{\partial u^2}, \tag{18.17}$$

we can express $\mathbf{w}$ in the v^i coordinate system as

$$\begin{aligned}
\mathbf{w} &= w^1\frac{\partial}{\partial u^1} + w^2\frac{\partial}{\partial u^2} \\
&= w^1\left(\left(-(v^1)^2+(v^2)^2\right)\frac{\partial}{\partial v^1} - 2v^1v^2\frac{\partial}{\partial v^2}\right) \\
&\quad + w^2\left(-2v^1v^2\frac{\partial}{\partial v^1} + \left((v^1)^2-(v^2)^2\right)\frac{\partial}{\partial v^2}\right) \\
&= \left(w^1\left(-(v^1)^2+(v^2)^2\right) - w^2\left(2v^1v^2\right)\right)\frac{\partial}{\partial v^1} \\
&\quad + \left(w^2\left((v^1)^2-(v^2)^2\right) - w^1\left(2v^1v^2\right)\right)\frac{\partial}{\partial v^2}
\end{aligned} \tag{18.18}$$

18.5 Differentiable Manifolds

Recall our formula

$$\mathbf{w} \;=\; w^1 \frac{(\varphi_{VU})^j}{\partial u^1} \frac{\partial}{\partial v^j} \;+\; \dots \;+\; w^n \frac{(\varphi_{VU})^j}{\partial u^n} \frac{\partial}{\partial v^j}. \tag{18.19}$$

Now the map φ_{VU} maps a region of $\mathbb{R}^n$ to another region of $\mathbb{R}^n$. We are taking partial derivatives of the map φ_{VU}. Therefore, in order for a vector to expressible in more than one chart, the transition function *must be differentiable.*

We are then motivated to make the following

Definition. A manifold M^n is called a *differentiable manifold* if it has an atlas (U_i, φ_i) in which the transition functions $\varphi_{ji} = \varphi_j \circ \varphi_i^{-1}$ are differentiable maps from the region $\varphi_i(U_i \cap U_j) \subseteq \mathbb{R}^n$ to the region $\varphi_j(U_i \cap U_j) \subseteq \mathbb{R}^n$.

18.6 Exercises

1) If $\mathbf{v} = v^1 \frac{\partial}{\partial x^1} + v^2 \frac{\partial}{\partial x^2}$ is a vector field expressed in rectangular coordinates, express $\mathbf{v}$ in polar coordinates.

2) Let $\varphi_U : \mathbb{S}^2 \setminus \{N\} \to \mathbb{R}^2$ be stereographic projection, where $\mathbb{R}^2$ has u^1-u^2 coordinates, and let $\varphi_V : \mathbb{S}^2 \setminus \{S\} \to \mathbb{R}^2$ be stereographic projection, where $\mathbb{R}^2$ has v^1-v^2 coordinates. Let $\mathbf{v}$ be the vector field on $\mathbb{S}^2$ given in u-coordinates by $\mathbf{v} = \frac{\partial}{\partial u^1}$.

   a) Find the vector field in terms of both the $\left\{\frac{\partial}{\partial u^i}\right\}$ and $\left\{\frac{\partial}{\partial v^i}\right\}$ bases.

   b) Make a graph of this vector field in the u-plane, and in the v-plane.

   c) Do your best to sketch this vector field on $\mathbb{S}^2$ itself. Does $\mathbf{v}$ have a singularity? Where?

3) The torus $\mathbb{T}^2$ can be expressed as the image of the map

$$\begin{aligned} \varphi : \mathbb{R}^2 \;&\longrightarrow\; \mathbb{R}^4 \\ \varphi(u,\, v) \;&=\; (\cos(2\pi u),\, \sin(2\pi u),\, \cos(2\pi v),\, \sin(2\pi v))^T . \end{aligned} \tag{18.20}$$

   This is not a parametrization, because it is not one-to-one. Find four domains $U_1, U_2, U_3, U_4 \subset \mathbb{R}^2$ so that the four parametrizations

$$(U_1,\, \varphi|_{U_1}),\; (U_2,\, \varphi|_{U_2}),\; (U_3,\, \varphi|_{U_3}),\; (U_4,\, \varphi|_{U_4}) \tag{18.21}$$

   constitute an atlas.

Lecture 19 - Vectors and Covectors

Tuesday April 15, 2014

19.1 Second definition of a vector at a point

Let $p \in M^n$ be a point in a manifold. A path η through p is a map

$$\eta : (-\epsilon,\, \epsilon) \;\longrightarrow\; M^n \tag{19.1}$$

so that $\eta(0) = p$. If (U, φ_U) is a coordinate chart we define

$$\begin{aligned} \gamma_U &: (-\epsilon,\, \epsilon) \;\longrightarrow\; \mathbb{R}^n \\ \gamma_U(t) &\;\triangleq\; (\gamma \circ \varphi_U)(t). \end{aligned} \tag{19.2}$$

A path η is called *differentiable* if the manifold M^n is differentiable, and η_U is differentiable for each domain U in the atlas.

Any path through p defines a differential operator at p. If $f : M^n \to \mathbb{R}$ is a real-valued function, then this operator is given by

$$\left.\frac{d}{dt}\right|_{t=0} (f \circ \gamma)\,(t) \tag{19.3}$$

If η is another path through p, then the operator it defines is of course

$$\left.\frac{d}{dt}\right|_{t=0} (f \circ \eta)\,(t). \tag{19.4}$$

Possibly these two operators are the same. We define an equivalence relation on differentiable paths through p by

$$\eta \;\sim\; \gamma \quad \text{if and only if } \left.\frac{d}{dt}\right|_{t=0} (f \circ \gamma)\,(t) \;=\; \left.\frac{d}{dt}\right|_{t=0} (f \circ \eta)\,(t). \tag{19.5}$$

The equivalence class of a differentiable path γ is denoted $[\gamma]$. Specifically,

$$\begin{aligned}[\gamma] \;=\; & \text{the set of all paths } \eta : (-\epsilon,\epsilon) \to M^n \text{ with } \eta(0) = p \\ & \text{and so that } \left.\frac{d}{dt}\right|_{t=0} (f\circ\eta)(t) \;=\; \left.\frac{d}{dt}\right|_{t=0} (f\circ\gamma)(t) \\ & \text{for all differntiable functions } f\end{aligned} \tag{19.6}$$

Now we can define the notion of vectors based at p.

Definition. A vector $\vec{v} \in T_pM$ based at p is the differential operator associated to any equivalence class of differentiable paths through p.

Specifically, if $[\gamma]$ is such a equivalence class, then the vector $\vec{v}$ it defined is the differentiable operator whose action on functions is

$$\vec{v}(f) \;=\; \left.\frac{d}{dt}\right|_{t=0} (f\circ\gamma)(t). \tag{19.7}$$

19.2 Relation between the new and the old notion of vectors

Let η be a vector, and let (U, φ_U) be a coordinate chart, where the coordinates are given the labels $u^1, \ldots, u^n$. We will see how to express the vector associated to $[\gamma]$ in terms of the u-coordinates.

Recall that our previous way of expressing a vector required a coordinate chart, and a generic vector had the form

$$\vec{v} \;=\; a^i \frac{\partial}{\partial u^i} \tag{19.8}$$

Define the real-valued function

$$f_U \triangleq f\circ\varphi_U^{-1} \,:\, \widetilde{U} \;\longrightarrow\; \mathbb{R} \tag{19.9}$$

and the path in $\mathbb{R}^n$

$$\begin{aligned}\gamma_U &\,:\, (-\epsilon,\,\epsilon) \;\longrightarrow\; \widetilde{U} \subset \mathbb{R}^n \\ \gamma_U &\;\triangleq\; \varphi_U\circ\gamma \\ \gamma_U(t) &\;=\; \begin{pmatrix} \gamma_U^1(t) \\ \vdots \\ \gamma_U^n(t) \end{pmatrix}\end{aligned} \tag{19.10}$$

Then we compute the action of $[\gamma]$ on the function $f : M^n \to \mathbb{R}$ by computing

$$\begin{aligned}
\vec{v}(f) &= \frac{d}{dt}\bigg|_{t=0} (f \circ \gamma)(t) && \text{Definition of the vector } \vec{v} \\
&= \frac{d}{dt}\bigg|_{t=0} \left(f \circ \varphi_U^{-1} \circ \varphi_U \circ \gamma\right)(t) && \text{Trivial} \\
&= \frac{d}{dt}\bigg|_{t=0} (f_U \circ \gamma_U)(t) && \text{Notice } f_U \text{ and } \gamma_U \text{ are simply objects defined on Euclidean spaces} \\
&= \frac{d\gamma_U^i}{dt}\frac{\partial f_U}{\partial u^i} && \text{Ordinary chain rule for functions in Euclidean space.} \\
&= \left(\frac{d\gamma_U^i}{dt}\frac{\partial}{\partial u^i}\right)(f_U)
\end{aligned} \tag{19.11}$$

Thus we have shown that the vector $\vec{v}$ is the operator

$$\vec{v} \;=\; \frac{d\gamma_U^i}{dt}\frac{\partial}{\partial u^i}. \tag{19.12}$$

From the expression (19.8) we see that the coefficients of the vector $\vec{v}$ that is represented by the path γ are $a^i = \frac{d\gamma_U^i}{dt}$.

19.3 Covectors

Given a point $p \in M^n$, the tangent space at p is the space of all vectors based at p. If we have a coordinate chart (U, φ_U), we know that

$$T_pM \;=\; \mathrm{span}_{\mathbb{R}}\left\{\frac{\partial}{\partial u^1}, \dots, \frac{\partial}{\partial u^n}\right\}. \tag{19.13}$$

This is a vector space. We know that every vector space automatically has a dual; the vector space that is dual to T_pM is usually denoted T_p^*M. If f is any differentiable function on M^n and $\vec{v} \in T_pM$ then we make the definition

$$df(\vec{v}) \;=\; \vec{v}(f). \tag{19.14}$$

In other words, the "d" map takes functions, and assigns them to covectors. Now consider the coordinate functions $u^1, \dots, u^n$. We have the basis

$$\mathcal{B} \;=\; \left\{\frac{\partial}{\partial u^1}, \dots, \frac{\partial}{\partial u^n}\right\} \;\subset\; T_pM. \tag{19.15}$$

Now by definition we have

$$du^i(\vec{v}) \;=\; \vec{v}(u^i) \tag{19.16}$$

so that if we set $\vec{v} = \frac{\partial}{\partial u^j}$ we get

$$du^i \left(\frac{\partial}{\partial u^j} \right) = \frac{\partial u^i}{\partial u^j} = \delta^i_j. \tag{19.17}$$

Therefore

$$\mathcal{B}^* = \left\{ du^1, \ldots, du^n \right\} \subset T^*_P M. \tag{19.18}$$

is the basis that is dual to $\mathcal{B}$.

19.4 Exercises

1) Consider the 2-sphere $\mathbb{S}^2$ given as usual by $x^2 + y^2 + z^2 = 1$. Let $U = \mathbb{S}^2 \setminus \{N\}$ and let $\varphi_U : U \to \mathbb{R}^2$ be northern-polar stereographic projection. Define the function $f : \mathbb{S}^2 \to \mathbb{R}$ given by the height function: $f(x, y, z) = z$.

   a) Express f in the u^1-u^2 coordinate system.

   b) Determine the covector df in terms of the basis $\{du^1, du^2\}$.

Lecture 20 - Push-Forwards, Pull-Backs, the Exterior Derivative, and the Metric

Thursday April 17, 2014

20.1 Vectors and the Push-Forward

20.1.1 Maps and diffeomorphisms

Let M^n and K^m be manifolds (not necessarily of the same dimension), and let

$$\psi : M^n \to K^m \tag{20.1}$$

be a map. Let $U \subset M^n$ and $V \subset K^m$ be charts. Assume $x^1, \ldots, x^n$ are coordinates on U and assume $y^1, \ldots, y^m$ are coordinates on V. Then, provided the image of U under φ lies in V, then ψ can be expressed in coordinates:

$$\psi(x^1, \ldots, x^n) = \begin{pmatrix} \psi^1(x^1, \ldots, x^n) \\ \vdots \\ \psi^m(x^1, \ldots, x^n) \end{pmatrix} \tag{20.2}$$

So if a point $p \in M^n$ has coordinates

$$p = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \tag{20.3}$$

then $\psi(p) \in K^m$ has y-coordinates

$$\psi(p) = \begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} \psi^1(x^1, \ldots, x^n) \\ \vdots \\ \psi^m(x^1, \ldots, x^n) \end{pmatrix} \tag{20.4}$$

Definition. The map $\psi : M^n \to K^m$ is called a *differentiable map* provided, whenever ψ is expressed in coordinates $\psi(x^1, \dots, x^n) = \big(\psi^1(x^1, \dots, x^n), \dots, \psi^m(x^1, \dots, x^n)\big)$, the functions ψ^i are differentiable.

Definition. A map $\psi : M^n \to K^m$ is called a *diffeomorphism* provided ψ is on-to-one and onto (so in particular $m = n$), ψ is a differentiable map, and the inverse ψ^{-1} is a differentiable map.

20.1.2 Push-Forwards of Vectors

Recall the intrinsic definition of a vector: a vector $\mathbf{v}$ based at $p \in M^n$ is the differential operator associated to any equivalence class of differentiable paths $[\gamma]$ through p.

Suppose the path γ represents the vector $\mathbf{v}$. This means that $\gamma : (-\epsilon, \epsilon) \to M^n$ has $\gamma(0) = p$ and that if $f : M^n \to \mathbb{R}$ is any differentiable function, then the action of $\mathbf{v}$ on f is

$$\mathbf{v}(f) \;=\; \frac{d}{dt}\bigg|_{t=0} (f \circ \gamma)(t). \tag{20.5}$$

Given a differentiable map $\psi : M^n \to K^m$ (not necessarily a diffeomorphism), the vector $\mathbf{v}$ can be transferred, in a natural way, to K^m. This is done simply by mapping the representative path $\gamma : (-\epsilon, \epsilon) \to M^n$ to the path $\psi \circ \gamma : (-\epsilon, \epsilon) \to K^m$. We call this the *push-forward* of the vector $\mathbf{v}$.

Definition. If $\mathbf{v} \in T_pM$ is a vector and $\psi : M^n \to K^m$ is a differentiable map, then the *push-forward* of $\mathbf{v}$ along ψ, denoted $\psi_*(\mathbf{v}) \in T_{\psi(p)}K$, is the differential operator given by the following action on differentiable functions $g : K^m \to \mathbb{R}$:

$$\psi_*(\mathbf{v})(g) \;=\; \frac{d}{dt}\bigg|_{t=0} (g \circ \psi \circ \gamma)(t) \tag{20.6}$$

Put simply, if γ represents $\mathbf{v} \in T_pM$, then $\psi \circ \gamma$ represents $\psi_*(\mathbf{v}) \in T_{\psi(p)}K$.

20.1.3 Expressing the push-forward of a vector in coordinates

Letting $\psi : M^n \to K^m$ be a differentiable map, let U and V be coordinate charts with $p \in U$ and $\psi(p) \in V$. Let the coordinates on U be $x^1, \dots, x^n$ and let the coordinates on V be $y^1, \dots, y^m$. Then the differentiable map ψ leads to the change-of-variables

$$y^i \;=\; \psi^i(x^1, \dots, x^n). \tag{20.7}$$

First we examine the vector associated to a path $\gamma(t)$. Given a function $f : M^n \to \mathbb{R}$, we use the chain rule to compute

$$\begin{aligned} \mathbf{v}(f) &= \left.\frac{d}{dt}\right|_{t=0} (f \circ \gamma)(t) \\ &= \frac{d\gamma^i}{dt}\frac{\partial f}{\partial x^i} \\ &= \left(\frac{d\gamma^i}{dt}\frac{\partial}{\partial x^i}\right) f. \end{aligned} \tag{20.8}$$

Thus $\mathbf{v}$ is the operator

$$\mathbf{v} = v^i \frac{\partial}{\partial x^i} \quad \text{where} \quad v^i = \frac{d\gamma^i}{dt}. \tag{20.9}$$

Now let's compute the push-forward. If $h : K \to \mathbb{R}$ is a function on K, then using the chain rule twice, we have

$$\begin{aligned} \psi_*(\mathbf{v})(f) &= \left.\frac{d}{dt}\right|_{t=0} (h \circ \psi \circ \gamma)(t) \\ &= \frac{\partial h}{\partial y^j}\frac{\partial \psi^j}{\partial x^i}\frac{d\gamma^i}{dt} \\ &= \left(\frac{\partial \psi^j}{\partial x^i}\frac{d\gamma^i}{dt}\frac{\partial}{\partial y^j}\right) h \\ &= \left(v^i \frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j}\right) h. \end{aligned} \tag{20.10}$$

We have shown that

$$\begin{aligned} \mathbf{v} &= v^i \frac{\partial}{\partial x^i} \\ \psi_*(\mathbf{v}) &= v^j \frac{\partial y^i}{\partial x^j}\frac{\partial}{\partial y^i} \end{aligned} \tag{20.11}$$

We have the following theorem

Theorem 20.1.1 *If $\psi : M^n \to K^m$ is a differentiable map and $p \in M$, then $\psi_* : T_pM \to T_{\psi(p)}K$ is a linear map. If $\{x^1, \ldots, x^n\}$ are coordinates on M near p and $\{y^1, \ldots, y^m\}$ are coordinates on K near $\psi(p)$, then*

$$\psi_*\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j}. \tag{20.12}$$

20.2 Pull-Back of a Covector

In the situation above, if $\psi : M^n \to K^m$ is a differentiable map, then for every $p \in M$, ψ induces a linear map

$$\psi_* : T_pM \to T_{\psi(p)}K. \tag{20.13}$$

If $\mathbf{v} \in T_pM$ is a vector then $\psi^*(\mathbf{v}) \in T_{\psi(p)}K$. Now if $\eta \in T_p^*K$ then we can actually define an action of η on vectors in T_pM. This action is called the *pull-back* action, and the associated covector on M is denoted $\psi_*(\eta) \in T_p^*M$. We define

$$\psi^*(\eta)(\mathbf{v}) \;=\; \eta\left(\psi^*(\mathbf{v})\right). \tag{20.14}$$

We can compute this in coordinates. We have bases $\{dx^1, \dots, dx^n\} \subset T_p^*M$ and $\{dy^1, \dots, dy^m\} \subset T_{\psi(p)}^*K$. Thus a covector $\eta \in T_{\psi(p)}^*M$ can be represented by $\eta = \eta_i dy^i$. We check the action of $\psi^*(\eta)$ on basis vectors $\frac{\partial}{\partial x^j} \in T_pM$:

$$\begin{aligned}
\psi^*\left(\eta_i dy^i\right)\left(\frac{\partial}{\partial x^j}\right) &= \eta_i dy^i\left(\psi_*\left(\frac{\partial}{\partial x^j}\right)\right) && \text{definition of } \psi^* \\
&= \eta_i dy^i\left(\frac{\partial y^k}{\partial x^j}\frac{\partial}{\partial y^k}\right) && \text{computation from above} \\
&= \frac{\partial y^k}{\partial x^j}\eta_i dy^i\left(\frac{\partial}{\partial y^k}\right) && \text{linearity of the covector operation} \\
&= \frac{\partial y^k}{\partial x^j}\eta_i\,\delta^i_k && \text{action of } dy^i \text{ on } \partial/\partial y^k \\
&= \frac{\partial y^i}{\partial x^j}\eta_i && \text{simplification}
\end{aligned} \tag{20.15}$$

On the other hand, a separate computation gives

$$\begin{aligned}
\frac{\partial y^i}{\partial x^k}\eta_i dx^k\left(\frac{\partial}{\partial x^j}\right) &= \frac{\partial y^i}{\partial x^k}\eta_i\delta^k_j \\
&= \frac{\partial y^i}{\partial x^j}\eta_i.
\end{aligned} \tag{20.16}$$

These computations imply the following transition under pullback:

$$\psi^*\left(dy^i\right) \;=\; \frac{\partial y^i}{\partial x^j}dx^j \tag{20.17}$$

20.2.1 Summary

Given a differentiable map: $\psi : M^n \to K^m$ and local coordinates $\{x^1, \dots, x^n\}$ on M and $\{y^1, \dots, y^m\}$ on K, the maps $\psi_* : T_pM \to T_{\psi(p)}K$ and $\psi^* : T_{\psi(p)}^*K \to T_pM$ are linear. We have the transitions

$$\begin{aligned}\psi_*\left(\frac{\partial}{\partial x^i}\right) &= \frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j} \\ \psi^*\left(dy^i\right) &= \frac{\partial y^i}{\partial x^j}\,dx^j.\end{aligned} \tag{20.18}$$

20.3 The tangent, cotangent, and tensor bundles

We have defined the tangent and cotangent spaces at a point. We can accumulate all tangent spaces at all points on a manifold, and the object is called a *bundle*.

$$\begin{aligned}\text{The Tangent Bundle}: \qquad TM &= \bigcup_{p\in M} T_pM \\ \text{The Cotangent Bundle}: \qquad T^*M &= \bigcup_{p\in M} T_p^*M.\end{aligned} \tag{20.19}$$

Similarly, we have tensor algebras at eat point: $\bigotimes^{*,*} T_pM$ which we can accumulate into the *tensor algebra*

$$\text{The Tensor Bundle}: \qquad \bigotimes\nolimits^{*,*}M = \bigcup\nolimits_{p\in M}\bigotimes\nolimits^{*,*}T_pM. \tag{20.20}$$

One particularly important bundle is the *exterior bundle*:

$$\text{The Exterior Bundle}: \qquad \bigwedge\nolimits^{*}M = \bigcup\nolimits_{p\in M}\bigotimes\nolimits^{*,*}T_p^*M. \tag{20.21}$$

The exterior bundle is only formed over the cotangent spaces, never tangent spaces, at each point.

20.4 The exterior derivative

If f is a differentiable function on a differentiable manifold, we have assigned a meaning to the symbol df. Namely, if $\mathbf{v}$ is a vector based at p, then

$$df\,(\mathbf{v}) = \mathbf{v}(f). \tag{20.22}$$

Notice that

$$df\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial f}{\partial x^i} \tag{20.23}$$

but also that

$$\begin{aligned}\frac{\partial f}{\partial x^j}dx^j\left(\frac{\partial}{\partial x^i}\right) &= \frac{\partial f}{\partial x^j}\left(\frac{\partial x^j}{\partial x^i}\right) \\ &= \frac{\partial f}{\partial x^j}\delta_i^j \\ &= \frac{\partial f}{\partial x^i}.\end{aligned} \tag{20.24}$$

Thus we have the formula

$$df = \frac{\partial f}{\partial x^i}dx^i. \tag{20.25}$$

The exterior derivative, so defined, is a map $d : \bigwedge^0 M \to \bigwedge^1 M$. We can extend this to a map $d : \bigwedge^k \to \bigwedge^{k+1}$ in an axiomatic way as such:

1) (Covector Rule) If f is a differentiable function and $\mathbf{v} \in T_pM$, any $p \in M$, then $df(\mathbf{v}) = \mathbf{v}(f)$.
2) (Composition rule) If f is any function, then $ddf = 0$.
3) (Sum and difference rule) If $\eta \in \bigwedge^k$ and $\omega \in \bigwedge^l$ then $d(\eta \pm \omega) = d\eta \pm d\omega$.
4) (Leibniz rule) If $\eta \in \bigwedge^k$ and $\omega \in \bigwedge^l$ then $d(\eta \wedge \omega) = (d\eta) \wedge \omega + (-1)^k \eta \wedge (d\omega)$.

20.5 Riemannian Metrics

Definition. A *Riemannian metric* is a choice of a symmetric $(0, 2)$-tensor g at each point with the property that, for any $\mathbf{v} \in T_pM$, we have $g(\mathbf{v}, \mathbf{v}) \geq 0$ with equality if and only if $\mathbf{v} = 0 \in T_pM$.

In simple language, a Riemannian metric is the assignment of an inner product to the tangent space at each point of a manifold.

Definition. A *Riemannian manifold* is a pair (M^n, g) where M^n is a differentiable manifold and g is a Riemannian metric.

In coordinates we have the basis $\{dx^1, \ldots, dx^n\}$ for the cotangent spaces, and we can therefore write

$$g = g_{ij}dx^i \otimes dx^j. \tag{20.26}$$

where of course, the coefficients g_{ij} are allowed to vary from point to point. That is, the symbols g_{ij} are themselves functions. Notice that

$$g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = g_{ij}. \tag{20.27}$$

Now given a path $\gamma : [a,b] \to M$, its tangent vector is

$$\dot{\gamma}(t) \;=\; \gamma_*\left(\frac{d}{dt}\right) \;=\; \frac{d\gamma^i}{dt}\frac{\partial}{\partial x^i}. \tag{20.28}$$

The norm-square of this vector is

$$\begin{aligned} |\dot{\gamma}|^2 \;&=\; g(\dot{\gamma},\dot{\gamma}) \\ &=\; g\left(\frac{d\gamma^i}{dt}\frac{\partial}{\partial x^i},\, \frac{d\gamma^j}{dt}\frac{\partial}{\partial x^j}\right) \\ &=\; \frac{d\gamma^i}{dt}\frac{d\gamma^j}{dt} g\left(\frac{\partial}{\partial x^i},\, \frac{\partial}{\partial x^j}\right) \\ &=\; \frac{d\gamma^i}{dt}\frac{d\gamma^j}{dt} g_{ij}. \end{aligned} \tag{20.29}$$

To compute the length of the path $\gamma : [a,b] \to M$ we use the standard formula

$$\mathcal{L}(\gamma) \;=\; \int_a^b |\dot{\gamma}|\; dt. \tag{20.30}$$

This lets us define the distance between any two points on a Riemannian manifold: if $p, q \in M^n$ then we define

$$\mathrm{dist}(p,q) \;=\; \inf_{\gamma}\, \mathcal{L}(\gamma) \tag{20.31}$$

where the infimum is taken over all possible paths γ with end points at p and q.

20.6 Exercises

1) Prove that $dd\eta = 0$ for any k-form $\eta \in \bigwedge^k$. To do this, you'll have to use an induction argument similar to the one you used to show that $i_{\mathbf{v}} i_{\mathbf{v}} \eta = 0$.

2) (The curl operator in 3-space) Recall the curl of a vector field in 3-space:

$$\begin{aligned} \mathbf{v} \;&=\; v^1\frac{\partial}{\partial x^1} + v^2\frac{\partial}{\partial x^2} + v^3\frac{\partial}{\partial x^3} \\ \nabla \times \mathbf{v} \;&=\; \left(\frac{\partial v^3}{\partial x^2} - \frac{\partial v^2}{\partial x^3}\right)\frac{\partial}{\partial x^1} + \left(\frac{\partial v^3}{\partial x^1} - \frac{\partial v^1}{\partial x^3}\right)\frac{\partial}{\partial x^2} + \left(\frac{\partial v^1}{\partial x^2} - \frac{\partial v^2}{\partial x^1}\right)\frac{\partial}{\partial x^3} \end{aligned} \tag{20.32}$$

If $\eta = \eta_1 dx^1 + \eta_2 dx_2 + \eta_3 dx^3$, then show that the curl is $*d\eta$.

3) (The divergence operator in n-space) If $\eta = \eta_1 dx^1 + \cdots + \eta_n dx^n$ is a 1-form in n-space, show that the divergence is $*d*\eta$.

4) Do 3.3.2

5) Do 3.4.1

6) Do 3.3.3

7) Do 4.3.2 (a), (b) in Lovett. He uses the notation $\partial_i = \frac{\partial}{\partial u^i}$. Notice that the coordinates here are spherical polar coordinates, that is, the coordinates of latitude and longitude.

8) Do 4.3.2

9) Do 4.3.6

10) Do 4.3.11

11) Do 4.3.14

Lecture 21 - Pullbacks, metrics, and the distance between two points on a manifold

Tueseday April 22, 2014

21.1 Example: The metric on $\mathbb{S}^2$

Consider our northern polar parametrization:

$$\psi : \mathbb{R}^2 \longrightarrow \mathbb{S}^2 \setminus \{N\} \tag{21.1}$$

$$\psi(u^1, u^2) = \begin{pmatrix} \frac{2u^1}{1+(u^1)^2+(u^2)^2} \\ \\ \frac{2u^2}{1+(u^1)^2+(u^2)^2} \\ \\ \frac{-1+(u^1)^2+(u^2)^2}{1+(u^1)^2+(u^2)^2} \end{pmatrix} \tag{21.2}$$

We compute the push-forwards of the coordinate fields $\frac{\partial}{\partial u^i}$. We place coordinates x^1, x^2, x^3 on $\mathbb{R}^3$, the ambient space of $\mathbb{S}^2$. From (21.2) we have

$$\begin{aligned} \psi^1(u^1, u^2) &= \frac{2u^1}{1+(u^1)^2+(u^2)^2} \\ \psi^2(u^1, u^2) &= \frac{2u^2}{1+(u^1)^2+(u^2)^2} \\ \psi^3(u^1, u^2) &= \frac{-1+(u^1)^2+(u^2)^2}{1+(u^1)^2+(u^2)^2} \end{aligned} \tag{21.3}$$

Therefore

$$\begin{aligned}
\psi_*\left(\frac{\partial}{\partial u^1}\right) &= \frac{\partial\psi^i}{\partial u^1}\frac{\partial}{\partial x^i} \\
&= \frac{\partial\psi^1}{\partial u^1}\frac{\partial}{\partial x^1} + \frac{\partial\psi^2}{\partial u^1}\frac{\partial}{\partial x^2} + \frac{\partial\psi^3}{\partial u^1}\frac{\partial}{\partial x^3} \\
&= \left(\frac{2(1-(u^1)^2+(u^2)^2)}{(1+(u^1)^2+(u^2)^2)^2}\right)\frac{\partial}{\partial x^1} - \left(\frac{4u^1u^2}{(1+(u^1)^2+(u^2)^2)^2}\right)\frac{\partial}{\partial x^2} + \left(\frac{4u^1}{(1+(u^1)^2+(u^2)^2)^2}\right)\frac{\partial}{\partial x^3}
\end{aligned} \tag{21.4}$$

and likewise

$$\begin{aligned}
\psi_*\left(\frac{\partial}{\partial u^2}\right) &= \frac{\partial\psi^1}{\partial u^2}\frac{\partial}{\partial x^1} + \frac{\partial\psi^2}{\partial u^2}\frac{\partial}{\partial x^2} + \frac{\partial\psi^3}{\partial u^2}\frac{\partial}{\partial x^3} \\
&= -\left(\frac{4u^1u^2}{(1+(u^1)^2+(u^2)^2)^2}\right)\frac{\partial}{\partial x^1} + \left(\frac{2(1-(u^1)^2+(u^2)^2)}{(1+(u^1)^2+(u^2)^2)^2}\right)\frac{\partial}{\partial x^2} + \left(\frac{4u^2}{(1+(u^1)^2+(u^2)^2)^2}\right)\frac{\partial}{\partial x^3}
\end{aligned} \tag{21.5}$$

This allows us to compute the pullback metric:

$$(\psi^*g)\left(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j}\right) = g\left(\psi_*\left(\frac{\partial}{\partial u^i}\right), \psi_*\left(\frac{\partial}{\partial u^j}\right)\right) \tag{21.6}$$

We have

$$\begin{aligned}
(\psi^*g)\left(\frac{\partial}{\partial u^1}, \frac{\partial}{\partial u^1}\right) &= \left(\frac{2(1-(u^1)^2+(u^2)^2)}{(1+(u^1)^2+(u^2)^2)^2}\right)^2 + \left(\frac{4u^1u^2}{(1+(u^1)^2+(u^2)^2)^2}\right)^2 + \left(\frac{4u^1}{(1+(u^1)^2+(u^2)^2)^2}\right)^2 \\
&= \frac{4\left(1+(u^1)^4+(u^2)^4-2(u^1)^2+2(u^2)^2-2(u^1)^2(u^2)^2+2(u^1)^2(u^2)^2\right)}{(1+(u^1)^2+(u^2)^2)^4} \\
&\quad + \frac{16(u^1)^2(u^2)^2}{(1+(u^1)^2+(u^2)^2)^4} + \frac{16(u^1)^2}{(1+(u^1)^2+(u^2)^2)^4} \\
&= \frac{4\left(1+(u^1)^4+(u^2)^4+2(u^1)^2+2(u^2)^2+2(u^1)^2(u^2)^2+2(u^1)^2(u^2)^2\right)}{(1+(u^1)^2+(u^2)^2)^4} \\
&= \frac{4\left(1+(u^1)^2+(u^2)^2\right)^2}{(1+(u^1)^2+(u^2)^2)^4} \\
&= \frac{4}{(1+(u^1)^2+(u^2)^2)^2}.
\end{aligned} \tag{21.7}$$

Similarly, we compute

$$\begin{aligned}
(\psi^*g)\left(\frac{\partial}{\partial u^2}, \frac{\partial}{\partial u^2}\right) &= \frac{4}{\left(1+(u^1)^2+(u^2)^2\right)^2} \\
(\psi^*g)\left(\frac{\partial}{\partial u^1}, \frac{\partial}{\partial u^2}\right) &= 0.
\end{aligned} \tag{21.8}$$

Therefore we have computed

$$\begin{aligned}
(\psi^*g) &= \frac{4}{\left(1+(u^1)^2+(u^2)^2\right)^2}\left(du^1\otimes du^1 + du^2\otimes du^2\right) \\
&= \frac{4}{\left(1+(u^1)^2+(u^2)^2\right)^2}\,\delta_{ij}\,du^i\otimes du^j.
\end{aligned} \tag{21.9}$$

21.2 The circumference of a great circle on the unit sphere

Consider the great circle

$$\mathbb{S}^1 \;=\; \left\{ \begin{pmatrix} \sin(t) \\ 0 \\ \cos(t) \end{pmatrix} \;\middle|\; t \in [0, 2\pi) \right\} \subset \mathbb{S}^2 \tag{21.10}$$

Under the stereographic projection

$$\varphi(x^1, x^2, x^3) \;=\; \begin{pmatrix} \frac{x^1}{1-x^3} \\ \frac{x^2}{1-x^3} \end{pmatrix} \tag{21.11}$$

this becomes the path

$$\gamma(t) \;=\; \begin{pmatrix} \gamma^1(t) \\ \\ \gamma^2(t) \end{pmatrix} \;=\; \begin{pmatrix} t \\ \\ 0 \end{pmatrix} \tag{21.12}$$

Then the tangent vector is

$$\begin{aligned} \dot{\gamma}(t) \;&=\; \frac{d\gamma^1}{dt}\frac{\partial}{\partial u^1} + \frac{d\gamma^2}{dt}\frac{\partial}{\partial u^2} \\ &=\; \frac{\partial}{\partial u^1} \end{aligned} \tag{21.13}$$

Then using our previous computation we obtain

$$|\dot{\gamma}(t)| \;=\; \frac{2}{1+t^2} \tag{21.14}$$

Then the total length of γ, as t varies from $-\infty$ to ∞ is

$$\begin{aligned} \mathcal{L}(\gamma) \;&=\; \int_{-\infty}^{\infty} |\dot{\gamma}| \; dt \\ &=\; \int_{-\infty}^{\infty} \frac{2}{1+t^2} \; dt \\ &=\; 2\tan^{-1}(t)\Big|_{t=-\infty}^{\infty} \\ &=\; 2\left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right) \;=\; 2\pi \end{aligned} \tag{21.15}$$

This computation verifies a fact we already knew: the circumference of a great circle on the unit sphere is 2π.

21.3 The Cigar Metric

Consider the metric

$$g = \frac{4}{1+(u^1)^2+(u^2)^2}\left(du^1 \otimes du^1 + du^2 \otimes du^2\right) \tag{21.16}$$

This is known as the *cigar metric*.

21.4 Shortest paths

21.4.1 The setup

Let p and q be points on the Riemannian manifold (M^n, g). We have defined the distance between p and q to be the infimum of the lengths of all paths between p and q.

We begin by assuming that some shortest path $\gamma : [a, b] \to M^n$ between p and q does indeed exist; later on we shall prove this, but for now, we just assume that such a path does indeed exist. That is to say $\gamma(a) = p$ and $\gamma(b) = q$.

Now we consider a *variation of paths*. Let $\gamma_s(t)$ be a family of paths, where $\gamma_0(t) = \gamma(t)$ is the original path. We also require the boundary conditions $\gamma_s(a) = p$ and $\gamma_s(b) = q$.

21.4.2 The variational field

Consider our family of paths γ_s. If we fix s and vary t, we are moving along the path. If we vary s, we are seeing how the paths themselves move. We define the variational field

$$\frac{\partial \gamma_s}{\partial s} = \frac{\partial \gamma_s^i}{\partial s}\frac{\partial}{\partial x^i} \tag{21.17}$$

We have the tangent field:

$$\mathbf{t} = \frac{\partial \gamma_s}{\partial t} = \frac{\partial \gamma_s^i}{\partial t}\frac{\partial}{\partial x^i} \tag{21.18}$$

We define this to be the *variational field*:

$$\mathbf{v} = \left.\frac{\partial \gamma_s}{\partial s}\right|_{s=0} = \left.\frac{\partial \gamma_s^i}{\partial s}\right|_{s=0}\frac{\partial}{\partial x^i} = v^i\frac{\partial}{\partial x^i}. \tag{21.19}$$

Now given the path $\gamma = \gamma_0$ we are free to choose any variation we wish provided the endpoints are fixed, which amounts to choosing an arbitrary variational field $\mathbf{v} = v^i \frac{\partial}{\partial x^i}$.

21.4.3 Energy

The length of path is defined to be

$$\mathcal{L}(\gamma) \;=\; \int_a^b |\dot{\gamma}| \; dt \tag{21.20}$$

and the *energy* of the path to be

$$\mathcal{E}(\gamma) \;=\; \int_a^b |\dot{\gamma}|^2 \; dt \tag{21.21}$$

Instead of minimizing the length, we will minimize the energy. Later we will show that minimizing either length or energy is equivalent.

Now if γ minimizes length, then given any variation γ_s whatsoever, we must have

$$\frac{d}{ds}\mathcal{E}(\gamma_s) \;=\; 0 \tag{21.22}$$

This means

$$\begin{aligned} 0 \;&=\; \frac{d}{ds}\mathcal{E}(\gamma_s) \\ &=\; \frac{d}{ds}\int_a^b |\dot{\gamma}|^2 \; dt \\ &=\; \frac{d}{ds}\int_a^b g\left(\frac{d\gamma_s^i}{dt}\frac{\partial}{\partial x^i}, \; \frac{d\gamma_s^j}{dt}\frac{\partial}{\partial x^j}\right) \\ &=\; \frac{d}{ds}\int_a^b g_{ij}\frac{d\gamma_s^i}{dt}\frac{d\gamma_s^j}{dt}\; dt \end{aligned} \tag{21.23}$$

21.4.4 The variational argument

The basic argument is to differentiate under the integral, commute partial derivatives, and use integration by parts.

$$\begin{aligned} 0 \;&=\; \frac{d}{ds}\int_a^b g_{ij}\frac{d\gamma_s^i}{dt}\frac{d\gamma_s^j}{dt}\; dt \\ &=\; \int_a^b \frac{dg_{ij}}{ds}\frac{d\gamma_s^i}{dt}\frac{d\gamma_s^j}{dt}\; dt + \int_a^b g_{ij}\frac{d^2\gamma_s^i}{dsdt}\frac{d\gamma_s^j}{dt}\; dt + \int_a^b g_{ij}\frac{d\gamma_s^i}{dt}\frac{d^2\gamma_s^j}{dsdt}\; dt \\ &=\; \int_a^b \frac{dg_{ij}}{ds}\frac{d\gamma_s^i}{dt}\frac{d\gamma_s^j}{dt}\; dt + \int_a^b g_{ij}\frac{d^2\gamma_s^i}{dtds}\frac{d\gamma_s^j}{dt}\; dt + \int_a^b g_{ij}\frac{d\gamma_s^i}{dt}\frac{d^2\gamma_s^j}{dtds}\; dt \qquad \text{Swap partials} \\ &=\; \int_a^b \frac{dg_{ij}}{ds}\frac{d\gamma_s^i}{dt}\frac{d\gamma_s^j}{dt}\; dt - \int_a^b \frac{d\gamma_s^i}{ds}\frac{d}{dt}\left(g_{ij}\frac{d\gamma_s^j}{dt}\right)\; dt - \int_a^b \frac{d\gamma_s^j}{ds}\frac{d}{dt}\left(g_{ij}\frac{d\gamma_s^i}{dt}\right)\; dt \qquad \text{Integration by parts} \end{aligned} \tag{21.24}$$

We can use that $\frac{d}{ds}$ is the derivative along the variational field, meaning

$$\frac{\partial g_{ij}}{\partial s} = \frac{\partial g_{ij}}{\partial x^k}\frac{\partial \gamma_s^k}{\partial s}$$

Then we simplify to obtain

$$\begin{aligned} 0 &= \int_a^b \frac{dg_{ij}}{dx^k}\frac{\partial \gamma_s^k}{\partial s}\frac{d\gamma_s^i}{dt}\frac{d\gamma_s^j}{dt}\,dt \\ &\quad - \int_a^b \frac{d\gamma_s^i}{ds}\frac{dg_{ij}}{dt}\frac{d\gamma_s^j}{dt}\,dt - \int_a^b \frac{d\gamma_s^i}{ds} g_{ij}\frac{d^2\gamma_s^j}{dt^2}\,dt \\ &\quad - \int_a^b \frac{d\gamma_s^j}{ds}\frac{dg_{ij}}{dt}\frac{d\gamma_s^i}{dt}\,dt - \int_a^b \frac{d\gamma_s^j}{ds} g_{ij}\frac{d^2\gamma_s^i}{dt^2}\,dt \end{aligned} \tag{21.25}$$

Using that $\frac{dg_{ij}}{dt} = \frac{\partial g_{ij}}{\partial x^k}\frac{d\gamma_s^k}{dt}$ we obtain

$$\begin{aligned} 0 &= -2\int_a^b \frac{d\gamma_s^i}{ds} g_{ij}\frac{d^2\gamma_s^j}{dt^2}\,dt + \int_a^b \frac{dg_{ij}}{dx^k}\frac{\partial \gamma_s^k}{\partial s}\frac{d\gamma_s^i}{dt}\frac{d\gamma_s^j}{dt}\,dt \\ &\quad - \int_a^b \frac{d\gamma_s^j}{ds}\frac{\partial g_{ij}}{\partial x^k}\frac{d\gamma_s^k}{dt}\frac{d\gamma_s^i}{dt}\,dt - \int_a^b \frac{d\gamma_s^i}{ds}\frac{\partial g_{ij}}{\partial x^k}\frac{d\gamma_s^k}{dt}\frac{d\gamma_s^j}{dt}\,dt \end{aligned} \tag{21.26}$$

Now we can use $\frac{\partial \gamma_s^i}{ds} = v^i$ along with a re-labeling of indices to obtain

$$0 = -2\int_a^b v^l g_{li}\left(\frac{d^2\gamma^i}{dt^2} + \left[\frac{1}{2}\left(\frac{\partial g_{js}}{\partial x^k} + \frac{\partial g_{ks}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^s}\right) g^{si}\right]\frac{d\gamma^j}{dt}\frac{d\gamma^k}{dt}\right)\,dt \tag{21.27}$$

The expression in the brackets occurs often enough that it is given a name:

$$\Gamma_{ij}^k = \frac{1}{2}\left(\frac{\partial g_{is}}{\partial x^j} + \frac{\partial g_{js}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right) g^{sk} \tag{21.28}$$

The Γ_{ij}^k are known as the *Christoffel symbols*. The integral can be expressed

$$0 = -2\int_a^b g_{il}v^l\left(\frac{d^2\gamma^i}{dt^2} + \Gamma_{jk}^i\frac{d\gamma^j}{dt}\frac{d\gamma^k}{dt}\right)\,dt \tag{21.29}$$

Now we can choose v^i arbitrarily, so we can let it be a multiple of $\frac{d^2\gamma^i}{dt^2} + \Gamma_{ij}^k\frac{d\gamma^i}{dt}\frac{d\gamma^j}{dt}$ itself. Then inside the integral is

$$0 = -2\int_a^b \varphi(t)\left|\left(\frac{d^2\gamma^i}{dt^2} + \Gamma_{jk}^i\frac{d\gamma^j}{dt}\frac{d\gamma^k}{dt}\right)\frac{\partial}{\partial x^i}\right|^2\,dt \tag{21.30}$$

The inside of the integral is always non-negative, and because the integral itself is zero, the inside must be exactly zero. We therefore conclude

$$\frac{d^2\gamma^i}{dt^2} + \Gamma_{jk}^i\frac{d\gamma^j}{dt}\frac{d\gamma^k}{dt} = 0 \tag{21.31}$$

This is a non-linear second order system of ODEs.

21.5 Exercises

1) Still to come.

Lecture 22 - Gradients, the Bracket, the Connection, and the Riemann Curvature Tensor

Thursday April 24, 2014

22.1 The Derivatives of Functions

We have the exterior derivative of a function

$$df \tag{22.1}$$

which is defined to be the covector field whose action on vectors is

$$df(\mathbf{v}) \;=\; \mathbf{v}(f). \tag{22.2}$$

If we have a coordinate system $x^1, \ldots, x^n$, then we have computed, in the coordinate basis $\left\{dx^i\right\}_i$ we have

$$df \;=\; \frac{\partial f}{\partial x^i}\,dx^i \tag{22.3}$$

We define the *gradient* of f, denoted ∇f to be

$$\nabla f \;=\; (df)^\sharp\,. \tag{22.4}$$

Recalling the definition of the $\sharp$ map, this means that

$$\langle \nabla f, \mathbf{v}\rangle \;=\; df(\mathbf{v}) \;=\; \vec{v}(f). \tag{22.5}$$

We already computed $(dx^i)^\sharp = g^{ij}\frac{\partial}{\partial x^j}$, and by the linearity of the $\sharp$ map, we have

$$(df)^\sharp \;=\; \left(\frac{\partial f}{\partial x^i}dx^i\right)^\sharp \;=\; \frac{\partial f}{\partial x^i}\left(dx^i\right)^\sharp \;=\; g^{ij}\frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^j} \tag{22.6}$$

As an aside, this corresponds well with the older notion of the directional derivative. In classical vector calculus, the directional derivative of f in the direction $\mathbf{v}$ is the dot product of the gradient with the vector: $\frac{df}{d\vec{v}} = \vec{v}\cdot\nabla f$.

22.2 The Bracket of Two Vector Fields

If X and Y are differentiable vector fields, we define the *bracket* of X and Y to be the differential operator

$$[X, Y](f) \;=\; X(Y(f)) \;-\; Y(X(f)) \tag{22.7}$$

where f is any twice-differentiable function. On the surface it seems that $[X, Y]$ is a second-order differentiable operator. However we can prove the $[X, Y]$ is actually a first order operator, and is therefore a directional derivative, and therefore a vector field.

Theorem 22.2.1 *If X, Y are differentiable vector fields, then the bracket $[X, Y]$ is a first order differentiable operator, and therefore a vector field.*

Proof.

To compute locally near a point p, we can assume there is a coordinate chart U with $p \in U$ and coordinates $x^1, \dots, x^n$. Then we can express X and Y in these coordinates as

$$X \;=\; a^i \frac{\partial}{\partial x^i}, \qquad\qquad Y \;=\; b^j \frac{\partial}{\partial x^j}. \tag{22.8}$$

Then letting f be a twice differentiable function, we have, by using the definition

$$\begin{aligned} [X, Y](f) \;&=\; X(Y(f)) \;-\; Y(X(f)) \\ &=\; a^i \frac{\partial}{\partial x^i}\left(b^j \frac{\partial f}{\partial x^j}\right) - b^i \frac{\partial}{\partial x^i}\left(a^j \frac{\partial f}{\partial x^j}\right) \\ &=\; a^i \frac{\partial b^j}{\partial x^i}\frac{\partial f}{\partial x^j} \;+\; a^i b^j \frac{\partial^2 f}{\partial x^i \partial x^j} \;-\; b^j \frac{\partial a^i}{\partial x^j}\frac{\partial f}{\partial x^i} \;-\; a^i b^j \frac{\partial^2 f}{\partial x^j \partial x^i} \end{aligned} \tag{22.9}$$

By the commutativity of mixed partials, we have $a^i b^j \frac{\partial^2 f}{\partial x^i \partial x^j} \;-\; a^i b^j \frac{\partial^2 f}{\partial x^j \partial x^i} \;=\; 0$. Therefore

$$\begin{aligned} [X, Y] \;&=\; \left(a^i \frac{\partial b^j}{\partial x^i}\frac{\partial}{\partial x^j} \;-\; b^j \frac{\partial a^i}{\partial x^j}\frac{\partial}{\partial x^i}\right) \\ &=\; \left(a^i \frac{\partial b^j}{\partial x^i} \;-\; b^i \frac{\partial a^j}{\partial x^i}\right)\frac{\partial}{\partial x^i} \end{aligned} \tag{22.10}$$

□

Incidentally, this gives us a formula for $[X, Y]$. We have

$$[X, Y] \;=\; \left(a^i \frac{\partial b^j}{\partial x^i} \;-\; b^i \frac{\partial a^j}{\partial x^i}\right)\frac{\partial}{\partial x^i} \tag{22.11}$$

where $X = a^i \frac{\partial}{\partial x^i}$ and $Y = b^j \frac{\partial}{\partial x^j}$.

22.3 The Covariant Derivative

In contrast to our ideas of derivatives of functions, the notion of a directional derivative of a vector field is fraught with substantial difficulties. Consider, for instance, what a constant function is as opposed to a "constant" vector: a constant function $f : M^n \to \mathbb{R}$ has the simple form $f(x) = c$ for all $x \in M^n$. But the idea of a constant vector field has no such interpretation. For instance if M^n were to be embedded in $\mathbb{R}^n$, then how can a tangent vector field be constant if the tangent space itself is everywhere changing?

So we take an axiomatic approach. Let X and Y be vector fields. By convention, the "directional derivative" of Y in the direction X is denoted

$$\nabla_X Y. \tag{22.12}$$

For historical (and other) reasons, the symbol ∇ is called *the connection.*

Theorem 22.3.1 *Given any (differentiable) vector fields X, Y, X_1, X^2, Y_1, Y_2, and Z the vector field $\nabla_X Y$ is uniquely defined by the following five axioms:*

i) (Linearity in the first variable) If f_1, f_2 are functions then

$$\nabla_{f_1 X_1 + f_2 X_2} Y \;=\; f_1 \nabla_{X_1} Y \;+\; f_2 \nabla_{X_2} Y. \tag{22.13}$$

ii) (Additivity in the second variable) We have

$$\nabla_X \left(Y_1 \pm Y_2\right) \;=\; \nabla_X Y_1 \;\pm\; \nabla_X Y_2. \tag{22.14}$$

iii) (The Leibniz Rule) If f is any differentiable function, we have

$$\nabla_X (fY) \;=\; X(f)\, Y \;+\; f \cdot \nabla_X Y \tag{22.15}$$

iv) (The Torsion Free Condition) We have

$$\nabla_X Y \;-\; \nabla_Y X \;=\; [X, Y] \tag{22.16}$$

v) (Compatibility with the metric) With the notation $\langle X, Y\rangle = g(X, Y)$ we have

$$X \left\langle Y,\, Z\right\rangle \;=\; \left\langle \nabla_X Y,\, Z\right\rangle \;+\; \left\langle Y,\, \nabla_X Z\right\rangle \tag{22.17}$$

Further, we have the implicit formula

$$\begin{aligned} \left\langle \nabla_X Y,\, Z\right\rangle \;=\; & \frac{1}{2} X \left\langle X,\, Y\right\rangle + \frac{1}{2} Y \left\langle X,\, Z\right\rangle \,-\, \frac{1}{2} Z \left\langle X,\, Y\right\rangle \\ & + \frac{1}{2} \left\langle [X, Y],\, Z\right\rangle \,-\, \frac{1}{2} \left\langle [Y, Z],\, X\right\rangle \,+\, \frac{1}{2} \left\langle [Z, X],\, Y\right\rangle. \end{aligned} \tag{22.18}$$

Proof. Using the axioms, we verify the formula, which is known as the *Koszul formula*. We write

$$\begin{aligned} 2\langle \nabla_X Y,\, Z\rangle \;=\; & X\langle Y,\, Z\rangle + Y\langle X,\, Z\rangle \,-\, Z\langle X,\, Y\rangle \\ & +\langle [X,Y],\, Z\rangle \,-\, \langle [Y,Z],\, X\rangle \,+\, \langle [Z,X],\, Y\rangle\,. \end{aligned} \tag{22.19}$$

Expanding the right side using compatibility with the metric we have

$$\begin{aligned} & \langle \nabla_X Y,\, Z\rangle + \langle Y,\, \nabla_X Z\rangle + \langle \nabla_Y X,\, Z\rangle + \langle X,\, \nabla_Y Z\rangle - \langle \nabla_Z X,\, Y\rangle - \langle X,\, \nabla_Z Y\rangle \\ & \quad +\langle [X,Y],\, Z\rangle \,-\, \langle [Y,Z],\, X\rangle \,+\, \langle [Z,X],\, Y\rangle\,. \end{aligned} \tag{22.20}$$

Using the torsion-free condition, we have

$$\begin{aligned} & \langle \nabla_X Y,\, Z\rangle + \langle Y,\, \nabla_X Z\rangle + \langle \nabla_Y X,\, Z\rangle + \langle X,\, \nabla_Y Z\rangle - \langle \nabla_Z X,\, Y\rangle - \langle X,\, \nabla_Z Y\rangle \\ & \quad +\langle \nabla_X Y,\, Z\rangle - \langle \nabla_Y X,\, Z\rangle \,-\, \langle \nabla_Y Z,\, X\rangle \,-\, \langle \nabla_Z Y,\, X\rangle \,+\, \langle \nabla_Z X,\, Y\rangle \,+\, \langle \nabla_X Z,\, Y\rangle\,. \end{aligned} \tag{22.21}$$

Canceling terms, we see that indeed

$$\begin{aligned} & \langle \nabla_X Y,\, Z\rangle + \langle Y,\, \nabla_X Z\rangle + \langle \nabla_Y X,\, Z\rangle + \langle X,\, \nabla_Y Z\rangle - \langle \nabla_Z X,\, Y\rangle - \langle X,\, \nabla_Z Y\rangle \\ & \quad +\langle \nabla_X Y,\, Z\rangle - \langle \nabla_Y X,\, Z\rangle \,-\, \langle \nabla_Y Z,\, X\rangle \,-\, \langle \nabla_Z Y,\, X\rangle \,+\, \langle \nabla_Z X,\, Y\rangle \,+\, \langle \nabla_X Z,\, Y\rangle \\ & \quad = \; 2\langle \nabla_X Y,\, Z\rangle\,. \end{aligned} \tag{22.22}$$

This formula unequivocally determines $\nabla_X Y$. □

Definition. The *covariant derivative* of Y in the direction of X is the field $\nabla_X Y$.

If $x^1, \ldots, x^n$ is a coordinate system, we can express the metric as

$$g \;=\; g_{ij}\, dx^i \otimes dx^j \tag{22.23}$$

where, of course, the coefficients $g_{ij} = g_{ij}(x^1, \ldots, x^n)$ are functions of the variables. Let's determine the covariant derivatives $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$. Since partials commute, we have $[\frac{\partial}{\partial x^s}, \frac{\partial}{\partial x^t}] = 0$ for any s, t. From the Koszul formula we have

$$\begin{aligned} 2\left\langle \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j},\, \frac{\partial}{\partial x^s}\right\rangle \;&=\; \frac{\partial}{\partial x^i}\left\langle \frac{\partial}{\partial x^j},\, \frac{\partial}{\partial x^s}\right\rangle + \frac{\partial}{\partial x^j}\left\langle \frac{\partial}{\partial x^i},\, \frac{\partial}{\partial x^s}\right\rangle - \frac{\partial}{\partial x^s}\left\langle \frac{\partial}{\partial x^i},\, \frac{\partial}{\partial x^j}\right\rangle \\ &=\; \frac{\partial}{\partial x^i} g_{js} + \frac{\partial}{\partial x^j} g_{is} - \frac{\partial}{\partial x^s} g_{ij}. \end{aligned} \tag{22.24}$$

Now $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$ is a vector field, so we can write $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = A^l \frac{\partial}{\partial x^l}$ for some unknown coefficients A^l. Therefore the left side can be written

$$2\left\langle \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j},\, \frac{\partial}{\partial x^s}\right\rangle \;=\; 2\left\langle A^l \frac{\partial}{\partial x^l},\, \frac{\partial}{\partial x^s}\right\rangle \;=\; 2A^l g_{sl}. \tag{22.25}$$

This gives

$$A^l g_{sl} \;=\; \frac{1}{2}\frac{\partial}{\partial x^i} g_{js} + \frac{1}{2}\frac{\partial}{\partial x^j} g_{is} - \frac{1}{2}\frac{\partial}{\partial x^s} g_{ij}. \tag{22.26}$$

so that multiplying both sides by the inverse g^{sk} we have

$$A^k \;=\; A^l g_{sl} g^{sk} \;=\; \frac{1}{2}\left(\frac{\partial}{\partial x^i} g_{js} + \frac{\partial}{\partial x^j} g_{is} - \frac{\partial}{\partial x^s} g_{ij}\right) g^{sk} \tag{22.27}$$

You will notice that the right side is a Christoffel symbol. We therefore have

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \;=\; \Gamma^k_{ij} \frac{\partial}{\partial x^k}. \tag{22.28}$$

22.4 An interpretation of the geodesic equation

Recall that a path

$$\gamma(t) \;=\; \begin{pmatrix} \gamma^1(t) \\ \vdots \\ \gamma^n(t) \end{pmatrix} \tag{22.29}$$

(expressed in coordinates) is a geodesic provided that the coefficients γ^i satisfies the second order system

$$\frac{d^s \gamma^k}{dt^2} \;+\; \Gamma^k_{ij} \frac{d\gamma^i}{dt} \frac{d\gamma^j}{dt} \;=\; 0. \tag{22.30}$$

Now the tangent vector is $\dot{\gamma} = \frac{d\gamma^i}{dt} \frac{\partial}{\partial x^i}$. We compute

$$\begin{aligned} \nabla_{\dot{\gamma}} \dot{\gamma} \;&=\; \nabla_{\frac{d\gamma^i}{dt} \frac{\partial}{\partial x^i}} \left(\frac{d\gamma^j}{dt} \frac{\partial}{\partial x^j}\right) \\ &=\; \frac{d\gamma^i}{dt} \nabla_{\frac{\partial}{\partial x^i}} \left(\frac{d\gamma^j}{dt} \frac{\partial}{\partial x^j}\right) \end{aligned} \tag{22.31}$$

Using the Leibniz rule, we have

$$\begin{aligned} \nabla_{\dot{\gamma}} \dot{\gamma} \;&=\; \frac{d\gamma^i}{dt} \frac{\partial}{\partial x^i} \left(\frac{d\gamma^j}{dt}\right) \frac{\partial}{\partial x^j} \;+\; \frac{d\gamma^i}{dt} \frac{d\gamma^j}{dt} \Gamma^k_{ij} \frac{\partial}{\partial x^k} \\ &=\; \frac{d^2\gamma^k}{dt^2} \frac{\partial}{\partial x^k} \;+\; \frac{d\gamma^i}{dt} \frac{d\gamma^j}{dt} \Gamma^k_{ij} \frac{\partial}{\partial x^k} \\ &=\; \left(\frac{d^2\gamma^k}{dt^2} \;+\; \frac{d\gamma^i}{dt} \frac{d\gamma^j}{dt} \Gamma^k_{ij}\right) \frac{\partial}{\partial x^k}. \end{aligned} \tag{22.32}$$

Therefore a path γ is a geodesic if and only if

$$\nabla_{\dot{\gamma}} \dot{\gamma} \;=\; 0. \tag{22.33}$$

22.5 The Riemann Curvature Tensor

As we know, partial derivatives commute. However, partial covariant derivatives of vector fields do not necessarily commute. We define the *Riemann curvature tensor* to be

$$\mathrm{Rm}(X, Y)Z \;=\; \nabla_X \nabla_Y Z \,-\, \nabla_Y \nabla_X Z \,-\, \nabla_{[X,Y]} Z. \tag{22.34}$$

This is a measure of the failure of mixed partial derivatives of vector fields to commute.

22.6 Exercises

1) Still to come.

Lecture 23 - Parallel Transport, the Riemann Tensor, the Ricci Tensor, and Scalar Curvature

Wednesday April 30, 2014

23.1 Notions of Curvature

We have come through three notions of curvature:

- Eulerian curvature: Product of principle curvatures (Valid for surfaces in $\mathbb{R}^3$)
- Gaussian Curvature: The Jacobian of the Gauss map, which is $det(II)/det(I)$ (Valid for surfaces in $\mathbb{R}^3$)
- Riemannian Curvature: The failure of partial derivatives of vector fields to commute (Valid for arbitrary Riemannian manifolds, whether embedded in an ambient space or not)

The Riemannian notion of curvature is the fully modern notion of curvature. This is not a stepping stone to something else; it is the way things continue to be done today.

The Riemann curvature tensor $\mathrm{Rm}(X,Y)Z$ is related to to Gaussian curvature via the notion of *sectional curvature*: If $X, Y \in T_pM$ are non-parallel vectors based at p, they span an "infinitesimal parallelogram." The curvature of the "germ of the plane" spanned by these vectors is

$$sec(X,\,Y) \;=\; \frac{\langle \mathrm{Rm}(X,Y)Y, X\rangle}{|X|^2|Y|^2 - \langle X,Y\rangle^2}. \tag{23.1}$$

This should be thought of as the Gaussian curvature of the infinitesimal part of the bent plane spanned by X and Y near p.

The infinitesimal manifestation of curvature is the failure of partials to commute when taking derivatives of vector fields. The large-scale manifestation of curvature is the phenomenon of *holonomy*. This occurs when a vector is transported along a curve in such a way that the vector field remains constant along the curve, but when the vector returns to the starting point, it is no longer pointing in the direction it was when it started.

23.2 Parallel Fields and Parallel Transport

Let $\gamma : [a, b] \to M^n$ be a parametrized path in M^n, with initial point $\gamma(a) = p$ and terminal point $\gamma(b) = q$. A vector field X is said to be *parallel* along γ if

$$\nabla_{\dot{\gamma}} X \;=\; 0. \tag{23.2}$$

Now using $\dot{\gamma} = \frac{d\gamma^i}{dt}\frac{\partial}{\partial x^i}$ and writing $X = a^i \frac{\partial}{\partial x^i}$, we compute (using linearity and Leibnitz rules)

$$\begin{aligned} \nabla_{\dot{\gamma}} X \;&=\; \nabla_{\frac{d\gamma^i}{dt}\frac{\partial}{\partial x^i}} \left(a^j \frac{\partial}{\partial x^j} \right) \\ &=\; \frac{d\gamma^i}{dt}\frac{\partial a^j}{\partial x^i}\frac{\partial}{\partial x^j} \;+\; a^i \frac{d\gamma^i}{dt} \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^i} \\ &=\; \frac{da^j}{dt}\frac{\partial}{\partial x^j} \;+\; a^i \frac{d\gamma^j}{dt} \Gamma_{ij}^k \frac{\partial}{\partial x^k} \\ &=\; \left(\frac{da^k}{dt} \;+\; a^i \frac{d\gamma^j}{dt} \Gamma_{ij}^k \right) \frac{\partial}{\partial x^k}. \end{aligned} \tag{23.3}$$

The Christoffel symbols are uniquely determined by the metric, and the path's velocity vector $\dot{\gamma}(t) = \frac{d\gamma^i}{dt}\frac{\partial}{\partial x^i}$ are given. Thus this is a *linear* system of ODEs for the coefficients a^j.

Therefore given a vector $X \in T_pM$, one can use this as an initial condition, and *uniquely* solve the system. This is known as the *parallel transport* of the vector X along γ.

Given a path $\gamma : [a, b] \to \mathbb{R}$ with $\gamma(a) = \gamma(b) = p$ (the end point and the starting point are the same), we can take a vector $X \in T_pM$, transport it around the path γ back to the starting point. This induces a transformation $T_pM \to T_pM$, called the *holonomy map*. If the manifold is flat, the holonomy map is always the identity map. If the manifold is curved, then the holonomy map associated to most paths will not be the identity. Holonomy is thus a manifestation of curvature.

23.3 The Riemann Tensor

Recall the Riamann curvature operator:

$$\mathrm{Rm}(X, Y)Z \;\triangleq\; \nabla_X \nabla_Y Z \;-\; \nabla_Y \nabla_X Z \;=\; \nabla_{[X,Y]} Z. \tag{23.4}$$

23.3.1 The Riemann operator is a tensor

This appears to be a second-order differential operator acting on Z. Surprisingly, it is not a differential operator, but indeed a linear operator. To wit, if f is any differentiable function on M, then we use the Leibniz rule to compute

$$\begin{aligned}
\mathrm{Rm}(X,Y)fZ &= \nabla_X\nabla_Y(fZ) - \nabla_Y\nabla_X(fZ) - \nabla_{[X,Y]}(fZ) \\
&= \nabla_X\left(f\nabla_Y Z + Y(f)Z\right) - \nabla_Y\left(f\nabla_X Z - X(f)Z\right) - f\nabla_{[X,Y]}Z - [X,Y](f)\,Z \\
&= X(f)\nabla_Y Z + f\nabla_X\nabla_Y Z + X(Y(f))Z + Y(f)\nabla_X Z \\
&\quad - Y(f)\nabla_X Z - Y(X(f))Z - f\nabla_Y\nabla_X Z - X(f)\nabla_Y Z \\
&\quad - f\nabla_{[X,Y]}Z - [X,Y](f)\,Z \\
&= f\left(\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z\right) + X(Y(f)) - Y(X(f)) - [X,Y](f) \\
&= f\,\mathrm{Rm}(X,\,Y)Z.
\end{aligned} \tag{23.5}$$

Similarly one can compute $\mathrm{Rm}(X,fY)Z = f\,\mathrm{Rm}(X,Y)Z$ and $\mathrm{Rm}(fX,Y)Z = f\,\mathrm{Rm}(X,Y)Z$. Therefore the Riemann operator Rm is actually a tensor.

23.3.2 Computation of Rm in terms of g

Since Rm is a tensor, in coordinates it can be expressed

$$\mathrm{Rm} = \mathrm{Rm}_{ijk}{}^{l}\,dx^i\otimes dx^j\otimes dx^k\otimes\frac{\partial}{\partial x^k} \tag{23.6}$$

By definition we have

$$\mathrm{Rm}_{ijk}{}^{l}\frac{\partial}{\partial x^l} = \mathrm{Rm}\left(\frac{\partial}{\partial x^i},\,\frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^k} \tag{23.7}$$

Then we compute

$$\begin{aligned}
\mathrm{Rm}\left(\frac{\partial}{\partial x^i},\,\frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^k} &= \nabla_{\frac{\partial}{\partial x^i}}\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^k} - \nabla_{\frac{\partial}{\partial x^j}}\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^k} \\
&= \nabla_{\frac{\partial}{\partial x^i}}\left(\Gamma^s_{jk}\frac{\partial}{\partial x^s}\right) - \nabla_{\frac{\partial}{\partial x^j}}\left(\Gamma^s_{ik}\frac{\partial}{\partial x^s}\right) \\
&= \frac{\partial\Gamma^l_{jk}}{\partial x^i}\frac{\partial}{\partial x^l} + \Gamma^s_{jk}\Gamma^l_{is}\frac{\partial}{\partial x^l} - \frac{\partial\Gamma^l_{ik}}{\partial x^j}\frac{\partial}{\partial x^l} - \Gamma^s_{ik}\Gamma^l_{js}\frac{\partial}{\partial x^l}
\end{aligned} \tag{23.8}$$

Factoring out the $\frac{\partial}{\partial x^l}$, we have

$$\begin{aligned}
\mathrm{Rm}_{ijk}{}^{l} &= \frac{\partial\Gamma^l_{jk}}{\partial x^i} - \frac{\partial\Gamma^l_{ik}}{\partial x^j} + \Gamma^s_{jk}\Gamma^l_{is} - \Gamma^s_{ik}\Gamma^l_{js}, \quad \textit{where} \\
\Gamma^k_{ij} &= \left(\frac{\partial g_{is}}{\partial x^j} + \frac{\partial g_{js}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^s}\right)g^{sk}.
\end{aligned} \tag{23.9}$$

This is a terribly nonlinear second order differential expression in the metric.

23.3.3 Properties of the Riemann Tensor

The $(4,0)$ Riemann tensor is simply

$$\mathrm{Rm}(X,Y,Z,W) \;=\; \langle \mathrm{Rm}(X,Y)Z,\, W\rangle\,. \tag{23.10}$$

In components, this is the lowering of an index: $\mathrm{Rm}_{ijkl} = \mathrm{Rm}_{ijk}{}^{s}g_{sl}$. We list the following elementary properties of the Riemann tensor. These are easily proven by following computations similar to those done above.

- Antisymmetry in the first two variables: $\mathrm{Rm}(X,Y,Z,W) = -\,\mathrm{Rm}(Y,X,Z,W)$
- Symmetry in the first two and the last two variables: $\mathrm{Rm}(X,Y,Z,W) = +\,\mathrm{Rm}(Z,W,X,Y)$
- Antisymmetry in the last two variables: $\mathrm{Rm}(X,Y,Z,W) = -\,\mathrm{Rm}(X,Y,W,Z)$
- The Bianchi identity: $\mathrm{Rm}(X,Y,Z,W) + \mathrm{Rm}(Y,Z,X,W) + \mathrm{Rm}(Z,X,Y,W) \;=\; 0$ (this is the analogue of the Jacobi identity)

23.4 The Ricci Tensor and Scalar Curvature

The *Ricci tensor* is a particular trace of the Riemann tensor:

$$\mathrm{Ric}_{ij} \;\triangleq\; \mathrm{Rm}\,{}_{sij}{}^{s} \;=\; g^{st}\,\mathrm{Rm}\,{}_{sijt} \tag{23.11}$$

In the case of an n-dimensional manifold M^n, this has the following interpretation. Let $X \in T_pM$ be a vector; then the perpendicular space has orthonormal span $\{\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}\}$. Then we have precisely

$$\mathrm{Ric}(X,X) \;=\; \sum_{i=1}^{n-1} \mathrm{Rm}(\mathbf{e}_i,\, X,\, X,\, \mathbf{e}_i) \;=\; \sum_{i=1}^{n-1} sec(X,\mathbf{e}_i). \tag{23.12}$$

This $\mathrm{Ric}(X,X)$ is the sum of sectional curvature amongst planes that pass through X.

The *scalar curvature* is the trace of the Ricci tensor:

$$s \;\triangleq\; g^{ij}\,\mathrm{Ric}\,{}_{ij} \;=\; g^{st}g^{ij}\,\mathrm{Rm}\,{}_{sijt}. \tag{23.13}$$

23.5 The Einstein Gravity Equations

The *gravity tensor* (also called the Einstein tensor) is

$$G_{ij} \;=\; \mathrm{Ric}\,{}_{ij} \;-\; \frac{1}{2}sg_{ij}. \tag{23.14}$$

In physics, all of the matter fields, energy fields, and various energy flows and energy flow torsions can be encoded into an object called the *stress-energy tensor*. We won't be able to discuss it here, except to say that it is the variational field associated to the matter Lagrangian of quantum mechanics. The Einstein Field equations are

$$G_{ij} \;=\; \frac{8\pi G}{c^4}\,T_{ij} \tag{23.15}$$

where c is the speed of light and G is Newton's gravitational constant. It is possible to append only one more term, while maintaining conservation of energy. If Λ is any constant, then the equations

$$G_{ij} \;+\; \Lambda g_{ij} \;=\; \frac{8\pi G}{c^4}\,T_{ij} \tag{23.16}$$

continue to satisfy conservation of energy. The constant Λ is known as the *cosmological constant*. Whether it is zero or nonzero has been a contentious issue in cosmology since it was introduced in a 1917 paper of Einstein's. The Ricci curvature is essentially a kind of non-linear Laplacian of the metric. The field equations are therefore analogous to a non-linear Poisson equation. (Indeed the signature of the metric of General Relativity is $(+,-,-,-)$, so the Laplacian is, instead, the D'Alembertian; the field equations are therefore better seen as a non-linear wave equation).

23.6 Warped Prodcuts

Consider the graph of a function $f(t)$, rotated about the x-axis. This lead to a surface of revolution that can be parametrized by the variable t and the rotational parameter α:

$$\psi(t,\alpha) \;=\; \left(\begin{array}{c} t \\ -\sin(\alpha)\,f(t) \\ \cos(\alpha)\,f(t) \end{array}\right) \tag{23.17}$$

23.6.1 The surface's metric

We compute the coordinate fields:

$$\begin{aligned}
\frac{\partial}{\partial t} \;&=\; \frac{\partial \psi^i}{\partial x^i} \;=\; \frac{dt}{dt}\frac{\partial}{\partial x} - \sin(\alpha)\frac{df}{dt}\frac{\partial}{\partial y} + \cos(\alpha)\frac{df}{dt}\frac{\partial}{\partial z} \\
&=\; \frac{\partial}{\partial x} - \sin(\alpha)f'(t)\frac{\partial}{\partial y} + \cos(\alpha)f'(t)\frac{\partial}{\partial z} \\
\frac{\partial}{\partial \alpha} \;&=\; \frac{\partial t}{\partial \alpha}\frac{\partial}{\partial x} - f(t)\frac{d\sin(\alpha)}{d\alpha}\frac{\partial}{\partial y} + f(t)\frac{d\cos(\alpha)}{d\alpha}\frac{\partial}{\partial z} \\
&=\; -f(t)\cos(\alpha)\frac{\partial}{\partial y} + f(t)\sin(\alpha)\frac{\partial}{\partial z}
\end{aligned} \tag{23.18}$$

From this easily compute

$$g\left(\frac{\partial}{\partial t},\frac{\partial}{\partial t}\right) \;=\; \left(1\;+\;(f'(t))^2\right) \tag{23.19}$$

$$g\left(\frac{\partial}{\partial t},\frac{\partial}{\partial \alpha}\right) \;=\; 0 \tag{23.20}$$

$$g\left(\frac{\partial}{\partial \alpha},\frac{\partial}{\partial \alpha}\right) \;=\; f(t)^2 \tag{23.21}$$

so that

$$g \;=\; \left(1+(f'(t))^2\right) dt\otimes dt \;+\; f(t)^2\; d\alpha\otimes d\alpha \tag{23.22}$$

This metric becomes simpler if we introduce the following change of coordinates:

$$\begin{aligned} r \;&=\; \int_0^t (1\;+\;f'(\tau))\,d\tau \\ \theta \;&=\; \alpha \end{aligned} \tag{23.23}$$

We therefore compute $dr = \left(1+(f'(t))^2\right) dt$ and $d\theta = d\alpha$. Therefore

$$g \;=\; dr\otimes dr \;+\; f(r)^2\, d\theta\otimes d\theta. \tag{23.24}$$

Whether or not this metric comes from a surface of revolution in $\mathbb{R}^3$, a metric in this form is called a *warped product metric*.

23.7 The Christoffel symbols

We compute, using the Koszul formula, that

$$\begin{aligned} 2\left\langle\nabla_{\frac{\partial}{\partial r}}\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right\rangle \;&=\; \frac{\partial}{\partial r}\left\langle\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right\rangle+\frac{\partial}{\partial r}\left\langle\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right\rangle-\frac{\partial}{\partial r}\left\langle\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right\rangle \\ &=\; \frac{\partial}{\partial r}1 \;=\; 0 \\ 2\left\langle\nabla_{\frac{\partial}{\partial r}}\frac{\partial}{\partial r},\frac{\partial}{\partial \theta}\right\rangle \;&=\; \frac{\partial}{\partial r}\left\langle\frac{\partial}{\partial r},\frac{\partial}{\partial \theta}\right\rangle+\frac{\partial}{\partial r}\left\langle\frac{\partial}{\partial r},\frac{\partial}{\partial \theta}\right\rangle-\frac{\partial}{\partial \theta}\left\langle\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right\rangle \\ &=\; 0. \end{aligned} \tag{23.25}$$

which gives $\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r} \; = \; 0$ Then we compute

$$\begin{aligned} 2\left\langle \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta}, \frac{\partial}{\partial r} \right\rangle \; &= \; \frac{\partial}{\partial r}\left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial r} \right\rangle + \frac{\partial}{\partial \theta}\left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \right\rangle - \frac{\partial}{\partial r}\left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta} \right\rangle \\ &= \; \frac{\partial}{\partial \theta} 1 \; = \; 0 \\ 2\left\langle \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta} \right\rangle \; &= \; \frac{\partial}{\partial r}\left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta} \right\rangle + \frac{\partial}{\partial \theta}\left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta} \right\rangle - \frac{\partial}{\partial \theta}\left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta} \right\rangle \\ &= \; \frac{\partial}{\partial r}(f(r))^2 \; = \; 2f(r)f'(r). \end{aligned} \tag{23.26}$$

which gives $\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta} \; = \; \frac{f'(r)}{f(r)} \frac{\partial}{\partial \theta}$ Finally we compute

$$\begin{aligned} 2\left\langle \nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial \theta}, \frac{\partial}{\partial r} \right\rangle \; &= \; \frac{\partial}{\partial \theta}\left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial r} \right\rangle + \frac{\partial}{\partial \theta}\left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial r} \right\rangle - \frac{\partial}{\partial r}\left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta} \right\rangle \\ &= \; -\frac{\partial}{\partial r}\left(f(r)\right)^2 \; = \; -2f(r)f'(r) \\ 2\left\langle \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta} \right\rangle \; &= \; \frac{\partial}{\partial \theta}\left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta} \right\rangle + \frac{\partial}{\partial \theta}\left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta} \right\rangle - \frac{\partial}{\partial \theta}\left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta} \right\rangle \\ &= \; \frac{\partial}{\partial \theta}(f(r))^2 \; = \; 0. \end{aligned} \tag{23.27}$$

which gives $\nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial \theta} \; = \; -f(r)f'(r)\frac{\partial}{\partial r}$.

To sum up, we have

$$\begin{aligned} \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r} \; &= \; 0 \\ \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta} \; &= \; \nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial r} \; = \; \frac{f'(r)}{f(r)} \frac{\partial}{\partial \theta} \\ \nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial \theta} \; &= \; -f(r)f'(r)\frac{\partial}{\partial r} \end{aligned} \tag{23.28}$$

23.7.1 The Riemann tensor and the sectional curvature of a warped product

We compute

$$\begin{aligned}
\mathrm{Rm}\left(\frac{\partial}{\partial\theta},\frac{\partial}{\partial r}\right)\frac{\partial}{\partial r} &= \nabla_{\frac{\partial}{\partial\theta}}\nabla_{\frac{\partial}{\partial r}}\frac{\partial}{\partial r} \;-\; \nabla_{\frac{\partial}{\partial r}}\nabla_{\frac{\partial}{\partial\theta}}\frac{\partial}{\partial r} \\
&= -\nabla_{\frac{\partial}{\partial r}}\left(\frac{f'(r)}{f(r)}\frac{\partial}{\partial\theta}\right) \\
&= -\frac{\partial}{\partial r}\left(\frac{f'(r)}{f(r)}\right)\frac{\partial}{\partial\theta} - \frac{f'(r)}{f(r)}\nabla_{\frac{\partial}{\partial r}}\frac{\partial}{\partial\theta} \\
&= -\frac{f''(r)}{f(r)}\frac{\partial}{\partial\theta} + \left(\frac{f'(r)}{f(r)}\right)^2\frac{\partial}{\partial\theta} - \left(\frac{f'(r)}{f(r)}\right)\frac{\partial}{\partial\theta}
\end{aligned} \tag{23.29}$$

Therefore

$$\mathrm{Rm}\left(\frac{\partial}{\partial\theta},\frac{\partial}{\partial r}\right)\frac{\partial}{\partial r} \;=\; \frac{f''(r)}{f(r)}\frac{\partial}{\partial\theta} \tag{23.30}$$

so that the sectional (that is, Gaussian) curvature is

$$\begin{aligned}
\sec\left(\frac{\partial}{\partial r},\frac{\partial}{\partial\theta}\right) &= \frac{\left\langle \mathrm{Rm}\left(\frac{\partial}{\partial\theta},\frac{\partial}{\partial r}\right)\frac{\partial}{\partial r},\,\frac{\partial}{\partial\theta}\right\rangle}{\left|\frac{\partial}{\partial r}\right|^2\left|\frac{\partial}{\partial\theta}\right|^2 \;-\; \left\langle\frac{\partial}{\partial r},\,\frac{\partial}{\partial\theta}\right\rangle^2} \\
&= \frac{\left\langle -\frac{f''(r)}{f(r)}\frac{\partial}{\partial\theta},\,\frac{\partial}{\partial r}\right\rangle}{f(r)^2} \\
&= -\frac{f''(r)}{f(r)}.
\end{aligned} \tag{23.31}$$

23.7.2 Spaces of constant curvature +1 and -1

We have found a new parametrization of the sphere: If we use $f(r) = \sin(r)$ then

$$g \;=\; dr\otimes dr \;+\; \sin^2(r)\,d\theta\otimes d\theta \tag{23.32}$$

then we find that $\sec \;=\; -\frac{1}{\sin^2(r)}\frac{d^2\sin(r)}{dr^2} \;=\; +1$, and the parametrization only proceeds for $r\in[0,\pi]$.

If we use $f(r) \;=\; \sinh(r)$ instead, then the parametrization proceeds for $r\in[0,\infty)$ and we find that

$$g \;=\; dr\otimes dr \;+\; \sinh^2(r)\,d\theta\otimes d\theta \tag{23.33}$$

has Gaussian curvature

$$\sec\left(\frac{\partial}{\partial r},\frac{\partial}{\partial\theta}\right) \;=\; -\frac{1}{\sinh(r)}\frac{d^2\sinh(r)}{dr^2} \;=\; -1. \tag{23.34}$$

23.8 Exercises

1) Still to come.