Differential Geometry in Physics

Gabriel Lugo
Department of Mathematical Sciences and Statistics
University of North Carolina at Wilmington

This document was reproduced by the University of North Carolina at Wilmington from a camera ready copy supplied by the authors. The text was generated on an desktop computer using LaTeX.

Preface

These notes were developed as a supplement to a course on Differential Geometry at the advanced undergraduate, first year graduate level, which the author has taught for several years. There are many excellent texts in Differential Geometry but very few have an early introduction to differential forms and their applications to Physics. It is the purpose of these notes to bridge some of these gaps and thus help the student get a more profound understanding of the concepts involved. We also provide a bridge between the very practical formulation of classical differential geometry and the more elegant but less intuitive modern formulation of the subject. In particular, the central topic of curvature is presented in three different but equivalent formalisms.

These notes should be accessible to students who have completed traditional training in Advanced Calculus, Linear Algebra, and Differential Equations. Students who master the entirety of this material will have gained enough background to begin a formal study of the General Theory of Relativity.

Gabriel Lugo, Ph. D.
Mathematical Sciences and Statistics
UNCW
Wilmington, NC 28403
lugo@uncw.edu

Contents

Chapter 1

Vectors and Curves

1.1 Tangent Vectors

1.1 Definition Euclidean n-space $\mathbf{R}^n$ is defined as the set of ordered n-tuples $\mathbf{p} = \langle p^1, \ldots, p^n \rangle$, where $p^i \in \mathbf{R}$, for each $i = 1, \ldots, n$.

We may associate $\mathbf{p}$ with the position vector of a point $p\,(p^1, \ldots, p^n)$ in n-space. Given any two n-tuples $\mathbf{p} = \langle p^1, \ldots, p^n \rangle$, $\mathbf{q} = \langle q^1, \ldots, q^n \rangle$ and any real number c, we define two operations:

$$\begin{aligned} \mathbf{p} + \mathbf{q} &= \langle p^1 + q^1, \ldots, p^n + q^n \rangle, \\ c\mathbf{p} &= \langle cp^1, \ldots, cp^n \rangle. \end{aligned} \tag{1.1}$$

These two operations of vector sum and multiplication by a scalar satisfy all the 8 properties needed to give $\mathbf{R}^n$ a natural structure of a vector space[1].

1.2 Definition Let x^i be the real valued functions in $\mathbf{R}^n$ such that $x^i(\mathbf{p}) = p^i$ for any point $\mathbf{p} = \langle p^1, \ldots, p^n \rangle$. The functions x^i are then called the natural *coordinates* of the the point $\mathbf{p}$. When the dimension of the space $n = 3$, we often write: $x^1 = x$, $x^2 = y$ and $x^3 = z$.

1.3 Definition A real valued function in $\mathbf{R}^n$ is of class C^r if all the partial derivatives of the function up to order r exist and are continuous. The space of infinitely differentiable (smooth) functions will be denoted by $C^\infty(\mathbf{R}^n)$ or $\mathscr{F}((\mathbf{R})^n)$.

In calculus, vectors are usually regarded as arrows characterized by a direction and a length. Vectors as thus considered as independent of their location in space. Because of physical and mathematical reasons, it is advantageous to introduce a notion of vectors that does depend on location. For example, if the vector is to represent a force acting on a rigid body, then the resulting equations of motion will obviously depend on the point at which the force is applied.

In a later chapter we will consider vectors on curved spaces. In these cases the positions of the vectors are crucial. For instance, a unit vector pointing north at the earth's equator is not at all the same as a unit vector pointing north at the tropic of Capricorn. This example should help motivate the following definition.

1.4 Definition A **tangent vector** X_p in $\mathbf{R}^n$, is an ordered pair $\{\mathbf{x}, \mathbf{p}\}$. We may regard $\mathbf{x}$ as an ordinary advanced calculus "arrow-vector" and $\mathbf{p}$ is the position vector of the foot of the arrow.

[1] In these notes we will use the following index conventions:
Indices such as i, j, k, l, m, n, run from 1 to n.
Indices such as μ, ν, ρ, σ, run from 0 to n.
Indices such as $\alpha, \beta, \gamma, \delta$, run from 1 to 2.

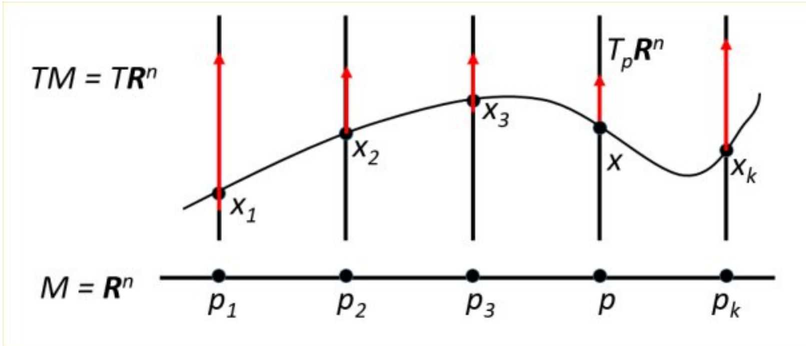


Fig. 1.1: Vector Field as section of the Tangent Bundle.

The collection of all tangent vectors at a point $\mathbf{p} \in \mathbf{R}^n$ is called the **tangent space** at $\mathbf{p}$ and will be denoted by $T_p(\mathbf{R}^n)$. Given two tangent vectors X_p, Y_p and a constant c, we can define new tangent vectors at $\mathbf{p}$ by $(X+Y)_p = X_p + Y_p$ and $(cX)_p = cX_p$. With this definition, it is clear that for each point $\mathbf{p}$, the corresponding tangent space $T_p(\mathbf{R}^n)$ at that point has the structure of a vector space. On the other hand, there is no natural way to add two tangent vectors at different points.

The set $T\mathbf{R}^n$ consisting of the union of all tangent spaces at all points in $\mathbf{R}^n$ is called the **tangent bundle**. This object is not a vector space, but as we will see later it has much more structure than just a set.

1.5 Definition A **vector field** X in $U \in \mathbf{R}^n$ is a smooth function from U to $T(U)$.

We may think of a vector field as a smooth assignment of a tangent vector X_p to each point in in U. More specifically, a vector field is a function from the base space into the tangent bundle - this is called a **section** of the bundle. The collection of all sections of the bundle is denoted by $\Gamma(T\mathbf{R}^n)$ or simply by the set of vector fields $\mathscr{X}(\mathbf{R}^n)$. Vector Fields are also called **contravariant tensor** fields of rank one, denoted by $\mathscr{X}(\mathbf{R}^n) = \mathscr{T}_0^1(\mathbf{R}^n)$.

The difference between a tangent vector and a vector field is that in the latter case, the coefficients a^i are smooth functions of x^i. Since in general there are not enough dimensions to depict a tangent bundle and vector fields as sections thereof, we use abstract diagrams such as shown Figure 1.1. In such a picture, the base space M (in this case $M = \mathbf{R}^n$) is compressed into the continuum at the bottom of the picture in which several points $\mathbf{p}_1 \ldots \mathbf{p}_k$ are shown. To each such point one attaches a tangent space. Here, the tangent spaces are just copies of $\mathbf{R}^n$ shown as vertical **fibers** in the diagram. The vector component $\mathbf{x}_p$ of a tangent vector at the point $\mathbf{p}$ is depicted as an arrow embedded in the fiber. The union of all such fibers constitutes the tangent bundle $TM = T\mathbf{R}^n$. A section of the bundle, that is, a function from the base space into the bundle, amounts to assigning a tangent vector to every point in the the base. It is required that such assignment of vectors is done in a smooth way so that there are no major "changes" of the vector field between nearby points.

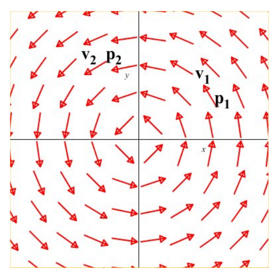


Fig. 1.2: Vector Field

Given any two vector fields X and Y and any smooth function f, we can define new vector fields $X+Y$ and fX by

$$\begin{aligned} (X+Y)_p &= X_p + Y_p \\ (fX)_p &= fX_p \end{aligned} \tag{1.2}$$

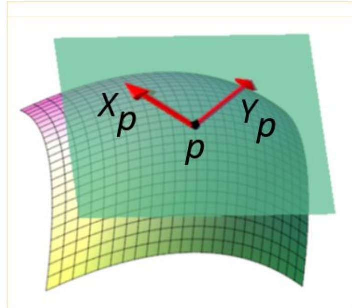


Fig. 1.3: Tangent vectors X_p, Y_p on a surface in $\mathbb{R}^3$.

Remark Since the space of smooth functions is not a field but only a ring, the operations above give the space of vector fields the structure of a ring module. The subscript notation X_p to indicate the location of a tangent vector is sometimes cumbersome but necessary to distinguish them from vector fields.

Vector fields are essential objects in physical applications. If we consider the flow of a fluid in a region, the velocity vector field indicates the speed and direction of the flow of the fluid at that point. Other examples of vector fields in classical physics are the electric, magnetic and gravitational fields. The vector field in figure 1.2 represents a magnetic field around an electrical wire pointing out of the page.

1.6 Definition Let $X_p = \{\mathbf{x}, \mathbf{p}\}$ be a tangent vector in an open neighborhood U of a point $\mathbf{p} \in \mathbf{R}^n$ and let f be a C^∞ function in U. The directional derivative of f at the point $\mathbf{p}$, in the direction of $\mathbf{x}$, is defined by

$$X_p(f) = \boldsymbol{\nabla} f(p) \cdot \mathbf{x}, \tag{1.3}$$

where $\boldsymbol{\nabla} f(p)$ is the gradient of the function f at the point $\mathbf{p}$.

The notation

$$X_p(f) \equiv \nabla_{X_p} f$$

is also used commonly. This notation emphasizes that in differential geometry, we may think of a tangent vector at a point as an operator on the space of smooth functions in a neighborhood of the point. The operator assigns to a function the directional derivative of that function in the direction of the vector. Here we need not assume as in calculus that the direction vectors have unit length.

It is easy to generalize the notion of directional derivatives to vector fields by defining

$$X(f) \equiv \nabla_X f = \boldsymbol{\nabla} f \cdot \mathbf{x}, \tag{1.4}$$

where the function f and the components of $\mathbf{x}$ depend smoothly on the points of $\mathbf{R}^n$.

The tangent space at a point p in $\mathbf{R}^n$ can be envisioned as another copy of $\mathbf{R}^n$ superimposed at the point p. Thus, at a point p in $\mathbf{R}^2$, the tangent space consist of the point p and a copy of the vector space $\mathbf{R}^2$ attached as a "tangent plane" at the point p. Since the base space is a flat 2-dimensional continuum, the tangent plane for each point appears indistinguishable from the base space as in figure 1.2.

Later we will define the tangent space for a curved continuum such as a surface in $\mathbf{R}^3$ as shown in figure 1.3. In this case, the tangent space at a point p consists of the vector space of all vectors actually tangent to the surface at the given point.

1.7 Proposition If $f, g \in \mathscr{F}(\mathbf{R}^n)$, $a, b \in \mathbf{R}$, and $X \in \mathscr{X}(\mathbf{R}^n)$ is a vector field, then

$$\begin{aligned} X(af+bg) &= aX(f)+bX(g) \\ X(fg) &= fX(g)+gX(f) \end{aligned} \tag{1.5}$$

Proof First, let us develop an mathematical expression for tangent vectors and vector fields that will facilitate computation.

Let $\mathbf{p} \in U$ be a point and let x^i be the coordinate functions in U. Suppose that $X_p = (\mathbf{x}, \mathbf{p})$, where the components of the Euclidean vector $\mathbf{x}$ are $\langle v^1, \ldots, v^n \rangle$. Then, for any function f, the tangent vector X_p operates on f according to the formula

$$X_p(f) = \sum_{i=1}^{n} v^i \left(\frac{\partial f}{\partial x^i} \right)(p). \tag{1.6}$$

It is therefore natural to identify the tangent vector X_p with the differential operator

$$\begin{aligned} X_p &= \sum_{i=1}^{n} v^i \left(\frac{\partial}{\partial x^i} \right)_p \\ X_p &= v^1 \left(\frac{\partial}{\partial x^1} \right)_p + \cdots + v^n \left(\frac{\partial}{\partial x^n} \right)_p . \end{aligned} \tag{1.7}$$

Notation: We will be using Einstein's convention to suppress the summation symbol whenever an expression contains a repeated index. Thus, for example, the equation above could be simply written as

$$X_p = v^i \left(\frac{\partial}{\partial x^i} \right)_p . \tag{1.8}$$

This equation implies that the action of the vector X_p on the coordinate functions x^i yields the components a^i of the vector. In elementary treatments, vectors are often identified with the components of the vector and this may cause some confusion.

The quantities

$$\left\{ \left(\frac{\partial}{\partial x^1} \right)_p, \ldots, \left(\frac{\partial}{\partial x^n} \right)_p \right\}$$

form a basis for the tangent space $T_p(\mathbf{R}^n)$ at the point $\mathbf{p}$, and any tangent vector can be written as a linear combination of these basis vectors. The quantities v^i are called the **contravariant** components of the tangent vector. Thus, for example, the Euclidean vector in $\mathbf{R}^3$

$$\mathbf{x} = 3\mathbf{i} + 4\mathbf{j} - 3\mathbf{k}$$

located at a point $\mathbf{p}$, would correspond to the tangent vector

$$X_p = 3 \left(\frac{\partial}{\partial x} \right)_p + 4 \left(\frac{\partial}{\partial y} \right)_p - 3 \left(\frac{\partial}{\partial z} \right)_p .$$

Let $X = v^i \dfrac{\partial}{\partial x^i}$ be an arbitrary vector field and let f and g be real-valued functions. Then

$$\begin{aligned} X(af+bg) &= v^i \frac{\partial}{\partial x^i}(af+bg) \\ &= v^i \frac{\partial}{\partial x^i}(af) + v^i \frac{\partial}{\partial x^i}(bg) \\ &= av^i \frac{\partial f}{\partial x^i} + bv^i \frac{\partial g}{\partial x^i} \\ &= aX(f)+bX(g). \end{aligned}$$

Similarly,

$$
\begin{aligned}
X(fg) &= v^i \frac{\partial}{\partial x^i}(fg) \\
&= v^i f \frac{\partial}{\partial x^i}(g) + v^i g \frac{\partial}{\partial x^i}(f) \\
&= f v^i \frac{\partial g}{\partial x^i} + g v^i \frac{\partial f}{\partial x^i} \\
&= f X(g) + g X(f).
\end{aligned}
$$

Any quantity in Euclidean space which satisfies relations 1.5 is a called a **linear derivation** on the space of smooth functions. The word *linear* here is used in the usual sense of a linear operator in linear algebra, and the word derivation means that the operator satisfies Leibnitz' rule.

The proof of the following proposition is slightly beyond the scope of this course, but the proposition is important because it characterizes vector fields in a coordinate-independent manner.

1.8 Proposition Any linear derivation on $\mathscr{F}(\mathbf{R}^n)$ is a vector field.

This result allows us to identify vector fields with linear derivations. This step is a big departure from the usual concept of a "calculus" vector. To a differential geometer, a vector is a linear operator whose inputs are functions and the output are functions which at each point represent the directional derivative in the direction of the Euclidean vector.

1.9 Example Given the point $p(1,1)$, the Euclidean vector $\mathbf{x} = \langle 3, 4 \rangle$ and the function $f(x,y) = x^2 + y^2$, we associate $\mathbf{x}$ with the tangent vector

$$
X_p = 3\frac{\partial}{\partial x} + 4\frac{\partial}{\partial y}.
$$

Then,

$$
\begin{aligned}
X_p(f) &= 3\left(\frac{\partial f}{\partial x}\right)_p + 4\left(\frac{\partial f}{\partial y}\right)_p, \\
&= 3(2x)|_p + 4(2y)|_p, \\
&= 3(2) + 4(2) = 14.
\end{aligned}
$$

1.10 Example Let $f(x,y,z) = xy^2z^3$ and $\mathbf{x} = \langle 3x, 2y, z \rangle$. Then

$$
\begin{aligned}
X(f) &= 3x\left(\frac{\partial f}{\partial x}\right) + 2y\left(\frac{\partial f}{\partial y}\right) + z\left(\frac{\partial f}{\partial z}\right) \\
&= 3x(y^2z^3) + 2y(2xyz^3) + z(3xy^2z^2), \\
&= 3xy^2z^3 + 4xy^2z^3 + 3xy^2z^3 = 10xy^2z^3.
\end{aligned}
$$

1.11 Definition Let X be a vector field in $\mathbf{R}^n$ and p be a point. A curve $\alpha(t)$ with $\alpha(0) = p$ is called an integral curve of X if $\alpha'(0) = X_p$.

In local coordinates, the expression defining integral curves constitutes a system of first order differential equations, so the existence and uniqueness of solutions apply locally.

Mappings

1.12 Definition Let $F : \mathbf{R}^n \to \mathbf{R}^m$ be a vector function defined by coordinate entries $F(\mathbf{p}) = \langle f^1(\mathbf{p}), f^2(\mathbf{p}), \dots f^m(\mathbf{p})\rangle$. The vector function is called a **mapping** if the coordinate functions are all differentiable. If the coordinate functions are C^∞, F is called a smooth mapping. If $\langle x^1, x^2, \dots x^n\rangle$ are local coordinates in $\mathbf{R}^n$ and $\langle y^1, y^2, \dots y^m\rangle$ local coordinates in $\mathbf{R}^m$, a mapping F is represented in advanced calculus by m functions of n variables

$$y^j = f^j(x^i), \quad i = 1 \dots n, \; j = 1..m \tag{1.9}$$

For each point $\mathbf{p} \in \mathbf{R}^n$, a mapping $F : \mathbf{R}^n \to \mathbf{R}^m$ induces a linear transformation F_* from the tangent space $T_p\mathbf{R}^n$ to the tangent space $T_{F(p)}\mathbf{R}^m$. This map is called the **Jacobian map**, or the **push-forward**. If we let X be a tangent vector in $\mathbf{R}^n$, then the tangent vector F_*X in $\mathbf{R}^m$ is defined by

$$F_*X(f) = X(f \circ F), \tag{1.10}$$

where $f \in \mathscr{F}(\mathbf{R}^m)$. (See fig 1.4)

$$\begin{tikzcd}
X \in T_p\mathbf{R}^n \arrow[r, "F_*"] \arrow[d] & T_{F(t)}\mathbf{R}^m \ni F_*X \arrow[d] & \\
\mathbf{R^n} \arrow[r, "F"] & \mathbf{R}^m \arrow[r, "f"] & \mathbf{R}
\end{tikzcd}$$

Fig. 1.4: Jacobian Map.

If we denote by $\{\frac{\partial}{\partial x^i}\}$ the basis for the tangent space at a point in $\mathbf{R}^n$ and by $\{\frac{\partial}{\partial y^j}\}$ the basis for the tangent space at the corresponding point in $\mathbf{R}^n$, we recognize the right hand side of the formula above as the chain rule

$$\begin{aligned}
\frac{\partial f}{\partial x^i} &= \frac{\partial f}{\partial y^j}\frac{\partial y^j}{\partial x^i} \\
\frac{\partial}{\partial x^i} &= \frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j}.
\end{aligned} \tag{1.11}$$

1.13 Inverse Function Theorem. When $m = n$, mappings are called change of coordinates. In the terminology of tangent spaces, the classical inverse function theorem states that if the Jacobian map F_* is a vector space isomorphism at a point, then there exists a neighborhood of the point in which F is a diffeomorphism.

1.14 Remarks

1. The equation above shows that under change of coordinates, basis tangent vectors and by linearity all tangent vectors change by multiplication by the matrix representation of the Jacobian. This is the source of the pedestrian definition in physics, that a contravariant tensor of rank one, is one that transforms like a contravariant tensor of rank one.

2. Many authors use the notation dF to denote the push-forward map F_*, but we prefer to avoid this notation for now because most students fresh out of advanced calculus have not yet been introduced to the interpretation of the differential as a linear map on tangent spaces.

3. If $F : \mathbf{R}^n \to \mathbf{R}^m$ and $G : \mathbf{R}^m \to \mathbf{R}^p$ are mappings, then $(G \circ F)_* = G_* F_*$. This follows from the standard result in linear algebra for the composition of linear transformations.

4. As we will see later, the concept of the push-forward extends to manifold mappings $F : M \to N$

1.2 Curves in $\mathbf{R}^3$

1.15 Definition A **curve** $\alpha(t)$ in $\mathbf{R}^3$ is a C^∞ map from an interval $I \in \mathbf{R}$ into $\mathbf{R}^3$. The curve assigns to each value of a parameter $t \in \mathbf{R}$, a point $(\alpha^1(t), \alpha^2(t), \alpha^3(t)) \in \mathbf{R}^3$.

$$\begin{array}{rcl} I \in \mathbf{R} & \overset{\alpha}{\longmapsto} & \mathbf{R}^3 \\ t & \longmapsto & \alpha(t) = (\alpha^1(t), \alpha^2(t), \alpha^3(t)) \end{array}$$

One may think of the parameter t as representing time, and the curve α as representing the trajectory of a moving point particle. In these notes we also use classical notation for the position vector

$$\mathbf{x}(t) = \langle x^1(t), x^2(t), x^3(t) \rangle, \tag{1.12}$$

which is more prevalent in vector calculus and elementary physics textbooks. Of course, what this notation really means is

$$x^i(t) = (x^i \circ \alpha)(t), \tag{1.13}$$

where x^i are the coordinate slot functions in an open set in $\mathbf{R}^3$.

1.16 Example Let

$$\alpha(t) = (a_1 t + b_1, a_2 t + b_2, a_3 t + b_3). \tag{1.14}$$

This equation represents a straight line passing through the point $\mathbf{p} = (b_1, b_2, b_3)$, in the direction of the vector $\mathbf{v} = (a_1, a_2, a_3)$.

1.17 Example Let

$$\alpha(t) = (a \cos \omega t, a \sin \omega t, bt). \tag{1.15}$$

This curve is called a circular helix. Geometrically, we may view the curve as the path described by the hypotenuse of a triangle with slope b, which is wrapped around a circular cylinder of radius a. The projection of the helix onto the xy-plane is a circle and the curve rises at a constant rate in the z-direction (See Figure 1.5a). Similarly, The equation $\alpha(t) = (a \cosh \omega t, a \sinh \omega t, bt)$ is called a hyperbolic "helix." It represents the graph of curve that wraps around a hyperbolic cylinder rising at a constant rate.

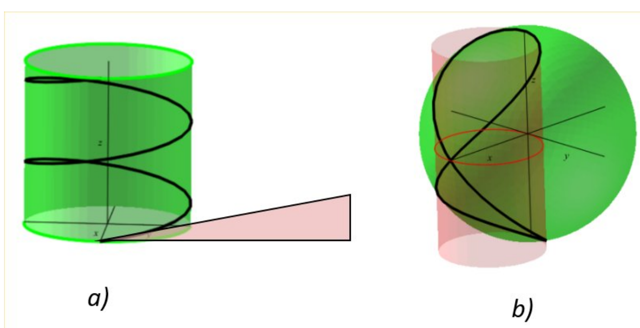


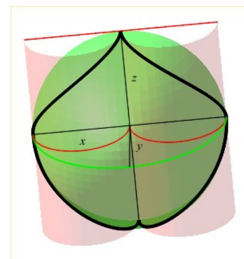
Fig. 1.5: a) Circular Helix. b) Temple of Viviani

1.18 Example Let

$$\alpha(t) = (a(1+\cos t), a\sin t, 2a\sin(t/2)). \tag{1.16}$$

This curve is called the Temple of Viviani. Geometrically, this is the curve of intersection of a sphere $x^2+y^2+z^2=4a^2$ of radius $2a$, and the cylinder $x^2+y^2=2ax$ of radius a with a generator tangent to the diameter of the sphere along the z-axis (See Figure 1.5b).

The Temple of Viviani is of historical interest in the development of calculus. The problem was posed anonymously by Viviani to Leibnitz, to determine on the surface of a semi-sphere, four identical windows, in such a way that the remaining surface be equivalent to a square. It appears as if Viviani was challenging the effectiveness of the new methods of calculus against the power of traditional geometry.



It is said that Leibnitz understood the nature of the challenge and solved the problem in one day. Not knowing the proposer of the enigma, he sent the solution to his Serenity Ferdinando, as he guessed that the challenge came from prominent Italian mathematicians. Upon receipt of the solution by Leibnitz, Viviani posted a mechanical solution solution without proof. He described it as using a boring device to remove from a semisphere, the surface area cut by two cylinders with half the radius, and which are tangential to a diameter of the base. Upon realizing this could not physically be rendered as a temple since the roof surface would rest on only four points, Viviani no longer spoke of a temple but referred to the shape as a "sail."

1.19 Definition Let $\alpha : I \to \mathbf{R}^3$ be a curve in $\mathbf{R}^3$ given in components as above $\alpha = (\alpha^1, \alpha^2, \alpha^3)$. For each point $t \in I$ we define the **velocity** of the curve to be the tangent vector

$$\alpha'(t) = \left(\frac{d\alpha^1}{dt}, \frac{d\alpha^2}{dt}, \frac{d\alpha^3}{dt}\right)_{\alpha(t)} \tag{1.17}$$

At each point of the curve, the velocity vector is tangent to the curve and thus the velocity constitutes a vector field representing the velocity flow along that curve. In a similar manner the second derivative $\alpha''(t)$ is a vector field called the **acceleration** along the curve. The length $v = \|\alpha'(t)\|$ of the velocity vector is called the speed of the curve. The classical components of the velocity vector are simply given by

$$\mathbf{v}(t) = \dot{\mathbf{x}} \equiv \frac{d\mathbf{x}}{dt} = \left(\frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}\right), \tag{1.18}$$

and the speed is

$$v = \sqrt{\left(\frac{dx^1}{dt}\right)^2 + \left(\frac{dx^2}{dt}\right)^2 + \left(\frac{dx^3}{dt}\right)^2}. \tag{1.19}$$

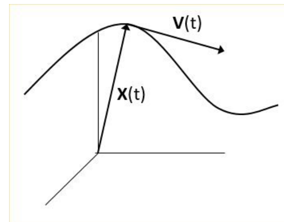
As is well known, the vector form of the equation of the line 1.14 can be written as $\mathbf{x}(t) = \mathbf{p} + t\mathbf{v}$, which is consistent with the Euclidean axiom stating that given a point and a direction, there is only one line passing through that point in that direction. In this case, the velocity $\dot{\mathbf{x}} = \mathbf{v}$ is constant and hence the acceleration $\ddot{\mathbf{x}} = 0$. This is as one would expect from Newton's law of inertia.
The differential $d\mathbf{x}$ of the position vector given by

$$d\mathbf{x} = \langle dx^1, dx^2, dx^3\rangle = \left(\frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}\right) dt \tag{1.20}$$

which appears in line integrals in advanced calculus is some sort of an **infinitesimal tangent vector**. The norm $\|d\mathbf{x}\|$ of this infinitesimal tangent vector is called the differential of arclength ds. Clearly, we have

$$ds = \|d\mathbf{x}\| = v dt. \tag{1.21}$$

If one identifies the parameter t as time in some given units, what this says is that for a particle moving along a curve, the speed is the rate of change of the arclength with respect to time. This is intuitively exactly what one would expect.



As we will see later in this text, the notion of infinitesimal objects needs to be treated in a more rigorous mathematical setting. At the same time, we must not discard the great intuitive value of this notion as envisioned by the masters who invented Calculus, even at the risk of some possible confusion! Thus, whereas in the more strict sense of modern differential geometry, the velocity is a tangent vector and hence it is a differential operator on the space of functions, the quantity $d\mathbf{x}$ can be viewed as a traditional vector which, at the infinitesimal level is a linear approximation to the curve, and points tangentially in the direction of $\mathbf{v}$.

For any smooth function $f : \mathbf{R}^3 \to \mathbf{R}$, we formally define the action of the velocity vector field $\alpha'(t)$ as a linear derivation by the formula

$$\alpha'(t)(f) \mid_{\alpha(t)} = \frac{d}{dt}(f \circ \alpha) \mid_t . \tag{1.22}$$

The modern notation is more precise, since it takes into account that the velocity has a vector part as well as point of application. Given a point on the curve, the velocity of the curve acting on a function, yields the directional derivative of that function in the direction tangential to the curve at the point in question. The diagram in figure 1.6 below provides a more visual interpretation of the velocity vector formula 1.22 as a linear mapping between tangent spaces.

$$\begin{tikzcd}
\frac{d}{dt} \in T_t\mathbf{R} \arrow[r, "\alpha_*"] \arrow[d] & T_{\alpha(t)}\mathbf{R}^3 \ni \alpha'(t) \arrow[d] & \\
\mathbf{R} \arrow[r, "\alpha"] & \mathbf{R}^3 \arrow[r, "f"] & \mathbf{R}
\end{tikzcd}$$

Fig. 1.6: Velocity Vector Operator

The map $\alpha(t)$ from $\mathbf{R}$ to $\mathbf{R}^3$ induces a push-forward map α_* from the tangent space of $\mathbf{R}$ to the tangent space of $\mathbf{R}^3$. The image $\alpha_*(\frac{d}{dt})$ in $T\mathbf{R}^3$ of the tangent vector $\frac{d}{dt}$ is what we call $\alpha'(t)$.

$$\alpha_*(\frac{d}{dt}) = \alpha'(t).$$

Since $\alpha'(t)$ is a tangent vector in $\mathbf{R}^3$, it acts on functions in $\mathbf{R}^3$. The action of $\alpha'(t)$ on a function f on $\mathbf{R}^3$ is the same as the action of $\frac{d}{dt}$ on the composition $f \circ \alpha$. In particular, if we apply $\alpha'(t)$ to the coordinate functions x^i, we get the components of the the tangent vector

$$\alpha'(t)(x^i) \mid_{\alpha(t)} = \frac{d}{dt}(x^i \circ \alpha) \mid_t . \tag{1.23}$$

To unpack the above discussion in the simplest possible terms, we associate with the classical velocity vector $\mathbf{v} = \dot{\mathbf{x}}$ a linear derivation $\alpha'(t)$ given by

$$\alpha'(t) = \frac{dx^1}{dt}\frac{\partial}{\partial x^1} + \frac{dx^2}{dt}\frac{\partial}{\partial x^2} + \frac{dx^3}{dt}\frac{\partial}{\partial x^3}.$$

So, given a real valued function on $\mathbf{R}^3$, the action of the velocity vector is given by the chain rule

$$\alpha'(t)(f) = \frac{\partial f}{\partial x^1}\frac{dx^1}{dt} + \frac{\partial f}{\partial x^2}\frac{dx^2}{dt} + \frac{\partial f}{\partial x^3}\frac{dx^3}{dt} = \boldsymbol{\nabla} f \cdot \mathbf{v}.$$

1.20 Definition

If $t = t(s)$ is a smooth, real valued function and $\alpha(t)$ is a curve in $\mathbf{R}^3$, we say that the curve $\beta(s) = \alpha(t(s))$ is a **reparametrization** of α.

A common reparametrization of curve is obtained by using the arclength as the parameter. Using this reparametrization is quite natural, since we know from basic physics that the rate of change of the arclength is what we call speed

$$v = \frac{ds}{dt} = \|\alpha'(t)\|. \tag{1.24}$$

The arc length is obtained by integrating the above formula

$$s = \int \|\alpha'(t)\|\, dt = \int \sqrt{\left(\frac{dx^1}{dt}\right)^2 + \left(\frac{dx^2}{dt}\right)^2 + \left(\frac{dx^3}{dt}\right)^2}\, dt \tag{1.25}$$

In practice it is typically difficult to actually find an explicit arc length parameterization of a curve since not only does one have calculate the integral, but also one needs to be able to find the inverse function t in terms of s. On the other hand, from a theoretical point of view, arc length parameterizations are ideal since any curve so parametrized has unit speed. The proof of this fact is a simple application of the chain rule and the inverse function theorem.

$$\begin{aligned}
\beta'(s) &= [\alpha(t(s))]' \\
&= \alpha'(t(s))t'(s) \\
&= \alpha'(t(s))\frac{1}{s'(t(s))} \\
&= \frac{\alpha'(t(s))}{\|\alpha'(t(s))\|},
\end{aligned}$$

and any vector divided by its length is a unit vector. Leibnitz notation makes this even more self evident

$$\begin{aligned}
\frac{d\mathbf{x}}{ds} &= \frac{d\mathbf{x}}{dt}\frac{dt}{ds} = \frac{\frac{d\mathbf{x}}{dt}}{\frac{ds}{dt}} \\
&= \frac{\frac{d\mathbf{x}}{dt}}{\|\frac{d\mathbf{x}}{dt}\|}
\end{aligned}$$

1.21 Example Let $\alpha(t) = (a\cos\omega t, a\sin\omega t, bt)$. Then

$$\mathbf{v}(t) = (-a\omega\sin\omega t, a\omega\cos\omega t, b),$$

$$\begin{aligned}
s(t) &= \int_0^t \sqrt{(-a\omega\sin\omega u)^2 + (a\omega\cos\omega u)^2 + b^2}\, du \\
&= \int_0^t \sqrt{a^2\omega^2 + b^2}\, du \\
&= ct, \quad \text{where,} \quad c = \sqrt{a^2\omega^2 + b^2}.
\end{aligned}$$

The helix of unit speed is then given by

$$\beta(s) = (a\cos\frac{\omega s}{c}, a\sin\frac{\omega s}{c}, b\frac{\omega s}{c}).$$

Frenet Frames

Let $\beta(s)$ be a curve parametrized by arc length and let T(s) be the vector

$$T(s) = \beta'(s). \tag{1.26}$$

The vector $T(s)$ is tangential to the curve and it has unit length. Hereafter, we will call T the **unit tangent** vector. Differentiating the relation

$$T \cdot T = 1, \tag{1.27}$$

we get

$$2T \cdot T' = 0, \tag{1.28}$$

so we conclude that the vector T' is orthogonal to T. Let N be a unit vector orthogonal to T, and let κ be the scalar such that

$$T'(s) = \kappa N(s). \tag{1.29}$$

We call N the **unit normal** to the curve, and κ the **curvature**. Taking the length of both sides of last equation, and recalling that N has unit length, we deduce that

$$\kappa = \|T'(s)\| \tag{1.30}$$

It makes sense to call κ the curvature since, if T is a unit vector, then $T'(s)$ is not zero only if the direction of T is changing. The rate of change of the direction of the tangent vector is precisely what one would expect to be a measure how much a curve is curving. In particular, it $T' = 0$ at a particular point, we expect that at that point, the curve is locally well approximated by a straight line.

We now introduce a third vector

$$B = T \times N, \tag{1.31}$$

which we will call the **binormal** vector. The triplet of vectors (T, N, B) forms an orthonormal set; that is,

$$\begin{aligned} T \cdot T = N \cdot N = B \cdot B = 1 \\ T \cdot N = T \cdot B = N \cdot B = 0. \end{aligned} \tag{1.32}$$

If we differentiate the relation $B \cdot B = 1$, we find that $B \cdot B' = 0$, hence B' is orthogonal to B. Furthermore, differentiating the equation $T \cdot B = 0$, we get

$$B' \cdot T + B \cdot T' = 0.$$

rewriting the last equation

$$B' \cdot T = -T' \cdot B = -\kappa N \cdot B = 0,$$

we also conclude that B' must also be orthogonal to T. This can only happen if B' is orthogonal to the TB-plane, so B' must be proportional to N. In other words, we must have

$$B'(s) = -\tau N(s) \tag{1.33}$$

for some quantity τ, which we will call the **torsion**. The torsion is similar to the curvature in the sense that it measures the rate of change of the binormal. Since the binormal also has unit length, the only way one can have a non-zero derivative is if B is changing directions. This means that if addition B did not change directions, the vector would truly be a constant vector, so the curve would be a flat curve embedded into the TN-plane.

The quantity B' then measures the rate of change in the up and down direction of an observer moving with the curve always facing forward in the direction of the tangent vector. The binormal B is something like the flag in the back of sand dune buggy.

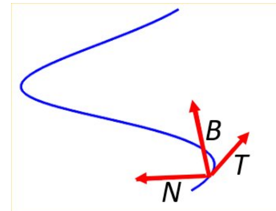


Fig. 1.7: Frenet Frame.

The set of basis vectors $\{T, N, B\}$ is called the **Frenet frame** or the **repère mobile** (moving frame). The advantage of this basis over the fixed $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ basis is that the Frenet frame is naturally adapted to the curve. It propagates along the curve with the tangent vector always pointing in the direction of motion, and the normal and binormal vectors pointing in the directions in which the curve is tending to curve. In particular, a complete description of how the curve is curving can be obtained by calculating the rate of change of the frame in terms of the frame itself.

1.22 Theorem Let $\beta(s)$ be a unit speed curve with curvature κ and torsion τ. Then

$$\begin{array}{rclcl} T' & = & & \kappa N & \\ N' & = & -\kappa T & & \tau B \; . \\ B' & = & & -\tau N & \end{array} \tag{1.34}$$

Proof: We need only establish the equation for N'. Differentiating the equation $N \cdot N = 1$, we get $2N \cdot N' = 0$, so N' is orthogonal to N. Hence, N' must be a linear combination of T and B.

$$N' = aT + bB.$$

Taking the dot product of last equation with T and B respectively, we see that

$$a = N' \cdot T, \text{ and } \quad b = N' \cdot B.$$

On the other hand, differentiating the equations $N \cdot T = 0$, and $N \cdot B = 0$, we find that

$$\begin{aligned} N' \cdot T = -N \cdot T' = -N \cdot (\kappa N) = -\kappa \\ N' \cdot B = -N \cdot B' = -N \cdot (-\tau N) = \tau. \end{aligned}$$

We conclude that $a = -\kappa$, $b = \tau$, and thus

$$N' = -\kappa T + \tau B.$$

The Frenet frame equations (1.34) can also be written in matrix form as shown below.

$$\begin{bmatrix} T \\ N \\ B \end{bmatrix}' = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} T \\ N \\ B \end{bmatrix}. \tag{1.35}$$

The group-theoretic significance of this matrix formulation is quite important and we will come back to this later when we talk about general orthonormal frames. At this time, perhaps it suffices to point out that the appearance of an antisymmetric matrix in the Frenet equations is not at all coincidental.

The following theorem provides a computational method to calculate the curvature and torsion directly from the equation of a given unit speed curve.

1.23 Proposition Let $\beta(s)$ be a unit speed curve with curvature $\kappa > 0$ and torsion τ. Then

$$\begin{aligned} \kappa &= \|\beta''(s)\| \\ \tau &= \frac{\beta' \cdot [\beta'' \times \beta''']}{\beta'' \cdot \beta''} \end{aligned} \tag{1.36}$$

Proof: If $\beta(s)$ is a unit speed curve, we have $\beta'(s) = T$. Then

$$\begin{aligned} T' = \beta''(s) &= \kappa N, \\ \beta'' \cdot \beta'' &= (\kappa N) \cdot (\kappa N), \\ \beta'' \cdot \beta'' &= \kappa^2 \\ \kappa^2 &= \|\beta''\|^2 \end{aligned}$$

$$\begin{aligned} \beta'''(s) &= \kappa' N + \kappa N' \\ &= \kappa' N + \kappa(-\kappa T + \tau B) \\ &= \kappa' N + -\kappa^2 T + \kappa\tau B. \end{aligned}$$

$$\begin{aligned} \beta' \cdot [\beta'' \times \beta'''] &= T \cdot [\kappa N \times (\kappa' N + -\kappa^2 T + \kappa\tau B)] \\ &= T \cdot [\kappa^3 B + \kappa^2 \tau T] \\ &= \kappa^2 \tau \\ \tau &= \frac{\beta' \cdot [\beta'' \times \beta''']}{\kappa^2} \\ &= \frac{\beta' \cdot [\beta'' \times \beta''']}{\beta'' \cdot \beta''} \end{aligned}$$

1.24 Example Consider a circle of radius r whose equation is given by

$$\alpha(t) = (r\cos t, r\sin t, 0).$$

Then,

$$\begin{aligned} \alpha'(t) &= (-r\sin t, r\cos t, 0) \\ \|\alpha'(t)\| &= \sqrt{(-r\sin t)^2 + (r\cos t)^2 + 0^2} \\ &= \sqrt{r^2(\sin^2 t + \cos^2 t)} \\ &= r. \end{aligned}$$

Therefore, $ds/dt = r$ and $s = rt$, which we recognize as the formula for the length of an arc of circle of radius t, subtended by a central angle whose measure is t radians. We conclude that

$$\beta(s) = (r\cos\frac{s}{r}, r\sin\frac{s}{r}, 0)$$

is a unit speed reparametrization. The curvature of the circle can now be easily computed

$$\begin{aligned} T &= \beta'(s) = (-\sin\frac{s}{r}, \cos\frac{s}{r}, 0) \\ T' &= (-\frac{1}{r}\cos\frac{s}{r}, -\frac{1}{r}\sin\frac{s}{r}, 0) \\ \kappa &= \|\beta''\| = \|T'\| \\ &= \sqrt{\frac{1}{r^2}\cos^2\frac{s}{r} + \frac{1}{r^2}\sin^2\frac{s}{r} + 0^2} \\ &= \sqrt{\frac{1}{r^2}(\cos^2\frac{s}{r} + \sin^2\frac{s}{r})} \\ &= \frac{1}{r} \end{aligned}$$

This is a very simple but important example. The fact that for a circle of radius r the curvature is $\kappa = 1/r$ could not be more intuitive. A small circle has large curvature and a large circle has small curvature. As the radius of the circle approaches infinity, the circle locally looks more and more like a straight line, and the curvature approaches 0. If one were walking along a great circle on a very large sphere (like the earth) one would be perceive the space to be locally flat.

1.25 Proposition Let $\alpha(t)$ be a curve of velocity $\mathbf{v}$, acceleration a, speed v and curvature κ, then

$$\begin{aligned} \mathbf{v} &= vT, \\ \mathbf{a} &= \frac{dv}{dt}T + v^2\kappa N. \end{aligned} \tag{1.37}$$

Proof: Let $s(t)$ be the arc length and let $\beta(s)$ be a unit speed reparametrization. Then $\alpha(t) = \beta(s(t))$ and by the chain rule

$$\begin{aligned} \mathbf{v} &= \alpha'(t) \\ &= \beta'(s(t))s'(t) \\ &= vT \end{aligned}$$

$$\begin{aligned} \mathbf{a} &= \alpha''(t) \\ &= \frac{dv}{dt}T + vT'(s(t))s'(t) \\ &= \frac{dv}{dt}T + v(\kappa N)v \\ &= \frac{dv}{dt}T + v^2\kappa N \end{aligned}$$

Equation 1.37 is important in physics. The equation states that a particle moving along a curve in space feels a component of acceleration along the direction of motion whenever there is a change of speed, and a centripetal acceleration in the direction of the normal whenever it changes direction. The **centripetal acceleration** and any point is

$$a = v^3\kappa = \frac{v^2}{r}$$

where r is the radius of a circle called the **osculating circle**. The osculating circle has maximal tangential contact with the curve at the point in question. This is called contact of order 3, in the sense that the circle passes through three "consecutive points" in the curve. The osculating circle can be envisioned by a limiting process similar to that of the tangent to a curve in differential calculus. Let p be point on the curve, and let q_1 and q_2 be two nearby points. The three points uniquely determine a circle. This circle is a "secant" approximation to the tangent circle. As the points q_1 and q_2 approach the point p, the "secant" circle approaches the osculating circle. The osculating circle, as shown in figure 1.8, always lies in the the TN-plane, which by analogy is called the **osculating plane**. The physics interpretation of equation 1.37 is that as a particle moves along a curve, in some sense at an infinitesimal level, it is moving tangential to a circle, and hence, the centripetal acceleration at each point coincides with the centripetal acceleration along the osculating circle. As the points move along, the osculating circles move along with them, changing their radii appropriately.

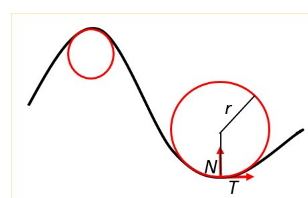


Fig. 1.8: Osculating Circle

1.26 Example (Helix)

$$
\begin{aligned}
\beta(s) &= (a\cos\frac{\omega s}{c}, a\sin\frac{\omega s}{c}, \frac{bs}{c}), \quad \text{where} \quad c=\sqrt{a^2\omega^2+b^2} \\
\beta'(s) &= (-\frac{a\omega}{c}\sin\frac{\omega s}{c}, \frac{a\omega}{c}\cos\frac{\omega s}{c}, \frac{b}{c}) \\
\beta''(s) &= (-\frac{a\omega^2}{c^2}\cos\frac{\omega s}{c}, -\frac{a\omega^2}{c^2}\sin\frac{\omega s}{c}, 0) \\
\beta'''(s) &= (\frac{a\omega^3}{c^3}\sin\frac{\omega s}{c}, -\frac{a\omega^3}{c^3}\cos\frac{\omega s}{c}, 0) \\
\kappa^2 &= \beta''\cdot\beta'' \\
&= \frac{a^2\omega^4}{c^4} \\
\kappa &= \pm\frac{a\omega^2}{c^2} \\
\tau &= \frac{(\beta'\beta''\beta''')}{\beta''\cdot\beta''} \\
&= \frac{b}{c}\begin{bmatrix} -\frac{a\omega^2}{c^2}\cos\frac{\omega s}{c} & -\frac{a\omega^2}{c^2}\sin\frac{\omega s}{c} \\ \frac{a\omega^3}{c^2}\sin\frac{\omega s}{c} & -\frac{a\omega^3}{c^2}\cos\frac{\omega s}{c} \end{bmatrix}\frac{c^4}{a^2\omega^4}. \\
&= \frac{b}{c}\frac{a^2\omega^5}{c^5}\frac{c^4}{a^2\omega^4}
\end{aligned}
$$

Simplifying the last expression and substituting the value of c, we get

$$
\begin{aligned}
\tau &= \frac{b\omega}{a^2\omega^2+b^2} \\
\kappa &= \pm\frac{a\omega^2}{a^2\omega^2+b^2}
\end{aligned}
$$

Notice that if $b=0$, the helix collapses to a circle in the xy-plane. In this case, the formulas above reduce to $\kappa = 1/a$ and $\tau = 0$. The ratio $\kappa/\tau = a\omega/b$ is particularly simple. Any curve where $\kappa/\tau = constant$ is called a helix, of which the circular helix is a special case.

1.27 Example (Plane curves) Let $\alpha(t) = (x(t), y(t), 0)$. Then

$$
\begin{aligned}
\alpha' &= (x', y', 0) \\
\alpha'' &= (x'', y'', 0) \\
\alpha''' &= (x''', y''', 0) \\
\kappa &= \frac{\|\alpha'\times\alpha''\|}{\|\alpha'\|^3} \\
&= \frac{\mid x'y''-y'x''\mid}{(x'^2+y'^2)^{3/2}} \\
\tau &= 0
\end{aligned}
$$

1.28 Example Let $\beta(s) = (x(s), y(s), 0)$, where

$$
\begin{aligned}
x(s) &= \int_0^s \cos\frac{t^2}{2c^2}\,dt \\
y(s) &= \int_0^s \sin\frac{t^2}{2c^2}\,dt. \qquad (1.38)
\end{aligned}
$$

Then, using the fundamental theorem of calculus, we have

$$\beta'(s) = (\cos\frac{s^2}{2c^2}, \sin\frac{t^2}{2c^2}, 0),$$

Since $\|\beta'\| = v = 1$, the curve is of unit speed, and s is indeed the arc length. The curvature is given by

$$\begin{aligned}
\kappa &= \|x'y'' - y'x''\| = (\beta' \cdot \beta')^{1/2} \\
&= \| -\frac{s}{c^2}\sin\frac{t^2}{2c^2}, \frac{s}{c^2}\cos\frac{t^2}{2c^2}, 0\| \\
&= \frac{s}{c^2}.
\end{aligned}$$

The integrals (1.38) are classical **Fresnel integrals** which we will discuss in more detail in the next section.

In cases where the given curve $\alpha(t)$ is not of unit speed, the following proposition provides formulas to compute the curvature and torsion in terms of α.

1.29 Proposition If $\alpha(t)$ is a regular curve in $\mathbf{R}^3$, then

$$\kappa^2 = \frac{\|\alpha' \times \alpha''\|^2}{\|\alpha'\|^6} \tag{1.39}$$

$$\tau = \frac{(\alpha'\alpha''\alpha''')}{\|\alpha' \times \alpha''\|^2}, \tag{1.40}$$

where $(\alpha'\alpha''\alpha''')$ is the triple vector product $[\alpha' \times \alpha''] \cdot \alpha'''$.
Proof:

$$\begin{aligned}
\alpha' &= vT \\
\alpha'' &= v'T + v^2\kappa N \\
\alpha''' &= (v^2\kappa)N'((s(t))s'(t) + \ldots \\
&= v^3\kappa N' + \ldots \\
&= v^3\kappa\tau B + \ldots
\end{aligned}$$

The other terms in α''' are unimportant here because $\alpha' \times \alpha''$ is proportional to B, so all we need is the B component.

$$\begin{aligned}
\alpha' \times \alpha'' &= v^3\kappa(T \times N) = v^3\kappa B \\
\|\alpha' \times \alpha''\| &= v^3\kappa \\
\kappa &= \frac{\|\alpha' \times \alpha''\|}{v^3} \\
(\alpha' \times \alpha'') \cdot \alpha''' &= v^6\kappa^2\tau \\
\tau &= \frac{(\alpha'\alpha''\alpha''')}{v^6\kappa^2} \\
&= \frac{(\alpha'\alpha''\alpha''')}{\|\alpha' \times \alpha''\|^2}
\end{aligned}$$

1.3 Fundamental Theorem of Curves

The fundamental theorem of curves basically states that prescribing a curvature and torsion as functions of some parameter s, completely determines up to position and orientation, a curve

$\beta(s)$ with that given curvature and torsion. Some geometrical insight into the significance of the curvature and torsion can be gained by considering the Taylor series expansion of an arbitrary unit speed curve $\beta(s)$ about $s = 0$.

$$\beta(s) = \beta(0) + \beta'(0)s + \frac{\beta''(0)}{2!}s^2 + \frac{\beta'''(0)}{3!}s^3 + \dots \tag{1.41}$$

Since we are assuming that s is an arc length parameter,

$$\begin{aligned} \beta'(0) &= T(0) = T_0 \\ \beta''(0) &= (\kappa N)(0) = \kappa_0 N_0 \\ \beta'''(0) &= (-\kappa^2 T + \kappa' N + \kappa\tau B)(0) = -\kappa_0^2 T_0 + \kappa_0' N_0 + \kappa_0 \tau_0 B_0 \end{aligned}$$

Keeping only the lowest terms in the components of T, N, and B, we get the first order Frenet approximation to the curve

$$\beta(s) \doteq \beta(0) + T_0 s + \frac{1}{2}\kappa_0 N_0 s^2 + \frac{1}{6}\kappa_0 \tau_0 B_0 s^3. \tag{1.42}$$

The first two terms represent the linear approximation to the curve. The first three terms approximate the curve by a parabola which lies in the osculating plane (TN-plane). If $\kappa_0 = 0$, then locally the curve looks like a straight line. If $\tau_0 = 0$, then locally the curve is a plane curve contained on the osculating plane. In this sense, the curvature measures the deviation of the curve from a straight line and the torsion (also called the second curvature) measures the deviation of the curve from a plane curve. As shown in figure 1.9 a non-planar space curve locally looks like a wire that has first been bent in a parabolic shape in the TN and twisted into a cubic along the B axis.

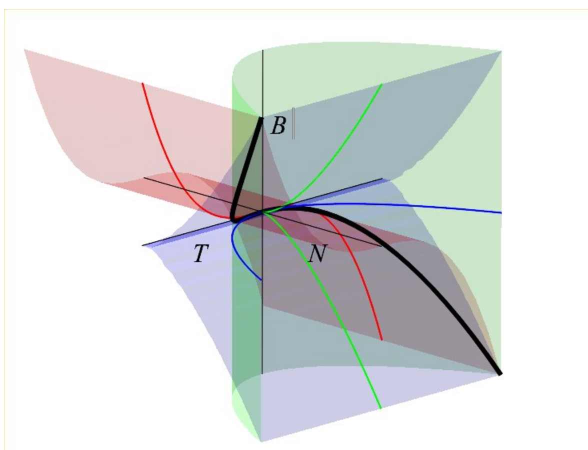


Fig. 1.9: Cubic Approximation to a Curve

So suppose that p is an arbitrary point on a curve $\beta(s)$ parametrized by arc length. We position the curve so that p is at the origin so that $\beta(0) = 0$ coincides with the point p. We chose the orthonormal basis vectors in $\mathbf{R}^3$ $\{e_1, e_2, e_3\}$ to coincide with the Frenet Frame T_0, N_0, B_0 at that point. then, the equation (1.42) provides a canonical representation of the curve near that point. This then constitutes a proof of the fundamental theorem of curves under the assumption the curve, curvature and torsion are analytic. (One could also treat the Frenet formulas as a system of differential equations and apply the conditions of existence and uniqueness of solutions for such systems.)

1.30 Proposition A curve with $\kappa = 0$ is part of a straight line.
If $\kappa = 0$ then $\beta(s) = \beta(0) + sT_0$.

1.31 Proposition A curve $\alpha(t)$ with $\tau = 0$ is a plane curve.
Proof: If $\tau = 0$, then $(\alpha'\alpha''\alpha''') = 0$. This means that the three vectors α', α'', and α''' are linearly dependent and hence there exist functions $a_1(s)$,$a_2(s)$ and $a_3(s)$ such that

$$a_3\alpha''' + a_2\alpha'' + a_1\alpha' = 0.$$

This linear homogeneous equation will have a solution of the form

$$\alpha = \mathbf{c}_1\alpha_1 + \mathbf{c}_2\alpha_2 + \mathbf{c}_3, \quad c_i = \text{constant vectors.}$$

This curve lies in the plane

$$(\mathbf{x} - \mathbf{c}_3) \cdot \mathbf{n} = 0, \text{ where } \quad \mathbf{n} = \mathbf{c}_1 \times \mathbf{c}_2$$

A consequence of the Frenet Equations is that given two curves in space C and C^* such that $\kappa(s) = \kappa^*(s)$ and $\tau(s) = \tau*(s)$, the two curves are the same up to their position in space. To clarify what we mean by their "position" we need to review some basic concepts of linear algebra leading the notion the of an isometry.

1.32 Definition Let $\mathbf{x}$ and $\mathbf{y}$ be two column vectors in $\mathbf{R}^n$ and let $\mathbf{x}^T$ represent the transpose row vector. To keep track on whether a vector is a row vector or a column vector, hereafter we write the components $\{x^i\}$ of a column vector with the indices up and the components $\{x_i\}$ of a row vector with the indices down. Similarly, if A is an $n \times n$ matrix, we write its components as $A = (a^i_j)$.

The standard **inner product** is given by matrix multiplication of the row and column vectors

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}, \tag{1.43}$$
$$= \langle \mathbf{y}, \mathbf{x} \rangle. \tag{1.44}$$

The inner product gives $\mathbf{R}^n$ the structure of a normed space by defining $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}$ and the structure of a metric space in which $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}.\|$. The real inner product is bilinear (linear in each slot), from which it follows that

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 \pm 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2, \tag{1.45}$$

and thus, we have the **polarization identity**

$$\langle \mathbf{x}, \mathbf{y} \rangle = \tfrac{1}{4}\|\mathbf{x} + \mathbf{y}\|^2 - \tfrac{1}{4}\|\mathbf{x} - \mathbf{y}\|^2. \tag{1.46}$$

The Euclidean inner product satisfies that relation

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cos\theta, \tag{1.47}$$

where θ is the angle subtended by the two vectors. Two vectors are called **orthogonal** if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, and a set of basis vectors $\mathscr{B} = \{\mathbf{e}_1, \ldots \mathbf{e}_n\}$ is called an **orthonormal** basis if $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$. Given an orthonormal basis, the dual basis is the set of linear functionals $\{\alpha^{\mathbf{i}}\}$ such that $\alpha^i(\mathbf{e}_j) = \delta^i_j$. In terms of basis components, column vectors are given by $\mathbf{x} = x^i\mathbf{e}_i$, row vectors by $\mathbf{x}^T = x_j\alpha^j$, and the inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \delta_{ij}x^iy^j = x_iy^j$.
Since $|\cos\theta| \leq 1$, there follows a special case of the **Schwarz Inequality**

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|, \tag{1.48}$$
$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle. \tag{1.49}$$

Let F be a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^n$ and $\mathscr{B} = \{\mathbf{e}_1, \dots \mathbf{e}_n\}$ be an orthonormal basis. Then, there exists a matrix $A = [F]_{\mathscr{B}}$ given by

$$A = (a^i_j) = \alpha^i(T(\mathbf{e}_j) \tag{1.50}$$

or in terms of the inner product,

$$A = (a_{ij}) = \langle \mathbf{e}_i, F(\mathbf{e})_j \rangle. \tag{1.51}$$

On the other hand, if A is a fixed $n \times n$ matrix, the map F defined by $F(\mathbf{x}) = A\mathbf{x}$ is a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^n$ whose matrix representation in the standard basis is the matrix A itself. It follows that given a linear transformation represented by a matrix A, we have

$$\begin{aligned} \langle \mathbf{x}, A\mathbf{y} \rangle &= \mathbf{x}^T A\mathbf{y}, && (1.52) \\ &= (A^T\mathbf{x})^{\mathbf{T}}\mathbf{y}, \\ &= \langle A^T\mathbf{x}, \mathbf{y} \rangle. && (1.53) \end{aligned}$$

1.33 Definition A real $n \times n$ matrix A is called orthogonal if $A^TA = AA^T = I$. The linear transformation represented by A is called an **orthogonal transformation**. Equivalently, the transformation represented by A is orthogonal if

$$\langle \mathbf{x}, A\mathbf{y} \rangle = \langle A\mathbf{x}, \mathbf{y} \rangle \tag{1.54}$$

Thus, real orthogonal transformations are represented by symmetric matrices (Hermitian in the complex case) and the condition $A^TA = I$ implies that $\det(A) = \pm 1$.

Theorem If A is an orthogonal matrix, then the transformation determined by A preserves the inner product and the norm.

Proof:

$$\begin{aligned} \langle A\mathbf{x}, A\mathbf{y} \rangle &= \langle A^TA\mathbf{x}, \mathbf{y} \rangle, \\ &= \langle \mathbf{x}, \mathbf{y} \rangle. \end{aligned}$$

Furthermore, setting $\mathbf{y} = \mathbf{x}$:

$$\begin{aligned} \langle A\mathbf{x}, A\mathbf{x} \rangle &= \langle x, x \rangle, \\ \|A\mathbf{x}\|^2 &= \|\mathbf{x}\|^2, \\ \|A\mathbf{x}\| &= \|\mathbf{x}\|. \end{aligned}$$

As a corollary, if $\{\mathbf{e}_i\}$ is an orthonormal basis, the so is $\{\mathbf{f}_i = A\mathbf{e}_i\}$. That is, an orthogonal transformation represents a rotation if $\det A = 1$ and a rotation with a reflection if $\det A = -1$.

1.34 Definition A mapping $F : \mathbf{R}^n \to \mathbf{R}^n$ called an **isometry** if it preserves distances. That is, if for all $\mathbf{x}, \mathbf{y}$

$$d(F(\mathbf{x}), F(\mathbf{y})) = d(\mathbf{x}, \mathbf{y}). \tag{1.55}$$

1.35 Example (Translations) Let $\mathbf{q}$ be fixed vector. The map $F(\mathbf{x}) = \mathbf{x} + \mathbf{q}$ is called a **translation**. It is clearly an isometry since $\|F(\mathbf{x}) - F(\mathbf{y})\| = \|\mathbf{x} + \mathbf{p} - (\mathbf{y} + \mathbf{p})\| = \|\mathbf{x} - \mathbf{y}\|$.

1.36 Theorem An orthogonal transformation is an isometry.
Proof: Let F be an isometry represented by an orthogonal matrix A. Then, since the transformation is linear and preserves norms, we have:

$$\begin{aligned} d(F(\mathbf{x}), F(\mathbf{x})) &= \|A\mathbf{x} - A\mathbf{y}\|, \\ &= \|A(\mathbf{x} - \mathbf{y})\|, \\ &= \|x - y\| \end{aligned}$$

The composition of two isometries is also an isometry. The inverse of a translation by $\mathbf{q}$ is a translation by $-\mathbf{q}$. The inverse of an orthogonal transformation represented by A is an orthogonal transformation represented by A^{-1}. Thus, the set of isometries consisting of translations and orthogonal transformations constitutes a group. Given an general isometry, we can use a translation to insure that $F(\mathbf{0}) = \mathbf{0}$. We now prove the following theorem.

1.37 Theorem If F is an isometry such that $F(\mathbf{0}) = \mathbf{0}$, then F is an orthogonal transformation.
Proof: We need to prove that F preserves the inner product and that it is linear. We first show that F preserves norms. In fact

$$\begin{aligned} \|F(\mathbf{x})\| &= d(F(\mathbf{x}), \mathbf{0}), \\ &= d(F(\mathbf{x}), F(\mathbf{0}), \\ &= d(\mathbf{x}, \mathbf{0}), \\ &= \|\mathbf{x} - \mathbf{0}\|, \\ &= \|\mathbf{x}\|. \end{aligned}$$

Now, using 1.45 and the norm preserving property above, we have:

$$\begin{aligned} d(F(\mathbf{x}), F(\mathbf{y})) &= d(\mathbf{x}, \mathbf{y}), \\ \|F(\mathbf{x}), F(\mathbf{y})\|^2 &= \|\mathbf{x} - \mathbf{y}\|^2, \\ \|F(\mathbf{x})\|^2 - 2\langle F(\mathbf{x}), F(\mathbf{y})\rangle + \|F(\mathbf{y})\|^2 &= \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2. \\ \langle F(\mathbf{x}), F(\mathbf{y})\rangle &= \langle \mathbf{x}, \mathbf{y}\rangle. \end{aligned}$$

To show F is linear, let $\mathbf{e}_i$ be an orthonormal basis, and therefore $\mathbf{f}_i = F(\mathbf{e}_1)$ is also an orthonormal basis. Then

$$\begin{aligned} F(a\mathbf{x} + b\mathbf{y}) &= \sum_{i=1}^{n} \langle F(a\mathbf{x} + b\mathbf{y}), \mathbf{f}_i\rangle \mathbf{f}_i, \\ &= \sum_{i=1}^{n} \langle F(a\mathbf{x} + b\mathbf{y}), F(\mathbf{e})_i\rangle \mathbf{f}_i, \\ &= \sum_{i=1}^{n} \langle (a\mathbf{x} + b\mathbf{y}), \mathbf{e}_i\rangle \mathbf{f}_i, \\ &= a\sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{e}_i\rangle \mathbf{f}_i + b\sum_{i=1}^{n} \langle \mathbf{y}, \mathbf{e}_i\rangle \mathbf{f}_i, \\ &= a\sum_{i=1}^{n} \langle F(\mathbf{x}), \mathbf{f}_i\rangle \mathbf{f}_i + b\sum_{i=1}^{n} \langle F(\mathbf{y}), \mathbf{f}_i\rangle \mathbf{f}_i, \\ &= aF(\mathbf{x}) + bF(\mathbf{y}). \end{aligned}$$

1.38 Theorem If $F : \mathbf{R}^n \to \mathbf{R}^n$ is an isometry then

$$F(\mathbf{x}) = A\mathbf{x} + \mathbf{q}, \tag{1.56}$$

where A is orthogonal.
Proof: If $F(\mathbf{0} = \mathbf{q}$, then $\widetilde{F} = F - \mathbf{q}$ is an isometry with $\widetilde{F}\mathbf{0} = \mathbf{0}$ and hence by the previous theorem $\widetilde{F}$ is an orthogonal transformation represented by an orthogonal matrix $\widetilde{F}\mathbf{x} = A\mathbf{x}$. It follows that $F(\mathbf{x}) = A\mathbf{x} + \mathbf{q}$.

We have just shown that any isometry is the composition of translation and an orthogonal transformation. The latter is the linear part of the isometry. The orthogonal transformation preserves the inner product, lengths, and maps orthonormal bases to orthonormal bases.

1.39 Theorem If α is a curve in $\mathbf{R}^n$ and β is the image of α under a mapping F, then vectors tangent to α get mapped to tangent vectors to β.
Proof: Let $\beta = F \circ \alpha$. The proof follows trivially from the properties of the Jacobian map $\beta_* = (F \circ \alpha)_* = F_* \alpha_*$ that takes tangent vectors to tangent vectors. If in addition F is an isometry, then F_* maps the Frenet frame of α to the Frenet frame of β.

We now have all the ingredients to prove the following:

1.40 Theorem (Fundamental theorem of Curves) If C and $\tilde{C}$ are space curves such that $\kappa(s) = \tilde{\kappa}(s)$, and $\tau(s) = \tilde{\tau}(s)$ for all s, the curves are isometric.
Proof: Given two such curves, we can perform a translation so that for some $s = s_0$ the corresponding points on C and $\tilde{C}$ are made to coincide. Without lost of generality we can make this point be the origin. Now we perform an orthogonal trasformation to make the Frenet frame $\{T_0, N_0, B_0\}$ of C coincide with the Frenet frame $\{\tilde{T}_0, \tilde{N}_0, \tilde{B}_0\}$ of $\tilde{C}$. By Schwarz inequality, the inner product of two unit vectors is 1 if and only if the vectors are equal. With this in mind, let

$$L = T \cdot \tilde{T} + N \cdot \tilde{N} + B \cdot \tilde{B}.$$

A simple computation using the Frenet equations shows that $L' = 0$, so $L =$ constant. But at $s = 0$ the Frenet frames of the two curves coincide, so the constant is 3 and this can only happen if for all s, $T = \tilde{T},\ N = \tilde{N},\ B = \tilde{B}$. Finally, since $T = \tilde{T}$, we have $\beta'(s) = \tilde{\beta}'(s)$, so $\beta(s) = \tilde{\beta}(s) +$ constant. But since $\beta(0) = \tilde{\beta}(0)$, the constant is 0 and $\beta(s) = \tilde{\beta}(s)$ for all s.

Natural Equations

The Fundamental Theorem of Curves states that up to an isometry, that is up to location and orientation, a curve is completely determined by the curvature and torsion. However The formulas for computing κ and τ are sufficiently complicated that solving the Frenet system of differential equations could be a daunting task indeed. However with the invention of modern computers, obtaining and plotting numerical solutions is a *routine* matter. There is a plethora of differential equations solvers available nowadays, including the solvers built-in into Maple, Mathematica and Matlab.

For plane curves which as we know are characterized by $\tau = 0$, it is possible to find an integral formula for the curve coordinates in terms of the curvature. Given a curve parametrized by arc length, consider an arbitrary point with position vector $\mathbf{x} = \langle x, y \rangle$ on the curve and let φ be the angle that the tangent vector T makes with the horizontal, as shown in figure 1.10. Then, the Euclidean vector components of the unit tangent vector are given by

$$\frac{d\mathbf{x}}{ds} = \mathbf{T} = \langle \cos\varphi, \sin\varphi \rangle.$$

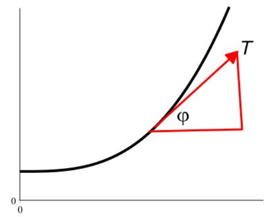


Fig. 1.10: Tangent

This means that

$$\frac{dx}{ds} = \cos\varphi, \quad \text{and} \quad \frac{dy}{ds} = \sin\varphi$$

.
From the first Frenet equation we also have

$$\frac{d\mathbf{T}}{ds} = \langle -\sin\varphi\frac{d\varphi}{ds}, \cos\varphi\frac{d\varphi}{ds}\rangle = \kappa\mathbf{N}$$

so that,

$$\|\frac{d\mathbf{T}}{ds}\| = \frac{d\varphi}{ds} = \kappa.$$

We conclude that

$$x(s) = \int \cos\varphi ds, \quad y(s) = \int \sin\varphi ds, \text{where,} \quad \varphi = \int \kappa ds. \tag{1.57}$$

Equations 1.57 are called the **Natural Equations** of a plane curve. Given the curvature κ, the equation of the curve can be obtained by "quadratures," the classical term for integrals.

1.41 Example Circle: $\kappa = 1/R$

The simplest natural equation is one where the curvature is constant. For obvious geometrical reasons we choose this constant to be $1/R$. Then, $\varphi = s/R$ and

$$\mathbf{x} = \langle R\sin\frac{s}{R}, -R\cos\frac{s}{R}\rangle,$$

which is the equation of a unit speed circle of radius R.

1.42 Example Cornu Spiral: $\kappa = \pi s$

This is the most basic linear natural equation, except for the scaling factor of π which is inserted for historical conventions. Then $\varphi = \frac{1}{2}\pi s^2$, and

$$x(s) = C(s) = \int \cos(\tfrac{1}{2}\pi s^2\, ds; \qquad y(s) = S(s) = \int \sin(\tfrac{1}{2}\pi s^2\, ds. \tag{1.58}$$

The functions $C(s)$ and $S(s)$ are called **Fresnel Integrals**. In the standard classical function libraries of Maple and Mathematica, they are listed as $FresnelC$ and $FresnelS$ respectively. The fast increasing frequency of oscillations of the integrands here make the computation prohibitive without the use of high speed computers. Graphing calculators are inadequate to render the rapid oscillations for s ranging from 0 to 15, for example, and simple computer programs for the trapezoidal rule as taught in typical calculus courses, completely falls apart in this range.

The **Cornu Spiral** is the curve $\mathbf{x}(s) = \langle x(s), y(s)\rangle$ parametrized by Fresnel integrals (See figure 1.11a). It is a tribute to the mathematicians of the 1800's that not only were they able to compute

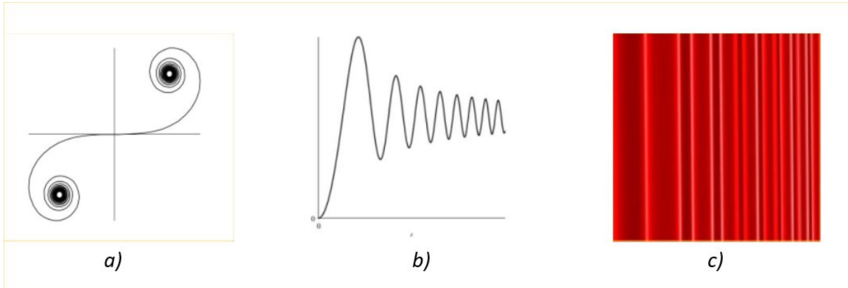


Fig. 1.11: Fresnel Diffraction

the values of the Fresnel integrals to 4 or 5 decimal places, but they did it for the range of s from 0 to 15 as mentioned above, producing remarkably accurate renditions of the spiral.

Fresnel integrals appear in the study of diffraction. If a coherent beam of light such as a laser beam, hits a sharp straight edge and a screen is placed behind, there will appear on the screen a pattern of diffraction fringes. The amplitude and intensity of the diffraction pattern can obtained by a geometrical construction involving the Fresnel integrals. First consider the function $\Psi(s) = \|\mathbf{x}\|$ that measures the distance from the origin to the points in the Cornu spiral in the first quadrant. The square of this function is then proportional to the intensity of the diffraction pattern, The graph of $|\Psi(s)|^2$ is shown in figure 1.11b. Translating this curve along an axis coinciding with that of the straight edge, generates a three dimensional surface as shown from "above" in figure 1.11c. A color scheme was used here to depict a model of the Fresnel diffraction by the straight edge.

1.43 Example Meandering Curves: $\kappa = \sin s$

A whole family of meandering curves are obtained by letting $\kappa = A \sin ks$. The meandering graph shown in picture 1.12 was obtained by numerical integration for $A = 2$ and "wave number" $k = 1$. The larger the value of A the larger the curvature of the "throats." If A is large enough, the "Throats" will overlap.

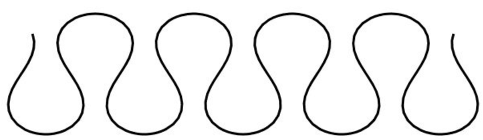


Fig. 1.12: Meandering Curve

Using superpositions of sine functions gives rise to a beautiful family of "multi-frequency" meanders with graphs that would challenge the most skillful calligraphists of the 1800's. Figure 1.13 shows a rendition with two sine functions with equal amplitude $A = 1.8$, and with $k_1 = 1,\ k_2 = 1.2$.

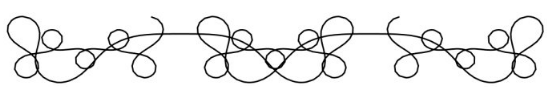


Fig. 1.13: Bimodal Meander

Chapter 2

Differential Forms

2.1 1-Forms

One of the most puzzling ideas in elementary calculus is that of the of the differential. In the usual definition, the differential of a dependent variable $y = f(x)$ is given in terms of the differential of the independent variable by $dy = f'(x)dx$. The problem is with the quantity dx. What does "dx" mean? What is the difference between Δx and dx? How much "smaller" than Δx does dx have to be? There is no trivial resolution to this question. Most introductory calculus texts evade the issue by treating dx as an arbitrarily small quantity (lacking mathematical rigor) or by simply referring to dx as an infinitesimal (a term introduced by Newton for an idea that could not otherwise be clearly defined at the time.)

In this section we introduce linear algebraic tools that will allow us to interpret the differential in terms of an linear operator.

2.1 Definition Let $\mathbf{p} \in \mathbf{R}^n$, and let $T_p(\mathbf{R}^n)$ be the tangent space at $\mathbf{p}$. A **1-form** at $\mathbf{p}$ is a linear map ϕ from $T_p(\mathbf{R}^n)$ into $\mathbf{R}$, in other words, a linear functional. We recall that such a map must satisfy the following properties:

$$\begin{array}{lll} \text{a)} & \phi(X_p) \in \mathbf{R}, & \forall X_p \in \mathbf{R}^n \\ \text{b)} & \phi(aX_p + bY_p) = a\phi(X_p) + b\phi(Y_p), & \forall a, b \in \mathbf{R},\ X_p, Y_p \in T_p(\mathbf{R}^n) \end{array} \tag{2.1}$$

A **1-form** is a smooth assignment of a linear map ϕ as above for each point in the space.

2.2 Definition Let $f : \mathbf{R}^n \to \mathbf{R}$ be a real-valued C^∞ function. We define the differential df of the function as the 1-form such that

$$df(X) = X(f) \tag{2.2}$$

for every vector field in X in $\mathbf{R}^n$.

In other words, at any point $\mathbf{p}$, the differential df of a function is an operator that assigns to a tangent vector X_p the directional derivative of the function in the direction of that vector.

$$df(X)(p) = X_p(f) = \boldsymbol{\nabla} f(p) \cdot \mathbf{X}(p) \tag{2.3}$$

In particular, if we apply the differential of the coordinate functions x^i to the basis vector fields, we get

$$dx^i(\frac{\partial}{\partial x^j}) = \frac{\partial x^i}{\partial x^j} = \delta^i_j \tag{2.4}$$

The set of all linear functionals on a vector space is called the **dual** of the vector space. It is a standard theorem in linear algebra that the dual of a vector space is also a vector space of the same dimension. Thus, the space $T_p^\star \mathbf{R}^n$ of all 1-forms at $\mathbf{p}$ is a vector space which is the dual of the tangent space $T_p\mathbf{R}^n$. The space $T_p^\star(\mathbf{R}^n)$ is called the **cotangent space** of $\mathbf{R}^n$ at the point $\mathbf{p}$. Equation (2.4) indicates that the set of differential forms $\{(dx^1)_p, \ldots, (dx^n)_p\}$ constitutes the basis of the cotangent space which is dual to the standard basis $\{(\frac{\partial}{\partial x^1})_p, \ldots (\frac{\partial}{\partial x^n})_p\}$ of the tangent space. The union of all the cotangent spaces as $\mathbf{p}$ ranges over all points in $\mathbf{R}^n$ is called the **cotangent bundle** $T^*(\mathbf{R}^n)$.

2.3 Proposition Let f be any smooth function in $\mathbf{R}^n$ and let $\{x^1, \ldots x^n\}$ be coordinate functions in a neighborhood U of a point $\mathbf{p}$. Then, the differential df is given locally by the expression

$$
\begin{aligned}
df &= \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} dx^i \\
&= \frac{\partial f}{\partial x^i} dx^i
\end{aligned} \tag{2.5}
$$

Proof: The differential df is by definition a 1-form, so, at each point, it must be expressible as a linear combination of the basis elements $\{(dx^1)_p, \ldots, (dx^n)_p\}$. Therefore, to prove the proposition, it suffices to show that the expression 2.5 applied to an arbitrary tangent vector coincides with definition 2.2. To see this, consider a tangent vector $X_p = v^j(\frac{\partial}{\partial x^j})_p$ and apply the expression above as follows:

$$
\begin{aligned}
(\frac{\partial f}{\partial x^i} dx^i)_p(X_p) &= (\frac{\partial f}{\partial x^i} dx^i)(v^j \frac{\partial}{\partial x^j})(p) \\
&= v^j(\frac{\partial f}{\partial x^i} dx^i)(\frac{\partial}{\partial x^j})(p) \\
&= v^j(\frac{\partial f}{\partial x^i} \frac{\partial x^i}{\partial x^j})(p) \\
&= v^j(\frac{\partial f}{\partial x^i} \delta^i_j)(p) \\
&= (\frac{\partial f}{\partial x^i} v^i)(p) \\
&= \boldsymbol{\nabla} f(p) \cdot \mathbf{x}(p) \\
&= df(X)(p)
\end{aligned} \tag{2.6}
$$

The definition of differentials as linear functionals on the space of vector fields is much more satisfactory than the notion of infinitesimals, since the new definition is based on the rigorous machinery of linear algebra. If α is an arbitrary 1-form, then locally

$$
\alpha = a_1(\mathbf{x}) dx^1 +, \ldots + a_n(\mathbf{x}) dx^n, \tag{2.7}
$$

where the coefficients a_i are C^∞ functions. Thus, a 1-form is a smooth section of the cotangent bundle and we refer to it as a **covariant tensor** of rank 1, or simply a **covector**. The collection of all 1-forms is denoted by $\mathscr{T}_1^0(\mathbf{R}^n)$. The coefficients $(a_1, \ldots, a_n)$ are called the **covariant** components of the covector. We will adopt the convention to always write the covariant components of a covector with the indices down. Physicists often refer to the covariant components of a 1-form as a covariant vector and this causes some confusion about the position of the indices. We emphasize that not all one forms are obtained by taking the differential of a function. If there exists a function f, such that $\alpha = df$, then the one form α is called **exact**. In vector calculus and elementary physics, exact forms are important in understanding the path independence of line integrals of conservative vector fields.

As we have already noted, the cotangent space $T_p^*(\mathbf{R}^n)$ of 1-forms at a point $\mathbf{p}$ has a natural vector space structure. We can easily extend the operations of addition and scalar multiplication to the space of all 1-forms by defining

$$\begin{aligned} (\alpha+\beta)(X) &= \alpha(X)+\beta(X) \\ (f\alpha)(X) &= f\alpha(X) \end{aligned} \tag{2.8}$$

for all vector fields X and all smooth functions f.

2.2 Tensors and Forms of Higher Rank

As we mentioned at the beginning of this chapter, the notion of the differential dx is not made precise in elementary treatments of calculus, so consequently, the differential of area $dxdy$ in $\mathbf{R}^2$, as well as the differential of surface area in $\mathbf{R}^3$ also need to be revisited in a more rigorous setting. For this purpose, we introduce a new type of multiplication between forms that not only captures the essence of differentials of area and volume, but also provides a rich algebraic and geometric structure generalizing cross products (which make sense only in $\mathbf{R}^3$) to Euclidean space of any dimension.

2.4 Definition A map $\phi : \mathscr{X}(\mathbf{R}^n)\times\mathscr{X}(\mathbf{R}^n) \longrightarrow \mathbf{R}$ is called a **bilinear** map vector fields, if it is linear on each slot. That is,

$$\begin{aligned} \phi(f^1X_1+f^2X_2,Y_1) &= f^1\phi(X_1,Y_1)+f^2\phi(X_2,Y_1) \\ \phi(X_1,f^1Y_1+f^2Y_2) &= f^1\phi(X_1,Y_1)+f^2\phi(X_1,Y_2), \quad \forall X_i,Y_i\in\mathscr{X}(\mathbf{R}^n),\ \ f^i\in\mathscr{F}\mathbf{R}^n. \end{aligned}$$

Tensor Products

2.5 Definition Let α and β be 1-forms. The **tensor product** of α and β is defined as the bilinear map $\alpha\otimes\beta$ such that

$$(\alpha\otimes\beta)(X,Y)=\alpha(X)\beta(Y) \tag{2.9}$$

for all vector fields X and Y.

Thus, for example, if $\alpha=a_i dx^i$ and $\beta=b_j dx^j$, then,

$$\begin{aligned} (\alpha\otimes\beta)(\frac{\partial}{\partial x^k},\frac{\partial}{\partial x^l}) &= \alpha(\frac{\partial}{\partial x^k})\beta(\frac{\partial}{\partial x^l}) \\ &= (a_i dx^i)(\frac{\partial}{\partial x^k})(b_j dx^j)(\frac{\partial}{\partial x^l}) \\ &= a_i\delta^i_k b_j\delta^j_l \\ &= a_k b_l. \end{aligned}$$

A quantity of the form $T=T_{ij}dx^i\otimes dx^j$ is called a **covariant tensor of rank 2,** and we may think of the set $\{dx^i\otimes dx^j\}$ as a basis for all such tensors. The space of covariant tensor fields of rank 2 is denoted $\mathscr{T}^0_2(\mathbf{R}^n)$. We must caution the reader again that there is possible confusion about the location of the indices, since physicists often refer to the components T_{ij} as a covariant tensor of rank two, as long is it satisfies some transformation laws.

In a similar fashion, one can define the tensor product of vectors X and Y as the bilinear map $X\otimes Y$ such that

$$(X\otimes Y)(f,g)=X(f)Y(g) \tag{2.10}$$

for any pair of arbitrary functions f and g.

If $X = a^i \frac{\partial}{\partial x^i}$ and $Y = b^j \frac{\partial}{\partial x^j}$, then the components of $X \otimes Y$ in the basis $\frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}$ are simply given by $a^i b^j$. Any bilinear map of the form

$$T = T^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \tag{2.11}$$

is called a contravariant tensor of rank 2 in $\mathbf{R}^n$.

The notion of tensor products can easily be generalized to higher rank, and in fact one can have tensors of mixed ranks. For example, a tensor of contravariant rank 2 and covariant rank 1 in $\mathbf{R}^n$ is represented in local coordinates by an expression of the form

$$T = T^{ij}{}_k \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \otimes dx^k.$$

This object is also called a tensor of type T^2_1. Thus, we may think of a tensor of type T^2_1 as map with three input slots. The map expects two functions in the first two slots and a vector in the third one. The action of the map is bilinear on the two functions and linear on the vector. The output is a real number. The set $T^r_s|_p(\mathbf{R}^n)$ of all tensors of type T^r_s at a point p has a vector space structure. The union of all such vector spaces is called the **tensor bundle**, and smooth sections of the bundle are called **tensor fields** $\mathscr{T}^r_s(\mathbf{R}^n)$; that is, a tensor field is a smooth assignment of a tensor to each point in $\mathbf{R}^n$.

Inner Products

Let $X = a^i \frac{\partial}{\partial x^i}$ and $Y = b^j \frac{\partial}{\partial x^j}$ be two vector fields and let

$$g(X,Y) = \delta_{ij} a^i b^j. \tag{2.12}$$

The quantity $g(X,Y)$ is an example of a bilinear map that the reader will recognize as the usual dot product.

2.6 Definition A bilinear map $g(X,Y)$ on vectors is called a real **inner product** if

1. $g(X,Y) = g(Y,X)$,
2. $g(X,X) \geq 0, \quad \forall X$,
3. $g(X,X) = 0$ iff $X = 0$.

Since we assume $g(X,Y)$ to be bilinear, an inner product is completely specified by its action on ordered pairs of basis vectors. The components g_{ij} of the inner product are thus given by

$$g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}) = g_{ij}, \tag{2.13}$$

where g_{ij} is a symmetric $n \times n$ matrix which we assume to be non-singular. By linearity, it is easy to see that if $X = a^i \frac{\partial}{\partial x^i}$ and $Y = b^j \frac{\partial}{\partial x^j}$ are two arbitrary vectors, then

$$g(X,Y) = g_{ij} a^i b^j.$$

In this sense, an inner product can be viewed as a generalization of the dot product. The standard Euclidean inner product is obtained if we take $g_{ij} = \delta_{ij}$. In this case, the quantity $g(X,X) = \| X \|^2$ gives the square of the length of the vector. For this reason g_{ij} is called a **metric** and g is called a **metric tensor**.

Another interpretation of the dot product can be seen if instead one considers a vector $X = a^i \frac{\partial}{\partial x^i}$ and a 1-form $\alpha = b_j dx^j$. The action of the 1-form on the vector gives

$$\begin{aligned}
\alpha(X) &= (b_j dx^j)(a^i \frac{\partial}{\partial x^i}) \\
&= b_j a^i (dx^j)(\frac{\partial}{\partial x^i}) \\
&= b_j a^i \delta^j_i \\
&= a^i b_i.
\end{aligned}$$

If we now define

$$b_i = g_{ij} b^j, \tag{2.14}$$

we see that the equation above can be rewritten as

$$a^i b_j = g_{ij} a^i b^j,$$

and we recover the expression for the inner product.

Equation (2.14) shows that the metric can be used as a mechanism to lower indices, thus transforming the contravariant components of a vector to covariant ones. If we let g^{ij} be the inverse of the matrix g_{ij}, that is

$$g^{ik} g_{kj} = \delta^i_j, \tag{2.15}$$

we can also raise covariant indices by the equation

$$b^i = g^{ij} b_j. \tag{2.16}$$

We have mentioned that the tangent and cotangent spaces of Euclidean space at a particular point p are isomorphic. In view of the above discussion, we see that the metric accepts a dual interpretation; one as a bilinear pairing of two vectors

$$g : T_p(\mathbf{R}^n) \times T_p(\mathbf{R}^n) \longrightarrow \mathbf{R},$$

and another as a linear isomorphism

$$g : T^\star_p(\mathbf{R}^n) \longrightarrow T_p(\mathbf{R}^n)$$

that maps vectors to covectors and vice-versa.

In elementary treatments of calculus, authors often ignore the subtleties of differential 1-forms and tensor products and define the differential of arclength as

$$ds^2 \equiv g_{ij} dx^i dx^j,$$

although what is really meant by such an expression is

$$ds^2 \equiv g_{ij} dx^i \otimes dx^j. \tag{2.17}$$

2.7 Example In cylindrical coordinates, the differential of arclength is

$$ds^2 = dr^2 + r^2 d\theta^2 + dz^2. \tag{2.18}$$

In this case, the metric tensor has components

$$g_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.19}$$

2.8 Example In spherical coordinates,

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta, \end{aligned} \tag{2.20}$$

and the differential of arclength is given by

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2. \tag{2.21}$$

In this case the metric tensor has components

$$g_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin\theta^2 \end{bmatrix}. \tag{2.22}$$

Minkowski Space

An important object in mathematical physics is the so-called **Minkowski space** which is defined as the pair $(\mathcal{M}_{1,3}, \eta)$, where

$$\mathcal{M}_{(1,3)} = \{(t, x^1, x^2, x^3) |\ t, x^i \in \mathbf{R}\} \tag{2.23}$$

and η is the bilinear map such that

$$\eta(X, X) = -t^2 - (x^1)^2 - (x^2)^2 - (x^3)^2. \tag{2.24}$$

The matrix representing Minkowski's metric η is given by

$$\eta = \mathrm{diag}(1, -1, -1, -1),$$

in which case, the differential of arclength is given by

$$\begin{aligned} ds^2 &= \eta_{\mu\nu} dx^\mu \otimes dx^\nu \\ &= dt \otimes dt - dx^1 \otimes dx^1 - dx^2 \otimes dx^2 - dx^3 \otimes dx^3 \\ &= dt^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2. \end{aligned} \tag{2.25}$$

Note: Technically speaking, Minkowski's metric is not really a metric since $\eta(X, X) = 0$ does not imply that $X = 0$. Non-zero vectors with zero length are called **Light-like** vectors and they are associated with with particles that travel at the speed of light (which we have set equal to 1 in our system of units.)

The Minkowski metric $g_{\mu\nu}$ and its matrix inverse $g^{\mu\nu}$ are also used to raise and lower indices in the space in a manner completely analogous to $\mathbf{R}^n$. Thus, for example, if A is a covariant vector with components

$$A_\mu = (\rho, A_1, A_2, A_3),$$

then the contravariant components of A are

$$\begin{aligned} A^\mu &= \eta^{\mu\nu} A_\nu \\ &= (-\rho, A_1, A_2, A_3) \end{aligned}$$

Wedge Products and n-Forms

2.9 Definition A map $\phi : T(\mathbf{R}^n) \times T(\mathbf{R}^n) \longrightarrow \mathbf{R}$ is called **alternating** if

$$\phi(X,Y) = -\phi(Y,X).$$

The alternating property is reminiscent of determinants of square matrices that change sign if any two column vectors are switched. In fact, the determinant function is a perfect example of an alternating bilinear map on the space $M_{2\times 2}$ of two by two matrices. Of course, for the definition above to apply, one has to view $M_{2\times 2}$ as the space of column vectors.

2.10 Definition A **2-form** ϕ is a map $\phi : T(\mathbf{R}^n) \times T(\mathbf{R}^n) \longrightarrow \mathbf{R}$ which is alternating and bilinear.

2.11 Definition Let α and β be 1-forms in $\mathbf{R}^n$ and let X and Y be any two vector fields. The **wedge product** of the two 1-forms is the map $\alpha \wedge \beta : T(\mathbf{R}^n) \times T(\mathbf{R}^n) \longrightarrow \mathbf{R}$, given by the equation

$$(\alpha \wedge \beta)(X,Y) = \alpha(X)\beta(Y) - \alpha(Y)\beta(X). \tag{2.26}$$

2.12 Theorem If α and β are 1-forms, then $\alpha \wedge \beta$ is a 2-form.
Proof: We break up the proof into the following two lemmas:

2.13 Lemma The wedge product of two 1-forms is alternating.
Proof: Let α and β be 1-forms in $\mathbf{R}^n$ and let X and Y be any two vector fields. Then

$$\begin{aligned}
(\alpha \wedge \beta)(X,Y) &= \alpha(X)\beta(Y) - \alpha(Y)\beta(X) \\
&= -(\alpha(Y)\beta(X) - \alpha(X)\beta(Y)) \\
&= -(\alpha \wedge \beta)(Y,X).
\end{aligned}$$

2.14 Lemma The wedge product of two 1-forms is bilinear.
Proof: Consider 1-forms, α, β, vector fields X_1, X_2, Y and functions f^1, f^2. Then, since the 1-forms are linear functionals, we get

$$\begin{aligned}
(\alpha \wedge \beta)(f^1X_1 + f^2X_2, Y) &= \alpha(f^1X_1 + f^2X_2)\beta(Y) - \alpha(Y)\beta(f^1X_1 + f^2X_2) \\
&= [f^1\alpha(X_1) + f^2\alpha(X_2)]\beta(Y) - \alpha(Y)[f^1\beta(X_1) + f^2\alpha(X_2)] \\
&= f^1\alpha(X_1)\beta(Y) + f^2\alpha(X_2)\beta(Y) + f^1\alpha(Y)\beta(X_1) + f^2\alpha(Y)\beta(X_2) \\
&= f^1[\alpha(X_1)\beta(Y) + \alpha(Y)\beta(X_1)] + f^2[\alpha(X_2)\beta(Y) + \alpha(Y)\beta(X_2)] \\
&= f^1(\alpha \wedge \beta)(X_1, Y) + f^2(\alpha \wedge \beta)(X_2, Y).
\end{aligned}$$

The proof of linearity on the second slot is quite similar and is left to the reader.

2.15 Corollary If α and β are 1-forms, then

$$\alpha \wedge \beta = -\beta \wedge \alpha. \tag{2.27}$$

This last result tells us that wedge products have characteristics similar to cross products of vectors in the sense that both of these products anti-commute. This means that we need to be careful to introduce a minus sign every time we interchange the order of the operation. Thus, for example, we have

$$dx^i \wedge dx^j = -dx^j \wedge dx^i$$

if $i \neq j$, whereas

$$dx^i \wedge dx^i = -dx^i \wedge dx^i = 0$$

since any quantity that equals the negative of itself must vanish. The similarity between wedge products is even more striking in the next proposition but we emphasize again that wedge products are much more powerful than cross products, because wedge products can be computed in any dimension.

2.16 Proposition Let $\alpha = A_i dx^i$ and $\beta = B_i dx^i$ be any two 1-forms in $\mathbf{R}^n$. Then

$$\alpha \wedge \beta = (A_i B_j) dx^i \wedge dx^j. \tag{2.28}$$

Proof: Let X and Y be arbitrary vector fields. Then

$$\begin{aligned}
(\alpha \wedge \beta)((X,Y) &= (A_i dx^i)(X)(B_j dx^j)(Y) - (A_i dx^i)(Y)(B_j dx^j)(X) \\
&= (A_i B_j)[dx^i(X) dx^j(Y) - dx^i(Y) dx^j(X)] \\
&= (A_i B_j)(dx^i \wedge dx^j)(X,Y).
\end{aligned}$$

Because of the antisymmetry of the wedge product, the last of the above equations can be written as

$$\alpha \wedge \beta = \sum_{i=1}^{n} \sum_{j<i}^{n} (A_i B_j - A_j B_i)(dx^i \wedge dx^j).$$

In particular, if $n = 3$, then the coefficients of the wedge product are the components of the cross product of $\mathbf{A} = A_1 \mathbf{i} + A_2 \mathbf{j} + A_3 \mathbf{k}$ and $\mathbf{B} = B_1 \mathbf{i} + B_2 \mathbf{j} + B_3 \mathbf{k}$.

2.17 Example Let $\alpha = x^2 dx - y^2 dy$ and $\beta = dx + dy - 2xydz$. Then

$$\begin{aligned}
\alpha \wedge \beta &= (x^2 dx - y^2 dy) \wedge (dx + dy - 2xydz) \\
&= x^2 dx \wedge dx + x^2 dx \wedge dy - 2x^3 y dx \wedge dz - y^2 dy \wedge dx - y^2 dy \wedge dy + 2xy^3 dy \wedge dz \\
&= x^2 dx \wedge dy - 2x^3 y dx \wedge dz - y^2 dy \wedge dx + 2xy^3 dy \wedge dz \\
&= (x^2 + y^2) dx \wedge dy - 2x^3 y dx \wedge dx + 2xy^3 dy \wedge dz.
\end{aligned}$$

2.18 Example Let $x = r\cos\theta$ and $y = r\sin\theta$. Then

$$\begin{aligned}
dx \wedge dy &= (-r\sin\theta d\theta + \cos\theta dr) \wedge (r\cos\theta d\theta + \sin\theta dr) \\
&= -r\sin^2\theta d\theta \wedge dr + r\cos^2\theta dr \wedge d\theta \\
&= (r\cos^2\theta + r\sin^2\theta)(dr \wedge d\theta) \\
&= r(dr \wedge d\theta). \qquad (2.29)
\end{aligned}$$

2.19 Remark

1. The result of the last example yields the familiar differential of area in polar coordinates.
2. The differential of area in polar coordinates is a special example of the change of coordinate theorem for multiple integrals. It is easy to establish that if $x = f^1(u,v)$ and $y = f^2(u,v)$, then $dx \wedge dy = \|J\| du \wedge dv$, where $\|J\|$ is the determinant of the Jacobian of the transformation.
3. Quantities such as $dxdy$ and $dydz$ which often appear in calculus, are not well defined. In most cases, these entities are actually wedge products of 1-forms.

4. We state without proof that all 2-forms ϕ in $\mathbf{R}^n$ can be expressed as linear combinations of wedge products of differentials such as

$$\phi = F_{ij} dx^i \wedge dx^j. \tag{2.30}$$

In a more elementary (ie, sloppier) treatment of this subject one could simply define 2-forms to be gadgets which look like the quantity in equation (2.30). With this motivation, we introduce the next definition.

2.20 Definition A 3-form ϕ in $\mathbf{R}^n$ is an object of the following type

$$\phi = A_{ijk} dx^i \wedge dx^j \wedge dx^k \tag{2.31}$$

where we assume that the wedge product of three 1-forms is associative but alternating in the sense that if one switches any two differentials, then the entire expression changes by a minus sign. We challenge the reader to come up with a rigorous definition of three forms (or an n-form, for that matter) in the spirit of multilinear maps. There is nothing really wrong with using definition (2.31). This definition however, is coordinate-dependent and differential geometers prefer coordinate-free definitions, theorems and proofs.

Now for a little combinatorics. Let us count the number of linearly independent differential forms in Euclidean space. More specifically, we want to find the vector space dimension of the space of k-forms in $\mathbf{R}^n$. We will think of 0-forms as being ordinary functions. Since functions are the "scalars", the space of 0-forms as a vector space has dimension 1.

$\mathbf{R}^2$	Forms	Dim
0-forms	f	1
1-forms	$f dx^1, g dx^2$	2
2-forms	$f dx^1 \wedge dx^2$	1

$\mathbf{R}^3$	Forms	Dim
0-forms	f	1
1-forms	$f_1 dx^1, f_2 dx^2, f_3 dx^3$	3
2-forms	$f_1 dx^2 \wedge dx^3,\ f_2 dx^3 \wedge dx^1,\ f_3 dx^1 \wedge dx^2$	3
3-forms	$f_1 dx^1 \wedge dx^2 \wedge dx^3$	1

The binomial coefficient pattern should be evident to the reader.

2.3 Exterior Derivatives

In this section we introduce a differential operator that generalizes the classical gradient, curl and divergence operators.

Denote by $\bigwedge^m_{(p)}(\mathbf{R}^n)$ the space of m-forms at $\mathbf{p} \in \mathbf{R}^n$. This vector space has dimension

$$\dim \bigwedge\nolimits^m_{(p)}(\mathbf{R}^n) = \binom{n}{m} = \frac{n!}{m!(n-m)!}$$

for $m \leq n$ and dimension 0 for $m > n$. We will identify $\bigwedge^0_{(p)}(\mathbf{R}^n)$ with the space of $\mathcal{C}^\infty$ functions at $\mathbf{p}$. The union of all $\bigwedge^m_{(p)}(\mathbf{R}^n)$ as $\mathbf{p}$ ranges through all points in $\mathbf{R}^n$ is called the **bundle of m-forms** and will be denoted by $\bigwedge^m(\mathbf{R}^n)$.

A section α of the bundle

$$\bigwedge\nolimits^m(\mathbf{R}^n) = \bigcup_p \bigwedge\nolimits^m_p(\mathbf{R}^n).$$

is called an **m-form** and it can be written as:

$$\alpha = A_{i_1,\ldots i_m}(x) dx^{i_1} \wedge \ldots dx^{i_m}. \tag{2.32}$$

.

2.21 Definition Let α be a one form in $\mathbf{R}^n$. The differential $d\alpha$ is the two form defined by

$$d\alpha(X,Y) = X(\alpha(Y)) - Y(\alpha(X)), \tag{2.33}$$

for any pair of vector fields X and Y

To explore the meaning of this definition in local coordinates, let $\alpha = f_i dx^i$ and let $X = \frac{\partial}{\partial x^j}$, $Y = \frac{\partial}{\partial x^k}$, then

$$\begin{aligned}
d\alpha(X,Y) &= \frac{\partial}{\partial x^j}\left[f_i dx^i\left(\frac{\partial}{\partial x^k}\right)\right] - \frac{\partial}{\partial x^k}\left[f_i dx^i\left(\frac{\partial}{\partial x^j}\right)\right], \\
&= \frac{\partial}{\partial x^j}(f_i \delta^i_k) - \frac{\partial}{\partial x^k}(f_i \delta^i_j), \\
d\alpha\left(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^j}\right) &= \frac{\partial f_k}{\partial x^j} - \frac{\partial f_j}{\partial x^k}
\end{aligned}$$

Therefore, taking into account the antisymmetry of wedge products, we have.

$$\begin{aligned}
d\alpha &= \frac{1}{2}\left(\frac{\partial f_k}{\partial x^j} - \frac{\partial f_j}{\partial x^k}\right) dx^j \wedge dx^k, \\
&= \frac{\partial f_i k}{\partial x^j} dx^j \wedge dx^k, \\
&= df_k \wedge dx^k.
\end{aligned}$$

We should mention that the definition of the differential of a 1-form above will need to be refined later to accommodate more general manifolds (see 6.19.) It is also possible to provide a coordinate-free definition of the differential of a an m-form, but again, we will leave this for the chapter on Riemannian manifolds. For now, it suffices to use the computation above to motivate the following coordinate dependent definition:

2.22 Definition Let α be an m-form (given in coordinates as in equation (2.32)). The **exterior derivative** of α is the (m+1-form) $d\alpha$ given by

$$\begin{aligned}
d\alpha &= dA_{i_1,\ldots i_m} \wedge dx^{i_1} \ldots dx^{i_m} \\
&= \frac{\partial A_{i_1,\ldots i_m}}{\partial x^{i_0}}(x) dx^{i_0} \wedge dx^{i_1} \ldots dx^{i_m}.
\end{aligned} \tag{2.34}$$

In the special case where α is a 0-form, that is, a function, we write

$$df = \frac{\partial f}{\partial x^i} dx^i.$$

2.23 Theorem

$$\begin{aligned}
&\text{a)} \quad d : \textstyle\bigwedge^m \longrightarrow \bigwedge^{m+1} \\
&\text{b)} \quad d^2 = d \circ d = 0 \\
&\text{c)} \quad d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta \qquad \forall \alpha \in \textstyle\bigwedge^p, \beta \in \bigwedge^q
\end{aligned} \tag{2.35}$$

Proof:

a) Obvious from equation (2.34).

b) First, we prove the proposition for $\alpha = f \in \bigwedge^0$. We have

$$\begin{aligned} d(d\alpha) &= d(\frac{\partial f}{\partial x^i}) \\ &= \frac{\partial^2 f}{\partial x^j \partial x^i} dx^j \wedge dx^i \\ &= \frac{1}{2}[\frac{\partial^2 f}{\partial x^j \partial x^i} - \frac{\partial^2 f}{\partial x^i \partial x^j}] dx^j \wedge dx^i \\ &= 0. \end{aligned}$$

Now, suppose that α is represented locally as in equation (2.32). It follows from equation 2.34 that

$$d(d\alpha) = d(dA_{i_1,\ldots i_m}) \wedge dx^{i_0} \wedge dx^{i_1} \ldots dx^{i_m} = 0.$$

c) Let $\alpha \in \bigwedge^p, \beta \in \bigwedge^q$. Then, we can write

$$\begin{aligned} \alpha &= A_{i_1,\ldots i_p}(x) dx^{i_1} \wedge \ldots dx^{i_p} \\ \beta &= B_{j_1,\ldots j_q}(x) dx^{j_1} \wedge \ldots dx^{j_q}. \end{aligned} \tag{2.36}$$

By definition,

$$\alpha \wedge \beta = A_{i_1 \ldots i_p} B_{j_1 \ldots j_q} (dx^{i_1} \wedge \ldots \wedge dx^{i_p}) \wedge (dx^{j_1} \wedge \ldots \wedge dx^{j_q}).$$

Now, we take take the exterior derivative of the last equation, taking into account that $d(fg) = fdg + gdf$ for any functions f and g. We get

$$\begin{aligned} d(\alpha \wedge \beta) &= [d(A_{i_1 \ldots i_p}) B_{j_1 \ldots j_q} + (A_{i_1 \ldots i_p} d(B_{j_1 \ldots j_q})](dx^{i_1} \wedge \ldots \wedge dx^{i_p}) \wedge (dx^{j_1} \wedge \ldots \wedge dx^{j_q}) \\ &= [dA_{i_1 \ldots i_p} \wedge (dx^{i_1} \wedge \ldots \wedge dx^{i_p})] \wedge [B_{j_1 \ldots j_q} \wedge (dx^{j_1} \wedge \ldots \wedge dx^{j_q})] + \\ &= [A_{i_1 \ldots i_p} \wedge (dx^{i_1} \wedge \ldots \wedge dx^{i_p})] \wedge (-1)^p [dB_{j_1 \ldots j_q} \wedge (dx^{j_1} \wedge \ldots \wedge dx^{j_q})] \\ &= d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta. \end{aligned} \tag{2.37}$$

The $(-1)^p$ factor comes into play since in order to pass the term $dB_{j_i \ldots j_p}$ through p 1-forms of the type dx^i, one must perform p transpositions.

2.24 Definition Let $F : \mathbf{R}^n \to \mathbf{R}^m$ be differentiable mapping and let α be a k-form in $\mathbf{R}^m$. Then, at each point $y \in \mathbf{R}^m$ with $y = F(x)$, the mapping F induces a map called the **pull-back** $F^* : \bigwedge^k_{(F(x))} \to \bigwedge^k_{(x)}$ defined by

$$(F^*\alpha)_x(X_1, \ldots X_k) = \alpha_{F(x)}(F_*X_1, \ldots F_*X_k), \tag{2.38}$$

for any tangent vectors $\{X_1, \ldots X_k\}$ in $\mathbf{R}^n$.
If g is a 0-form, namely a function, $F^*(g) = g \circ F$. we have the following theorem.

2.25 Theorem

$$\begin{array}{lrcl} \text{a)} & F^*(g\alpha_1) &=& (g \circ F)\, F^*\alpha, \\ \text{b)} & F^*(\alpha_1 + \alpha_2) &=& F * \alpha_1 + F^*\alpha_2, \\ \text{c)} & F^*(\alpha \wedge \beta) &=& F^*\alpha \wedge F^*\beta, \\ \text{d)} & F^*(d\alpha) &=& d(F^*\alpha.) \end{array} \tag{2.39}$$

$$
\begin{tikzcd}
\bigwedge^k(\mathbf{R}^n) \arrow[d, "d"] & \bigwedge^k(\mathbf{R}^m) \arrow[l, "F^*"'] \arrow[d, "d"] \\
\bigwedge^{k+1}(\mathbf{R}^n) & \bigwedge^{k+1}(\mathbf{R}^m) \arrow[l, "F^*"']
\end{tikzcd}
$$

Fig. 2.1: $d\,F^* = F^*\,d$

Part (d) is encapsulated in the commuting diagram in figure 2.1.

Proof: Part (a) is basically the definition for the case of 0-forms and part (b) is clear from the linearity of the push-forward. We leave part (c) as an exercise and prove part (d). In the case of a 0-form, let g, be a function and X a vector field in $\mathbf{R}^m$. By a simple computation that amounts to recycling definitions, we have:

$$
\begin{aligned}
d(F^*g) &= d(g \circ F), \\
(F^*dg)(X) &= dg(F_*X) = (F_*X)(g), \\
&= X(g \circ F) = d(g \circ F)(X), \\
F^*dg &= d(g \circ F),
\end{aligned}
$$

so, $d(F^*g) = F^*(dg)$ is true by the composite mapping theorem. Let α be a k-form

$$\alpha = A_{i_1,\ldots i_k}\, dy^{i_1} \wedge \ldots dy^{i_k},$$

so that

$$d\alpha = (dA_{i_1,\ldots i_k}) \wedge dy^{i_1} \ldots dy^{i_k}.$$

Then, by part (c),

$$
\begin{aligned}
F^*\alpha &= (F^*A_{i_1,\ldots i_k})\, f^*dy^{i_1} \wedge \ldots f^*dy^{i_k}, \\
d(F^*\alpha) &= dF^*(A_{i_1,\ldots i_k}) \wedge f^*dy^{i_1} \wedge \ldots f^*dy^{i_k}, \\
&= F^*(dA_{i_1,\ldots i_k}) \wedge f^*dy^{i_1} \wedge \ldots f^*dy^{i_k}, \\
&= F^*(d\alpha).
\end{aligned}
$$

So again, the result rests on the chain rule.

To connect with advanced calculus, suppose that locally the mapping F is given by $y^k = f^k(x^i)$. Recalling as in 1.11 that in coordinates the push-forward on vectors is just the linear transformation represented by the Jacobian, we see that the pullback for the basis 1-forms dy^k is another manifestation of the chain rule

$$dy^k = \frac{\partial y^k}{\partial x^i}\, dx^i. \tag{2.40}$$

In particular, if $m = n$ and Ω is the volume form $\Omega = dy^1 \wedge \ldots \wedge dy^n$, the pullback $F^*\Omega$ gives rise to the Jacobian factor that appears in the chain of variables theorem for integration.

If $F : \mathbf{R}^n \to \mathbf{R}^n$ is a diffeomorphism, the push-forward is defined by $F_* = (F^{-1})^*$.

2.26 Example Let $x = r\cos\theta$ and $y = r\sin\theta$. Then

$$\begin{aligned}
dx \wedge dy &= (-r\sin\theta d\theta + \cos\theta dr) \wedge (r\cos\theta d\theta + \sin\theta dr) \\
&= -r\sin^2\theta d\theta \wedge dr + r\cos^2\theta dr \wedge d\theta \\
&= (r\cos^2\theta + r\sin^2\theta)(dr \wedge d\theta) \\
&= r(dr \wedge d\theta). \qquad (2.41)
\end{aligned}$$

2.27 Remark

1. The result of the last example yields the familiar differential of area in polar coordinates.
2. The differential of area in polar coordinates is of course a special example of the change of coordinate theorem for multiple integrals as indicated above. Using the chain rule and the antisymmetric properties of the wedge product, it follows immediately that if $x = f^1(u,v)$ and $y = f^2(u,v)$, then $dx \wedge dy = \|J\| du \wedge dv$, where $\|J\|$ is the determinant of the Jacobian of the transformation.
3. Quantities such as $dxdy$ and $dydz$ which often appear in calculus, are not really well defined. What is meant by them are actually wedge products of 1-forms, but in reversing the order of integration, the antisymmetry of the wedge product is ignored. In performing surface integrals, however, the surfaces must be considered oriented surfaces and one has to insert a negative sign in the differential of surface area component in the xz-plane as show later in 2.43.

2.28 Example Let $\alpha = P(x,y)\,dx + Q(x,y)\,dy$. Then,

$$\begin{aligned}
d\alpha &= (\tfrac{\partial P}{\partial x} dx + \tfrac{\partial P}{\partial y}) \wedge dx + (\tfrac{\partial Q}{\partial x} dx + \tfrac{\partial Q}{\partial y}) \wedge dy \\
&= \tfrac{\partial P}{\partial y}\, dy \wedge dx + \tfrac{\partial Q}{\partial x}\, dx \wedge dy \\
&= (\tfrac{\partial Q}{\partial x} - \tfrac{\partial P}{\partial y})\, dx \wedge dy. \qquad (2.42)
\end{aligned}$$

This example is related to Green's theorem in $\mathbf{R}^2$.

2.29 Example Let $\alpha = M(x,y)dx + N(x,y)dy$, and suppose that $d\alpha = 0$. Then, by the previous example,

$$d\alpha = (\tfrac{\partial N}{\partial x} - \tfrac{\partial M}{\partial y})\, dx \wedge dy.$$

Thus, $d\alpha = 0$ iff $N_x = M_y$, which implies that $N = f_y$ and M_x for some function $f(x,y)$. Hence,

$$\alpha = f_x\, dx + f_y\, dy = df.$$

The reader should also be familiar with this example in the context of exact differential equations of first order and conservative force fields.

2.30 Definition A differential form α is called **closed** if $d\alpha = 0$.

2.31 Definition A differential form α is called **exact** if there exists a form β such that $\alpha = d\beta$.

Since $d \circ d = 0$, it is clear that an exact form is also closed. For the converse to be true, one must require a topological condition that the space be contractible. A contractible space is one that can be deformed continuously to an interior point.

2.32 Poincare's Lemma In a contractible space (such as $\mathbf{R}^n$), if a differential is closed, then it is exact.

The standard counterexample showing that the topological condition in Poincare's Lemma is needed is the form $d\theta$ where θ is the polar coordinate angle in the plane. It is not hard to prove that

$$d\theta = \frac{-ydx + xdy}{x^2 + y^2}$$

is closed, but not exact in the punctured plane $\{\mathbf{R}^2 \setminus \{0\}\}$. We postpone the proof of Poincare's lemma to the chapter on Riemannian manifolds.

2.4 The Hodge $\star$ Operator

An important lesson students learn in linear algebra is that all vector spaces of finite dimension n are isomorphic to each other. Thus, for instance, the space P_3 of all real polynomials in x of degree 3, and the space $\mathcal{M}_{2\times 2}$ of real 2 by 2 matrices are, in terms of their vector space properties, basically no different from the Euclidean vector space $\mathbf{R}^4$. A good example of this is the tangent space $T_p\mathbf{R}^3$ which has dimension 3. The process of replacing $\frac{\partial}{\partial x}$ by $\mathbf{i}$, $\frac{\partial}{\partial y}$ by $\mathbf{j}$ and $\frac{\partial}{\partial z}$ by $\mathbf{k}$ is a linear, 1-1 and onto map that sends the "vector" part of a tangent vector $a^1\frac{\partial}{\partial x} + a^2\frac{\partial}{\partial y} + a^3\frac{\partial}{\partial z}$ to regular Euclidean vector $\langle a^1, a^2, a^3\rangle$.

We have also observed that the tangent space $T_p\mathbf{R}^n$ is isomorphic to the cotangent space $T_p^\star\mathbf{R}^n$. In this case, the vector space isomorphism maps the standard basis vectors $\{\frac{\partial}{\partial x^i}\}$ to their duals $\{dx^i\}$. This isomorphism then transforms a contravariant vector to a covariant vector. In terms of components, the isomorphism is provided by the Euclidean metric that maps the components of a contravariant vector with indices up to a covariant vector with indices down.

Another interesting example is provided by the spaces $\bigwedge_p^1(\mathbf{R}^3)$ and $\bigwedge_p^2(\mathbf{R}^3)$, both of which have dimension 3. It follows that these two spaces must be isomorphic. In this case the isomorphism is given as follows:

$$\begin{aligned} dx &\longmapsto dy \wedge dz \\ dy &\longmapsto -dx \wedge dz \\ dz &\longmapsto dx \wedge dy \end{aligned} \tag{2.43}$$

More generally, we have seen that the dimension of the space of m-forms in $\mathbf{R}^n$ is given by the binomial coefficient $\binom{n}{m}$. Since

$$\binom{n}{m} = \binom{n}{n-m} = \frac{n!}{(n-m)!},$$

it must be true that

$$\textstyle\bigwedge_p^m(\mathbf{R}^n) \cong \bigwedge_p^{n-m}(\mathbf{R}^n). \tag{2.44}$$

To describe the isomorphism between these two spaces, we will first need to introduce the totally antisymmetric **Levi-Civita** permutation symbol defined as follows:

$$\epsilon_{i_1\dots i_m} = \begin{cases} +1 & \text{if } (i_1,\dots,i_m) \text{ is an even permutation of}(1,\dots,m) \\ -1 & \text{if } (i_1,\dots,i_m) \text{ is an odd permutation of}(1,\dots,m) \\ 0 & \text{otherwise} \end{cases} \tag{2.45}$$

In dimension 3, there are only 3 ($3! = 6$) non-vanishing components of $\epsilon_{i,j,k}$ in

$$\begin{aligned} \epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1 \\ \epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1 \end{aligned} \tag{2.46}$$

The permutation symbols are useful in the theory of determinants. In fact, if $A = (a^i_j)$ is a 3×3 matrix, then, using equation (2.46), the reader can easily verify that

$$\det A = \|A\| = \epsilon_{i_1 i_2 i_3} a^{i_1}_1 a^{i_2}_2 a^{i_3}_3. \tag{2.47}$$

This formula for determinants extends in an obvious manner to $n \times n$ matrices. A more thorough discussion of the Levi-Civita symbol will appear later in these notes.

In $\mathbf{R}^n$ the Levi-Civita symbol with some or all the indices up is numerically equal to the permutation symbol will all the indices down,

$$\epsilon_{i_1 \ldots i_m} = \epsilon^{i_1 \ldots i_m},$$

since the Euclidean metric used to raise and lower indices is the Kronecker δ_{ij}.

On the other hand, in Minkowski space, raising an index with a value of 0 costs a minus sign, because $\eta_{00} = \eta^{00} = -1$. Thus, in $\mathcal{M}_{(1,3)}$

$$\epsilon_{i_0 i_1 i_2 i_3} = -\epsilon^{i_0 i_1 i_2 i_3},$$

since any permutation of $\{0, 1, 2, 3\}$ must contain a 0.

2.33 Definition The Hodge $\star$ operator is a linear map $\star : \bigwedge^m_p(\mathbf{R}^n) \longrightarrow \bigwedge^{n-m}_p(\mathbf{R}^n)$ defined in standard local coordinates by the equation,

$$\star\,(dx^{i_1} \wedge \ldots \wedge dx^{i_m}) = \frac{1}{(n-m)!} \epsilon^{i_1 \ldots i_m}{}_{i_{m+1} \ldots i_n} dx^{i_{m+1}} \wedge \ldots \wedge dx^{i_n}, \tag{2.48}$$

Since the forms $dx^{i_1} \wedge \ldots \wedge dx^{i_m}$ constitute a basis of the vector space $\bigwedge^m_p(\mathbf{R}^n)$ and the $\star$ operator is assumed to be a linear map, equation (2.48) completely specifies the map for all m-forms.

2.34 Example Consider the dimension $n = 3$ case. then

$$\begin{aligned}
\star dx^1 &= \epsilon^1{}_{jk} dx^j \wedge dx^k \\
&= \frac{1}{2!}[\epsilon^1{}_{23} dx^2 \wedge dx^3 + \epsilon^1{}_{32} dx^3 \wedge dx^2] \\
&= \frac{1}{2!}[dx^2 \wedge dx^3 - dx^3 \wedge dx^2] \\
&= \frac{1}{2!}[dx^2 \wedge dx^3 + dx^2 \wedge dx^3] \\
&= dx^2 \wedge dx^3.
\end{aligned}$$

We leave it to the reader to complete the computation of the action of the $\star$ operator on the other basis forms. The results are

$$\begin{aligned}
\star dx^1 &= +dx^2 \wedge dx^3 \\
\star dx^2 &= -dx^1 \wedge dx^3 \\
\star dx^3 &= +dx^1 \wedge dx^2,
\end{aligned} \tag{2.49}$$

$$\begin{aligned}
\star(dx^2 \wedge dx^3) &= dx^1 \\
\star(-dx^3 \wedge dx^1) &= dx^2 \\
\star(dx^1 \wedge dx^2) &= dx^3,
\end{aligned} \tag{2.50}$$

and

$$\star\,(dx^1 \wedge dx^2 \wedge dx^3) = 1. \tag{2.51}$$

In particular, if $f : \mathbf{R}^3 \longrightarrow \mathbf{R}$ is any 0-form (a function), then,

$$\begin{aligned} \star f &= f(dx^1 \wedge dx^2 \wedge dx^3) \\ &= f dV, \end{aligned} \tag{2.52}$$

where dV is the differential of volume, also called the **volume form**.

2.35 Example Let $\alpha = A_1 dx^1 A_2 dx^2 + A_3 dx^3$, and $\beta = B_1 dx^1 B_2 dx^2 + B_3 dx^3$. Then,

$$\begin{aligned} \star(\alpha \wedge \beta) &= (A_2B_3 - A_3B_2) \star (dx^2 \wedge dx^3) + (A_1B_3 - A_3B_1) \star (dx^1 \wedge dx^3) + \\ &\quad (A_1B_2 - A_2B_1) \star (dx^1 \wedge dx^2) \\ &= (A_2B_3 - A_3B_2)dx^1 + (A_1B_3 - A_3B_1)dx^2 + (A_1B_2 - A_2B_1)dx^3 \\ &= (\vec{\mathbf{A}} \times \vec{\mathbf{B}})_i dx^i \end{aligned} \tag{2.53}$$

The previous examples provide some insight on the action of the $\wedge$ and $\star$ operators. If one thinks of the quantities dx^1, dx^2 and dx^3 as analogous to $\vec{\mathbf{i}}$, $\vec{\mathbf{j}}$ and $\vec{\mathbf{k}}$, then it should be apparent that equations (2.49) are the differential geometry versions of the well known relations

$$\begin{aligned} \mathbf{i} &= \mathbf{j} \times \mathbf{k} \\ \mathbf{j} &= -\mathbf{i} \times \mathbf{k} \\ \mathbf{k} &= \mathbf{i} \times \mathbf{j}. \end{aligned}$$

This is even more evident upon inspection of equation (2.53), which relates the $\wedge$ operator to the Cartesian cross product.

One can get a good sense of the geometrical significance and the motivation for the creation of wedge products by considering a classical analogy in the language of vector calculus. As shown in figure 2.2, let us consider infinitesimal arc length vectors $\mathbf{i}\ dx$, $\mathbf{j}\ dy$ and $\mathbf{k}\ dz$ pointing along the coordinate axes. Recall from the definition, that the cross product of two vectors is a new vector whose magnitude is the area of the parallelogram subtended by the two vectors and which points in the direction of a unit vector perpendicular to the plane containing the two vectors, oriented according to the right hand rule. Since $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are mutually orthogonal vectors, the cross product of any pair is again a unit vector pointed in the direction of the third or the negative thereof. Thus, for example, in the xy-plane the differential of area is really an oriented quantity that can computed by the cross product $(\mathbf{i}\ dx \times \mathbf{j}\ dy) = dx\ dy\ \mathbf{k}$. A similar computation yields the differential of areas in the other two coordinate planes, except that in the xz-plane, the cross product needs to be taken in the reverse order. In terms of wedge products, the differential of area in the xy-plane is $(dx \wedge dy)$, so that the oriented nature of the surface element is built-in. Technically, when reversing the order of variables in a double integral one should introduce a minus sign. This is typically ignored in basic calculus computations of double and triple integrals, but it cannot be ignored in vector calculus in the context of flux of a vector field through a surface.

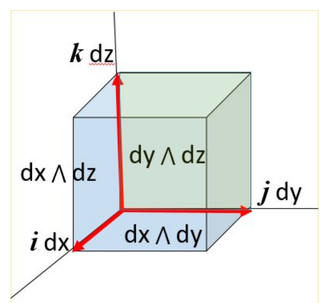


Fig. 2.2: Area Forms

2.36 Example In Minkowski space the collection of all 2-forms has dimension $\binom{4}{2} = 6$. The Hodge $\star$ operator in this case splits $\bigwedge^2(\mathcal{M}_{1,3})$ into two 3-dim subspaces $\bigwedge^2_{\pm}$, such that $\star : \bigwedge^2_{\pm} \longrightarrow \bigwedge^2_{\mp}$. More specifically, $\bigwedge^2_{+}$ is spanned by the forms $\{dx^0 \wedge dx^1, dx^0 \wedge dx^2, dx^0 \wedge dx^3\}$, and $\bigwedge^2_{-}$ is spanned by the forms $\{dx^2 \wedge dx^3, -dx^1 \wedge dx^3, dx^1 \wedge dx^2\}$. The action of $\star$ on $\bigwedge^2_+$ is

$$\begin{aligned}
\star(dx^0 \wedge dx^1) &= \tfrac{1}{2}\epsilon^{01}{}_{kl}dx^k \wedge dx^l = & -dx^2 \wedge dx^3 \\
\star(dx^0 \wedge dx^2) &= \tfrac{1}{2}\epsilon^{02}{}_{kl}dx^k \wedge dx^l = & +dx^1 \wedge dx^3 \\
\star(dx^0 \wedge dx^3) &= \tfrac{1}{2}\epsilon^{03}{}_{kl}dx^k \wedge dx^l = & -dx^1 \wedge dx^2,
\end{aligned}$$

and on $\bigwedge^2_-$,

$$\begin{aligned}
\star(+dx^2 \wedge dx^3) &= \tfrac{1}{2}\epsilon^{23}{}_{kl}dx^k \wedge dx^l = & dx^0 \wedge dx^1 \\
\star(-dx^1 \wedge dx^3) &= \tfrac{1}{2}\epsilon^{13}{}_{kl}dx^k \wedge dx^l = & dx^0 \wedge dx^2 \\
\star(+dx^1 \wedge dx^2) &= \tfrac{1}{2}\epsilon^{12}{}_{kl}dx^k \wedge dx^l = & dx^0 \wedge dx^3.
\end{aligned}$$

In verifying the equations above, we recall that the Levi-Civita symbols that contain an index with value 0 in the up position have an extra minus sign as a result of raising the index with η^{00}. If $F \in \bigwedge^2(\mathcal{M})$, we will formally write $F = F_+ + F_-$, where $F_\pm \in \bigwedge^2_\pm$. We would like to note that the action of the dual operator on $\bigwedge^2(\mathcal{M})$ is such that $\star \bigwedge^2(\mathcal{M}) \longrightarrow \bigwedge^2(\mathcal{M})$, and $\star^2 = -1$. In a vector space a map like $\star$ with the property $\star^2 = -1$ is a **linear involution** of the space. In the case in question $\bigwedge^2_\pm$ are the eigenspaces corresponding to the +1 and -1 eigenvalues of this involution.

It is also worthwhile to calculate the duals of 1-forms in $\mathcal{M}_{1,3}$. The results are

$$\begin{aligned}
\star dt &= -dx^1 \wedge dx^2 \wedge dx^3 \\
\star dx^1 &= +dx^2 \wedge dt \wedge dx^3 \\
\star dx^2 &= +dt \wedge dx^1 \wedge dx^3 \\
\star dx^3 &= +dx^1 \wedge dt \wedge dx^2. \qquad (2.54)
\end{aligned}$$

Gradient, Curl and Divergence

Classical differential operators that enter in Green's and Stokes' Theorems are better understood as special manifestations of the exterior differential and the Hodge-$\star$ operators in $\mathbf{R}^3$. Here is precisely how this works:

1. Let $f : \mathbf{R}^3 \longrightarrow \mathbf{R}$ be a $\mathcal{C}^\infty$ function. Then

$$df = \frac{\partial f}{\partial x^j} dx^j = \nabla f \cdot \mathbf{dx} \qquad (2.55)$$

2. Let $\alpha = A_i dx^i$ be a 1-form in $\mathbf{R}^3$. Then

$$\begin{aligned}
(\star d)\alpha &= \frac{1}{2}(\frac{\partial A_i}{\partial x^j} - \frac{\partial A_i}{\partial x^j}) \star (dx^i \wedge dx^j) \\
&= (\nabla \times \mathbf{A}) \cdot \mathbf{dx} \qquad (2.56)
\end{aligned}$$

3. Let $\alpha = B_1 dx^2 \wedge dx^3 + B_2 dx^3 \wedge dx^1 + B_3 dx^1 \wedge dx^2$ be a 2-form in $\mathbf{R}^3$. Then

$$\begin{aligned}
d\alpha &= (\frac{\partial B_1}{\partial x^1} + \frac{\partial B_2}{\partial x^2} + \frac{\partial B_3}{\partial x^3})\, dx^1 \wedge dx^2 \wedge dx^3 \\
&= (\nabla \cdot \mathbf{B})\, dV \qquad (2.57)
\end{aligned}$$

4. Let $\alpha = B_i dx^i$, then

$$(\star d \star)\, \alpha = \nabla \cdot \mathbf{B} \qquad (2.58)$$

5. Let f be a real valued function. Then the **Laplacian** is given by:

$$(\star d \star)\, df = \nabla \cdot \nabla f = \nabla^2 f \tag{2.59}$$

The results above can be summarized in terms of short exact sequence called the **De Rham Complex** as shown in figure 2.3. The sequence is called exact because successive application of the differential operator gives zero. That is, $d \circ d = 0$. Since there are no 4-forms in $\mathbf{R}^3$, the sequence terminates as shown. If one starts with a function in $\bigwedge^0(\mathbf{R}^3)$, then $(d \circ d)f = 0$ just says that

$$\begin{tikzcd} \bigwedge^0(\mathbf{R}^3) \arrow[r, "d", "\mathbf{Grad}"'] & \bigwedge^1(\mathbf{R}^3) \arrow[r, "d", "\mathbf{Curl}"'] \arrow[r, bend left, "\star"] & \bigwedge^2(\mathbf{R}^3) \arrow[r, "d", "\mathbf{Div}"'] & \bigwedge^3(\mathbf{R}^3) \end{tikzcd}$$

Fig. 2.3: De Rham Complex

$\nabla \times \nabla f = 0$ as in the case of conservative vector fields. If instead, one starts with a one form α in $\bigwedge^1(\mathbf{R}^3)$, corresponding to a vector field $\mathbf{A}$, then $(d \circ d)f = 0$ says that $\nabla \cdot (\nabla \times \mathbf{A}) = 0$ as in the case of incompressible vector fields. If one starts with a function, but instead of applying the differential twice consecutively, one "hops" in between with the Hodge operator, the result is the Laplacian of the function.

If denote by R a closed region in Euclidean space whose boundary is δR , then in terms of forms, the the fundamental theorem, Stoke's Theorem and the Divergence theorem can be summarized in $\mathbf{R}^3$ by a single Stokes theorem

$$\int_{\delta R} \omega = \int \int_R d\omega. \tag{2.60}$$

It is also possible to define and manipulate formulas of classical vector calculus using the permutation symbols. For example, let $\mathbf{a} = (A_1, A_2, A_3)$ and $\mathbf{B} = (B_1, B_2, B_3)$ be any two Euclidean vectors. Then it is easy to see that

$$(\mathbf{A} \times \mathbf{B})_k = \epsilon^{ij}{}_k A_i B_j,$$

and

$$(\nabla \times \mathbf{B})_k = \epsilon^{ij}{}_k \frac{\partial A_i}{\partial x^j},$$

To derive many classical vector identities in this formalism, it is necessary to first establish the following identity:

$$\epsilon^{ijm} \epsilon_{klm} = \delta^i_k \delta^j_l - \delta^i_l \delta^j_k. \tag{2.61}$$

2.37 Example

$$\begin{aligned}
[\mathbf{A} \times (\mathbf{B} \times \mathbf{C})]_l &= \epsilon_l^{mn} A\ _m (\mathbf{B} \times \mathbf{C})_m \\
&= \epsilon_l^{mn} A_m (\epsilon^{jk}{}_n B_j C_k) \\
&= \epsilon^{mn}_l \epsilon^{jk}{}_n A_m B_j C_k) \\
&= \epsilon_{mnl} \epsilon^{jkn} A^m B_j C_k) \\
&= (\delta^j_l \delta^k_m - \delta^k_l \delta^j_n) A^m B_j C_k \\
&= B_l A^m C_m - C_l A^m b_n,
\end{aligned}$$

Or, rewriting in vector form

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}). \tag{2.62}$$

Maxwell Equations

The classical equations of Maxwell describing electromagnetic phenomena are

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 4\pi\rho & \nabla \times \mathbf{B} &= 4\pi\mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} \\ \nabla \cdot \mathbf{B} &= 0 & \nabla \times \mathbf{E} &= -\frac{\partial B}{\partial t} \end{aligned} \tag{2.63}$$

We would like to formulate these equations in the language of differential forms. Let $x^\mu = (t, x^1, x^2, x^3)$ be local coordinates in Minkowski's space $\mathcal{M}_{1,3}$. Define the Maxwell 2-form F by the equation

$$F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu, \quad (\mu, \nu = 0, 1, 2, 3), \tag{2.64}$$

where

$$F_{\mu\nu} = \begin{bmatrix} 0 & -E_x & -E_y & -E_y \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{bmatrix}. \tag{2.65}$$

Written in complete detail, Maxwell's 2-form is given by

$$\begin{aligned} F \quad = \quad & -E_x dt \wedge dx^1 - E_y dt \wedge dx^2 - E_z dt \wedge dx^3 + \\ & B_z dx^1 \wedge dx^2 - B_y dx^1 \wedge dx^3 + B_x dx^2 \wedge dx^3. \end{aligned} \tag{2.66}$$

We also define the source current 1-form

$$J = J_\mu dx^\mu = \rho dt + J_1 dx^1 + J_2 dx^2 + J_3 dx^3. \tag{2.67}$$

2.38 Proposition Maxwell's Equations (2.63) are equivalent to the equations

$$\begin{aligned} dF &= 0, \\ d \star F &= 4\pi \star J. \end{aligned} \tag{2.68}$$

Proof: The proof is by direct computation using the definitions of the exterior derivative and the Hodge-⋆ operator.

$$\begin{aligned} dF \quad = \quad & -\frac{\partial E_x}{\partial x^2} \wedge dx^2 \wedge dt \wedge dx^1 - \frac{\partial E_x}{\partial x^3} \wedge dx^3 \wedge dt \wedge dx^1 + \\ & -\frac{\partial E_y}{\partial x^1} \wedge dx^1 \wedge dt \wedge dx^2 - \frac{\partial E_y}{\partial x^3} \wedge dx^3 \wedge dt \wedge dx^2 + \\ & -\frac{\partial E_z}{\partial x^1} \wedge dx^1 \wedge dt \wedge dx^3 - \frac{\partial E_z}{\partial x^2} \wedge dx^2 \wedge dt \wedge dx^3 + \\ & \frac{\partial B_z}{\partial t} \wedge dt \wedge dx^1 \wedge dx^2 - \frac{\partial B_z}{\partial x^3} \wedge dx^3 \wedge dx^1 \wedge dx^2 - \\ & \frac{\partial B_y}{\partial t} \wedge dt \wedge dx^1 \wedge dx^3 - \frac{\partial B_y}{\partial x^2} \wedge dx^2 \wedge dx^1 \wedge dx^3 + \\ & \frac{\partial B_x}{\partial t} \wedge dt \wedge dx^2 \wedge dx^3 + \frac{\partial B_x}{\partial x^1} \wedge dx^1 \wedge dx^2 \wedge dx^3. \end{aligned}$$

Collecting terms and using the antisymmetry of the wedge operator, we get

$$\begin{aligned} dF &= (\frac{\partial B_x}{\partial x^1}+\frac{\partial B_y}{\partial x^2}+\frac{\partial B_z}{\partial x^3})dx^1\wedge dx^2\wedge dx^3 + \\ &(\frac{\partial E_y}{\partial x^3}-\frac{\partial E_z}{\partial x^2}-\frac{\partial B_x}{\partial t})dx^2\wedge dt\wedge dx^3 + \\ &(\frac{\partial E_z}{\partial x^1}-\frac{\partial E_x}{\partial dx^3}-\frac{\partial B_y}{\partial t})dt\wedge dx^1\wedge x^3 + \\ &(\frac{\partial E_x}{\partial x^2}-\frac{\partial E_y}{\partial x^1}-\frac{\partial B_z}{\partial t})dx^1\wedge dt\wedge dx^2. \end{aligned}$$

Therefore, $dF = 0$ iff

$$\frac{\partial B_x}{\partial x^1}+\frac{\partial B_y}{\partial x^2}+\frac{\partial B_y}{\partial x^3}=0,$$

which is the same as

$$\nabla\cdot\mathbf{B}=0,$$

and

$$\begin{aligned} \frac{\partial E_y}{\partial x^3}-\frac{\partial E_z}{\partial x^2}-\frac{\partial B_x}{\partial x^1}&=0,\\ \frac{\partial E_z}{\partial x^1}-\frac{\partial E_x}{\partial x^3}-\frac{\partial B_y}{\partial x^2}&=0,\\ \frac{\partial E_x}{\partial x^2}-\frac{\partial E_y}{\partial x^1}-\frac{\partial B_z}{\partial x^3}&=0, \end{aligned}$$

which means that

$$-\nabla\times\mathbf{E}-\frac{\partial B}{\partial t}=0. \tag{2.69}$$

To verify the second set of Maxwell equations, we first compute the dual of the current density 1-form (2.67) using the results from example 2.4. We get

$$\star J=-\rho dx^1\wedge dx^2\wedge dx^3+J_1dx^2\wedge dt\wedge dx^3+J_2dt\wedge dx^1\wedge dx^3+J_3dx^1\wedge dt\wedge dx^2. \tag{2.70}$$

We could now proceed to compute $d{\star}F$, but perhaps it is more elegant to notice that $F\in\bigwedge^2(\mathcal{M})$, and so, according to example (2.4), F splits into $F=F_+ + F_-$. In fact, we see from (2.65) that the components of F_+ are those of $-\mathbf{E}$ and the components of F_- constitute the magnetic field vector $\mathbf{B}$. Using the results of example (2.4), we can immediately write the components of ${\star}F$:

$$\begin{aligned} \star F &= B_xdt\wedge dx^1+B_ydt\wedge dx^2+B_zdt\wedge dx^3+ \\ &E_zdx^1\wedge dx^2-E_ydx^1\wedge dx^3+E_xdx^2\wedge dx^3, \end{aligned} \tag{2.71}$$

or equivalently,

$$F_{\mu\nu}=\begin{bmatrix} 0 & B_x & B_y & B_y\\ -B_x & 0 & E_z & -E_y\\ -B_y & -E_z & 0 & E_x\\ -B_z & E_y & -E_x & 0 \end{bmatrix}. \tag{2.72}$$

Since the effect of the dual operator amounts to exchanging

$$\begin{aligned} \mathbf{E} &\longmapsto -\mathbf{B}\\ \mathbf{B} &\longmapsto +\mathbf{E}, \end{aligned}$$

we can infer from equations (2.69) and (2.70) that

$$\nabla \cdot \mathbf{E} = 4\pi\rho$$

and

$$\nabla \times \mathbf{B} - \frac{\partial E}{\partial t} = 4\pi\mathbf{J}.$$

Chapter 3

Connections

3.1 Frames

As we have already noted in Chapter 1, the theory of curves in $\mathbf{R}^3$ can be elegantly formulated by introducing orthonormal triplets of vectors which we called Frenet frames. The Frenet vectors are adapted to the curves in such a manner that the rate of change of the frame gives information about the curvature of the curve. In this chapter we will study the properties of arbitrary frames and their corresponding rates of change in the direction of the various vectors in the frame. These concepts will then be applied later to special frames adapted to surfaces.

3.1 Definition A coordinate **frame** in $\mathbf{R}^n$ is an n-tuple of vector fields $\{e_1, \ldots, e_n\}$ which are linearly independent at each point $\mathbf{p}$ in the space.

In local coordinates $\{x^1, \ldots, x^n\}$, we can always express the frame vectors as linear combinations of the standard basis vectors

$$e_i = \partial_j A^j_{\ i}, \tag{3.1}$$

where $\partial_j = \frac{\partial}{\partial x^1}$. We assume the matrix $A = (A^j_{\ i})$ to be nonsingular at each point. In linear algebra, this concept is referred to as a change of basis, the difference being that in our case, the transformation matrix A depends on the position. A frame field is called **orthonormal** if at each point,

$$< e_i, e_j >= \delta_{ij}. \tag{3.2}$$

Throughout this chapter, we will assume that all frame fields are orthonormal. Whereas this restriction is not necessary, it is convenient because it it results in considerable simplification in computing the components of an arbitrary vector in the frame.

3.2 Proposition If $\{e_1, \ldots, e_n\}$ is an orthonormal frame, then the transformation matrix is orthogonal (ie, $AA^T = I$)

Proof: The proof is by direct computation. Let $e_i = \partial_j A^j_{\ i}$. Then

$$\begin{aligned}
\delta_{ij} &= < e_i, e_j > \\
&= < \partial_k A^k_{\ i}, \partial_l A^l_{\ j} > \\
&= A^k_{\ i} A^l_{\ j} < \partial_k, \partial_l > \\
&= A^k_{\ i} A^l_{\ j} \delta_{kl} \\
&= A^k_{\ i} A_{kj} \\
&= A^k_{\ i} (A^T)_{jk}.
\end{aligned}$$

Hence

$$\begin{aligned}(A^T)_{jk}A^k_i &= \delta_{ij}\\ (A^T)^j{}_k A^k_i &= \delta^j_i\\ A^T A &= I\end{aligned}$$

Given a frame vectors e_i, we can also introduce the corresponding dual coframe forms θ^i by requiring that

$$\theta^i(e_j) = \delta^i_j. \tag{3.3}$$

Since the dual coframe is a set of 1-forms, they can also be expressed in local coordinates as linear combinations

$$\theta^i = B^i{}_k dx^k.$$

It follows from equation(3.3), that

$$\begin{aligned}\theta^i(e_j) &= B^i_k dx^k(\partial_l A^l{}_j)\\ &= B^i_k A^l{}_j dx^k(\partial_l)\\ &= B^i_k A^l{}_j \delta^k_l\\ \delta^i{}_j &= B^i_k A^k_j.\end{aligned}$$

Therefore, we conclude that $BA = I$, so $B = A^{-1} = A^T$. In other words, when the frames are orthonormal, we have

$$\begin{aligned}e_i &= \partial_k A^k_i\\ \theta^i &= A^i{}_k dx^k.\end{aligned} \tag{3.4}$$

3.3 Example Consider the transformation from Cartesian to cylindrical coordinates:

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = z. \tag{3.5}$$

Using the chain rule for partial derivatives, we have

$$\begin{aligned}\frac{\partial}{\partial r} &= \cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}\\ \frac{\partial}{\partial\theta} &= -r\sin\theta\frac{\partial}{\partial x} + r\cos\theta\frac{\partial}{\partial y}\\ \frac{\partial}{\partial z} &= \frac{\partial}{\partial z}\end{aligned}$$

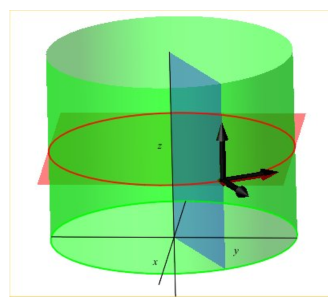


Fig. 3.1: Frame

The vectors $\frac{\partial}{\partial r}$, and $\frac{\partial}{\partial z}$ are clearly unit vectors. To make the vector $\frac{\partial}{\partial\theta}$ a unit vector, it suffices to divide it by its length r. We can then compute the dot products of each pair of vectors and easily verify that the quantities

$$e_1 = \frac{\partial}{\partial r}, \quad e_2 = \frac{1}{r}\frac{\partial}{\partial\theta}, \quad e_3 = \frac{\partial}{\partial z}, \tag{3.6}$$

are a triplet of mutually orthogonal unit vectors and thus constitute an orthonormal frame. The surfaces with constant value for the coordinates r, θ and z respectively, represent a set of mutually orthogonal surfaces at each point. The frame vectors at a point are normal to these surfaces as shown in figure 3.1. Physicists often refer to these frame vectors as $\{\hat{\mathbf{r}}, \hat{\theta}, \hat{\mathbf{z}}\}$, or as $\{e_r, e_\theta, e_z.\}$.

3.4 Example For spherical coordinates (2.20)

$$
\begin{aligned}
x &= r\sin\theta\cos\phi \\
y &= r\sin\theta\sin\phi \\
z &= r\cos\theta,
\end{aligned}
$$

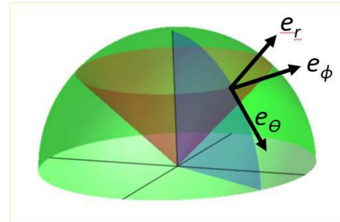


Fig. 3.2: Frame

the chain rule leads to

$$
\begin{aligned}
\frac{\partial}{\partial r} &= \sin\theta\cos\phi\frac{\partial}{\partial x} + \sin\theta\sin\phi\frac{\partial}{\partial y} + \cos\theta\frac{\partial}{\partial z}, \\
\frac{\partial}{\partial \theta} &= r\cos\theta\cos\phi\frac{\partial}{\partial x} + r\cos\theta\sin\phi\frac{\partial}{\partial y} - r\sin\theta\frac{\partial}{\partial z}, \\
\frac{\partial}{\partial \phi} &= -r\sin\theta\sin\phi\frac{\partial}{\partial x} + r\sin\theta\cos\phi\frac{\partial}{\partial y}.
\end{aligned}
$$

The vector $\frac{\partial}{\partial r}$ is of unit length but the other two need to be normalised. As before, all we need to do is divide the vectors by their magnitude. For $\frac{\partial}{\partial \theta}$ we divide by r and for $\frac{\partial}{\partial \phi}$, we divide by $r\sin\theta$. Taking the dot products of all pairs and using basic trigonometric identities, one can verify that we again obtain an orthonormal frame.

$$
e_1 = e_r = \frac{\partial}{\partial r}, \quad e_2 = e_\theta = \frac{1}{r}\frac{\partial}{\partial \theta}, \quad e_3 = e_\varphi = \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}, \tag{3.7}
$$

Furthermore, the frame vectors are normal to triply orthogonal surfaces, which in this case are spheres, cones and planes, as shown in figure 3.2. The fact that the chain rule in the two situations above leads to orthonormal frames is not coincidental. The results are related to the orthogonality of the level surfaces $x^i = constant$. Since the level surfaces are orthogonal whenever they intersect, one expects the gradients of the surfaces to also be orthogonal. Transformations of this type are called triply orthogonal systems.

3.2 Curvilinear Coordinates

Orthogonal transformations such as spherical and cylindrical coordinates appear ubiquitously in mathematical physics because the geometry of a large number of problems in this discipline exhibit symmetry with respect to an axis or to the origin. In such situations, transformations to the appropriate coordinate system often result in considerable simplification of the field equations involved in the problem. It has been shown that the Laplace operator that appears in potential, heat, and wave equations, and in Schrödinger field equations is separable in twelve orthogonal coordinate systems. A simple and efficient method to calculate the Laplacian in orthogonal coordinates can be implemented using differential forms.

3.5 Example In spherical coordinates the differential of arc length is given by (see equation 2.21) the metric:

$$
ds^2 = dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2.
$$

Let

$$
\begin{aligned}
\theta^1 &= dr \\
\theta^2 &= r d\theta \\
\theta^3 &= r\sin\theta d\phi.
\end{aligned}
\tag{3.8}
$$

Note that these three 1-forms constitute the dual coframe to the orthonormal frame derived in equation(3.7). Consider a scalar field $f = f(r, \theta, \phi)$. We now calculate the Laplacian of f in spherical coordinates using the methods of section 2.4. To do this, we first compute the differential df and express the result in terms of the coframe.

$$\begin{aligned} df &= \frac{\partial f}{\partial r} dr + \frac{\partial f}{\partial \theta} d\theta + \frac{\partial f}{\partial \phi} d\phi \\ &= \frac{\partial f}{\partial r} \theta^1 + \frac{1}{r} \frac{\partial f}{\partial \theta} \theta^2 + \frac{1}{r \sin\theta} \frac{\partial f}{\partial \phi} \theta^3 \end{aligned}$$

The components df in the coframe represent the gradient in spherical coordinates. Continuing with the scheme of section 2.4, we next apply the Hodge-$\star$ operator. Then, we rewrite the resulting 2-form in terms of wedge products of coordinate differentials so that we can apply the definition of the exterior derivative.

$$\begin{aligned} \star df &= \frac{\partial f}{\partial r} \theta^2 \wedge \theta^3 - \frac{1}{r} \frac{\partial f}{\partial \theta} \theta^1 \wedge \theta^3 + \frac{1}{r \sin\theta} \frac{\partial f}{\partial \phi} \theta^1 \wedge \theta^2 \\ &= r^2 \sin\theta \frac{\partial f}{\partial r} d\theta \wedge d\phi - r \sin\theta \frac{1}{r} \frac{\partial f}{\partial \theta} dr \wedge d\phi + r \frac{1}{r \sin\theta} \frac{\partial f}{\partial \phi} dr \wedge d\theta \\ &= r^2 \sin\theta \frac{\partial f}{\partial r} d\theta \wedge d\phi - \sin\theta \, \frac{\partial f}{\partial \theta} dr \wedge d\phi + \frac{1}{\sin\theta} \frac{\partial f}{\partial \phi} dr \wedge d\theta \\ d \star df &= \frac{\partial}{\partial r}(r^2 \sin\theta \frac{\partial f}{\partial r}) dr \wedge d\theta \wedge d\phi - \frac{\partial}{\partial \theta}(\sin\theta \frac{\partial f}{\partial \theta}) d\theta \wedge dr \wedge d\phi + \frac{1}{\sin\theta} \frac{\partial}{\partial \phi}(\frac{\partial f}{\partial \phi}) d\phi \wedge dr \wedge d\theta \\ &= \left[\sin\theta \frac{\partial}{\partial r}(r^2 \frac{\partial f}{\partial r}) + \frac{\partial}{\partial \theta}(\sin\theta \frac{\partial f}{\partial \theta}) + \frac{1}{\sin\theta} \frac{\partial^2 f}{\partial \phi^2} \right] dr \wedge d\theta \wedge d\phi. \end{aligned}$$

Finally, rewriting the differentials back in terms of the the coframe, we get

$$d \star df = \frac{1}{r^2 \sin\theta} \left[\sin\theta \frac{\partial}{\partial r}(r^2 \frac{\partial f}{\partial r}) + \frac{\partial}{\partial \theta}(\sin\theta \frac{\partial f}{\partial \theta}) + \frac{1}{\sin\theta} \frac{\partial^2 f}{\partial \phi^2} \right] \theta^1 \wedge \theta^2 \wedge \theta^3.$$

So, the Laplacian of f is given by

$$\nabla^2 f = \frac{1}{r^2} \frac{\partial}{\partial r} \left[r^2 \frac{\partial f}{\partial r} \right] + \frac{1}{r^2} \left[\frac{1}{\sin\theta} \frac{\partial}{\partial \theta}(\sin\theta \frac{\partial f}{\partial \theta}) + \frac{1}{\sin^2\theta} \frac{\partial^2 f}{\partial \phi^2} \right] \tag{3.9}$$

The derivation of the expression for the spherical Laplacian through the use of differential forms is elegant and leads naturally to the operator in Sturm Liouville form.

The process above can be carried out for general orthogonal transformations. A change of coordinates $x^i = x^i(u^k)$ leads to an orthogonal transformation if in the new coordinate system u^k, the line metric

$$ds^2 = g_{11}(du^1)^2 + g_{22}(du^2)^2 + g_{33}(du^3)^2 \tag{3.10}$$

only has diagonal entries. In this case, we choose the coframe

$$\begin{aligned} \theta^1 &= \sqrt{g_{11}} du^1 = h_1 du^1 \\ \theta^2 &= \sqrt{g_{22}} du^2 = h_2 du^2 \\ \theta^3 &= \sqrt{g_{33}} du^3 = h_3 du^3 \end{aligned}$$

The quantities $\{h_1, h_2, h_3\}$ are classically called the weights. Please note that, in the interest of connecting to classical terminology, we have exchanged two indices for one and this will cause small discrepancies with the index summation convention. We will revert to using a summation symbol when these discrepancies occur. To satisfy the duality condition $\theta^i(e_j) = \delta^i_j$, we must choose the corresponding frame vectors e_i as follows:

$$e_1 = \frac{1}{\sqrt{g_{11}}}\frac{\partial}{\partial u^1} = \frac{1}{h_1}\frac{\partial}{\partial u^1}$$
$$e_2 = \frac{1}{\sqrt{g_{22}}}\frac{\partial}{\partial u^2} = \frac{1}{h_2}\frac{\partial}{\partial u^2}$$
$$e_3 = \frac{1}{\sqrt{g_{33}}}\frac{\partial}{\partial u^3} = \frac{1}{h_3}\frac{\partial}{\partial u^3}$$

Gradient. Let $f = f(x^i)$ and $x^i = x^i(u^k)$. Then

$$\begin{aligned}
df &= \frac{\partial f}{\partial x^k}dx^k \\
&= \frac{\partial f}{\partial u^i}\frac{\partial u^i}{\partial x^k}dx^k \\
&= \frac{\partial f}{\partial u^i}du^i \\
&= \sum_i \frac{1}{h^i}\frac{\partial f}{\partial u^i}\theta^i \\
&= e_i(f)\theta^i.
\end{aligned}$$

As expected, the components of the gradient in the coframe θ^i are the just the frame vectors.

$$\nabla = \left(\frac{1}{h_1}\frac{\partial}{\partial u^1}, \frac{1}{h_2}\frac{\partial}{\partial u^2}, \frac{1}{h_3}\frac{\partial}{\partial u^3}\right) \tag{3.11}$$

Curl. Let $F = (F_1, F_2, F_3)$ be a classical vector field. Construct the corresponding 1-form $F = F_i\theta^i$ in the coframe. We calculate the curl using the dual of the exterior derivative.

$$\begin{aligned}
F &= F_1\theta^1 + F_2\theta^2 + F_3\theta^3 \\
&= (h_1F_1)du^1 + (h_2F_2)du^2 + (h_3F_3)du^3 \\
&= (hF)_i du^i, \quad \text{where } (hF)_i = h_iF_i \\
dF &= \frac{1}{2}\left[\frac{\partial(hF)_i}{\partial u^j} - \frac{\partial(hF)_j}{\partial u^i}\right] du^i \wedge du^j \\
&= \frac{1}{h_ih_j}\left[\frac{\partial(hF)_i}{\partial u^j} - \frac{\partial(hF)_j}{\partial u^i}\right] d\theta^i \wedge d\theta^j \\
\star dF &= \epsilon^{ij}{}_k\left[\frac{1}{h_ih_j}[\frac{\partial(hF)_i}{\partial u^j} - \frac{\partial(hF)_j}{\partial u^i}]\right]\theta^k = (\nabla \times F)_k\theta^k.
\end{aligned}$$

Thus, the components of the curl are

$$\left(\frac{1}{h_2h_3}[\frac{\partial(h_3F_3)}{\partial u^2} - \frac{\partial(h_2F_2)}{\partial u^3}], \frac{1}{h_1h_3}[\frac{\partial(h_3F_3)}{\partial u^1} - \frac{\partial(h_1F_1)}{\partial u^3}], \frac{1}{h_1h_2}[\frac{\partial(h_1F_1)}{\partial u^2} - \frac{\partial(h_2F_2)}{\partial u^1}].\right) \tag{3.12}$$

Divergence. As before, let $F = F_i\theta^i$ and recall that $\nabla \cdot F = \star d \star F$. The computation yields

$$\begin{aligned}
F &= F_1\theta^1 + F_2\theta^2 + F_3\theta^3 \\
\star F &= F_1\theta^2 \wedge \theta^3 + F_2\theta^3 \wedge \theta^1 + F_3\theta^1 \wedge \theta^2 \\
&= (h_2h_3F_1)du^2 \wedge du^3 + (h_1h_3F_2)du^3 \wedge du^1 + (h_1h_2F_3)du^1 \wedge du^2 \\
d \star dF &= \left[\frac{\partial(h_2h_3F_1)}{\partial u^1} + \frac{\partial(h_1h_3F_2)}{\partial u^2} + \frac{\partial(h_1h_2F_3)}{\partial u^3}\right] du^1 \wedge du^2 \wedge du^3.
\end{aligned}$$

Therefore,

$$\nabla \cdot F = \star d \star F = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial(h_2 h_3 F_1)}{\partial u^1} + \frac{\partial(h_1 h_3 F_2)}{\partial u^2} + \frac{\partial(h_1 h_2 F_3)}{\partial u^3}\right]. \tag{3.13}$$

3.3 Covariant Derivative

In this section we introduce a generalization of directional derivatives. The directional derivative measures the rate of change of a function in the direction of a vector. What we want is a quantity which measures the rate of change of a vector field in the direction of another.

3.6 Definition Given a pair (X, Y) of arbitrary vector field in $\mathbf{R}^n$, we associate a new vector field $\overline{\nabla}_X Y$, so that $\overline{\nabla}_X : \mathscr{X}(\mathbf{R}^n) \longrightarrow \mathscr{X}(\mathbf{R}^n)$. The quantity $\overline{\nabla}$ called a **Koszul connection** if it satisfies the following properties:

1. $\overline{\nabla}_{fX}(Y) = f\overline{\nabla}_X Y$,
2. $\overline{\nabla}_{(X_1+X_2)}Y = \overline{\nabla}_{X_1}Y + \overline{\nabla}_{X_2}Y$,
3. $\overline{\nabla}_X(Y_1 + Y_2) = \overline{\nabla}_X Y_1 + \overline{\nabla}_X Y_2$,
4. $\overline{\nabla}_X fY = X(f)Y + f\overline{\nabla}_X Y$,

for all vector fields $X, X_1, X_2, Y, Y_1, Y_2 \in \mathscr{X}(\mathbf{R}^n)$ and all smooth functions f. The definition states that the map $\overline{\nabla}_X$ is linear on X but behaves as a linear derivation on Y. For this reason, the quantity $\overline{\nabla}_X Y$ is called the **covariant derivative** of Y in the direction of X.

3.7 Proposition Let $Y = f^i \frac{\partial}{\partial x^i}$ be a vector field in $\mathbf{R}^n$, and let X another C^∞ vector field. Then the operator given by

$$\overline{\nabla}_X Y = X(f^i)\frac{\partial}{\partial x^i} \tag{3.14}$$

defines a Koszul connection.
Proof: The proof just requires verification that the four properties above are satisfied, and it is left as an exercise.

The operator defined in this proposition is called the **standard connection** compatible with the standard Euclidean metric. The action of this connection on a vector field Y yields a new vector field whose components are the directional derivatives of the components of Y.

3.8 Example Let

$$X = x\frac{\partial}{\partial x} + xz\frac{\partial}{\partial y},\ Y = x^2\frac{\partial}{\partial x} + xy^2\frac{\partial}{\partial y}.$$

Then,

$$\begin{aligned}
\overline{\nabla}_X Y &= X(x^2)\frac{\partial}{\partial x} + X(xy^2)\frac{\partial}{\partial y} \\
&= [x\frac{\partial}{\partial x}(x^2) + xz\frac{\partial}{\partial y}(x^2)]\frac{\partial}{\partial x} + [x\frac{\partial}{\partial x}(xy^2) + xz\frac{\partial}{\partial y}(xy^2)]\frac{\partial}{\partial y} \\
&= 2x^2\frac{\partial}{\partial x} + (xy^2 + 2x^2yz)\frac{\partial}{\partial y}.
\end{aligned}$$

3.9 Definition A Koszul connection $\overline{\nabla}_X$ is compatible with the metric $g(Y, Z) = \langle Y, Z\rangle$ if

$$\overline{\nabla}_X\langle Y, Z\rangle = \langle\overline{\nabla}_X Y, Z\rangle + \langle Y, \overline{\nabla}_X Z\rangle. \tag{3.15}$$

if $F : \mathbf{R}^n \to \mathbf{R}^n$ is an isometry so that $\langle F_*X, F_*Y\rangle = \langle X, Y\rangle$, then the quantity defined by $F_*\nabla$

In Euclidean space, the components of the standard frame vectors are constant, and thus their rates of change in any direction vanish. Let e_i be arbitrary frame field with dual forms θ^i. The covariant derivatives of the frame vectors in the directions of a vector X will in general yield new vectors. The new vectors must be linear combinations of the the basis vectors as follows:

$$
\begin{aligned}
\overline{\nabla}_X e_1 &= \omega^1{}_1(X)e_1 + \omega^2{}_1(X)e_2 + \omega^3{}_1(X)e_3 \\
\overline{\nabla}_X e_2 &= \omega^1{}_2(X)e_1 + \omega^2{}_2(X)e_2 + \omega^3{}_2(X)e_3 \\
\overline{\nabla}_X e_3 &= \omega^1{}_3(X)e_1 + \omega^2{}_3(X)e_2 + \omega^3{}_3(X)e_3
\end{aligned} \tag{3.16}
$$

The coefficients can be more succinctly expressed using the compact index notation,

$$\overline{\nabla}_X e_i = e_j \omega^j{}_i(X). \tag{3.17}$$

It follows immediately that

$$\omega^j{}_i(X) = \theta^j(\overline{\nabla}_X e_i). \tag{3.18}$$

Equivalently, one can take the inner product of both sides of equation (3.17) with e_k to get

$$
\begin{aligned}
< \overline{\nabla}_X e_i, e_k > &= < e_j \omega^j{}_i(X), e_k > \\
&= \omega^j{}_i(X) < e_j, e_k > \\
&= \omega^j{}_i(X) g_{jk}
\end{aligned}
$$

Hence,

$$< \overline{\nabla}_X e_i, e_k >= \omega_{ki}(X) \tag{3.19}$$

The left hand side of the last equation is the inner product of two vectors, so the expression represents an array of functions. Consequently, the right hand side also represents an array of functions. In addition, both expressions are linear on X, since by definition $\overline{\nabla}_X$ is linear on X. We conclude that the right hand side can be interpreted as a matrix in which each entry is a 1-forms acting on the vector X to yield a function. The matrix valued quantity $\omega^i{}_j$ is called the **connection form**. At the risk of being inconsistent with formalism of differential forms, but with the goal of connecting to classical notation we sometimes will write the above equation as

$$< d\mathbf{e}_i, \mathbf{e}_k >= \omega_{ki}, \tag{3.20}$$

where $\{\mathbf{e}_i\}$ are vector calculus vectors forming an orthonormal basis.

3.10 Definition Let $\overline{\nabla}_X$ be a Koszul connection and let $\{e_i\}$ be a frame. The **Christoffel** symbols associated with the connection in the given frame are the functions $\Gamma^k{}_{ij}$ given by

$$\overline{\nabla}_{e_i} e_j = \Gamma^k{}_{ij} e_k \tag{3.21}$$

The Christoffel symbols are the coefficients that give the representation of the rate of change of the frame vectors in the direction of the frame vectors themselves. Many physicists therefore refer to the Christoffel symbols as the connection once again giving rise to possible confusion. The precise relation between the Christoffel symbols and the connection 1-forms is captured by the equations,

$$\omega^k{}_i(e_j) = \Gamma^k{}_{ij}, \tag{3.22}$$

or equivalently

$$\omega^k{}_i = \Gamma^k{}_{ij}\theta^j. \tag{3.23}$$

In a general frame in $\mathbf{R}^n$ there are n^2 entries in the connection 1-form and n^3 Christoffel symbols. The number of independent components is reduced if one assumes that the frame is orthonormal.

If $X = v^i e_i$ is a general vector field, then

$$\begin{aligned}\nabla_{e_j} X &= \nabla_{e_j}(v^i e_i)\\ &= v^i_{,j}\, e_i + v^i \Gamma^k{}_{ji} e_k\\ &= (v^i_{,j} + v^k \Gamma^i{}_{jk}) e_i \qquad (3.24)\end{aligned}$$

Which is denoted classically as the covariant derivative

$$v^i_{\|j} = v^i_{,j} + v^k \Gamma^i{}_{jk} \qquad (3.25)$$

3.11 Proposition Let and e_i be an orthonormal frame and $\overline{\nabla}_X$ be a Koszul connection compatible with the metric . Then

$$\omega_{ji} = -\omega_{ij} \qquad (3.26)$$

Proof: Since it is given that $< e_i, e_j >= \delta_{ij}$, we have

$$\begin{aligned}0 &= \overline{\nabla}_X < e_i, e_j >\\ &= < \overline{\nabla}_X e_i, e_j > + < e_i, \overline{\nabla}_X e_j >\\ &= < \omega^k_i e_k, e_j > + < e_i, \omega^k_j e_k >\\ &= \omega^k_i < e_k, e_j > + \omega^k_j < e_i, e_k >\\ &= \omega^k_i g_{kj} + \omega^k_j g_{ik}\\ &= \omega_{ji} + \omega_{ij}.\end{aligned}$$

thus proving that ω is indeed antisymmetric.

In an orthonormal frame in $\mathbf{R}^n$ the number of independent coefficients of the connection 1-form is $(1/2)n(n-1)$ since by antisymmetry, the diagonal entries are zero, and one only needs to count the number of entries in the upper triangular part of the $n \times n$ matrix ω_{ij}. Similarly, the number of independent Christoffel symbols gets reduced to $(1/2)n^2(n-1)$. Raising one index with g^{ij}, we find that $\omega^i{}_j$ is also antisymmetric, so in $\mathbf{R}^3$ the connection equations become

$$\overline{\nabla}_X [e_1, e_2, e_3] = [e_1, e_2, e_3] \begin{bmatrix} 0 & \omega^1{}_2(X) & \omega^1{}_3(X)\\ -\omega^1{}_2(X) & 0 & \omega^2{}_3(X)\\ -\omega^1{}_3(X) & -\omega^2{}_3(X) & 0 \end{bmatrix} \qquad (3.27)$$

Comparing the Frenet frame equation (1.35), we notice the obvious similarity to the general frame equations above. Clearly, the Frenet frame is a special case in which the basis vectors have been adapted to a curve resulting in a simpler connection in which some of the coefficients vanish. A further simplification occurs in the Frenet frame since here the equations represent the rate of change of the frame only along the direction of the curve rather than an arbitrary direction vector X. To unpack this transition from classical to modern notation, consider a unit speed curve $\beta(s)$. Then, as we discussed in section 1.15, we associate with the classical tangent vector $\mathbf{T} = \frac{d\mathbf{x}}{ds}$ the vector field $T = \beta'(s) = \frac{dx^i}{ds}\frac{\partial}{\partial x^i}$. Let $W = W(\beta(s)) = w^j(s)\frac{\partial}{\partial x^j}$ be an arbitrary vector field constrained to the curve. The rate of change of W along the curve is given by

$$\begin{aligned}
\overline{\nabla}_T W &= \overline{\nabla}_{(\frac{dx^i}{ds}\frac{\partial}{\partial x^i})}(w^j \frac{\partial}{\partial x^j}),\\
&= \frac{dx^i}{ds}\overline{\nabla}_{\frac{\partial}{\partial x^i}}(w^j \frac{\partial}{\partial x^j})\\
&= \frac{dx^i}{ds}\frac{\partial w^j}{\partial x^i}\frac{\partial}{\partial x^j}\\
&= \frac{dw^j}{ds}\frac{\partial}{\partial x^j}\\
&= W'(s).
\end{aligned}$$

3.4 Cartan Equations

Perhaps the most important contribution to the development of Differential Geometry is the work of Cartan culminating into the famous equations of structure discussed in this chapter.

First Structure Equation

3.12 Theorem Let $\{e_i\}$ be a frame with connection $\omega_j^{\ i}$ and dual coframe θ^i. Then

$$\Theta^i \equiv d\theta^i + \omega^i_{\ j} \wedge \theta^j = 0 \tag{3.28}$$

Proof: Let

$$e_i = \partial_j A^j_{\ i}.$$

be a frame, and let θ^i be the corresponding coframe. Since $\theta^i(e_j)$, we have

$$\theta^i = (A^{-1})^i_{\ j} dx^j.$$

Let X be an arbitrary vector field. Then

$$\begin{aligned}
\overline{\nabla}_X e_i &= \overline{\nabla}_X(\partial_j A^j_{\ i})\\
e_j \omega^j_{\ i}(X) &= \partial_j X(A^j_{\ i})\\
&= \partial_j d(A^j_{\ i})(X)\\
&= e_k (A^{-1})^k_{\ j} d(A^j_{\ i})(X)\\
\omega^k_{\ i}(X) &= (A^{-1})^k_{\ j} d(A^j_{\ i})(X).
\end{aligned}$$

Hence,

$$\omega^k_{\ i} = (A^{-1})^k_{\ j} d(A^j_{\ i}),$$

or, in matrix notation,

$$\omega = A^{-1} dA. \tag{3.29}$$

On the other hand, taking the exterior derivative of θ^i, we find that

$$\begin{aligned}
d\theta^i &= d(A^{-1})^i_{\ j} \wedge dx^j\\
&= d(A^{-1})^i_{\ j} \wedge A^j_{\ k}\theta^k\\
d\theta &= d(A^{-1})A \wedge \theta.
\end{aligned}$$

However, since $A^{-1}A = I$, we have $d(A^{-1})A = -A^{-1}dA = -\omega$, hence

$$d\theta = -\omega \wedge \theta. \tag{3.30}$$

In other words

$$d\theta^i + \omega^i{}_j \wedge \theta^j = 0.$$

3.13 Example $SO(2)$

Consider the polar coordinates part of the transformation in equation 3.5. Then the frame equations 3.6 in matrix form are given by:

$$\begin{bmatrix} e_1, e_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{3.31}$$

Thus, the attitude matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{3.32}$$

in this case, is a rotation matrix in $\mathbf{R}^2$. The set of all such matrices forms a continuous group (a **Lie Group**) called $SO(2)$. In such cases, the matrix $\omega = A^{-1}dA$ in equation 3.29 is called the **Maurer-Cartan** form of the group. An easy computation shows that for the rotation group $SO(2)$, the connection form is

$$\omega = \begin{bmatrix} 0 & -d\theta \\ d\theta & 0. \end{bmatrix} \tag{3.33}$$

Second Structure Equation

Let θ^i be a coframe in $\mathbf{R}^n$ with connection $\omega^i{}_j$. Taking the exterior derivative of the first equation of structure and recalling the properties (2.35), we get

$$\begin{aligned} d(d\theta^i) + d(\omega^i{}_j \wedge \theta^j) &= 0 \\ d\omega^i{}_j \wedge \theta^j - \omega^i{}_j \wedge d\theta^j &= 0. \end{aligned}$$

Substituting recursively from the first equation of structure, we get

$$\begin{aligned} d\omega^i{}_j \wedge \theta^j - \omega^i{}_j \wedge (-\omega^j{}_k \wedge \theta^k) &= 0 \\ d\omega^i{}_j \wedge \theta^j + \omega^i{}_k \wedge \omega^k{}_j \wedge \theta^j &= 0 \\ (d\omega^i{}_j + \omega^i{}_k \wedge \omega^k{}_j) \wedge \theta^j &= 0 \\ d\omega^i{}_j + \omega^i{}_k \wedge \omega^k{}_j &= 0. \end{aligned}$$

3.14 Definition The curvature Ω of a connection ω is the matrix valued 2-form,

$$\Omega^i{}_j \equiv d\omega^i{}_j + \omega^i{}_k \wedge \omega^k{}_j. \tag{3.34}$$

3.15 Theorem Let θ be a coframe with connection ω in $\mathbf{R}^n$. Then the curvature form vanishes:

$$\Omega = d\omega + \omega \wedge \omega = 0. \tag{3.35}$$

Proof: Given that there is a non-singular matrix A such that $\theta = A^{-1}dx$ and $\omega = A^{-1}dA$, we have

$$d\omega = d(A^{-1}) \wedge dA.$$

On the other hand,

$$\begin{aligned}\omega \wedge \omega &= (A^{-1}dA) \wedge (A^{-1}dA)\\ &= -d(A^{-1})A \wedge A^{-1}dA\\ &= -d(A^{-1})(AA^{-1}) \wedge dA\\ &= -d(A^{-1}) \wedge dA.\end{aligned}$$

Therefore, $d\omega = -\omega \wedge \omega$.
We should point out the slight abuse of the wedge notation. The connection ω is matrix valued, so the symbol $\omega \wedge \omega$ is really a composite of matrix and wedge multiplication.

3.16 Example Sphere Frame
The frame for spherical coordinates 3.7 in matrix form is

$$[e_r, e_\theta, e_\phi] = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right] \begin{bmatrix} \sin\theta\cos\phi & \cos\theta\cos\phi & -\sin\phi \\ \sin\theta\sin\phi & \cos\theta\sin\phi & \cos\phi \\ \cos\theta & -\sin\theta & 0 \end{bmatrix}.$$

Hence,

$$A^{-1} = \begin{bmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0, \end{bmatrix}$$

and

$$dA = \begin{bmatrix} \cos\theta\cos\phi\, d\theta - \sin\theta\sin\phi\, d\phi & -\sin\theta\cos\phi\, d\theta - \cos\theta\sin\phi\, d\phi & -\cos\phi\, d\phi \\ \cos\theta\sin\phi\, d\theta + \sin\theta\cos\phi\, d\phi & -\sin\theta\sin\phi\, d\theta + \cos\theta\cos\phi\, d\phi & -\sin\phi\, d\phi \\ -\sin\theta\, d\theta & -\cos\theta\, d\theta & 0 \end{bmatrix}.$$

Since the $\omega = A^{-1}dA$ is antisymmetric, it suffices to compute:

$$\begin{aligned}\omega^1{}_2 &= [-\sin^2\theta\cos^2\phi - \sin^2\theta\sin^2\phi \; - \cos^2\theta]\, d\theta\\ &\qquad + [\sin\theta\cos\theta\cos\phi\sin\phi - \sin\theta\cos\theta\cos\phi\sin\phi]\, d\phi\\ &= -d\theta\\ \omega^1{}_3 &= [-\sin\theta\cos^2\phi - \sin\theta\sin^2\phi]\, d\phi = -\sin\theta\, d\phi\\ \omega^2{}_3 &= [-\cos\theta\cos^2\phi - \cos\theta\sin^2\phi]\, d\phi = -\cos\theta\, d\phi\end{aligned}$$

We conclude that the matrix-valued connection one form is

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

A slicker computation of the connection form can be obtained by a method of educated guessing working directly from the structure equations. We have that the dual one forms are:

$$\begin{aligned}\theta^1 &= dr\\ \theta^2 &= r\, d\theta\\ \theta^3 &= r\sin\theta\, d\phi.\end{aligned}$$

Then

$$
\begin{aligned}
d\theta^2 &= -d\theta \wedge dr \\
&= -\omega^2{}_1 \wedge \theta^1 - \omega^2{}_3 \wedge \theta^3.
\end{aligned}
$$

So, on a first iteration we guess that $\omega^2{}_1 = d\theta$. The component $\omega^2{}_3$ is not necessarily 0 because it might contain terms with $d\phi$. Proceeding in this manner, we compute:

$$
\begin{aligned}
d\theta^3 &= \sin\theta \, dr \wedge d\phi + r\cos\theta \, d\theta \wedge d\phi \\
&= -\sin\theta \, d\phi \wedge dr - \cos\theta \, d\phi \wedge r \, d\theta \\
&= -\omega^3{}_1 \wedge dr \wedge \theta^1 - \omega^3{}_2 \wedge \theta^2.
\end{aligned}
$$

Now we guess that $\omega^3{}_1 = \sin\theta \, d\phi$, and $\omega^3{}_2 = \cos\theta \, d\phi$. Finally, we insert these into the full structure equations and check to see if any modifications need to be made. In this case, the forms we have found are completely compatible with the first equation of structure, so these must be the forms. The second equations of structure are much more straight-forward to verify. For example

$$
\begin{aligned}
d\omega^2{}_3 &= d(-\cos\theta \, d\phi) \\
&= \sin\theta \, d\theta \wedge d\phi \\
&= -d\theta \wedge (-\sin\theta \, d\phi) \\
&= -\omega^2{}_1 \wedge \omega^1{}_3.
\end{aligned}
$$

Change of Basis

We briefly explore the behavior of the quantities Θ^i and $\Omega^i{}_j$ under a change of basis.

Let e_i be frame with dual forms θ^i, and let $\overline{e}_i$ be another frame related to the first frame by an invertible transformation.

$$\overline{e}_i = e_j B^j{}_i, \tag{3.36}$$

which we will write in matrix notation as $\overline{e} = eB$. Referring back to the definition of connections (3.17), we introduce the **covariant differential** $\overline{\nabla}$ by the formula

$$
\begin{aligned}
\overline{\nabla} e_i &= e_j \otimes \omega^j{}_i \\
&= e_j \omega^j{}_i \\
\overline{\nabla} e &= e \, \omega
\end{aligned} \tag{3.37}
$$

where, once again, we have simplified the equation by using matrix notation. This definition is elegant because it does not explicitly show the dependence on X in the connection (3.17). The idea of switching from derivatives to differentials is familiar from basic calculus. Consistent with equation 3.20, the vector calculus notation for equation 3.37 would be

$$d\mathbf{e_i} = \mathbf{e_j} \, \omega^j{}_i. \tag{3.38}$$

However, we point out that in the present context, the situation is much more subtle. The operator $\overline{\nabla}$ here maps a vector field to a matrix-valued tensor of rank $T^{1,1}$. Another way to view the covariant differential is to think of $\overline{\nabla}$ as an operator such that if e is a frame, and X a vector field, then $\overline{\nabla} e(X) = \overline{\nabla}_X e$. If f is a function, then $\overline{\nabla} f(X) = \overline{\nabla}_X f = df(X)$, so that $\overline{\nabla} f = df$. In other words,

$\overline{\nabla}$ behaves like a covariant derivative on vectors, but like a differential on functions. We require $\overline{\nabla}$ to behave like a derivation on tensor products:

$$\overline{\nabla}(T_1 \otimes T_2) = \overline{\nabla}T_1 \otimes T_2 + T_1 \otimes \overline{\nabla}T_2. \tag{3.39}$$

Taking the exterior differential of (3.36) and using (3.37) recursively, we get

$$\begin{aligned}
\overline{\nabla}\overline{e} &= \overline{\nabla}e \otimes B + e \otimes \overline{\nabla}B \\
&= (\overline{\nabla}e)B + e(dB) \\
&= e\,\omega B + e(dB) \\
&= \overline{e}B^{-1}\omega B + \overline{e}B^{-1}dB \\
&= \overline{e}[B^{-1}\omega B + B^{-1}dB] \\
&= \overline{e}\,\overline{\omega}
\end{aligned}$$

provided that the connection $\overline{\omega}$ in the new frame $\overline{e}$ is related to the connection ω by the transformation law,

$$\overline{\omega} = B^{-1}\omega B + B^{-1}dB. \tag{3.40}$$

It should be noted than if e is the standard frame $e_i = \partial_i$ in $\mathbf{R}^n$, then $\overline{\nabla}e = 0$, so that $\omega = 0$. In this case, the formula above reduces to $\overline{\omega} = B^{-1}dB$, showing that the transformation rule is consistent with equation (3.29).

3.17 Example Suppose that B is a change of basis consisting of a rotation by an angle θ about e_3 The transformation is a an isometry that can be represented by the orthogonal rotation matrix

$$B = \begin{bmatrix} \cos\theta & \sin -\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.41}$$

Carrying out the computation for the change of basis 3.40, we find:

$$\begin{aligned}
&\overline{\omega}^1{}_2 = \omega^1{}_2 - d\theta, \\
&\overline{\omega}^1{}_3 = \cos\theta\,\omega^1{}_3 + \sin\theta\,\omega^2{}_3, \\
&\overline{\omega}^2{}_3 = -\sin\theta\,\omega^1{}_3 + \cos\theta\,\omega^2{}_3.
\end{aligned} \tag{3.42}$$

The $B^{-1}dB$ part of the transformation only affects the $\omega^1{}_2$ term, and the effect is just adding $d\theta$ much like the case of the Maurer-Cartan form for SO_2 above.

Chapter 4

Theory of Surfaces

4.1 Manifolds

4.1 Definition A coordinate chart in $\mathcal{M} \in \mathbf{R}^3$ is a differentiable map $\mathbf{x}$ from an open subset V of $\mathbf{R}^2$ onto a set $U \in \mathcal{M}$.

$$\begin{aligned} \mathbf{x} : V \in \mathbf{R}^2 &\longrightarrow \mathbf{R}^3 \\ (u,v) &\overset{\mathbf{x}}{\longmapsto} (x(u,v), y(u,v), z(u,v)) \end{aligned} \tag{4.1}$$

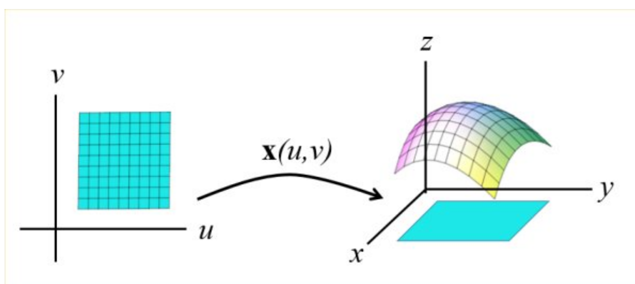


Fig. 4.1: Surface

Each set $U = \mathbf{x}(V)$ is called a **Coordinate Neighborhood** of $\mathcal{M}$. We require that the the Jacobian of the map has maximal rank. In local coordinates, a coordinate chart is represented by three equations in two variables:

$$x^i = f^i(u^\alpha), \text{ where } i = 1,2,3, \quad \alpha = 1,2. \tag{4.2}$$

The local coordinate representation allows us to use the tensor index formalism introduced in earlier chapters. The assumption that the Jacobian $J = (\partial x^i / \partial u^\alpha)$ be of maximal rank allows one to invoke the Implicit Function Theorem. Thus, in principle, one can locally solve for one of the coordinates, say x^3, in terms of the other two, like so:

$$x^3 = f(x^1, x^2). \tag{4.3}$$

The locus of points in $\mathbf{R}^3$ satisfying the equations $x^i = f^i(u^\alpha)$ can also be locally represented by an expression of the form

$$F(x^1, x^2, x^3) = 0 \tag{4.4}$$

4.2 Definition Let U_i and U_j be two coordinate neighborhoods of a point $\mathbf{p} \in \mathcal{M}$ with corresponding charts $\mathbf{x}(u^1, u^2) : V_i \longrightarrow U_i \in \mathbf{R}^3$ and $\mathbf{y}(v^1, v^2) : V_j \longrightarrow U_j \in \mathbf{R}^3$ with a non-empty intersection $U_i \cap U_j \neq \emptyset$. On the overlaps, the maps $\phi_{ij} = \mathbf{x}^{-1}\mathbf{y}$ are called transition functions or coordinate transformations. (see Figure 4.2)

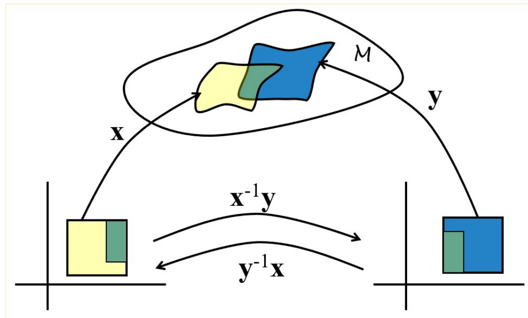


Fig. 4.2: Coordinate Charts

4.3 Definition A **differentiable manifold** of dimension 2, is a space $\mathcal{M}$ together with an indexed collection $\{U_\alpha\}_{\alpha \in I}$ of coordinate neighborhoods satisfying the following properties:

1. The neighborhoods $\{U_\alpha\}$ constitute an open cover $\mathcal{M}$. That is, if $\mathbf{p} \in \mathcal{M}$, then $\mathbf{p}$ belongs to some chart.
2. For any pair of coordinate neighborhoods U_i and U_j with $U_i \cap U_j \neq \emptyset$, the transition maps ϕ_{ij} and their inverses are differentiable.
3. An indexed collection satisfying the conditions above is called an **Atlas**. We require the atlas to be maximal in the sense that it contains all possible coordinate neighborhoods.

The overlapping coordinate patches represent different parametrizations for the same set of points in $\mathbf{R}^3$. Part (2) of the definition insures that on the overlap, the coordinate transformations are invertible. Part (3) is included for technical reasons, although in practice the condition is superfluous. A family of coordinate neighborhoods satisfying conditions (1) and (2) can always be extended to a maximal atlas. This can be shown from the fact that $\mathcal{M}$ inherits a subspace topology consisting of open sets which are defined by the intersection of open sets in $\mathbf{R}^3$ with $\mathcal{M}$.

If the coordinate patches in the definition map from $\mathbf{R}^n$ to $\mathbf{R}^m$ $n < m$ we say the $\mathcal{M}$ is a n-dimensional submanifold embedded in $\mathbf{R}^m$. In fact, one could define an an abstract manifold without the reference to the embedding space by starting with a topological space $\mathcal{M}$ that is locally Euclidean via homeomorphic coordinate patches and has a differentiable structure as in the definition above. However, it turns out that any differentiable manifold of dimension n can be embedded in $\mathbf{R}^{2n}$, as proved by Whitney in a theorem that is beyond the scope of these notes.

A 2-dimensional manifold embedded in $\mathbf{R}^3$ in which the transition functions are $C^\infty$is called a smooth surface. The first condition in the definition states that each coordinate neighborhood looks locally like a subset of $\mathbf{R}^2$. The second differentiability condition indicates that the patches are joined together smoothly as some sort of quilt. We summarize this notion by saying that a manifold is a space that is **locally Euclidean** and has a **differentiable structure** so that the notion of differentiation makes sense. Of course, $\mathbf{R}^n$ is itself an n dimensional manifold.

The smoothness condition on the coordinate component functions $x^i(u^\alpha)$ implies that at any point $x^i(u_0^\alpha + h^\alpha)$ near a point $x^i(u_0^\alpha) = x^i(u_0, v_0)$, the functions admit a Taylor expansion

$$x^i(u_0^\alpha + h^\alpha) = x^i(u_0^\alpha) + h^\alpha \left(\frac{\partial x^i}{\partial u^\alpha}\right)_0 + \frac{1}{2!} h^\alpha h^\beta \left(\frac{\partial^2 x^i}{\partial u^\alpha \partial u^\beta}\right)_0 + \ldots \tag{4.5}$$

Since the parameters u^α must enter independently, the Jacobian matrix

$$J \equiv \left[\frac{\partial x^i}{\partial u^\alpha}\right] = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{bmatrix}$$

must have maximal rank. At points where J has rank 0 or 1, there is a singularity in the coordinate patch.

4.4 Example Consider the local coordinate chart for the unit sphere obtained by setting $r = 1$ in the equations for spherical coordinates 2.20

$$\mathbf{x}(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

The vector equation is equivalent to three scalar functions in two variables:

$$\begin{aligned} x &= \sin\theta\cos\phi, \\ y &= \sin\theta\sin\phi, \\ z &= \cos\phi. \end{aligned} \tag{4.6}$$

Clearly, the surface represented by this chart is part of the sphere $x^2 + y^2 + z^2 = 1$. The chart cannot possibly represent the whole sphere because, although a sphere is locally Euclidean, (the earth is locally flat) there is certainly a topological difference between a sphere and a plane. Indeed, if one analyzes the coordinate chart carefully, one will note that at the North pole ($\theta = 0,\ z = 1$, the coordinates become singular. This happens because $\theta = 0$ implies that $x = y = 0$ regardless of the value of ϕ, so that the North pole has an infinite number of labels. In this coordinate patch, the Jacobian at the North Pole does not have maximal rank. To cover the entire sphere, one would need at least two coordinate patches. In fact, introducing an exactly analogous patch $\mathbf{y}(u.v)$ based on South pole would suffice, as long as in overlap around the equator functions $\mathbf{x}^{-1}\mathbf{y}$, and $\mathbf{y}^{-1}\mathbf{x}$ are smooth. One could conceive more elaborate coordinate patches such as those used in baseball and tennis balls.

The fact that it is required to have two parameters to describe a patch on a surface in $\mathbf{R}^3$ is a manifestation of the 2-dimensional nature of of the surfaces. If one holds one of the parameters constant while varying the other, then the resulting 1-parameter equation describes a curve on the surface. Thus, for example, letting $\phi = \textit{constant}$ in equation (4.6), we get the equation of a meridian great circle.

4.5 Example Surface of Revolution
Given a function $f(r)$, the coordinate chart

$$\mathbf{x}(r, \phi) = (r\cos\phi, r\sin\phi, f(r)) \tag{4.7}$$

represents a surface of revolution around the z-axis in which the cross section profile has the shape of the function. Horizontal cross-sections are circles of radius r. In figure 4.3, we have chosen the function $f(r) = e^{-r^2}$ to be a Gaussian, so the surface of revolution is bell-shaped. A lateral curve profile for $\phi = \pi/4$ is shown in black.

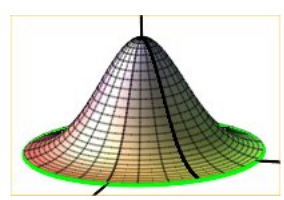


Fig. 4.3: Bell

We should point out that this parametrization of surfaces of revolution is fairly constraining because of the requirement of $z = f(r)$ to be a function. Thus, for instance, the parametrization will not work for surfaces of revolution generated by closed curves. In the next example, we illustrate how one easily get around this constraint

4.6 Example Torus

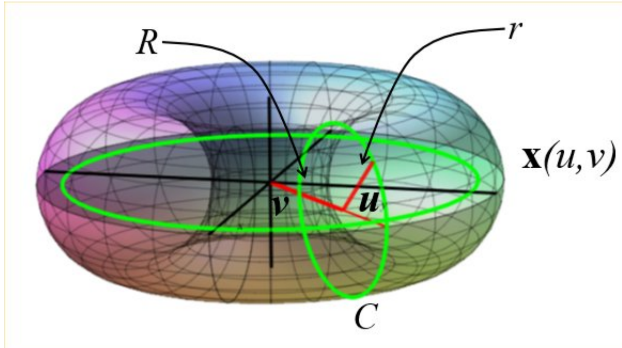


Fig. 4.4: Torus

Consider the surface of revolution generated by rotating a circle C of radius r around a parallel axis located a distance R from its center as shown in figure 4.4
The resulting surface called a Torus can be parametrized by the coordinate patch

$$\mathbf{x}(u, v) = ((R + r\cos u)\cos v, (R + r\cos u)\sin v, r\sin u). \tag{4.8}$$

Here the angle u traces points around the z-axis, whereas the angle v traces points around the circle C. (At the risk of some confusion in notation, (the parameters in the figure are bold-faced; this is done solely for the purpose of visibility.) The projection of a point in the surface of the Torus onto the xy-plane is located at a distance $(R + r\cos u)$ from the origin. Thus, the x and y coordinates of the point in the torus are just the polar coordinates of the projection of the point in the plane. The z-coordinate corresponds to the height of a right triangle with radius r and opposite angle u.

4.7 Example Cartesian Equation
Surfaces in $\mathbf{R}^3$ are first introduced in vector calculus by a function of two variables $z = f(x, y)$. We will find it useful for consistency to use the obvious parametrization

$$\mathbf{x}(u, v) = (u, v, f(u, v)). \tag{4.9}$$

4.8 Notation Given a parametrization of a surface in a local chart $\mathbf{x}(u, v) = \mathbf{x}(u^1, u^2) = \mathbf{x}(u^\alpha)$, we will denote the partial derivatives by any of the following notations:

$$\begin{aligned}
\mathbf{x}_u = \mathbf{x}_1 &= \frac{\partial \mathbf{x}}{\partial u}, & \mathbf{x}_{uu} = \mathbf{x}_{11} &= \frac{\partial^2 \mathbf{x}}{\partial u^2} \\
\mathbf{x}_v = \mathbf{x}_2 &= \frac{\partial \mathbf{x}}{\partial v}, & \mathbf{x}_{vv} = \mathbf{x}_{22} &= \frac{\partial^2 \mathbf{x}}{\partial v^2} \\
\mathbf{x}_\alpha &= \frac{\partial \mathbf{x}}{\partial u^\alpha} & \mathbf{x}_{\alpha\beta} &= \frac{\partial^2 \mathbf{x}}{\partial u^\alpha \partial v^\beta}
\end{aligned}$$

4.2 The First Fundamental Form

Let $x^i(u^\alpha)$ be a local parametrization of a surface. Then, the Euclidean inner product in $\mathbf{R}^3$ induces an inner product in the space of tangent vectors at each point in the surface. This metric on the surface is obtained as follows:

$$\begin{aligned} dx^i &= \frac{\partial x^i}{\partial u^\alpha} du^\alpha \\ ds^2 &= \delta_{ij} dx^i dx^j \\ &= \delta_{ij} \frac{\partial x^i}{\partial u^\alpha} \frac{\partial x^j}{\partial u^\beta} du^\alpha du^\beta. \end{aligned}$$

Thus,

$$ds^2 = g_{\alpha\beta} du^\alpha du^\beta, \tag{4.10}$$

where

$$g_{\alpha\beta} = \delta_{ij} \frac{\partial x^i}{\partial u^\alpha} \frac{\partial x^j}{\partial u^\beta}. \tag{4.11}$$

We conclude that the surface, by virtue of being embedded in $\mathbf{R}^3$, inherits a natural metric (4.10) which we will call the **induced metric**. A pair $\{\mathcal{M}, g\}$, where $\mathcal{M}$ is a manifold and $g = g_{\alpha\beta} du^\alpha \otimes du^\beta$ is a metric is called a **Riemannian manifold** if considered as an entity in itself, and a Riemannian submanifold of $\mathbf{R}^n$ if viewed as an object embedded in Euclidean space. An equivalent version of the metric (4.10) can be obtained by using a more traditional calculus notation:

$$\begin{aligned} d\mathbf{x} &= \mathbf{x}_u du + \mathbf{x}_v dv \\ ds^2 &= d\mathbf{x} \cdot d\mathbf{x} \\ &= (\mathbf{x}_u du + \mathbf{x}_v dv) \cdot (\mathbf{x}_u du + \mathbf{x}_v dv) \\ &= (\mathbf{x}_u \cdot \mathbf{x}_u) du^2 + 2(\mathbf{x}_u \cdot \mathbf{x}_v) du dv + (\mathbf{x}_v \cdot \mathbf{x}_v) dv^2. \end{aligned}$$

We can rewrite the last result as

$$ds^2 = E du^2 + 2F du dv + G dv^2, \tag{4.12}$$

where

$$\begin{aligned} E &= g_{11} = \mathbf{x}_u \cdot \mathbf{x}_u \\ F &= g_{12} = \mathbf{x}_u \cdot \mathbf{x}_v \\ &= g_{21} = \mathbf{x}_v \cdot \mathbf{x}_u \\ G &= g_{22} = \mathbf{x}_v \cdot \mathbf{x}_v. \end{aligned}$$

That is

$$g_{\alpha\beta} = \mathbf{x}_\alpha \cdot \mathbf{x}_\beta = < \mathbf{x}_\alpha, \mathbf{x}_\beta > .$$

4.9 Definition The element of arclength,

$$ds^2 = g_{\alpha\beta} du^\alpha \otimes du^\beta, \tag{4.13}$$

is also called the **first fundamental form**. We must caution the reader that this quantity is not a form in the sense of differential geometry since ds^2 involves the symmetric tensor product rather than the wedge product.

The first fundamental form plays such a crucial role in the theory of surfaces that we will find it convenient to introduce a third, more modern version. Following the same development as in the theory of curves, consider a surface $\mathcal{M}$ defined locally by a function $\mathbf{q} = (u^1, u^2) \longmapsto \mathbf{p} = \alpha(u^1, u^2)$. We say that a quantity X_p is a tangent vector at a point $\mathbf{p} \in \mathcal{M}$ if X_p is a linear derivation on the space of C^∞ real-valued functions $\mathcal{F} = \{f | f : \mathcal{M} \longrightarrow \mathbf{R}\}$ on the surface. The set of all tangent vectors at a point $\mathbf{p} \in \mathcal{M}$ is called the **tangent space** $T_p\mathcal{M}$. As before, a vector field X on the surface is a smooth choice of a tangent vector at each point on the surface and the union of all tangent spaces is called the **tangent bundle** $T\mathcal{M}$.

The coordinate chart map $\alpha : \mathbf{R}^2 \longrightarrow \mathcal{M} \in \mathbf{R}^3$ induces a **Push-Forward** map $\alpha_* : T\mathbf{R}^2 \longrightarrow T\mathcal{M}$ which maps a vector field V in $T \in \mathbf{R}^2$ into a vector field $X = \alpha_*(V) \in T\mathcal{M}$, as illustrated in the diagram 4.5

$$\begin{tikzcd} V \in T\mathbf{R}^2 \arrow[r, "\alpha_*"] \arrow[d] & T\mathcal{M} \arrow[d] & \\ q \in \mathbf{R}^2 \arrow[r, "\alpha"] & \mathcal{M} \in \mathbf{R}^3 \arrow[r, "f"] & \mathbf{R} \end{tikzcd}$$

Fig. 4.5: Push-Forward

The action of the Push-Forward is defined by

$$\alpha_*(V)(f) \mid_{\alpha(q)} = V(f \circ \alpha) \mid_q . \tag{4.14}$$

Just as in the case of curves, when we revert back to classical notation to describe a surface as $x^i(u^\alpha)$, what we really mean is $(x^i \circ \alpha)(u^\alpha)$, where x^i are the coordinate functions in $\mathbf{R}^3$. Particular examples of tangent vectors on $\mathcal{M}$ are given by the push-forward of the standard basis of $T\mathbf{R}^2$. These tangent vectors which earlier we called $\mathbf{x}_\alpha$ are defined by

$$\alpha_*(\frac{\partial}{\partial u^\alpha})(f) \mid_{\alpha(u^\alpha)} = \frac{\partial}{\partial u^\alpha}(f \circ \alpha) \mid_{u^\alpha}$$

In this formalism, the first fundamental form I is just the symmetric bilinear tensor defined by the induced metric,

$$I(X, Y) = g(X, Y) = < X, Y >, \tag{4.15}$$

where X and Y are any pair of vector fields in $T\mathcal{M}$.

Orthogonal Parametric Curves

Let V and W be vectors tangent to a surface $\mathcal{M}$ defined locally by a chart $\mathbf{x}(u^\alpha)$. Since the vectors $\mathbf{x}_\alpha$ span the tangent space of $\mathcal{M}$ at each point, the vectors V and W can be written as linear combinations,

$$\begin{aligned} V &= V^\alpha \mathbf{x}_\alpha \\ W &= W^\alpha \mathbf{x}_\alpha. \end{aligned}$$

The functions V^α and W^α are called the **curvilinear coordinates** of the vectors. We can calculate the length and the inner product of the vectors using the induced Riemannian metric as follows:

$$\begin{aligned} \|V\|^2 &= < V, V > = < V^\alpha \mathbf{x}_\alpha, V^\beta \mathbf{x}_\beta > = V^\alpha V^\beta < \mathbf{x}_\alpha, \mathbf{x}_\beta > \\ \|V\|^2 &= g_{\alpha\beta} V^\alpha V^\beta \\ \|W\|^2 &= g_{\alpha\beta} W^\alpha W^\beta, \end{aligned}$$

and

$$\begin{aligned} < V, W > &= < V^\alpha \mathbf{x}_\alpha, W^\beta \mathbf{x}_\beta > = V^\alpha W^\beta < \mathbf{x}_\alpha, \mathbf{x}_\beta > \\ &= g_{\alpha\beta} V^\alpha W^\beta. \end{aligned}$$

The angle θ subtended by the the vectors V and W is the given by the equation

$$\begin{aligned}\cos\theta &= \frac{<V,W>}{\|V\|\cdot\|W\|}\\ &= \frac{I(V,W)}{\sqrt{I(V,V)}\sqrt{I(W,W)}}\\ &= \frac{g_{\alpha_1\beta_1}V^{\alpha_1}W^{\beta_1}}{g_{\alpha_2\beta_2}V^{\alpha_2}V^{\beta_2}\cdot g_{\alpha_3\beta_3}W^{\alpha_3}W^{\beta_3}},\end{aligned}\tag{4.16}$$

where the numerical subscripts are needed for the α and β indices to comply with Einstein's summation convention.

Let $u^\alpha = \phi^\alpha(t)$ and $u^\alpha = \psi^\alpha(t)$ be two curves on the surface. Then the total differentials

$$du^\alpha = \frac{d\phi^\alpha}{dt}dt, \quad\text{and}\quad \delta u^\alpha = \frac{d\psi^\alpha}{dt}\delta t$$

represent infinitesimal tangent vectors (1.20) to the curves. Thus, the angle between two infinitesimal vectors tangent to two intersecting curves on the surface satisfies the equation:

$$\cos\theta = \frac{g_{\alpha_1\beta_1}du^{\alpha_1}\delta u^{\beta_1}}{\sqrt{g_{\alpha_2\beta_2}du^{\alpha_2}du^{\beta_2}}\sqrt{g_{\alpha_3\beta_3}\delta u^{\alpha_3}\delta u^{\beta_3}}}\tag{4.17}$$

In particular, if the two curves happen to be the parametric curves, $u^1 =$ const. and $u^2 =$const., then along one curve we have $du^1 = 0$, with du^2 arbitrary, and along the second δu^1 is arbitrary and $\delta u^2 = 0$. In this case, the cosine of the angle subtended by the infinitesimal tangent vectors reduces to:

$$\cos\theta = \frac{g_{12}\delta u^1 du^2}{\sqrt{g_{11}(\delta u^1)^2}\sqrt{g_{22}(du^2)^2}} = \frac{g_{12}}{g_{11}g_{22}} = \frac{F}{\sqrt{EG}}.\tag{4.18}$$

A simpler way to obtain this result is to recall that that parametric directions are given by $\mathbf{x}_u$ and $\mathbf{x}_v$, so

$$\cos\theta = \frac{<\mathbf{x}_u,\mathbf{x}_v>}{\|\mathbf{x}_u\|\cdot\|\mathbf{x}_v\|} = \frac{F}{\sqrt{EG}}.\tag{4.19}$$

It follows immediately from the equation above that:

4.10 Proposition The parametric curves are orthogonal if $F = 0$.

Orthogonal parametric curves are an important class because locally the coordinate grid on the surface is similar to coordinate grids in basic calculus, such as in polar coordinates for which $ds^2 = dr^2 + r^2 d\theta^2$.

4.11 Examples a) Sphere

$$\begin{aligned}\mathbf{x} &= (a\sin\theta\cos\phi, a\sin\theta\sin\phi, a\cos\theta)\\ \mathbf{x}_\theta &= (a\cos\theta\cos\phi, a\cos\theta\sin\phi, -a\sin\theta)\\ \mathbf{x}_\phi &= (-a\sin\theta\sin\phi, a\sin\theta\cos\phi, 0)\\ E &= \mathbf{x}_\theta\cdot\mathbf{x}_\theta = a^2\\ F &= \mathbf{x}_\theta\cdot\mathbf{x}_\phi = 0\\ G &= \mathbf{x}_\phi\cdot\mathbf{x}_\phi = a^2\sin^2\theta\\ ds^2 &= a^2 d\theta^2 + a^2\sin^2\theta\, d\phi^2 \qquad (4.20)\end{aligned}$$

There are many interesting curves on a sphere, but amongst these the **Loxodromes** have a special role in history. A loxodrome is a curve that winds around a sphere making a constant angle with the meridians. In this sense, it is the spherical analog of a cylndrical helix and as such it is often called a spherical helix. The curves were significant in early navigation where they are referred as **Rhumb** lines. As people in the late 1400's began to rediscover that earth was not flat, cartographers figured out methods to render maps on flat paper surfaces. One such technique is called the **Mercator** projection which is obtained by projecting the sphere onto a plane that wraps around the sphere as a cylinder tangential to the sphere along the equator

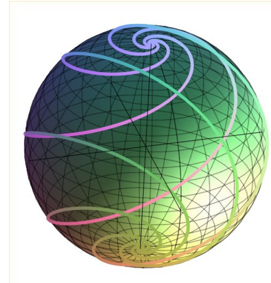


Fig. 4.6: Loxodrome

As we will discuss in more detail later, a navigator travelling a constant bearing would be moving a on path the is a straight line on the Mercator projection map, but on the sphere it would be spiraling ever faster as one approached the poles. Thus it became important to understand the nature of such paths. It appears as if the first quantitative treatise of loxodromes was carried in the mid 1500's by the Portuguese applied mathematician Pedro Nuñes, who was chair of the department at the University of Coimbra.

As an application, we will derive the equations of loxodromes and compute the arc length. A general spherical curve can be parametrized in the form $\gamma(t) = \mathbf{x}(\theta(t), \phi(t))$. Let σ be the angle the curve makes with the meridians $\phi =$ constant. Then, recalling that $< \mathbf{x}_\theta, \mathbf{x}_\phi >= F = 0$, we have:

$$
\begin{aligned}
\gamma' &= \mathbf{x}_\theta \frac{d\theta}{dt} + \mathbf{x}_\phi \frac{d\phi}{dt}. \\
\cos\sigma &= \frac{< \mathbf{x}_\theta, \gamma' >}{\|\mathbf{x}_\theta\| \cdot \|\gamma'\|} = \frac{E\frac{d\theta}{dt}}{\sqrt{E}\frac{ds}{dt}} = a\frac{d\theta}{ds}. \\
a^2 d\theta^2 &= \cos^2\sigma \; ds^2, \\
a^2 \sin^2\sigma \; d\theta^2 &= a^2 \cos^2\sigma \sin^2\theta \; d\phi^2, \\
\sin\sigma \; d\theta &= \pm\cos\sigma\sin\theta \; d\phi, \\
\csc\theta \; d\theta &= \pm\cot\sigma \; d\phi.
\end{aligned}
$$

The convention use throughout history by cartographers, is to measure the angle θ from the equations. To better adhere to the history, but at the same time avoiding confusion, we replace θ with $\vartheta = \frac{\pi}{2} - \theta$, so that $\vartheta = 0$ corresponds to the equator. Integrating the last equation with this change, we get

$$
\begin{aligned}
\sec\vartheta \; d\vartheta &= \pm\cot\sigma \; d\phi \\
\ln\tan(\tfrac{\vartheta}{2} + \tfrac{\pi}{4}) &= \pm\cot\sigma(\phi - \phi_0).
\end{aligned}
$$

Thus, we conclude that the equations of loxodromes and their arc lengths are given by

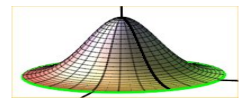
$$\phi = \pm(\tan\sigma)\ln\tan(\tfrac{\vartheta}{2} + \tfrac{\pi}{4}) + \phi_0 \tag{4.21}$$

$$s = a(\theta - \theta_0)\sec\sigma, \tag{4.22}$$

where θ_0 and ϕ_0 are the coordinates of the initial position. Figure 4.6 shows four loxodromes equally distributed around the sphere.

Loxodromes were the bases for a number beautiful drawings and woodcuts by M. C. Escher. We will show later under the discussion of conformal (angle preserving) maps, that loxodromes map into straight lines making a constant angle with meridians in the Mercator projection.

b) Surface of Revolution



$$
\begin{aligned}
\mathbf{x} &= (r\cos\theta, r\sin\theta, f(r)) \\
\mathbf{x}_r &= (\cos\theta, \sin\theta, f'(r)) \\
\mathbf{x}_\theta &= (-r\sin\theta, r\cos\theta, 0) \\
E &= \mathbf{x}_r \cdot \mathbf{x}_r = 1 + f'^2(r) \\
F &= \mathbf{x}_r \cdot \mathbf{x}_\theta = 0 \\
G &= \mathbf{x}_\theta \cdot \mathbf{x}_\theta = r^2 \\
ds^2 &= [1 + f'^2(r)]dr^2 + r^2 d\theta^2
\end{aligned}
$$

As in 4.7, we have chosen a Gaussian profile to illustrate a surface of revolution. Since $F = 0$ the parametric lines are orthogonal. The picture shows that this is indeed the case. At any point of the surface, the analogs of meridians and parallels intersect at right angles.

c) Pseudosphere

$$
\begin{aligned}
\mathbf{x} &= (a\sin u\cos v, a\sin u\sin v, a(\cos u + \ln(\tan\frac{u}{2})) \\
E &= a^2\cot^2 u \\
F &= = 0 \\
G &= a^2\sin^2 u \\
ds^2 &= a^2\cot^2 u\; du^2 + a^2\sin^2 u\; dv^2
\end{aligned}
$$

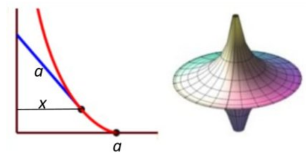


Fig. 4.7: Pseudosphere

The pseudosphere is a surface of revolution in which the profile curve is a **tractrix**. The tractrix curve was originated by a problem posed by Leibnitz to the effect of finding the path traced by a point initially placed on the horizontal axis at a distance a from the origin, as it was pulled along the vertical axis by a taught string of constant length a, as shown in figure 4.7. The tractrix was later studied by Huygens in 1692. Colloquially this is the path of a reluctant dog at $(a, 0)$ dragged by a man walking up the z-axis. The tangent segment is the hypothenuse of a right triangle with base x and height $\sqrt{a^2 - x^2}$, so the slope is $dz/dx = -\sqrt{a^2 - x^2}/x$. Using the trigonometric substitution $x = a\sin u$, we get $z = a\int(\cos^2 u/\sin u)\; du$ which leads to the appropriate form for the profile of the surface of revolution. The pseudosphere was studied by Beltrami in 1868, who discovered that in spite of the surface extending asymptotically to infinity, the surface area was finite with $S = 4\pi a^2$ as in a sphere of the same radius, and the volume enclosed was half that sphere. We will have much more to say about this surface.

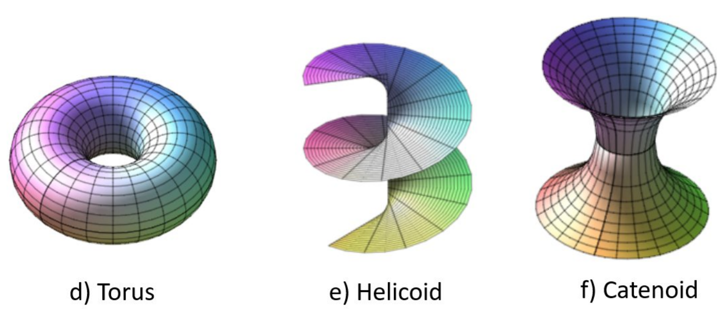


Fig. 4.8: Examples of Surfaces

d) Torus

$$\begin{aligned}
\mathbf{x} &= ((b+a\cos u)\cos v, (b+a\cos u)\sin v, a\sin u) \quad \text{See 4.8.} \\
E &= a^2 \\
F &= 0 \\
G &= (b+a\cos u)^2 \\
ds^2 &= a^2du^2 + (b+a\cos u)^2dv^2 \qquad (4.23)
\end{aligned}$$

e) Helicoid

$$\begin{aligned}
\mathbf{x} &= (u\cos v, u\sin v, av) \quad \text{Coordinate curves } u = c \text{ are helices.} \\
E &= 1, \\
F &= 0, \\
G &= u^2 + a^2, \\
ds^2 &= du^2 + (u^2+a^2)dv^2. \qquad (4.24)
\end{aligned}$$

f) Catenoid

$$\begin{aligned}
\mathbf{x} &= (u\cos v, u\sin v, c\cosh^{-1}\frac{u}{c}) \quad \text{This is a catenary of revolution.} \\
E &= \frac{u^2}{u^2-c^2}, \\
F &= 0, \\
G &= u^2, \\
ds^2 &= \frac{u^2}{u^2-c^2}du^2 + u^2dv^2, \qquad (4.25)
\end{aligned}$$

g) Cone and Conical Helix The equation $z^2 = \cot^2\alpha(x^2+y^2)$, represents a circular cone whose generator makes an angle α with the z-axis. In parametric form,

$$\begin{aligned}
\mathbf{x} &= (r\cos\phi, r\sin\phi, r\cot\alpha) \\
E &= \csc^2\alpha, \\
F &= 0, \\
G &= r^2, \\
ds^2 &= \csc^2\alpha\; dr^2 + r^2d\phi^2. \qquad (4.26)
\end{aligned}$$

A conical helix a curve $\gamma(t) = \mathbf{x}(r(t), \phi(t))$, that makes a constant angle σ with the generators of the cone. Similar to the case of loxodromes, we have

$$\begin{aligned}
\gamma' &= \mathbf{x}_r\frac{dr}{dt} + \mathbf{x}_\phi\frac{d\phi}{dt}. \\
\cos\sigma &= \frac{<\mathbf{x}_r, \gamma'>}{\|\mathbf{x}_r\|\cdot\|\gamma'\|} = \frac{E\frac{dr}{dt}}{\sqrt{E}\frac{ds}{dt}} = \sqrt{E}\frac{dr}{ds}. \\
E\; dr^2 &= \cos^2\sigma\; ds^2, \\
\csc^2\alpha\; dr^2 &= \cos^2\sigma(\csc^2\alpha\; dr^2 + r^2d\phi^2), \\
\csc^2\alpha\sin^2\sigma\; dr^2 &= r^2\cos^2\sigma\; d\phi^2, \\
\frac{1}{r}\,dr &= \cot\sigma\sin\alpha\; d\phi.
\end{aligned}$$

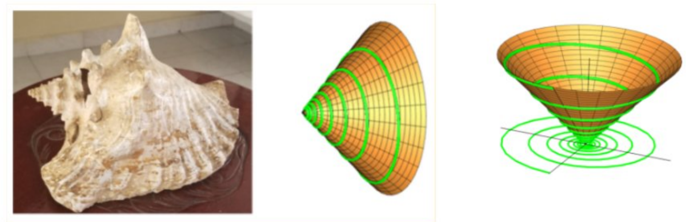


Fig. 4.9: Conical Helix.

Therefore, the equations of a conical helix are given by

$$r = ce^{\cot\sigma \sin\alpha\phi}. \tag{4.27}$$

As shown in figure 4.9, a conical helix projects into the plane as a logarithmic spiral. Many sea shells and other natural objects in nature exhibit neatly such conical spirals. The picture shown here is that of Lobatus gigas or Caracol Pala, previously known as Strombus gigas. This particular one is included here with certain degree of nostalgia for it has been a decorative item for decades in our family. The shell was probably found in Santa Cruz del Islote, Archipelago de San Bernardo, located in the Gulf of Morrosquillo in the Caribbean coast of Colombia. In this densely populated island paradise, which then enjoyed the pulchritude of enchanting coral reefs, the shells are now virtually extinct as the coral has succumbed to bleaching with rising temperatures of the waters. The shell shows a cut in the spire which the island natives use to sever the columellar muscle and thus release the edible snail.

4.3 The Second Fundamental Form

Let $\mathbf{x} = \mathbf{x}(u^\alpha)$ be a coordinate patch on a surface $\mathcal{M}$. Since $\mathbf{x}_u$ and $\mathbf{x}_v$ are tangential to the surface, we can construct a unit normal $\mathbf{n}$ to the surface by taking

$$\mathbf{n} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|} \tag{4.28}$$

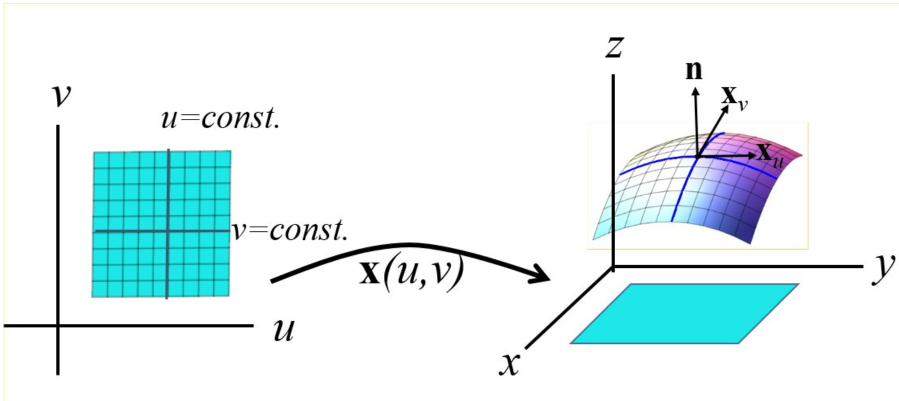


Fig. 4.10: Surface Normal

Now, consider a curve on the surface given by $u^\beta = u^\beta(s)$. Without loss of generality, we assume that the curve is parametrized by arc length s so that the curve has unit speed. Let $e = \{\mathbf{T}, \mathbf{N}, \mathbf{B}\}$ be the Frenet Frame of the curve. Recall that the rate of change $\overline{\nabla}_T W$ of a vector field W along the curve correspond to the classical vector $\mathbf{w}' = \frac{d\mathbf{w}}{ds}$, so $\overline{\nabla} W$ is associated with the vector $d\mathbf{w}$. Thus the connection equation $\overline{\nabla} e = e\omega$ is given by:

$$d[\mathbf{T}, \mathbf{N}, \mathbf{B}] = [\mathbf{T}, \mathbf{N}, \mathbf{B}] \begin{bmatrix} 0 & -\kappa\, ds & 0 \\ \kappa\, ds & 0 & -\tau\, ds \\ 0 & \tau\, ds & 0 \end{bmatrix} s \tag{4.29}$$

Following ideas first introduced by Darboux and subsequntly perfected by Cartan, we introduce a new orthonormal frame $f = \{\mathbf{T}, \mathbf{g}, \mathbf{n},\}$ adapted to the surface, where at each point, $\mathbf{T}$ is the common tangent to the surface and to the curve on the surface, $\mathbf{n}$ is the unit normal to the surface and $\mathbf{g} = \mathbf{n} \times \mathbf{T}$. Since the two orthonormal frames must be related by a rotation that leaves the $\mathbf{T}$ vector fixed, we have $f = eB$, where B is a matrix of the form

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}. \tag{4.30}$$

We wish to find $\overline{\nabla} f = f\,\overline{\omega}$. A short computation using the change of basis equations $\overline{\omega} = B^{-1}\omega B + B^{-1} dB$ (see equations 3.40 and 3.42) gives:

$$d[\mathbf{T}, \mathbf{g}, \mathbf{n}] = [\mathbf{T}, \mathbf{g}, \mathbf{n}] \begin{bmatrix} 0 & -\kappa\cos\theta\, ds & -\kappa\sin\theta\, ds \\ \kappa\cos\theta\, ds & 0 & -\tau\, ds + d\theta \\ \kappa\sin\theta\, ds & \tau\, ds - d\theta & 0 \end{bmatrix}, \tag{4.31}$$

$$= [\mathbf{T}, \mathbf{g}, \mathbf{n}] \begin{bmatrix} 0 & -\kappa_g\, ds & \kappa_n\, ds \\ \kappa_g\, ds & 0 & -\tau_g\, ds \\ \kappa_n\, ds & \tau_g\, ds & 0 \end{bmatrix}, \tag{4.32}$$

where:

$\kappa_n = \kappa \sin\theta$ is called the **Normal Curvature**,

$\kappa_g = \kappa \cos\theta$ is called the **Geodesic Curvature**; $\mathbf{K}_g = \kappa_g \mathbf{g}$ the geodesic curvature vector, and

$\tau_g = \tau - d\theta/ds$ is called the **Geodesic Torsion**.

We conclude that we can decompose $\mathbf{T}'$ and the curvature κ into their normal and surface tangent space components (see figure 4.11)

$$T' = \kappa_n \mathbf{n} + \kappa_g \mathbf{g}, \tag{4.33}$$
$$\kappa^2 = \kappa_n^2 + \kappa_g^2. \tag{4.34}$$

The normal curvature κ_n measures the curvature of $\mathbf{x}(u^\alpha(s))$ resulting from the constraint of the curve to lie on a surface. The geodesic curvature κ_g measures the "sideward" component of the curvature in the tangent plane to the surface. Thus, if one draws a straight line on a flat piece of paper and then smoothly bends the paper into a surface, the straight line would now acquire some curvature. Since the line was originally straight, there is no sideward component of curvature so $\kappa_g = 0$ in this case. This means that the entire contribution to the curvature comes from the normal component, reflecting the fact that the only reason there is curvature here is due to the bend in the surface itself. In this sense, a curve on a surface for which the geodesic curvature vanishes at all points reflects locally the shortest path between two points. These curves are therefore called **geodesics** of the surface. The property of minimizing the path between two points is a local property. For example, on a sphere one would expect the geodesics to be great circles. However travelling from NY to SF along one such great circle, there is a short path and a very long one that goes around the earth.

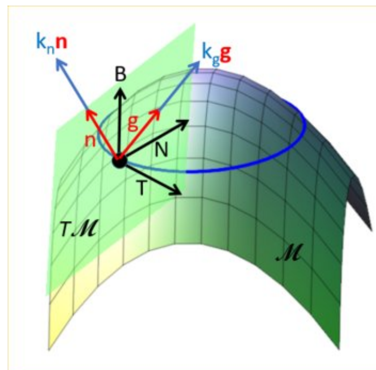


Fig. 4.11: Curvature

If one specifies a point $\mathbf{p} \in \mathcal{M}$ and a direction vector $X_p \in T_p\mathcal{M}$, one can geometrically envision the normal curvature by considering the equivalence class of all unit speed curves in $\mathcal{M}$ that contain the point $\mathbf{p}$ and whose tangent vectors line up with the direction of X. Of course, there are infinitely many such curves, but at an infinitesimal level, all these curves can be obtained by intersecting the surface with a "vertical" plane containing the vector X and the normal to $\mathcal{M}$. All curves in this equivalence class have the same normal curvature and their geodesic curvatures vanish. In this sense, the normal curvature is more of a property pertaining to a direction on the surface at a point, whereas the geodesic curvature really depends on the curve itself. It might be impossible for a hiker walking on the undulating hills of the Ozarks to find a straight line trail, since the rolling hills of the terrain extend in all directions. It might be possible, however, for the hiker to walk on a path with zero geodesic curvature as long the same compass direction is maintained. We will come back to the Cartan Structure equations associated with the Darboux frame, but for computational purposes, the classical approach is very practical.

Using the chain rule, we se that the unit tangent vector $\mathbf{T}$ to the curve is given by

$$\mathbf{T} = \frac{d\mathbf{x}}{ds} = \frac{d\mathbf{x}}{du^\alpha}\frac{du^\alpha}{ds} = \mathbf{x}_\alpha \frac{du^\alpha}{ds} \tag{4.35}$$

To find an explicit formula for the normal curvature we first differentiate equation (4.35)

$$\begin{aligned} \mathbf{T}' &= \frac{dT}{ds} \\ &= \frac{d}{ds}(\mathbf{x}_\alpha \frac{du^\alpha}{ds}) \\ &= \frac{d}{ds}(\mathbf{x}_\alpha)\frac{du^\alpha}{ds} + \mathbf{x}_\alpha \frac{d^2u^\alpha}{ds^2} \\ &= (\frac{d\mathbf{x}_\alpha}{du^\beta}\frac{du^\beta}{ds})\frac{du^\alpha}{ds} + \mathbf{x}_\alpha \frac{d^2u^\alpha}{ds^2} \\ &= \mathbf{x}_{\alpha\beta}\frac{du^\alpha}{ds}\frac{du^\beta}{ds} + \mathbf{x}_\alpha \frac{d^2u^\alpha}{ds^2}. \end{aligned}$$

Taking the inner product of the last equation with the normal and noticing that $< \mathbf{x}_\alpha, \mathbf{n} >= 0$, we get

$$\begin{aligned} \kappa_n &= < \mathbf{T}', \mathbf{n} >=< \mathbf{x}_{\alpha\beta}, \mathbf{n} > \frac{du^\alpha}{ds}\frac{du^\beta}{ds} \\ &= \frac{b_{\alpha\beta}du^\alpha du^\beta}{g_{\alpha\beta}du^\alpha du^\beta}, \end{aligned} \tag{4.36}$$

where

$$b_{\alpha\beta} =< \mathbf{x}_{\alpha\beta}, \mathbf{n} > \tag{4.37}$$

4.12 Definition The expression

$$II = b_{\alpha\beta}\, du^\alpha \otimes du^\beta \tag{4.38}$$

is called the **second fundamental form** .

4.13 Proposition The second fundamental form is symmetric.
Proof: In the classical formulation of the second fundamental form, the proof is trivial. We have $b_{\alpha\beta} = b_{\beta\alpha}$, since for a C^∞ patch $\mathbf{x}(u^\alpha)$, we have $\mathbf{x}_{\alpha\beta} = \mathbf{x}_{\beta\alpha}$, because the partial derivatives commute. We will denote the coefficients of the second fundamental form as follows:

$$\begin{aligned} e &= b_{11} =< \mathbf{x}_{uu}, \mathbf{n} > \\ f &= b_{12} =< \mathbf{x}_{uv}, \mathbf{n} > \\ &= b_{21} =< \mathbf{x}_{vu}, \mathbf{n} > \\ g &= b_{22} =< \mathbf{x}_{vv}, \mathbf{n} >, \end{aligned}$$

so that equation (4.38) can be written as

$$II = edu^2 + 2fdudv + gdv^2. \tag{4.39}$$

It follows that the equation for the normal curvature (4.36) can be written explicitly as

$$\kappa_n = \frac{II}{I} = \frac{edu^2 + 2fdudv + gdv^2}{Edu^2 + 2Fdudv + Gdv^2}. \tag{4.40}$$

It should pointed out that just as the first fundamental form can be represented as

$$I =< d\mathbf{x}, d\mathbf{x} >,$$

we can represent the second fundamental form as

$$II = - < d\mathbf{x}, d\mathbf{n} >$$

To see this, it suffices to note that differentiation of the identity, $< \mathbf{x}_\alpha, \mathbf{n} >= 0$, implies that

$$< \mathbf{x}_{\alpha\beta}, \mathbf{n} >= - < \mathbf{x}_\alpha, \mathbf{n}_\beta > .$$

Therefore,

$$\begin{aligned} < d\mathbf{x}, d\mathbf{n} > &= < \mathbf{x}_\alpha du^\alpha, \mathbf{n}_\beta du^\beta > \\ &= < \mathbf{x}_\alpha du^\alpha, \mathbf{n}_\beta du^\beta > \\ &= < \mathbf{x}_\alpha, \mathbf{n}_\beta > du^\alpha du^\beta \\ &= - < \mathbf{x}_{\alpha\beta}, \mathbf{n} > du^\alpha du^\beta \\ &= -II \end{aligned}$$

4.14 Definition Directions on a surface along which the second fundamental form

$$e\, du^2 + 2f\, du\, dv + g\, dv^2 = 0 \tag{4.41}$$

vanishes are called **asymptotic directions**, and curves having these directions are called **asymptotic curves**. This happens for example when there are straight lines on the surface as in the case of the intersection of the saddle $z = xy$ with the the plane $z = 0$.

We state without elaboration for now, that one can also define the third fundamental form by

$$III =< d\mathbf{n}, d\mathbf{n} >=< \mathbf{n}_\alpha, \mathbf{n}_\beta > du^\alpha du^\beta. \tag{4.42}$$

From a computational point a view, a more useful formula for the coefficients of the second fundamental formula can be derived by first applying the classical vector identity

$$(A \times B) \cdot (C \times D) = \begin{vmatrix} A \cdot C & A \cdot D \\ B \cdot C & B \cdot D \end{vmatrix} \tag{4.43}$$

to compute

$$\begin{aligned} \|\mathbf{x}_u \times \mathbf{x}_v\|^2 &= (\mathbf{x}_u \times \mathbf{x}_v) \cdot (\mathbf{x}_u \times \mathbf{x}_v) \\ &= det \begin{bmatrix} \mathbf{x}_u \cdot \mathbf{x}_u & \mathbf{x}_u \cdot \mathbf{x}_v \\ \mathbf{x}_v \cdot \mathbf{x}_u & \mathbf{x}_v \cdot \mathbf{x}_v \end{bmatrix} \\ &= EG - F^2. \end{aligned} \tag{4.44}$$

Consequently, the normal vector can be written as

$$\mathbf{n} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\sqrt{EG - F^2}}$$

Thus, we can write the coefficients $b_{\alpha\beta}$ directly as triple products involving derivatives of $(\mathbf{x})$. The expressions for these coefficients are

$$\begin{aligned} e &= \frac{(\mathbf{x}_u \mathbf{x}_v \mathbf{x}_{uu})}{\sqrt{EG - F^2}} \\ f &= \frac{(\mathbf{x}_u \mathbf{x}_v \mathbf{x}_{uv})}{\sqrt{EG - F^2}} \\ g &= \frac{(\mathbf{x}_u \mathbf{x}_v \mathbf{x}_{vv})}{\sqrt{EG - F^2}} \end{aligned} \tag{4.45}$$

4.15 Example Sphere
Going back to example 4.20, we have:

$$\begin{aligned}
\mathbf{x}_{\theta\theta} &= (a\sin\theta\cos\phi, -a\sin\theta\sin\phi, -a\cos\theta) \\
\mathbf{x}_{\theta\phi} &= (-a\cos\theta\sin\phi, a\cos\theta\cos\phi, 0) \\
\mathbf{x}_{\phi\phi} &= (-a\sin\theta\cos\phi, -a\sin\theta\sin\phi, 0) \\
\mathbf{n} &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \\
e &= \mathbf{x}_{\theta\theta}\cdot\mathbf{n} = -a \\
f &= \mathbf{x}_{\theta\phi}\cdot\mathbf{n} = 0 \\
g &= \mathbf{x}_{\phi\phi}\cdot\mathbf{n} = -a\sin^2\theta \\
II &= \frac{1}{a^2} I
\end{aligned}$$

The first fundamental form on a surface measures the square of the distance between two infinitesimally separated points. There is a similar interpretation of the second fundamental form as we show below. The second fundamental form measures the distance from a point on the surface to the tangent plane at a second infinitesimally separated point. To see this simple geometrical interpretation, consider a point $\mathbf{x}_0 = \mathbf{x}(u_0^\alpha) \in \mathcal{M}$ and a nearby point $\mathbf{x}(u_0^\alpha + du^\alpha)$. Expanding on a Taylor series, we get

$$\mathbf{x}(u_0^\alpha + du^\alpha) = \mathbf{x}_0 + (\mathbf{x}_0)_\alpha du^\alpha + \frac{1}{2}(\mathbf{x}_0)_{\alpha\beta} du^\alpha du^\beta + \ldots .$$

We recall that the distance formula from a point $\mathbf{x}$ to a plane which contains $\mathbf{x}_0$ is just the scalar projection of $(\mathbf{x} - \mathbf{x}_0)$ onto the normal. Since the normal to the plane at $\mathbf{x}_0$ is the same as the unit normal to the surface and $< \mathbf{x}_\alpha, \mathbf{n} >= 0$, we find that the distance D is

$$\begin{aligned}
D &= < \mathbf{x} - \mathbf{x}_0, \mathbf{n} > \\
&= \frac{1}{2} < (\mathbf{x}_0)_{\alpha\beta}, \mathbf{n} > du^\alpha du^\beta \\
&= \frac{1}{2} II_0
\end{aligned}$$

The first fundamental form (or, rather, its determinant) also appears in calculus in the context of calculating the area of a parametrized surface. If one considers an infinitesimal parallelogram subtended by the vectors $\mathbf{x}_u du$ and $\mathbf{x}_v dv$, then the differential of surface area is given by the length of the cross product of these two infinitesimal tangent vectors. That is,

$$\begin{aligned}
dS &= \|\mathbf{x}_u \times \mathbf{x}_v\| \, dudv \\
S &= \int\!\!\int \sqrt{EG - F^2} \, dudv
\end{aligned}$$

The second fundamental form contains information about the shape of the surface at a point. For example, the discussion above indicates that if $b = |b_{\alpha\beta}| = eg - f^2 > 0$ then all the neighboring points lie on the same side of the tangent plane, and hence, the surface is concave in one direction. If at a point on a surface $b > 0$, the point is called an elliptic point, if $b < 0$, the point is called hyperbolic or a saddle point, and if $b = 0$, the point is called parabolic.

4.4 Curvature

The concept of curvature and its relation to the fundamental forms, constitute the central object of study in differential geometry. One would like to be able to answer questions such as "what quantities remain invariant as one surface is smoothly changed into another?" There is certainly something intrinsically different between a cone, which we can construct from a flat piece of paper, and a sphere, which we cannot. What is it that makes these two surfaces so different? How does one calculate the shortest path between two objects when the path is constrained to lie on a surface?

These and questions of similar type can be quantitatively answered through the study of curvature. We cannot overstate the importance of this subject; perhaps it suffices to say that, without a clear understanding of curvature, there would be no general theory of relativity, no concept of black holes, and even more disastrous, no Star Trek.

Studying the curvature of a hypersurface in $\mathbf{R}^n$ (a surface of dimension $n-1$) begins by trying to understand the covariant derivative of the normal to the surface. If the normal to a surface is constant, then the surface is a flat hyperplane. Variations in the normal are what indicates the presence of curvature. For simplicity, we constrain our discussion to surfaces in $\mathbf{R}^3$, but the formalism we use is applicable to any dimension. We will also introduce in the modern version of the second fundamental form.

4.4.1 Classical Formulation of Curvature

The normal curvature κ_n at any point on a surface measures the deviation from flatness as one moves along a direction tangential to the surface at that point. The direction can be taken as the unit tangent vector to a curve on the surface. We seek the directions in which the normal curvature attains the extrema. For this purpose, let the curve on the surface be given by $v = v(u)$ and let $\lambda = \frac{dv}{du.}$ Then we can write the normal curvature 4.40 in the form

$$\kappa_n = \frac{II^*}{I^*} = \frac{e + 2f\lambda + g\lambda^2}{E + 2F\lambda + G\lambda^2}, \tag{4.46}$$

where II' and I' are the numerator and denominator respectively. To find the extrema, we take the derivative with respect to λ and set it equal to zero. The resulting fraction is zero only when the numerator is zero, so from the quotient rule we get

$$I^*(2f + 2g\lambda) - II^*(2F + 2G\lambda) = 0.$$

It follows that,

$$\kappa_n = \frac{II^*}{I^*} = \frac{f + g\lambda}{F + G\lambda}. \tag{4.47}$$

On the other hand, combining with equation 4.46 we have,

$$\kappa_n = \frac{(e + f\lambda) + \lambda(f + g\lambda)}{(E + F\lambda) + \lambda(F + G\lambda)} = \frac{f + g\lambda}{F + G\lambda}.$$

This can only happen if

$$\kappa_n = \frac{f + g\lambda}{F + G\lambda} = \frac{e + f\lambda}{E + F\lambda}. \tag{4.48}$$

Equation 4.48 contains a wealth of information. On one hand, we can eliminate κ_n which leads to the quadratic equation for λ

$$(Fg - gF)\lambda^2 + (Eg - Ge)\lambda + (Ef - Fe) = 0.$$

Recalling that $\lambda = dv/du$, and noticing that the coefficients resemble minors of a 3×3 matrix, we can elegantly rewrite the equation as

$$\begin{Vmatrix} du^2 & -du\,dv & dv^2 \\ E & F & G \\ e & f & g \end{Vmatrix} = 0. \tag{4.49}$$

Equation 4.49 determines two directions $\frac{du}{dv}$ along which the normal curvature attains the extrema, except for special cases when either $b_{\alpha\beta} = 0$, or $b_{\alpha\beta}$ and $g_{\alpha\beta}$ are proportional, which would cause the determinant to be identically zero. These two directions are called **principal directions** of curvature, each associated with an extremum of the normal curvature. We will have much more to say about these shortly.

On the other hand, we can write equations 4.48 in the form

$$\begin{cases} (e - E\kappa_n) + \lambda(f - F\kappa_n) = 0 \\ (f - F\kappa_n) + \lambda(g - G\kappa_n) = 0 \end{cases}$$

Solving each equation for λ we can eliminate λ instead, and we are lead to a quadratic equation for κ_n which we can write as

$$\begin{Vmatrix} e - E\kappa_n & f - F\kappa_n \\ f - F\kappa_n & g - G\kappa_n \end{Vmatrix} = 0. \tag{4.50}$$

It is interesting to note that equation 4.50 can be written as

$$\left\| \begin{bmatrix} e & f \\ f & g \end{bmatrix} - \kappa_n \begin{bmatrix} E & F \\ F & G \end{bmatrix} \right\| = 0.$$

In other words, the extrema for the values of the normal are the solutions of the equation

$$\| b_{\alpha\beta} - \kappa_n g_{\alpha\beta} \| = 0. \tag{4.51}$$

Had it been the case that $g_{\alpha\beta} = \delta_{\alpha\beta}$, the reader would recognize this as a eigenvalue equation for a symmetric matrix giving rise to two invariants, that is, the trace and the determinant of the matrix. We will treat this formally in the next section. The explicit quadratic expression for the extrema of κ_n

$$(EG - F^2)\kappa_n^2 - (Eg - 2Ff + Ge)\kappa_n + (eg - f^2) = 0.$$

We conclude there are two solutions κ_1 and κ_2 such that

$$\kappa_1 \kappa_2 = \frac{eg - f^2}{EG - F^2}, \tag{4.52}$$

and

$$\tfrac{1}{2}(\kappa_1 + \kappa_2) = \frac{1}{2} \frac{Eg - 2Ff + Ge}{EF - G^2}.. \tag{4.53}$$

To understand better the deep significance of the last two equations, we introduce the modern formulation which will allow is to draw conclusions from the inextricable connection of these results with the center piece of the spectral theorem for symmetric operators in linear algebra.

4.4.2 Covariant Derivative Formulation of Curvature

4.16 Definition Let X be a vector field on a surface M in $\mathbf{R}^3$, and let N be the normal vector. The map L, given by

$$LX = -\overline{\nabla}_X N \tag{4.54}$$

is called the **Weingarten map**.

Here, we have adopted the convention to overline the operator $\overline{\nabla}$ when it refers to the ambient space. The Weingarten map is natural to consider, since it represents the rate of change of the normal in an arbitrary direction tangential to the surface, which is what we wish to quantify.

4.17 Definition The **Lie bracket** $[X, Y]$ of two vector fields X and Y on a surface $\mathcal{M}$ is defined as the commutator,

$$[X, Y] = XY - YX, \tag{4.55}$$

meaning that if f is a function on $\mathcal{M}$, then $[X, Y](f) = X(Y(f)) - Y(X(f))$.

4.18 Proposition The Lie bracket of two vectors $X, Y \in T(\mathcal{M})$ is another vector in $T(\mathcal{M})$.
Proof: If suffices to prove that the bracket is a linear derivation on the space of C^∞ functions. Consider vectors $X, Y \in T(\mathcal{M})$ and smooth functions f, g in $\mathcal{M}$. Then,

$$\begin{aligned}
[X, Y](f+g) &= X(Y(f+g)) - Y(X(f+g)) \\
&= X(Y(f) + Y(g)) - Y(X(f) + X(g)) \\
&= X(Y(f)) - Y(X(f)) + X(Y(g)) - Y(X(g)) \\
&= [X, Y](f) + [X, Y](g),
\end{aligned}$$

and

$$\begin{aligned}
[X, Y](fg) &= X(Y(fg)) - Y(X(fg)) \\
&= X[fY(g) + gY(f)] - Y[fX(g) + gX(f)] \\
&= X(f)Y(g) + fX(Y(g)) + X(g)Y(f) + gX(Y(f)) \\
&\quad -Y(f)X(g) - f(Y(X(g)) - Y(g)X(f) - gY(X(f)) \\
&= f[X(Y(g)) - (Y(X(g))] + g[X(Y(f)) - Y(X(f))] \\
&= f[X, Y](g) + g[X, Y](f).
\end{aligned}$$

4.19 Proposition The Weingarten map is a linear transformation on $T(\mathcal{M})$.
Proof: Linearity follows from the linearity of $\overline{\nabla}$, so it suffices to show that $L : X \longrightarrow LX$ maps $X \in T(\mathcal{M})$ to a vector $LX \in T(\mathcal{M})$. Since N is the unit normal to the surface, $< N, N >= 1$, so any derivative of $< N, N >$ is 0. Assuming that the connection is compatible with the metric,

$$\begin{aligned}
\overline{\nabla}_X < N, N > &= < \overline{\nabla}_X N, > + < N, \overline{\nabla}_X N > \\
&= 2 < \overline{\nabla}_X N, N > \\
&= 2 < -LX, N >= 0.
\end{aligned}$$

Therefore, LX is orthogonal to N; hence, it lies in $T(\mathcal{M})$.

In the previous section, we gave two equivalent definitions $< d\mathbf{x}, d\mathbf{x} >$, and $< X, Y >$) of the first fundamental form. We will now do the same for the second fundamental form.

4.20 Definition The **second fundamental form** is the bilinear map

$$II(X,Y) = < LX, Y > . \tag{4.56}$$

4.21 Remark It should be noted that the two definitions of the second fundamental form are consistent. This is easy to see if one chooses X to have components $\mathbf{x}_\alpha$ and Y to have components $\mathbf{x}_\beta$. With these choices, LX has components $-\mathbf{n}_a$ and $II(X,Y)$ becomes $b_{\alpha\beta} = - < \mathbf{x}_\alpha, \mathbf{n}_\beta >$.

We also note that there is a third fundamental form defined by

$$III(X,Y) = < LX, LY > . \tag{4.57}$$

In classical notation, the third fundamental form would be denoted by $< d\mathbf{n}, d\mathbf{n} >$. As one would expect, the third fundamental form contains third order Taylor series information about the surface.

4.22 Definition The **torsion** of a connection $\overline{\nabla}$ is the operator T such that $\forall X, Y$,

$$T(X,Y) = \overline{\nabla}_X Y - \overline{\nabla}_Y X - [X,Y]. \tag{4.58}$$

A connection is called **torsion-free** if $T(X,Y) = 0$. In this case,

$$\overline{\nabla}_X Y - \overline{\nabla}_Y X = [X,Y].$$

We will say more later about the torsion and the importance of torsion-free connections. For the time being, it suffices to assume that for the rest of this section, all connections are torsion-free. Using this assumption, it is possible to prove the following important theorem.

4.23 Theorem The Weingarten map is a self-adjoint operator on $T\mathcal{M}$.
Proof: We have already shown that $L : T\mathcal{M} \longrightarrow T\mathcal{M}$ is a linear map. Recall that an operator L on a linear space is self adjoint if $< LX, Y > = < X, LY >$, so the theorem is equivalent to proving that the second fundamental form is symmetric ($II[X,Y] = II[Y,X]$). Computing the difference of these two quantities, we get

$$\begin{aligned} II[X,Y] - II[Y,X] &= < LX, Y > - < LY, X > \\ &= < -\overline{\nabla}_X N, Y > - < -\overline{\nabla}_Y N, X > . \end{aligned}$$

Since $< X, N > = < Y, N > = 0$ and the connection is compatible with the metric, we know that

$$\begin{aligned} < -\overline{\nabla}_X N, Y > &= < N, \overline{\nabla}_X Y > \\ < -\overline{\nabla}_Y N, X > &= < N, \overline{\nabla}_Y X >, \end{aligned}$$

hence,

$$\begin{aligned} II[X,Y] - II[Y,X] &= < N, \overline{\nabla}_Y X > - < N, \overline{\nabla}_X Y >, \\ &= < N, \overline{\nabla}_Y X - \overline{\nabla}_X Y > \\ &= < N, [X,Y] > \\ &= 0 \qquad \text{(iff } [X,Y] \in T(\mathcal{M})). \end{aligned}$$

The central theorem of linear algebra is the Spectral Theorem. In the case of self-adjoint operators, the Spectral Theorem states that given the eigenvalue equation for a symmetric operator on a vector space with an inner product,

$$LX = \kappa X, \tag{4.59}$$

the eigenvalues are always real and eigenvectors corresponding to different eigenvalues are orthogonal. Here, the vector spaces in question are the tangent spaces at each point of a surface in $\mathbf{R}^3$, so the dimension is 2. Hence, we expect two eigenvalues and two eigenvectors:

$$\begin{aligned} LX_1 &= \kappa_1 X_1 \qquad (4.60)\\ LX_2 &= \kappa_1 X_2. \qquad (4.61)\end{aligned}$$

4.24 Definition The eigenvalues κ_1 and κ_2 of the Weingarten map L are called the **principal curvatures** and the eigenvectors X_1 and X_2 are called the **principal directions**.

Several possible situations may occur, depending on the classification of the eigenvalues at each point $\mathbf{p}$ on a given surface:

1. If $\kappa_1 \neq \kappa_2$ and both eigenvalues are positive, then $\mathbf{p}$ is called an **elliptic point**.
2. If $\kappa_1\kappa_2 < 0$, then $\mathbf{p}$ is called a **hyperbolic point**.
3. If $\kappa_1 = \kappa_2 \neq 0$, then $\mathbf{p}$ is called an **umbilic point**.
4. If $\kappa_1\kappa_2 = 0$, then $\mathbf{p}$ is called a **parabolic point**.

It is also known from linear algebra, that the the determinant and the trace of a self-adjoint operator in a vector space of dimension two are the only invariants under an adjoint (similarity) transformation. Clearly, these invariants are important in the case of the operator L, and they deserve special names.

4.25 Definition The determinant $K = \det(L)$ is called the **Gaussian curvature** of $\mathcal{M}$ and $H = \frac{1}{2}\mathrm{Tr}(L)$ is called the **mean curvature** .

Since any self-adjoint operator is diagonalizable and in a diagonal basis the matrix representing L is $\mathrm{diag}(\kappa_1, \kappa_2)$, if follows immediately that

$$\begin{aligned} K &= \kappa_1\kappa_2\\ H &= \frac{1}{2}(\kappa_1 + \kappa_2). \qquad (4.62)\end{aligned}$$

4.26 Proposition Let X and Y be any linearly independent vectors in $T(\mathcal{M})$. Then

$$\begin{aligned} LX \times LY &= K(X \times Y)\\ (LX \times Y) + (X \times LY) &= 2H(X \times Y) \qquad (4.63)\end{aligned}$$

Proof: Since $LX, LY \in T(\mathcal{M})$, they can be expressed as linear combinations of the basis vectors X and Y.

$$\begin{aligned} LX &= a_1X + b_1Y\\ LY &= a_2X + b_2Y.\end{aligned}$$

computing the cross product, we get

$$\begin{aligned} LX \times LY &= \begin{vmatrix} a_1 & b_1\\ a_2 & b_2 \end{vmatrix} X \times Y\\ &= \det(L)(X \times Y).\end{aligned}$$

Similarly

$$\begin{aligned}(LX \times Y) + (X \times LY) &= (a_1 + b_2)(X \times Y)\\ &= \mathrm{Tr}(L)(X \times Y)\\ &= (2H)(X \times Y).\end{aligned}$$

4.27 Proposition

$$\begin{aligned}K &= \frac{eg - f^2}{EG - F^2}\\ H &= \frac{1}{2}\frac{Eg - 2Ff + eG}{EG - F^2}\textbf{??}\end{aligned} \tag{4.64}$$

Proof: Starting with equations (4.63), take the inner product of both sides with $X \times Y$ and use the vector identity (4.43). We immediately get

$$K = \frac{\begin{vmatrix} < LX, X > & < LX, Y > \\ < LY, X > & < LX, X > \end{vmatrix}}{\begin{vmatrix} < X, X > & < X, Y > \\ < Y, X > & < Y, Y > \end{vmatrix}} \tag{4.65}$$

$$2H = \frac{\begin{vmatrix} < LX, X > & < LX, Y > \\ < Y, X > & < Y, Y > \end{vmatrix} + \begin{vmatrix} < X, X > & < X, Y > \\ < LY, X > & < LY, Y > \end{vmatrix}}{\begin{vmatrix} < X, X > & < X, Y > \\ < Y, X > & < Y, Y > \end{vmatrix}} \tag{4.66}$$

The result follows by taking $X = \mathbf{x}_u$ and $Y = \mathbf{x}_v$. Not surprisingly, this is in complete agreement with the classical formulas for the Gausssian curvature (equation 4.52) and for the Mean curvature (equation 4.53. The result is also consistent with the well known fact from Linear Algebra, that the only similarity transformation invariants of a symmetry quadratic form are the determinant and the trace.

4.28 Example Sphere

From equations 4.20 and 4.3 we see that $K = 1/a^2$ and $H = 1/a$. This is totally intuitive since one would expect $\kappa_1 = \kappa_2 = 1/a$ since the normal curvature in any direction should equal the curvature of great circle. This means that a sphere is a surface of constant curvature and every point of a sphere is an umbilic point. This another way to think of the symmetry of the sphere in the sense that an observer at any point sees the same normal curvature in all directions.

The next theorem due to Euler gives a characterization of the normal curvature in the direction of an arbitrary unit vector X tangent to the surface surface M at a given point.

4.29 Theorem (Euler) Let X_1 and X_2 be unit eigenvectors of L and let $X = (\cos\theta)X_1 + (\sin\theta)X_2$. Then

$$II(X, X) = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta. \tag{4.67}$$

Proof: Simply compute $II(X, X) =< LX, X >$, using the fact the $LX_1 = \kappa_1 X_1$, $LX_2 = \kappa_2 X_2$, and noting that the eigenvectors are orthogonal. We get

$$\begin{aligned}< LX, X > &= < (\cos\theta)\kappa_1 X_1 + (\sin\theta)\kappa_2 X_2, (\cos\theta)X_1 + (\sin\theta)X_2 >\\ &= \kappa_1 \cos^2\theta < X_1, X_1 > + \kappa_2 \sin^2\theta < X_2, X_2 >\\ &= \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta.\end{aligned}$$

4.30 Theorem The first, second and third fundamental forms satisfy the equation

$$III - 2H\,II + KI = 0 \tag{4.68}$$

Proof: The proof follows immediately from the fact that for a symmetric 2 by 2 matrix A, the characteristic polynomial $\kappa^2 - \mathrm{tr}(A)\kappa + \det(A) = 0$, and from the Cayley-Hamilton theorem stating that the matrix is annihilated by its characteristic polynomial.

4.5 Fundamental Equations

4.5.1 Gauss-Weingarten Equations

As we have just seen for example, the Gaussian curvature of sphere of radius a is $1/a^2$. To compute this curvature we had to first compute the coefficients of the second fundamental form, and therefore, we first needed to compute the normal to the surface in $\mathbf{R}^3$. The computation therefore depended on the particular coordinate chart parametrizing the surface.

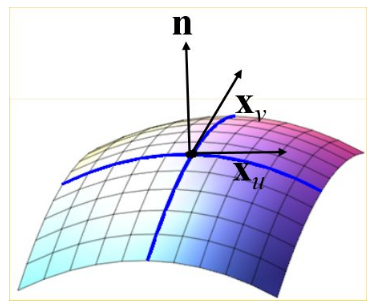


Fig. 4.12: Surface Frame

However, it would reasonable to conclude that the curvature of the sphere is an intrinsic quantity, independent of the embedding in $\mathbf{R}^3$. After all, a "two-dimensional" creature such as ant moving on the surface of the sphere would be constrained by the curvature of the sphere independent of the higher dimension on which the surface lives. This mode of thinking lead the brilliant mathematicians Gauss and Riemann to question if the coefficients of the second fundamental form were functionally computable from the coefficients of the first fundamental form. To explore this idea, consider again the basis vectors at each point of a surface consisting of two tangent vectors and the normal, as shown in figure 4.12. Given a coordinate chart $\mathbf{x}(u^\alpha)$, the vectors $\mathbf{x}_\alpha$ live on the tangent space, but this is not necessarily true for the second derivative vectors $\mathbf{x}_{\alpha\beta}$. However, since the triplet of vectors $\{\mathbf{x}_\alpha, \mathbf{n}\}$ constitutes a basis for $\mathbf{R}^3$ at each point on the surface, we can express the vectors $\mathbf{x}_{\alpha\beta}$ as linear combinations of the bases vectors. Therefore, there exist coefficients $\Gamma^\gamma_{\alpha\beta}$ and $c_{\alpha\beta}$ such that

$$\mathbf{x}_{\alpha\beta} = \Gamma^\gamma_{\alpha\beta}\mathbf{x}_\gamma + c_{\alpha\beta}\mathbf{n}. \tag{4.69}$$

Taking the inner product of equation 4.69 with $\mathbf{n}$, noticing that the latter is a unit vector orthogonal to $\mathbf{x}_\gamma$, we find that $c_{\alpha\beta} =< \mathbf{x}_{\alpha\beta}, \mathbf{n} >$, and hence these are just the coefficients of the second fundamental form. In other words, equation 4.69 can be written as

$$\mathbf{x}_{\alpha\beta} = \Gamma^\gamma_{\alpha\beta}\mathbf{x}_\gamma + b_{\alpha\beta}\mathbf{n}. \tag{4.70}$$

Equation 4.70 together with equation 4.73 below, are called the formulae of **Gauss**. The covariant derivative formulation of the equation of Gauss follows in a similar fashion. Let X and Y be vector fields tangent to the surface. We can the decompose the covariant derivative of Y in the direction of X into its tangential and normal components

$$\overline{\nabla}_X Y = \nabla_X Y + h(X,Y)N.$$

But then,

$$
\begin{aligned}
h(X,Y) &= \langle \overline{\nabla}_X Y, N\rangle,\\
&= -\langle Y, \overline{\nabla}_X N\rangle,\\
&= -\langle Y, LX, \rangle,\\
&= -\langle LX, Y\rangle,\\
&= II(X,Y).
\end{aligned}
$$

Thus, the coordinate independent formulation of the equation of Gauss reads:

$$
\overline{\nabla}_X Y = \nabla_X Y + II(X,Y)N. \tag{4.71}
$$

The quantity $\nabla_X Y$ represents a covariant derivative on the surface, so in that sense, it is intrinsic to the surface. If $\alpha(s)$ is a curve on the surface with tangent $T = \alpha'(s)$, we say that a vector field Y is **parallel-transported** along the curve if $\nabla_T Y = 0$. This notion of parallelism refers to parallelism on the surface, not the ambient space. To illustrate by example, Figure 4.13 shows a vector field Y tangent to a sphere along the circle with azimuthal angle $\theta = \pi/3$. The circle has unit tangent $T = \alpha'(s)$, and at each point on the circle, the vector Y points North. To the inhabitants of the sphere, the vector Y appears parallel-transported on the surface along the curve, that is $\nabla_T Y = 0$. However, Y is clearly not parallel-transported in the ambient $\mathbf{R}^3$ space with respect to the connection $\overline{\nabla}$.

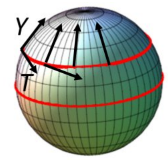


Fig. 4.13:

The torsion T of the connection ∇ is defined exactly as before (See equation 4.58).

$$
T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y].
$$

Also, as in definition 3.15, the connection is compatible with the metric on the surface if

$$
\nabla_X \langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle.
$$

A torsion-free connection that is compatible with the metric is called a **Levi-Civita** connection.

4.31 Proposition A Levi-Civita connection preserves length and angles under parallel transport.
Proof: Let $T = \alpha'(t)$ be tangent to curve $\alpha(T)$, and X and Y be parallel-transported along α. By definition, $\nabla_T X = \nabla_T Y = 0$. Then

$$
\begin{aligned}
\nabla_T\langle X, X\rangle &= \langle \nabla_T X, X\rangle + \langle X, \nabla_T X\rangle,\\
&= 2\langle \nabla_T X, X\rangle = 0 \Rightarrow \|X\| = \text{constant}.\\
\nabla_T\langle X, Y\rangle &= \langle \nabla_T X, Y\rangle + \langle X, \nabla_T Y\rangle = 0, \Rightarrow \langle X, Y\rangle = \text{constant, so}\\
\cos\theta &= \frac{\langle X, Y\rangle}{\|X\|\cdot\|Y\|} = \text{constant}.
\end{aligned}
$$

If one takes $\{e_\alpha\}$ to be a basis of tangent space, the components of the connection in that basis are given by the familiar equation

$$
\nabla_{e_\alpha e_\beta} = \Gamma^\gamma{}_{\alpha\beta} e_\gamma.
$$

The Γ's here are of course the same Christoffel symbols in the equation of Gauss 4.70. We have the following important result:

4.32 Theorem In a surface $\{\mathcal{M}, g\}$ with metric g (the first fundamental form), there exists a unique Levi-Civita connection.

The proof is implicit in the computations that follow leading to equation 4.73, which express the components uniquely in terms of the metric. The entire equation (4.70) must be symmetric on the indices $\alpha\beta$, since $\mathbf{x}_{\alpha\beta} = \mathbf{x}_{\beta\alpha}$, so $\Gamma^{\gamma}_{\alpha\beta} = \Gamma^{\gamma}_{\alpha\beta}$ is also symmetric on the lower indices. These quantities are called the **Christoffel symbols of the first kind**. Now we take the inner product with $\mathbf{x}_\sigma$ to deduce that

$$\begin{aligned} < \mathbf{x}_{\alpha\beta}, \mathbf{x}_\sigma > &= \Gamma^{\gamma}_{\alpha\beta} < \mathbf{x}_\gamma, \mathbf{x}_\sigma >, \\ &= \Gamma^{\gamma}_{\alpha\beta} g_{\gamma\sigma}, \\ &= \Gamma_{\alpha\beta\sigma}; \end{aligned}$$

where we have lowered the third index with the metric on the right hand side of the last equation. The quantities $\Gamma_{\alpha\beta\sigma}$ are called **Christoffel symbols of the second kind**. Here we must note that not all indices are created equal. The Christoffel symbols of the second kind are only symmetric on the first two indices. The notation $\Gamma_{\alpha\beta\sigma} = [\alpha\beta, \sigma]$ is also used in the literature.

The Christoffel symbols can be expressed in terms of the metric by first noticing that the derivative of the first fundamental form is given by

$$\begin{aligned} g_{\alpha\gamma,\beta} &= \frac{\partial}{\partial u^\beta} < \mathbf{x}_\alpha, \mathbf{x}_\gamma >, \\ &=< \mathbf{x}_{\alpha\beta}, \mathbf{x}_\gamma > + < \mathbf{x}_\alpha, \mathbf{x}_{\gamma\beta}, > \\ &= \Gamma_{\alpha\beta\gamma} + \Gamma_{\gamma\beta\alpha} \end{aligned}$$

Taking other cyclic permutations of this equation, we get

$$\begin{aligned} g_{\alpha\gamma,\beta} &= \Gamma_{\alpha\beta\gamma} + \Gamma_{\gamma\beta\alpha}, \\ g_{\beta\gamma,\alpha} &= \Gamma_{\alpha\beta\gamma} + \Gamma_{\gamma\alpha\beta}, \\ g_{\alpha\beta,\gamma} &= \Gamma_{\alpha\gamma\beta} + \Gamma_{\gamma\beta\alpha}. \end{aligned}$$

Adding the first two and subtracting the third of the equations above, and recalling that the Γ's are symmetric on the first two indices, we obtain the formula

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}(g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha} - g_{\alpha\beta,\gamma}.) \tag{4.72}$$

Raising the third index with the inverse of the metric, we also have the following formula for the Christoffel symbols of the first kind (hereafter, Christoffel symbols refer to the symbols of the first kind, unless otherwise specified.)

$$\Gamma^{\sigma}_{\alpha\beta} = \frac{1}{2} g^{\sigma\gamma}(g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha} - g_{\alpha\beta,\gamma}.) \tag{4.73}$$

The Christoffel symbols are clearly symmetric in the lower indices

$$\Gamma^{\sigma}_{\alpha\beta} = \Gamma^{\sigma}_{\beta\alpha}. \tag{4.74}$$

Unless otherwise specified, a connection on $\{\mathcal{M}, g\}$ refers to the unique Levi-Civita connection.

4.33 Example

As an example we unpack the formula for Γ^1_{11}. Let $g = \|g_{\alpha\beta}\| = EG - F^2$. From equation 4.73 we have

$$\begin{aligned}
\Gamma^1_{11} &= \frac{1}{2} g^{1\gamma}(g_{1\gamma,1} + g_{1\gamma,1} - g_{11,\gamma}),\\
&= \frac{1}{2} g^{1\gamma}(2g_{1\gamma,1} - g_{11,\gamma}),\\
&= \frac{1}{2}[g^{11}(2g_{11,1} - g_{11,1}) + g^{12}(2g_{12,1} - g_{11,2})],\\
&= \frac{1}{2g}[GE_u - F(2F_u - FE_v)],\\
&= \frac{GE_u - 2FF_u + FE_v}{2(EG - F^2)}.
\end{aligned}$$

Due to symmetry, there are five other similar equations for the other Γ's. Proceeding as above, we can derive the entire set.

$$\begin{aligned}
\Gamma^1_{11} &= \frac{GE_u - 2FF_u + FE_v}{2(EG - F^2)} & \Gamma^2_{11} &= \frac{2EF_u - EE_v - FE_u}{2(EG - F^2)}\\
\Gamma^1_{12} &= \frac{GE_v - FG_u}{2(EG - F^2)} & \Gamma^2_{12} &= \frac{EG_u - FE_v}{2(EG - F^2)}\\
\Gamma^1_{22} &= \frac{2GF_v - GG_u - FG_v}{2(EG - F^2)} & \Gamma^2_{22} &= \frac{EG_v - 2FF_v + FG_u}{2(EG - F^2)}.
\end{aligned} \tag{4.75}$$

They are a bit messy, but they all simplify considerably for orthogonal systems, in which case $F = 0$. Another reason why we like those coordinate systems.

We can also write $\mathbf{n}_\alpha$ in terms of the frame vectors. This is by far easier since $< \mathbf{n}, \mathbf{n} >= 1$ implies that $< \mathbf{n}_\alpha, \mathbf{n} >= 0$, so $\mathbf{n}_\alpha$ lies on the tangent plane and it is therefore a linear combination the tangent vectors. As before, we easily verify that the coefficients are the second fundamental form with a raised index

$$\mathbf{n}_\alpha = -b^\gamma_\alpha \mathbf{x}_\gamma. \tag{4.76}$$

These are called the formulae of **Weingarten**.

4.5.2 Curvature Tensor, Gauss's Theorema Egregium

A fascinating set of relations can be obtained simply by equating $\mathbf{x}_{\beta\gamma\delta} = \mathbf{x}_{\beta\delta\gamma}$. Differentiating the equation of Gauss and recursively using the formualae of Gauss and Weingarten to write all the components in terms of the frame, we get

$$\begin{aligned}
\mathbf{x}_{\beta\delta} &= \Gamma^\alpha_{\beta\delta}\mathbf{x}_\alpha + b_{\beta\delta}\mathbf{n}\\
\mathbf{x}_{\beta\delta\gamma} &= \Gamma^\alpha_{\beta\delta,\gamma}\mathbf{x}_\alpha + \Gamma^\alpha_{\beta\delta}\mathbf{x}_{\alpha\gamma} + b_{\beta\delta,\gamma}\mathbf{n} + b_{\beta\delta}\mathbf{n}_\gamma\\
&= \Gamma^\alpha_{\beta\delta,\gamma}\mathbf{x}_\alpha + \Gamma^\alpha_{\beta\delta}[\Gamma^\mu_{\alpha\gamma}\mathbf{x}_\mu + b_{\alpha\gamma}\mathbf{n}] + b_{\beta\delta,\gamma}\mathbf{n} - b_{\beta\delta}b^\alpha_\gamma\mathbf{x}_\alpha\\
\mathbf{x}_{\beta\delta\gamma} &= [\Gamma^\alpha_{\beta\delta,\gamma} + \Gamma^\mu_{\beta\delta}\Gamma^\alpha_{\mu\gamma} - b_{\beta\delta}b^\alpha_\gamma]\mathbf{x}_\alpha + [\Gamma^\alpha_{\beta\delta}b_{\alpha\gamma} + b_{\beta\delta,\gamma}]\mathbf{n}, && (4.77)\\
\mathbf{x}_{\beta\gamma\delta} &= [\Gamma^\alpha_{\beta\gamma,\delta} + \Gamma^\mu_{\beta\gamma}\Gamma^\alpha_{\mu\delta} - b_{\beta\gamma}b^\alpha_\delta]\mathbf{x}_\alpha + [\Gamma^\alpha_{\beta\gamma}b_{\alpha\delta} + b_{\beta\gamma,\delta}]\mathbf{n}. && (4.78)
\end{aligned}$$

The last equation above was obtained from the preceding one just by permuting δ and γ. Subtracting that last two equations and setting the tangential component to zero we get

$$R^{\alpha}{}_{\beta\gamma\delta} = b_{\beta\delta} b^{\alpha}_{\gamma} - b_{\beta\gamma} b^{\alpha}_{\delta}, \tag{4.79}$$

where the components of the **Riemann tensor** R are defined by

$$R^{\alpha}_{\beta\gamma\delta} = \Gamma^{\alpha}_{\beta\delta,\gamma} - \Gamma^{\alpha}_{\beta\gamma,\delta} + \Gamma^{\mu}_{\beta\delta}\Gamma^{\alpha}_{\gamma\mu} - \Gamma^{\mu}_{\beta\gamma}\Gamma^{\alpha}_{\delta\mu}. \tag{4.80}$$

Technically we are not justified at this point in calling R a tensor since we have not established yet the appropriate multi-linear features that a tensor must exhibit. We address this point at a later chapter. Lowering the index above we get

$$R_{\alpha\beta\gamma\delta} = b_{\beta\delta} b_{\alpha\gamma} - b_{\beta\gamma} b_{\alpha\delta}. \tag{4.81}$$

4.34 Theorema Egregium

$$K = \frac{R_{1212}}{g}. \tag{4.82}$$

Proof: Let $\alpha = \gamma = 1$ and $\beta = \delta = 2$ above. The equation then reads

$$\begin{aligned} R_{1212} &= b_{22} b_{11} - b_{21} b_{12}, \\ &= (eg - f^2), \\ &= K(EF - G^2), \\ &= Kg \end{aligned}$$

The remarkable result is that the Riemann tensor and hence the Gaussian curvature does not depend on the second fundamental form but only on the coefficients of the metric. Thus, the Gaussian curvature is an intrinsic quantity independent of the embedding, so that two surfaces that have the same first fundamental form have the same curvature. In this sense, the Gaussian curvature is a bending invariant!

Setting the normal components equal to zero gives

$$\Gamma^{\alpha}_{\beta\delta} b_{\alpha\gamma} - \Gamma^{\alpha}_{\beta\gamma} b_{\alpha\delta} + b_{\beta\delta,\gamma} - b_{\beta\gamma,\delta} = 0 \tag{4.83}$$

These are called the **Codazzi** (or Codazzi-Mainardi) equations.

Computing the Riemann tensor is labor intensive since one must first obtain all the non=zero Christoffel symbols as shown in the example above. Considerable gain in efficiency results from a form computation. For this purpose, let $\{e_1, e_2, e_3\}$ be a Darboux frame adapted to the surface M, with $e_3 = \mathbf{n}$. Let $\{\theta^1, \theta^2, \theta^3\}$ be the corresponding orthonormal dual basis. Since at every point, a tangent vector $X \in TM$ is a linear combination of $\{e_1, e_2\}$, we see that $\theta^3(X) = 0$ for all such vectors. That is, $\theta^3 = 0$ on the surface. As a consequence, the entire set of the structure equations is

$$\begin{aligned} d\theta^1 &= -\omega^1{}_2 \wedge \theta^2, && \qquad (4.84) \\ d\theta^2 &= -\omega^2{}_1 \wedge \theta^1, && \qquad (4.85) \\ d\theta^3 &= -\omega^3{}_1 \wedge \theta^1 - \omega^3{}_2 \wedge \theta^2 = 0, && \qquad (4.86) \\ d\omega^1{}_2 &= -\omega^1{}_3 \wedge \omega^3{}_2, \qquad \text{Gauss Equation} && \qquad (4.87) \\ d\omega^1{}_3 &= -\omega^1{}_2 \wedge \omega^2{}_3, \qquad \text{Codazzi Equations} && \qquad (4.88) \\ d\omega^2{}_3 &= -\omega^2{}_1 \wedge \omega^1{}_3. && \qquad (4.89) \end{aligned}$$

The key result is the following theorem

4.35 Curvature Form Equations

$$d\omega^1{}_2 = K\,\theta^1 \wedge \theta^2, \tag{4.90}$$
$$\omega^1{}_3 \wedge \theta^2 + \omega^2{}_3 \wedge \theta^1 = -2H\;\theta^1 \wedge \theta^2. \tag{4.91}$$

Proof: By applying the Weingarten map to the basis vector $\{e_1, e_2\}$ of TM, we find a matrix representation of the linear transformation:

$$\begin{aligned} Le_1 &= -\nabla_{e_1} e_3 = -\omega^1{}_3(e_1)e_1 - \omega^2{}_3(e_1)e_2, \\ Le_2 &= -\nabla_{e_2} e_3 = -\omega^1{}_3(e_2)e_1 - \omega^2{}_3(e_2)3_2. \end{aligned}$$

Recalling that ω is antisymmetric, we find:

$$\begin{aligned} K = \det(L) &= -[\omega^1{}_3(e_1)\omega^3{}_2(e_1) - \omega^1{}_3(e_2)\omega^3{}_2(e_1)] \\ &= -(\omega^1{}_3 \wedge \omega^3{}_2)(e_1, e_2) \\ &= d\omega^1{}_2(e_1, e_2). \end{aligned}$$

Hence

$$d\omega^1{}_2 = K\,\theta^1 \wedge \theta^2.$$

Similarly, recalling that $\theta^1(e_j) = \delta^i_j$, we have

$$\begin{aligned} (\omega^1{}_3 \wedge \theta^2 + \omega^3{}_2 \wedge \theta^1))(e_1, e_2) &= \omega^1{}_3(e_1) - \omega^3{}_2(e_2) \\ &= \omega^1{}_3(e_1) + \omega^2{}_3(e_2) \\ &= \mathrm{Tr}(L) = -2H \end{aligned}$$

4.36 Definition A point of a surface at which $K = 0$ is called a **planar point**. A surface with $K = 0$ at all points is called a **flat** or **Gaussian flat** surface. A surface on which $H = 0$ at all points is called a **minimal** surface.

4.37 Example Sphere
Since the first fundamental form is $I = a^2\,d\theta^2 + a^2 \sin^2\theta\,d\phi^2$, we have

$$\begin{aligned} \theta^1 &= a\,d\theta \\ \theta^2 &= a \sin\theta\,d\phi \\ d\theta^2 &= a\cos\theta\,d\theta \wedge d\phi \\ &= -\cos\theta\,d\phi \wedge \theta^1 = -\omega^2{}_1 \wedge \theta^1 \\ \omega^2{}_1 &= \cos\theta\,d\phi = -\omega^1{}_2 \\ d\omega^1{}_2 &= \sin\theta\,d\theta \wedge d\phi = \frac{1}{a^2}\,(a\,d\theta) \wedge (a\sin\theta\,d\phi) \\ &= \frac{1}{a^2}\theta^1 \wedge \theta^2 \\ K &= \frac{1}{a^2} \end{aligned}$$

4.38 Example Torus

Using the the parametrization (See 4.23),

$$\mathbf{x} = ((b + a\cos\theta)\cos\phi, (b + a\cos\theta)\sin\phi, a\sin u)$$

the first fundamental form is

$$ds^2 = a^2 d\theta^2 + (b + a\cos u)^2 d\phi^2.$$

Thus, we have:

$$\begin{aligned}
\theta^1 &= a\,d\theta \\
\theta^2 &= (b + a\cos\theta)\,d\phi \\
d\theta^2 &= -a\sin\theta\,d\theta \wedge d\phi \\
&= \sin\theta\,d\phi \wedge \theta^1 = -\omega^2{}_1 \wedge \theta^1 \\
\omega^2{}_1 &= -\sin\theta\,d\phi = -\omega^1{}_2 \\
d\omega^1{}_2 &= \cos\theta\,d\theta \wedge d\phi = \frac{\cos\theta}{a(b + a\cos\theta)}\,(a\,d\theta) \wedge [(a + b\cos\theta)\,d\phi] \\
&= \frac{\cos\theta}{a(b + a\cos\theta)}\theta^1 \wedge \theta^2 \\
K &= \frac{\cos\theta}{a(b + a\cos\theta)}.
\end{aligned}$$

This result makes intuitive sense.

When $\theta = 0$, the points lie on the outer equator, so $K = \dfrac{1}{a(b+a)} > 0$ is the product of the curvatures of the generating circle and the outer equator circle. The points are elliptic.

When $\theta = \pi/2$, the points lie on the top of the torus, so $K = 0$. The points are parabolic.

When $\theta = \pi$, the points lie on the inner equator, so $K = \dfrac{-1}{a(b-a)} < 0$ is the product of the curvatures of the generating circle and the inner equator circle. The points are hyperbolic.

4.39 Example Orthogonal Parametric Curves

The examples above have the common feature that the parametric curves are orthogonal and hence $F = 0$. Using the same method, we can find a general formula for such cases. Since the first fundamental form is given by

$$I = Edu^2 + Gdv^2$$

We have:

$$
\begin{aligned}
\theta^1 &= \sqrt{E}\, du \\
\theta^2 &= \sqrt{G}\, dv \\
d\theta^1 &= (\sqrt{E})_v\, dv \wedge du = -(\sqrt{E})_v\, du \wedge dv \\
&= -\frac{(\sqrt{E})_v}{\sqrt{G}}\, du \wedge \theta^2 = -\omega^1{}_2 \wedge \theta^2 \\
d\theta^2 &= (\sqrt{G})_u\, du \wedge dv = -(\sqrt{G})_u\, dv \wedge du \\
&= -\frac{(\sqrt{G})_u}{\sqrt{E}}\, dv \wedge \theta^2 = -\omega^2{}_1 \wedge \theta^1 \\
\omega^1{}_2 &= \frac{(\sqrt{E})_v}{\sqrt{G}}\, du - \frac{(\sqrt{G})_u}{\sqrt{E}}\, dv \\
d\omega^1{}_2 &= -\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right) du \wedge dv + \frac{\partial}{\partial v}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right) dv \wedge du \\
&= -\frac{1}{\sqrt{EG}}\left[\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right)\right] \theta^1 \wedge \theta^2.
\end{aligned}
$$

Therefore, the Gaussian Curvature of a surface mapped by a coordinate patch in which the parametric lines are orthogonal is given by:

$$
K = -\frac{1}{\sqrt{EG}}\left[\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right)\right] \tag{4.92}
$$

Again, to connect with more classical notation, if a surface described by a coordinate patch $\mathbf{x}(u, v)$ has first fundamental for given by $I = E\ du^2 + G\ dv^2$, then

$$
\begin{aligned}
d\mathbf{x} &= \mathbf{x}_u\ du + \mathbf{x}_v\ dv, \\
&= \frac{\mathbf{x}_u}{\sqrt{E}}\sqrt{E}\ du + \frac{\mathbf{x}_v}{\sqrt{G}}\sqrt{G}\ dv, \\
&= \frac{\mathbf{x}_u}{\sqrt{E}}\theta^1 + \frac{\mathbf{x}_v}{\sqrt{G}}\theta^2, \\
d\mathbf{x} &= \mathbf{e}_1\ \theta^1 + \mathbf{e}_2\ \theta^2,
\end{aligned} \tag{4.93}
$$

where

$$
\mathbf{e}_1 = \frac{\mathbf{x}_u}{\sqrt{E}}, \quad \mathbf{e}_2 = \frac{\mathbf{x}_v}{\sqrt{G}}.
$$

Thus, when the parametric curves are orthogonal, the triplet $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e_3} = \mathbf{n}\}$ constitutes a moving orthonormal frame adapted to the surface. The awkwardness of combining calculus vectors and differential forms in the same equation is mitigated by the ease of jumping back and forth between the classical and the modern formalism. Thus, for example, covariant differential of the normal in 4.89 can be rewritten without the arbitrary vector in the operator LX as shown:

$$
\overline{\nabla}_X e_3 = \omega^1{}_3(X)\, e_1 + \omega^2{}_3(X)\, e_2, \tag{4.94}
$$

$$
d\mathbf{e}_3 = \mathbf{e}_1\ \omega^1{}_3 + \mathbf{e}_2\ \omega^2{}_3 = 0, \tag{4.95}
$$

The equation just expresses the fact that the components of the Weingarten map, that is, the second fundamental form in this basis, can be written as some symmetric matrix given by:

$$\begin{aligned}\omega^1{}_3 &= l\theta^1 + m\theta^2,\\ \omega^2{}_3 &= m\theta^1 + n\theta^2.\end{aligned} \tag{4.96}$$

In the case that $E = 1$, so that the metric is

$$ds^2 = du^2 + G(u,v)dv^2, \tag{4.97}$$

the equation for curvature above reduces even further to:

$$K = -\frac{1}{\sqrt{G}}\frac{\partial^2\sqrt{G}}{\partial u^2}. \tag{4.98}$$

The case is not as special as it appears at first. The change of parameters

$$\hat{u}' = \int_0^u \sqrt{E}\, du$$

results on $d\hat{u}^2 = E\, du^2$, and thus it transforms an orthogonal system to one with $E = 1$. The parameters are reminiscent of polar coordinates $ds^2 = dr^2 + r^2\, d\phi^2$. We will have more to say about this in the context of geodesics.

A slick proof of the Teorema Egregium can be obtained by differential forms. Let $F : \mathcal{M} \to \widetilde{\mathcal{M}}$ be an isometry between two surfaces with metrics g and $\tilde{g}$ respectively. Let $\{e_\alpha\}$ be an orthonormal basis for dual basis $\{\theta^\alpha\}$. Define $\tilde{e}_\alpha = F_* e_\alpha$. Recalling that isometries preserve inner products, we have

$$\langle \tilde{e}_\alpha, \tilde{e}_\beta\rangle = \langle F_* e_\alpha, F_* e_\beta\rangle = \langle e_\alpha, e_\beta\rangle = \delta_{\alpha\beta}.$$

Thus, $\{\tilde{e}_\alpha\}$ is also an orthonormal basis of the tangent space of $\widetilde{\mathcal{M}}$. Let $\tilde{\theta}^\alpha$ be the dual forms and denote with tilde's the connection forms and Gaussian curvature of $\widetilde{\mathcal{M}}$.

4.40 Theorem (Teorema Egregium)
a) $F^*\tilde{\theta}_\alpha = \theta_\alpha$,
b) $F^*\tilde{\omega}^\alpha{}_\beta = \omega^\alpha{}_\beta$,
c) $F^*\tilde{K} = K$.

Proof:

a) It suffices to show that the forms agree on basis vectors. We have

$$\begin{aligned}F^*\tilde{\theta}_\alpha(e_\beta) &= \tilde{\theta}_\alpha(F_* e_\beta),\\ &= \tilde{\theta}_\alpha(\tilde{e}_\beta),\\ &= \delta^\alpha{}_\beta,\\ &= \theta(e_\beta).\end{aligned}$$

b) We compute the pull-back of the first structure equation in $\widetilde{\mathcal{M}}$:

$$\begin{aligned}d\tilde{\theta}^\alpha + \tilde{\omega}^\alpha{}_\beta \wedge \tilde{\theta}^\beta &= 0,\\ F^* d\tilde{\theta}^\alpha + F^*\tilde{\omega}^\alpha{}_\beta \wedge F^*\tilde{\theta}^\beta &= 0,\\ d\theta^\alpha + F^*\tilde{\omega}^\alpha{}_\beta \wedge \theta^\beta &= 0,\end{aligned}$$

The connection forms are defined uniquely by the first structure equation, so

$$F^*\tilde{\omega}^\alpha{}_\beta = \omega^\alpha{}_\beta$$

.

c) In a similar manner, we compute the pull-back of the curvature equation:

$$\begin{aligned}
d\tilde{\omega}^1{}_2 &= \tilde{K}\,\tilde{\theta}^1 \wedge \tilde{\theta}^2, \\
F^* d\tilde{\omega}^1{}_2 &= (F^*\tilde{K})\,F^*\tilde{\theta}^1 \wedge F^*\tilde{\theta}^2, \\
dF^*\tilde{\omega}^1{}_2 &= (F^*\tilde{K})\,F^*\tilde{\theta}^1 \wedge F^*\tilde{\theta}^2, \\
d\omega^1{}_2 &= (F^*K)\,\theta^1 \wedge \theta^2,
\end{aligned}$$

So again by uniqueness, $F^*K = K$.

4.41 Example Catenoid - Helicoid

Perhaps the most celebrated classical is that of mapping between a helicoid $\mathcal{M}$ and a catenoid $\widetilde{\mathcal{M}}$. Let $a = c$, and label the coordinate patch for the former as $\mathbf{x}(u^\alpha)$ and $\mathbf{y}(\tilde{u}^\alpha)$ for the latter. The first fundamental forms are given as in 4.24and 4.25.

$$\begin{aligned}
ds^2 &= du^2 + (u^2+a^2)\,dv^2, & & E = 1 \quad G = u^2 + a^2, \\
d\tilde{s}^2 &= \frac{\tilde{u}^2}{\tilde{u}^2 - a^2}\,d\tilde{u}^2 + \tilde{u}^2\,d\tilde{v}^2 & \text{with} \quad & \tilde{E} = \frac{\tilde{u}^2}{\tilde{u}^2 - a^2} \quad \tilde{G} = \tilde{u}^2.
\end{aligned}$$

Let $F : \mathcal{M} \to \widetilde{\mathbf{M}}$ be the mapping $\mathbf{y} = F\mathbf{x}$, defined by $\tilde{u}^2 = u^2 + a^2$ and $\tilde{v} = v$. Since $\tilde{u}\,d\tilde{u} = u\,du$, we have $\tilde{u}^2\,d\tilde{u}^2 = u^2\,du^2$ which shows that the mapping preserves the metric and hence it is an isometry. The Gaussian curvatures K and $\tilde{K}$ follow from an easy computation using formula 4.92.

$$K = \frac{-1}{\sqrt{u^2+a^2}}\frac{\partial}{\partial u}\left(\frac{\partial}{\partial u}\sqrt{u^2+a^2}\right) = \frac{a^2}{(u^2+a^2)^2}, \tag{4.99}$$

$$\tilde{K} = -\frac{\sqrt{\tilde{u}^2-a^2}}{\tilde{u}^2}\;\frac{\partial}{\partial \tilde{u}}\left(\frac{\sqrt{\tilde{u}^2-a^2}}{\tilde{u}}\right) = -\frac{a^2}{\tilde{u}^4} \tag{4.100}$$

It is immediately evident by substitution that as expected $F^*\tilde{K} = K$. Figure 4.14 shows several stages of a one-parameter family $\mathcal{M}_t$ of isometries deforming a catenoid into a helicoid. The one-parameter family of coordinate patches chosen is $\mathbf{z}_t = (\cos t)\,\mathbf{x} + (\sin t)\,\mathbf{y}$

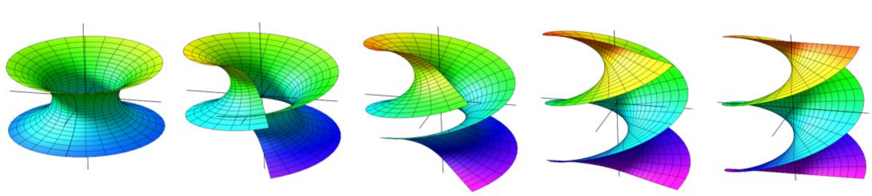


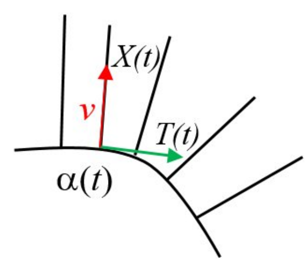
Fig. 4.14: Catenoid - Helicoid isometry

Chapter 5

Geometry of Surfaces

5.1 Surfaces of Constant Curvature

5.1.1 Ruled and Developable Surfaces



We present a brief discussion of surfaces of constant curvature $K = 0$. since $K = k_1 k_2$, a surface with zero Gaussian curvature at each point must have a principal direction with zero normal curvature, that is, either $k_1 = 0$ or $k_2 = 0$. It is therefore a necessary condition for a surface to have $K = 0$, that at each point there be a principal direction which is a straight line. A surface having this property of containing a straight line or segment of a straight line at each point is called a ruled surface. We may think of ruled surface as a surface generated by the motion of a straight line. Given a point p on a ruled surface, let $\alpha(t)$ be curve with $\alpha(0) = p$, and let $X(t)$ be a unit vector field on the curve and pointing along the lines at their intersection points with the curve. On can then parametrize the surface near p by a coordinate patch $\mathbf{y}(t, v) = \alpha(t) + vX(t)$ as shown in the figure.

Having a straight line passing through each point in the surface is a necessary but not sufficient condition to insure that $K = 0$, as illustrated by the following examples

1) Saddle. Consider the saddle $z = xy$ which is trivially parametrized by the coordinate patch $\mathbf{y}(u, v) = \langle u, v, uv\rangle$. The patch can be written as $\mathbf{y}(u, v) = \langle u, 0, v\rangle + v\langle 0, 1, u\rangle$ or as $\mathbf{y}(u, v) = \langle v, 0, u\rangle + u\langle 1, 0, v\rangle$, so that the surface is doubly-ruled as A shown if figure 5.1(a). The rulings are the coordinate curves u= constant, and v= constant. This neat fact is reflected in some architectural designs of simple structures with roofs made of straight slabs arranged in the shape of a hyperbolic paraboloid. A short computation gives $K = -(1 + u^2 + v^2)^{-2} = -(1 + x^2 + y^2)^{-2}$

2) Hyperboloid. A common calculus example of a doubly ruled surface is given by circular hyperboloids of one sheet. Consider the circle$\alpha(u) = \langle \cos u, \sin u, 0\rangle$, and the vector field $X(u) = \dot{\alpha}+\mathbf{k}$ which points at a constant skew angle $\pi/4$. Then the $\langle x, y, z\rangle$ coordinates in the parametrization

$$\begin{aligned}
\mathbf{y}(u, v) &= \alpha(u) + vX(u),\\
&= \langle \cos u, \sin u, 0\rangle + v\langle -\sin u, \cos u, 1\rangle,\\
&= \langle \cos u - v \sin u, \sin u + v \cos u, v\rangle,
\end{aligned}$$

satisfy the equation $x^2 + y^2 - z^2 = 1$; that is, the surface is a circular hyperboloid of one sheet. The coordinate curves s =constant are straight line generators. If instead, we choose $X(u) = -\dot{\alpha} + \mathbf{k}$, we get the same surface, but with an orthogonal set of line-generators as shown in figure 5.1(b). This is an example of a surface in which the asymptotic curves are orthogonal at each point. Tangent planes to the surface at any point in the circle to the circle $x^2+y^2=1$ at the throat intersect the hyperboloid in intersecting pair of line generators. The Gaussian curvature is also negative and is given by $K=-(1+2v^2)^2=-(1+2z^2)^2$. The double-ruled nature of the circular hyperboloid has been exploited by civil engineers in the design of heavy-duty gears with long teeth engaging along the lines. The double-ruling is also advantageous for the construction of the metal frame to support the cooling towers of nuclear reactors.

3) Möbius Band. Another famous example is given by the Möbius strip for which we choose the coordinate patch

$$\begin{aligned}
\mathbf{y}(u,v) &= \alpha(u)+vX(u),\\
\alpha(t) &= \langle \cos 2u, \sin 2u, 0\rangle,\\
X(u) &= \langle \cos u\cos 2u, \cos u \sin 2u, \sin u\rangle,\\
\mathbf{x}(u,v) &= \langle \cos 2u + v\cdot \cos u\cos 2u, \sin 2u + v\cdot \cos u\sin 2u, v\sin u\rangle,
\end{aligned}$$

The curve $\alpha(u)$ is a circle, and the vector $X(u)$ on the circle points in a direction that winds around by an angle π in one revolution. The parameter v here is restricted to $[-0.2, 0.2]$, so the generating line segment creates a non-orientable surface as shown in figure 5.1. The Gaussian curvature is somewhat messy, but the computation shows that K is negative everywhere.

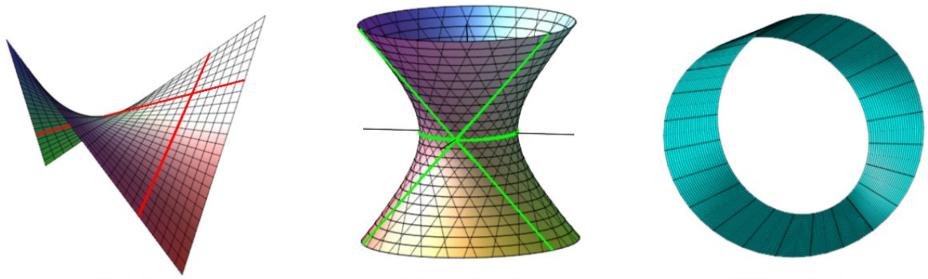


Fig. 5.1: Examples of Ruled Surfaces

Let $\mathcal{M}$ be a ruled surface with unit normal$N$ and let X be a unit vector field tangent along the straight lines that generate the surface. The straight lines are geodesic in $\mathbf{R}^3$ so $\overline{\nabla}_X X=0$. By Gauss's equation 4.71, $\nabla_X X=0$, so the line is also a geodesic on the surface and $\langle LX, X\rangle = 0$, that is, the generator lines are asymptotic. Let Y be a unit tangent vector field orthogonal to X so that the pair constitutes an orthogonal basis of the tangent space at each point. Then

$$\begin{aligned}
K &= \langle LX, X\rangle\langle LY, Y\rangle - \langle LX, Y\rangle^2,\\
&= -\langle LX, Y\rangle^2 = -f^2,
\end{aligned}$$

so we conclude that $K\leq 0$. If the vectors X and Y are not an orthogonal basis, then by equation 4.65, written in classical notation, the curvature of a ruled surface is

$$K=-\frac{f^2}{EG-F^2}. \tag{5.1}$$

The general formula for the Gaussian curvature of a ruled surface is obtained by a straight forward computation. We have:

$$
\begin{aligned}
\mathbf{y}_u &= \alpha' + vX', \\
\mathbf{y}_v &= X, \\
\mathbf{y}_u \times \mathbf{y}_v &= (\alpha' + vX') \times X, \\
EG - F^2 &= \|\mathbf{y}_u \times \mathbf{y}_v\| = \|X \times (\alpha' + vX')\|, \mathbf{y}_{uv} \qquad\qquad = X',\ \mathbf{y}_{vv} = 0.
\end{aligned}
$$

Hence, $g = \langle \mathbf{y}_{vv}, N\rangle = 0$, and $f = \langle \mathbf{y}_{uv}, N\rangle = (\alpha' X X')/\sqrt{EG - F^2}$, where we are using the notation for the triple product. The resulting curvature is:

$$
K - \frac{(\alpha' X X')}{\|X \times (\alpha' + vX')\|^4} \tag{5.2}
$$

We may choose the orthogonal trajectories $\alpha(u)$ to be integral curves of Y parametrized by arc length, so that $Y = \alpha'(u) = T$. Since by choice, T and X are orthogonal unit vectors tangent to the surface, $N = T \times X$ is normal to the surface and we have an orthonormal frame at each point. The covariant derivatives of the frame with respect to T along the one-parameter curves α is just derivative with respect to the parameter, so we we get Frenet-like frame

$$
\begin{array}{lcrrcl}
T' & = & & c_1 X & + & c_2 N \\
X' & = & -c_1 T & & + & c_3 N \\
N' & = & -c_2 T & -c_3 X & &
\end{array} . \tag{5.3}
$$

A one-line computation gives:

$$
c_3 = -\langle X, N'\rangle = \langle N, X'\rangle = (TXX').
$$

The function

$$
p(u) = \frac{(TXX')}{\|X'\|^2} = \frac{c_3}{{c_1}^2 + {c_3}^2},
$$

is called the **distribution parameter**. Substituting 5.3 into 5.2 we rewrite the Gaussian curvature as

$$
\begin{aligned}
K &= -\frac{(TXX')}{\|X \times (T + vX')\|^4}, \\
&= -\frac{{c_3}^2}{[1 - 2c_1 v + {c_1}^2 v^2 + v^2 {c_3}^2]^2}.
\end{aligned}
$$

The special curve along which $a_1 = 0$ that is, $\langle T', X\rangle = 0$ is called the stricture curve. Using the parametrization with the base curve being the stricture curve, $p(u) = 1/c_3$ and

$$
K = -\frac{{c_3}^2}{(1 + v^2 c_3(2)} = \frac{p^2(u)}{(p^2(u) + v^2))}. \tag{5.4}
$$

A beautiful example is the hyperboloid of revolution in figure 5.1(b). The circle $\alpha(t)$ at the throat used to generate the surface is the stricture curve. It turns out that X does not need to be orthogonal to T as it is the case here, as long as $\|x\| = 1$ and $\langle T', X\rangle = 0$.

A ruled surface is called a **developable surface** if in addition, $LX = \overline{\nabla}_X N = 0$, that is, the normal vector is parallel along the generating lines. Then equality holds and we have the following theorem

5.1 Theorem A necessary and sufficient condition for surface to be developable is to have Gaussian curvature $K = 0$.

It is surprising that general case of a closed and connected surface with $K = 0$ to be developable was not proved until 1961 by proved by Massey.

A particularly interesting developable surfaces are those in which the vector X is taken to be the tangent vector T of the curve α itself. A surface with this property is called a **tangential developable**.

5.2 Example Developable Helicoid. A helicoid $\mathbf{x}(u, v) = \langle 0, 0, v\rangle$ can be written in the form $\mathbf{x} = \alpha(v) + uX(v)$, where $\alpha(v) = \langle \cos v, \sin v, 0\rangle$ and $X(v) = \langle \cos v, \sin v, 0\rangle$, so it is a ruled surface. The surface has negative curvature as computed in 4.99 and the stricture curve is the z-axis. A neat related surface is obtained by the tangential developable of a helix. We choose $\alpha(u) = \langle \cos u, \sin u, u\rangle$ and $X = T = \langle -\sin u, \cos u1\rangle$ so that

$$\mathbf{x}(u, v) = \langle \cos u - v \sin u, \sin u + v \cos u, u + v\rangle. \tag{5.5}$$

Since this is flat surface having $K = 0$ it is isometric to a plane. Indeed, if one takes a thin cardboard annulus with a slit in the xy-plane with the appropriate radius, one can bend the annulus around a cylinder by lifting one edge of the slit, thus creating a ribbon that wraps around the cylinder as shown in figure 5.2. A magnificent architectural example is exhibited by the spiral staircase near the pyramid of the Louvre museum. For the maple-generated image, a small numerical computation was carried to figure out the vertical shift and radius of the helicoid so that the staircase and the supporting developable match at the helix of intersection.

Fig. 5.2: Developable Helicoid

5.1.2 Surfaces of Constant Positive Curvature

In this section we prove a few global theorems. We assume the reader is acquainted with the notion of a compact space. In particular, in $\mathbf{R}^n$ a compact set is closed and bounded so it is contained in a ball of sufficiently large radius, centered at the origin. We are concerned with compact manifolds, which by definition are locally Euclidean and have a differentiable structure. Thus a compact surface in $\mathbf{R}^3$ can not have any edges or creases, and the tangent space is well defined all points.

5.3 Theorem In any compact surface in $\mathbf{R}^3$ there exists at least one point p at which $K(p) > 0$.
Proof: Let $\mathcal{M}$ be compact. Consider the function $f : \mathcal{M} \to \mathbf{R}$ defined by $f = \|\mathbf{x}\}$, where $\mathbf{x}$ are local coordinates of a point on the surface. This is continuous function on compact space, so it attains a maximum at a least one point p. The geometric interpretation of p is that is farthest away from the origin. The intuition about this theorem is simply that near the point, the surface is entirely on one one side of the tangent plane as shown in figure 5.3, so the principal curvatures have

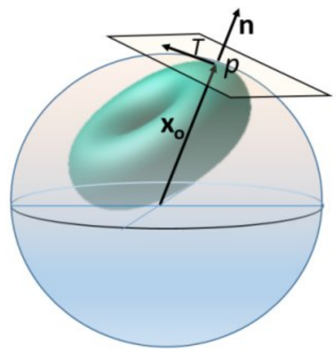


Fig. 5.3: Compact Surface

the same sign. We make this formal. Let R be the distance from p to the origin, and construct a sphere of radius R centered at the origin. The sphere will be tangential to the surface at p. Given a unit tangent T, let $\alpha(t)$ be unit speed integral curve near p. That is $\alpha(0) = p$ and unit tangent vector $T = \alpha'(0)$. The composite function $f(\alpha(t))$ also has a maximum at p, so by the second derivative test, we have $[f(\alpha)]'(0) = 0$, and $[f(\alpha)]''(0) < 0$. By the definition of f, $f(\alpha) = \|\alpha^2\| = \langle \alpha, \alpha \rangle$, so $[f(\alpha)]'(0) = 2\langle \alpha, \alpha' \rangle(0) = 0$. We conclude that the position vector $\mathbf{x}_o = \alpha(0)$ of the point p is orthogonal to T. Since this is true for any such T, the vector $\mathbf{x}_o$ is also normal to the surface, so that the unit normal is $\mathbf{n} = \mathbf{x}_o/R$. Computing the second derivative we get

$$\begin{aligned} \tfrac{1}{2}[f(\alpha)]''(0) &= \langle \alpha, \alpha' \rangle'(0), \\ &= \langle T, T \rangle + \langle \mathbf{x}_o, \alpha''(0) \rangle, \\ &= 1 + R\langle \mathbf{n}, \alpha''(0) \rangle < 0, \end{aligned}$$

, But clearly $\langle \mathbf{n}, \alpha''(0) \rangle$ is the normal curvature along T so

$$k_n(p) < -\frac{1}{R},$$

Again, since T was arbitrary, the normal curvature is less that $-1/R$ in any direction, a geometric indication that the surface is bending more than the sphere as intuitively shown by the picture. Therefore

$$K(p) = \kappa_1 \kappa_2 > \frac{1}{R^2} > 0.$$

5.4 Theorem In a surface in which the coordinate directions are chosen to be the principal directions of curvature, the Codazzi equations are

$$\begin{aligned} \frac{\partial \kappa_1}{\partial u} &= \frac{1}{2}\frac{E_v}{E}(\kappa_2 - \kappa_1), \\ \frac{\partial \kappa_2}{\partial u} &= \frac{1}{2}\frac{G_u}{G}(\kappa_1 - \kappa_2) \end{aligned} \tag{5.6}$$

Proof: Let X and Y be eigenvectors of L, so that

$$LX = \kappa_1 X, \quad \kappa_1 = \frac{\langle LX, X\rangle}{\langle X, X\rangle} = \frac{e}{E},$$
$$LY = \kappa_2 Y, \quad \kappa_1 = \frac{\langle LY, Y\rangle}{\langle Y, Y\rangle} = \frac{g}{G}.$$

If $\kappa_1 \neq \kappa_2$, the eigenvectors are orthogonal, so taken them as the parametric directions means that $F = f = 0$. The Codazzi equations 4.83 are obtained by setting to zero the normal component of $\mathbf{x}_{\alpha\beta\gamma} - \mathbf{x}_{\beta\alpha\gamma} = 0$. In terms of the covariant derivative formulation of the Gauss 4.71 with $X = e_1 = \mathbf{x}_u$, $Y = e_2 = \mathbf{x}_v$, $Z = e_\gamma$, the normal component of $(\overline{\nabla}_X\overline{\nabla}_Y - \overline{\nabla}_Y\overline{\nabla}_X)Z = 0$ result in the equations of Codazzi in the form: (See 6.13)

$$\langle \overline{\nabla}_X LY - \overline{\nabla}_X LX, Z\rangle = 0,$$
$$\overline{\nabla}_X LY - \overline{\nabla}_X LX = 0$$

We proceed to expand this equation:

$$\overline{\nabla}_{e_1} Le_2 - \overline{\nabla}_{e_2} Le_1 = 0,$$
$$\overline{\nabla}_{e_1}(\kappa_2 e_2) - \overline{\nabla}_{e_2}(\kappa_1 e_1) = 0,$$
$$\frac{\partial \kappa_2}{\partial u} e_2 + \kappa_2 \Gamma^{\alpha}{}_{12} e_\alpha - \frac{\partial \kappa_1}{\partial v} e_1 - \kappa_1 \Gamma^{\alpha}{}_{21} e_\alpha = 0.$$

Setting the e_1 and e_2 components to zero, we get:

$$\frac{\partial \kappa_2}{\partial_u} = \kappa_1 \Gamma^2{}_{21} - \kappa_2 \Gamma^2{}_{12} = (\kappa_1 - \kappa_2)\Gamma^2{}_{12},$$
$$\frac{\partial \kappa_1}{\partial_v} = \kappa_2 \Gamma^1{}_{12} - \kappa_1 \Gamma^1{}_{21} = (\kappa_2 - \kappa_1)\Gamma^1{}_{21}.$$

The result follows immediately from the expressions for the Christoffel symbols 4.75 after setting $F = 0$.

5.5 Proposition (Hilbert) Let p be a non-umbilic point and $\kappa_1(p) > \kappa_2(p)$. If κ_1 has a local maximum at p and κ_2 has a local minimum at p, the $K(p) < 0$.
Proof: Take the asymptotic curves as parametric curves as in the preceding proposition. Suppose $\kappa_1(p) > \kappa_2(p)$ and that the principal curvatures are local extrema. Then $(\kappa_1)_u = (\kappa_2)_v = 0$ so by equation 5.6, we have $E_v = G_u = 0$. Applying the second derivative test by differentiating 5.6 at p we get

$$(\kappa_1)_{vv} = \frac{1}{2}\frac{E_{vv}}{E}(\kappa_2 - \kappa_1) \leq 0,$$
$$(\kappa_2)_{uu} = \frac{1}{2}\frac{G_{uu}}{G}(\kappa_1 - \kappa_2) \geq 0,$$

which implies that $E_{vv} \geq 0$, end $G_{uu} \geq 0$. On the other hand, noting as above that $E_v = G_u = 0$, the Gaussian curvature formula 4.92 gives

$$K = -\frac{1}{2EF}(E_{vv} + G_{uu}) \leq 0.$$

5.6 Theorem (Liebmann) A compact manifold $\mathcal{M}$ in $\mathbf{R}^3$ of constant Gaussian curvature K is a sphere of radius R with $K = 1/R^2$.

Proof: Since $\mathcal{M}$ is compact, there is at least one point at which $K > 0$ and since K is constant, $K > 0$ everywhere. We prove by contradiction that all points are umbilic. Suppose there exists an non-umbilic point. Without loss of generality we assume that the larger principal curvature is κ_1. The principal curvatures are continuous functions in a compact space, so there is a point p at which κ_1 is maximum. Since $K = \kappa_1\kappa_2$ =constant then at p, κ_2 is a minimum. By Hilbert's theorem above, $K(p) < 0$ which is a contradiction. So $\mathcal{M}$ is a sphere so some radius R and $K = 1/R^2$.

5.1.3 Surfaces of Constant Negative Curvature

The geometry of surfaces of constant negative curvature is very rich and is has a number of neat applications to physics. If $K < 0$, then it must be the case that the principal curvatures κ_1 and κ_2 have different signs. All points on the surface are hyperbolic, and by Hilbert's theorem there are no compact surfaces of constant negative curvature. In addition, since $\kappa_1 \neq \kappa_2$, there always exist orthogonal asymptotic curves with asymptotic directions in along the eigenvectors of the the second fundamental form. The prototype of a surface of constant negative curvature is the pseudosphere introduced in equation 4.23 which we repeat here for convenience.

$$\mathbf{x}(u, v) = \langle a \sin u \cos v, a \sin u \sin v, a(\cos u + \ln(\tan \frac{u}{2}))\rangle.$$

To compute the Gaussian curvature we first verify that the first fundamental form is as stated in 4.23. We have:

$$\begin{aligned}
\mathbf{x}_u &= \langle a \cos u \cos v, a \cos u \sin v, a\frac{\cos^2 u}{\sin u}\rangle, \\
\mathbf{x}_v &= \langle -a \sin u \sin v, a \sin u \cos v, 0\rangle, \\
E &= a^2 \cos^2 u + a^2 \frac{\cos^4 u}{\sin^2 u}, \\
&= a^2 \cos^2 u (1 + \frac{\cos^2 u}{\sin^2 u}), \\
&= a^2 \cot^2 u, \\
F &= 0, \\
G &= a^2 \sin^2 u.
\end{aligned}$$

So the parametric curves are orthogonal, and

$$I = a^2 \cot^2 u \, du^2 + a^2 \sin^2 u \, dv^2.$$

Inserting into formula 4.92, we get

$$\begin{aligned}
K &= -\frac{1}{a^2 \cos u}\left[\frac{\partial}{\partial u}\left(\frac{\sin u}{a \cos u}\frac{\partial}{\partial u}(a \sin u)\right)\right], \\
&= -\frac{1}{a^2 \cos u}\frac{\partial}{\partial u}(\sin u), \\
&= -\frac{1}{a^2}.
\end{aligned} \tag{5.7}$$

Another common parametrization of the pseudosphere is obtained by the substitution

$$\mu = \ln \tan(\tfrac{u}{2}). \tag{5.8}$$

so that $e^{\mu} = \tan(u/2)$. The substitution is somewhat related to the classical Gudermannian. We have:

$$\begin{aligned} \operatorname{sech}\mu &= \frac{2}{e^{u/2}+e^{-u/2}}, \\ &= \frac{2}{\tan(u/2)+\cot(u/2)}, \\ &= 2\sin(u/2)\cos(u/2), \\ &= \sin u, \end{aligned} \qquad \text{and} \qquad \begin{aligned} \tanh\mu &= \frac{e^{u/2}-e^{-u/2}}{e^{u/2}+e^{-u/2}}, \\ &= \frac{\tan(u/2)-\cot(u/2)}{\tan(u/2)+\cot(u/2)}, \\ &= \sin^2(u/2)-\cos^2(u/2), \\ &= -\cos u \end{aligned}$$

In simplifying the equations above we multiplied top and bottom of the fractions by $\sin(u/2)\cos(u/2)$. In terms of the parameter μ the coordinate patch for the pseudosphere becomes

$$\mathbf{x}(u,v) = \langle a\operatorname{sech}\mu\cos v, a\operatorname{sech}\mu\sin v, a(\mu-\tanh\mu)\rangle, \tag{5.9}$$

and the fundamental forms are:

$$\begin{aligned} I &= \tanh^2\mu\, d\mu^2 + \operatorname{sech}^2\mu\, dv^2, \\ II &= -\operatorname{sech}\mu\, d\mu^2 + \operatorname{sech}\mu\tanh\mu\, dv^2. \end{aligned}$$

Using this latter parametrization we compute the surface area and volume of the top half of the pseudosphere.

$$\begin{aligned} S &= \int\int \sqrt{EG-F^2}, \\ &= \int_0^{\pi}\int_0^{\infty} (a\operatorname{sech}\mu)(a\tanh\mu)\, d\mu\, dv, \\ &= 4\pi a^2, & (5.10) \\ V &= \pi a^3 \int_{-\infty}^{\infty} \operatorname{sech}^2\mu\tanh^2\mu\, d\mu, & (5.11) \\ &= \tfrac{2}{3}\pi a^3. & (5.12) \end{aligned}$$

It is interestng to note that the surface area is which is exactly the same as that of a sphere of radius a. wheraas the volume of revolution is half the volume of the sphere. Without loss of understanding of the geometry, for the rest of this section we set $a = 1$, so that $K = -1$

We have the following theorem:

5.7 Theorem Let $\mathcal{M}$ be a surface has constant negative curvature $K = -1$. If the parametric curves are chosen to be the asymptotic directions, there exists some quantity ω so that the first fundamental form can be written as:

$$I = \cos^2\omega\, du^2 + \sin^2\omega\, dv^2, \tag{5.13}$$

Proof: The proof amounts to analyzing the integrability conditions represented by the Codazzi equations. In the brilliant book by Eisenhart, the author writes down the Codazzi equations and notes that the choice of E and G above are solutions of the equations. We take a more modest approach and show the steps on how to find a solution using the Cartan formalism. we have seen that using the asymptotic curves as parametric curves means that $F = f = 0$, $\kappa_1 = E/e$ and $\kappa_1 = G/e$. Let $E = \alpha^2$ and $G = \beta^2$ so that the first fundamental form is:

$$I = \alpha^2\, du^2 + \beta^2\, dv^2.$$

We choose as $\theta^1 = \alpha\ du$ and $\theta^2 = \beta\ dv$ as basis for the cotangent space dual to $\{e_1, e_1\}$. Let e_3 be the unit normal to the surface. We have:

$$\begin{aligned} Le_1 = \kappa_1 e_1 = \overline{\nabla}_{e_1} e_3 = \omega^i{}_3(e_1)e_i, \\ Le_2 = \kappa_2 e_2 = \overline{\nabla}_{e_2} e_3 = \omega^i{}_3(e_2)e_i, \end{aligned}$$

so,

$$\begin{aligned} \omega^1{}_3 = \kappa_1\theta^1 = \kappa_1\alpha\ du, \\ \omega^2{}_3 = \kappa_2\theta^2 = \kappa_2\beta\ dv. \end{aligned}$$

On the other hand,

$$\begin{aligned} d\theta^1 &= \alpha_v\ dv \wedge du = -(\frac{\alpha_v}{\beta}\ du) \wedge \theta^2 = -\omega^1{}_2 \wedge \theta^2, \\ d\theta^2 &= \beta_u\ du \wedge dv = -(\frac{\beta_u}{\alpha}\ dv) \wedge \theta^1 = -\omega^2{}_1 \wedge \theta^1, \\ \omega^1{}_2 &= \frac{\alpha_v}{\beta}\ du - \frac{\beta_u}{\alpha}\ dv. \end{aligned} \tag{5.14}$$

Recall the Codazzi equations **??**

$$\begin{aligned} d\omega^1_3 = -\omega^1_2 \wedge \omega^2_3, \\ d\omega^2_3 = -\omega^2_1 \wedge \omega^1_3. \end{aligned}$$

Inserting the connection forms into the first Codazzi equation gives:

$$(\kappa_1\alpha)_v\ dv \wedge du + [\frac{\alpha_v}{\beta}\ du - \frac{\beta_u}{\alpha}\ dv] \wedge \kappa_2\beta\ dv = 0.$$

Since $\kappa_1\kappa_2 = -1$, We can thus eliminate κ_2 and solve for α.

$$\begin{aligned} (\kappa_1\alpha)_v - \alpha_v\kappa_2 &= 0, \\ (\kappa_1 - \kappa_2)\alpha_v + (\kappa_1)_v\alpha &= 0, \\ \frac{\alpha_v}{\alpha} &= -\frac{(\kappa_1)_v}{\kappa_1 - \kappa_2}, \\ &= -\frac{(\kappa_1)_v}{\kappa_1 + (1/\kappa_1)}, \\ &= -\frac{\kappa_1(\kappa_1)_v}{\kappa_1{}^2 + 1}, \\ \frac{\partial}{\partial v}(\ln\alpha) &= -\frac{\partial}{\partial v}\ln[(\kappa_1{}^2 + 1)^{1/2}]. \end{aligned}$$

We may set

$$\begin{aligned} \kappa_1 &= \tan\omega, \\ \kappa_2 &= -\cot\omega \end{aligned} \tag{5.15}$$

, for some ω. Then $(\kappa_1{}^2 + 1)^{1/2} = \sec\omega$, so

$$\frac{\partial}{\partial v}(\ln \alpha) = -\frac{\partial}{\partial v}\ln(\sec\omega) = \frac{\partial}{\partial v}\ln(\cos\omega).$$

We choose the simplest solution $\alpha = \cos\omega$. By a completely analogous computation Using the second Codazzi equation, we get $\beta = \sin\omega$ and that proves the theorem.

5.8 Theorem If surface with $K = -1$ has first fundamental form written as $I = \cos^2\omega\ du^2 + \sin^2\omega\ du^2$, then ω satisfies the so-called sine-Gordon equation:

$$\omega_{uu} - \omega_{vv} = \sin\omega\cos\omega. \tag{5.16}$$

Proof: Here $E = \cos^2\omega$ and $G = \sin^2\omega$. The theorem follows immediately from inserting these into the Gauss curvature equation in orthogonal coordinates 4.92 and setting $K = -1$

$$K = -\frac{\omega_{uu} - \omega_{vv}}{\sin\omega\cos\omega} = -1$$

The following transformation is often made:

$$u = \hat{u} + \hat{v} \qquad u = \hat{u} - \hat{v}.$$

A quick computation yields a transformed fundamental form

$$\hat{I} = d\hat{u}^2 + 2\cos\hat{\omega}\ d\hat{u}d\hat{v} + d\hat{v}^2, \tag{5.17}$$

where $\omega = \hat{\omega}/2$. A coordinate system in which the first fundamental form is of this type is called a **Tchebychev Patch** (eventually one has to make a choice on how to transliterate from the Cyrillic alphabet). The corresponding curvature equation is

$$\hat{\omega}_{uv} = \sin\hat{\omega} \tag{5.18}$$

The sine-Gordon equation is one of class of very special type of nonlinear partial differential equations which admit **soliton** solutions. This is an incredibly rich area of research that would take us into whole new branch of mathematics. We constrain our discussion to certain transformations that allow one to obtain new solutions from known solutions, and associate these with pseudospherical surfaces, that is, surfaces in $\mathbf{R}^3$ with constant negative curvature. We note that if in the Sine-Gordon equation 5.16 one sets $v = t$ where t is a time parameter, what we have is a non-linear wave equation with speed $v = 1$. The reader will then recognize the transformation $u = \hat{u} + \hat{t}\ u = \hat{u} - \hat{t}$ as the equations of characteristics. It is thus not surprising that the equation has solutions of the form $f(u - t)$.

5.1.4 Bäcklund Transforms

5.9 Definition Let $\mathcal{M}$ by a surface with $K = -1$ and let $F : \mathcal{M} \to \hat{\mathcal{M}}$ be a map to another surface $\hat{\mathcal{M}}$. Let $\hat{p} = F(p)$ and $N(p)$ and $\hat{N}(\hat{p})$ be the unit normals at p and $\hat{p}$ respectively. $\hat{\mathcal{M}}$ is called a **Bäcklund Transform** (BT) of $\mathcal{M}$ with constant angle of inclination σ, if for all p:

a) The angle between N and $\hat{N}$ is σ;

b) The distance λ between p and $\hat{p}$ is $\sin\sigma$

c) the segment $\overline{p\hat{p}}$ is tangent to $\mathcal{M}$ at p

Bäcklund proved in 1883 that F maps pseudospherical surfaces to pseudospherical surfaces and asymptotic lines to asymptotic lines. He also shows that given any unit tangent vector at p which is

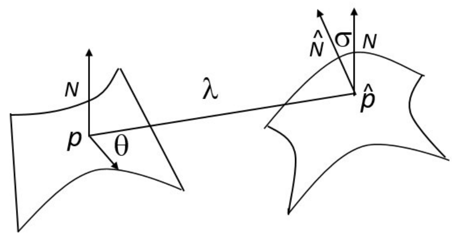


Fig. 5.4: Bäcklund Transform

not an asymptotic direction, a BT exists with $\overline{p\hat{p}}$ in the direction of that tangent. The idea behind the proof of the BT theorem is basically to find the quantify the transformation from orthonormal frame at p to the orhtonormal frame at $\hat{p}$, find the conditions required for $\hat{K} = -1$, and write down the integrability conditions for the Cartan structure equations. The transformation consists of a rotation by an angle σ, a translation from p to $\hat{p}$ and a rotation by an angle θ to align the frame with segment $\overline{p\hat{p}}$. This could be done all at once, but we prefer to do carry out the process in two stages. In the first stage, we apply the translation of the frame assumming that the segment joining p and $\hat{p}$ is parallel to the basis vector e_1 at p followed by a rotation by an angle σ around the e_1 direction. We use this seek conditions for guarantee that $\hat{K} = -1$. In stage two, we apply a rotation by an angle θ in the tangent plane.

5.10 Theorem Let $\mathcal{M}$ have Gaussian curvature $K = -1$, and let $\mathbf{x}(u, v)$ a coordinate patch for $\mathcal{M}$ so that $I = E\ du^2 + G\ dv^2$. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 = \mathbf{n}\}$ be an orthonormal frame aligned with the asymptotic directions. Denote the frame and Cartan forms at $\hat{p}$ with hats. Consider the transformation $\hat{\mathbf{x}} = \mathbf{x} + \lambda \mathbf{e}_1$ along with rotation by an angle σ around the e_1 axis. Then $\hat{K} = -1$ if and only if $\lambda = \sin\sigma$.
Proof: A rotation by an angle σ around bde_1 leaves the tangent vector e_1 and its dual form θ^1) fixed. The rotation has a matrix representation as shown below.

$$\begin{bmatrix} \hat{\mathbf{e}_1} \\ \hat{\mathbf{e}_2} \\ \hat{\mathbf{e}_3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\sigma & -\sin\sigma \\ 0 & \sin\sigma & \cos\sigma \end{bmatrix} \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix}. \tag{5.19}$$

We compute the Cartan frame equations. We have: The coframe forms are $\theta^1 = \sqrt{E}\ du$, $\theta^2 = \sqrt{G}\ dv$ and $\theta^3 = 0$ on $\mathcal{M}$. Thus

$$\begin{aligned} d\mathbf{x} &= \mathbf{x}_u\ du + \mathbf{x}_v\ dv, \\ &= \frac{\mathbf{x}_u}{\sqrt{E}} \sqrt{e}\ du + \frac{\mathbf{x}_v}{\sqrt{G}} \sqrt{e}\ dv, \\ &= \frac{\mathbf{x}_u}{\sqrt{E}}\ \theta^1 + \frac{\mathbf{x}_v}{\sqrt{G}}\ \theta^2, \\ &= \mathbf{e}_1\theta^1 + \mathbf{e}_2\theta^2. \end{aligned}$$

We now compute $d\hat{\mathbf{x}}$ taking into account the rotation and then the translation:

$$\begin{aligned} d\hat{\mathbf{x}} &= \hat{\mathbf{e}}_1\hat{\theta}^1 + \hat{\mathbf{e}}_2\hat{\theta}^2, \\ &= \mathbf{e}_1\theta^1 + [\cos\sigma\ \mathbf{e}_2 - \sin\sigma \mathbf{e}_3]\ \hat{\theta}^2, \\ d\hat{\mathbf{x}} &= d\mathbf{x} + \lambda d\mathbf{e}_1, \\ &= \mathbf{e}_1\theta^1 + \mathbf{e}_2\theta^2 + \lambda(\mathbf{e}_2\,\omega^2{}_1 + \mathbf{e}_3\,\omega^3{}_1) \end{aligned}$$

Equating the coefficients of $\mathbf{e}_2$ and $\mathbf{e}_3$ in two equations above, we get

$$\begin{aligned} \cos\sigma\,\hat{\theta}^2 &= \theta^2 + \lambda\omega^2{}_1, \\ -\sin\sigma\,\hat{\theta}^2 &= \lambda\,\omega^3{}_1. \end{aligned} \tag{5.20}$$

Recall from 4.96 that $\omega^1{}_3 = l\theta^1 + m\theta^2$ and $\omega^2{}_3 = m\theta^1 + n\theta^2$ yields symmetric matrix components of the second fundamental form in the given basis. Using this fact and wedging with $\hat{\theta}^1$ the second equation above, we get

$$\begin{aligned} -\sin\sigma\,\hat{\theta}^1 \wedge \hat{\theta}^2 &= \lambda\,\hat{\theta}^1 \wedge \omega^3{}_1, \\ &= \lambda\,\theta^1 \wedge (l\,\theta^1 + m\,\theta^2), \\ \hat{\theta}^1 \wedge \hat{\theta}^2 &= -\frac{\lambda m}{\sin\sigma}\theta^1 \wedge \theta^2. \end{aligned} \tag{5.21}$$

Multiplying the first equation in 5.20 by $\sin\sigma$, the second by $\cos\sigma$, and adding, we get

$$\sin\sigma[\theta^2 + \lambda\,\omega^2{}_1] = -\lambda\ \cos\sigma\,\omega^3{}_1, \tag{5.22}$$

$$\theta^2 = -\frac{\lambda}{\sin\sigma}\ [\sin\sigma\,\omega^2{}_1 + \cos\sigma\,\omega^3{}_1]. \tag{5.23}$$

Next, we compute $\hat{\omega}_{32}$:

$$\begin{aligned} \hat{\omega}_{32}(X) &= \langle \overline{\nabla}_X \hat{e}_2, \hat{e}_3 \rangle, \\ &= \langle \cos\sigma \overline{\nabla}_X e_3 - \sin\sigma \overline{\nabla}_X e_3, \sin\sigma\ e_2 + \cos\sigma\ e_3 \rangle, \\ &= \cos^2\sigma \langle \overline{\nabla}_X e_2, e_3 \rangle - \sin^2\sigma \langle \overline{\nabla}_X e_3, e_2 \rangle, \\ &= (\cos^2\sigma + \sin^2\sigma) \langle \overline{\nabla}_X e_2, e_3 \rangle, \\ &= \omega_{32}(X), \\ \hat{\omega}^3_{\ 2} &= \omega^3{}_2 = -m\ \theta^1 - n\ \theta^2 \end{aligned}$$

By the same process, we calculate $\hat{\omega}_{31}$:

$$\begin{aligned} \hat{\omega}_{31}(X) &= \langle \overline{\nabla}_X \hat{e}_1, \hat{e}_3 \rangle, \\ &= \langle \overline{\nabla}_X e^3 + \overline{\nabla}_X e^3, \sin\sigma\ e^2 + \cos\sigma\ e^3 \rangle, \\ &= \langle \omega^2{}_1(X) e_2 + \omega^3{}_1(X) e_3, \sin\sigma\ e_2 + \cos\sigma\ e_3 \rangle, \\ &= \sin\sigma \omega_{21}(X) + \sin\sigma \omega_{31}(X), \\ \hat{\omega}^3_{\ 1} &= \frac{\sin\sigma}{\lambda}\theta^2. \end{aligned}$$

Finally, putting these results together, we get

$$\begin{aligned} d\hat{\omega}^1_{\ 2} &= \hat{\omega}^1_{\ 3} \wedge \hat{\omega}^3_{\ 2}, \\ &= -\left[\frac{m\sin\sigma}{\lambda}\right]\ \theta^1 \wedge \theta^2, \\ &= \left[\frac{m\sin\sigma}{\lambda}\right]\left[\frac{\sin\sigma}{m\ \lambda}\right]\hat{\theta}^1 \wedge \hat{\theta}^2, \\ &= \left[\frac{\sin\sigma}{\lambda}\right]^2\ \hat{\theta}^1 \wedge \hat{\theta}^2. \end{aligned}$$

Hence

$$K = -\left[\frac{\sin\sigma}{\lambda}\right]^2 = -1$$

if and only if $\lambda = \pm\sin\sigma$. We choose $\lambda = \sin\sigma$

The conclusion of the theorem explains the condition in the definition of a BT that requires this equation to hold. With this condition, equation 5.22 takes the form:

$$\omega^1{}_2 = \cot\sigma\,\omega^3{}_1 + \csc\sigma\,\theta^2. \tag{5.24}$$

We move to stage two of the BT process.

5.11 Theorem Let $\mathcal{M}$ be a pseudospherical surface with first fundmamental form as in 5.13, and let $F : \mathcal{M} \to \hat{\mathcal{M}}$ be a BT with angle of inclination σ. If the segment $\overline{p\hat{p}}$ makes a constant angle α with the basis vector e_1 at each point $p \in \mathcal{M}$, then

$$\begin{aligned} \sin\sigma(\theta_u + \omega_v) &= \quad\sin\theta\cos\omega - \cos\sigma\cos\theta\sin\omega, \\ \sin\sigma(\theta_v + \omega_u) &= -\cos\theta\sin\omega + \cos\sigma\sin\theta\cos\omega. \end{aligned} \tag{5.25}$$

Proof: Suppose $\overline{p\hat{p}}$ makes an angle θ with the tangent vector e_1. In this case we first perform a rotation of axis around the normal vector e_3 to align the frame with e_1. The rotation can be represented by a matrix $\overline{e}_i = e_j\,B^j{}_i$

$$B = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.26}$$

The effect on the Cartan frame is much easier establish since all we have to do is apply the change of basis formula 3.40 as shown in 3.42.

$$\begin{aligned} \overline{\omega}^1{}_2 &= \omega^1{}_2 - d\theta, \\ \overline{\omega}^1{}_3 &= \cos\theta\,\omega^1{}_3 + \sin\theta\,\omega^2{}_3, \\ \overline{\omega}^2{}_3 &= -\sin\theta\,\omega^1{}_3 + \cos\theta\,\omega^2{}_3. \end{aligned} \tag{5.27}$$

The dual forms transform with $A = B^{-1} = B^T$, that is $\overline{\theta}^i = A^i{}_j\theta^j$. In particular,

$$\overline{\theta}^2 = -\sin\theta\,\theta^1 + \cos\theta\,\theta^2 \tag{5.28}$$

If we start with a pseudospherical surface with $I = \cos^2\omega\;du^2 + \sin^2\omega\;dv^2$, the Cartan forms are:

$$\begin{aligned} \theta^1 &= \cos\omega\;du, \quad \theta^2 = \sin\omega\;dv, \\ \omega^1{}_3 &= \sin\omega\;du, \quad \omega^2{}_3 = -\cos\omega\;dv \\ \omega^1{}_2 &= -\omega_v\;du - \omega_u\;dv. \end{aligned}$$

The BT trasformation is the composition of the stage one and stage two. This means that we must subject the Cartan forms to the change of basis by a rotation by θ, as described by equations 5.27 and 5.28, followed by substitution into equation 5.24. We get:

$$(\omega^1{}_2 - d\theta) = \cot\sigma\;(\cos\theta\,\omega^1{}_3 + \sin\theta\,\omega^2{}_3) + \csc\sigma(-\sin\theta\,\theta^1 + \cos\theta\,\theta^2).$$

We extract two formulas obtained by equating the coefficients of du and dv respectively.

$$
\begin{aligned}
-\omega_v - \theta_u &= \quad \cot\sigma\cos\theta\sin\omega - \csc\sigma\sin\theta\cos\omega,\\
-\omega_u - \theta_v &= -\cot\sigma\sin\theta\cos\omega + \csc\sigma\cos\theta\sin\omega.
\end{aligned}
$$

The theorem follows by multiplying these equations by $\sin\sigma$, and rearranging terms. The system of equations 5.25 is the classical **Bäcklund Transform**. In the special case in which $\sigma = \pi/2$, the angle between the normals e_3 and $\hat{e}_3$ is a right angle, so $\hat{e}_3$ is parallel to a tangent vector of $\mathcal{M}$. This is called a **Bianchi Transform**. Equations 5.25 the reduce to the much simpler system:

$$
\begin{aligned}
\theta_u + \omega_v &= \quad \sin\theta\cos\omega,\\
\theta_v + \omega_u &= -\cos\theta\sin\omega. \qquad (5.29)
\end{aligned}
$$

We can rewrite the BT-equations in the so called asymptotic coordinates. Let

$$
\begin{aligned}
u &= x + t,\\
v &= x - t,
\end{aligned}
\qquad \text{so that} \qquad
\begin{aligned}
x &= \frac{1}{2}(u+v),\\
t &= \frac{1}{2}(u-v).
\end{aligned}
$$

By the chain rule, we have:

$$
\begin{aligned}
\theta_u &= \frac{1}{2}(\theta_x + \theta_t),\\
\theta_v &= \frac{1}{2}(\theta_x - \theta_t),
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
\omega_u &= \frac{1}{2}(\omega_x + \omega_t),\\
\omega_v &= \frac{1}{2}(\theta_x - \theta_t).
\end{aligned}
$$

Adding and subtraction equations 5.25, the system reduces to

$$
\begin{aligned}
\theta_x + \omega_x &= \frac{1+\cos\sigma}{\sin\sigma}\sin(\theta - \omega),\\
\theta_t - \omega_t &= \frac{1-\cos\sigma}{\sin\sigma}\sin(\theta + \omega),
\end{aligned}
$$

which we can rewrite as

$$
\begin{aligned}
\theta_t &= \quad \omega_t + s\sin(\theta + \omega),\\
\theta_x &= -\omega_x + \frac{1}{s}\sin(\theta - \omega), \qquad (5.30)\\
& \qquad (5.31)
\end{aligned}
$$

where $s = \tan(\sigma/2.)$ As we have stated, a BT trasform $F = F(\omega, \theta, \sigma)$ with angle σ, allows one to create a new solution of the Sine-Gordon equations from a given solution. Of course, the process can be iterated, neat thing is that further iterations can be carried out algebraically. This remarkable result is encoded into the

5.12 Theorem (Bianchi Permutability) Suppose that $\sin^2\theta_1 \neq \sin^2\theta_2$, and ϕ is a solution of the BT systems

$$
\begin{aligned}
\phi_t &= \quad (\theta_1)_t + s_1\sin(\phi + \theta_1)\\
\phi_x &= -(\theta_1)_x + \frac{1}{s_1}\sin(\phi - \theta_1)
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
\phi_t &= \quad (\theta_2)_t + s_2\sin(\phi + \theta_2)\\
\phi_x &= -(\theta_2)_x + \frac{1}{s_2}\sin(\phi - \theta_2)
\end{aligned}
$$

Then,

$$\tan\left(\frac{\phi-\omega}{2}\right)=\frac{s_1+s_2}{s_1-s_2}\tan\left(\frac{\theta_1-\theta_2}{2}\right) \tag{5.32}$$

Thus, if one begins with the trivial solution $\omega=0$ of the Sine-Gordon equation 5.18, solving the BT equations gives the one-soliton solution

$$\theta(x,t)=2\tan^{-1}(e^{sx+\frac{1}{s}t}),$$

The permutability theorem then gives the two-soliton solution:

$$\phi=2\tan^{-1}\left[\frac{s_1+s_2}{s_1-s_2}\frac{e^{s_1x+\frac{1}{s_1}t}-e^{s_2x+\frac{1}{s_2}t}}{1+e^{(s_1+s_2)x+(\frac{1}{s_1}+\frac{1}{s_2})t}}\right].$$

It is easy to write coordinate patch equations for a BT. The vector $\hat{\mathbf{x}}-\mathbf{x}$ must be a vector of length $\sin\sigma$ which is tangent to $\mathcal{M}$ at each point p and makes and angle θ with e_1. Therefore, we must have

$$\hat{\mathbf{x}}-\mathbf{x}=\sin\sigma[\cos\theta\; e_1+\sin\theta\; e_2].$$

But, $e_1=\mathbf{x}_u/\sqrt{E}$ and $e_2=\mathbf{x}_v/\sqrt{G}$, so we have:

$$\hat{\mathbf{x}}=\mathbf{x}+\sin\sigma\left[\frac{\cos\theta}{\cos\omega}\mathbf{x}_u+\frac{\sin\theta}{\sin\omega}\mathbf{x}_v\right]. \tag{5.33}$$

5.13 Example Kuen Surface

In this example, we perform a Bianchi transformation of a pseudosphere to obtain a Kuen surface. We begin with the parametrization of a pseudosphere given by 5.9 with $a=1$

$$\mathbf{x}(u,v)=\langle\operatorname{sech}\mu\cos v,\operatorname{sech}\mu\sin v,(\mu-\tanh\mu)\rangle.$$

Let $\mu=\ln\tan(\omega/2)$. Then

$$\begin{aligned}\omega &= 2\tan^{-1}(e^{\mu}),\\ \sin\omega &= \operatorname{sech}\mu,\\ \cos\omega &= -\tanh\mu.\end{aligned}$$

We will find θ by solving the Bianchi equations 5.29. We compute:

$$\begin{aligned}\omega_\mu &= \frac{2e^{\mu}}{1+e^{2\mu}}=\frac{2}{e^{\mu}+e^{\mu}}=\operatorname{sech}\mu,\\ \omega_v &= 0.\end{aligned}$$

Substituting into the Bianchi equations, we get:

$$\begin{cases}\theta_\mu=-\sin\theta\tanh\mu,\\ \theta_v=-\cos\theta\operatorname{sech}\mu-\operatorname{sech}\mu=-\operatorname{sech}\mu(1+\cos\theta)=-2\cos^2(\frac{\theta}{2})\operatorname{sech}\mu.\end{cases}$$

Separate variables

$$\begin{cases}\csc\theta\;\theta_\mu=-\tanh\mu,\\ \frac{1}{2}\sec^2(\frac{\theta}{2})\;\theta_v=-\operatorname{sech},\end{cases}$$

and integrate. The result is:

$$\begin{cases}\tan(\frac{\theta}{2})=-h_1(v)\operatorname{sech}\mu,\\ \tan(\frac{\theta}{2})=-v\operatorname{sech}\mu+h_2(\mu),\end{cases}$$

where h_1 and h_2 are the arbitrary functions of integration. Consistency of the equations requires $h_1 = 1$ and $h_2 = 0$. The solution is therefore

$$\begin{aligned}\tan(\tfrac{\theta}{2}) &= -v \operatorname{sech}\mu = \frac{-v}{\cosh\mu},\\ \theta &= 2\tan^{-1}(-v\operatorname{sech}\mu).\end{aligned}$$

Only the cosine and the sine of the angle θ enter into the Bianchi coordinate patch. Thinking of $\tan(\frac{\theta}{2})$ as the ratio of the opposite over the adjacent side of the righr triangle with hypothenuse $\sqrt{\cosh^2\mu + v^2}$, we can compute these from the double angle formulas:

$$\begin{aligned}\cos\theta &= \cos^2(\tfrac{\theta}{2}) - \sin^2(\tfrac{\theta}{2}) = \frac{\cosh^2\mu - v^2}{\cosh^2\mu + v^2},\\ \sin\theta &= 2\sin(\tfrac{\theta}{2})\cos(\tfrac{\theta}{2}) = \frac{-2v\cosh\mu}{\cosh^2\mu + v^2}.\end{aligned}$$

It remains to go through the algebraic gymnastics of computing the coordinate patch:

$$\begin{aligned}\hat{\mathbf{x}} &= \mathbf{x} + \frac{\cos\theta}{\cos\omega}\mathbf{x}_u + \frac{\sin\theta}{\sin\omega}\mathbf{x}_v,\\ &= \langle\, \operatorname{sech}\mu\cos v, \operatorname{sech}\mu\sin v, (\mu - \tanh\mu)\,\rangle\\ &\quad - \frac{\cos\theta}{\tanh\mu}\langle\, -\operatorname{sech}\mu\tanh\mu\cos v, -\operatorname{sech}\tanh\mu\sin v, \tanh^2\mu\,\rangle\\ &\quad + \frac{\sin\theta}{\operatorname{sech}\mu}\langle\, -\operatorname{sech}\mu\sin v, \operatorname{sech}\mu\cos v, 0\,\rangle,\\ &= \langle\, \operatorname{sech}\mu\cos v, \operatorname{sech}\mu\sin v, (\mu - \tanh\mu)\,\rangle\\ &\quad + \cos\theta\langle\, \operatorname{sech}\mu\cos v, \operatorname{sech}\sin v, -\tanh\mu\,\rangle\\ &\quad + \sin\theta\langle\, -\sin v, \cos v, 0\,\rangle,\end{aligned}$$

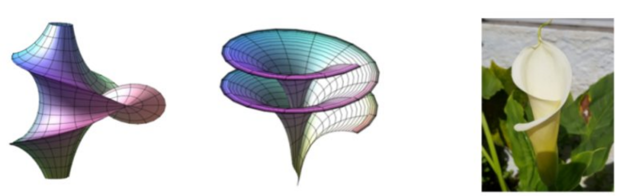

b) Kuen b) Dini c) Calla Lily

Fig. 5.5: Surfaces with K= -1

The x-component of $\mathbf{x}$ is:

$$
\begin{aligned}
\mathbf{x}_{(x)} &= (1+\cos\theta)\,\mathrm{sech}\,\mu\cos v - \sin\theta\sin v \\
&= \left[1+\frac{\cosh^2\mu - v^2}{\cosh^2\mu + v^2}\right]\cos v + \left[\frac{2v\cosh\mu}{\cosh^2\mu+v^2}\right]\sin v, \\
&= \left[\frac{2\cosh^2\mu}{\cosh^2\mu+v^2}\right]\cos v + \left[\frac{2v\cosh\mu}{\cosh^2\mu+v^2}\right]\sin v, \\
&= \frac{2\cosh\mu(\cos v + v\sin v)}{\cosh^2\mu+v^2}.
\end{aligned}
$$

The computation of the other two components is left as an exercise. The result is

$$
\mathbf{x}(u,v) = \left\langle \frac{2\cosh\mu(\cos v - v\sin v)}{\cosh^2\mu+v^2}, \frac{2\cosh\mu(\sin v - v\cos v)}{\cosh^2\mu+v^2}, \mu - \frac{2\sinh 2\mu}{\cosh^2\mu+v^2} \right\rangle \tag{5.34}
$$

The Kuen surface in figure 5.5 is plotted with parameters $u \in [-1.4, 1.4]$ and $v \in [-4, 4]$.

Another well known surface that can be easily obtained by BT that results in 'twisting' the pseudosphere in a helicoidal manner. We will not discuss the details for now, but instead we simply quote a coordinate patch in the form

$$
\mathbf{x}(u,v) = \langle\, a\cos u\cos v, a\cos u\sin v, a(\cos u + \ln\tan(\frac{x}{2})) + bv\,\rangle. \tag{5.35}
$$

This surface has curvature $K=-1$ when $a^2+b^2=1$. This is called the bf Dini surface and it has an unfolding infundibular shape as shown in figure 5.5, with parameters $u \in [0,2],\ v \in [0,4\pi]$ and $a=1,\ b=0.2$.

Chapter 6

Riemannian Geometry

6.1 Riemannian Manifolds

In sections 4.1 and 4.2 we defined a Riemannian manifold to be a manifold with a metric. A **Riemannian submanifold** is subset that is also a Riemannian manifold. The most natural example is that of a surface in $\mathbf{R}^n$. If $(x^1, x^2 \ldots x^n)$ are local coordinates in $\mathbf{R}^n$ with the standard metric, and the surface $\mathscr{M}$ is defined locally by functions $x^i = x^i u^\alpha$, then $\mathscr{M}$ together with the induced first fundamental form, has a canonical Riemannian structure. We will continue to use the notation $\overline{\nabla}$ for a connection in the ambient space and ∇ for the connection on the surface induced by the tangential component of the covariant derivative

$$\overline{\nabla}_X Y = \nabla_X Y + H(X, Y), \tag{6.1}$$

where $H(X, yY)$ is the component in the normal space. In the case of a hypersurface, we have the classical Gauss equation 4.71

$$\begin{aligned} \overline{\nabla}_X Y &= \nabla_X Y + II(X, Y)N && (6.2)\\ &= \nabla_x Y + \langle LX, Y\rangle N, && (6.3) \end{aligned}$$

where $LX = -\overline{\nabla}_X N$ is the Weingarten map. If $\mathscr{M}$ is a submanifold of codimension $n-k$, then there are k normal vectors N_k and k classical second fundamental forms $II_k(X, Y)$, so that $H(X, Y) = \sum_k II_k(X, Y)N_k$.

As shown by the Teorema Egregium, the curvature of a surface in $\mathbf{R}^3$ depends only on the first fundamental form, so the definition of Gaussian curvature as the determinant of the second fundamental form does not even make sense intrinsically. One could redefine K by Cartan's second structure equation as it was used to compute curvatures in Chapter 4, but what we need is a more general definition that is applicable to any Riemannian manifold. The concept leading to the equations of the Teorema Egregium involved calculation of the difference of second derivatives of tangent vectors. At the risk of being somewhat misleading, figure 4.80 illustrates the concept. In this figure, the vector field X consists of unit vectors tangent to parallels on the sphere, and the vector field Y are unit tangents to meridians. If an arbitrary tangent vector Z is parallel-transported by from one point on an spherical triangle to the diagonally opposed point, the result depends on the path taken. Parallel transport of Z along X followed by Y, would yield a different outcome that parallel transport along Y followed by parallel transport along X. The failure of the covariant derivatives to commute is reflection of the existence of curvature. clearly, the analogous parallel transport by two different paths on a rectangle in $\mathbf{R}^n$ yield the same result. This fact is the reason why in elementary calculus, vectors depend on direction and length. As indicated, the picture is misleading, because, covariant derivatives, as is the case with any other type of derivative, involves

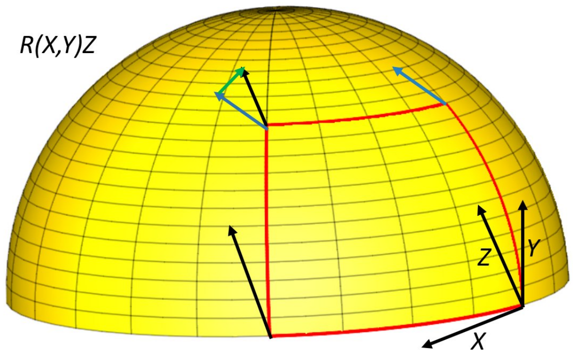


Fig. 6.1: R(X,Y)Z

comparing the change of a vector under infinitesimal parallel transport. Still, the figure should help motivate the definition that follows.

6.1 Definition On a Riemannian manifold with connection ∇, the curvature R and the torsion T are defined by:

$$R(X,Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]} \tag{6.4}$$

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]. \tag{6.5}$$

6.2 Theorem The Curvature R is a tensor. At each point $p \in \mathscr{M}$, $R(X,Y)$ assigns to each pair of tangent vectors, a linear transformation from $T_p\mathscr{M}$ into itself.

Proof: Let $X, Y, Z \in \mathscr{X}$ be vector fields on $\mathscr{M}$. We need to establish that R is muiltilinear over the ring $\mathscr{M}$. Since clearly $R(X,Y) = -R(Y,X)$, we only need to establish linearity on two slots. Let f be a C^∞ function. Then,

$$\begin{aligned}
R(fX,Y) &= \nabla_{fX}\nabla_Y Z - \nabla_Y \nabla_{fX} Z - \nabla_{[fX,Y]}Z,\\
&= f\nabla_X\nabla_Y Z - \nabla_Y(f\nabla_X)Z - \nabla_{[fXY - Y(fX)]}Z,\\
&= f\nabla_X\nabla_Y Z - Y(f)\nabla_X)Z - f\nabla_Y\nabla_X Z - \nabla_{fXY}Z + \nabla_{(Y(f)X + fYX)}Z,\\
&= f\nabla_X\nabla_Y Z - Y(f)\nabla_X)Z - f\nabla_Y\nabla_X Z - f\nabla_{XY}Z + \nabla_{Y(f)X}Z + \nabla_{fYX}Z,\\
&= f\nabla_X\nabla_Y Z - Y(f)\nabla_X)Z - f\nabla_Y\nabla_X Z - f\nabla_{XY}Z + Y(f)\nabla_X Z + f\nabla_{YX}Z,\\
&= f\nabla_X\nabla_Y Z - f\nabla_Y\nabla_X Z - f(\nabla_{XY}Z - \nabla_{YX})Z,\\
&= fR(X,Y)Z.
\end{aligned}$$

Similarly, recalling that $[X,Y] \in \mathscr{X}$, we get:

$$\begin{aligned}
R(X,Y)(fZ) &= \nabla_X\nabla_Y(fZ) - \nabla_Y\nabla_X(fZ) - \nabla_{[X,Y]}(fZ),\\
&= \nabla_X(Y(f)Z) + f\nabla_Y Z) - \nabla_Y(X(f)Z + f\nabla_X Z) - [X,Y](f)Z) - f\nabla_{[X,Y]}Z,\\
&= XY(f)Z) + Y(f)\nabla_X Z + X(f)\nabla_Y Z + f\nabla_X\nabla_Y Z -\\
&\quad\; YX(f)Z) - X(f)\nabla_Y Z - Y(f)\nabla_X Z - f\nabla_Y\nabla_X Z -\\
&\quad\; [X,Y](f)Z) - f\nabla_{[X,Y]}(Z),\\
&= fR(X,Y)Z.
\end{aligned}$$

We leave it as an almost trivial exercise to check linearity over addition in all slots.

6.3 Theorem The torsion T is also a tensor.

Proof: Since $T(X,Y) = -T(Y,X)$, it suffices to prove linearity on one slot. Thus,

$$\begin{aligned} T(fX,Y) &= \nabla fXY - \nabla_Y(fX) - [fX,Y], \\ &= f\nabla_X Y - Y(f)X - f\nabla_Y X - fXY + Y(fX), \\ &= f\nabla_X Y - Y(f)X - f\nabla_Y X - fXY + Y(f)X + fYX, \\ &= f\nabla_X Y - f\nabla Y X - f[X,Y], \\ &= fT(X,Y). \end{aligned}$$

Again, linearity over sums is clear.

6.4 Theorem In a Riemannian manifold there exist a unique torsion free connection called the **Levi-Civita connection**, that is compatible with the metric. That is:

$$[X,Y] = \nabla_X Y - \nabla_Y X, \tag{6.6}$$

$$X\langle Y,Z\rangle = \nabla_X\langle Y,Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle. \tag{6.7}$$

Proof: The proof parallels the computation leading to equation 4.73 taking cyclic derivatives of the inner product.

$$\begin{aligned} \nabla_X\langle Y,Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, \\ \nabla_Z\langle X,Y\rangle &= \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle, \\ \nabla_Y\langle Z,X\rangle &= \langle \nabla_Y Z, X\rangle + \langle Z, \nabla_Y X\rangle, \\ \nabla_X\langle Y,Z\rangle - \nabla_Z\langle X,Y\rangle - \nabla_Y\langle Z,X &= \langle [X,Y],Z\rangle + \langle [X,Z],Y\rangle - \langle \nabla_Y Z + \nabla_Z Y, Z\rangle, \\ &= \langle [X,Y],Z\rangle + \langle [X,Z],Y\rangle + \langle [Y,Z],X\rangle - 2\langle \nabla_Y Z, X\rangle, \end{aligned}$$

Therefore:

$$\langle \nabla_Y Z, X\rangle = \tfrac{1}{2}\{\langle [X,Y],Z\rangle + \langle [X,Z],Y\rangle + \langle [Y,Z],X\rangle - X\langle Y,Z\rangle - Z\langle X,Y\rangle - Y\langle Z,X\rangle\}. \tag{6.8}$$

Since the inner product is non-degenerate and this equation is true for arbitrary vector fields, it determines $\nabla_Y Z$ completely in terms of the metric. In disguise, this is the formula for the Christoffel symbols 4.73

As before, if $\{e_\alpha\}$ is a frame with dual frame $\{\theta^a\}$, we define the connection forms ω, Christoffel symbols Γ and Torsion components by

$$\nabla_X e_\beta = \omega^\gamma{}_\beta(X)\, e_\gamma, \tag{6.9}$$

$$\nabla_{e_\alpha} e_\beta = \Gamma^\gamma_{\alpha\beta}\, e_\gamma, \tag{6.10}$$

$$T(e_\alpha, e_\beta) = T^\gamma{}_{\alpha\beta}\, e_\gamma. \tag{6.11}$$

As was pointed out in the previous chapter, If the frame is an orthonormal coordinate frame so that the bracket is zero, then $T = 0$ implies that the Christoffel symbols are symmetric in the lower indices.

$$T^\gamma{}_{\alpha\beta} = \Gamma^\gamma_{\alpha\beta} - \Gamma^\gamma_{\beta\alpha} = 0.$$

For such a coordinate frame, we can compute the components of the Riemann tensors as follows:

$$\begin{aligned} R(e_\gamma, e_\delta)\, e_\beta &= \nabla_{e_\gamma}\nabla_{e_\delta} e_\beta - \nabla_{e_\delta}\nabla_{e_\gamma} e_\beta, \\ &= \nabla_{e_\gamma}(\Gamma^\alpha_{\delta\beta} e_\alpha) - \nabla_{e_\delta}(\Gamma^\alpha_{\gamma\beta} e_\alpha), \\ &= \Gamma^\alpha_{\delta\beta,\gamma} e_\alpha + \Gamma^\alpha_{\delta\beta}\Gamma^\mu_{\gamma\alpha} e_\mu - \Gamma^\alpha_{\gamma\beta,\delta} e_\alpha - \Gamma^\alpha_{\gamma\beta}\Gamma^\mu_{\delta\alpha} e_\mu, \\ &= [\Gamma^\alpha_{\beta\delta,\gamma} - \Gamma^\alpha_{\beta\gamma,\delta} + \Gamma^\mu_{\beta\delta}\Gamma^\alpha_{\gamma\mu} - \Gamma^\mu_{\beta\gamma}\Gamma^\alpha_{\delta\mu}] e_\alpha, \\ &= R^\alpha{}_{\beta\gamma\delta}\, e_\alpha, \end{aligned}$$

where the components of the Riemann Tensor are defined by:

$$R^{\alpha}{}_{\beta\gamma\delta} = \Gamma^{\alpha}_{\beta\delta,\gamma} - \Gamma^{\alpha}_{\beta\gamma,\delta} + \Gamma^{\mu}_{\beta\delta}\Gamma^{\alpha}_{\gamma\mu} - \Gamma^{\mu}_{\beta\gamma}\Gamma^{\alpha}_{\delta\mu}. \tag{6.12}$$

The generalization of the Teorema Egregium to manifolds comes from the same principle of splitting the curvature tensor of the ambient space into the tangential on normal components. In the case of a hypersurface with normal N and tangent vectors X, Y, Z, we have:

$$\begin{aligned}
\overline{R}(X,Y)Z &= \overline{\nabla}_X\overline{\nabla}_Y Z - \overline{\nabla}_Y\overline{\nabla}_X Z - \overline{\nabla}_{[X,Y]}Z,\\
&= \overline{\nabla}_X(\nabla_y Z + \langle LY, Z\rangle N) - \overline{\nabla}_(\nabla_X Z + \langle LX, Z\rangle N) - \overline{\nabla}_{[X,Y]}Z,\\
&= \nabla_X\nabla_Y Z + \langle LX, \nabla_Y Z\rangle N + \nabla_X\langle LY, Z\rangle N + \langle LY, Z\rangle LX -\\
&\quad \nabla_Y\nabla_X Z - \langle LY, \nabla_Y Z\rangle N - \nabla_Y\langle LX, Z\rangle N - \langle LX, Z\rangle LY -\\
&\quad \nabla_{[X,Y]}Z - \langle L([X,Y]), Z\rangle N,\\
&= \nabla_X\nabla_Y Z + \langle LX, \nabla_Y Z\rangle N + \nabla_X\langle LY, Z\rangle N + \langle LY, Z\rangle LX -\\
&\quad \nabla_Y\nabla_X Z - \langle LY, \nabla_Y Z\rangle N - \nabla_Y\langle LX, Z\rangle N - \langle LX, Z\rangle LY -\\
&\quad \nabla_{[X,Y]}Z - \langle L([X,Y]), Z\rangle N,\\
&= \nabla_X\nabla_Y Z + \langle LX, \nabla_Y Z\rangle N + \langle \nabla_X LY, Z\rangle N + \langle LY, \nabla_X Z\rangle N + \langle LY, Z\rangle LX -\\
&\quad \nabla_Y\nabla_X Z - \langle LY, \nabla_Y Z\rangle N - \langle \nabla_Y LX, Z\rangle N - \langle LX, \nabla_Y Z\rangle N - \langle LX, Z\rangle LY -\\
&\quad \nabla_{[X,Y]}Z - \langle L([X,Y]), Z\rangle N,\\
&= R(X,Y)Z + \langle LY, Z\rangle LX - \langle LX, Z\rangle LY +\\
&\quad \{\langle \nabla_X LY, Z\rangle - \langle \nabla_Y LX, Z\rangle - \langle L([X,Y]), Z\rangle\}N.
\end{aligned}$$

If the ambient space is $\mathbf{R}^n$, the curvature tensor $\overline{R}$ is zero, so we can set the horizontal and normal components in the right to zero. Noting that the normal component is zero for all Z, we get:

$$\begin{aligned}
R(X,Y)Z + \langle LY, Z\rangle LX - \langle LX, Z\rangle LY &= 0,\\
\nabla_X LY - \nabla_Y LX - L([X,Y]) &= 0.
\end{aligned} \tag{6.13}$$

In particular, if $n = 3$, and at each point in the surface, the vectors X and Y constitute an a basis of the tangent space, we get the coordinate-free Teorema Egregium

$$K = \langle R(X,Y)X, Y\rangle = \langle LX, X\rangle\langle LY, Y\rangle - \langle LY, X\rangle\langle LX, Y\rangle = \det(L). \tag{6.14}$$

The expression 6.13 is the coordinate-independent version of the equation of Codazzi.

We expect the covariant definition of the torsion and curvature tensors to be consistent with the formalism of Cartan.

6.5 Theorem Equations of Structure.

$$\Theta^{\alpha} = d\theta^{\alpha} + \omega^{\alpha}{}_{\beta}\wedge\theta^{\beta}, \tag{6.15}$$

$$\Omega^{\alpha}{}_{\beta} = d\omega^{\alpha}{}_{\beta} + \omega^{\alpha}{}_{\gamma}\wedge\omega^{\gamma}{}_{\beta}. \tag{6.16}$$

To verify this is the case, we define:

$$T(X,Y) = \Theta^{\alpha}(X,Y)e_{\alpha}, \tag{6.17}$$

$$R(X,Y)e_{\beta} = \Omega^{\alpha}_{\beta}(X,Y)e_{\alpha}. \tag{6.18}$$

Recalling that any tangent vector X can be expressed in terms of the basis as $X = \theta^\alpha(X)\ e_\beta$, we can carry out a straight-forward computation:

$$\begin{aligned}
\Theta^\alpha(X,Y)\ e_\alpha &= T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y],\\
&= \nabla_X(\theta^\alpha(Y))\ e_\alpha - \nabla_Y(\theta^\alpha(X))\ e_\alpha - \theta^\alpha([X,Y])\ e_\alpha,\\
&= X(\theta^\alpha(Y))\ e_\alpha + \theta^\alpha(Y)\ \omega^\beta{}_\alpha(X)\ e_\beta - Y(\theta^\alpha(X))\ e_\alpha - \theta^\alpha(X)\ \omega^\beta{}_\alpha(Y)\ e_\beta - \theta^\alpha([X,Y])\ e_\alpha,\\
&= \{X(\theta^\alpha(Y)) - Y(\theta^\alpha(X)) - \theta^\alpha([X,Y]) + \omega^\alpha{}_\beta(X)(\theta^\beta(Y) - \omega^\alpha{}_\beta(Y)(\theta^\beta(X)\}e_\alpha,\\
&= \{(d\theta^\alpha + \omega^\alpha{}_\beta \wedge \theta^\beta)(X,Y)\}e_\alpha,
\end{aligned}$$

where we have introduced a coordinate-free definition of the differential of the one form θ by

$$d\theta(X,Y) = X(\theta(Y)) - y(\theta(X)) - \theta([X,Y]) \tag{6.19}$$

It is easy to verify that this definition of the differential of a one form satisfies all the required properties of the differential, and that it is consistent with the coordinate version of the differential introduced in Chapter 2. We conclude that

$$\Theta^\alpha = d\theta^\alpha + \omega^\alpha{}_\beta \wedge \theta^\beta, \tag{6.20}$$

which is indeed the first Cartan equation of structure.

Proceeding along the same lines, we compute:

$$\begin{aligned}
\Omega^\alpha{}_\beta(X,Y)\ e_\alpha &= \nabla_X\nabla_Y\ e_\beta - \nabla_Y\nabla_X\ e_\beta - \nabla_{[X,Y]}\ e_\beta,\\
&= \nabla_X(\omega^\alpha{}_\beta(Y)\ e_\alpha) - \nabla_Y(\omega^\alpha{}_\beta(X)\ e_\alpha) - \omega^\alpha{}_\beta([X,Y])\ e_\alpha,\\
&= X(\omega^\alpha{}_\beta(Y))\ e_\alpha + \omega^\alpha{}_\beta(Y)\ \omega^\gamma{}_\alpha(X)\ e_\gamma - Y(\omega^\alpha{}_\beta(X))\ e_\alpha - \omega^\alpha{}_\beta(X)\ \omega^\gamma{}_\alpha(Y)\ e_\gamma - \omega^\alpha{}_\beta([X,Y])\ e_\alpha\\
&= \{(d\omega^\alpha{}_\beta + \omega^\alpha{}_\gamma \wedge \omega^\gamma{}_\beta)(X,Y)\}e_\alpha,
\end{aligned}$$

thus arriving at the second equation of structure

$$\Omega^\alpha{}_\beta = d\omega^\alpha{}_\beta + \omega^\alpha{}_\gamma \wedge \omega^\gamma{}_\beta. \tag{6.21}$$

The relation between the full Torsion and Riemann tensor components with the corresponding differential forms are given by:

$$\Theta^\alpha = \tfrac{1}{2}T^\alpha{}_{\gamma\delta}\,\theta^\gamma \wedge \theta^\delta, \tag{6.22}$$
$$\Omega^\alpha{}_\beta = \tfrac{1}{2}R^\alpha{}_{\beta\gamma\delta}\,\theta^\gamma \wedge \theta^\delta. \tag{6.23}$$

In the case of a non-coordinate frame in which the Lie bracket of frame vectors does not vanish, we first write them as linear combinations of the frame

$$[e_\beta, e_\gamma] = C^\alpha_{\beta\gamma}\ e_\alpha. \tag{6.24}$$

The components of the Torsion and Riemann tensors are then given by:

$$T^\alpha{}_{\beta\gamma} = \Gamma^\alpha_{\beta\gamma} - \Gamma^\alpha_{\gamma\beta} - C^\alpha_{\beta\gamma} \tag{6.25}$$
$$R^\alpha{}_{\beta\gamma\delta} = \Gamma^\alpha_{\beta\delta,\gamma} - \Gamma^\alpha_{\beta\gamma,\delta} + \Gamma^\mu_{\beta\delta}\Gamma^\alpha_{\gamma\mu} - \Gamma^\mu_{\beta\gamma}\Gamma^\alpha_{\delta\mu} - \Gamma^\alpha_{\beta\mu}C^\mu_{\gamma\delta}. \tag{6.26}$$

The Riemann tensor has the following symmetries;

$$R(X,Y) = -R(Y,X), \tag{6.27}$$
$$\langle R(X,Y)Z, W\rangle = -\langle R(X,Y)W, Z\rangle \tag{6.28}$$
$$R(X,Y)Z + R(Z,X)Y + R(Y,Z)X = 0. \tag{6.29}$$

The last cyclic equation is called the first Bianchi Identity, and we will derive it later. In terms of components, the Riemann Tensor symmetries can be expressed as:

$$R_{\alpha\beta\gamma\delta} = -R_{\alpha\beta\delta\gamma} = -R_{\beta\alpha\gamma\delta}, \tag{6.30}$$

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}, \tag{6.31}$$

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\delta\beta\gamma} = 0 \tag{6.32}$$

The symmetries reduce the number of independent components in an n-dimensional manifold from n^4 to $n^2(n^2-1)/12$. Thus for 4-dimensional space, there are at most 20 independent components.

The formalism above refers to Riemannian manifolds, for which the metric is positive definite, but it applies just as well to pseudo-Riemannian manifolds for which is not. A 4-dimensional manifold $\{\mathscr{M}, g\}$ is called a **Lorentzian manifold** if the metric has signature $(+---)$. Locally, a Lorentzian manifold is diffeomorphic to Minkowski's space which is the model space introduced in section 2.2. Some authors use signature $(-+++)$.

For the purposes of general relativity we introduce the symmetric tensor **Ricci Tensor** by the the contraction

$$R_{\beta\delta} = R^{\alpha}{}_{\beta\alpha\delta}, \tag{6.33}$$

$R_{\beta\gamma}$ and the **scalar curvature** R

$$R = R^{\alpha}{}_{\beta}. \tag{6.34}$$

The traceless part of the Ricci tensor

$$G_{\alpha\beta} = R_{\alpha\beta} - \tfrac{1}{2}Rg_{\alpha\beta}, \tag{6.35}$$

is called **Einstein's tensor.** The Einstein field equations (without a cosmological constant) are

$$G_{\alpha\beta} = \frac{8\pi G}{c^4}T_{\alpha\beta}, \tag{6.36}$$

where T is the **stress energy tensor** and G is the gravitational constant. As I first learned from one of my professors Arthur Fischer, the equation states that curvature indicates the presence of matter, and maatter tells the space how to curve.

Fig. 6.2: Gravity

A space time which satisfies

$$R_{\alpha\beta} = 0 \tag{6.37}$$

is called Ricci-Flat.

6.6 Example: Vaidya Metric

This example of a curvature computation in four dimensional space-time is due to W. Israel. It appears in his 1978 notes on Differential Forms in General Relativity, but the author indicates the work arose 10 years earlier from a seminar at the Dublin Institute for Advanced Studies. The Vaidya metric

$$ds^2 = 2drdu + [1 - \tfrac{2m(u)}{r}]\, du^2 - r^2 d\theta^2 - r^2 \sin^2\theta\, d\phi^2, \tag{6.38}$$

where $m(u)$ is an arbitrary function. The geometry described by the Vaidya solution to Einstein equations, represents the gravitational field in the exterior of a radiating, spherically symmetric star. The variable u is the retarded time and $m(u)$ is the mass. If $m(u)$ isa constant, this is equivalent to a Schwarzschild metric. The computation will be carried out using Cartan's structure equations. The first step in the computation involves picking out a basis of one forms. The idea is to pick out the forms so that in the new basis, the metric has constant coefficients. One possible choice of 1-forms is

$$\begin{aligned}
\theta^0 &= du,\\
\theta^1 &= dr + \tfrac{1}{2}[1 - \tfrac{2m(u)}{r}]\, du\\
\theta^2 &= r\, d\theta\\
\theta^3 &= r \sin\theta\, d\phi.
\end{aligned} \tag{6.39}$$

In terms of these forms, the line element becomes

$$ds^2 = g_{\alpha\beta}\theta^\alpha\theta^\beta = 2\theta^0\theta^1 - (\theta^2)^2 - (\theta^3)^2.$$

where

$$g_{01} = g_{10} = -g_{22} = -g_{33} = 1,$$

while all the other $g_{\alpha\beta} = 0$. In all our previous Gaussian curvature computations the metric has been diagonal. This is an instructive example of non-diagonal metric which in this frame the metric and its inverse have components:

$$g_{\alpha\beta} = \begin{bmatrix} 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{6.40}$$

Since the coefficients of the metric are constant, the components $\omega_{\alpha\beta}$ of the connection will be antisymmetric. This means that

$$\omega_{00} = \omega_{11} = \omega_{22} = \omega_{33} = 0.$$

We thus conclude that

$$\begin{aligned}
{\omega^1}_0 &= g^{10}\omega_{00} = 0\\
{\omega^0}_1 &= g^{01}\omega_{11} = 0\\
{\omega^2}_2 &= g^{22}\omega_{22} = 0\\
{\omega^3}_3 &= g^{33}\omega_{33} = 0.
\end{aligned}$$

To compute the connection, we take the exterior derivative of the basis 1-forms. The result of this computation is

$$\begin{aligned}
d\theta^0 &= 0\\
d\theta^1 &= -d[\tfrac{m}{r} du] = \frac{m}{r^2}\, dr \wedge du = \frac{m}{r^2}\, \theta^1 \wedge \theta^0\\
d\theta^2 &= dr \wedge d\theta = \frac{1}{r}\theta^1 \wedge \theta^2 - \frac{1}{2r}[1 - \tfrac{2m}{r}]\, \theta^0 \wedge \theta^2\\
d\theta^3 &= \sin\theta\, dr \wedge d\phi + r\cos\theta\, d\theta \wedge d\phi\\
&= \frac{1}{r}\theta^1 \wedge \theta^3 - \frac{1}{2}[1 - \tfrac{2m}{r}]\, \theta^0 \wedge \theta^3 + \frac{1}{r}\cot\theta\, \theta^2 \wedge \theta^3
\end{aligned} \tag{6.41}$$

For convenience, we write below the first equation of structure [6.15] in complete detail.

$$\begin{aligned}
d\theta^0 &= \omega^0{}_0 \wedge \theta^0 + \omega^0{}_1 \wedge \theta^1 + \omega^0{}_2 \wedge \theta^2 + \omega^0{}_3 \wedge \theta^3, \\
d\theta^1 &= \omega^1{}_0 \wedge \theta^0 + \omega^1{}_1 \wedge \theta^1 + \omega^1{}_2 \wedge \theta^2 + \omega^1{}_3 \wedge \theta^3, \\
d\theta^2 &= \omega^2{}_0 \wedge \theta^0 + \omega^2{}_1 \wedge \theta^1 + \omega^2{}_2 \wedge \theta^2 + \omega^2{}_3 \wedge \theta^3, \\
d\theta^3 &= \omega^3{}_0 \wedge \theta^0 + \omega^3{}_1 \wedge \theta^1 + \omega^3{}_2 \wedge \theta^2 + \omega^3{}_3 \wedge \theta^3. \qquad (6.42)
\end{aligned}$$

Since the the ω's are one forms, they must be linear combinations of the θ's. Comparing Cartan's first structural equation with the exterior derivatives of the coframe, we can start with the initial guess for the connection coefficients below:

$$\omega^1{}_0 = 0, \qquad \omega^1{}_1 = \frac{m}{r^2}\,\theta^0, \qquad \omega^1{}_2 = A\,\theta^2, \qquad \omega^1{}_3 = B\,\theta^3, \qquad (6.43)$$

$$\omega^2{}_0 = -\frac{1}{2}[1-\frac{2m}{r}]\,\theta^2, \qquad \omega^2{}_1 = \frac{1}{r}\,\theta^2, \qquad \omega^2{}_2 = 0, \qquad \omega^2{}_3 = C\,\theta^3, \qquad (6.44)$$

$$\omega^3{}_0 = -\frac{1}{2}[1-\frac{2m}{r}]\,\theta^3, \qquad \omega^3{}_1 = \frac{1}{r}\,\theta^3, \qquad \omega^3{}_2 = \frac{1}{r}\cot\theta\,\theta^3, \qquad \omega^3{}_3 = 0, \qquad (6.45)$$

Here, the quantities A, B,and C are unknowns to be determined. Observe that these are not the most general choices for the ω's. For example, we could have added a term proportional to θ^1 in the expression for $\omega^1{}_1$, without affecting the validity of the first structure equation for $d\theta^1$. The strategy is to interactively tweak the expressions until we we set of forms completely consistent with Cartan's structure equations.

We now take advantage of the skewsymmetry of $\omega_{\alpha\beta}$, to determine the other components. The find A, B and C, we note that

$$\begin{aligned}
\omega^1{}_2 &= g^{10}\omega_{02} = -\omega_{20} = \omega^2{}_0 \\
\omega^1{}_3 &= g^{10}\omega_{03} = -\omega_{30} = \omega^3{}_0. \\
\omega^2{}_3 &= g^{22}\omega_{23} = \omega_{32} = -\omega^3{}_2,
\end{aligned}$$

From equations [6.43], [6.44], and [6.45], we find that

$$A = -\frac{1}{2}[1-\frac{2m}{r}], \quad B = -\frac{1}{2}[1-\frac{2m}{r}].C = -\frac{1}{r}\cot\theta. \qquad (6.46)$$

Similarly, we have

$$\begin{aligned}
\omega^0{}_0 &= -\omega^1{}_1, \\
\omega^0{}_2 &= \omega^2{}_1, \\
\omega^0{}_3 &= \omega^3{}_1,
\end{aligned}$$

hence,

$$\begin{aligned}
\omega^0{}_0 &= -\frac{m}{r^2}\theta^0, \\
\omega^0{}_2 &= -\frac{1}{r}\theta^2, \\
\omega^0{}_3 &= \frac{1}{r}\theta^3.
\end{aligned}$$

By inspection we see that our choices for the ω's are consistent with first structure equations, so by uniqueness, these must be the right values.

There is no guesswork in obtaining the curvature forms. All we do is take the exterior derivative of the connection forms, and pick out the components of the curvature from the second Cartan equations [6.16]. Thus, for example, to obtain $\Omega^1{}_1$, we proceed as follows.

$$\begin{aligned}
\Omega^1{}_1 &= d\omega^1{}_1 + \omega^1{}_1 \wedge \omega^1{}_1 + \omega^1{}_2 \wedge \omega^2{}_1 + \omega^1{}_3 \wedge \omega^3{}_1, \\
&= d[\frac{m}{r^2}\theta^0] + 0 - \frac{1}{2r^2}[1 - \frac{2m}{r}]\,\omega^1{}_3 \wedge \omega^3{}_1 + (\theta^2 \wedge \theta^2 + \theta^3 \wedge \theta^3), \\
&= -\frac{2m}{r^3} dr \wedge \theta^0, \\
&= -\frac{2m}{r^3}\theta^1 \wedge \theta^0.
\end{aligned}$$

The computation of the other components is straight forward and we just present the results.

$$\begin{aligned}
\Omega^1{}_2 &= -\frac{1}{r^2}\frac{dm}{du}\theta^2 \wedge \theta^0 - \frac{m}{r^3}\theta^1 \wedge \theta^2, \\
\Omega^1{}_3 &= -\frac{1}{r^2}\frac{dm}{du}\theta^3 \wedge \theta^0 - \frac{m}{r^3}\theta^1 \wedge \theta^3, \\
\Omega^2{}_1 &= \frac{m}{r^3}\theta^2 \wedge \theta^0, \\
\Omega^3{}_1 &= \frac{m}{r^3}\theta^3 \wedge \theta^0, \\
\Omega^2{}_3 &= \frac{2m}{r^3}\theta^2 \wedge \theta^3.
\end{aligned}$$

The other components of the curvature are determined by observing that like the connection forms, the curvature forms are also antisymmetric. We can also read the components of the full Riemann curvature tensor from the definition

$$\Omega^\alpha{}_\beta = \frac{1}{2}R^\alpha{}_{\beta\gamma\delta}\theta^\gamma \wedge \theta^\delta. \tag{6.47}$$

Thus, for example, we have

$$\Omega^1{}_1 = \frac{1}{2}R^1{}_{1\gamma\delta}\theta^\gamma \wedge \theta^\delta,$$

hence

$$R^1{}_{101} = -R^1{}_{110} = \frac{2m}{r^3}; \text{ other } \; R^1{}_{1\gamma\delta} = 0.$$

Using the antisymmetry of the curvature forms, we see, that for the Vaidya metric $\Omega^1{}_0 = \Omega_{00} = 0$, $\Omega^2{}_0 = -\Omega^1{}_2$, etc., so that

$$\begin{aligned}
R_{00} &= R^2{}_{020} + R^3{}_{030} \\
&= R^1{}_{220} + R^1{}_{330}
\end{aligned}$$

Substituting the relevant components of the curvature tensor read from the Ω's we find that

$$R_{00} = 2\frac{1}{r^2}\frac{dm}{du} \tag{6.48}$$

while all the other components of the Ricci tensor vanish. As stated earlier, if m is constant, we get the Ricci flat Schwarzschild metric.

6.2 Geodesics

6.7 Definition Let $u^\alpha(t)$ be a curve on a surface $\mathbf{x} = \mathbf{x}(u^\alpha)$, and let $V = \alpha'(t) = \alpha_*(\frac{d}{dt})$ be the velocity vector as defined in 1.22 and illustrated in figure 1.6. A vector field Y is called **parallel** along α if $\nabla_V Y = 0$. The vector field $\nabla_V V$ is called the **Geodesic Vector Field**, and its magnitude is called the geodesic curvature κ_g of α.

As usual, we define the speed v of the curve by $\|V\|$ and the unit tangent $T = V/\|V\|$, so that $V = vT$. We assume $v > 0$ so that T is defined on the domain of the curve. The arc length s along the curve is the related the speed by the equation $v = ds/dt$. or

6.8 Definition A curve $\alpha(t)$ with velocity vector $V = \alpha'(t)$ is called a **geodesic** if $\nabla_V V = 0$.

6.9 Theorem A curve $\alpha(t)$ is geodesic iff a) $v = \|V\|$ is constant along the curve and either 1) $\nabla_T T = 0$, or 2) $\kappa_g = 0$.

Proof: Expanding the definition of the geodesic vector field:

$$\begin{aligned}
\nabla_V V &= \nabla_{vT}(vT),\\
&= v\nabla_T(vT),\\
&= v\frac{dv}{dt}T + v^2\nabla_T T,\\
&= \tfrac{1}{2}\frac{d}{dt}(v^2)T + v^2\nabla_T T
\end{aligned}$$

We have $\langle T,T\rangle = 1$, so $2\langle \nabla_T T, T\rangle = 0$ which shows that $\nabla_T T$ is orthogonal to T. We also have $v > 0$. Since both the tangential and the normal components need to vanish, the theorem follows.

If $\mathscr{M}$ is a hypersurface in $\mathbf{R}^n$ with unit normal $\mathbf{n}$, we gain more insight on the geometry of geodesics as a direct consequence of the discussion above. Without real loss of generality consider the geometry in the case of $n = 3$. Since $\|\alpha'\|^2 = \langle \alpha', \alpha'\rangle$ =constant, differentiation gives $\langle \alpha', \alpha''\rangle = 0$, so that the acceleration α'' is orthogonal to α'. Comparing with equation 4.33 we see that $T' = \kappa_n \mathbf{n}$, which reinforces the fact that the entire curvature of the curve is due to the normal curvature of the surface as a submanifold of the ambient space. In this sense, inhabitants constrained to live on the surface would be unaware of this curvature, and to them, geodesics would appear locally as the straightest path to travel. Thus, for a sphere in $\mathbf{R}^3$ of radius a centered a the origin, the acceleration α'' of a geodesic only has a normal component, and the normal curvature is $1/a$. That is, the geodesic must lie along a great circle.

6.10 Theorem Let $\alpha(t)$ by curve with velocity V. For each vector Y in the tangent space restricted to the curve, there is a unique vector field $Y(t)$ locally obtained by parallel transport.

Proof: We choose local coordinates with frame field $\{e_\alpha = \frac{\partial}{\partial u^\alpha}\}$. We write the components of the vector fields in terms of the frame

$$\begin{aligned}
Y &= a^\beta \frac{\partial}{\partial u^\beta}, \\
V &= \frac{du^\alpha}{dt}\frac{\partial}{\partial u^\alpha}. \quad \text{then,} \\
\nabla_T V &= \nabla_{\dot{u^\alpha} e_\alpha}(a^\beta e_\beta), \\
&= \dot{u^\alpha}\nabla_{e^\alpha}(a^\beta e_\beta), \\
&= \frac{du^\alpha}{dt}\frac{\partial a^\beta}{\partial u^\alpha} + \dot{u^\alpha} a^\beta \Gamma^\gamma{}_{\alpha\beta} e_\gamma, \\
&= [\frac{da^\gamma}{dt} + a^\beta \frac{du^\alpha}{dt}\Gamma^\gamma{}_{\alpha\beta}]e_\gamma.
\end{aligned}$$

So, Y is parallel along the curve iff

$$\frac{da^\gamma}{dt} + a^\beta \frac{du^\alpha}{dt}\Gamma^\gamma{}_{\alpha\beta} = 0 \tag{6.49}$$

The existence and uniqueness of the coefficients a^β that define Y are guaranteed by the theorem on existence and uniqueness of differential equations with appropriate initial conditions.

At the risk of sounding a bit repetitions, we derive the equations of geodesics by an almost identical computation.

$$\begin{aligned}
\nabla_V V &= \nabla_{\dot{u}^\alpha e_\alpha}[\dot{u^\beta} e_\beta], \\
&= \dot{u}^\alpha \nabla_{e_\alpha}[\dot{u}^\beta e_\beta], \\
&= \dot{u}^\alpha [\frac{\partial \dot{u}^\beta}{\partial u^\alpha} e_\beta + \dot{u}^\beta \nabla_{e_\alpha} e_\beta], \\
&= \dot{u}^\alpha \frac{\partial \dot{u}^\beta}{\partial u^\alpha} e_\beta + \dot{u}^\alpha \dot{u}^\beta \nabla_{e_\alpha} e_\beta, \\
&= \frac{du^\alpha}{dt}\frac{\partial \dot{u}^\beta}{\partial u^\alpha} e_\beta + \dot{u}^\alpha \dot{u}^\beta \Gamma^\sigma_{\alpha\beta} e_\sigma, \\
&= \ddot{u}^\beta e_\beta + \dot{u}^\alpha \dot{u}^\beta \Gamma^\sigma_{\alpha\beta} e_\sigma, \\
&= [\ddot{u}^\sigma + \dot{u}^\alpha \dot{u}^\beta \Gamma^\sigma_{\alpha\beta}] e_\sigma.
\end{aligned}$$

Thus, the equation for geodesics becomes:

$$\ddot{u}^\sigma + \Gamma^\sigma_{\alpha\beta}\dot{u}^\alpha \dot{u}^\beta = 0. \tag{6.50}$$

The existence and uniqueness theorem for solutions of differential equations leads to the following theorem

6.11 Theorem Let p be a point in $\mathscr{M}$ and V a vector $T_p\mathscr{M}$. Then, for any real number t_0, there exists a number δ and a curve $\alpha(t)$ defined on $[t_0 - \delta, t_0 + \delta]$, such that $\alpha(t_0) = p$, $\alpha'(t_0) = V$, and α is a geodesic.

Whereas it is possible to compute all the Christoffel symbols starting with the metric as in equation 4.73, this is most inefficient, as it is often the case that many of the Christoffel symbols vanish. Instead, we show next how to obtain the geodesic equations by using variational principles

$$\delta \int L(u^\alpha, \dot{u}^\alpha, s)\, ds = 0, \tag{6.51}$$

to minimize the arc length. Following the standard methods of Lagrangian Mechanics, we let u^α and $\dot{u}^\alpha$ be treated as independent (canonical) coordinates and choose the Langrangian in this case to be

$$L = g_{\alpha\beta}\dot{u}^\alpha \dot{u}^\beta. \tag{6.52}$$

The choice will actually result in minimizing the square of the arclength, but clearly this is an equivalent problem. It should be observed that the lagrangian is basically a multiple of the kinetic energy $\frac{1}{2}mv^2$. The motion dynamics are given by the Euler-Lagrange equations.

$$\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{u}^\gamma}\right) - \frac{\partial L}{\partial u^\gamma} = 0. \tag{6.53}$$

Applying this equations keeping in mind that $g_{\alpha\beta}$ is the only quantity that depends on u^α, we get:

$$\begin{aligned}
0 &= \tfrac{d}{ds}[g_{\alpha\beta}\delta^\alpha_\gamma \dot{u}^\beta + g_{\alpha\beta}\dot{u}^\alpha \delta^\beta_\gamma] - g_{\alpha\beta,\gamma}\dot{u}^\alpha \dot{u}^\beta \\
&= \tfrac{d}{ds}[g_{\gamma\beta}\dot{u}^\beta + g_{\alpha\gamma}\dot{u}^\alpha] - g_{\alpha\beta,\gamma}\dot{u}^\alpha \dot{u}^\beta \\
&= g_{\gamma\beta}\ddot{u}^\beta + g_{\alpha\gamma}\ddot{u}^\alpha + g_{\gamma\beta,\alpha}\dot{u}^\alpha \dot{u}^\beta + g_{\alpha\gamma,\beta}\dot{u}^\beta \dot{u}^\alpha - g_{\alpha\beta,\gamma}\dot{u}^\alpha \dot{u}^\beta \\
&= 2g_{\gamma\beta}\ddot{u}^\beta + [g_{\gamma\beta,\alpha} + g_{\alpha\gamma,\beta} - g_{\alpha\beta,\gamma}]\dot{u}^\alpha \dot{u}^\beta \\
&= \delta^\sigma_\beta \ddot{u}^\beta + \tfrac{1}{2}g^{\gamma\sigma}[g_{\gamma\beta,\alpha} + g_{\alpha\gamma,\beta} - g_{\alpha\beta,\gamma}]\dot{u}^\alpha \dot{u}^\beta
\end{aligned}$$

where the last equation was obtained contracting with $\frac{1}{2}g^{\gamma\sigma}$ to raise indices. Comparing with the expression for the Christoffel symbols found in equation REF, we get

$$\ddot{u}^\sigma + \Gamma^\sigma_{\alpha\beta}\dot{u}^\alpha \dot{u}^\beta = 0$$

which is exactly the equations of geodesics 6.50.

6.12 Example Geodesics of Sphere

Let S^2 be a sphere of radius a so that the metric is given by

$$ds^2 = a^2 d\theta^2 + a^2 \sin^2\theta \, d\phi^2.$$

Then the Lagrangian is

$$L = a^2\dot{\theta}^2 + a^2\sin^2\theta \, \dot{\phi}^2.$$

The Euler-Lagrange equation for the ϕ coordinate is

$$\begin{aligned}
\frac{d}{ds}(\frac{\partial L}{\partial \dot{\phi}}) - \frac{\partial L}{\partial \phi} &= 0 \\
\frac{d}{ds}(2a^2\sin^2\theta\dot{\phi}) &= 0
\end{aligned}$$

and therefore the equation integrates to a constant

$$\sin^2\theta \, \dot{\phi} = k.$$

Rather than trying to solve the second Euler-Lagrange equation for θ, we evoke a standard trick that involves reusing the metric. It goes as follows:

$$
\begin{aligned}
\sin^2\theta\,\frac{d\phi}{ds} &= k,\\
\sin^2\theta\; d\phi &= k\; ds,\\
\sin^4\theta\; d\phi^2 &= k^2 ds^2,\\
\sin^4\theta\; d\phi^2 &= k^2(a^2 d\theta^2 + a^2\sin^2\theta\; d\phi^2),\\
(\sin^4\theta - k^2a^2\sin^2\theta)\, d\phi^2 &= a^2k^2 d\theta^2.
\end{aligned}
$$

The last equation above is separable and it can be integrated using the substitution $u = \cot\theta$.

$$
\begin{aligned}
d\phi &= \frac{ak}{\sin\theta\sqrt{\sin^2\theta - a^2k^2}}\, d\theta,\\
&= \frac{ak}{\sin^2\theta\sqrt{1 - a^2k^2\csc^2\theta}}\, d\theta,\\
&= \frac{ak}{\sin^2\theta\sqrt{1 - a^2k^2(1+\cot^2\theta)}}\, d\theta,\\
&= \frac{ak\csc^2\theta}{\sqrt{1 - a^2k^2(1+\cot^2\theta)}}\, d\theta,\\
&= \frac{ak\csc^2\theta}{\sqrt{(1 - a^2k^2) - a^2k^2\cot^2\theta}}\, d\theta,\\
&= \frac{\csc^2\theta}{\sqrt{\frac{1-a^2k^2}{a^2k^2} - \cot^2\theta}}\, d\theta,\\
&= \frac{-1}{\sqrt{c^2 - u^2}}\, du, \quad \text{where(} c^2 = \tfrac{1-a^2k^2}{a^2k^2}).\\
\phi &= -\sin^{-1}(\tfrac{1}{c}\cot\theta) + \phi_0.
\end{aligned}
$$

Here, ϕ_0 is the constant of integration. To get a geometrical sense of the geodesics equations we have just derived, we rewrite the equations as follows:

$$
\begin{aligned}
\cot\theta &= c\sin(\phi_0 - \phi),\\
\cos\theta &= c\sin\theta(\sin\phi_0\cos\phi - \cos\phi_0\sin\phi),\\
a\cos\theta &= (c\sin\phi_0)(a\sin\theta\cos\phi) - (c\cos\phi_0)(a\sin\theta\sin\phi.)\\
z &= Ax - By, \quad \text{where } A = c\sin\phi_0,\ B = c\cos\phi_0.
\end{aligned}
$$

We conclude that the geodesics of the sphere are great circles determined by the intersections with planes through the origin.

6.13 Example Geodesics of Surface of Revolution
The first fundamental form a surface of revolution $z = f(r)$ in cylindrical coordinates as in 4.7, is

$$
ds^2 = (1 + f'^2)\, dr^2 + r^2\, d\phi^2, \tag{6.54}
$$

so the Lagrangian becomes:

$$
L = (1 + f'^2)\,\dot r^2 + r^2\dot\phi^2.
$$

Since there is no dependance on ϕ, the Euler-Lagrange equation on ϕ gives rise to a conserved quantity.

$$\frac{d}{ds}(2r^2\dot{\phi}) = 0,$$
$$r^2\dot{\phi} = c \tag{6.55}$$

where c is a constant of integration. If the geodesic $\alpha(s) = \alpha(r(s), \phi(s))$ represents the path of a free particle constrained to move in the surface, this conserved quantity is essentially the angular momentum. A neat result can be obtained by considering the angle σ that the tangent vector $V = \alpha'$ makes with a meridian. Recall that the length of V along the geodesic is constant, so let's set $\|V\| = k$ From the chain rule we have

$$\alpha'(t) = \mathbf{x}_r\frac{dr}{ds} + \mathbf{x}_\phi\frac{d\phi}{ds}.$$

Then

$$\begin{aligned}\cos\sigma &= \frac{\langle\alpha', \mathbf{x}_\phi\rangle}{\|\alpha'\|\cdot\|\mathbf{x}_\phi\|} = \frac{G\frac{d\phi}{ds}}{k\sqrt{G}},\\ &= \frac{1}{k}\sqrt{G}\frac{d\phi}{ds} = \frac{1}{k}r\dot{\phi}.\end{aligned}$$

We conclude from 6.55, that for a surface of revolution, the geodesics make an angle σ with meridians that satisfies the equation

$$r\cos\sigma = \text{constant}. \tag{6.56}$$

This result is called **Clairaut's relation**. Writing equation 6.55 in terms of differentials, and reusing the metric as we did in the computation of the geodesics for a sphere, we get:

$$\begin{aligned}r^2\ d\phi &= c\ ds,\\ r^4\ d\phi^2 &= c^2\ ds^2,\\ &= c^2[(1+f'^2)\ dr^2 + r^2\ d\phi^2],\\ (r^4 - c^2r^2)\ d\phi^2 &= c^2[(1+f'^2)\ dr^2,\\ r\sqrt{r^2-c^2}\ d\phi &= c\sqrt{1+f'^2}\ dr,\end{aligned}$$

So

$$\phi = \pm c\int\frac{\sqrt{1+f'^2}}{r\sqrt{r^2-c^2}}\ dr. \tag{6.57}$$

If $c = 0$, then the first equation above gives ϕ =constant, so the meridians are geodesics. The parallels r =constant are geodesics when $f'(r) = \infty$ in which case the tangent bundle restricted to the parallel is a cylinder with a vertical generator.

In the particular case of a cone of revolution with a generator that makes an angle α with the z-axis, $f(r) = \cot(\alpha)r$, equation 6.57 becomes:

$$\phi = \pm c\int\frac{\sqrt{1+\cot^2\alpha}}{r\sqrt{r^2-c^2}}\ dr$$

which can be immediately integrated to yield:

$$\phi = \pm\csc\alpha\sec^{-1}(r/c) \tag{6.58}$$

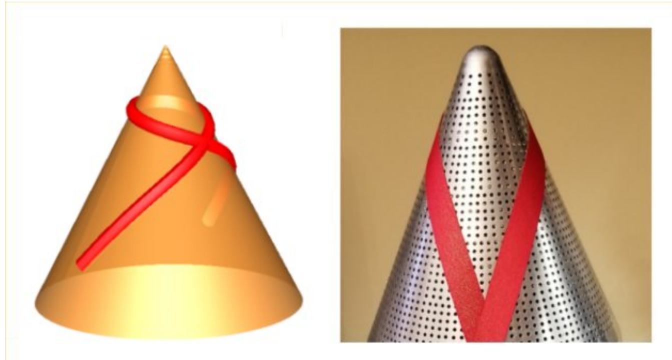


Fig. 6.3: Geodesics on a Cone.

As shown in figure 6.3, a ribbon laid flatly around a cone follows the path of a geodesic. None of the parallels, which in this case are the generators of the cone, are geodesics.

6.14 Example Morris-Thorne (MT) Wormhole
In 1987, Michael Morris and Kip Thorne from the California Institute of Technology proposed a tantalizing simple model for teaching general relativity by alluding to interspace travel in a geometry of traversible wormhole. We constraint the discussion purely to geometrical aspects of the model and not the physics of stress and strains of a "traveler" traversing the wormhole. The MT metric for this spherically symmetric geometry is:

$$ds^2 = -c^2 dt^2 + dl^2 + (b_0^2 + l^2)\,(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{6.59}$$

where b_0 is a constant. The obvious choice for a coframe is

$$\theta^0 = c\,dt \qquad \theta^2 = \sqrt{b_o^2 + l^2}\,d\theta$$
$$\theta^1 = dl \qquad \theta^3 = \sqrt{b_o^2 + l^2}\,\sin\theta\,d\phi.$$

We have $d\theta^0 = d\theta^1 = 0$. To find the connection forms we compute $d\theta^2$ and $d\theta^3$, and rewrite in terms of the coframe. We get

$$\begin{aligned}
d\theta^2 &= \frac{l}{\sqrt{b_o^2 + l^2}}\,dl \wedge d\theta = -\frac{l}{\sqrt{b_o^2 + l^2}}\,d\theta \wedge dl,\\
&= -\frac{l}{b_o^2 + l^2}\,\theta^2 \wedge \theta^1,\\
d\theta^3 &= \frac{l}{\sqrt{b_o^2 + l^2}}\sin\theta\,dl \wedge d\phi + \cos\theta\sqrt{b_o^2 + l^2}d\theta \wedge d\phi,\\
&= -\frac{l}{b_o^2 + l^2}\,\theta^3 \wedge \theta^1 - \frac{\cot\theta}{\sqrt{b_o^2 + l^2}}\,\theta^3 \wedge \theta^2.
\end{aligned}$$

Comparing with the first equation of structure, we start with simplest guess for the connection forms ω's. We set

$$\begin{aligned}
\omega^2{}_1 &= \frac{l}{b_o^2 + l^2}\,\theta^2,\\
\omega^3{}_1 &= \frac{l}{b_o^2 + l^2}\,\theta^3,\\
\omega^3{}_2 &= \frac{\cot\theta}{\sqrt{b_o^2 + l^2}}\,\theta^3.
\end{aligned}$$

Using the antisymmetry of the ω's and the diagonal metric, we have $\omega^2{}_1 = -\omega^1{}_2$, $\omega^1{}_3 = -\omega^3{}_1$, and $\omega^2{}_3 = -\omega^3{}_2$. This choice of connection coefficients turns out to be completely compatible with the entire set of Cartan's first equation of structure, so, these are the connection forms, all other ω's are zero. We can then proceed to evaluate the curvature forms. There is no guess-work here. An straight-forward calculus computation which results in some pleasing cancellations, yields:

$$\begin{aligned}
\Omega^1{}_2 &= d\omega^1{}_2 + \omega^2{}_1 \wedge \omega^2{}_1 = -\frac{b_o^2}{(b_o^2 + l^2)^2}\,\theta^1 \wedge \theta^2, \\
\Omega^1{}_3 &= d\omega^1{}_3 + \omega^1{}_2 \wedge \omega^2{}_3 = -\frac{b_o^2}{(b_o^2 + l^2)^2}\,\theta^1 \wedge \theta^3, \\
\Omega^2{}_3 &= d\omega^2{}_3 + \omega^2{}_1 \wedge \omega^1{}_3 = \frac{b_o^2}{(b_o^2 + l^2)^2}\,\theta^2 \wedge \theta^3.
\end{aligned}$$

Thus, from equation 6.26, other than permutations of the indices, the only independent components of the Riemann tensor are

$$R_{2323} = -R_{1212} = R_{1313} = \frac{b_o^2}{(b_o^2 + l^2)^2},$$

and the only non-zero component of the Ricci tensor is

$$R_{11} = -2\frac{b_o^2}{(b_o^2 + l^2)^2}$$

Of course, this space is a 4-dimensional continuum, but since the space is spherically symmetric, we may get a good sense of the geometry by taking a slice with $\theta = \pi/2$ at a fixed value of time. The resulting metric ds_2 for the surface is

$$ds_2{}^2 = dl^2 + (b_o^2 + l^2)\,d\phi^2. \tag{6.60}$$

Let $r^2 = b_o^2 + l^2$. Then $dl^2 = (r^2/l^2)\,dr^2$ and the metric becomes

$$\begin{aligned}
ds_2{}^2 &= \frac{r^2}{r^2 - b_o^2}\,dr^2 + r^2\,d\phi^2, && (6.61)\\
&= \frac{1}{1 - \frac{b_o^2}{r^2}}\,dr^2 + r^2\,d\phi^2. && (6.62)
\end{aligned}$$

Comparing to 4.25 we recognize this to be a Catenoid of revolution, so the equations of geodesics are given by **??** with $f(r) = b_0 \cosh^{-1}(r/b_0)$. Substituting this value of f into the geodesic equation, we get

$$\phi = \pm c \int \frac{1}{\sqrt{r^2 - b_o^2}\sqrt{r^2 - c^2}} dr. \tag{6.63}$$

There are three cases. If $c = b_o$, the integral gives immediately $\phi = \pm(c/b_0)\tanh^{-1}(r/b_0)$. We consider the case $c > b_0$. The remaining case can be treated in a similar fashion. Let $r = c/\sin\beta$. Then $\sqrt{r^2 - c^2} = r\cos\beta$ and $dr = -r\cot\beta\,d\beta$, so, assuming the initial condition $\phi(0) = 0$, the the

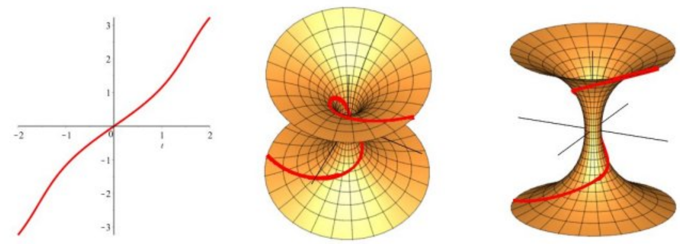


Fig. 6.4: Geodesics on Catenoid.

substitution leads to the integral

$$
\begin{aligned}
\phi &= \pm c \int_0^s \frac{1}{r\cos\beta\sqrt{\frac{c^2}{\sin^2\beta} - b_o^2}} \frac{(-r\cos\beta)}{\sin\beta}\, d\beta, \\
&= \pm c \int_0^s \frac{1}{\sqrt{c^2 - b_o^2 \sin^2\beta}}\, d\beta, \\
&= \pm \int_0^s \frac{1}{\sqrt{1 - k^2\sin^2\beta}}\, d\beta, \quad (k = b_o/c) \qquad (6.64)\\
&= F(s,k), \qquad (6.65)
\end{aligned}
$$

where $F(s,k)$ is the well known incomplete elliptic integral of the first kind. Elliptic integrals are standard functions implemented in computer algebra systems, so it is easy to render some geodesics as shown in figure 6.4. The plot of the elliptic integral shown here is for $k = 0.9$. The plot shows clearly that this is a 1-1, so if one one wishes to express r in terms of ϕ one just finds the inverse of the elliptic integral which yields a Jacobi elliptic function. Thomas Muller has created a neat Wolfram-Demonstration that allows the user to play with MT wormhole geodesics with parameters controlled by sliders.