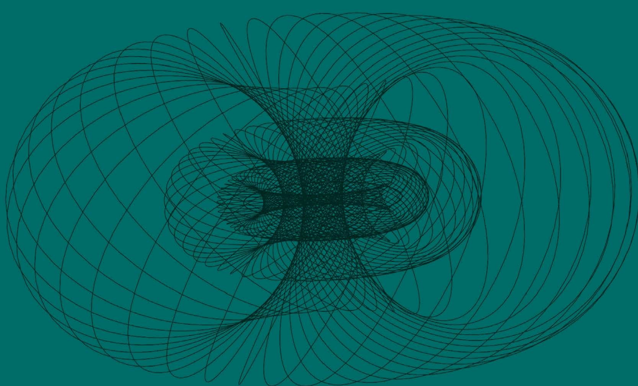
Gabriel Lugo

Differential Geometry in Physics



Second Edition

Draft

Suggested citation: Lugo, Gabriel. Differential Geometry in Physics. Wilmington. University of North Carolina Wilmington William Madison Randall Library, 2021.
doi: https: //doi.org/10.5149/9781469669267_Lugo

ISBN 978-1-4696-6924-3 (cloth: alk. paper)
ISBN 978-1-4696-6925-0 (paperback: alk. paper)
ISBN 978-1-4696-6926-7 (open access ebook)

Cover illustration: The Hopf fibration.

Published by the UNC Wilmington William Madison Randall Library

Distributed by UNC Press
www.uncpress.org

Publication of this book was supported by a grant from the Thomas W. Ross Fund from the University of North Carolina Press.

Escher detail on following page from "Belvedere" plate 230, [1]. All other figures and illustrations were produced by the author or taken from photographs taken by the author and family members. Exceptions are included with appropriate attribution.

This book is dedicated to my family, for without their love and support, this work would have not been possible. Most importantly, let us not forget that our greatest gift are our children and our children's children, who stand as a reminder, that curiosity and imagination is what drives our intellectual pursuits, detaching us from the mundane and confinement to Earthly values, and bringing us closer to the divine and the infinite.

G. Lugo (2021)

Contents

Preface

These notes were developed as part of a course on differential geometry which the author has taught for many years at UNCW. The first five chapters plus chapter six, constitute the foundation of the three-hour course. The course is cross-listed at the level of seniors and first year graduate students. In addition to applied mathematics majors, the class usually attracts a good cohort of double majors in mathematics and physics. Material from other chapters have inspired a number of honors and master level theses. This book 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 insight on very powerful tools in mathematical physics at the graduate level.

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:

1. Provide a bridge between the very practical formulation of classical differential geometry created by early masters of the late 1800's, and the more elegant but less intuitive modern formulation in terms of manifolds, bundles and differential forms. In particular, the central topic of curvature is presented in three different but equivalent formalisms.

2. Present the subject of differential geometry with an emphasis on making the material readable to physicists who may have encountered some of the concepts in the context of classical or quantum mechanics, but wish to strengthen the rigor of the mathematics. A source of inspiration for this goal is rooted in the shock to this author as a graduate student in the 70's at Berkeley, at observing the gasping failure of communications between the particle physicists working on gauge theories and differential geometers working on connection on fiber bundles. They seemed to be completely unaware at the time, that they were working on the same subject.

3. Make the material as readable as possible for those who stand at the boundary between theoretical physics and applied mathematics. For this reason, it will be occasionally necessary to sacrifice some mathematical rigor or depth of physics, in favor of ease of comprehension.

4. Provide the formal geometrical background for the mathematical theory of general relativity.

5. Introduce examples of other applications of differential geometry to physics that might not appear in traditional texts used in courses for mathematics students. For example, several students at UNCW have written masters' theses in the theory of solitons, but usually they have followed the path of Lie symmetries in the style of Olver. We hope that the elegance of Bäcklund transforms will attract students to a geometric approach to the subject. The book is also a stepping stone to other interconnected areas of mathematics such as representation theory, complex variables and algebraic topology.

G. Lugo (2021)

The main change in this second edition is the inclusion of exercises and projects suitable for a course using this textbook. The edition includes corrections of errors and misprints that seem to have a way of infiltrating on the first pass of most mathematics books. A list of known errors and misprints is found at the course web site,

http://people.uncw.edu/lugo/courses/DiffGeom/index.htm.

The author is grateful to any readers pointing out other needed corrections and welcomes suggestions for revisions that would improve the content.

G. Lugo (2022)

Chapter 1

Vectors and Curves

1.1 Tangent Vectors

1.1.1 Definition Euclidean n-space $\mathbf{R}^n$ is defined as the set of ordered n-tuples $p(p^1, \ldots, p^n)$, where $p^i \in \mathbf{R}$, for $i = 1, \ldots, n$. We may associate a position vector $\mathbf{p} = (p^1, \ldots, p^n)$ with any given point p in n-space. Given any two n-tuples $\mathbf{p} = (p^1, \ldots, p^n)$, $\mathbf{q} = (q^1, \ldots, q^n)$ and any real number c, we define two operations:

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

These two operations of vector sum and multiplication by a scalar satisfy all the 8 properties needed to give the set $V = \mathbf{R}^n$ a natural structure of a vector space. It is common to use the same notation $\mathbf{R}^n$ for the space of n-tuples and for the vector space of position vectors. Technically, we should write $p \in \mathbf{R}^n$ when we think of $\mathbf{R}^n$ as a metric space and $\mathbf{p} \in \mathbf{R}^n$ when we think of it as vector space, but as most authors, we will freely abuse the notation. [1]

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 $\mathbf{p} = (p^1, \ldots, p^n)$. The functions x^i are then called the natural *coordinate functions.* When convenient, we revert to the usual names for the coordinates, $x^1 = x$, $x^2 = y$ and $x^3 = z$ in $\mathbf{R}^3$. A small awkwardness might

[1] In these notes we will use the following index conventions:

- In $\mathbf{R}^n$, indices such as i, j, k, l, m, n, run from 1 to n.
- In space-time, indices such as μ, ν, ρ, σ, run from 0 to 3.
- On surfaces in $\mathbf{R}^3$, indices such as $\alpha, \beta, \gamma, \delta$, run from 1 to 2.
- Spinor indices such as $A, B, \dot{A}, \dot{B}$ run from 1 to 2.

occur in the transition to modern notation. In classical vector calculus, a point in $\mathbf{R}^n$ is often denoted by $\mathbf{x}$, in which case, we pick up the coordinates with the slot projection functions $u^i : \mathbf{R}^n \to \mathbf{R}$ defined by

$$u^i(\mathbf{x}) = x^i.$$

1.1.3 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 $\mathscr{F}(\mathbf{R}^n)$.

1.1.4 Definition Let V and V' be finite dimensional vector spaces such as $V = \mathbf{R}^k$ and $V' = \mathbf{R}^n$, and let $\mathbf{L}(V, V')$ be the space of linear transformations from V to V'. The set of linear functionals $\mathbf{L}(V, \mathbf{R})$ is called the *dual vector space* V^*. This space has the same dimension as V.

In calculus, vectors are usually regarded as arrows characterized by a direction and a length. Thus, vectors are 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 later chapters, 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.

1.1.5 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.

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 $T(\mathbf{R}^n)$ (or simply $T\mathbf{R}^n$) 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.

1.1.6 Definition A *vector field* X in $U \subset \mathbf{R}^n$ is a *section of the tangent bundle*, that is, a smooth function from U to $T(U)$. The space of sections $\Gamma(T(U)$ is also denoted by $\mathscr{X}(U)$.

The difference between a tangent vector and a vector field is that in the latter case, the coefficients v^i of $\mathbf{x}$ are smooth functions of x^i. Since in general

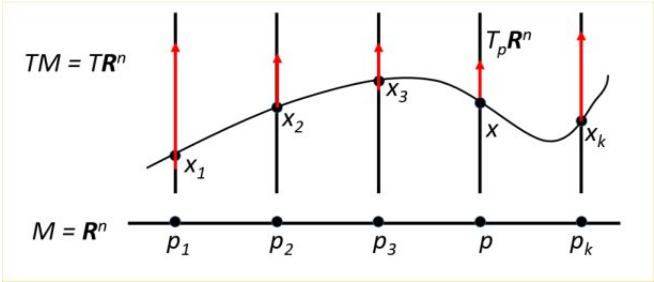


Fig. 1.1: Tangent Bundle

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 $\mathbf{p}_1, \dots, \mathbf{p}_k$ 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 amounts to assigning a tangent vector to every point in 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.

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

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

so that $\mathscr{X}(U)$ has the structure of a vector space over $\mathbf{R}$. The subscript notation X_p indicating the location of a tangent vector is sometimes cumbersome, but necessary to distinguish them from vector fields.

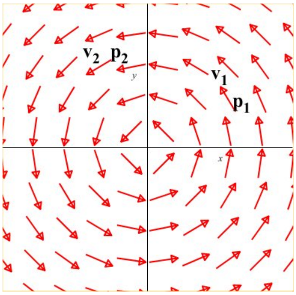


Fig. 1.2: Vector Field

Vector fields are essential objects in physical applications. If we consider the flow of a fluid in a region, the velocity vector field represents 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.

1.1.7 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 commonly used. 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 f, 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.

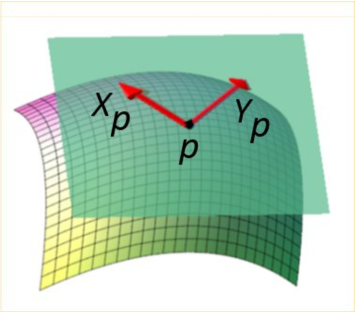


Fig. 1.3: Tangent vectors X_p, Y_p on a surface in $\mathbb{R}^3$.

1.1.8 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

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

1.1.9 Remark The space of smooth functions is a ring, ignoring a small technicality with domains. An operator such as a vector field with the properties above, is called a *linear derivation* on $\mathscr{F}(\mathbf{R}^n)$.

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 $(v^1, \ldots, v^n)$. 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

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

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 v^i of the vector. In elementary treatments, vectors are often identified with the components of the vector, and this may cause some confusion.

The operators

$$\{e_1, \ldots, e_k\}|_p = \left\{ \left(\frac{\partial}{\partial x^1}\right)_p , \ldots, \left(\frac{\partial}{\partial x^n}\right)_p \right\}$$

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 $X = v^i \frac{\partial}{\partial x^i}$ 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,

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

To re-emphasize, 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.

1.1.10 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 whose output are functions that at each point represent the directional derivative in the direction of the Euclidean vector.

1.1.11 Example Given the point $p(1,1)$, the Euclidean vector $\mathbf{x} = (3,4)$, 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,

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

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

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

1.1.13 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 $\alpha'(0) = X_p$, and, whenever $\alpha(t)$ is the domain of the vector field, $\alpha'(t) = X_{\alpha(t)}$.

In elementary calculus and differential equations, the families of integral curves of a vector field are called the streamlines, suggesting the trajectories of a fluid with velocity vector X. In figure 1.2, the integral curves would be circles that fit neatly along the *flow* of the vector field. In local coordinates, the expression defining integral curves of X constitutes a system of first order differential equations, so the existence and uniqueness of solutions apply locally. We will treat this in more detail in subsection **??**

1.2 Differentiable Maps

1.2.1 Definition Let $F : \mathbf{R}^n \to \mathbf{R}^m$ be a vector function defined by coordinate entries $F(\mathbf{p}) = (f^1(\mathbf{p}), f^2(\mathbf{p}), \ldots f^m(\mathbf{p}))$. 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 $(x^1, x^2, \ldots x^n)$ are local coordinates in $\mathbf{R}^n$ and $(y^1, y^2, \ldots y^m)$ local coordinates in $\mathbf{R}^m$, a map $\mathbf{y} = F(\mathbf{x})$ is represented in advanced calculus by m functions of n variables

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

A map $F : \mathbf{R}^n \to \mathbf{R}^m$ is differentiable at a point $\mathbf{p} \in \mathbf{R}^n$ if there exists a linear transformation $DF(\mathbf{p}) : \mathbf{R}^n \to \mathbf{R}^m$ such that

$$\lim_{\mathbf{h}\to 0} \frac{|F(\mathbf{p}+\mathbf{h}) - F(\mathbf{p}) - DF(\mathbf{p})(\mathbf{h})|}{|\mathbf{h}|} = 0 \tag{1.10}$$

The linear transformation $DF(\mathbf{p})$is called the *Jacobian*. A differentiable map that is invertible and the inverse is differentiable, is called a *diffeomorphism*.

Remarks

1. A differentiable mapping $F : I \in \mathbf{R} \to \mathbf{R}^n$ is what we called a curve. If $t \in I = [a, b]$, the mapping gives a parametrization $\mathbf{x}(t)$, as we discussed in the previous section.

2. A differentiable mapping $F : R \in \mathbf{R}^n \to \mathbf{R}^n$ is called a coordinate transformation. Thus, for example, the mapping $F : (u, v) \in \mathbf{R}^2 \to (x, y) \in$

$\mathbf{R}^2$, given by functions $x = x(u, v),\ y = y(u, v)$, would constitute a change of coordinates from (u, v) to (x, y). The most familiar case is the polar coordinates transformation $x = r\cos\theta,\ \ y = r\sin\theta$.

3. A differentiable mapping $F : R \in \mathbf{R}^2 \to \mathbf{R}^3$ is what in calculus we called a parametric surface. Typically, one assumes that R is a simple closed region, such as a rectangle. If one denotes the coordinates in $\mathbf{R}^2$ by $(u, v) \in R$, and $\mathbf{x} \in \mathbf{R}^3$, the parametrization is written as $\mathbf{x}(u, v) = (x(u, v), y(u, v), z(u, v))$. The treatment of surfaces in $\mathbf{R}^3$ is presented in chapter 4. If $\mathbf{R}^3$ is replaced by $\mathbf{R}^n$, the mapping locally represents a 2-dimensional surface in a space of n dimensions.

For each point $\mathbf{p} \in \mathbf{R}^n$, we say that the Jacobian induces a linear transformation F_* from the tangent space $T_p\mathbf{R}^n$ to the tangent space $T_{F(p)}\mathbf{R}^m$. In differential geometry we this Jacobian map is also called 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

$$F_*X(f) = X(f \circ F), \tag{1.11}$$

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

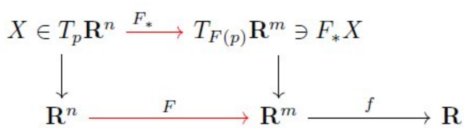


Fig. 1.4: Jacobian Map.

As shown in the diagram, $F_*X(f)$ is evaluated at $F(p)$ whereas X is evaluated at p. So, to be precise, equation 1.11 should really be written as

$$F_*X(f)(F(p)) = X(f \circ F)(p), \tag{1.12}$$

$$F_*X(f) \circ F = X(f \circ F), \tag{1.13}$$

As we have learned from linear algebra, to find a matrix representation of a linear map in a particular basis, one applies the map to the basis vectors. If we denote by $\{\frac{\partial}{\partial x^i}\}$ the basis for the tangent space at a point $p \in \mathbf{R}^n$ and by $\{\frac{\partial}{\partial y^j}\}$ the basis for the tangent space at the corresponding point $F(p) \in \mathbf{R}^m$ with coordinates given by $y^j = f^j(x^i)$, the push-forward definition reads,

$$\begin{aligned} F_*(\frac{\partial}{\partial x^i})(f) &= \frac{\partial}{\partial x^i}(f \circ F), \\ &= \frac{\partial f}{\partial y^j}\frac{\partial y^j}{\partial x^i}, \\ F_*(\frac{\partial}{\partial x^i}) &= \frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j}. \end{aligned}$$

In other words, the matrix representation of F_* in standard basis is in fact the Jacobian matrix. In classical notation, we simply write the Jacobian map in the familiar form,

$$\frac{\partial}{\partial x^i} = \frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j}. \tag{1.14}$$

1.2.2 Theorem If $F : \mathbf{R}^n \to \mathbf{R}^m$ and $G : \mathbf{R}^m \to \mathbf{R}^p$ are mappings, then $(G \circ F)_* = G_* \circ F_*$.

Proof Let $X \in T_p(\mathbf{R})^n$, and f be a smooth function $f : \mathbf{R}^p \to \mathbf{R}$. Then,

$$\begin{aligned}
(G \circ F)_*(X)(f) &= X(f \circ (G \circ F), \\
&= X((f \circ G) \circ F), \\
&= F_*(X)(f \circ G), \\
&= G_*(F_*(X)(f)), \\
&= (G_* \circ F_*)(X)(f).
\end{aligned}$$

1.2.3 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.

1.2.4 Remarks

1. Equation 1.14 shows that under change of coordinates, basis tangent vectors and by linearity all tangent vectors transform by multiplication by the matrix representation of the Jacobian. This is the source of the almost tautological 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_*.

3. If $F : \mathbf{R}^n \to \mathbf{R}^m$ and $G : \mathbf{R}^m \to \mathbf{R}^p$ are mappings, we leave it as an exercise for the reader to verify that the formula $(G \circ F)_* = G_* \circ F_*$ for the composition of linear transformations corresponds to the classical chain rule.

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

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

1.3.1 Parametric Curves

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

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

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

$$\mathbf{x}(t) = (x^1(t), x^2(t), x^3(t)), \tag{1.15}$$

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

$$x^i(t) = (u^i \circ \alpha)(t), \tag{1.16}$$

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

1.3.2 Example Let

$$\alpha(t) = (a_1 t + b_1, a_2 t + b_2, a_3 t + b_3). \tag{1.17}$$

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

1.3.3 Example Let

$$\alpha(t) = (a \cos \omega t, a \sin \omega t, bt). \tag{1.18}$$

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.

1.3.4 Example Let

$$\alpha(t) = (a(1 + \cos t), a \sin t, 2a \sin(t/2)). \tag{1.19}$$

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

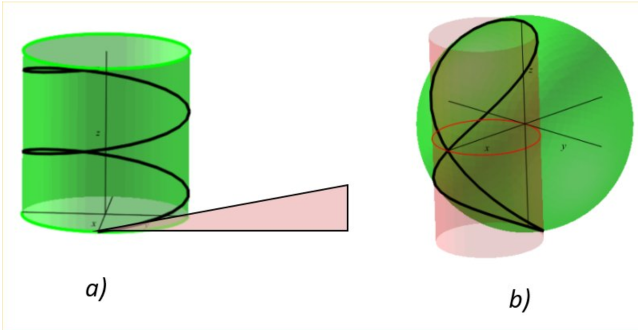
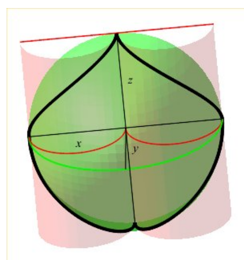


Fig. 1.5: a) Circular Helix. b) Temple of Viviani

$x^2 + y^2 = 2ax$ 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 must have originated from prominent Italian mathematicians. Upon receipt of the solution by Leibnitz, Viviani posted a mechanical 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."

1.3.5 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* or *tangent vector* of the curve by

$$\alpha'(t) = \left(\frac{d\alpha^1}{dt}, \frac{d\alpha^2}{dt}, \frac{d\alpha^3}{dt} \right)_{\alpha(t)}. \tag{1.20}$$

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

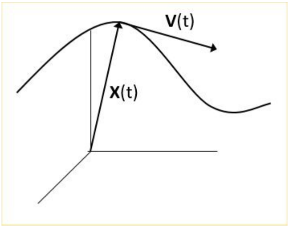
$$\mathbf{v}(t) = \dot{\mathbf{x}} \equiv \frac{d\mathbf{x}}{dt} = \left(\frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}\right), \tag{1.21}$$

and the speed is

$$v = \sqrt{\left(\frac{dx^1}{dt}\right)^2 + \left(\frac{dx^2}{dt}\right)^2 + \left(\frac{dx^3}{dt}\right)^2}. \tag{1.22}$$

The notation $T(t)$ or $T_\alpha(t)$ is also used for the tangent vector $\alpha'(t)$, but for now, we reserve $T(t)$ for the unit tangent vector to be introduced in section 1.3.3 on Frenet frames.

As is well known, the vector form of the equation of the line 1.17 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

$$d\mathbf{x} = (dx^1, dx^2, dx^3) = \left(\frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}\right) dt \tag{1.23}$$

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 arc length ds. Clearly, we have

$$ds = \|d\mathbf{x}\| = v\, dt. \tag{1.24}$$

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 arc length with respect to time. This is intuitively exactly what one would expect.

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, represents a linear approximation to the curve and points tangentially in the direction of $\mathbf{v}$.

1.3.2 Velocity

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

$$\alpha'(t)(f) \mid_{\alpha(t)} = \frac{d}{dt}(f \circ \alpha) \mid_t . \tag{1.25}$$

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 1.25, as a linear mapping between tangent spaces.

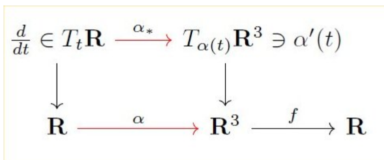


Fig. 1.6: Velocity Vector Operator

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 $\alpha'(t)$.

$$\alpha_*(d/dt) = \alpha'(t).$$

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 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 tangent vector

$$\alpha'(t)(x^i) \mid_{\alpha(t)} = \frac{d}{dt}(x^i \circ \alpha)|_t. \tag{1.26}$$

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

$$\begin{aligned} \alpha'(t) &= \frac{d}{dt}(x^i \circ \alpha)_t (\partial/\partial x^i)_{\alpha(t)}, \\ &= \frac{dx^1}{dt}\frac{\partial}{\partial x^1} + \frac{dx^2}{dt}\frac{\partial}{\partial x^2} + \frac{dx^3}{dt}\frac{\partial}{\partial x^3}. \end{aligned} \tag{1.27}$$

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

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

If $\alpha(t)$ is a curve in $\mathbf{R}^n$ with tangent vector $X = \alpha'(t)$, and $F : \mathbf{R}^n \to \mathbf{R}^m$ is differentiable map, then $F_* X$ is a tangent vector to the curve $F \circ \alpha$ in $\mathbf{R}^m$. That is, F_* maps tangent vectors of α to tangent vectors of $F \circ \alpha$.

1.3.6 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 arc length as the parameter. Using this reparametrization is quite natural, since we know from basic physics that the rate of change of the arc length is what we call speed

$$v = \frac{ds}{dt} = \|\alpha'(t)\|. \tag{1.28}$$

The arc length is obtained by integrating the above formula

$$s = \int \|\alpha'(t)\| \, dt = \int \sqrt{\left(\frac{dx^1}{dt}\right)^2 + \left(\frac{dx^2}{dt}\right)^2 + \left(\frac{dx^3}{dt}\right)^2} \, dt \tag{1.29}$$

In practice, it is typically difficult to find an explicit arc length parametrization of a curve since not only does one have to 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

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

1.3.7 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

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

1.3.3 Frenet Frames

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

$$T(s) = \beta'(s). \tag{1.30}$$

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

$$T \cdot T = 1, \tag{1.31}$$

we get

$$2\, T \cdot T' = 0, \tag{1.32}$$

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

$$T'(s) = \kappa N(s). \tag{1.33}$$

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

$$\kappa = \|T'(s)\|. \tag{1.34}$$

It makes sense to call κ the curvature because, 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 measure how much a curve is curving. We now introduce a third vector

$$B = T \times N, \tag{1.35}$$

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

$$
\begin{aligned}
T \cdot T = N \cdot N = B \cdot B = 1, \\
T \cdot N = T \cdot B = N \cdot B = 0.
\end{aligned} \tag{1.36}
$$

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

$$B'(s) = -\tau N(s), \tag{1.37}$$

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 in 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.

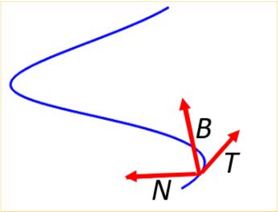


Fig. 1.7: Frenet Frame.

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.

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

$$\begin{array}{rcrcrcr} T' & = & & & \kappa N & & \\ N' & = & -\kappa T & & & & \tau B \\ B' & = & & & -\tau N & & \end{array} . \tag{1.38}$$

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 (1.38) can also be written in matrix form as shown below.

$$\begin{bmatrix} T \\ N \\ B \end{bmatrix}' = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} T \\ N \\ B \end{bmatrix}. \tag{1.39}$$

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. Presently, 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.

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

$$\begin{aligned} \kappa &= \|\beta''(s)\| \\ \tau &= \frac{\beta' \cdot [\beta'' \times \beta''']}{\beta'' \cdot \beta''} \end{aligned} \tag{1.40}$$

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

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

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

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

1.3.10 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 r, 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.

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

$$\begin{aligned} \mathbf{v} &= vT, \\ \mathbf{a} &= \frac{dv}{dt}T + v^2\kappa N. \end{aligned} \tag{1.41}$$

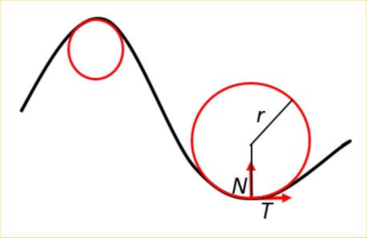


Fig. 1.8: Osculating Circle

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 1.41 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

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

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 2, in the sense that the circle passes through two nearby 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. If the three points are not collinear, they uniquely determine a circle. The center of this circle is located at the intersection of the perpendicular bisectors of the segments joining two consecutive points. 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 TN-plane, which by analogy is called the *osculating plane.* If $T' = 0$, then $\kappa = 0$ and the osculating circle degenerates into a circle of infinite radius, that is, a straight line. The physics interpretation of equation 1.41 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.

1.3.12 Example (Helix)

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

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 for which $\kappa/\tau =$ constant, is called a helix; the circular helix is a special case.

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

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

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

$$
\begin{aligned}
x(s) &= \int_0^s \cos \frac{t^2}{2c^2}\, dt, \\
y(s) &= \int_0^s \sin \frac{t^2}{2c^2}\, dt.
\end{aligned} \tag{1.42}
$$

Then, using the fundamental theorem of calculus, we have

$$\beta'(s) = (\cos \frac{s^2}{2c^2}, \sin \frac{s^2}{2c^2}, 0),$$

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

$$
\begin{aligned}
\kappa &= \|x'y'' - y'x''\| = (\beta' \cdot \beta')^{1/2}, \\
&= \| -\frac{s}{c^2} \sin \frac{s^2}{2c^2}, \frac{s}{c^2} \cos \frac{s^2}{2c^2}, 0\|, \\
&= \frac{s}{c^2}.
\end{aligned}
$$

The functions (1.42) are the 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.

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

$$\kappa^2 = \frac{\|\alpha' \times \alpha''\|^2}{\|\alpha'\|^6}, \tag{1.43}$$

$$\tau = \frac{(\alpha'\alpha''\alpha''')}{\|\alpha' \times \alpha''\|^2}, \tag{1.44}$$

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

Proof

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

As the computation below shows, the other terms in α''' are unimportant here because $\alpha' \times \alpha''$ is proportional to B, so all we need is the B component to solve for τ.

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

1.4 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 $s = 0$.

$$\beta(s) = \beta(0) + \beta'(0)s + \frac{\beta''(0)}{2!}s^2 + \frac{\beta'''(0)}{3!}s^3 + \dots \tag{1.45}$$

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

$$\beta(s) \doteq \beta(0) + T_0 s + \frac{1}{2}\kappa_0 N_0 s^2 + \frac{1}{6}\kappa_0\tau_0 B_0 s^3. \tag{1.46}$$

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 into a parabolic shape in the TN and twisted into a cubic along the B axis. So suppose that p

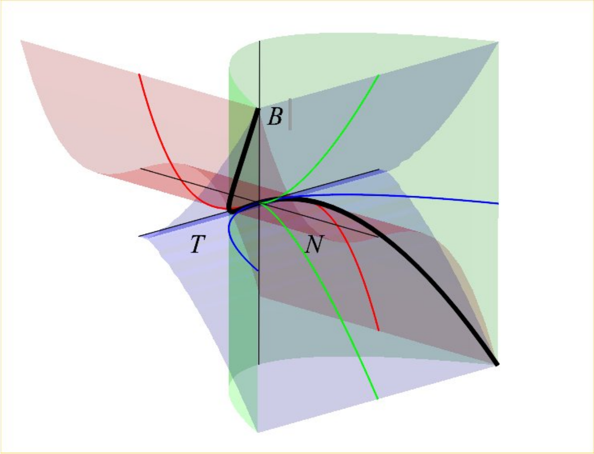


Fig. 1.9: Cubic Approximation to a Curve

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 $\{e_1, e_2, e_3\}$ in $\mathbf{R}^3$ to coincide with the Frenet Frame T_0, N_0, B_0 at that point. then, the equation (1.46) 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.)

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

1.4.2 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

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

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 to the notion of isometries.

1.4.1 Isometries

1.4.3 Definition Let $\mathbf{x}$ and $\mathbf{y}$ be two column vectors in $\mathbf{R}^n$ and let $\mathbf{x}^T$ represent the transposed 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

$$\begin{aligned} < \mathbf{x}, \mathbf{y} > &= \mathbf{x}^T\mathbf{y}, && (1.47)\\ &=< \mathbf{y}, \mathbf{x} > . && (1.48) \end{aligned}$$

The inner product gives $\mathbf{R}^n$ the structure of a normed space by defining $\|\mathbf{x}\| =< \mathbf{x}, \mathbf{x} >^{1/2}$ 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

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 \pm 2 < \mathbf{x}, \mathbf{y} > + \|\mathbf{y}\|^2. \tag{1.49}$$

Thus, we have the *polarization identity*

$$< \mathbf{x}, \mathbf{y} >= \tfrac{1}{4}\|\mathbf{x} + \mathbf{y}\|^2 - \tfrac{1}{4}\|\mathbf{x} - \mathbf{y}\|^2. \tag{1.50}$$

The Euclidean inner product satisfies the relation

$$< \mathbf{x}, \mathbf{y} >= \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cos\theta, \tag{1.51}$$

where θ is the angle subtended by the two vectors.

Two vectors $\mathbf{x}$ and $\mathbf{y}$ are called *orthogonal* if $< \mathbf{x}, \mathbf{y} >= 0$, and a set of basis vectors $\mathscr{B} = \{\mathbf{e}_1, \ldots \mathbf{e}_n\}$ is called an *orthonormal* basis if $< \mathbf{e}_i, \mathbf{e}_j >= \delta_{ij}$. 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

$$\begin{aligned} <\mathbf{x}, \mathbf{y}> &= \mathbf{x}^T \mathbf{y}, \\ &= (x_i \alpha^i)(y^j \mathbf{e}_j), \\ &= (x_i y^j) \alpha^i (\mathbf{e}_j) = (x_i y^j) \delta^i_j. \\ &= x_i y^i, \\ &= \begin{bmatrix} x_1 & x_2 \ldots & x_n \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \\ \ldots \\ y^n \end{bmatrix} \end{aligned}$$

Since $|\cos\theta| \leq 1$, it follows from equation 1.51, a special case of the *Schwarz inequality*

$$| <\mathbf{x}, \mathbf{y}> | \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|. \tag{1.52}$$

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

$$A = (a^i_j) = \alpha^i(F(\mathbf{e}_j)), \tag{1.53}$$

or in terms of the inner product,

$$A = (a_{ij}) = < \mathbf{e}_i, F(\mathbf{e}_j) > . \tag{1.54}$$

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

$$\begin{aligned} <\mathbf{x}, A\mathbf{y}> &= \mathbf{x}^T A\mathbf{y}, && (1.55) \\ &= (A^T \mathbf{x})^{\mathbf{T}} \mathbf{y}, \\ &= < A^T \mathbf{x}, \mathbf{y} > . && (1.56) \end{aligned}$$

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

$$<\mathbf{x}, A\mathbf{y}> = < A^{-1}\mathbf{x}, \mathbf{y} > . \tag{1.57}$$

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

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

Proof

$$\begin{aligned} < A\mathbf{x}, A\mathbf{y} > &= < A^T A\mathbf{x}, \mathbf{y} >, \\ &= < \mathbf{x}, \mathbf{y} > . \end{aligned}$$

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

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

As a corollary, if $\{\mathbf{e}_i\}$ is an orthonormal basis, then 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 $\det A = -1$.

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

$$d(F(\mathbf{x}), F(\mathbf{y})) = d(\mathbf{x}, \mathbf{y}). \tag{1.58}$$

1.4.7 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 $\|F(\mathbf{x}) - F(\mathbf{y})\| = \|\mathbf{x} + \mathbf{p} - (\mathbf{y} + \mathbf{p})\| = \|\mathbf{x} - \mathbf{y}\|$.

1.4.8 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:

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

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}. Consequently, the set of isometries consisting of translations and orthogonal transformations constitutes a group. Given a general isometry, we can use a translation to insure that $F(\mathbf{0}) = \mathbf{0}$. We now prove the following theorem.

1.4.9 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 1.49 and the norm preserving property above, we have:

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

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

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

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

$$F(\mathbf{x}) = A\mathbf{x} + \mathbf{q}, \tag{1.59}$$

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.

1.4.11 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:

1.4.12 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 loss of generality, we can make this point be the origin. Now we perform an orthogonal transformation 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 also a unit vector, 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.

1.4.2 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. 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 are characterized

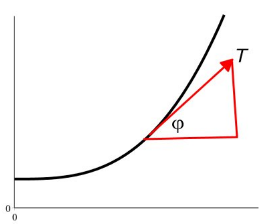


Fig. 1.10: Tangent

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 $\mathbf{x} = (x, y)$ 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

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

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

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

so that,

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

We conclude that

$$x(s) = \int \cos\varphi \, ds, \quad y(s) = \int \sin\varphi \, ds, \text{where,} \quad \varphi = \int \kappa \, ds. \tag{1.60}$$

Equations 1.60 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.

1.4.13 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

$$\mathbf{x} = (R\sin\frac{s}{R}, -R\cos\frac{s}{R}),$$

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

1.4.14 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

$$x(s) = C(s) = \int \cos(\tfrac{1}{2}\pi s^2) \, ds; \qquad y(s) = S(s) = \int \sin(\tfrac{1}{2}\pi s^2) \, ds. \tag{1.61}$$

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

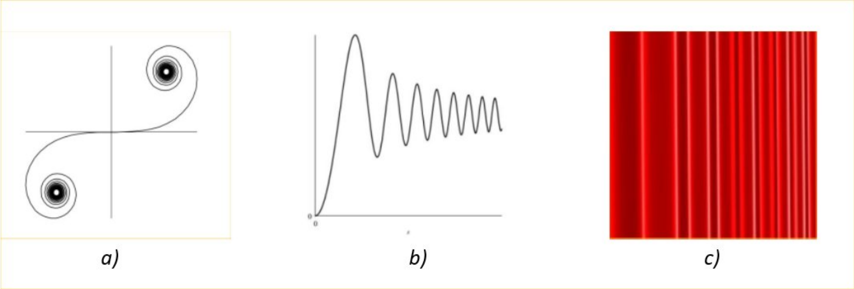


Fig. 1.11: Fresnel Diffraction

trapezoidal rule as taught in typical calculus courses, completely fall apart in this range. The *Cornu spiral* is the curve $\mathbf{x}(s) = (x(s), y(s))$ 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 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 be 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.

1.4.15 Example Logarithmic Spiral $\kappa = 1/(as + b)$
A logarithmic spiral is a curve in which the position vector $\mathbf{x}$ makes a constant angle with the tangent vector, as shown in figure 1.12. A formula for the curve can be found easily if one uses the calculus formula in polar coordinates

$$\tan\psi = \frac{r}{dr/d\theta}. \tag{1.62}$$

Here, ψ is the angle between the polar direction and the tangent. If ψ is constant, then one can immediately integrate the equation to get the exponential function below, in which k is the constant of integration

$$r(\theta) = ke^{(\cot\psi)\theta} \tag{1.63}$$

Derivation of formula 1.62 has fallen through the cracks in standard fat calculus textbooks, at best relegated to an advanced exercise which most students

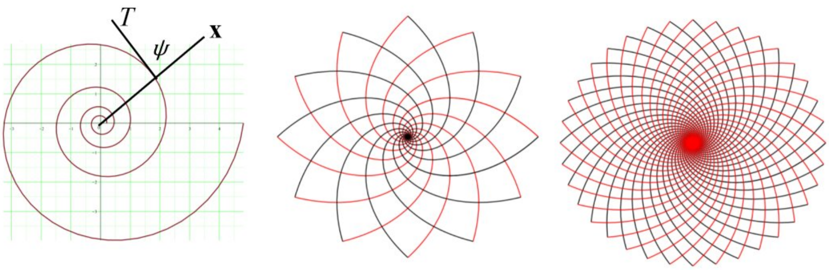


Fig. 1.12: Logarithmic Spiral

will not do. Perhaps the reason is that the section on polar coordinates is typically covered in Calculus II, so students have not yet been exposed to the tools of vector calculus that facilitate the otherwise messy computation. To fill-in this gap, we present a short derivation of this neat formula. For a plane curve in parametric polar coordinates, we have

$$\begin{aligned} \mathbf{x}(t) &= (r(t)\cos\theta(t), r(t)\sin\theta(t)), \\ \dot{\mathbf{x}} &= (\dot{r}\cos\theta - r\sin\theta\,\dot{\theta}, \dot{r}\sin\theta + r\cos\theta\,\dot{\theta}). \end{aligned}$$

A direct computation of the dot product gives,

$$| < \mathbf{x}, \dot{\mathbf{x}} > |^2 = (r\dot{r})^2.$$

On the other hand,

$$\begin{aligned} | < \mathbf{x}, \dot{\mathbf{x}} > |^2 &= \|\mathbf{x}\|^2 \, \|\dot{\mathbf{x}}\|^2 \, \cos^2\psi, \\ &= r^2(\dot{r}^2 + r^2\dot{\theta}^2)\cos^2\psi. \end{aligned}$$

Equating the two, we find,

$$\begin{aligned} \dot{r}^2 &= (\dot{r}^2 + r^2\dot{\theta}^2)\cos^2\psi, \\ (\sin^2\psi)\dot{r}^2 &= r^2\dot{\theta}^2\cos^2\psi, \\ (\sin\psi)\,dr &= r\cos\psi\,d\theta, \\ \tan\psi &= \frac{r}{dr/d\theta}. \end{aligned}$$

We leave it to the reader to do a direct computation of the curvature. Instead, we prove that if $\kappa = 1/(as+b)$, where a and b are constant, then the curve is a logarithmic spiral. From the natural equations, we have,

$$\begin{aligned}
\frac{d\theta}{ds} &= \kappa = \frac{1}{as+b}, \\
\theta &= \frac{1}{a}\ln(as+b) + C, \quad C = \text{const}, \\
e^{a\theta} &= B(as+b), \quad B = e^{aC} = 1/A, \\
\frac{1}{\kappa} &= Ae^{a\theta} = \frac{ds}{d\theta}, \\
ds &= Ae^{as}\, d\theta.
\end{aligned}$$

Back to the natural equations, the x and y coordinates are obtained by integrating,

$$\begin{aligned}
x &= \int Ae^{a\theta}\cos\theta\, d\theta, \\
y &= \int Ae^{a\theta}\sin\theta\, d\theta.
\end{aligned}$$

We can avoid the integrations by parts by letting $z = x + iy = re^{i\theta}$. We get

$$\begin{aligned}
z &= A\int e^{a\theta}e^{i\theta}\, d\theta, \\
&= A\int e^{(a+i)\theta}\, d\theta, \\
&= \frac{A}{a+i}e^{(a+i)\theta}, \\
&= \frac{A}{a+i}e^{a\theta}e^{i\theta}.
\end{aligned}$$

Extracting the real part $\|z\| = r$, we get

$$r = \frac{A}{\sqrt{a^2+1}}e^{a\theta}, \tag{1.64}$$

which is the equation of a logarithmic spiral with $a = \cot\psi$. As shown in figure 1.12, families of concentric logarithmic spirals are ubiquitous in nature as in flowers and pine cones, in architectural designs. The projection of a conical helix as in figure 4.8 onto the plane through the origin, is a logarithmic spiral. The analog of a logarithmic spiral on a sphere is called a loxodrome as depicted in figure 4.2.

1.4.16 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 1.13 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.

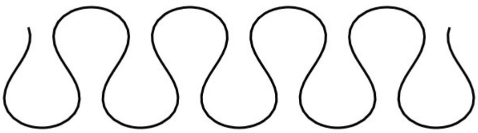


Fig. 1.13: Meandering Curve

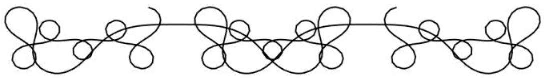


Fig. 1.14: Bimodal Meander

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 1.14 shows a rendition with two sine functions with equal amplitude $A = 1.8$, and with $k_1 = 1,\ k_2 = 1.2$.

EXERCISES

1. Show that the space $T_p\mathbf{R^3} = \{a^1 \frac{\partial}{\partial x^1}|_p + a^2 \frac{\partial}{\partial x^2}|_p + a^3 \frac{\partial}{\partial x^3}|_p \quad : \ a^1, a^2, a^3 \in \mathbf{R}\}$, has the structure of a vector space.

2. Let $X_p = -2\frac{\partial}{\partial x^1}|_p + \frac{\partial}{\partial x^2}|_p - \frac{\partial}{\partial x^3}|_p$, and $Y_p = \frac{\partial}{\partial x^2}|_p + 3\frac{\partial}{\partial x^3}|_p$. Express$6\frac{\partial}{\partial x^1}|_p + 5\frac{\partial}{\partial x^2}|_p + 5\frac{\partial}{\partial x^3}|_p$ as a linear combination of X_p and Y_p.

3. Let X_p be the tangent vector for which $\mathbf{X} = \langle 1, -2, 2\rangle$ and $\mathbf{p} = \langle 3, 4, -5\rangle$. Compute $X_p(f)$, where $f = x^3 - y^2z^2 + 2$.

4. Let X_p be the tangent vector for which $\mathbf{X} = \langle 3, -4, 5\rangle$ and $\mathbf{p} = \langle 1, -2, 4\rangle$. Compute $X_p(f)$, where $f = x^2y^3 - yz^3 + 5y$.

5. Given the vector field $X = xy\frac{\partial}{\partial x} + y^2\frac{\partial}{\partial z}$ and the function $f = x^2yz^4$, compute $X(f)$

6. Let $\mathbf{v}$ be the Euclidean vector $\langle v^1, v^2, v^3\rangle$. Starting from the alternative definition $V_p(f) = \frac{d}{dt}((f(p + t\mathbf{v}))|t = 0$, show that $V_p = v^i\frac{\partial}{\partial x^i}$.

7. Given the vector fields $V = y^2\frac{\partial}{\partial x} - x^2\frac{\partial}{\partial z}$ and $W = x^2\frac{\partial}{\partial x} - z\frac{\partial}{\partial y}$, find functions f and g so that the vector field $fV + gW$ can be expressed in term of $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$ only.

8. Given the vector field $X = z\frac{\partial}{\partial x} + y^2z\frac{\partial}{\partial z}$ and functions $f = x^2yz^4,\ g = x + z$, compute:

   a) $X(f)$ *b)* $X(fg)$
   c) $X(X(f))$ *d)* $gX(f) - fX(g)$

9. Given vector fields X and Y and any function f, define $\pounds_X Y(f) = [X, Y](f)$, where $[X, Y](f) = X(Y(f)) - Y(X(f))$. Show that:

   a)$\pounds_X Y(f) = [X, Y](f)$ is a linear derivation.

   b)$([[X, Y], Z] + [[Z, X], Y] + [[Y, Z], X])(f) = 0$

10. Show that (Hint, first prove Schaum's, Theorem 1.8, pg 10,)

$$(\mathbf{p}\times\mathbf{q})\cdot(\mathbf{p}\times\mathbf{q}) = \det\begin{pmatrix} \mathbf{p}\cdot\mathbf{p} & \mathbf{q}\cdot\mathbf{p} \\ \mathbf{p}\cdot\mathbf{q} & \mathbf{q}\cdot\mathbf{q} \end{pmatrix} = \|\mathbf{p}\|^2\|\mathbf{q}\|^2 - (\mathbf{q}\cdot\mathbf{p})^2$$

11. Let X and Y be vector fields in $\mathbf{R}^3$ and define $\mathscr{L}_X Y \equiv [X,Y] = XY - YX$.

    a) Show that $\mathscr{L}_X Y$ is a linear derivation on the space of functions and thus it is also a vector field.

    b) Show that $\mathscr{L}_X[Y,Z] = [\mathscr{L}_X Y, Z] + [Y, \mathscr{L}_X Z]$.

12. Let $P_i = -i\hbar\frac{\partial}{\partial x_i}$ $i = 1..3$, and consider the vector fields

    $L_1 = x^2P_3 - x^3P_2,\;\; L_2 = x^3P_1 - x^1P_3,\;\; L_3 = x^1P_2 - x^2P_1.$

    Show that $[L_i, L_j] = i\hbar\epsilon_{ij}{}^k L_k$.

13. Define $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

    Define $[A,B] = AB - BA$. Show that $\frac{\hbar}{2}[\sigma_i, \sigma_j] = i\hbar\epsilon_{ij}{}^k\sigma_k$

14. Show that $\kappa^2 = \tau^2$ along the curve (Struik, pg 21)

$$\alpha(t) = \langle a(3t - t^3), 3at^2, a(3t + t^3)\rangle$$

15. Show that the natural equations for the epicycloid (Struik, pg. 26)

$$\alpha(t) = \langle[(a+b)\cos\phi - b\cos(\frac{a+b}{b})\phi],\;[(a+b)\sin\phi - b\sin(\frac{a+b}{b})\phi]\rangle$$

    is of the form

$$\frac{s^2}{A^2} + \frac{R^2}{B^2} = 1,$$

    where A and B are constant. Plot using a CAS

16. Find the cartesian equation of the curve C with natural equations $Rs = a^2$ (a=const), (Struik, pg. 201) taking the inflection point at the origin and the tangent at this point to be the x-axis. Plot the curve C. Plot the curve C and the curve Γ representing the distance from the origin to a point on the curve C (restricted to the first quadrant,) as a function of arclength. (Note: The curve C is called a clothoid, It appears in the theory of diffraction and in the design of roller-coaster loops)

17. Find the unit tangent vector of the curve of intersection of two surfaces (Struik, pg. 22) $F_1(x,y,z) = 0$ and $F_2(x,y,z) = 0$.

18. Find the curvature and torsion of the curve (Struik, pg 21.)

$$\mathbf{x}(t) = \langle a(3t - t^3), 3at^2, a(3t + t^3)\rangle.$$

19. The Temple of Viviani curve is defined by the equation $\mathbf{x}(t) = \langle a(1 + \cos t), a\sin t, 2a\sin t/2) \rangle$.

    a) Show that the Temple of Viviani is the curve of intersection of the sphere $x^2 + y^2 + z^2 = 2a^2$ and the cylinder $(x - a)^2 + y^2 = a^2$.

    b) Show that the curve is also the intersection of the cone $(x-2a)^2+y^2 = z^2$, and the parabolic cylinder $z^2 = 2a(2a - x)$. Use maple to graph the curve of intersection for $a = 1/2$.

    c) Compute the curvature $\kappa(t)$ and evaluate at $t = 0$

20. Find the natural equation of the catenary $\mathbf{x}(t) = \langle t, a\cosh(t/a) \rangle$.

21. (Loxodromes) Let $\mathbf{x}(t) = \langle \cos t \cos \sigma, \sin t \cos \sigma, -\sin \sigma \rangle$, where $\sigma = \tan^{-1}(at)$.

    a) Show that this curve lies on a unit sphere.

    b) Use maple to render this curve on the sphere, for $a = \pi/3, \pi/6, \pi/12$.

22. The tangent vectors $\mathbf{T}$ along a unit speed curve $\beta(t)$ generate a curve $\Sigma : \ \beta_\sigma(s) = T(s)$ (called the spherical indicatrix) on the unit sphere. Show that the curvature κ_σ and torsion τ_σ of Σ satisfy

$$\begin{aligned} \kappa_\sigma^2 &= \frac{\kappa^2 + \tau^2}{\kappa^2} \\ \tau_\sigma &= \frac{\tau\kappa' - \kappa\tau'}{\kappa(\kappa^2 + \tau^2)}. \end{aligned}$$

23. (Spherical Curve) Let $\beta(s)$ be a unit speed curve with $\kappa > 0, \tau \neq 0$ that lies on a sphere of radius a, and let $R = 1/\kappa$, $T = 1/\tau$. Show that $R^2 + (TR')^2 = a^2$

24. Use the Frenet frame formulas to show that a curve $\mathbf{x}(s)$ with prescribed curvature and torsion κ and τ satisfies the differential equation:

    $\mathbf{x}^{(iv)} - (2\frac{\kappa'}{\kappa} + \frac{\tau'}{\tau})\mathbf{x}''' + (\kappa^2 + \tau^2 - \frac{\kappa\kappa'' - 2\kappa'^2}{\kappa^2} + \frac{\kappa'\tau'}{\kappa\tau})\mathbf{x}'' + \kappa^2(\frac{\kappa'}{\kappa} - \frac{\tau'}{\tau})\mathbf{x}' = 0.$

25. (Rigid Body Motion) Considered as a rigid body, the moving trihedron moving along a curve $\mathbf{x}(s)$ rotates around an axis on the direction of the vector $\mathbf{R} = \tau\mathbf{T} + \kappa\mathbf{B}$. Show that:

    $\mathbf{T}' = \mathbf{R} \times \mathbf{T}, \quad \mathbf{N}' = \mathbf{R} \times \mathbf{N}, \quad \mathbf{B}' = \mathbf{R} \times \mathbf{B},$

26. Find the unit tangent vector of the curve of intersection of two surfaces (Struik, pg. 22) $F_1(x, y, z) = 0$ and $F_2(x, y, z) = 0$.

Chapter 2

Differential Forms

2.1 One-Forms

The concept of the differential of a function is one of the most puzzling ideas in elementary calculus. 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 a linear operator.

2.1.1 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:

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

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

2.1.2 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 finite dimensional vector space is also a vector space of the same dimension. Thus, the space $T^\star_p(\mathbf{R}^n)$ 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 $T^\star_p(\mathbf{R}^n)$ 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* $T^*(\mathbf{R}^n)$.

2.1.3 Proposition Let f be a 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 $\{(dx^1)_p, \ldots, (dx^n)_p\}$. 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 $\Omega^1(\mathbf{R}^n) = \mathscr{T}_1^0(\mathbf{R}^n)$. 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

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.

2.2.1 Definition A map $\phi : \mathscr{X}(\mathbf{R}^n) \times \mathscr{X}(\mathbf{R}^n) \longrightarrow \mathbf{R}$ is called a *bilinear* map of vector fields, if it is linear on each slot. That is, $\forall X_i, Y_i \in \mathscr{X}(\mathbf{R}^n),\ \ f^i \in \mathscr{F}(\mathbf{R}^n)$, we have

$$\begin{aligned} \phi(f^1X_1 + f^2X_2, Y_1) &= f^1\phi(X_1, Y_1) + f^2\phi(X_2, Y_1) \\ \phi(X_1, f^1Y_1 + f^2Y_2) &= f^1\phi(X_1, Y_1) + f^2\phi(X_1, Y_2). \end{aligned}$$

2.2.1 Tensor Products

2.2.2 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 $\binom{2}{1}$. Thus, we may think of a tensor of type $\binom{2}{1}$ as a 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.

A tensor of type $\binom{r}{s}$ is written in local coordinates as

$$T = T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots dx^{j_s} \tag{2.12}$$

The tensor components are given by

$$T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} = T(dx^{i_1},\ldots,dx^{i_r},\frac{\partial}{\partial x^{j_1}},\ldots,\frac{\partial}{\partial x^{j_s}}). \tag{2.13}$$

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 $\mathbf{R}^n$.

2.2.2 Inner Product

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

$$g(X,Y) = \delta_{ij} a^i b^j. \tag{2.14}$$

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

2.2.3 Definition A bilinear map $g(X,Y) \equiv < X, Y >$ on vectors is called a real *inner product* if

1. $g(X,Y) = g(Y,X)$,
2. $g(X,X) \geq 0, \;\; \forall X$,
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

$$g(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}) = g_{ij}, \tag{2.15}$$

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

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

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

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

If we now define

$$b_i = g_{ij} b^j, \tag{2.16}$$

we see that the equation above can be rewritten as

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

and we recover the expression for the inner product.

Equation (2.16) 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

$$g^{ik} g_{kj} = \delta^i_j, \tag{2.17}$$

we can also raise covariant indices by the equation

$$b^i = g^{ij} b_j. \tag{2.18}$$

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 g can be interpreted on one hand as a bilinear pairing of two vectors

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

and on the other, as inducing a linear isomorphism

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

defined by

$$G_\flat X(Y) = g(X, Y), \tag{2.19}$$

that maps vectors to covectors. To verify this definition is consistent with the action of lowering indices, let $X = a^i \frac{\partial}{\partial x^i}$ and $Y = b^j \frac{\partial}{\partial x^j}$. We show that that $G_\flat X = a_i \, dx^i$. In fact,

$$\begin{aligned}G_\flat X(Y) &= (a_i \, dx^i)(b^j \tfrac{\partial}{\partial x^j}), \\ &= a_i b^j dx^i(\tfrac{\partial}{\partial x^j}), \\ &= a_i b^j \delta^i{}_j, \\ &= a_i b^i = g_{ij} a^j b^i, \\ &= g(X, Y).\end{aligned}$$

The inverse map $G^\sharp : T_p^\star(\mathbf{R}^n) \longrightarrow T_p(\mathbf{R}^n)$ is defined by

$$< G^\sharp\alpha, X >= \alpha(X), \tag{2.20}$$

for any 1-form α and tangent vector X. In quantum mechanics, it is common to use Dirac's notation, in which a linear functional α on a vector space $\mathcal{V}$ is called a *bra-vector* denoted by $\langle\alpha|$, and a vector $X \in \mathcal{V}$ is called a *ket-vector*, denoted by $|X\rangle$. The, action of a bra-vector on a ket-vector is defined by the *bracket*,

$$\langle\alpha|X\rangle = \alpha(X). \tag{2.21}$$

Thus, if the vector space has an inner product as above, we have

$$\langle\alpha|X\rangle =< G^\sharp\alpha, X >= \alpha(X). \tag{2.22}$$

The mapping $C : T_p^*(\mathbf{R}^n) \to \mathbf{R}$ given by $(\alpha, X) \mapsto \langle\alpha|X\rangle = \alpha(X)$ is called a *contraction*. In passing, we introduce a related concept called the *interior product*, or contraction of a vector and a form. If α is a $(k+1)$-form and X a vector, we define

$$i_X\alpha(X_1, \ldots, X_k) = \alpha(X, X_1, \ldots, X_k). \tag{2.23}$$

In particular, for a one form, we have

$$i_X\alpha = \langle\alpha|X\rangle = \alpha(X).$$

If T is a type $\binom{1}{1}$ tensor, that is,

$$T = T^i{}_j dx^j \otimes \frac{\partial}{\partial x^i},$$

The contraction of the tensor is given by

$$\begin{aligned} C(T) &= T^i{}_j \langle dx^j | \tfrac{\partial}{\partial x^i}\rangle, \\ &= T^i{}_j\ dx^j(\tfrac{\partial}{\partial x^i}), \\ &= T^i{}_j\ \delta^j_i, \\ &= T^i{}_i. \end{aligned}$$

In other words, the contraction of the tensor is the trace of the $n \times n$ array that represents the tensor in the given basis. The notion of raising and lowering indices as well as contractions can be extended to tensors of all types. Thus, for example, we have

$$g^{ij}T_{iklm} = T^i{}_{klm}.$$

A contraction between the indices i and l in the tensor above could be denoted by the notation

$$C^1_2(T^i{}_{klm}) = T^i{}_{kim} = T_{km}.$$

This is a very simple concept, but the notation for a general contraction is a bit awkward because one needs to keep track of the positions of the indices contracted. Let T be a tensor of type $\binom{r}{s}$. A contraction C^k_l yields a tensor of type $\binom{r-1}{s-1}$. Let T be given in the form 2.12. Then,

$$C^l_k(T) = T^{i_1\ldots i_{l-1},m,i_{i+1}\ldots i_r}_{j_1\ldots j_{k-1},m,j_{k+1}\ldots j_s}\frac{\partial}{\partial x^{i_1}}\otimes\ldots\otimes\widehat{\frac{\partial}{\partial x^{i_l}}}\otimes\ldots\otimes\frac{\partial}{\partial x^{i_r}}\otimes dx^{j_1}\otimes\ldots\otimes\widehat{dx^{j_k}}\otimes\ldots dx^{j_s},$$

where the "hat" means that these are excluded. Here is a very neat and most useful result. If S is a 2-tensor with symmetric components $T_{ij} = T_{ji}$ and A is a 2-tensor with antisymmetric components $A^{ij} = -A^{ji}$, then the contraction

$$S_{ij}A^{ij} = 0 \tag{2.24}$$

The short proof uses the fact that summation indices are dummy indices and they can be relabeled at will by any other index that is not already used in an expression. We have

$$S_{ij}A^{ij} = S_{ji}A^{ij} = -S_{ji}A^{ji} = -S_{kl}A^{kl} = -S_{ji}A^{ij} = 0,$$

since the quantity is the negative of itself.

In terms of the vector space isomorphism between the tangent and cotangent space induced by the metric, the gradient of a function f, viewed as a differential geometry vector field, is given by

$$\text{Grad } f = G^\sharp df, \tag{2.25}$$

or in components

$$(\nabla f)^i \equiv \nabla^i f = g^{ij} f_{,j}\,, \tag{2.26}$$

where $f_{,j}$ is the commonly used abbreviation for the partial derivative with respect to x^j.

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

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

although what is really meant by such an expression is

$$ds^2 \equiv g_{ij}dx^i \otimes dx^j. \tag{2.27}$$

2.2.4 Example In cylindrical coordinates, the differential of arc length is

$$ds^2 = dr^2 + r^2 d\theta^2 + dz^2. \tag{2.28}$$

In this case, the metric tensor has components

$$g_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.29}$$

2.2.5 Example In spherical coordinates,

$$\begin{aligned} x &= r \sin\theta \cos\phi, \\ y &= r \sin\theta \sin\phi, \\ z &= r \cos\theta, \end{aligned} \tag{2.30}$$

and the differential of arc length is given by

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2. \tag{2.31}$$

In this case the metric tensor has components

$$g_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin\theta^2 \end{bmatrix}. \tag{2.32}$$

2.2.3 Minkowski Space

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

$$M_{(1,3)} = \{(t, x^1, x^2, x^3) |\; t, x^i \in \mathbf{R}\} \tag{2.33}$$

and η is the bilinear map such that

$$\eta(X, X) = t^2 - (x^1)^2 - (x^2)^2 - (x^3)^2. \tag{2.34}$$

The matrix representing Minkowski's metric η is given by

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

in which case, the differential of arc length is given by

$$\begin{aligned} ds^2 &= \eta_{\mu\nu} dx^\mu \otimes dx^\nu \\ &= dt \otimes dt - dx^1 \otimes dx^1 - dx^2 \otimes dx^2 - dx^3 \otimes dx^3 \\ &= dt^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2. \end{aligned} \tag{2.35}$$

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 particles that travel at the speed of light (which we have set equal to 1 in our system of units.)

The Minkowski metric $\eta_{\mu\nu}$ and its matrix inverse $\eta^{\mu\nu}$ 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

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

2.2.4 Wedge Products and 2-Forms

2.2.6 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 model 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.

2.2.7 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.

2.2.8 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

$$\begin{aligned}(\alpha \wedge \beta)(X, Y) &= \alpha(X)\beta(Y) - \alpha(Y)\beta(X), \\ &= \begin{bmatrix} \alpha(X) & \alpha(Y) \\ \beta(X) & \beta(Y) \end{bmatrix}\end{aligned} \tag{2.36}$$

2.2.9 Theorem If α and β are 1-forms, then $\alpha \wedge \beta$ is a 2-form.

Proof Let α and β be 1-forms in $\mathbf{R}^n$ and let X and Y be any two vector fields. Then

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

Thus, the wedge product of two 1-forms is alternating.

To show that the wedge product of two 1-forms is bilinear, 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

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

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

The wedge product of two 1-forms has 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.

2.2.10 Example Consider the case of $\mathbf{R}^2$. Let

$$\begin{aligned} \alpha &= a\, dx + b\, dy, \\ \beta &= c\, dx + d\, dy. \end{aligned}$$

since $dx \wedge dx = dy \wedge dy = 0$, and $dx \wedge dy = -dy \wedge dx$, we get,

$$\begin{aligned} \alpha \wedge \beta &= ad\, dx \wedge dy + bc\, dy \wedge dx, \\ &= ad\, dx \wedge dy - bc\, dx \wedge dy, \\ &= \begin{vmatrix} a & b \\ c & d \end{vmatrix} dx \wedge dy. \end{aligned}$$

The similarity between wedge products is even more striking in the next example, but we emphasize again that wedge products are much more powerful than cross products, because wedge products can be computed in any dimension.

2.2.11 Example For combinatoric reasons, it is convenient to label the coordinates as $\{x^1, x^2, x^3\}$. Let

$$\begin{aligned} \alpha &= a_1\, dx^1 + a_2\, dx^2 + a_3\, dx^3, \\ \beta &= b_1\, dx^1 + b_2\, dx^2 + b_3\, dx^3, \end{aligned}$$

There are only three independent basis 2-forms, namely

$$\begin{aligned} dy \wedge dz &= dx^2 \wedge dx^3, \\ dx \wedge dz &= -dx^1 \wedge dx^3, \\ dx \wedge dy &= dx^1 \wedge dx^2. \end{aligned}$$

Computing the wedge products in pairs, we get

$$\alpha \wedge \beta = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} dx^2 \wedge dx^3 + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} dx^1 \wedge dx^3 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} dx^1 \wedge dx^2.$$

If we consider vectors $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$, we see that the result above can be written as

$$\alpha \wedge \beta = (\mathbf{a} \times \mathbf{b})_1\, dx^2 \wedge dx^3 - (\mathbf{a} \times \mathbf{b})_2\, dx^1 \wedge dx^3 + (\mathbf{a} \times \mathbf{b})_3\, dx^1 \wedge dx^2 \quad (2.37)$$

It is worthwhile noticing that if one thinks of the indices in the formula above as permutations of the integers $\{1, 2, 3\}$, the signs of the three terms correspond to the signature of the permutation. In particular, the middle term indices constitute an odd permutation, so the signature is minus one. 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 2.1, 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.

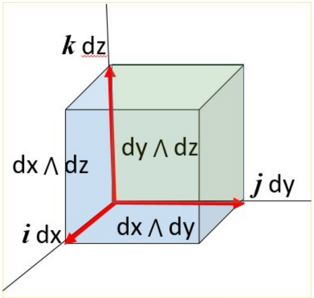


Fig. 2.1: Area Forms

2.2.12 Example One could of course compute wedge products by just using the linearity properties. It would not be as efficient as grouping into pairs, but it would yield the same result. For example, let
$\alpha = x^2 dx - y^2 dy$ and $\beta = dx + dy - 2xydz$. Then,

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

In local coordinates, a 2-form can always be written in components as

$$\phi = F_{ij}\ dx^i \wedge dx^j \tag{2.38}$$

If we think of F as a matrix with components F_{ij}, we know from linear algebra that we can write F uniquely as a sum of a symmetric and an antisymmetric matrix, namely,

$$\begin{aligned} F &= S + A, \\ &= \frac{1}{2}(F + F^T) + \frac{1}{2}(F - F^T), \\ F_{ij} &= F_{(ij)} + F_{[ij]}, \end{aligned}$$

where,

$$\begin{aligned} F_{(ij)} &= \frac{1}{2}(F_{ij} + F_{ji}), \\ F_{[ij]} &= \frac{1}{2}(F_{ij} - F_{ji}), \end{aligned}$$

are the completely symmetric and antisymmetric components. Since $dx^i \wedge dx^j$ is antisymmetric, and the contraction of a symmetric tensor with an antisymmetric tensor is zero, one may assume that the components of the 2-form in equation 2.38 are antisymmetric as well. With this mind, we can easily find a formula using wedges that generalizes the cross product to any dimension.

Let $\alpha = a_i dx^i$ and $\beta = b_i dx^i$ be any two 1-forms in $\mathbf{R}^n$, and 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

$$\begin{aligned} \alpha \wedge \beta &= \sum_{i=1}^{n} \sum_{j<i}^{n} (a_i b_j - a_j b_i)(dx^i \wedge dx^j), \\ &= \frac{1}{2}(a_i b_j - a_j b_i)(dx^i \wedge dx^j). \end{aligned}$$

In particular, if $n = 3$, the reader will recognize the coefficients of the wedge product as the components of the cross product of $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, as shown earlier.

Remark Quantities such as $dx\, dy$ and $dy\, dz$ which often appear in calculus II, 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 shown later in equation 2.83.

2.2.5 Determinants

The properties of n-forms are closely related to determinants, so it might be helpful to digress a bit and review the fundamentals of determinants, as found in any standard linear algebra textbook such as [16]. Let $A \in M_n$ be an $n \times n$ matrix with column vectors

$$A = [\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n]$$

2.2.13 Definition A function $f : M_n \to \mathbf{R}$ is called multilinear if it is linear on each slot; that is,

$$f[\mathbf{v}_1, \ldots, a_1\mathbf{v}_i + a_2\mathbf{v}_j, \ldots, \mathbf{v}_n] = a_1 f[\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n] + a_2 f[\mathbf{v}_1, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_n].$$

2.2.14 Definition A function $f : M_n \to R$ is called alternating if it changes sign whenever any two columns are switched; that is,

$$f[\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_n] = -f[\mathbf{v}_1, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_i, \ldots \mathbf{v}_n]$$

2.2.15 Definition A determinant function is a map $D : M_n \to \mathbf{R}$ that is
  a) Multilinear,
  b) Alternating,
  c) $D(I) = 1$.

One can then prove that this defines the function uniquely. In particular, if $A = (a^i{}_j)$, the determinant can be expressed as

$$\det(A) = \sum_{\pi} \text{sgn}(\pi)\; a^1_{\pi(1)} a^2_{\pi(2)} \ldots a^n_{\pi(n)}, \tag{2.39}$$

where the sum is over all the permutations of $\{1, 2 \ldots, n\}$. The determinant can also be calculated by the cofactor expansion formula of Laplace. Thus, for example, the cofactor expansion along the entries on the first row $(a^1{}_j)$, is given by

$$\det(A) = \sum_k a^1{}_k \Delta^k{}_1, \tag{2.40}$$

where Δ is the cofactor matrix.
At this point it is convenient to introduce the totally antisymmetric *Levi-Civita* permutation symbol defined as follows:

$$\epsilon_{i_1 i_2 \ldots i_k} = \begin{cases} +1 & \text{if } (i_1, i_2, \ldots i_k) \text{ is an even permutation of } (1, 2, \ldots, k) \\ -1 & \text{if } (i_1, i_2, \ldots, i_k) \text{ is an odd permutation of } (1, 2, \ldots, k) \\ 0 & \text{otherwise} \end{cases} \tag{2.41}$$

In dimension 3, there are only 6 ($3! = 6$) non-vanishing components of ϵ_{ijk}, namely,

$$\begin{aligned} \epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1 \\ \epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1 \end{aligned} \tag{2.42}$$

We set the Levi-Civita symbol with some or all the indices up, numerically equal to the permutation symbol will all the indices down. The permutation symbols are useful in the theory of determinants. In fact, if $A = (a^i{}_j)$ is an $n \times n$ matrix, then, equation (2.39) can be written as,

$$\det A = |A| = \epsilon^{i_1 i_2 \dots i_n} a^1{}_{i_1} a^2{}_{i_2} \dots a^n{}_{i_n}. \tag{2.43}$$

Thus, for example, for a 2×2 matrix,

$$\begin{aligned} A &= \begin{vmatrix} a^1{}_1 & a^1{}_2 \\ a^2{}_1 & a^2{}_2 \end{vmatrix}, \\ \det(A) &= \epsilon^{ij} a^1{}_i a^2{}_j, \\ &= \epsilon^{12} a^1{}_1 a^2{}_2 + \epsilon^{21} a^1{}_2 a^2{}_1, \\ &= a^1{}_1 a^2{}_2 - a^1{}_2 a^2{}_1. \end{aligned}$$

We also introduce the generalized Kronecker delta symbol

$$\delta^{i_1 i_2 \dots i_k}_{j_1 j_2 \dots j_k} = \begin{cases} +1 & \text{if } (i_1, i_2, \dots, i_k) \text{ is an even permutation of } (j_1, j_2, \dots, j_k) \\ -1 & \text{if } (i_1, i_2, \dots, i_k) \text{ is an odd permutation of } (j_1, j_2, \dots, j_k) \\ 0 & \text{otherwise} \end{cases} \tag{2.44}$$

If one views the indices i_k as labelling rows and j_k as labelling columns of a matrix, we can represent the completely antisymmetric symbol by the determinant,

$$\delta^{i_1 i_2 \dots i_k}_{j_1 j_2 \dots j_k} = \begin{vmatrix} \delta^{i_1}_{j_1} & \delta^{i_1}_{j_2} & \dots & \delta^{i_1}_{j_k} \\ \delta^{i_2}_{j_1} & \delta^{i_2}_{j_2} & \dots & \delta^{i_2}_{j_k} \\ \dots & \dots & \dots & \dots \\ \delta^{i_k}_{j_1} & \delta^{i_k}_{j_2} & \dots & \delta^{i_k}_{j_k} \end{vmatrix} \tag{2.45}$$

Not surprisingly, the generalized Kronecker delta is related to a product of Levi-Civita symbols by the equation

$$\epsilon^{i_1 i_2 \dots i_k} \epsilon_{j_1 j_2 \dots j_k} = \delta^{i_1 i_2 \dots i_k}_{j_1 j_2 \dots j_k}, \tag{2.46}$$

which is evident since both sides are completely antisymmetric. In dimension 3, the only non-zero components of δ^{ij}_{kl} are,

$$\begin{aligned} \delta^{12}_{12} = \delta^{13}_{13} = \delta^{23}_{23} = 1 \qquad \delta^{12}_{21} = \delta^{13}_{31} = \delta^{23}_{32} = -1 \\ \delta^{21}_{21} = \delta^{31}_{31} = \delta^{32}_{32} = 1 \qquad \delta^{21}_{12} = \delta^{31}_{13} = \delta^{32}_{23} = -1. \end{aligned}$$

2.2.16 Proposition In dimension 3 the following identities hold

a) $\epsilon^{ijk}_{imn} = \delta^{jk}_{mn} = \delta^j_m \delta^k_n - \delta^j_n \delta^k_m$,

b) $\epsilon^{ijk}_{ijn} = 2\delta^k_n$,

c) $\epsilon^{ijk}_{ijk} = 3!$

Proof For part (a), we compute the determinant by cofactor expansion on the first row

$$\begin{aligned}
\epsilon^{ijk}_{imn} &= \begin{vmatrix} \delta^i_i & \delta^i_m & \delta^i_n \\ \delta^j_i & \delta^j_m & \delta^j_n \\ \delta^k_i & \delta^k_m & \delta^k_n \end{vmatrix} \\
&= \delta^i_i \begin{vmatrix} \delta^j_m & \delta^j_n \\ \delta^k_m & \delta^k_n \end{vmatrix} - \delta^i_m \begin{vmatrix} \delta^j_i & \delta^j_n \\ \delta^k_i & \delta^k_n \end{vmatrix} + \delta^i_n \begin{vmatrix} \delta^j_i & \delta^j_m \\ \delta^k_i & \delta^k_m \end{vmatrix} \\
&= 3 \begin{vmatrix} \delta^j_m & \delta^j_n \\ \delta^k_m & \delta^k_n \end{vmatrix} - \begin{vmatrix} \delta^j_m & \delta^j_n \\ \delta^k_m & \delta^k_n \end{vmatrix} + \begin{vmatrix} \delta^j_n & \delta^j_m \\ \delta^k_n & \delta^k_m \end{vmatrix} \\
&= (3-1-1) \begin{vmatrix} \delta^j_m & \delta^j_n \\ \delta^k_m & \delta^k_n \end{vmatrix} = \begin{vmatrix} \delta^j_m & \delta^j_n \\ \delta^k_m & \delta^k_n \end{vmatrix}
\end{aligned}$$

Here we used the fact that the contraction δ^i_i is just the trace of the identity matrix and the observation that we had to transpose columns in the last determinant in the next to last line. for part (b) follows easily for part(a), namely,

$$\begin{aligned}
\epsilon^{ijk}_{inj} &= \delta^{jk}_{jn}, \\
&= \delta^j_j \delta^k_n - \delta^j_n \delta^k_j, \\
&= 3\delta^k_n - \delta^k_n, \\
&= 2\delta^k_n.
\end{aligned}$$

From this, part (c) is obvious. With considerably more effort, but inductively following the same scheme, one can establish the general formula,

$$\epsilon^{i_1 \ldots i_k, i_{k+1} \ldots i_n} \epsilon_{i_1 \ldots i_k, j_{k+1} \ldots j_n} = k! \delta^{i_{k+1} \ldots i_n}_{j_{k+1} \ldots j_n}. \tag{2.47}$$

2.2.6 Vector Identities

The permutation symbols are very useful in establishing and manipulating classical vector formulas. We present here a number of examples. For this purpose, let,

$$\begin{aligned}
\mathbf{a} &= a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k}, \\
\mathbf{b} &= b_1 \mathbf{i} + b_2 \mathbf{j} + b_3 \mathbf{k}, \\
\mathbf{c} &= c_1 \mathbf{i} + c_2 \mathbf{j} + c_3 \mathbf{k}, \\
\mathbf{d} &= d_1 \mathbf{i} + d_2 \mathbf{j} + d_3 \mathbf{k},
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
\alpha &= a_1\, dx^1 + a_2\, dx^2 + a_3\, dx^3, \\
\beta &= b_1\, dx^1 + b_2\, dx^2 + b_3\, dx^3, \\
\gamma &= c_1\, dx^1 + c_2\, dx^2 + c_3\, dx^3, \\
\delta &= d_1\, dx^1 + d_2\, dx^2 + d_3\, dx^3,
\end{aligned}$$

1. Dot product and cross product

$$\mathbf{a} \cdot \mathbf{b} = \delta^{ij} a_i b_j = a_i b^i, \qquad (\mathbf{a} \times \mathbf{b})_k = \epsilon_k{}^{ij} a_i b_j \tag{2.48}$$

2. Wedge product

$$\alpha \wedge \beta = \epsilon^k{}_{ij} (\mathbf{a} \times \mathbf{b})_k \, dx^i \wedge dx^j. \tag{2.49}$$

3. Triple product

$$
\begin{aligned}
\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= \delta_{ij}a^i(\mathbf{b}\times\mathbf{c})^l, \\
&= \delta_{ij}a^i\epsilon^j{}_{kl}b^kc^l, \\
&= \epsilon_{ikl}a^ib^kc^l, \\
\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= \det([\mathbf{abc}]), && (2.50) \\
&= (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c} && (2.51)
\end{aligned}
$$

4. Triple cross product: bac-cab identity

$$
\begin{aligned}
[\mathbf{a}\times(\mathbf{b}\times\mathbf{c})]_l &= \epsilon_l{}^{mn}a_m(\mathbf{b}\times\mathbf{c})_n \\
&= \epsilon_l{}^{mn}a_m(\epsilon_n{}^{jk}b_jc_k) \\
&= \epsilon_l{}^{mn}\epsilon_n{}^{jk}a_mb_jc_k) \\
&= \epsilon_{mnl}\epsilon^{jkn}a^mb_jc_k) \\
&= (\delta^k_m\delta^j_l - \delta^j_l\delta^k_m)a^mb_jc_k \\
&= b_la^mc_m - c_la^mb_m.
\end{aligned}
$$

Rewriting in vector form

$$
\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b}). \tag{2.52}
$$

5. Dot product of cross products

$$
\begin{aligned}
(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) &= \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}\times\mathbf{d}), \\
&= \mathbf{a}\cdot[\mathbf{c}(\mathbf{b}\cdot\mathbf{d}) - \mathbf{d}(\mathbf{b}\cdot\mathbf{c})] \\
&= (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}),
\end{aligned}
$$

$$
(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) = \begin{vmatrix} \mathbf{a}\cdot\mathbf{c} & \mathbf{a}\cdot\mathbf{d} \\ \mathbf{b}\cdot\mathbf{c} & \mathbf{b}\cdot\mathbf{d} \end{vmatrix} \tag{2.53}
$$

6. Norm of cross-product

$$
\begin{aligned}
\|\mathbf{a}\times\mathbf{b}\|^2 &= (\mathbf{a}\times\mathbf{b})\cdot(\mathbf{a}\times\mathbf{b}), \\
&= \begin{vmatrix} \mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b} \\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b}, \end{vmatrix}, \\
&= \|\mathbf{a}||^2\|\mathbf{b}\| - (\mathbf{a}\cdot\mathbf{b})^2 && (2.54)
\end{aligned}
$$

7. More wedge products. Let $C = c^k\frac{\partial}{\partial x^k}, \quad D = d^m\frac{\partial}{\partial x^m}$. Then,

$$
\begin{aligned}
(\alpha\wedge\beta)(C,D) &= \begin{vmatrix} \alpha(C) & \alpha(D) \\ \beta(C) & \beta(D) \end{vmatrix}, \\
&= \begin{vmatrix} \mathbf{a}\cdot\mathbf{c} & \mathbf{a}\cdot\mathbf{d} \\ \mathbf{b}\cdot\mathbf{c} & \mathbf{b}\cdot\mathbf{d} \end{vmatrix} && (2.55)
\end{aligned}
$$

8. Grad, Curl, Div in $\mathbf{R}^3$

   Let $\nabla_i = \frac{\partial}{\partial x^i}$, $\nabla^i = \delta^{ij}\nabla_j$, $\mathbf{A} = \mathbf{a}$ and define

   ◇ $(\nabla f)_i = \nabla_i f$

   ◇ $(\nabla \times \mathbf{A})_i = \epsilon_i{}^{jk}\nabla_j a_k$

   ◇ $\nabla \cdot \mathbf{A} = \delta^{ij}\nabla_i a_j = \nabla^j a_j$

   ◇ $\nabla \cdot \nabla(f) \equiv \nabla^2 f = \nabla^i \nabla_i f$

   (a)

$$\begin{aligned}(\nabla \times \nabla f)_i &= \epsilon_i{}^{jk}\nabla_j \nabla_f = 0,\\ \nabla \times \nabla f &= 0\end{aligned} \tag{2.56}$$

   (b)

$$\begin{aligned}\nabla \cdot (\nabla \times \mathbf{A}) &= \delta^{ij}\nabla_i(\nabla \times \mathbf{a})_j,\\ &= \delta^{ij}\nabla_i \epsilon_j{}^{kl}\nabla_k a_l,\\ &= \epsilon^{jkl}\nabla_i \nabla_j a_k,\\ \nabla \cdot (\nabla \times \mathbf{A}) &= 0\end{aligned} \tag{2.57}$$

   where in the last step in the two items above we use the fact that a contraction of two symmetric with two antisymmetric indices is always 0.

   (c) The same steps as in the bac-cab identity give

$$\begin{aligned}[\nabla \times (\nabla \times \mathbf{A})]_l &= \nabla_l(\nabla^m a_m) - \nabla^m \nabla_m a_l,\\ \nabla \times (\nabla \times \mathbf{A}) &= \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A},\end{aligned}$$

   where $\nabla^2 \mathbf{A}$ means the Laplacian of each component of $\mathbf{A}$.

   This last equation is crucial in the derivation of the wave equation for light from Maxwell's equations for the electromagnetic field.

2.2.7 n-Forms

2.2.17 Definition Let $\alpha^1, \alpha^2, \alpha^2$, be one forms, and $X_1, X_2, X_3 \in \mathscr{X}$. Let π be the set of permutations of $\{1, 2, 3\}$. Then

$$\begin{aligned}(\alpha^1 \wedge \alpha^2 \wedge \alpha^3)(X_1, X_2, X_3) &= \sum_\pi \operatorname{sign}(\pi)\alpha^1(X_{\pi(1)})\alpha^2(X_{\pi(2)})\alpha^3(X_{\pi(3)}),\\ &= \epsilon^{ijk}\alpha^1(X_i)\alpha^2(X_j)\alpha^3(X_k).\end{aligned}$$

This trilinear map is an example of a alternating covariant 3-tensor.

2.2.18 Definition A 3-form ϕ in $\mathbf{R}^n$ is an alternating, covariant 3-tensor. In local coordinates, a 3-from can be written as an object of the following type

$$\phi = A_{ijk} dx^i \wedge dx^j \wedge dx^k \tag{2.58}$$

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. There is nothing really wrong with using definition (2.58). This definition however, is coordinate-dependent and differential geometers prefer coordinate-free definitions, theorems and proofs.

We can easily extend the concepts above to higher order forms.

2.2.19 Definition Let $T^0_k(\mathbf{R}^n)$ be the set multilinear maps

$$t : \underbrace{T(\mathbf{R}) \times ... \times T(\mathbf{R})}_{\text{k times}} \to \mathbf{R}$$

from k copies of $T(\mathbf{R})$ to $\mathbf{R}$. The map t is called *skew-symmetric* if

$$t(e_1, \dots, e_k) = \text{sign}(\pi) t(e_{\pi(1)}, \dots, e_{\pi(k)}), \tag{2.59}$$

where π is the set of permutations of $\{1, \dots, k\}$. A skew-symmetry covariant tensor of rank k at p, is called a *k-form* at p. denote by $\Lambda^k_{(p)}(\mathbf{R}^n)$ the space of k-forms at $p \in \mathbf{R}^n$. This vector space has dimension

$$\dim \Lambda^k_p(\mathbf{R}^n) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

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

$$\Lambda^k(\mathbf{R}^n) = \bigcup_p \Lambda^k_p(\mathbf{R}^n).$$

Sections of the bundle are called *k-forms* and the space of all sections is denoted by

$$\Omega^k(\mathbf{R}^n) = \Gamma(\Lambda^k(\mathbf{R}^n)).$$

A section $\alpha \in \Omega^k$ of the bundle technically should be called k-form field, but the consensus in the literature is to call such a section simply a k-form. In local coordinates, a k-form can be written as

$$\alpha = A_{i_1, \dots i_k}(x) dx^{i_1} \wedge \dots dx^{i_k}. \tag{2.60}$$

2.2.20 Definition The *alternation map* $A : T^0_k(\mathbf{R}^n) \to T^0_k(\mathbf{R}^n)$ is defined by

$$At(e_1, \dots, e_k) = \frac{1}{k!} \sum_\pi (\text{sign}\pi) t(e_{\pi(1)}, \dots, e_{\pi(k)}).$$

2.2.21 Definition If $\alpha \in \Omega^k(\mathbf{R}^n)$ and $\beta \in \Omega^l(\mathbf{R}^n)$, then

$$\alpha \wedge \beta = \frac{(k+l)!}{k!l!} A(\alpha \otimes \beta) \tag{2.61}$$

If α is a k-form and β an l-form, we have

$$\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha. \tag{2.62}$$

Now, for a little combinatorics. Factorials are unavoidable due to the permutation attributes of the wedge product. The convention here follows Marsden [20] and Spivak [34], which reduces proliferation of factorials later. Let us count the number of linearly independent differential forms in Euclidean space. More specifically, we want to find a basis for the vector space of k-forms in $\mathbf{R}^3$. As stated above, 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.
It is possible define *tensor-valued differential forms*. Let $E = T^r_s(\mathbf{R^n})$ be the tensor bundle. A tensor-valued p-form is defined as a section

$$T \in \Omega^p(\mathbf{R}^n, E) = \Gamma(E \otimes \Lambda^p(\mathbf{R}^n)).$$

In local coordinates, a tensor-valued k-form is a $\binom{r}{s+p}$ tensor

$$T = T^{i_1,\dots i_r}{}_{j_1,\dots j_s,k_1,\dots,k_p} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s} \wedge dx^{k_1} \wedge \dots \wedge dx^{k_p}. \tag{2.63}$$

Thus, for example, the quantity

$$\Omega^i{}_j = \tfrac{1}{2} R^i{}_{jkl}\, dx^k \wedge dx^l$$

would be called the components of a $\binom{1}{1}$-valued 2-form

$$\Omega = \Omega^i{}_j \frac{\partial}{\partial x^i} \otimes dx^j.$$

The notion of the wedge product can be extended to tensor-valued forms using tensor products on the tensorial indices and wedge products on the differential form indices.

2.3 Exterior Derivatives

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

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

$$d\alpha(X,Y) = X(\alpha(Y)) - Y(\alpha(X)), \tag{2.64}$$

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.

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

The definition 2.64 of a differential of a 1-form can be refined to provide a coordinate-free definition in general manifolds (see 6.28,) and it can be extended to differentials of m-forms. For now, the computation immediately above suffices to motivate the following coordinate dependent definition (for a coordinate-free definition for general manifolds, see (**??**):

2.3.2 Definition Let α be an m-form, given in coordinates as in equation (2.60). The *exterior derivative* of α is the $(m+1)$-form $d\alpha$ given by

$$\begin{aligned} d\alpha &= dA_{i_1,\ldots i_m} \wedge dx^{i_1} \ldots dx^{i_m} \\ &= \frac{\partial A_{i_1,\ldots i_m}}{\partial x^{i_0}}(x) dx^{i_0} \wedge dx^{i_1} \ldots dx^{i_m}. \end{aligned} \tag{2.65}$$

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

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

2.3.3 Theorem

$$
\begin{array}{ll}
\text{a)} & d : \Omega^m \longrightarrow \Omega^{m+1} \\
\text{b)} & d^2 = d \circ d = 0 \\
\text{c)} & d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta \qquad \forall \alpha \in \Omega^p, \beta \in \Omega^q
\end{array}
\tag{2.66}
$$

Proof

a) Obvious from equation (2.65).

b) First, we prove the proposition for $\alpha = f \in \Omega^0$. 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 (2.60). It follows from equation 2.65, 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 $\alpha \in \Omega^p, \beta \in \Omega^q$. Then, we can write

$$
\begin{aligned}
\alpha &= A_{i_1,\ldots i_p}(x) dx^{i_1} \wedge \ldots dx^{i_p} \\
\beta &= B_{j_1,\ldots j_q}(x) dx^{j_1} \wedge \ldots dx^{j_q}.
\end{aligned}
\tag{2.67}
$$

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

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

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

2.3.1 Pull-back

2.3.4 Definition Let $F : \mathbf{R}^n \to \mathbf{R}^m$ be a 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* $F^* : \Omega^k_{(F(x))} \to \Omega^k_{(x)}$ defined by

$$(F^*\alpha)_x(X_1, \ldots X_k) = \alpha_{F(x)}(F_*X_1, \ldots F_*X_k), \tag{2.68}$$

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.

2.3.5 Theorem

$$\begin{array}{lrcl} \text{a)} & F^*(g\alpha_1) & = & (g \circ F)\, F^*\alpha, \\ \text{b)} & F^*(\alpha_1 + \alpha_2) & = & F * \alpha_1 + F^*\alpha_2, \\ \text{c)} & F^*(\alpha \wedge \beta) & = & F^*\alpha \wedge F^*\beta, \\ \text{d)} & F^*(d\alpha) & = & d(F^*\alpha.) \end{array} \tag{2.69}$$

Part (d) is encapsulated in the commuting diagram in figure 2.2.

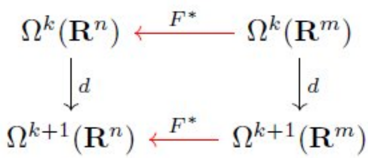


Fig. 2.2: $d\, F^* = F^*\, d$

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, $F^*(dg) = d(F^*g)$ 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 $y^k = f^k(x^i)$. Then the pullback of the form dg given the formula above $F^*dg = d(g\circ F)$ is given in local coordinates by the chain rule

$$F^*dg = \frac{\partial g}{\partial x^j}dx^j.$$

In particular, the pull-back of local coordinate functions is given by

$$F^*(dy^i) = \frac{\partial y^i}{\partial x^j}dx^j. \tag{2.70}$$

Thus, pullback for the basis 1-forms dy^k is yet another manifestation of the differential as a linear map represented by the Jacobian

$$dy^k = \frac{\partial y^k}{\partial x^i}\,dx^i. \tag{2.71}$$

In particular, if $m = n$,

$$\begin{aligned} d\Omega &= dy^1\wedge dy^2\wedge\ldots\wedge dy^n,\\ &= \frac{\partial y^1}{\partial x^{i_1}}\frac{\partial y^2}{\partial x^{i_2}}\cdots\frac{\partial y^n}{\partial x^{i_n}}\,dx^{i_1}\wedge dx^{i_2}\wedge\ldots dx^{i_n},\\ &= \epsilon^{i_1 i_2\ldots i_n}\frac{\partial y^1}{\partial x^{i_1}}\frac{\partial y^2}{\partial x^{i_2}}\cdots\frac{\partial y^n}{\partial x^{i_n}}\,dx^1\wedge dx^2\wedge\ldots dx^n,\\ &= |J|\wedge dx^1\wedge\ldots dx^n. \end{aligned} \tag{2.72}$$

So, the pull-back of the volume form,

$$F^*d\,\Omega = |J|\;dx^1\wedge\ldots\wedge dx^n,$$

gives rise to the integrand that appears in the change of variables theorem for integration. More explicitly, let $R\in\mathbf{R}^n$ be a simply connected region, F be a mapping $F: R\in\mathbf{R}^n\to\mathbf{R}^m$, with $m\geq n$. If ω is a k- form in $\mathbf{R}^m$, then

$$\int_{F(R)}\omega = \int_R F^*\omega \tag{2.73}$$

We refer to this formulation of the change of variables theorem as integration by pull-back.

If $F : \mathbf{R}^n \to \mathbf{R}^n$ is a diffeomorphism, one can push-forward forms with the inverse of the pull-back $F_* = (F^{-1})^*$.

2.3.6 Example Line Integrals
Let $\omega = f_i\, dx^i$ be a one form in $\mathbf{R}^3$ and let C be the curve given by the mapping $\phi : I = t \in [a,b] \to \mathbf{x}(t) \in \mathbf{R}^3$. We can write $\omega = \mathbf{F} \cdot d\mathbf{x}$, where $\mathbf{F} = (f_1, f_2, f_3)$ is a vector field. Then the integration by pull-back equation 2.73 reads,

$$\begin{aligned}
\int_C \mathbf{F} \cdot d\mathbf{x} &= \int_C \omega, \\
&= \int_{\phi(I)} \omega, \\
&= \int_I \phi^* \omega, \\
&= \int_I f^i(\mathbf{x}(t)) \frac{dx^i}{dt}\, dt, \\
&= \int_I \mathbf{F}(\mathbf{x}(\mathbf{t})) \frac{d\mathbf{x}}{dt}\, dt
\end{aligned}$$

This coincides with the definition of line integrals as introduced in calculus.

2.3.7 Example Polar Coordinates
Let $x = r\cos\theta$ and $y = r\sin\theta$ and $f = f(x,y)$. Then

$$\begin{aligned}
dx \wedge dy &= (-r\sin\theta d\theta + \cos\theta dr) \wedge (r\cos\theta d\theta + \sin\theta dr), \\
&= -r\sin^2\theta d\theta \wedge dr + r\cos^2\theta dr \wedge d\theta, \\
&= (r\cos^2\theta + r\sin^2\theta)(dr \wedge d\theta), \\
&= r(dr \wedge d\theta). \\
\int\!\!\int f(x,y) dx \wedge dy &= \int\!\!\int f(x(r,\theta), y(r,\theta))\, r(dr \wedge d\theta). \qquad (2.74)
\end{aligned}$$

In this case, the element of arc length is diagonal

$$ds^2 = dr^2 + r^2 d\theta^2,$$

as it should be for an orthogonal change of variables. The differential of area is

$$\begin{aligned}
dA &= \sqrt{\det g}\, dr \wedge d\theta, \\
&= r(dr \wedge d\theta)
\end{aligned}$$

If the polar coordinates map is denoted by $F : \mathbf{R}^2 \to \mathbf{R}^2$, then equation 2.74 is just the explicit expression for the pullback of $F^*(\, f\, dA)$.

2.3.8 Example Polar coordinates are just a special example of the general transformation in $\mathbf{R}^2$ given by,

$$x = x(u,v), \qquad dx = \frac{\partial x}{\partial u}\,du + \frac{\partial x}{\partial u}\,dv,$$
$$y = y(u,v), \qquad dy = \frac{\partial y}{\partial u}\,du + \frac{\partial y}{\partial u}\,dv,$$

for which

$$\phi * (dx \wedge dy) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} du \wedge dv \tag{2.75}$$

2.3.9 Example Surface Integrals

Let $R \in \mathbf{R}^2$ be a simply connected region with boundary δR and let the mapping

$$\phi : (u,v) \in R \longrightarrow \mathbf{x}(u^\alpha) \in \mathbf{R}^2$$

describe a surface S with boundary $C = \phi(\delta R)$. Here, $\alpha = 1, 2$, with $u = u^1,\ v = u^2$. Given a vector field $\mathbf{F} = (f_1, f_2, f_3)$, we assign to it the 2-form

$$\begin{aligned} \omega &= \mathbf{F} \cdot dS, \\ &= f_1\,dx^2 \wedge dx^3 - f_2\,dx^1 \wedge dx^3 + f_3\,dx^1 \wedge dx^2, \\ &= \epsilon^i_{jk} f_i\,dx^j \wedge dx^k. \end{aligned}$$

Then,

$$\begin{aligned} \int\!\!\int_S \mathbf{F} \cdot d\mathbf{S} &= \int\!\!\int_S \omega, \\ &= \int\!\!\int_R \phi^* \omega, \\ &= \int\!\!\int_R \epsilon^i{}_{jk} f_i \frac{\partial x^j}{\partial u^\alpha} du^\alpha \wedge \frac{\partial x^k}{\partial u^\beta} du^\beta, \\ &= \int\!\!\int_R \mathbf{F} \cdot (\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial u})\,du \wedge dv \end{aligned}$$

We elaborate a bit on this slick computation, for the benefit of those readers who may have gotten got lost in the index manipulation.

$$\begin{aligned} \int\!\!\int_S \mathbf{F} \cdot d\mathbf{S} &= \int\!\!\int_S \omega, \\ &= \int\!\!\int_R \phi^* \omega, \\ &= \int\!\!\int_R \left[f_1\,\phi*(dx^2 \wedge dx^3) - f_2\,\phi*(dx^1 \wedge dx^3) + f_3\,\phi*(dx^1 \wedge dx^2) \right], \\ &= \int\!\!\int_R \left[f_1 \begin{vmatrix} \frac{\partial x^2}{\partial u} & \frac{\partial x^2}{\partial v} \\ \frac{\partial x^3}{\partial u} & \frac{\partial x^3}{\partial v} \end{vmatrix} - f_2 \begin{vmatrix} \frac{\partial x^1}{\partial u} & \frac{\partial x^1}{\partial v} \\ \frac{\partial x^3}{\partial u} & \frac{\partial x^3}{\partial v} \end{vmatrix} + f_3 \begin{vmatrix} \frac{\partial x^1}{\partial u} & \frac{\partial x^1}{\partial v} \\ \frac{\partial x^2}{\partial u} & \frac{\partial x^2}{\partial v} \end{vmatrix} \right] du \wedge dv \\ &= \int\!\!\int_R \mathbf{F} \cdot \left(\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial u} \right) du \wedge dv \end{aligned}$$

This pull-back formula for surface integrals is how most students are introduced to this subject in the third semester of calculus.

2.3.10 Remark

1. 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.

2. As shown in equation 2.32 the metric in spherical coordinates is given by

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

so the differential of volume is

$$\begin{aligned} dV &= \sqrt{\det g}\; dr \wedge d\theta \wedge d\phi, \\ &= r^2 \sin\theta\; dr \wedge d\theta \wedge d\phi. \end{aligned}$$

2.3.2 Stokes' Theorem in $\mathbf{R}^n$

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

$$\begin{aligned} d\alpha &= (\tfrac{\partial P}{\partial x} dx + \tfrac{\partial P}{\partial y}) \wedge dx + (\tfrac{\partial Q}{\partial x} dx + \tfrac{\partial Q}{\partial y}) \wedge dy \\ &= \tfrac{\partial P}{\partial y}\, dy \wedge dx + \tfrac{\partial Q}{\partial x}\, dx \wedge dy \\ &= (\tfrac{\partial Q}{\partial x} - \tfrac{\partial P}{\partial y})\, dx \wedge dy. \end{aligned} \tag{2.76}$$

This example is related to Green's theorem in $\mathbf{R}^2$. For convenience, we include here a proof of Green's Theorem in a special case. We say that a region D

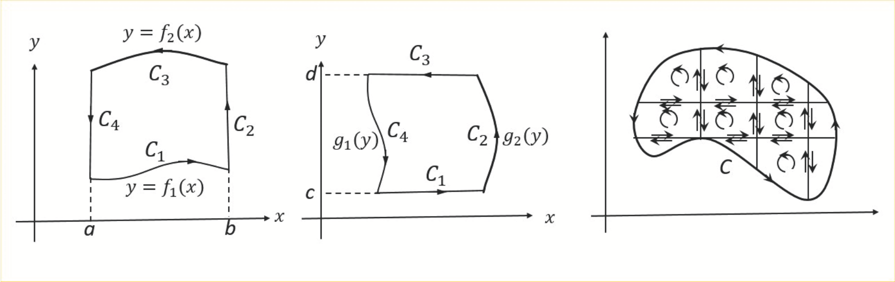


Fig. 2.3: Simple closed curve.

in the plane is of *type I* if it is enclosed between the graphs of two continuous functions of x. The region inside the simple closed curve in figure 2.3 bounded by $f_1(x)$ and $f_2(x)$, between a and b, is a region of type I. A region in the plane is of *type II* if it lies between two continuous functions of y. The region in 2.3 bounded between $c \leq y \leq d$, would be a region of type II.

2.3.11 Green's theorem

Let C be a simple closed curve in the xy-plane and let $\partial P/\partial x$ and $\partial Q/\partial y$ be continuous functions of (x, y) inside and on C. Let R be the region inside the closed curve so that the boundary $\delta R = C$. Then

$$\oint_{\delta R} P\ dx + Q\ dy = \int\int_R \left[\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right] dA. \tag{2.77}$$

We first prove that for a type I region such as the one bounded between a and b shown in 2.3, we have

$$\oint_C P\ dx = -\int\int_D \frac{\partial P}{\partial y}\ dA \tag{2.78}$$

Where C comprises the curves C_1, C_2, C_3 and C_4. By the fundamental theorem of calculus, we have on the right,

$$\begin{aligned}\int\int_D \frac{\partial P}{\partial y}\ dA &= \int_a^b \int_{f_1(x)}^{f_2(x)} \frac{\partial P}{\partial y}\ dydx,\\ &= \int_a^b [P(x, f_2(x)) - P(x, f_1(x))]\ dx.\end{aligned}$$

On the left, the integrals along C_2 and C_4 vanish, since there is no variation on x. The integral along C_3 is traversed in opposite direction of C_1, so we have,

$$\begin{aligned}\oint_C P(x,y)\ dx &= \int_{C_1} + \int_{C_2} + \int_{C_3} + \int_{C_4} P(x,y)\ dx,\\ &= \int_{C_1} P(x,y)\ dx - \int_{C_3} P(x,y)\ dx,\\ &= \int_a^b P(x, f_1(x))\ dx - \int_a^b P(x, f_2(x))\ dx\end{aligned}$$

This establishes the veracity of equation 2.78 for type I regions. By a completely analogous process on type II regions, we find that

$$\oint_C Q\ dy = \int\int_D \frac{\partial Q}{\partial x}\ dA. \tag{2.79}$$

The theorem follows by subdividing R into a grid of regions of both types, all oriented in the same direction as shown on the right in figure 2.3. Then one applies equations 2.78 or 2.79, as appropriate, for each of the subdomains. All contributions from internal boundaries cancel since each is traversed twice, each in opposite directions. All that remains of the line integrals is the contribution along the boundary δR.

Let $\alpha = P\ dx + Q\ dy$. Comparing with equation 2.76, we can write Green's theorem in the form

$$\int_C \alpha = \int\int_D d\alpha. \tag{2.80}$$

It is possible to extend Green's Theorem to more complicated regions that are not simple connected. Green's theorem is a special case in dimension of two of Stoke's theorem.

2.3.12 Stokes' theorem

If ω is a C^1 one form in $\mathbf{R}^n$ and S is C^2 surface with boundary $\delta S = C$, then

$$\int_{\delta S} \omega = \int\int_S d\omega. \tag{2.81}$$

Proof The proof can be done by pulling back to the uv-plane and using the chain rule, thus allowing us to use Green's theorem. Let $\omega = f_i\, dx^i$ and S be parametrized by $x^i = x^i(u^\alpha)$, where $(u^1, u^2) \in R \subset \mathbf{R}^2$. We assume that the boundary of R is a simple closed curve. Then

$$\begin{aligned}
\int_C \omega &= \int_{\delta S} f_i\, dx^i, \\
&= \int_{\delta R} f_i \frac{\partial x^i}{\partial u^\alpha}\, du^\alpha, \\
&= \int\int_R \frac{\partial}{\partial u^\beta}(f_i \frac{\partial x^i}{\partial u^\alpha})\, du^\beta \wedge du^\alpha, \\
&= \int\int_R \left[\frac{\partial f_i}{\partial x^k}\frac{\partial x^k}{\partial u^\beta}\frac{\partial x^i}{\partial u^\alpha} + f_i \frac{\partial^2 x^i}{\partial u^\beta \partial u^\alpha}\right] du^\beta \wedge du^\alpha, \\
&= \int\int_R \left[\frac{\partial f_i}{\partial x^k}\frac{\partial x^k}{\partial u^\beta}\frac{\partial x^i}{\partial u^\alpha}\right] du^\beta \wedge du^\alpha, \\
&= \int\int_R \left[\frac{\partial f_i}{\partial x^k}\frac{\partial x^k}{\partial u^\beta}\right] du^\beta \wedge \left[\frac{\partial x^i}{\partial u^\alpha}\right] du^\alpha \\
&= \int\int_S \frac{\partial f_i}{\partial x^k}\, dx^k \wedge dx^i = \int\int_S df_i \wedge dx^i \\
&= \int\int_S d\omega.
\end{aligned}$$

We present a less intuitive but far more elegant proof. The idea is formally the same, namely, we pull-back to the plane by formula 2.73, apply Green's theorem in the form given in equation 2.80, and then use the fact that the pull-back commutes with the differential as in theorem 2.69.

Let $\phi : R \subset \mathbf{R}^2 \to S$ denote the surface parametrization map. Assume that $\phi^{-1}(\delta S) = \delta(\phi^{-1} S)$, that is, the inverse of the boundary of S is the boundary of the domain R. Then,

$$\begin{aligned}
\int_{\delta S} \omega &= \int_{\phi^{-1}(\delta S)} \phi^* \omega = \int_{\delta(\phi^{-1} S)} \phi^* \omega, \\
&= \int \int_{\phi^{-1} S} d(\phi^* \omega), \\
&= \int \int_{\phi^{-1} S} \phi^* (d\omega), \\
&= \int_S d\omega.
\end{aligned}$$

The proof of Stokes' theorem presented here is one of those cases mentioned in the preface, where we have simplified the mathematics for the sake of clarity. Among other things, a rigorous proof requires one to quantify what is meant by the boundary (δS) of a region. The process involves either introducing *simplices* (generalized segments, triangles, tetrahedra...) or *singular cubes* (generalized segments, rectangles, cubes...). The former are preferred in the treatment of homology in algebraic topology, but the latter are more natural to use in the context of integration on manifolds with boundary. A singular n-cube in $\mathbf{R}^n$ is the image under a continuous map,

$$I^n : [0,1]^n \to \mathbf{R}^n,$$

of the Cartesian product of n copies of the unit interval $[0,1]$. The idea is to divide the region S into formal finite sums of singular cubes, called *chains*. One then introduces a boundary operator δ, that maps a singular n-cube and hence n-chain, into an $(n-1)$-singular cube or $(n-1)$-chain. Thus, in $\mathbf{R}^3$ for example, the boundary of a cube, is the sum $\sum c_i F_i$ of the six faces with a judicious choice of coefficients $c_i \in \{-1, 1\}$. With an appropriate scheme to label faces of singular cube and a corresponding definition of the boundary map, one proves that $\delta \circ \delta = 0$. For a thorough treatment, see the beautiful book *Calculus on Manifolds* by M. Spivak [33].

Closed and Exact forms

2.3.13 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.

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

2.3.15 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. The converse need not be true. The standard counterexample is the form,

$$\omega = \frac{-y\,dx + x\,dy}{x^2 + y^2} \tag{2.82}$$

A short computation shows that $d\omega = 0$, so ω is closed. Let $\theta = \tan^{-1}(y/x)$ be the angle in polar coordinates. One can recognize that $\omega = d\theta$, but this is only true in $\mathbf{R}^2 - L$, where L is the non-negative x-axis, $L = \{(x, 0) \in \mathbf{R}^2 | x \geq 0\}$. If one computes the line integral from $(-1, 0)$ to $(1, 0)$ along the top half of the unit circle, the result is π. But the line integral along the bottom half of the unit circle gives $-\pi$. The integral is therefore not path independent, so $\omega \neq d\theta$ on any region that contains the origin. If one tries to find another C^1 function f such that $\omega = df$, one can easily show that $f = \theta + \text{const}$, which is not possible along L.

On the other hand, if one imposes the topological condition that the space is contractible, then the statement is true. A *contractible* space is one that can be deformed continuously to an interior point. We have the following,

2.3.16 Poincaré Lemma. In a contractible space (such as $\mathbf{R}^n$), if a differential form is closed, then it is exact.

To prove this lemma we need much more machinery than we have available at this point. We present the proof in **??**.

2.4 The Hodge $\star$ Operator

2.4.1 Dual Forms

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 $\mathbf{R}^4$. As a good example of this, consider the tangent space $T_p\mathbf{R}^3$. 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 a regular Euclidean vector (a^1, a^2, a^3).

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

$$\begin{aligned} dx &\longmapsto dy \wedge dz \\ dy &\longmapsto -dx \wedge dz \\ dz &\longmapsto dx \wedge dy \end{aligned} \tag{2.83}$$

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

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

it must be true that

$$\Lambda_p^k(\mathbf{R}^n) \cong \Lambda_p^{n-k}(\mathbf{R}^n). \tag{2.84}$$

To describe the isomorphism between these two spaces, we introduce the following generalization of determinants,

2.4.1 Definition . Let $\phi : \mathbf{R}^n \to \mathbf{R}^n$ be a linear map. The unique constant $\det \phi$ such that,

$$\phi* : \Lambda^n(\mathbf{R}^n) \to \Lambda^n(\mathbf{R}^n)$$

satisfies,

$$\phi^* \omega = (\det \phi)\, \omega, \tag{2.85}$$

for all n-forms, is called the *determinant* of ϕ. This is congruent with the standard linear algebra formula 2.43, since in a particular basis, the Jacobian of a linear map is the same as the matrix the represents the linear map in that basis. Let, $g(X, Y)$ be an inner product and $\{e_1, \ldots, e_n\}$ be an orthonormal basis with dual forms $\{\theta^1, \ldots \theta^n\}$. The element of arc length is, the bilinear symmetric tensor

$$ds^2 = g_{ij}\, \theta^i \otimes \theta^j.$$

The metric then induces an n-form

$$d\Omega = \theta^1 \wedge \theta^2 \ldots \wedge \theta^n,$$

called the *volume* element. With this choice of form, the reader will recognize equation 2.85 as the integrand in the change of variables theorem for multiple integration, as in example 2.74. More generally, if $\{f_1, \ldots f_n\}$ is a positively oriented basis with dual basis $\{\phi^1, \ldots \phi^n\}$, then,

$$d\Omega = \sqrt{\det g}\, \phi^1 \wedge \ldots \wedge \phi^n. \tag{2.86}$$

2.4.2 Definition Let g be the matrix representing the components of the metric in $\mathbf{R}^n$. The *Hodge ⋆ operator* is the linear isomorphism $\star : \Lambda_p^n(\mathbf{R}^n) \longrightarrow \Lambda_p^{n-k}(\mathbf{R}^n)$ defined in standard local coordinates by the equation,

$$\star\,(dx^{i_1} \wedge \ldots \wedge dx^{i_k}) = \frac{\sqrt{\det g}}{(n-k)!} \epsilon^{i_1 \ldots i_k}{}_{i_{k+1} \ldots i_n}\, dx^{i_{k+1}} \wedge \ldots \wedge dx^{i_n}, \tag{2.87}$$

For flat Euclidean space $\sqrt{\det g} = 1$, so the factor in the definition may appear superfluous. However, when we consider more general Riemannian manifolds, we will have to be more careful with raising and lowering indices with the metric, and take into account that the Levi-Civita symbol is not a tensor but something slightly more complicated called a tensor density. Including the $\sqrt{\det g}$ is done in anticipation of this more general setting later. Since the forms $dx^{i_1} \wedge \ldots \wedge dx^{i_k}$ constitute a basis of the vector space $\Lambda^k_p(\mathbf{R}^n)$ and the $\star$ operator is assumed to be a linear map, equation (2.87) completely specifies the map for all k-forms. In particular, if the components of a dual of a form are equal to the components of the form, the tensor is called *self-dual.* Of course, this can only happen if the tensor and its dual are of the same rank.

A metric g on $\mathbf{R}^n$ induces an inner product on $\Lambda^k(\mathbf{R}^n)$ as follows. Let $\{e_1, , \ldots e_n\}$ by an orthonormal basis with dual basis $\theta^1, \ldots, \theta^n$. If $\alpha, \beta \in \Lambda^k(\mathbf{R}^n)$, we can write

$$\begin{aligned} \alpha &= a_{i_1 \ldots i_k}\ \theta^{i_1} \wedge, \ldots \theta^{i_k}, \\ \beta &= b_{j_1 \ldots j_k}\ \theta^{j_1} \wedge, \ldots \theta^{j_k} \end{aligned}$$

The induced inner product is defined by

$$< \alpha, \beta >^{(k)} = \frac{1}{k!} a_{i_1 \ldots i_k} b^{i_1 \ldots i_k}. \tag{2.88}$$

If $\alpha, \beta \in \Lambda^k(\mathbf{R}^n)$, then $\star\beta \in \Lambda^{n-k}(\mathbf{R}^n)$, so $\alpha \wedge \star\beta$ must be a multiple of the volume form. The Hodge $\star$ operator is the unique isomorphism such that

$$\alpha \wedge \star\beta = < \alpha, \beta >^{(k)}\ d\,\Omega. \tag{2.89}$$

Clearly,

$$\alpha \wedge \star\beta = \star\alpha \wedge \beta$$

When it is evident that the inner product is the induced inner product on $\Lambda^k(\mathbf{R}^n)$ the indicator (k) is often suppressed. An equivalent definition of the induced inner product of two k-forms is given by

$$< \alpha, \beta > = \int (\alpha \wedge \star\beta)\ d\,\Omega. \tag{2.90}$$

If α is a k-form and β is a $(k-1)$-form, one can define the adjoint or *co-differential* by

$$< \delta\alpha, \beta > = < \alpha, d\beta > . \tag{2.91}$$

The adjoint is given by

$$\delta = (-1)^{nk+n+1} \star d \star . \tag{2.92}$$

In particular,

$$\delta = \begin{cases} - \star d \star & \text{if } n \text{ is even} \\ (-1)^k \star\ d \star & \text{if } n \text{ is odd} \end{cases} \tag{2.93}$$

The differential maps $(k-1)$-forms to k-forms, and the co-differential maps k-forms to $(k-1)$-forms. It is also the case that $\delta \circ \delta = 0$ The combination,

$$\Delta = (d + \delta)^2 = d\delta + \delta d \tag{2.94}$$

extends the Laplacian operator to forms. It maps k-forms to k-forms. A central result in harmonic analysis is the *Hodge decomposition* theorem, that states that given any k-form ω, can be split uniquely as

$$\omega = d\alpha + \delta\beta + \gamma, \tag{2.95}$$

where $\alpha \in \Omega^{k-1}$, $\beta \in \Omega^{k+1}$, and $\Delta\gamma = 0$

2.4.3 Example Hodge operator in $\mathbf{R}^2$

In $\mathbf{R}^2$,

$$\star dx = dy \qquad \star\, dy = -dx,$$

or, if one thinks of a matrix representation of $\star : \Omega(\mathbf{R}^2) \to \Omega(\mathbf{R}^2)$ in standard basis, we can write the above as

$$\star \begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix}.$$

The reader might wish to peek at the symplectic matrix **??** in the discussion in chapter 5 on conformal mappings. Given functions $u = u(x, y)$ and $v = v(x, y)$, let $\omega = u\, dx - v\, dy$. Then,

$$\begin{aligned} d\omega &= -(u_y + v_x)\, dx \wedge dy, \\ d \star \omega &= \quad (u_x - u_y)\, dx \wedge dy, \end{aligned} \qquad \text{hence} \qquad \begin{aligned} d\omega = 0 &\Rightarrow u_y = -v_x, \\ \star d\omega = 0 &\Rightarrow u_x = v_y. \end{aligned} \tag{2.96}$$

Thus, the equations $d\omega = 0$ and $d{\star}\omega = 0$ are equivalent to the Cauchy-Riemann equations for a holomorphic function $f(z) = u(x, y) + iv(x, y)$. On the other hand,

$$\begin{aligned} du &= u_x\, dx + u_y\, dy, \\ dv &= v_x\, dx + u_y\, dy, \end{aligned}$$

so the determinant of the Jacobian of the transformation $T : (x, y) \to (u, v))$, with the condition above on ω, is given by,

$$|J| = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} = u_x^2 + u_y^2 = v_x^2 + v_y^2.$$

If $|J| \neq 0$, we can set $u_x = R\cos\phi, \quad u_y = R\sin\phi$, for some R and some angle ϕ. Then,

$$|J| = \begin{vmatrix} R & 0 \\ 0 & R \end{vmatrix} \begin{vmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{vmatrix}.$$

Thus, the transformation is given by the composition of a dilation and a rotation. A more thorough discussion of this topic is found in the section of conformal maps in chapter 5.

2.4.4 Example Hodge operator in $\mathbf{R}^3$

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

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

$$\begin{aligned}
\star dx^1 &= +dx^2 \wedge dx^3, \\
\star dx^2 &= -dx^1 \wedge dx^3, \\
\star dx^3 &= +dx^1 \wedge dx^2,
\end{aligned} \tag{2.97}$$

$$\begin{aligned}
\star(dx^2 \wedge dx^3) &= dx^1, \\
\star(-dx^3 \wedge dx^1) &= dx^2, \\
\star(dx^1 \wedge dx^2) &= dx^3,
\end{aligned} \tag{2.98}$$

and

$$\star\,(dx^1 \wedge dx^2 \wedge dx^3) = 1. \tag{2.99}$$

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

$$\begin{aligned}
\star f &= f(dx^1 \wedge dx^2 \wedge dx^3), \\
&= f dV,
\end{aligned} \tag{2.100}$$

where dV is the volume form.

2.4.5 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,

$$\begin{aligned}
\star(\alpha \wedge \beta) &= (a_2 b_3 - a_3 b_2) \star (dx^2 \wedge dx^3) + (a_1 b_3 - a_3 b_1) \star (dx^1 \wedge dx^3) + \\
&\quad (a_1 b_2 - a_2 b_1) \star (dx^1 \wedge dx^2), \\
&= (a_2 b_3 - a_3 b_2) dx^1 + (a_1 b_3 - a_3 b_1) dx^2 + (a_1 b_2 - a_2 b_1) dx^3, \\
&= (\mathbf{a} \times \mathbf{b})_i \; dx^i.
\end{aligned} \tag{2.101}$$

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 (2.97) are the differential geometry versions of the well-known relations

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

2.4.6 Example In Minkowski space the collection of all 2-forms has dimension $\binom{4}{2} = 6$. The Hodge ⋆ operator in this case splits $\Omega^2(M_{1,3})$ into two 3-dim subspaces $\Omega^2_\pm$, such that $\star : \Omega^2_\pm \longrightarrow \Omega^2_\mp$.
More specifically, Ω^2_+ is spanned by the forms $\{dx^0 \wedge dx^1, dx^0 \wedge dx^2, dx^0 \wedge dx^3\}$, and Ω^2_- is spanned by the forms $\{dx^2 \wedge dx^3, -dx^1 \wedge dx^3, dx^1 \wedge dx^2\}$. The action of ⋆ on Ω^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 Ω^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 $F \in \Omega^2(M)$, we will formally write $F = F_+ + F_-$, where $F_\pm \in \Omega^2_\pm$. We would like to note that the action of the dual operator on $\Omega^2(M)$ is such that $\star : \Omega^2(M) \longrightarrow \Omega^2(M)$, and $\star^2 = -1$. In a vector space a map like ⋆, with the property $\star^2 = -1$ is called a *linear involution* of the space. In the case in question, $\Omega^2_\pm$ 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 $M_{1,3}$. The results are,

$$\begin{aligned}
\star dt &= -dx^1 \wedge dx^2 \wedge dx^3,\\
\star dx^1 &= +dx^2 \wedge dt \wedge dx^3,\\
\star dx^2 &= +dt \wedge dx^1 \wedge dx^3,\\
\star dx^3 &= +dx^1 \wedge dt \wedge dx^2.
\end{aligned} \tag{2.102}$$

2.4.2 Laplacian

Classical differential operators that enter in Green's and Stokes' theorems are better understood as special manifestations of the exterior differential and the Hodge ⋆ 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

$$df = \frac{\partial f}{\partial x^j} dx^j = \nabla f \cdot d\mathbf{x}. \tag{2.103}$$

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

$$\begin{aligned}
(\star d)\alpha &= \frac{1}{2}(\frac{\partial A_i}{\partial x^j} - \frac{\partial A_i}{\partial x^j}) \star (dx^i \wedge dx^j)\\
&= (\nabla \times \mathbf{A}) \cdot d\mathbf{S}.
\end{aligned} \tag{2.104}$$

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

$$\begin{aligned} d\alpha &= (\frac{\partial B_1}{\partial x^1} + \frac{\partial B_2}{\partial x^2} + \frac{\partial B_3}{\partial x^3})\, dx^1 \wedge dx^2 \wedge dx^3 \\ &= (\nabla \cdot \mathbf{B})\, dV. \end{aligned} \tag{2.105}$$

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

$$(\star d \star)\ \alpha = \nabla \cdot \mathbf{B}. \tag{2.106}$$

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

$$(\star\ d\ \star)\ df = \nabla \cdot \nabla f = \nabla^2 f. \tag{2.107}$$

The Laplacian definition here is consistent with 2.94 because in the case of a function f, that is, a 0-form, $\delta f = 0$ so $\Delta f = \delta df$. The results above can be summarized in terms of short exact sequence called the *de Rham complex* as shown in figure 2.4. 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

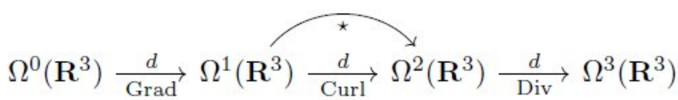


Fig. 2.4: de Rham Complex in $\mathbf{R}^3$

in $\Omega^0(\mathbf{R}^3)$, then $(d \circ d)f = 0$ just says that $\nabla \times \nabla f = 0$, as in the case of conservative vector fields. If instead, one starts with a one form α in $\Omega^1(\mathbf{R}^3)$, corresponding to a vector field $\mathbf{A}$, then $(d \circ d)\alpha = 0$ 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 we denote by R a simply connected closed region in Euclidean space whose boundary is δR , then in terms of forms, the fundamental theorem of calculus, Stokes' theorem (See ref 2.81), and the divergence theorem in $\mathbf{R}^3$ can be expressed by a single generalized Stokes' theorem.

$$\int_{\delta R} \omega = \int \int_R d\omega. \tag{2.108}$$

We find it irresistible to point out that if one defines a complex one-form,

$$\omega = f(z)\, dz, \tag{2.109}$$

where $f(z) = u(x, y) + iv(x, y)$, and where one assumes that u, v are differentiable with continuous derivatives, then the conditions introduced in equation

2.96 are equivalent to requiring that $d\omega = 0$. In other words, if the form is closed, then u and v satisfy the Cauchy-Riemann equations. Stokes' theorem then tells us that in a contractible region with boundary C, the line integral

$$\int_C \omega = \int_C f(z)\, dz = 0.$$

This is Cauchy's integral theorem. We should also point out the tantalizing resemblance of equations 2.96 to Maxwell's equations in the section that follows.

2.4.3 Maxwell Equations

The classical equations of Maxwell describing electromagnetic phenomena are

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 4\pi\rho & \qquad \nabla \times \mathbf{B} &= 4\pi \mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} \\ \nabla \cdot \mathbf{B} &= 0 & \qquad \nabla \times \mathbf{E} &= -\frac{\partial B}{\partial t}, \end{aligned} \tag{2.110}$$

where we are using Gaussian units with $c = 1$. 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 $M_{1,3}$. Define the Maxwell 2-form F by the equation

$$F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu, \quad (\mu, \nu = 0, 1, 2, 3), \tag{2.111}$$

where

$$F_{\mu\nu} = \begin{bmatrix} 0 & -E_x & -E_y & -E_y \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{bmatrix}. \tag{2.112}$$

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

$$\begin{aligned} F \quad = \quad & -E_x dt \wedge dx^1 - E_y dt \wedge dx^2 - E_z dt \wedge dx^3 + \\ & B_z dx^1 \wedge dx^2 - B_y dx^1 \wedge dx^3 + B_x dx^2 \wedge dx^3. \end{aligned} \tag{2.113}$$

We also define the source current 1-form

$$J = J_\mu dx^\mu = \rho dt + J_1 dx^1 + J_2 dx^2 + J_3 dx^3. \tag{2.114}$$

2.4.7 Proposition Maxwell's Equations (2.110) are equivalent to the equations

$$\begin{aligned} dF &= 0, \\ d \star F &= 4\pi \star J. \end{aligned} \tag{2.115}$$

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

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

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

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

which means that

$$-\nabla \times \mathbf{E} - \frac{\partial \mathbf{B}}{\partial t} = 0. \tag{2.116}$$

To verify the second set of Maxwell equations, we first compute the dual of the current density 1-form (2.114) using the results from example 2.4.1. We get

$$\star J = [-\rho dx^1 \wedge dx^2 \wedge dx^3 + J_1 dx^2 \wedge dt \wedge dx^3 + J_2 dt \wedge dx^1 \wedge dx^3 + J_3 dx^1 \wedge dt \wedge dx^2]. \tag{2.117}$$

We could now proceed to compute $d \star F$, but perhaps it is more elegant to notice that $F \in \Omega^2(M)$, and so, according to example (2.4.1), F splits into $F = F_+ + F_-$. In fact, we see from (2.112) 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 (2.4.1), we can immediately write the components of $\star F$:

$$\begin{aligned} \star F &= \frac{1}{2!} B_x dt \wedge dx^1 + B_y dt \wedge dx^2 + B_z dt \wedge dx^3 + \\ &\quad E_z dx^1 \wedge dx^2 - E_y dx^1 \wedge dx^3 + E_x dx^2 \wedge dx^3], \end{aligned} \tag{2.118}$$

or equivalently,

$$F^{\star}_{\mu\nu} = \begin{bmatrix} 0 & B_x & B_y & B_y \\ -B_x & 0 & E_z & -E_y \\ -B_y & -E_z & 0 & E_x \\ -B_z & E_y & -E_x & 0 \end{bmatrix}. \tag{2.119}$$

Effectively, the dual operator amounts to exchanging

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

in the left hand side of the first set of Maxwell equations. We infer from equations (2.116) and (2.117) that

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

and

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

Most standard electrodynamic textbooks carry out the computation entirely tensor components, To connect with this approach, we should mention that it $F^{\mu\nu}$ represents the electromagnetic tensor, then the dual tensor is

$$F^{\star}_{\mu\nu} = \frac{\sqrt{\det g}}{2} \epsilon_{\mu\nu\sigma\tau} F^{\sigma\tau}. \tag{2.120}$$

Since $dF = 0$, in a contractible region there exists a one form A such that $F = dA$. The form A is called the *4-vector potential.* The components of A are,

$$\begin{aligned} A &= A_\mu \, dx^\mu, \\ A_\mu &= (\phi, \mathbf{A}) \end{aligned} \tag{2.121}$$

where ϕ is the electric potential and $\mathbf{A}$ the magnetic vector potential. The components of the electromagnetic tensor F are given by

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}. \tag{2.122}$$

The classical electromagnetic Lagrangian is

$$L_{EM} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + J^\mu A_\mu, \tag{2.123}$$

with corresponding Euler-Lagrange equations

$$\frac{\partial}{\partial x^\mu}\left[\frac{\partial L}{\frac{\partial A_\mu}{\partial A_\nu}}\right] - \frac{\partial L}{\partial A_\mu} = 0. \tag{2.124}$$

To carry out the computation we first use the Minkowski to write the Lagrangian with the indices down. The key is to keep in mind that $A_{\mu,\nu}$ are treated as independent variables, so the derivatives of $A_{\alpha,\beta}$ vanish unless $\mu = \alpha$ and $\nu = \beta$. We get,

$$\begin{aligned}
\frac{\partial L}{\partial(A_{\mu,\nu})} &= -\frac{1}{4}\frac{\partial L}{\partial(A_{\mu,\nu})}(F_{\alpha\beta}F^{\alpha\beta}),\\
&= -\frac{1}{4}\frac{\partial L}{\partial(A_{\mu,\nu})}(F_{\alpha\beta}F_{\lambda\sigma}\eta^{\alpha\lambda}\eta^{\beta\sigma}),\\
&= -\frac{1}{4}\eta^{\alpha\lambda}\eta^{\beta\sigma}[F_{\alpha\beta}(\delta^\mu_\lambda\delta^\mu_\sigma - \delta^\mu_\sigma\delta^\mu_\lambda) + F_{\lambda\sigma}(\delta^\mu_\alpha\delta^\mu_\beta - \delta^\mu_\beta\delta^\mu_\alpha),\\
&= -\frac{1}{4}[\eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta} + \eta^{\mu\lambda}\eta^{\nu\sigma}F_{\lambda\sigma} - \eta^{\alpha\nu}\eta^{\beta\mu}F_{\alpha\beta} - \eta^{\nu\lambda}\eta^{\mu\sigma}(F_{\lambda\sigma}],\\
&= -\frac{1}{4}[F^{\mu\nu} + F^{\mu\nu} - F^{\nu\mu} - F^{\nu\mu}],\\
&= -F^{\mu\nu}.
\end{aligned}$$

On the other hand,

$$\frac{\partial L}{\partial A_\mu} = J^\nu.$$

Therefore, the field equations are

$$\frac{\partial}{\partial x^\mu}F^{\mu\nu} = J^\mu. \tag{2.125}$$

The dual equations equivalent to the other pair of Maxwell equations is

$$\frac{\partial}{\partial x^\mu}\star F^{\mu\nu} = 0.$$

In the gauge theory formulation of classical electrodynamics, the invariant expression for the Lagrangian is the square of the norm of the field F under the induced inner product

$$< F, F >= -\int (F \wedge \star F)\, d\,\Omega. \tag{2.126}$$

This the starting point to generalize to non-Abelian gauge theories.

Chapter 3

Connections

3.1 Frames

This chapter is dedicated to professor Arthur Fischer. In my second year as an undergraduate at Berkeley, I took the undergraduate course in differential geometry which to this day is still called Math 140. The driving force in my career was trying to understand the general theory of relativity, which was only available at the graduate level. However, the graduate course (Math 280 at the time) read that the only prerequisite was Math 140. So I got emboldened and enrolled in the graduate course taught that year by Dr. Fischer. The required book for the course was the classic by Adler, Bazin, Schiffer. I loved the book; it was definitely within my reach and I began to devour the pages with the great satisfaction that I was getting a grasp of the mathematics and the physics. On the other hand, I was completely lost in the course. It seemed as if it had nothing to do with the material I was learning on my own. Around the third week of classes, Dr. Fischer went through a computation with these mysterious operators, and upon finishing the computation he said if we were following, he had just derived the formula for the Christoffel symbols. Clearly, I was not following, they looked nothing like the Christoffel symbols I had learned from the book. So, with great embarrassment I went to his office and explained my predicament. He smiled, apologized when he did not need to, and invited me to 1-1 sessions for the rest of the two-semester course. That is how I got through the book he was really using, namely Abraham-Marsden. I am forever grateful.

As 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.

3.1.1 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

$$e_i = \sum_{j=1}^{n} A^j i \frac{\partial}{\partial x^j} = \partial_j A^j{}_i, \tag{3.1}$$

where $\partial_j = \frac{\partial}{\partial x^j}$. Placing the basis vectors ∂_j on the left is done to be consistent with the summation convention, keeping in mind that the differential operators do not act on the matrix elements. We assume the matrix $A = (A^j{}_i)$ to be nonsingular at each point. In linear algebra, this concept is called 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 results in considerable simplification in computions.

3.1.2 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 $\{e_i\}$, 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

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

3.1.3 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

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

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 $\{e_r, e_\theta, e_z.\}$.

3.1.4 Example For spherical coordinates (2.30)

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

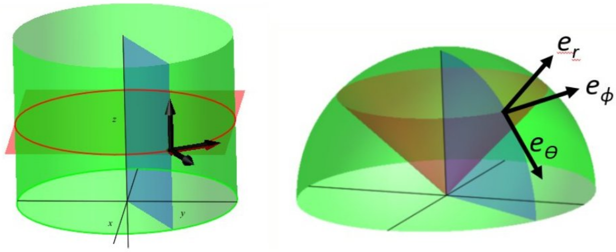


Fig. 3.1: Cylindrical and Spherical Frames.

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 normalized. 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 3.1. 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 many 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 the potential, heat, wave, and Schrödinger field equations, is separable in exactly twelve orthogonal coordinate systems. A simple and efficient method to calculate the Laplacian in orthogonal coordinates can be implemented using differential forms.

3.2.1 Example In spherical coordinates the differential of arc length is given by (see equation 2.31) 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 2.4.2. 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 2.4.2, 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.

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

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

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

Therefore, 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 by 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

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

Classically, the quantities $\{h_1, h_2, h_3\}$ are 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:

$$\begin{aligned}e_1 &= \frac{1}{\sqrt{g_{11}}}\frac{\partial}{\partial u^1} = \frac{1}{h_1}\frac{\partial}{\partial u^1},\\ e_2 &= \frac{1}{\sqrt{g_{22}}}\frac{\partial}{\partial u^2} = \frac{1}{h_2}\frac{\partial}{\partial u^2},\\ e_3 &= \frac{1}{\sqrt{g_{33}}}\frac{\partial}{\partial u^3} = \frac{1}{h_3}\frac{\partial}{\partial u^3}.\end{aligned}$$

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.

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

Thus, the components of the curl are

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

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,

$$\nabla \cdot F = \star d \star F = \frac{1}{h_1h_2h_3}\left[\frac{\partial(h_2h_3F_1)}{\partial u^1} + \frac{\partial(h_1h_3F_2)}{\partial u^2} + \frac{\partial(h_1h_2F_3)}{\partial u^3}\right]. \tag{3.12}$$

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. We seek a quantity which measures the rate of change of a vector field in the direction of another.

3.3.1 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. Implicit in the properties, we set $\overline{\nabla}_X f = X(f)$. 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.

3.3.2 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

$$\overline{\nabla}_X Y = X(f^i)\frac{\partial}{\partial x^i} \tag{3.13}$$

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 the standard connection compatible with the 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.

3.3.3 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,

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

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

$$\overline{\nabla}_X < Y, Z > = < \overline{\nabla}_X Y, Z > + < Y, \overline{\nabla}_X Z > . \tag{3.14}$$

if $F : \mathbf{R}^n \to \mathbf{R}^n$ is an isometry so that $< F_*X, F_*Y > = < X, Y >$, then it is connection preserving in the sense

$$F_*(\overline{\nabla}_X Y) = \overline{\nabla}_{F_*X} F_* y. \tag{3.15}$$

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 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.* Sacrificing some inconsistency with the formalism of differential forms for the sake of connecting to classical notation, we sometimes 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.

3.3.5 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 Γ^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, resulting in 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 $T = T^i e_i$ is a general vector field, then

$$\begin{aligned}\overline{\nabla}_{e_j} T &= \overline{\nabla}_{e_j}(T^i e_i) \\ &= T^i_{\ ,j}\, e_i + T^i \Gamma^k_{\ ji} e_k \\ &= (T^i_{\ ,j} + T^k \Gamma^i_{\ jk}) e_i,\end{aligned} \tag{3.24}$$

which is denoted classically as the covariant derivative

$$T^i_{\ \|j} = T^i_{\ ,j} + \Gamma^i_{\ jk} T^k. \tag{3.25}$$

Here, the comma in the subscript means regular derivative. The equation above is also commonly written as

$$\overline{\nabla}_{e_j} T^i = \overline{\nabla}_j T^i = T^i_{\ ,j} + \Gamma^i_{\ jk} T^k,$$

We should point out the accepted but inconsistent use of terminology. What is meant by the notation $\overline{\nabla}_j T^i$ above is not the covariant derivative of the vector but the tensor components of the covariant derivative of the vector; one more reminder that most physicists conflate a tensor with its components.

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

$$\omega_{ji} = -\omega_{ij} \tag{3.26}$$

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.
The covariant derivative can be extended to the full tensor field $\mathscr{T}^r_s(\mathbf{R}^n)$ by requiring that

a) $\overline{\nabla}_X : \mathscr{T}^r_s(\mathbf{R}^n) \to \mathscr{T}^r_s(\mathbf{R}^n)$,
b) $\overline{\nabla}_X(T_1 \otimes T_2) = \overline{\nabla}_X T_1 \otimes T_2 + T_1 \otimes \overline{\nabla}_X T_2$,
c) $\overline{\nabla}_X$ commutes with all contractions, $\overline{\nabla}_X(CT) = C(\overline{\nabla}_X)$.

Let us compute the covariant derivative of a one-form ω with respect to vector field X. The contraction of $\omega \otimes Y$ is the function $i_Y\omega = \omega(Y)$. Taking the covariant derivative, we have,

$$\begin{aligned}\overline{\nabla}_X(\omega(Y)) &= (\overline{\nabla}_X\omega)(Y) - \omega(\overline{\nabla}_X Y),\\ X(\omega(Y)) &= (\overline{\nabla}_X\omega)(Y) - \omega(\overline{\nabla}_X Y).\end{aligned}$$

Hence, the coordinate-free formula for the covariant derivative of one-form is,

$$(\overline{\nabla}_X\omega)(Y) = X(\omega(Y)) - \omega(\overline{\nabla}_X Y). \tag{3.27}$$

Let θ^i be the dual forms to e_i. We have

$$\overline{\nabla}_X(\theta^i \otimes e_j) = \overline{\nabla}_X\theta^i \otimes e_j + \theta^i \otimes \overline{\nabla}_X e_j.$$

The contraction of $i_{e_j}\theta^i = \theta^i(e_j) = \delta^i_j$, Hence, taking the contraction of the equation above, we see that the left-hand side becomes 0, and we conclude that,

$$(\overline{\nabla}_X\theta^i)(e_j) = -\theta^i(\overline{\nabla}_X e_j). \tag{3.28}$$

Let $\omega = T_i\theta^i$. Then,

$$\begin{aligned}(\overline{\nabla}_X\omega)(e_j) &= (\overline{\nabla}_X(T_i\theta^i))(e_j),\\ &= X(T_i)\theta^i(e_j) + T_i(\overline{\nabla}_X\theta^i)(e_j),\\ &= X(T_j) - T_i\theta^i(\overline{\nabla}_X e_j).\end{aligned} \tag{3.29}$$

If we now set $X = e_k$, we get,

$$\begin{aligned}(\overline{\nabla}_{e_k}\omega)(e_i) &= T_{j,k} - T_i\theta^i(\Gamma^l{}_{kj}e_l),\\ &= T_{j,k} - T_i\delta^i_l\Gamma^l{}_{kj},\\ &= T_{j,k} - \Gamma^i{}_{jk}T_i.\end{aligned}$$

Classically, we write

$$\overline{\nabla}_k T_j = T_{j\|k} = T_{j,k} - \Gamma^i{}_{jk}T_i. \tag{3.30}$$

In general, let T be a tensor of type $\binom{r}{s}$,

$$T = T^{i_1,\dots,i_r}_{j_1,\dots,j_s} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes \theta^{j_1} \otimes \dots \theta^{j_s}. \tag{3.31}$$

Since we know how to take the covariant derivative of a function, a vector, and a one form, we can use Leibnitz rule for tensor products and property of the covariant derivative commuting with contractions, to get by induction, a formula for the covariant derivative of an $\binom{r}{s}$-tensor,

$$\begin{aligned}(\overline{\nabla}_X T)(\theta^{i_1}, ..., \theta^{i_r}, e_{j_1}, ...e_{j_s}) &= X(T(\theta^{i_1}, ..., \theta^{i_r}, e_{j_1}, ..., e_{j_s}))\\ &\quad - T(\overline{\nabla}_X\theta^{i_1}, ..., \theta^{i_r}, e_{j_1}, ..., e_{j_s}) - \dots - T(\theta^{i_1}, ..., \overline{\nabla}_X\theta^{i_r}, e_{j_1}, ..., e_{j_s})...\\ &\quad - T(\theta^{i_1}, ..., \theta^{i_r}, \overline{\nabla}_X e_{j_1}, ..., e_{j_s}) - \dots - T(\theta^{i_1}, ..., \theta^{i_r}, e_{j_1}, ..., \overline{\nabla}_X e_{j_s}).\end{aligned} \tag{3.32}$$

The covariant derivative picks up a term with a positive Christoffel symbol factor for each contravariant index and a term with a negative Christoffel symbol factor for each covariant index. Thus, for example, for a $\binom{1}{2}$ tensor, be components of the covariant derivative in classical notation are

$$\nabla_l T^i{}_{jk} = T^i{}_{jk\|l} = T^i{}_{jk,l} + \Gamma^i{}_{lh} T^h{}_{jk} - \Gamma^h{}_{jl} T^i{}_{hk} - \Gamma^h{}_{kl} T^i{}_{hj}. \tag{3.33}$$

In particular, if g is the metric tensor and X, Y, Z vector fields, we have

$$(\overline{\nabla}_X g)(Y, Z) = X(g(X, Y)) - g(\overline{\nabla}_X Y, Z) - g(X, \overline{\nabla}_X Z).$$

Thus, if we impose the condition $\overline{\nabla}_X g = 0$, the equation above reads

$$\overline{\nabla}_X < Y, Z > = < \overline{\nabla}_X Y, Z > + < Y, \overline{\nabla}_X Z > . \tag{3.34}$$

In other words, a connection is compatible with the metric just means that the metric is covariantly constant along any vector field.

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

$$\overline{\nabla}_X\,[e_1, e_2, e_3] = [e_1, e_2, e_3] \begin{bmatrix} 0 & \omega^1{}_2(X) & \omega^1{}_3(X) \\ -\omega^1{}_2(X) & 0 & \omega^2{}_3(X) \\ -\omega^1{}_3(X) & -\omega^2{}_3(X) & 0 \end{bmatrix} \tag{3.35}$$

Comparing the Frenet frame equation (1.39), 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 in this case 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 elaborate on 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 modern differential geometry, is the work of Cartan, culminating into the famous equations of structure discussed in this chapter.

First Structure Equation

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

$$\Theta^i \equiv d\theta^i + {\omega^i}_j \wedge \theta^j = 0. \tag{3.36}$$

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,

$$\omega = A^{-1} dA. \tag{3.37}$$

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

$$d\theta = -\omega \wedge \theta. \tag{3.38}$$

In other words

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

3.4.2 Example $SO(2, \mathbf{R})$

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

$$\begin{bmatrix} e_1, e_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{3.39}$$

Thus, the attitude matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{3.40}$$

is a rotation matrix in $\mathbf{R}^2$. The set of all such matrices forms a continuous group (*Lie group*) called $SO(2, \mathbf{R})$. In such cases, the matrix

$$\omega = A^{-1}dA \tag{3.41}$$

in equation 3.37 is called the *Maurer-Cartan* form of the group. An easy computation shows that for the rotation group $SO(2)$, the connection form is

$$\omega = \begin{bmatrix} 0 & -d\theta \\ d\theta & 0 \end{bmatrix}. \tag{3.42}$$

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 (2.66), 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

$$\begin{aligned} d\omega^i_{\ j} \wedge \theta^j - \omega^i_{\ j} \wedge (-\omega^j_{\ k} \wedge \theta^k) &= 0, \\ d\omega^i_{\ j} \wedge \theta^j + \omega^i_{\ k} \wedge \omega^k_{\ j} \wedge \theta^j &= 0, \\ (d\omega^i_{\ j} + \omega^i_{\ k} \wedge \omega^k_{\ j}) \wedge \theta^j &= 0, \\ d\omega^i_{\ j} + \omega^i_{\ k} \wedge \omega^k_{\ j} &= 0. \end{aligned}$$

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

$$\Omega^i_{\ j} \equiv d\omega^i_{\ j} + \omega^i_{\ k} \wedge \omega^k_{\ j}. \tag{3.43}$$

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

$$\Omega = d\omega + \omega \wedge \omega = 0. \tag{3.44}$$

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, $d\omega = -\omega \wedge \omega$.

There is a slight abuse of the wedge notation here. The connection ω is matrix valued, so the symbol $\omega \wedge \omega$ is really a composite of matrix and wedge multiplication.

3.4.5 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 in $M = \mathbf{R^n}$ with dual forms θ^i, and let $\overline{e}_i$ be another frame related to the first frame by an invertible transformation.

$$\overline{e}_i = e_j B^j{}_i, \tag{3.45}$$

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 $\overline{\nabla}$ which maps vectors into vector-valued forms,

$$\overline{\nabla} : \Omega^0(M, TM) \to \Omega^1(M, TM)$$

given by the formula

$$\begin{aligned} \overline{\nabla} e_i &= e_j \otimes \omega^j{}_i \\ &= e_j \omega^j{}_i \\ \overline{\nabla} e &= e\,\omega \end{aligned} \tag{3.46}$$

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 3.46 would be

$$d\mathbf{e_i} = \mathbf{e_j}\, \omega^j{}_i. \tag{3.47}$$

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 $\binom{1}{1}$. 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. The action of the covariant differential also extends to the entire tensor algebra, but we do not need that formalism for now, and we delay discussion to section 6.4 on connections on vector bundles. Taking the exterior differential of (3.45) and using (3.46) recursively, we get

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

provided that the connection $\overline{\omega}$ in the new frame $\overline{e}$ is related to the connection ω by the transformation law, (See 6.62)

$$\overline{\omega} = B^{-1}\omega B + B^{-1}dB. \tag{3.48}$$

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 (3.37). The transformation law for the curvature forms is,

$$\overline{\Omega} = B^{-1}\Omega B. \tag{3.49}$$

A quantity transforming as in 3.49 is said to be a *tensorial form of adjoint type.*

3.4.6 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

$$B = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.50}$$

Carrying out the computation for the change of basis 3.48, we find:

$$\begin{aligned}
\overline{\omega}^1{}_2 &= \omega^1{}_2 - d\theta, \\
\overline{\omega}^1{}_3 &= \cos\theta\,\omega^1{}_3 + \sin\theta\,\omega^2{}_3, \\
\overline{\omega}^2{}_3 &= -\sin\theta\,\omega^1{}_3 + \cos\theta\,\omega^2{}_3.
\end{aligned} \tag{3.51}$$

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

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

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

Each set $U = \mathbf{x}(V)$ is called a *coordinate neighborhood* of M. We require that

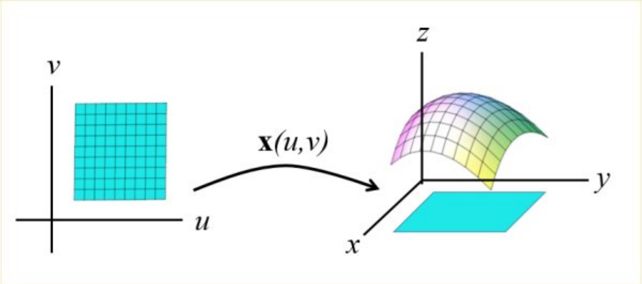


Fig. 4.1: Surface

the Jacobian of the map has maximal rank. In local coordinates, a coordinate chart is represented by three equations in two variables:

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

It will be convenient to use the tensor index formalism when appropriate, so that we can continue to take advantage of the Einstein summation convention. 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, to get an explicit function

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

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

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

4.1.2 Definition Let U_i and U_j be two coordinate neighborhoods of a point $p \in M$ with corresponding charts $\mathbf{x}(u^1, u^2) : V_i \longrightarrow U_i \subset \mathbf{R}^3$ and $\mathbf{y}(v^1, v^2) : V_j \longrightarrow U_j \subset \mathbf{R}^3$ 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)

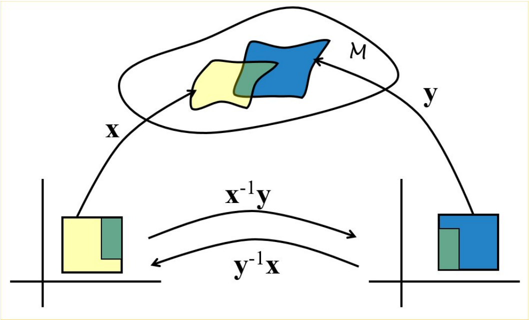


Fig. 4.2: Coordinate Charts

4.1.3 Definition A *differentiable manifold* of dimension 2, is a space 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 M. That is, if $p \in M$, then 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 M inherits a subspace topology consisting of open sets which are defined by the intersection of open sets in $\mathbf{R}^3$ with M.

If the coordinate patches in the definition map from $\mathbf{R}^n$ to $\mathbf{R}^m$ $n < m$ we say that M is a n-dimensional *submanifold* embedded in $\mathbf{R}^m$. In fact, one could define an abstract manifold without the reference to the embedding space by starting with a topological space 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^∞, 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.

4.1.4 Example Consider the local coordinate chart for the unit sphere obtained by setting $r = 1$ in the equations for spherical coordinates 2.30

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

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 soccer 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 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 $\phi =$ constant in equation (4.6), we get the equation of a meridian great circle.

4.1.5 Example *Surface of revolution*
Given a function $f(r)$, the coordinate chart

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

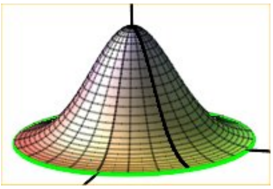


Fig. 4.3: Bell

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. 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.

4.1.6 Example *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

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

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

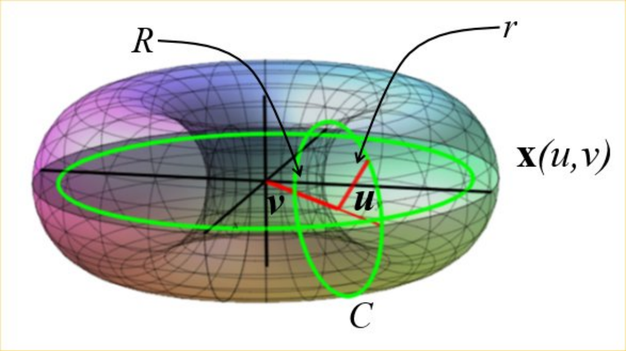


Fig. 4.4: Torus

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.

4.1.7 Example *Monge patch*
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 called an Monge patch

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

4.1.8 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:

$$\mathbf{x}_u = \mathbf{x}_1 = \frac{\partial \mathbf{x}}{\partial u}, \qquad \mathbf{x}_{uu} = \mathbf{x}_{11} = \frac{\partial^2 \mathbf{x}}{\partial u^2}$$
$$\mathbf{x}_v = \mathbf{x}_2 = \frac{\partial \mathbf{x}}{\partial v}, \qquad \mathbf{x}_{vv} = \mathbf{x}_{22} = \frac{\partial^2 \mathbf{x}}{\partial v^2},$$

or more succinctly,

$$\mathbf{x}_\alpha = \frac{\partial \mathbf{x}}{\partial u^\alpha}, \quad \mathbf{x}_{\alpha\beta} = \frac{\partial^2 \mathbf{x}}{\partial u^\alpha \partial u^\beta} \tag{4.10}$$

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,

$$ds^2 = g_{\alpha\beta} du^\alpha du^\beta, \tag{4.11}$$

where

$$g_{\alpha\beta} = \delta_{ij} \frac{\partial x^i}{\partial u^\alpha} \frac{\partial x^j}{\partial u^\beta}. \tag{4.12}$$

We conclude that the surface, by virtue of being embedded in $\mathbf{R}^3$, inherits a natural metric (4.11) which we will call the *induced metric*. A pair $\{M, g\}$, where 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 (4.11) 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

$$ds^2 = E du^2 + 2F du dv + G dv^2, \tag{4.13}$$

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

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

4.2.1 Definition *First fundamental form*
The element of arc length,

$$ds^2 = g_{\alpha\beta} du^\alpha \otimes du^\beta, \tag{4.14}$$

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 more modern version. Following the same development as in the theory of curves, consider a surface M defined locally by a function $q = (u^1, u^2) \longmapsto p = \alpha(u^1, u^2)$. We say that a quantity X_p is a tangent vector at a point $p \in M$ if X_p is a linear derivation on the space of C^∞ real-valued functions $\mathscr{F} = \{f | f : M \longrightarrow \mathbf{R}\}$ on the surface. The set of all tangent vectors at a point $p \in M$ is called the *tangent space* T_pM. 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* TM. Sections of the tangent bundle of M are consistently denoted by $\mathscr{X}(M)$. The coordinate chart map $\alpha : \mathbf{R}^2 \longrightarrow M \subset \mathbf{R}^3$ induces a *push-forward* map $\alpha_* : T\mathbf{R}^2 \longrightarrow TM$ which maps a vector V at each point in $T_q(\mathbf{R}^2)$ into a vector $V_{\alpha(q)} = \alpha_*(V_q)$ in $T_{\alpha(q)}M$, as illustrated in the diagram 4.5

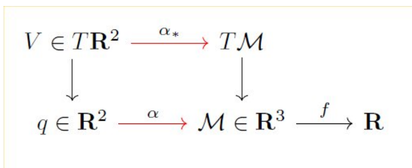


Fig. 4.5: Push-Forward

The action of the push-forward is defined by

$$\alpha_*(V)(f) \mid_{\alpha(q)} = V(f \circ \alpha) \mid_q . \tag{4.15}$$

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 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,

$$I(X, Y) = g(X, Y) = < X, Y >, \tag{4.16}$$

where X and Y are any pair of vector fields in $\mathscr{X}(M)$.

Orthogonal Parametric Curves

Let V and W be vectors tangent to a surface M defined locally by a chart $\mathbf{x}(u^\alpha)$. Since the vectors $\mathbf{x}_\alpha$ span the tangent space of 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 the curvilinear components 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 vectors V and W is the given by the equation

$$\begin{aligned} \cos\theta &= \frac{<V,W>}{\|V\| \cdot \|W\|}, \\ &= \frac{I(V,W)}{\sqrt{I(V,V)}\sqrt{I(W,W)}}, \\ &= \frac{g_{\alpha_1\beta_1} V^{\alpha_1} W^{\beta_1}}{\sqrt{g_{\alpha_2\beta_2} V^{\alpha_2} V^{\beta_2}}\sqrt{g_{\alpha_3\beta_3} W^{\alpha_3} W^{\beta_3}}}, \end{aligned} \tag{4.17}$$

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

$$\cos\theta = \frac{g_{\alpha_1\beta_1} du^{\alpha_1} \delta u^{\beta_1}}{\sqrt{g_{\alpha_2\beta_2} du^{\alpha_2} du^{\beta_2}}\sqrt{g_{\alpha_3\beta_3} \delta u^{\alpha_3} \delta u^{\beta_3}}}. \tag{4.18}$$

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:

$$\cos\theta = \frac{g_{12} \delta u^1 du^2}{\sqrt{g_{11}(\delta u^1)^2}\sqrt{g_{22}(du^2)^2}} = \frac{g_{12}}{g_{11} g_{22}} = \frac{F}{\sqrt{EG}}. \tag{4.19}$$

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

$$\cos\theta = \frac{<\mathbf{x}_u, \mathbf{x}_v>}{\|\mathbf{x}_u\| \cdot \|\mathbf{x}_v\|} = \frac{F}{\sqrt{EG}}. \tag{4.20}$$

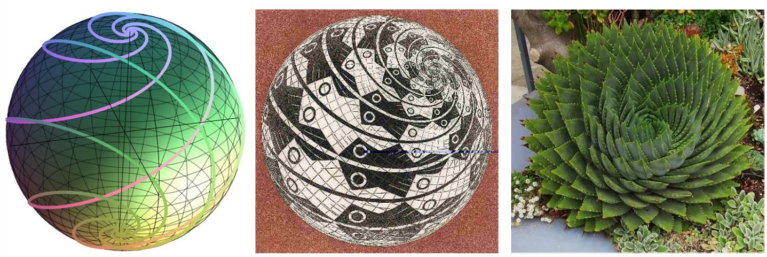
It follows immediately from the equation above that:

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

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

4.2.3 Examples a) Sphere

$$\begin{aligned}
\mathbf{x} &= (a\sin\theta\cos\phi, a\sin\theta\sin\phi, a\cos\theta),\\
\mathbf{x}_\theta &= (a\cos\theta\cos\phi, a\cos\theta\sin\phi, -a\sin\theta),\\
\mathbf{x}_\phi &= (-a\sin\theta\sin\phi, a\sin\theta\cos\phi, 0),\\
E &= \mathbf{x}_\theta\cdot\mathbf{x}_\theta = a^2,\\
F &= \mathbf{x}_\theta\cdot\mathbf{x}_\phi = 0,\\
G &= \mathbf{x}_\phi\cdot\mathbf{x}_\phi = a^2\sin^2\theta,\\
ds^2 &= a^2 d\theta^2 + a^2\sin^2\theta\; d\phi^2. \qquad (4.21)
\end{aligned}$$



a) Loxodromes b) Escher Drawing c) Aloe Polyphylla

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 cylindrical 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.

As we will discuss in more detail later, a navigator travelling a constant bearing would be moving on a straight path 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 used by cartographers, is to measure the angle θ from the equator. 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

$$\phi = \pm(\tan\sigma)\ln\tan(\tfrac{\vartheta}{2} + \tfrac{\pi}{4}) + \phi_0 \tag{4.22}$$

$$s = a(\theta - \theta_0)\sec\sigma, \tag{4.23}$$

where θ_0 and ϕ_0 are the coordinates of the initial position. Figure 4.2 shows four loxodromes equally distributed around the sphere.

Loxodromes were the bases for a number of beautiful drawings and woodcuts by M. C. Escher. figure 4.2 also shows one more beautiful manifestation of geometry in nature in a plant called Aloe Polyphylla. Not surprisingly, the plant has 5 loxodromoes which is a Fibonacci number. We will show later under the discussion of conformal (angle preserving) maps in section 5.2.2, that loxodromes map into straight lines making a constant angle with meridians in the Mercator projection (See Figure **??**).

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 figure 4.6, 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.

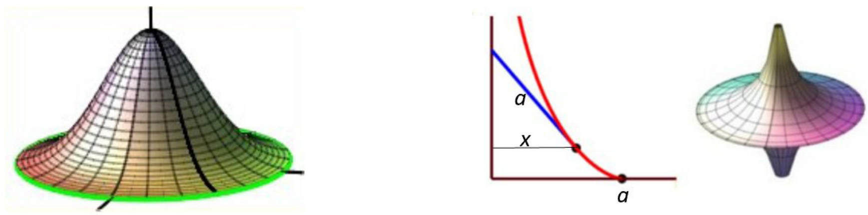


Fig. 4.6: Surface of Revolution and Pseudosphere

c) Pseudosphere

$$
\begin{aligned}
\mathbf{x} &= (a\sin u\cos v, a\sin u\sin v, a(\cos u + \ln(\tan\frac{u}{2})), \\
E &= a^2\cot^2 u, \\
F &= = 0 \\
G &= a^2\sin^2 u, \\
ds^2 &= a^2\cot^2 u\; du^2 + a^2\sin^2 u\; dv^2.
\end{aligned}
$$

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 4.6. 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. He discovered that in spite of the surface extending asymptotically to infinity, the surface area is finite with $S = 4\pi a^2$ as in a sphere of the same radius, and the volume enclosed is half that sphere. We will have much more to say about this surface.

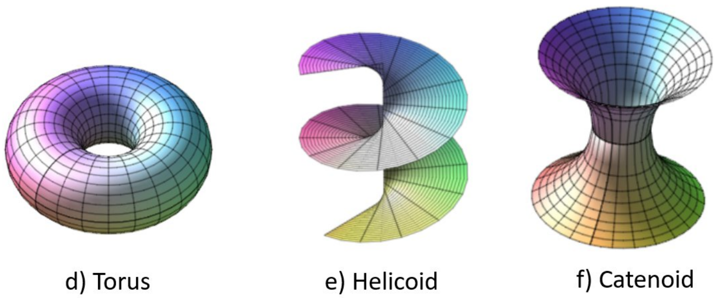


Fig. 4.7: Examples of Surfaces

d) Torus

$$\begin{aligned}
\mathbf{x} &= ((b + a\cos u)\cos v, (b + a\cos u)\sin v, a\sin u) \quad \text{(See 4.8)},\\
E &= a^2,\\
F &= 0,\\
G &= (b + a\cos u)^2,\\
ds^2 &= a^2du^2 + (b + a\cos u)^2dv^2. \qquad (4.24)
\end{aligned}$$

e) Helicoid

$$\begin{aligned}
\mathbf{x} &= (u\cos v, u\sin v, av) \quad \text{Coordinate curves } u = c. \quad \text{are helices.}\\
E &= 1,\\
F &= 0,\\
G &= u^2 + a^2,\\
ds^2 &= du^2 + (u^2 + a^2)dv^2. \qquad (4.25)
\end{aligned}$$

f) Catenoid

$$\begin{aligned}
\mathbf{x} &= (u\cos v, u\sin v, c\cosh^{-1}\frac{u}{c}), \quad \text{This is a catenary of revolution.}\\
E &= \frac{u^2}{u^2 - c^2},\\
F &= 0,\\
G &= u^2,\\
ds^2 &= \frac{u^2}{u^2 - c^2}du^2 + u^2dv^2, \qquad (4.26)
\end{aligned}$$

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,

$$\begin{aligned}
\mathbf{x} &= (r\cos\phi, r\sin\phi, r\cot\alpha),\\
E &= \csc^2\alpha,\\
F &= 0,\\
G &= r^2,\\
ds^2 &= \csc^2\alpha\, dr^2 + r^2 d\phi^2. \quad (4.27)
\end{aligned}$$

A conical helix is 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

$$\begin{aligned}
\gamma' &= \mathbf{x}_r\frac{dr}{dt} + \mathbf{x}_\phi\frac{d\phi}{dt}.\\
\cos\sigma &= \frac{<\mathbf{x}_r,\gamma'>}{\|\mathbf{x}_r\|\cdot\|\gamma'\|} = \frac{E\frac{dr}{dt}}{\sqrt{E}\frac{ds}{dt}} = \sqrt{E}\frac{dr}{ds}.\\
E\,dr^2 &= \cos^2\sigma\, ds^2,\\
\csc^2\alpha\, dr^2 &= \cos^2\sigma(\csc^2\alpha\, dr^2 + r^2 d\phi^2),\\
\csc^2\alpha\sin^2\sigma\, dr^2 &= r^2\cos^2\sigma\, d\phi^2,\\
\frac{1}{r}\,dr &= \cot\sigma\sin\alpha\, d\phi.
\end{aligned}$$

Therefore, the equations of a conical helix are given by

$$r = c\, e^{\cot\sigma\sin\alpha\phi}. \tag{4.28}$$

As shown in figure 4.8, 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. The 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.

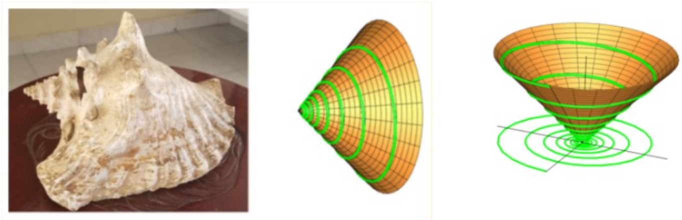


Fig. 4.8: Conical Helix.

4.3 The Second Fundamental Form

Let $\mathbf{x} = \mathbf{x}(u^\alpha)$ be a coordinate patch on a surface 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

$$\mathbf{n} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|}. \tag{4.29}$$

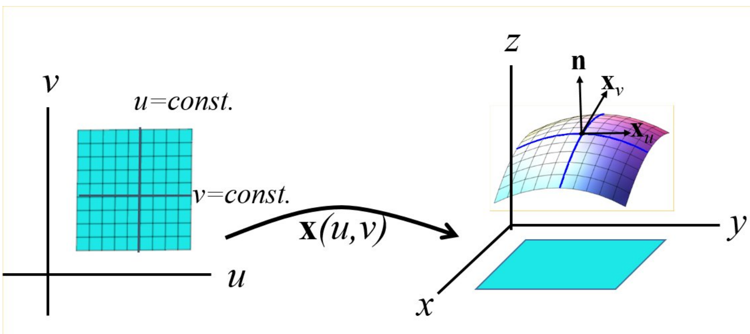


Fig. 4.9: Surface Normal

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

$$d[\mathbf{T}, \mathbf{N}, \mathbf{B}] = [\mathbf{T}, \mathbf{N}, \mathbf{B}] \begin{bmatrix} 0 & -\kappa\, ds & 0 \\ \kappa\, ds & 0 & -\tau\, ds \\ 0 & \tau\, ds & 0. \end{bmatrix} \tag{4.30}$$

Following ideas first introduced by Darboux and subsequently 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

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}. \tag{4.31}$$

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 3.48 and 3.51) gives:

$$d[\mathbf{T},\mathbf{g},\mathbf{n}] = [\mathbf{T},\mathbf{g},\mathbf{n}] \begin{bmatrix} 0 & -\kappa\cos\theta\,ds & -\kappa\sin\theta\,ds \\ \kappa\cos\theta\,ds & 0 & -\tau\,ds + d\theta \\ \kappa\sin\theta\,ds & \tau\,ds - d\theta & 0 \end{bmatrix}, \tag{4.32}$$

$$= [\mathbf{T},\mathbf{g},\mathbf{n}] \begin{bmatrix} 0 & -\kappa_g\,ds & -\kappa_n\,ds \\ \kappa_g\,ds & 0 & -\tau_g\,ds \\ \kappa_n\,ds & \tau_g\,ds & 0 \end{bmatrix}, \tag{4.33}$$

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 4.10)

$$\mathbf{T}' = \kappa_n\mathbf{n} + \kappa_g\mathbf{g}, \tag{4.34}$$

$$\kappa^2 = \kappa_n^2 + \kappa_g^2. \tag{4.35}$$

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 line acquires 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 Los Angeles to San Francisco along one such great circle, there is a short path and a very long one that goes around the earth.

If one specifies a point $\mathbf{p} \in M$ and a direction vector $X_p \in T_pM$, one can geometrically envision the normal curvature by considering the equivalence class of all unit speed curves in 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 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.

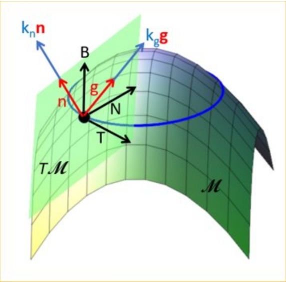


Fig. 4.10: Curvature

Using the chain rule, we se that the unit tangent vector $\mathbf{T}$ to the curve is given by

$$\mathbf{T} = \frac{d\mathbf{x}}{ds} = \frac{d\mathbf{x}}{du^\alpha}\frac{du^\alpha}{ds} = \mathbf{x}_\alpha \frac{du^\alpha}{ds}. \tag{4.36}$$

To find an explicit formula for the normal curvature we first differentiate equation (4.36)

$$\begin{aligned}
\mathbf{T}' &= \frac{d\mathbf{T}}{ds},\\
&= \frac{d}{ds}(\mathbf{x}_\alpha \frac{du^\alpha}{ds}),\\
&= \frac{d}{ds}(\mathbf{x}_\alpha)\frac{du^\alpha}{ds} + \mathbf{x}_\alpha \frac{d^2u^\alpha}{ds^2},\\
&= (\frac{d\mathbf{x}_\alpha}{du^\beta}\frac{du^\beta}{ds})\frac{du^\alpha}{ds} + \mathbf{x}_\alpha \frac{d^2u^\alpha}{ds^2},\\
&= \mathbf{x}_{\alpha\beta}\frac{du^\alpha}{ds}\frac{du^\beta}{ds} + \mathbf{x}_\alpha \frac{d^2u^\alpha}{ds^2}.
\end{aligned}$$

Taking the inner product of the last equation with the normal and noticing that $< \mathbf{x}_\alpha, \mathbf{n} >= 0$, we get

$$\begin{aligned}
\kappa_n &= < \mathbf{T}', \mathbf{n} >=< \mathbf{x}_{\alpha\beta}, \mathbf{n} > \frac{du^\alpha}{ds}\frac{du^\beta}{ds},\\
&= \frac{b_{\alpha\beta}du^\alpha du^\beta}{g_{\alpha\beta}du^\alpha du^\beta},
\end{aligned} \tag{4.37}$$

where

$$b_{\alpha\beta} =< \mathbf{x}_{\alpha\beta}, \mathbf{n} > \tag{4.38}$$

4.3.1 Definition The expression

$$II = b_{\alpha\beta}\, du^\alpha \otimes du^\beta \tag{4.39}$$

is called the *second fundamental form* .

4.3.2 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 (4.39) can be written as

$$II = edu^2 + 2fdudv + gdv^2. \tag{4.40}$$

It follows that the equation for the normal curvature (4.37) can be written explicitly as

$$\kappa_n = \frac{II}{I} = \frac{edu^2 + 2fdudv + gdv^2}{Edu^2 + 2Fdudv + Gdv^2}. \tag{4.41}$$

We 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,

$$\begin{aligned} < d\mathbf{x}, d\mathbf{n} > &= < \mathbf{x}_\alpha du^\alpha, \mathbf{n}_\beta du^\beta >, \\ &= < \mathbf{x}_\alpha du^\alpha, \mathbf{n}_\beta du^\beta >, \\ &= < \mathbf{x}_\alpha, \mathbf{n}_\beta > du^\alpha du^\beta, \\ &= - < \mathbf{x}_{\alpha\beta}, \mathbf{n} > du^\alpha du^\beta, \\ &= -II. \end{aligned}$$

4.3.3 Definition Directions on a surface along which the second fundamental form

$$e\, du^2 + 2f\, du\, dv + g\, dv^2 = 0 \tag{4.42}$$

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 plane $z = 0$.

For now, we state without elaboration, that one can also define the third fundamental form by

$$III =< d\mathbf{n}, d\mathbf{n} >=< \mathbf{n}_\alpha, \mathbf{n}_\beta > du^\alpha du^\beta. \tag{4.43}$$

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

$$(A \times B) \cdot (C \times D) = \begin{vmatrix} A \cdot C & A \cdot D \\ B \cdot C & B \cdot D \end{vmatrix}, \tag{4.44}$$

to compute

$$\begin{aligned} \|\mathbf{x}_u \times \mathbf{x}_v\|^2 &= (\mathbf{x}_u \times \mathbf{x}_v) \cdot (\mathbf{x}_u \times \mathbf{x}_v), \\ &= \det \begin{bmatrix} \mathbf{x}_u \cdot \mathbf{x}_u & \mathbf{x}_u \cdot \mathbf{x}_v \\ \mathbf{x}_v \cdot \mathbf{x}_u & \mathbf{x}_v \cdot \mathbf{x}_v \end{bmatrix}, \\ &= EG - F^2. \end{aligned} \tag{4.45}$$

Consequently, the normal vector can be written as

$$\mathbf{n} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\sqrt{EG - F^2}}.$$

It follows that we can write the coefficients $b_{\alpha\beta}$ directly as triple products involving derivatives of $(\mathbf{x})$. The expressions for these coefficients are

$$\begin{aligned} e &= \frac{(\mathbf{x}_u \mathbf{x}_v \mathbf{x}_{uu})}{\sqrt{EG - F^2}}, \\ f &= \frac{(\mathbf{x}_u \mathbf{x}_v \mathbf{x}_{uv})}{\sqrt{EG - F^2}}, \\ g &= \frac{(\mathbf{x}_u \mathbf{x}_v \mathbf{x}_{vv})}{\sqrt{EG - F^2}}. \end{aligned} \tag{4.46}$$

4.3.4 Example Sphere
Going back to example 4.21, 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 $\mathbf{x}_0 = \mathbf{x}(u_0^\alpha) \in M$ and a nearby point $\mathbf{x}(u_0^\alpha + du^\alpha)$. Expanding on a Taylor series, we get

$$\mathbf{x}(u_0^\alpha + du^\alpha) = \mathbf{x}_0 + (\mathbf{x}_0)_\alpha du^\alpha + \frac{1}{2}(\mathbf{x}_0)_{\alpha\beta}du^\alpha du^\beta + \dots.$$

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.

The notion of curvature of a hypersurface in $\mathbf{R}^n$ (a surface of dimension $n-1$) begins by studying 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 indicative of 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 4.41 in the form

$$\kappa_n = \frac{II^*}{I^*} = \frac{e + 2f\lambda + g\lambda^2}{E + 2F\lambda + G\lambda^2}, \tag{4.47}$$

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,

$$\kappa_n = \frac{II^*}{I^*} = \frac{f + g\lambda}{F + G\lambda}. \tag{4.48}$$

On the other hand, combining with equation 4.47 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

$$\kappa_n = \frac{f + g\lambda}{F + G\lambda} = \frac{e + f\lambda}{E + F\lambda}. \tag{4.49}$$

Equation 4.49 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

$$\begin{Vmatrix} du^2 & -du\,dv & dv^2 \\ E & F & G \\ e & f & g \end{Vmatrix} = 0. \tag{4.50}$$

Equation 4.50 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 4.49 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

$$\begin{Vmatrix} e - E\kappa_n & f - F\kappa_n \\ f - F\kappa_n & g - G\kappa_n \end{Vmatrix} = 0. \tag{4.51}$$

It is interesting to note that equation 4.51 can be written as

$$\left\Vert \begin{bmatrix} e & f \\ f & g \end{bmatrix} - \kappa_n \begin{bmatrix} E & F \\ F & G \end{bmatrix} \right\Vert = 0.$$

In other words, the extrema for the values of the normal are the solutions of the equation

$$\Vert b_{\alpha\beta} - \kappa_n g_{\alpha\beta} \Vert = 0. \tag{4.52}$$

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 is

$$(EG - F^2)\kappa_n^2 - (Eg - 2Ff + Ge)\kappa_n + (eg - f^2) = 0.$$

We conclude there are two solutions κ_1 and κ_2 such that

$$K = \kappa_1\kappa_2 = \frac{eg - f^2}{EG - F^2}, \tag{4.53}$$

and

$$M = \tfrac{1}{2}(\kappa_1 + \kappa_2) = \frac{1}{2}\frac{Eg - 2Ff + Ge}{EF - G^2}.. \tag{4.54}$$

The quantity K is called the *Gaussian curvature* and M is called the *mean curvature*. 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 linear algebra spectral theorem for symmetric operators.

4.4.2 Covariant Derivative Formulation of Curvature

4.4.1 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

$$LX = -\overline{\nabla}_X N, \tag{4.55}$$

is called the *Weingarten map*. Some authors call this the *shape operator*. The same definition applies if M is an n-dimensional hypersurface in $\mathbf{R}^{n+1}$.

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.

4.4.2 Definition The *Lie bracket* $[X, Y]$ of two vector fields X and Y on a surface M is defined as the commutator,

$$[X, Y] = XY - YX, \tag{4.56}$$

meaning that if f is a function on M, then $[X, Y](f) = X(Y(f)) - Y(X(f))$.

4.4.3 Proposition The Lie bracket of two vectors $X, Y \in \mathscr{X}(M)$ is another vector in $\mathscr{X}(M)$.

Proof If suffices to prove that the bracket is a linear derivation on the space of C^∞ functions. Consider vectors $X, Y \in \mathscr{X}(M)$ and smooth functions f, g in 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

$$\begin{aligned}
[X,Y](fg) &= X(Y(fg)) - Y(X(fg)),\\
&= X[fY(g) + gY(f)] - Y[fX(g) + gX(f)],\\
&= X(f)Y(g) + fX(Y(g)) + X(g)Y(f) + gX(Y(f)),\\
&\quad -Y(f)X(g) - f(Y(X(g)) - Y(g)X(f) - gY(X(f)),\\
&= f[X(Y(g)) - (Y(X(g))] + g[X(Y(f)) - Y(X(f))],\\
&= f[X,Y](g) + g[X,Y](f).
\end{aligned}$$

4.4.4 Proposition The Weingarten map is a linear transformation on $\mathscr{X}(M)$.
Proof Linearity follows from the linearity of $\overline{\nabla}$, so it suffices to show that $L : X \longrightarrow LX$ maps $X \in \mathscr{X}(M)$ to a vector $LX \in \mathscr{X}(M)$. 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 $\mathscr{X}(M)$.

In the preceding 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.

4.4.5 Definition The *second fundamental form* is the bilinear map

$$II(X,Y) =< LX, Y > . \tag{4.57}$$

4.4.6 Remark 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

$$III(X,Y) =< LX, LY >=< L^2X, Y > . \tag{4.58}$$

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.

4.4.7 Definition The *torsion* of a connection $\overline{\nabla}$ is the operator T such that $\forall X, Y$,

$$T(X,Y) = \overline{\nabla}_X Y - \overline{\nabla}_Y X - [X,Y]. \tag{4.59}$$

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 elaborate later on 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.

4.4.8 Theorem The Weingarten map is a self-adjoint endomorphism on $\mathscr{X}(M)$.

Proof We have already shown that $L : \mathscr{X}M \longrightarrow \mathscr{X}M$ 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,

$$\begin{aligned} II(X,Y) - II(Y,X) &= < N, \overline{\nabla}_Y X > - < N, \overline{\nabla}_X Y >, \\ &= < N, \overline{\nabla}_Y X - \overline{\nabla}_X Y >, \\ &= < N, [X,Y] >, \\ &= 0 \qquad (\text{iff } \ [X,Y] \in T(M)). \end{aligned}$$

The central theorem of linear algebra is the spectral theorem. In the case of real, self-adjoint operators, the spectral theorem states that given the eigenvalue equation for a symmetric operator

$$LX = \kappa X, \tag{4.60}$$

on a vector space with a real inner product, 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:

$$LX_1 = \kappa_1 X_1 \tag{4.61}$$
$$LX_2 = \kappa_2 X_2. \tag{4.62}$$

4.4.9 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 p on a given surface:

1. If $\kappa_1 \neq \kappa_2$ and both eigenvalues are positive, then p is called an *elliptic point.*

2. If $\kappa_1\kappa_2 < 0$, then p is called a *hyperbolic point.*

3. If $\kappa_1 = \kappa_2 \neq 0$, then p is called an *umbilic point.*

4. If $\kappa_1\,\kappa_2 = 0$, then p is called a *parabolic point.*

It is also known from linear algebra, that in a vector space of dimension two, the determinant and the trace of a self-adjoint operator 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. In the case of a hypersurface of n-dimensions, there would n eigenvalues, counting multiplicities, so the classification of the points would be more elaborate

4.4.10 Definition The determinant $K = \det(L)$ is called the *Gaussian curvature* of 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

$$\begin{aligned} K &= \kappa_1\kappa_2, \\ H &= \frac{1}{2}(\kappa_1 + \kappa_2). \end{aligned} \tag{4.63}$$

An alternative definition of curvature is obtained by considering the unit normal as a map $N : M \to S^2$, which maps each point p on the surface M, to the point on the sphere corresponding to the position vector N_p. The map is called the *Gauss map.*

4.4.11 Examples

1. The Gauss map of a plane is constant. The image is a single point on S^2.

2. The image of the Gauss map of a circular cylinder is a great circle on S^2.

3. The Gauss map of the top half of a circular cone sends all points on the cone into a circle. We may envision this circle as the intersection of the cone and a unit sphere centered at the vertex.

4. The Gauss map of a circular hyperboloid of one sheet misses two antipodal spherical caps with boundaries corresponding to the circles of the asymptotic cone.

5. The Gauss map of a catenoid misses two antipodal points.

The Weingarten map is minus the derivative $N_* = dN$ of the Gauss map. That is, $LX = -N_*(X)$.

4.4.12 Proposition Let X and Y be any linearly independent vectors in $\mathscr{X}(M)$. Then

$$\begin{aligned} LX \times LY &= K(X \times Y), \\ (LX \times Y) + (X \times LY) &= 2H(X \times Y). \end{aligned} \tag{4.64}$$

Proof Since $LX, LY \in \mathscr{X}(M)$, they can be expressed as linear combinations of the basis vectors X and Y.

$$\begin{aligned} LX &= a_1 X + b_1 Y, \\ LY &= a_2 X + b_2 Y. \end{aligned}$$

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

$$\begin{aligned} (LX \times Y) + (X \times LY) &= (a_1 + b_2)(X \times Y), \\ &= \mathrm{Tr}(L)(X \times Y), \\ &= (2H)(X \times Y). \end{aligned}$$

4.4.13 Proposition

$$\begin{aligned} K &= \frac{eg - f^2}{EG - F^2}, \\ H &= \frac{1}{2}\frac{Eg - 2Ff + eG}{EG - F^2}. \end{aligned} \tag{4.65}$$

Proof Starting with equations (4.64), take the inner product of both sides with $X \times Y$ and use the vector identity (4.44). We immediately get

$$K = \frac{\begin{vmatrix} < LX, X > & < LX, Y > \\ < LY, X > & < LX, X > \end{vmatrix}}{\begin{vmatrix} < X, X > & < X, Y > \\ < Y, X > & < Y, Y > \end{vmatrix}}, \tag{4.66}$$

$$2H = \frac{\begin{vmatrix} < LX, X > & < LX, Y > \\ < Y, X > & < Y, Y > \end{vmatrix} + \begin{vmatrix} < X, X > & < X, Y > \\ < LY, X > & < LY, Y > \end{vmatrix}}{\begin{vmatrix} < X, X > & < X, Y > \\ < Y, X > & < Y, Y > \end{vmatrix}}. \tag{4.67}$$

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 4.53) and for the mean curvature (equation 4.54.

If we denote by g and b the matrices of the fundamental forms whose components are $g_{\alpha\beta}$ and $b_{\alpha\beta}$ respectively, we can write the equations for the curvatures as:

$$K = \det\left(\frac{b}{g}\right) = \det(g^{-1}b), \tag{4.68}$$

$$2H = \mathrm{Tr}\left(\frac{b}{g}\right) = \mathrm{Tr}(g^{-1}b) \tag{4.69}$$

4.4.14 Example Sphere
From equations 4.21 and 4.3 we see that $K = 1/a^2$ and $H = 1/a$. This is totally intuitive since one would expect $\kappa_1 = \kappa_2 = 1/a$ because 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 is 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 M at a given point.

4.4.15 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

$$II(X,X) = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta. \tag{4.70}$$

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

$$\begin{aligned} < LX, X > &= < (\cos\theta)\kappa_1 X_1 + (\sin\theta)\kappa_2 X_2, (\cos\theta)X_1 + (\sin\theta)X_2 > \\ &= \kappa_1 \cos^2\theta < X_1, X_1 > + \kappa_2 \sin^2\theta < X_2, X_2 > \\ &= \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta. \end{aligned}$$

4.4.16 Theorem The first, second and third fundamental forms satisfy the equation

$$III - 2H\,II + KI = 0 \tag{4.71}$$

Proof The proof follows immediately from the fact that for a symmetric 2 by 2 matrix A, the characteristic polynomial is $\kappa^2 - \mathrm{tr}(A)\kappa + \det(A) = 0$, and from the Cayley-Hamilton theorem stating that the matrix is annihilated by its characteristic polynomial.

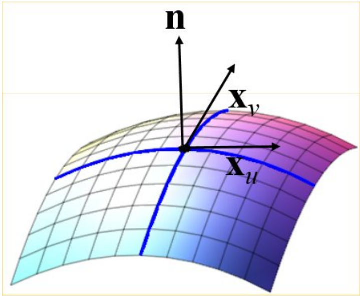


Fig. 4.11: Surface Frame.

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.

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 4.11. 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 $\mathbf{x}_{\alpha\beta}$. Here, $\mathbf{x}(u^\alpha)$ could refer to a coordinate patch in any number of dimensions, so all the tensor index formulas that follow, apply to surfaces of codimension 1 in $\mathbf{R}^n$. The set of vectors $\{\mathbf{x}_\alpha, \mathbf{n}\}$ constitutes a basis for $\mathbf{R}^n$ at each point on the surface, we can express the vectors $\mathbf{x}_{\alpha\beta}$ as linear combinations of the basis vectors. Therefore, there exist coefficients $\Gamma^\gamma_{\alpha\beta}$ and $c_{\alpha\beta}$ such that,

$$\mathbf{x}_{\alpha\beta} = \Gamma^\gamma_{\alpha\beta}\mathbf{x}_\gamma + c_{\alpha\beta}\mathbf{n}. \tag{4.72}$$

Taking the inner product of equation 4.72 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 4.72 can be written as

$$\mathbf{x}_{\alpha\beta} = \Gamma^\gamma_{\alpha\beta}\mathbf{x}_\gamma + b_{\alpha\beta}\mathbf{n}. \tag{4.73}$$

Equation 4.73 together with equation 4.76 below, are called the formulæ 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 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,

$$\begin{aligned} h(X,Y) &= < \overline{\nabla}_X Y, N >, \\ &= - < Y, \overline{\nabla}_X N >, \\ &= - < Y, LX, >, \\ &= - < LX, Y >, \\ &= II(X,Y). \end{aligned}$$

Thus, the coordinate independent formulation of the equation of Gauss reads

$$\overline{\nabla}_X Y = \nabla_X Y + II(X,Y)N. \tag{4.74}$$

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 4.12 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 $\overline{\nabla}$.

The torsion T of the connection ∇ is defined exactly as before (See equation 4.59).

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y].$$

Also, as in definition 3.14, the connection is compatible with the metric on the surface if

$$\nabla_X < Y, Z > = < \nabla_X Y, Z > + < Y, \nabla_X Z > .$$

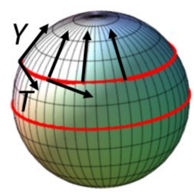


Fig. 4.12:

A torsion-free connection that is compatible with the metric is called a *Levi-Civita* connection.

4.5.1 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

$$\begin{aligned}
\nabla_T < X, X > &=< \nabla_T X, X > + < X, \nabla_T X >,\\
&= 2 < \nabla_T X, X >= 0,\\
&\Rightarrow \|X\| = \text{constant}.\\
\nabla_T < X, Y > &=< \nabla_T X, Y > + < X, \nabla_T Y >= 0,\\
&\Rightarrow< X, Y >= \text{constant}. \qquad \text{So,}\\
\cos\theta &= \frac{< X, Y >}{\|X\| \cdot \|Y\|} = \text{constant}.
\end{aligned}$$

If one takes $\{e_\alpha\}$ to be a basis of the 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 4.73. We have the following important result:

4.5.2 Theorem In a manifold $\{M, g\}$ with metric g, there exists a unique Levi-Civita connection.

The proof is implicit in the computations that follow leading to equation 4.76, which express the components uniquely in terms of the metric. The entire equation (4.73) 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 (see equation 3.34)

$$\begin{aligned}
g_{\alpha\gamma,\beta} &= \frac{\partial}{\partial u^\beta} < \mathbf{x}_\alpha, \mathbf{x}_\gamma >,\\
&=< \mathbf{x}_{\alpha\beta}, \mathbf{x}_\gamma > + < \mathbf{x}_\alpha, \mathbf{x}_{\gamma\beta}, >,\\
&= \Gamma_{\alpha\beta\gamma} + \Gamma_{\gamma\beta\alpha}.
\end{aligned}$$

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

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}(g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha} - g_{\alpha\beta,\gamma}). \tag{4.75}$$

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.)

$$\Gamma^{\sigma}_{\alpha\beta} = \frac{1}{2}g^{\sigma\gamma}(g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha} - g_{\alpha\beta,\gamma}). \tag{4.76}$$

The Christoffel symbols are clearly symmetric in the lower indices

$$\Gamma^{\sigma}_{\alpha\beta} = \Gamma^{\sigma}_{\beta\alpha}. \tag{4.77}$$

Unless otherwise specified, a connection on $\{M, g\}$ refers to the unique Levi-Civita connection.

We derive a well-known formula for the Christoffel symbols for the case $\Gamma^{\alpha}{}_{\alpha\beta}$. From 4.76 we have:

$$\Gamma^{\alpha}_{\alpha\beta} = \frac{1}{2}g^{\alpha\gamma}(g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha} - g_{\alpha\beta,\gamma}).$$

On the other hand,

$$g^{\alpha\gamma}g_{\beta\gamma,\alpha} = g^{\alpha\gamma}g_{\alpha\beta,\gamma}$$

as can be seen by switching the repeated indices of summation α and σ, and using the symmetry of the metric. The equation reduces to

$$\Gamma^{\alpha}_{\alpha\beta} = \frac{1}{2}g^{\alpha\gamma}g_{\alpha\gamma,\beta}$$

Let A be the cofactor transposed matrix of g. From the linear algebra formula for the expansion of a determinant in terms of cofactors we can get an expression an expression for the inverse of the metric as follows:

$$\begin{aligned}
\det(g) &= g_{\alpha\gamma}A^{\alpha\gamma},\\
\frac{\partial \det(g)}{\partial g_{\alpha\gamma}} &= A^{\alpha\gamma},\\
g^{\alpha\gamma} &= \frac{A^{\alpha\gamma}}{\det(g)},\\
&= \frac{1}{\det(g)}\frac{\partial \det(g)}{\partial g_{\alpha\gamma}}
\end{aligned}$$

so that

$$
\begin{aligned}
\Gamma^{\alpha}_{\alpha\beta} &= \frac{1}{2\det(g)} \frac{\partial \det(g)}{\partial g_{\alpha\gamma}} \frac{\partial}{\partial u^{\beta}} g_{\alpha\gamma}, \\
&= \frac{1}{2\det(g)} \frac{\partial}{\partial u^{\beta}} (\det(g)). \qquad (4.78)
\end{aligned}
$$

Using this result we can also get a tensorial version of the divergence of the vector field $X = v^{\alpha} e_{\alpha}$ on the manifold. Using the classical covariant derivative formula 3.25 for the components v^{α}, we define:

$$
\operatorname{Div} X = \nabla \cdot X = v^{\alpha}{}_{\|\alpha} \qquad (4.79)
$$

We get

$$
\begin{aligned}
\operatorname{Div} X &= v^{\alpha}{}_{,\alpha} + \Gamma^{\alpha}{}_{\alpha\gamma}\, v^{\gamma}, \\
&= \frac{\partial}{\partial u^{\alpha}} v^{\alpha} + \frac{1}{2\det(g)} \frac{\partial}{\partial u^{\gamma}} (\det(g)) v^{\gamma}, \\
&= \frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial u^{\alpha}} (\sqrt{\det(g)} v^{\alpha}). \qquad (4.80)
\end{aligned}
$$

If f is a function on the manifold, $df = f_{,\beta}\, du^{\beta}$ so the contravariant components of the gradient are

$$
(\nabla f)^{\alpha} = g^{\alpha\beta} f_{,\beta}\,. \qquad (4.81)
$$

Combining with equation above, we get a second order operator

$$
\begin{aligned}
\Delta f &= \operatorname{Div}(\operatorname{Grad} f), \\
&= \frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial u^{\alpha}} (\sqrt{\det(g)} g^{\alpha\beta} f_{,\beta}) \qquad (4.82)
\end{aligned}
$$

The quantity Δ is called the *Laplace-Beltrami* operator on a function and it generalizes the Laplacian of functions in $\mathbf{R}^n$ to functions on manifolds.

4.5.3 Example Laplacian in Spherical Coordinates

The metric in spherical coordinates is $ds^2 = dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\phi^2$, so

$$
g_{\alpha\beta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{bmatrix}, \qquad g^{\alpha\beta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{i}{r^2\sin^2\theta} \end{bmatrix}, \qquad \sqrt{\det(g)} = r^2 \sin\theta.
$$

The Laplace-Beltrami formula gives,

$$
\begin{aligned}
\Delta f &= \frac{1}{\sqrt{\det g}} \left[\frac{\partial}{\partial u^1} (\sqrt{\det g}\; g^{11} \frac{\partial f}{\partial u^1}) + \frac{\partial}{\partial u^2} (\sqrt{\det g}\; g^{22} \frac{\partial f}{\partial u^2}) + \frac{\partial}{\partial u^3} (\sqrt{\det g}\; g^{33} \frac{\partial f}{\partial u^3}) \right], \\
&= \frac{1}{r^2 \sin\theta} \left[\frac{\partial}{\partial r} (r^2 \sin\theta \frac{\partial f}{\partial r}) + \frac{\partial}{\partial \theta} (r^2 \sin\theta \frac{1}{r^2} \frac{\partial f}{\partial \theta}) + \frac{\partial}{\partial \phi} (r^2 \sin\theta \frac{1}{r^2 \sin^2\theta} \frac{\partial f}{\partial \phi}) \right], \\
&= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2 f}{\partial \phi^2}.
\end{aligned}
$$

The result is the same as the formula for the Laplacian 3.9 found by differential form methods.

4.5.4 Example

As an example we unpack the formula for Γ^1_{11}. First, note that $\det(g) = \|g_{\alpha\beta}\| = EG - F^2$. From equation 4.76 we have

$$\begin{aligned}
\Gamma^1_{11} &= \frac{1}{2} g^{1\gamma}(g_{1\gamma,1} + g_{1\gamma,1} - g_{11,\gamma}),\\
&= \frac{1}{2} g^{1\gamma}(2g_{1\gamma,1} - g_{11,\gamma}),\\
&= \frac{1}{2}[g^{11}(2g_{11,1} - g_{11,1}) + g^{12}(2g_{12,1} - g_{11,2})],\\
&= \frac{1}{2\det(g)}[GE_u - F(2F_u - FE_v)],\\
&= \frac{GE_u - 2FF_u + FE_v}{2(EG - F^2)}.
\end{aligned}$$

Due to symmetry, there are five other similar equations for the other Γ's. Proceeding as above, we can derive the entire set.

$$\begin{aligned}
\Gamma^1_{11} &= \frac{GE_u - 2FF_u + FE_v}{2(EG-F^2)} & \Gamma^2_{11} &= \frac{2EF_u - EE_v - FE_u}{2(EG-F^2)}\\
\Gamma^1_{12} &= \frac{GE_v - FG_u}{2(EG-F^2)} & \Gamma^2_{12} &= \frac{EG_u - FE_v}{2(EG-F^2)}\\
\Gamma^1_{22} &= \frac{2GF_v - GG_u - FG_v}{2(EG-F^2)} & \Gamma^2_{22} &= \frac{EG_v - 2FF_v + FG_u}{2(EG-F^2)}.
\end{aligned} \tag{4.83}$$

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.

4.5.5 Example Harmonic functions.

A function h on a surface in $\mathbf{R}^3$ is called *harmonic* if it satisfies:

$$\Delta\, h = 0. \tag{4.84}$$

Noticing that the matrix components of the inverse of the metric are given by

$$g^{\alpha\beta} = \frac{1}{\det(g)}\begin{bmatrix} G & -F \\ -F & E \end{bmatrix} \tag{4.85}$$

we get immediately from 4.82, the classical Laplace-Beltrami equation for surfaces,

$$\Delta h = \frac{1}{\sqrt{EG-F^2}}\left\{\frac{\partial}{\partial u}\left[\frac{Gh_{,u} - Fh_{,v}}{\sqrt{EG-F^2}}\right] + \frac{\partial}{\partial v}\left[\frac{Eh_{,v} - Fh_{,u}}{\sqrt{EG-F^2}}\right]\right\} = 0. \tag{4.86}$$

If the coordinate patch is orthogonal so that $F = 0$, the equation reduces to:

$$\frac{\partial}{\partial u}\left[\frac{\sqrt{G}}{\sqrt{E}}\frac{\partial h}{\partial u}\right] + \frac{\partial}{\partial v}\left[\frac{\sqrt{E}}{\sqrt{G}}\frac{\partial h}{\partial v}\right] = 0 \tag{4.87}$$

If in addition $E = G = \lambda^2$ so that the metric has the form,

$$ds^2 = \lambda^2\,(du^2 + dv^2), \tag{4.88}$$

then,

$$\Delta h = \frac{1}{\lambda^2}\left[\frac{\partial^2 h}{\partial u^2} + \frac{\partial^2 h}{\partial v^2}\right]. \tag{4.89}$$

Hence, $\Delta^2 h = 0$ is equivalent to $\nabla^2 h = 0$, where ∇^2 is the Euclidean Laplacian. (Please compare to the discussion on the isothermal coordinates example 4.5.14.) Two metrics that differ by a multiplicative factor are called conformally related. The result here means that the Laplacian is conformally invariant under this conformal transformation. This property is essential in applying the elegant methods of complex variables and conformal mappings to solve physical problems involving the Laplacian in the plane.

For a surface $z = f(x, y)$, which we can write as a Monge patch $\mathbf{x} =< x, y, f(x, y) >$, we have $E = 1 + f_x^2$, $F = 2f_x f_y$ and $G = 1 + f_y^2$=0. A short computation shows that in this case, the Laplace-Beltrami equation can be written as, (compare to equation **??**)

$$\Delta h = \frac{1}{\sqrt{1 + f_x^2 + f_y^2}}\left\{\frac{\partial}{\partial x}\left[\frac{f_x}{\sqrt{1 + f_x^2 + f_y^2}}\right] + \frac{\partial}{\partial y}\left[\frac{f_y}{\sqrt{1 + f_x^2 + f_y^2}}\right]\right\} = 0,$$

or in terms of the Euclidean $\mathbf{R}^2$ del operator $\nabla =< \frac{\partial}{\partial x}, \frac{\partial}{\partial y} >$,

$$\frac{\partial}{\partial x}\left[\frac{f_x}{\sqrt{1 + f_x^2 + f_y^2}}\right] + \frac{\partial}{\partial y}\left[\frac{f_y}{\sqrt{1 + f_x^2 + f_y^2}}\right] = 0,$$

$$\nabla \cdot \left[\frac{\nabla f}{\sqrt{1 + \|\nabla f\|^2}}\right] = 0. \tag{4.90}$$

4.5.2 Curvature Tensor, Gauss's Theorema Egregium

A fascinating set of relations can be obtained simply by equating $\mathbf{x}_{\beta\gamma\delta} = \mathbf{x}_{\beta\delta\gamma}$. First notice that 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

$$\mathbf{n}_\alpha = -b_\alpha^\gamma \mathbf{x}_\gamma. \tag{4.91}$$

These are called the formulæ of *Weingarten.*

Differentiating the equation of Gauss and recursively using the formulas of Gauss 4.73 and Weingarten 4.91 to write all the components in terms of the frame, we get

$$\begin{aligned}
\mathbf{x}_{\beta\delta} &= \Gamma^{\alpha}_{\beta\delta}\mathbf{x}_{\alpha} + b_{\beta\delta}\mathbf{n}, \\
\mathbf{x}_{\beta\delta\gamma} &= \Gamma^{\alpha}_{\beta\delta,\gamma}\mathbf{x}_{\alpha} + \Gamma^{\alpha}_{\beta\delta}\mathbf{x}_{\alpha\gamma} + b_{\beta\delta,\gamma}\mathbf{n} + b_{\beta\delta}\mathbf{n}_{\gamma} \\
&= \Gamma^{\alpha}_{\beta\delta,\gamma}\mathbf{x}_{\alpha} + \Gamma^{\alpha}_{\beta\delta}[\Gamma^{\mu}_{\alpha\gamma}\mathbf{x}_{\mu} + b_{\alpha\gamma}\mathbf{n}] + b_{\beta\delta,\gamma}\mathbf{n} - b_{\beta\delta}b^{\alpha}_{\gamma}\mathbf{x}_{\alpha} \\
\mathbf{x}_{\beta\delta\gamma} &= [\Gamma^{\alpha}_{\beta\delta,\gamma} + \Gamma^{\mu}_{\beta\delta}\Gamma^{\alpha}_{\mu\gamma} - b_{\beta\delta}b^{\alpha}_{\gamma}]\mathbf{x}_{\alpha} + [\Gamma^{\alpha}_{\beta\delta}b_{\alpha\gamma} + b_{\beta\delta,\gamma}]\mathbf{n}, && (4.92) \\
\mathbf{x}_{\beta\gamma\delta} &= [\Gamma^{\alpha}_{\beta\gamma,\delta} + \Gamma^{\mu}_{\beta\gamma}\Gamma^{\alpha}_{\mu\delta} - b_{\beta\gamma}b^{\alpha}_{\delta}]\mathbf{x}_{\alpha} + [\Gamma^{\alpha}_{\beta\gamma}b_{\alpha\delta} + b_{\beta\gamma,\delta}]\mathbf{n}. && (4.93)
\end{aligned}$$

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

$$R^{\alpha}{}_{\beta\gamma\delta} = b_{\beta\delta}b^{\alpha}_{\gamma} - b_{\beta\gamma}b^{\alpha}_{\delta}, \tag{4.94}$$

where the components of the *Riemann tensor* R are defined by

$$R^{\alpha}_{\beta\gamma\delta} = \Gamma^{\alpha}_{\beta\delta,\gamma} - \Gamma^{\alpha}_{\beta\gamma,\delta} + \Gamma^{\mu}_{\beta\delta}\Gamma^{\alpha}_{\gamma\mu} - \Gamma^{\mu}_{\beta\gamma}\Gamma^{\alpha}_{\delta\mu}. \tag{4.95}$$

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 in a later chapter. Lowering the index above we get

$$R_{\alpha\beta\gamma\delta} = b_{\beta\delta}b_{\alpha\gamma} - b_{\beta\gamma}b_{\alpha\delta}. \tag{4.96}$$

4.5.6 *Theorema egregium* Let M be a smooth surface in $\mathbf{R}^3$. Then,

$$K = \frac{R_{1212}}{\det(g)}. \tag{4.97}$$

Proof Let $\alpha = \gamma = 1$ and $\beta = \delta = 2$ above. The equation then reads

$$\begin{aligned}
R_{1212} &= b_{22}b_{11} - b_{21}b_{12}, \\
&= (eg - f^2), \\
&= K(EF - G^2), \\
&= K\det(g)
\end{aligned}$$

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

$$\Gamma^{\alpha}_{\beta\delta} b_{\alpha\gamma} - \Gamma^{\alpha}_{\beta\gamma} b_{\alpha\delta} + b_{\beta\delta,\gamma} - b_{\beta\gamma,\delta} = 0 \tag{4.98}$$

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

$$
\begin{aligned}
d\theta^1 &= -\omega^1{}_2 \wedge \theta^2, && (4.99)\\
d\theta^2 &= -\omega^2{}_1 \wedge \theta^1, && (4.100)\\
d\theta^3 &= -\omega^3{}_1 \wedge \theta^1 - \omega^3{}_2 \wedge \theta^2 = 0, && (4.101)\\
d\omega^1{}_2 &= -\omega^1{}_3 \wedge \omega^3{}_2, \qquad \text{Gauss Equation} && (4.102)\\
d\omega^1{}_3 &= -\omega^1{}_2 \wedge \omega^2{}_3, \qquad \text{Codazzi Equations} && (4.103)\\
d\omega^2{}_3 &= -\omega^2{}_1 \wedge \omega^1{}_3. && (4.104)
\end{aligned}
$$

The key result is the following theorem

4.5.7 *Curvature form equations*

$$
\begin{aligned}
d\omega^1{}_2 &= K\,\theta^1 \wedge \theta^2, && (4.105)\\
\omega^1{}_3 \wedge \theta^2 + \omega^2{}_3 \wedge \theta^1 &= -2H\;\theta^1 \wedge \theta^2. && (4.106)
\end{aligned}
$$

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:

$$
\begin{aligned}
K = \det(L) &= -[\omega^1{}_3(e_1)\omega^3{}_2(e_2) - \omega^1{}_3(e_2)\omega^3{}_2(e_1)],\\
&= -(\omega^1{}_3 \wedge \omega^3{}_2)(e_1, e_2),\\
&= d\omega^1{}_2(e_1, e_2).
\end{aligned}
$$

Hence

$$d\omega^1{}_2 = K\,\theta^1 \wedge \theta^2.$$

Similarly, recalling that $\theta^1(e_j) = \delta^i_j$, we have

$$\begin{aligned}
(\omega^1{}_3 \wedge \theta^2 + \omega^3{}_2 \wedge \theta^1)(e_1, e_2) &= \omega^1{}_3(e_1) - \omega^3{}_2(e_2), \\
&= \omega^1{}_3(e_1) + \omega^2{}_3(e_2), \\
&= \mathrm{Tr}(L) = -2H.
\end{aligned}$$

4.5.8 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.

4.5.9 Example Sphere Since the first fundamental form is $I = a^2\, d\theta^2 + a^2 \sin^2\theta\, d\phi^2$, we have

$$\begin{aligned}
\theta^1 &= a\, d\theta, \\
\theta^2 &= a \sin\theta\, d\phi, \\
d\theta^2 &= a \cos\theta\, d\theta \wedge d\phi, \\
&= -\cos\theta\, d\phi \wedge \theta^1 = -\omega^2{}_1 \wedge \theta^1, \\
\omega^2{}_1 &= \cos\theta\, d\phi = -\omega^1{}_2, \\
d\omega^1{}_2 &= \sin\theta\, d\theta \wedge d\phi = \frac{1}{a^2}\,(a\, d\theta) \wedge (a \sin\theta\, d\phi), \\
&= \frac{1}{a^2}\theta^1 \wedge \theta^2, \\
K &= \frac{1}{a^2}.
\end{aligned}$$

4.5.10 Example Torus

Using the the parametrization (See 4.24),

$$\mathbf{x} = ((b + a\cos\theta)\cos\phi, (b + a\cos\theta)\sin\phi, a\sin\theta),$$

the first fundamental form is

$$ds^2 = a^2 d\theta^2 + (b + a\cos\theta)^2 d\phi^2.$$

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.

4.5.11 **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:

$$
K = -\frac{1}{\sqrt{EG}}\left[\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\right)\right]. \tag{4.107}
$$

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

$$
\begin{aligned}
d\mathbf{x} &= \mathbf{x}_u\, du + \mathbf{x}_v\, dv, \\
&= \frac{\mathbf{x}_u}{\sqrt{E}}\sqrt{E}\, du + \frac{\mathbf{x}_v}{\sqrt{G}}\sqrt{G}\, dv, \\
&= \frac{\mathbf{x}_u}{\sqrt{E}}\theta^1 + \frac{\mathbf{x}_v}{\sqrt{G}}\theta^2, \\
d\mathbf{x} &= \mathbf{e}_1\, \theta^1 + \mathbf{e}_2\, \theta^2,
\end{aligned} \tag{4.108}
$$

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 4.104 can be rewritten without the arbitrary vector in the operator LX as shown:

$$\overline{\nabla}_X e_3 = \omega^1{}_3(X)\, e_1 + \omega^2{}_3(X)\, e_2, \tag{4.109}$$

$$d\mathbf{e}_3 = \mathbf{e}_1\, \omega^1{}_3 + \mathbf{e}_2\, \omega^2{}_3 = 0, \tag{4.110}$$

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:

$$\begin{aligned} \omega^1{}_3 &= l\, \theta^1 + m\, \theta^2, \\ \omega^2{}_3 &= m\, \theta^1 + n\, \theta^2. \end{aligned} \tag{4.111}$$

If $E = 1$, we say that the metric

$$ds^2 = du^2 + G(u,v)dv^2, \tag{4.112}$$

is in *geodesic coordinates.* In this case, the equation for curvature reduces even further to:

$$K = -\frac{1}{\sqrt{G}} \frac{\partial^2 \sqrt{G}}{\partial u^2}. \tag{4.113}$$

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 $ds^2 = dr^2 + r^2\, d\phi^2$. Equation 4.113 is called *Jacobi's differential equation* for geodesic coordinates.

A slick proof of the theorema egregium can be obtained by differential forms. Let $F : M \to \widetilde{\mathcal{M}}$ 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

$$< \tilde{e}_\alpha, \tilde{e}_\beta >=< F_* e_\alpha, F_* e_\beta >=< e_\alpha, e_\beta >= \delta_{\alpha\beta}.$$

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 $\widetilde{\mathcal{M}}$.

4.5.12 Theorem (Theorema 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 $\widetilde{M}$:

$$\begin{aligned} d\tilde{\theta}^{\alpha} + \tilde{\omega}^{\alpha}{}_{\beta} \wedge \tilde{\theta}^{\beta} &= 0, \\ F^{*} d\tilde{\theta}^{\alpha} + F^{*}\tilde{\omega}^{\alpha}{}_{\beta} \wedge F^{*}\tilde{\theta}^{\beta} &= 0, \\ d\theta^{\alpha} + F^{*}\tilde{\omega}^{\alpha}{}_{\beta} \wedge \theta^{\beta} &= 0, \end{aligned}$$

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, $F^{*}K = K$.

4.5.13 Example Catenoid - Helicoid

Perhaps the most celebrated manifestation of the theorema egregium, is that of mapping between a helicoid M and a catenoid $\widetilde{M}$. 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 4.25 and 4.26.

$$\begin{aligned} ds^2 &= du^2 + (u^2 + a^2)\, dv^2, & & E = 1, \quad G = u^2 + a^2, \\ d\tilde{s}^2 &= \frac{\tilde{u}^2}{\tilde{u}^2 - a^2}\, d\tilde{u}^2 + \tilde{u}^2\, d\tilde{v}^2 & \text{with} \quad & \tilde{E} = \frac{\tilde{u}^2}{\tilde{u}^2 - a^2}, \quad \tilde{G} = \tilde{u}^2. \end{aligned}$$

Let $F : M \to \widetilde{\mathbf{M}}$ 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 4.107.

$$K = \frac{-1}{\sqrt{u^2 + a^2}} \frac{\partial}{\partial u}\left(\frac{\partial}{\partial u}\sqrt{u^2 + a^2}\right) = \frac{a^2}{(u^2 + a^2)^2}, \tag{4.114}$$

$$\tilde{K} = -\frac{\sqrt{\tilde{u}^2 - a^2}}{\tilde{u}^2} \frac{\partial}{\partial \tilde{u}}\left(\frac{\sqrt{\tilde{u}^2 - a^2}}{\tilde{u}}\right) = -\frac{a^2}{\tilde{u}^4}. \tag{4.115}$$

It is immediately evident by substitution that as expected $F^{*}\tilde{K} = K$. Figure 4.13 shows several stages of a one-parameter family M_t of isometries deforming a catenoid into a helicoid. The one-parameter family of coordinate patches chosen is

$$\mathbf{z}_t = (\cos t)\, \mathbf{x} + (\sin t)\, \mathbf{y} \tag{4.116}$$

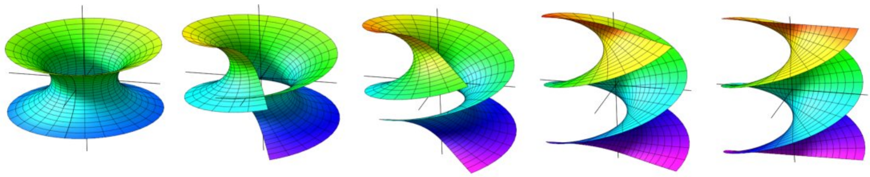


Fig. 4.13: Catenoid - Helicoid isometry

Writing the equation of the coordinate patch $\mathbf{z}_t$ in complete detail, one can compute the coefficients of the fundamental forms and thus establish the family of surfaces has mean curvature H independent of the parameter t, and in fact $H = 0$ for each member of the family. We will discuss at a later chapter the geometry of surfaces of zero mean curvature.

4.5.14 Example Isothermal coordinates.
Consider the case in which the metric has the form

$$ds^2 = \lambda^2\,(du^2 + dv^2), \tag{4.117}$$

so that $E = G = \lambda^2,\ F = 0$. A metric in this form is said to be in *isothermal coordinates.* Inserting into equation 4.107, we get

$$\begin{aligned} K &= -\frac{1}{\lambda^2}\left[\frac{\partial}{\partial u}\left(\frac{1}{\lambda}\frac{\partial \lambda}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{1}{\lambda}\frac{\partial \lambda}{\partial v}\right)\right], \\ &= -\frac{1}{\lambda^2}\left[\frac{\partial}{\partial u}\frac{\partial}{\partial u}(\ln\lambda) + \frac{\partial}{\partial v}\frac{\partial}{\partial v}(\ln\lambda)\right]. \end{aligned}$$

Hence,

$$K = -\frac{1}{\lambda^2}\nabla^2(\ln\lambda). \tag{4.118}$$

The tantalizing appearance of the Laplacian in this coordinate system gives an inkling that there is some complex analysis lurking in the neighborhood. Readers acquainted with complex variables will recall that the real and imaginary parts of holomorphic functions satisfy Laplace's equations and that any holomorphic function in the complex plane describes a conformal map. In anticipation of further discussion on this matter, we prove the following:

4.5.15 Theorem Define the mean curvature vector $\mathbf{H} = H\mathbf{n}$. If $\mathbf{x}(u, v)$ is an isothermal parametrization of a surface, then

$$\mathbf{x}_{uu} + \mathbf{x}_{vv} = 2\lambda^2\mathbf{H}. \tag{4.119}$$

Proof Since the coordinate patch is isothermal, $E = G = \lambda^2$ and $F = 0$. Specifically, we have $<\mathbf{x}_u, \mathbf{x}_u> = <\mathbf{x}_v, \mathbf{x}_v>$, and $<\mathbf{x}_u, \mathbf{x}_v> = 0$. Differentiation then gives:

$$
\begin{aligned}
< \mathbf{x}_{uu}, \mathbf{x}_u > &=< \mathbf{x}_v, \mathbf{x}_{vu} >,\\
&= - < \mathbf{x}_{vv}, \mathbf{x}_u >,\\
< \mathbf{x}_{uu} + \mathbf{x}_{vv}, \mathbf{x}_u > &= 0.
\end{aligned}
$$

In the same manner,

$$
\begin{aligned}
< \mathbf{x}_{vv}, \mathbf{x}_v > &=< \mathbf{x}_u, \mathbf{x}_{uv} >,\\
&= - < \mathbf{x}_{uu}, \mathbf{x}_v >,\\
< \mathbf{x}_{uu} + \mathbf{x}_{vv}, \mathbf{x}_v > &= 0.
\end{aligned}
$$

If follows that $\mathbf{x}_{uu} + \mathbf{x}_{vv}$ is orthogonal to the surface and points in the direction of the normal $\mathbf{n}$. On the other hand,

$$
\begin{aligned}
\frac{Eg + Ge}{2EG} &= H,\\
\frac{g + e}{2\lambda^2} &= H,\\
e + g &= 2\lambda^2 H,\\
< \mathbf{x}_{uu} + \mathbf{x}_{vv}, \mathbf{n} > &= 2\lambda^2 H,\\
\mathbf{x}_{uu} + \mathbf{x}_{vv} &= 2\lambda^2 \mathbf{H}.
\end{aligned}
$$

Chapter 5

Geometry of Surfaces

Chapter 6

Riemannian Geometry

6.1 Riemannian Manifolds

In the definition of manifolds introduced in section 4.1, it was implicitly assumed manifolds were embedded (or immersed) in $\mathbf{R}^n$. As such, they inherited a natural metric induced by the standard Euclidean metric of $\mathbf{R}^n$, as shown in section 4.2. For general manifolds it is more natural to start with a topological space M, and define the coordinate patches as pairs $\{U_i, \phi_i\}$, where $\{U_i\}$ is an open cover of M with local homeomorphisms

$$\phi_i : U_i \subset M \to \mathbf{R}^n.$$

If $p \in U_i \cap U_j$ is a point in the non-empty intersection of two charts, we require that the overlap map $\phi_{ij} = \phi_i \phi_j^{-1} : \mathbf{R}^n \to \mathbf{R^n}$ be a diffeomorphism. The local coordinates on patch $\{U, \phi\}$ are given by $(x^1, \ldots, x^n)$, where

$$x^i = u^i \circ \phi,$$

and $u^i : \mathbf{R}^n \to \mathbf{R}$ are the projection maps on each slot. The concept is the same as in figure 4.2, but, as stated, we are not assuming a priori that M is embedded (or immersed) in Euclidean space. If in addition the space is equipped with a metric, the space is called a Riemannian manifold. If the signature of the metric is of type $g = \text{diag}(1, 1, \ldots, -1, -1)$, with p '+' entries and q '-' entries, we say that M is a *pseudo-Riemannian* manifold of type (p, q). As we have done with Minkowski's space, we switch to Greek indices x^μ for local coordinates of curved space-times. We write the Riemannian metric as

$$ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu. \tag{6.1}$$

We will continue to be consistent with earlier notation and denote the tangent space at a point $p \in M$ as T_pM, the tangent bundle as TM, and the space of vector fields as $\mathscr{X}(M)$. Similarly, we denote the space of differential k-forms by $\Omega^k(M)$, and the set of type $\binom{r}{s}$ tensor fields by $\mathscr{T}^r_s(M)$.

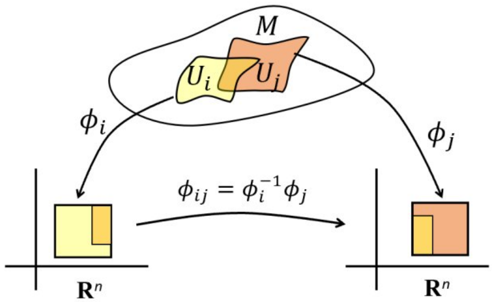


Fig. 6.1: Coordinate Charts

Product Manifolds

Suppose that M_1 and M_2 are differentiable manifolds of dimensions m_1 and m_2 respectively. Then, $M_1 \times M_2$ can be given a natural manifold structure of dimension $n = m_1 + m_2$ induced by the product of coordinate charts. That is, if (ϕ_{i_1}, U_{i_1}) is a chart in M_1 in a neighborhood of $p_1 \in M_1$, and (ϕ_{i_2}, U_{i_2}) is a chart in a neighborhood of $p_2 \in M_2$ in M_2, then the map

$$\phi_{i_1} \times \phi_{i_2} : U_{i_1} \times U_{i_2} \to \mathbf{R}^n$$

defined by

$$(\phi_{i_1} \times \phi_{i_2})(p_1, p_2) = (\phi_{i_1}(p_1), \phi_{i_2}(p_2)),$$

is a coordinate chart in the product manifold. An atlas constructed from such charts, gives the differentiable structure. Clearly, $M_1 \times M_2$ is locally diffeomorphic to $\mathbf{R}^{m_1} \times \mathbf{R}^{m_2}$. To discuss the tangent space of a product manifold, we recall from linear algebra, that given two vector spaces V and W, the *direct sum* $V \oplus W$ is the vector space consisting of the set of ordered pairs

$$V \oplus W = \{(v, w) : v \in V, \quad w \in W\},$$

together with the vector operations

$$\begin{aligned} (v_1, w_1) + (v_2, w_2) &= (v_1 + v_2, w_1 + w_2), \quad \text{for all, } \; v_1, v_2 \in V; \; w_1, w_2 \in W, \\ k(v, w) &= (kv, kw), \quad \text{for all } \; k \in \mathbf{R} \end{aligned}$$

People often say that one cannot add apples and peaches, but this is not a problem for mathematicians. For example, 3 apples and 2 peaches plus 4 apples and 6 peaches is 7 apples and 8 peaches. This is the basic idea behind the direct sum. We now have the following theorem:

6.1.1 Theorem Let $(p_1, p_2) \in M_1 \times M_2)$, then there is a vector space isomorphism

$$T_{(p_1, p_2)}(M_1 \times M_2) \cong T_{p_1} M_1 \oplus T_{p_2} M_2.$$

Proof The proof is adapted from [18]. Given $X_1 \in T_{p_1}M_1$ and $X_2 \in T_{p_2}M_2$, let

$$\begin{aligned} &x_1(t) \text{ be a curve with } x_1(0) = p_1 \text{ and } x_1'(0) = X_1, \\ &x_2(t) \text{ be a curve with } x_2(0) = p_2 \text{ and } x_2'(0) = X_2. \end{aligned}$$

Then, we can associate

$$(X_1, X_2) \in T_{p_1}M_1 \oplus T_{p_2}M_2$$

with the vector $X \in T_{(p_1,p_2)}(M_1 \times M_2)$, which is tangent to the curve $x(t) = (x_1(t), x_2(t))$, at the point (p_1, p_2). In the simplest possible case where the product manifold is $\mathbf{R}^2 = \mathbf{R}^1 \times \mathbf{R}^1$, the vector X would be the velocity vector $X = x'(t)$ of the curve $x(t)$. It is convenient to introduce the inclusion maps

$$\begin{array}{lll} i_{p_2}\colon M_1 & & \\ & \searrow & \\ & & (M_1 \times M_2) \\ & \nearrow & \\ i_{p_1}\colon M_2, & & \end{array}$$

defined by,

$$\begin{aligned} i_{p_2}(p) &= (p, p_2), \quad \text{for } p \in M_1, \\ i_{p_1}(q) &= (p_1, q), \quad \text{for } q \in M_2 \end{aligned}$$

The image of the vectors X_1 and X_2 under the push-forward of the inclusion maps

$$\begin{array}{lll} i_{p_2*} : T_{p_1}M_1 & & \\ & \searrow & \\ & & T_{(p_1,p_2)}(M_1 \times M_2) \\ & \nearrow & \\ i_{p_1*} : T_{p_2}M_2, & & \end{array}$$

yield vectors $\overline{X}_1$ and $\overline{X}_2$, given by,

$$\begin{aligned} i_{p_2*}(X_1) &= \overline{X}_1 = (x_1'(t), p_2), \\ i_{p_1*}(X_2) &= \overline{X}_1 = (p_1, x_2'(t)). \end{aligned}$$

Then, it is easy to show that,

$$X = i_{p_2*}(X_1) + i_{p_1*}(X_2).$$

Indeed, if f is a smooth function $f : M_1 \times M_2 \to \mathbf{R}$, we have,

$$\begin{aligned} X(f) &= \frac{d}{dt}(f(x_1(t), x_2(t))|_{t=0}, \\ &= \frac{d}{dt}(f(x_1(t), x_2(0))|_{t=0} + \frac{d}{dt}(f(x_1(0), x_2(t))|_{t=0}, \\ &= \overline{X}_1(f) + \overline{X}_2(f). \end{aligned}$$

More generally, if

$$\varphi : M_1 \times M_2 \to N$$

is a smooth manifold mapping, then we have a type of product rule formula for the Jacobian map,

$$\begin{aligned} \varphi_* X &= \varphi_*(i_{p_2 *}(X_1)) + \varphi_*(i_{p_1 *}(X_2)), \\ &= (\varphi \circ i_{p_2})_* X_1 + (\varphi \circ i_{p_1})_* X_2 \end{aligned} \tag{6.2}$$

This formula will be useful in the treatment of principal fiber bundles, in which case we have a bundle space E, and a Lie group G acting on the right by a product manifold map $\mu : E \times G \to E$.

6.2 Submanifolds

A *Riemannian submanifold* is a subset of a Riemannian manifold that is also Riemannian. The most natural example is a hypersurface in $\mathbf{R}^n$. If $(x^1, x^2 \dots x^n)$ are local coordinates in $\mathbf{R}^n$ with the standard metric, and the surface M is defined locally by functions $x^i = x^i(u^\alpha)$, then M together with the induced first fundamental form 4.12, 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

$$\overline{\nabla}_X Y = \nabla_X Y + H(X, Y), \tag{6.3}$$

where $H(X, Y)$ is the component in the normal space. In the case of a hypersurface, we have the classical Gauss equation 4.74

$$\overline{\nabla}_X Y = \nabla_X Y + II(X, Y)N \tag{6.4}$$

$$= \nabla_X Y + < LX, Y > N, \tag{6.5}$$

where $LX = -\overline{\nabla}_X N$ is the Weingarten map. If 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 theorema 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 of curvature that is applicable to any Riemannian manifold. The concept leading to the equations of the theorema egregium involved calculation of the difference of second derivatives of tangent vectors. At the risk of being somewhat misleading, figure 4.95 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 from one point on an spherical triangle to the diagonally opposed point, the result depends on the path taken. Parallel transport of Z

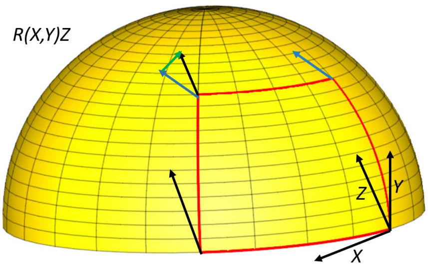


Fig. 6.2: R(X,Y)Z

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 a 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 are defined as quantities that depend only on direction and length. As indicated, the picture is misleading, because, covariant derivatives, as is the case with any other type of derivative, involve comparing the change of a vector under infinitesimal parallel transport. The failure of a vector to return to itself when parallel-transported along a closed path is measured by an entity related to the curvature called the *holonomy* of the connection. Still, the figure should help motivate the definition that follows.

6.2.1 Definition On a Riemannian manifold with connection ∇, the curvature R and the torsion T are defined by:

$$R(X,Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}, \tag{6.6}$$
$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]. \tag{6.7}$$

6.2.2 Theorem The Curvature R is a tensor. At each point $p \in M$, $R(X,Y)$ assigns to each pair of tangent vectors, a linear transformation from T_pM into itself.

Proof Let $X, Y, Z \in \mathscr{X}(M)$ be vector fields on M. We need to establish that R is muiltilinear. 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.

6.2.3 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,

$$\begin{aligned}
T(fX,Y) &= \nabla fXY - \nabla_Y (fX) - [fX,Y],\\
&= f\nabla_X Y - Y(f)X - f\nabla_Y X - fXY + Y(fX),\\
&= f\nabla_X Y - Y(f)X - f\nabla_Y X - fXY + Y(f)X + fYX,\\
&= f\nabla_X Y - f\nabla YX - f[X,Y],\\
&= fT(X,Y).
\end{aligned}$$

Again, linearity over sums is clear.

6.2.4 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:

$$[X,Y] = \nabla_X Y - \nabla_Y X, \tag{6.8}$$

$$\nabla_X < Y,Z >= < \nabla_X Y, Z > + < Y, \nabla_X Z > . \tag{6.9}$$

Proof The proof parallels the computation leading to equation 4.76.Let ∇ be a connection compatible with the metric. By taking the three cyclic derivatives of the inner product, and subtracting the third from the sum of the first two

$$\begin{aligned}
&(a) \quad \nabla_X < Y,Z >= < \nabla_X Y, Z > + < Y, \nabla_X Z >,\\
&(b) \quad \nabla_Y < X,Z >= < \nabla_Y X, Z > + < X, \nabla_Y Z >,\\
&(c) \quad \nabla_Z < X,Y >= < \nabla_Z X, Y > + < X, \nabla_Z Y >,\\
&\qquad (a)+(b)-(c) = < \nabla_X Y, Z > + < \nabla_Y X, Z > + < [X,Z], Y > + < [Y,Z], X >\\
&\qquad\qquad = 2 < \nabla_X Y, Z > + < [Y,X], Z > + < [X,Z], Y > + < [Y,Z], X >
\end{aligned}$$

Therefore:

$$\begin{aligned}
< \nabla_X Y, Z >= \tfrac{1}{2}\{&\nabla_X < Y,Z > + \nabla_Y < X,Z > - \nabla_Z < X,Y >\\
&+ < [X,Y], Z > + < [Z,X], Y > + < [Z,Y], X >\}.
\end{aligned} \tag{6.10}$$

The bracket of any two vector fields is a vector field, so the connection is unique since it is completely determined by the metric. In disguise, this is the formula in local coordinates for the Christoffel symbols 4.76. This follows immediately by choosing $X = \partial/\partial x^\alpha$, $Y = \partial/\partial x^\beta$ and $Z = \partial/\partial x^\gamma$. Conversely, if one defines $\nabla_X Y$ by equation 6.10, a long but straightforward computation with lots of cancellations, shows that this defines a connection compatible with the metric.

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 in the frame by

$$\nabla_X e_\beta = \omega^\gamma{}_\beta(X)\, e_\gamma, \tag{6.11}$$

$$\nabla_{e_\alpha} e_\beta = \Gamma^\gamma_{\alpha\beta}\, e_\gamma, \tag{6.12}$$

$$T(e_\alpha, e_\beta) = T^\gamma{}_{\alpha\beta}\, e_\gamma. \tag{6.13}$$

As was pointed out in the previous chapter, if the frame is an orthonormal frame such as the coordinate frame $\{\partial/\partial x^\mu\}$ for which 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:

$$\begin{aligned} R(e_\gamma, e_\beta)\, e_\delta &= \nabla_{e_\gamma}\nabla_{e_\beta} e_\delta - \nabla_{e_\beta}\nabla_{e_\gamma} e_\delta, \\ &= \nabla_{e_\gamma}(\Gamma^\alpha_{\beta\delta} e_\alpha) - \nabla_{e_\beta}(\Gamma^\alpha_{\gamma\delta} e_\alpha), \\ &= \Gamma^\alpha_{\beta\delta,\gamma} e_\alpha + \Gamma^\alpha_{\beta\delta}\Gamma^\mu_{\gamma\alpha} e_\mu - \Gamma^\alpha_{\gamma\delta,\beta} e_\alpha - \Gamma^\alpha_{\gamma\delta}\Gamma^\mu_{\beta\alpha} e_\mu, \\ &= [\Gamma^\alpha_{\beta\delta,\gamma} - \Gamma^\alpha_{\beta\gamma,\delta} + \Gamma^\mu_{\beta\delta}\Gamma^\alpha_{\gamma\mu} - \Gamma^\mu_{\beta\gamma}\Gamma^\alpha_{\delta\mu}] e_\alpha, \\ &= R^\alpha{}_{\beta\gamma\delta}\, e_\alpha, \end{aligned}$$

where the components of the Riemann Tensor are defined by:

$$R^\alpha{}_{\beta\gamma\delta} = \Gamma^\alpha_{\beta\delta,\gamma} - \Gamma^\alpha_{\beta\gamma,\delta} + \Gamma^\mu_{\beta\delta}\Gamma^\alpha_{\gamma\mu} - \Gamma^\mu_{\beta\gamma}\Gamma^\alpha_{\delta\mu}. \tag{6.14}$$

Let $X = X^\mu e_\mu$ be and $\alpha = X_\mu \theta^\mu$ be a covariant and a contravariant vector field respectively. Using the notation $\nabla_\alpha = \nabla_{e_\alpha}$ it is almost trivial to compute the covariant derivatives. The results are,

$$\begin{aligned} \nabla_\beta X &= (X^\mu_{,\beta} + X^\nu \Gamma^\mu{}_{\beta\nu}) e_\mu, \\ \nabla_\beta \alpha &= (X_{\mu,\beta} - X^\nu \Gamma^\nu{}_{\beta\mu}) \theta^\mu, \end{aligned} \tag{6.15}$$

We show the details of the first computation, and leave the second one as an easy exercise

$$\nabla_\beta X = \nabla_\beta (X^\mu e_\mu), \tag{6.16}$$

$$= X^\mu_{,\beta} e_\mu + X^\mu \Gamma^\delta_{\beta\mu} e_\delta, \tag{6.17}$$

$$= (X^\mu_{,\beta} + X^\nu \Gamma^\mu{}_{\beta\nu}) e_\mu. \tag{6.18}$$

In classical notation, the covariant derivatives $X^\mu{}_{\|\beta}$ and $X_{\mu\|\beta}$ are given in terms of the tensor components,

$$\begin{aligned} X^\mu{}_{\|\beta} &= X^\mu_{,\beta} + X^\nu \Gamma^\mu{}_{\beta\nu}, \\ X_{\mu\|\beta} &= X_{\mu,\beta} - X_\nu \Gamma^\nu{}_{\beta\mu}. \end{aligned} \tag{6.19}$$

It is also straightforward to establish the *Ricci identities*

$$\begin{aligned} X^\mu{}_{\|\alpha\beta} - X^\mu{}_{\|\beta\alpha} &= \quad X^\nu R^\mu{}_{\nu\alpha\beta}, \\ X_{\mu\|\alpha\beta} - X_{\mu\|\beta\alpha} &= -X_\nu R^\nu{}_{\mu\alpha\beta}. \end{aligned} \tag{6.20}$$

Again, we show the computation for the first identity and leave the second as a exercise. We take the second derivative. and then reverse the order,

$$\begin{aligned} \nabla_\alpha \nabla_\beta X &= \nabla_\alpha (X^\mu_{,\beta} e_\mu + X^\nu \Gamma^\mu{}_{\beta\nu} e_\mu), \\ &= X^\mu_{,\beta\alpha} e_\mu + X^\mu_{,\beta} \Gamma^\delta_{\alpha\mu} e_\delta + X^{\nu,\alpha} \Gamma^\mu_{\beta\mu} e_\nu + X^\nu \Gamma^\mu_{\beta\nu,\alpha} e_\mu + X^\nu \Gamma^\mu_{\beta\nu} \Gamma^\delta_{\alpha\nu} e_\delta, \\ \nabla_\alpha \nabla_\beta X &= (X^\mu_{,\beta\alpha} + X^\nu_{,\beta} \Gamma^\mu_{\alpha\nu} + X^\nu_{,\alpha} \Gamma^\nu_{\beta\nu} + X^\nu \Gamma^\mu_{\beta\nu,\alpha} + X^\nu \Gamma^\delta_{\beta\nu} \Gamma^\mu_{\alpha\delta}) e_\mu, \\ \nabla_\beta \nabla_\alpha X &= (X^\mu_{,\alpha\beta} + X^\nu_{,\alpha} \Gamma^\mu_{\beta\nu} + X^\nu_{,\beta} \Gamma^\nu_{\alpha\nu} + X^\nu \Gamma^\mu_{\alpha\nu,\beta} + X^\nu \Gamma^\delta_{\alpha\nu} \Gamma^\mu_{\beta\delta}) e_\mu. \end{aligned}$$

Subtracting the last two equations, only the last two terms of each survive, and we get the desired result,

$$\begin{aligned} 2\nabla_{[\alpha} \nabla_{\beta]}(X) &= X^\nu (\Gamma^\mu_{\beta\nu,\alpha} - \Gamma^\mu_{\alpha\nu,\beta} + \Gamma^\delta_{\beta\nu} \Gamma^\mu_{\alpha\delta} - \Gamma^\delta_{\alpha\nu} \Gamma^\mu_{\beta\delta}) e_\mu, \\ 2\nabla_{[\alpha} \nabla_{\beta]}(X^\mu e_\mu) &= \quad (X^\nu R^\mu{}_{\nu\alpha\beta}) e_\mu. \end{aligned}$$

In the lietrature, many authors use the notation $\nabla_\beta X^\mu$ to denote the covariant derivative $X^\mu{}_{\|\beta}$, but it is really an (excusable) abuse of notation that arises from thinking of tensors as the components of the tensors. The Ricci identities are the basis for the notion of holonomy, namely, the simple interpretation that the failure of parallel transport to commute along the edges of an rectangle, indicates the presence of curvature. With more effort with repeated use of Leibnitz rule, one can establish more elaborate Ricci identities for higher order tensors. If one assumes zero torsion, the Ricci identities of higher order tensors just involve more terms with the curvature. It the torsion is not zero, there are additional terms involving the torsion tensor; in this case it is perhaps a bit more elegant to use the covariant differential introduced in the next section, so we will postpone the computation until then.

The generalization of the theorema 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:

$$\begin{aligned}
\overline{R}(X,Y)Z &= \overline{\nabla}_X \overline{\nabla}_Y Z - \overline{\nabla}_Y \overline{\nabla}_X Z - \overline{\nabla}_{[X,Y]} Z, \\
&= \overline{\nabla}_X (\nabla_Y Z + < LY, Z > N) - \overline{\nabla}_Y (\nabla_X Z + < LX, Z > N) - \overline{\nabla}_{[X,Y]} Z, \\
&\quad \nabla_X \nabla_Y Z + < LX, \nabla_Y Z > N + X < LY, Z > N + < LY, Z > LX - \\
&\quad \nabla_Y \nabla_X Z - < LY, \nabla_Y Z > N - Y < LX, Z > N - < LX, Z > LY - \\
&\quad \nabla_{[X,Y]} Z - < L([X,Y]), Z > N, \\
&\quad \nabla_X \nabla_Y Z + < LX, \nabla_Y Z > N + X < LY, Z > N + < LY, Z > LX - \\
&\quad \nabla_Y \nabla_X Z - < LY, \nabla_Y Z > N - Y < LX, Z > N - < LX, Z > LY - \\
&\quad \nabla_{[X,Y]} Z - < L([X,Y]), Z > N, \\
&= \nabla_X \nabla_Y Z + < LX, \nabla_Y Z > N + < \nabla_X LY, Z > N + < LY, \nabla_X Z > N + < LY, Z > LX - \\
&\quad \nabla_Y \nabla_X Z - < LY, \nabla_Y Z > N - < \nabla_Y LX, Z > N - < LX, \nabla_Y Z > N - < LX, Z > LY - \\
&\quad \nabla_{[X,Y]} Z - < L([X,Y]), Z > N, \\
&= R(X,Y)Z + < LY, Z > LX - < LX, Z > LY + \\
&\quad \{ < \nabla_X LY, Z > - < \nabla_Y LX, Z > - < L([X,Y]), Z > \} N.
\end{aligned}$$

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:

$$R(X,Y)Z + < LY, Z > LX - < LX, Z > LY = 0, \tag{6.21}$$
$$\nabla_X LY - \nabla_Y LX - L([X,Y]) = 0. \tag{6.22}$$

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 theorema egregium

$$K =< R(X,Y)X, Y >=< LX, X >< LY, Y > - < LY, X >< LX, Y >= \det(L). \tag{6.23}$$

The expression 6.22 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.

6.2.5 Theorem Equations of Structure.

$$\Theta^\alpha = d\theta^\alpha + \omega^\alpha{}_\beta \wedge \theta^\beta, \tag{6.24}$$
$$\Omega^\alpha{}_\beta = d\omega^\alpha{}_\beta + \omega^\alpha{}_\gamma \wedge \omega^\gamma{}_\beta. \tag{6.25}$$

To verify this is the case, we define:

$$T(X,Y) = \Theta^\alpha(X,Y) e_\alpha, \tag{6.26}$$
$$R(X,Y) e_\beta = \Omega^\alpha{}_\beta(X,Y) e_\alpha. \tag{6.27}$$

Recalling that any tangent vector X can be expressed in terms of the basis as

$X = \theta^\alpha(X)\, e_\alpha$, we can carry out a straight-forward computation:

$$\begin{aligned}
\Theta^\alpha(X,Y)\, e_\alpha = T(X,Y) &= \nabla_X Y - \nabla_Y X - [X,Y], \\
&= \nabla_X(\theta^\alpha(Y))\, e_\alpha - \nabla_Y(\theta^\alpha(X))\, e_\alpha - \theta^\alpha([X,Y])\, e_\alpha, \\
&= X(\theta^\alpha(Y))\, e_\alpha + \theta^\alpha(Y)\, \omega^\beta{}_\alpha(X)\, e_\beta - Y(\theta^\alpha(X))\, e_\alpha \\
&\quad - \theta^\alpha(X)\, \omega^\beta{}_\alpha(Y)\, e_\beta - \theta^\alpha([X,Y])\, e_\alpha, \\
&= \{X(\theta^\alpha(Y)) - Y(\theta^\alpha(X)) - \theta^\alpha([X,Y]) + \omega^\alpha{}_\beta(X)(\theta^\beta(Y) - \omega^\alpha{}_\beta(Y)(\theta^\beta(X)\}e_\alpha, \\
&= \{(d\theta^\alpha + \omega^\alpha{}_\beta \wedge \theta^\beta)(X,Y)\}e_\alpha,
\end{aligned}$$

where we have introduced a coordinate-free definition of the differential of the one form θ by

$$d\theta(X,Y) = X(\theta(Y)) - y(\theta(X)) - \theta([X,Y]). \tag{6.28}$$

It is easy to verify that this definition of the differential of a one form satisfies all the required properties of the exterior derivative, and that it is consistent with the coordinate version of the differential introduced in Chapter 2. We conclude that

$$\Theta^\alpha = d\theta^\alpha + \omega^\alpha{}_\beta \wedge \theta^\beta, \tag{6.29}$$

which is indeed the first Cartan equation of structure. Proceeding along the same lines, we compute:

$$\begin{aligned}
\Omega^\alpha{}_\beta(X,Y)\, e_\alpha &= \nabla_X \nabla_Y\, e_\beta - \nabla_Y \nabla_X\, e_\beta - \nabla_{[X,Y]}\, e_\beta, \\
&= \nabla_X(\omega^\alpha{}_\beta(Y)\, e_\alpha) - \nabla_Y(\omega^\alpha{}_\beta(X)\, e_\alpha) - \omega^\alpha{}_\beta([X,Y])\, e_\alpha, \\
&= X(\omega^\alpha{}_\beta(Y))\, e_\alpha + \omega^\alpha{}_\beta(Y)\, \omega^\gamma{}_\alpha(X)\, e_\gamma - Y(\omega^\alpha{}_\beta(X))\, e_\alpha \\
&\quad - \omega^\alpha{}_\beta(X)\, \omega^\gamma{}_\alpha(Y)\, e_\gamma - \omega^\alpha{}_\beta([X,Y])\, e_\alpha \\
&= \{(d\omega^\alpha{}_\beta + \omega^\alpha{}_\gamma \wedge \omega^\gamma{}_\beta)(X,Y)\}e_\alpha,
\end{aligned}$$

thus arriving at the second equation of structure

$$\Omega^\alpha{}_\beta = d\omega^\alpha{}_\beta + \omega^\alpha{}_\gamma \wedge \omega^\gamma{}_\beta. \tag{6.30}$$

The quantities connection and curvature forms are matrix-valued. Using matrix multiplication notation, we can abbreviate the equations of structure as

$$\begin{aligned}
\Theta &= d\theta + \omega \wedge \theta, \\
\Omega &= d\omega + \omega \wedge \omega.
\end{aligned} \tag{6.31}$$

Taking the exterior derivative of the structure equations gives some interesting results. Here is the first computation,

$$\begin{aligned}
d\Theta &= d\omega \wedge \theta - \omega \wedge d\theta, \\
&= d\omega \wedge \theta - \omega \wedge (\Theta - \omega \wedge \theta), \\
&= d\omega \wedge \theta - \omega \wedge \Theta + \omega \wedge \omega\theta, \\
&= (d\omega + \omega \wedge \omega) \wedge \theta - \omega \wedge \Theta, \\
&= \Omega \wedge \theta - \omega \wedge \Theta,
\end{aligned}$$

so,

$$d\Theta + \omega \wedge \Theta = \Omega \wedge \theta. \tag{6.32}$$

Similarly, taking d of the second structure equation we get,

$$\begin{aligned} d\Omega &= d\omega \wedge \omega + \omega \wedge d\omega, \\ &= (\Omega - \omega \wedge \omega) \wedge \omega + \omega \wedge (\Omega - \omega \wedge \omega). \end{aligned}$$

Hence,

$$d\Omega = \Omega \wedge \omega - \omega \wedge \Omega. \tag{6.33}$$

Equations 6.32 and 6.33 are called the first and second Bianchi identities. The relationship between the torsion and Riemann tensor components with the corresponding differential forms are given by

$$\begin{aligned} \Theta^\alpha &= \tfrac{1}{2} T^\alpha{}_{\gamma\delta}\, \theta^\gamma \wedge \theta^\delta, \\ \Omega^\alpha{}_\beta &= \tfrac{1}{2} R^\alpha{}_{\beta\gamma\delta}\, \theta^\gamma \wedge \theta^\delta. \end{aligned} \tag{6.34}$$

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

$$[e_\beta, e_\gamma] = C^\alpha_{\beta\gamma}\, e_\alpha. \tag{6.35}$$

The components of the torsion and Riemann tensors are then given by

$$\begin{aligned} T^\alpha{}_{\beta\gamma} &= \Gamma^\alpha_{\beta\gamma} - \Gamma^\alpha_{\gamma\beta} - C^\alpha_{\beta\gamma}, \\ R^\alpha{}_{\beta\gamma\delta} &= \Gamma^\alpha_{\beta\delta,\gamma} - \Gamma^\alpha_{\beta\gamma,\delta} + \Gamma^\mu_{\beta\delta}\Gamma^\alpha_{\gamma\mu} - \Gamma^\mu_{\beta\gamma}\Gamma^\alpha_{\delta\mu} - \Gamma^\alpha_{\beta\mu}C^\mu_{\gamma\delta} - \Gamma^\alpha{}_{\sigma\beta}C^\sigma{}_{\gamma\delta}. \end{aligned} \tag{6.36}$$

The Riemann tensor for a torsion-free connection has the following symmetries;

$$\begin{aligned} R(X,Y) &= -R(Y,X), \\ < R(X,Y)Z, W > &= - < R(X,Y)W, Z >, \\ R(X,Y)Z + R(Z,X)Y + R(Y,Z)X &= 0. \end{aligned} \tag{6.37}$$

In terms of components, the Riemann Tensor symmetries can be expressed as

$$\begin{aligned} R_{\alpha\beta\gamma\delta} &= -R_{\alpha\beta\delta\gamma} = -R_{\beta\alpha\gamma\delta}, \\ R_{\alpha\beta\gamma\delta} &= R_{\gamma\delta\alpha\beta}, \\ R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\delta\beta\gamma} &= 0. \end{aligned} \tag{6.38}$$

The last cyclic equation is the tensor version of the first Bianchi Identity with 0 torsion. It follows immediately from setting $\Omega \wedge \theta = 0$ and taking a cyclic permutation of the antisymmetric indices $\{\beta, \gamma, \delta\}$ of the Riemann tensor. 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 a 4-dimensional space, there are at most 20 independent components. The derivation of the tensor version of the second Bianchi identity from the elegant differential forms version, takes a bit more effort. In components the formula

$$d\Omega = \Omega \wedge \omega - \omega \wedge \Omega$$

reads,

$$R^{\alpha}{}_{\beta\kappa\lambda;\mu}\,\theta^{\mu} \wedge \theta^{\kappa} \wedge \theta^{\lambda} = (\Gamma^{\rho}{}_{\mu\beta}R^{\alpha}{}_{\rho\kappa\lambda} - \Gamma^{\alpha}{}_{\mu\rho}R^{\rho}{}_{\beta\kappa\lambda})\theta^{\mu} \wedge \theta^{\kappa} \wedge \theta^{\lambda},$$

where we used the notation,

$$\nabla_{\mu}R^{\alpha}{}_{\beta\kappa\lambda} = R^{\alpha}{}_{\beta\kappa\lambda;\mu}.$$

Taking a cyclic permutation on the antisymmetric indices κ, λ, μ, and using some index gymnastics to show that the right hand becomes zero, the tensor version of the second Bianchi identity for zero torsion becomes

$$R^{\alpha}{}_{\beta[\kappa\lambda;\mu]} = 0 \tag{6.39}$$

6.3 Sectional Curvature

Let $\{M, g\}$ be a Riemannian manifold with Levi-Civita connection ∇ and curvature tensor $R(X, Y)$. In local coordinates at a point $p \in M$ we can express the components

$$R = R_{\mu\nu\rho\sigma}\, dx^{\mu}dx^{\nu}dx^{\rho}dx^{\sigma}$$

of a covariant tensor of rank 4. With this in mind, we define a multilinear function

$$R : T_p(M) \otimes T_p(M) \otimes T_p(M) \otimes T_p(M) \to \mathbf{R},$$

by

$$R(W, Y, X, Z) =< W, R(X, Y)Z > \tag{6.40}$$

In this notation, the symmetries of the tensor take the form,

$$\begin{aligned} R(W, X, Y, Z) &= -R(W, Y, X, Z), \\ R(W, X, Y, Z) &= -R(Z, Y, X, W) \\ R(W, X, Y, Z) + R(W, Z, X, Y) + R(X, Y, Z, X) &= 0. \end{aligned} \tag{6.41}$$

From the metric, we can also define a multilinear function

$$G(W, Y, X, Z) =< Z, Y >< X, W > - < Z, X >< Y, W > .$$

Now, consider any 2-dimensional plane $V_p \subset T_p(M)$ and let $X, Y \in V$ be linearly independent. Then,

$$G(X, Y, X, Y) =< X, X >< Y, Y > - < X, Y >^2$$

is a bilinear form that represents the area of the parallelogram spanned by X and Y. If we perform a linear, non-singular change of of coordinates,

$$X' = aX + bY, \quad y' = cX + dY, \quad ad - bc \neq 0,$$

then both, $G(X, Y.X, Y)$ and $R(X, Y, X, Y)$ transform by the square of the determinant $D = ad - bc$, so the ratio is independent of the choice of vectors. We define the *sectional curvature* of the subspace V_p by

$$\begin{aligned} K(V_p) &= \frac{R(X,Y,X,Y)}{G(X,Y,X,Y)}, \\ &= \frac{R(X,Y,X,Y)}{< X, X >< Y, Y > - < X, Y >^2} \end{aligned} \tag{6.42}$$

The set of values of the sectional curvatures for all planes at $T_p(M)$ completely determines the Riemannian curvature at p. For a surface in $\mathbf{R}^3$ the sectional curvature is the Gaussian curvature, and the formula is equivalent to the theorema egregium. If $K(V_p)$ is constant for all planes $V_p \in T_p(M)$ and for all points $p \in M$, we say that M is a space of *constant curvature.* For a space of constant curvature k, we have

$$R(X, Y)Z = k(< Z, Y > X - < Z, X > Y) \tag{6.43}$$

In local coordinates, the equation gives

$$R_{\mu\nu\rho\sigma} = k(g_{\nu\sigma} g_{\mu\rho} - g_{\nu\gamma} g_{\mu\sigma}). \tag{6.44}$$

6.3.1 Example

The model space of manifolds of constant curvature is a quadric hypersurface M of $\mathbf{R}^{n+1}$ with metric

$$ds^2 = \epsilon k^2 dt^2 + (dy^1)^2 + \cdots + dy^n)^2,$$

given by the equation

$$M : \epsilon k^2 t^2 + (y^1)^2 + \ldots (y^n)^2 = \epsilon k^2, \quad t \neq 0,$$

where k is a constant and $\epsilon = \pm 1$. For the purposes of this example it will actually be simpler to completely abandon the summation convention. Thus, we write the quadric as

$$\epsilon k^2 t^2 + \Sigma_i (y^i)^2 = \epsilon k^2.$$

If $k = 0$, the space is flat. If $\epsilon = 1$, let $(y^0)^2 = k^2 t^2$ and the quadric is isometric to a sphere of constant curvature $1/k^2$. If $\epsilon = -1$, $\Sigma_i (x^i)^2 = -k^2(1 - t^2) > 0$, then $t^2 < 1$ and the surface is a hyperboloid of two sheets. Consider the mapping from $(\mathbf{R})^{\mathbf{n+1}}$ to $\mathbf{R}^n$ given by

$$x^i = y^i / t.$$

We would like to compute the induced metric on the the surface. We have

$$-k^2t^2 + \Sigma_i(y^i)^2 = -k^2t^2 + t^2\Sigma_i(x^i)^2 = -k^2$$

so

$$t^2 = \frac{-k^2}{-k^2 + \Sigma_i(x^i)^2}.$$

Taking the differential, we get

$$t\ dt = \frac{-k^2\Sigma_i(x^i dx^i)}{-k^2 + \Sigma_i(x^i)^2}.$$

Squaring and dividing by t^2 we also have

$$dt^2 = \frac{-k^2(\Sigma_i x^i\, dx^i)^2}{(-k^2 + \Sigma_i(x^i)^2)^3}.$$

From the product rule, we have $dy^i = x^i\, dt + t\, dx^i$, so the metric is

$$\begin{aligned} ds^2 &= -k^2\, dt^2 + [\Sigma(x^i)^2]dt^2 + 2t\, dt\, \Sigma_i(x^i\, dx^i) + t^2\Sigma(dx^i)^2, \\ &= [-k^2 + \Sigma_i(x^i)^2]\, dt^2 + 2t\, dt\, \Sigma_i(x^i\, dx^i) + t^2\Sigma(dx^1)^2, \\ &= \frac{-k^2[\Sigma_i(x^i dx^i)]^2}{[-k^2 + \Sigma_i(x^i)^2]^2} + \frac{2k^2[\Sigma_i(x^i dx^i)]^2}{[-k^2 + \Sigma_i(x^i)^2]^2} + \frac{-k^2\Sigma_i(dx^i)^2}{-k^2 + \Sigma_i(dx^i)^2}, \\ &= k^2\frac{[k^2 - \Sigma_i(x^i)^2]\Sigma_i(dx^i)^2 - (\Sigma_i(x^i\, dx^i))^2}{[k^2 - \Sigma_i(x^i)^2]^2} \end{aligned}$$

It is not obvious, but in fact, the space is also of constant curvature $(-1/k^2)$. For an elegant proof, see [18]. When $n = 4$ and $\epsilon = -1$, the group leaving the metric

$$ds^2 = -k^2dt^2 + (dy^1)^2 + (dy^2)^2 + (dy^3)^2 + (dy^4)^2$$

invariant, is the Lorentz group $O(1,4)$. With a minor modification of the above, consider the quadric

$$M : -k^2t^2 + (y^1)^2 + \ldots (y^4)^2 = k^2.$$

In this case, the quadric is a hyperboloid of one sheet, and the submanifold with the induced metric is called the *de Sitter* space. The isotropy subgroup that leaves $(1,0,0,0,0)$ fixed is $O(1,3)$ and the manifold is diffeomorphic to $O(1,4)/O(1,3)$. Many alternative forms of the de Sitter metric exist in the literature. One that is particularly appealing is obtained as follows. Write the metric in ambient space as

$$ds^2 = -(dy^0)^2 + (dy^1)^2 + (dy^2)^2 + (dy^3)^2 + (dy^4)^2$$

with the quadric given by

$$M : -(y^0)^2 + (y^1)^2 + \ldots (y^4)^2 = k^2.$$

Let

$$\sum_{i=1}^{4}(x^i)^2 = 1$$

so M represents a unit sphere S^3. Introduce the coordinates for M

$$\begin{aligned} y^0 &= k\sinh(\tau/k), \\ y^i &= k\cosh(\tau/k). \end{aligned}$$

Then, we have

$$\begin{aligned} dy^0 &= \cosh(\tau/k)\,d\tau, \\ dy^i &= \sinh(\tau/k)x^i\,d\tau + k\cosh(\tau/k)\,dx^i. \end{aligned}$$

The induced metric on M becomes,

$$\begin{aligned} ds^2 &= -[\cosh^2(\tau/k) - \sinh^2(\tau/k)\Sigma_i(x^i)^2]\,d\tau + \cosh^2(\tau/k)\Sigma_i(dx^i)^2, \\ &= -d\tau^2 + \cosh^2(\tau/k)d\Omega^2, \end{aligned}$$

where $d\Omega$ is the volume form for S^3. The most natural coordinates for the volume form are the Euler angles and Cayley-Klein parameters. The interpretation of this space-time is that we have a spatial 3-sphere which propagates in time by shrinking to a minimum radius at the throat of the hyperboloid, followed by an expansion. Being a space of constant curvature, the Ricci tensor is proportional to the metric, so this is an Einstein manifold.

6.4 Big D

In this section we discuss the notion of a connection on a vector bundle E. Let M be a smooth manifold and as usual we denote by $T^r_s(p)$ the vector space of type $\binom{r}{s}$ tensors at a point $p \in M$. The formalism applies to any vector bundle, but in this section we are primarily concerned with the case where E is the tensor bundle $E = T^r_s(M)$. Sections $\Gamma(E) = \mathscr{T}^r_s(M)$ of this bundle are called tensor fields on M. For general vector bundles, we use the notation $s \in \Gamma(E)$ for the sections of the bundle. The section that maps every point of M to the zero vector, is called the *zero section*. Let $\{e_\alpha\}$ be an orthonormal frame with dual forms $\{\theta^\alpha\}$. We define the space $\Omega^p(M, E)$ tensor-valued p-form as sections of the bundle,

$$\Omega^p(M,E) = \Gamma(E \otimes \Lambda^p(M)). \tag{6.45}$$

As in equation 2.63, a tensor-valued p form is a tensor of type $\binom{r}{s+p}$ with components,

$$T = T^{\alpha_1,\ldots\alpha_r}_{\beta_1,\ldots\beta_s,\gamma_1,\ldots,\gamma_p} e_{\alpha_1} \otimes \ldots e_{\alpha_r} \otimes \theta^{\beta_1} \otimes \cdots \otimes \theta^{\beta_s} \wedge \theta^{\gamma_1} \wedge \ldots \wedge \theta^{\gamma_p}. \tag{6.46}$$

A tensor-valued 0-form is just a regular tensor field $T \in \mathscr{T}^r_s(M)$ The main examples of tensor-valued forms are the torsion and the curvature forms

$$\begin{aligned}\Theta &= \Theta^\alpha \otimes e_\alpha,\\ \Omega &= \Omega^\alpha{}_\beta \otimes e_\alpha \otimes \theta^\beta.\end{aligned} \tag{6.47}$$

The tensorial components of the torsion tensor, would then be written as

$$\begin{aligned}T &= T^\alpha{}_{\beta\gamma} e_\alpha \otimes \theta^\beta \otimes \theta^\gamma,\\ &= \frac{1}{2} T^\alpha{}_{\beta\gamma} e_\alpha \otimes \theta^\beta \wedge \theta^\gamma,\\ &= e_\alpha \otimes (\tfrac{1}{2} T^\alpha{}_{\beta\gamma} \theta^\beta \wedge \theta^\gamma).\end{aligned}$$

since the tensor is antisymmetric in the lower indices. Similarly, the tensorial components of the curvature are

$$\begin{aligned}\Omega &= \frac{1}{2} R^\alpha{}_{\beta\gamma\delta}\, e_\alpha \otimes \theta^\beta \otimes \theta^\gamma \wedge \theta^\delta,\\ &= \ e_\alpha \otimes \theta^\beta \otimes (\tfrac{1}{2} R^\alpha{}_{\beta\gamma\delta} \theta^\gamma \wedge \theta^\delta).\end{aligned}$$

The connection forms

$$\omega = \omega^\alpha{}_\beta \otimes e_\alpha \otimes \theta^\beta \tag{6.48}$$

are matrix-valued, but they are not tensorial forms. If T is a type $\binom{r}{s}$ tensor field, and α a p-form, we can write a tensor-valued p-form as $T \otimes \alpha \in \Omega^p(M, E)$ is. We seek an operator that behaves like a covariant derivative ∇ for tensors and exterior derivative d for forms.

6.4.1 Linear Connections

Given a vector field X and a smooth function f, we define a *linear connection* as a map

$$\nabla_X : \Gamma(T^r_s) \to \Gamma(T^r_s)$$

with the following properties

1) $\nabla_X(f) = X(f)$,
1) $\nabla_{fX} T = f D_X T$,
2) $\nabla_{X+Y} T = \nabla_X T + \nabla_Y T$, for all $X, Y \in \mathscr{X}(M)$,
3) $\nabla_X(T_1 + T_2) = \nabla_X T_1 + \nabla_X T_2$, for $T_1, T_2 \in \Gamma(T^r_s)$,
4) $\nabla_X(fT) = X(f)T + f\nabla_X T$.

If instead of the tensor bundle we have a general vector bundle E, we replace the tensor fields in the definition above by sections $s \in \Gamma(E)$ of the vector bundle. The definition induces a derivation on the entire tensor algebra satisfying the additional conditions,

5) $\nabla_X(T_1 \otimes T_2) = \nabla_X T_1 \otimes T_2 + T_1 \otimes \nabla_X T_2$,
6) $\nabla_X \circ C = C \circ \nabla_X$, for any contraction C.

The properties are the same as a Koszul connection, or covariant derivative for tensor-valued 0 forms T. Given an orthonormal frame, consider the identity tensor,

$$I = \delta^\alpha{}_\beta\, e_\alpha \otimes \theta^\beta, \tag{6.49}$$

and take the covariant derivative ∇_X. We get

$$\begin{aligned}
\nabla_X e_\alpha \otimes \theta^\alpha + e_\alpha \otimes \nabla_X \theta^\alpha &= 0, \\
e_\alpha \otimes \nabla_X \theta^\alpha &= -\nabla_X e_\alpha \otimes \theta^\alpha, \\
&= -e_\beta\, \omega^\beta{}_\alpha(X) \otimes \theta^\alpha, \\
e_\beta \otimes \nabla_X \theta^\beta &= -e_\beta\, \omega^\beta{}_\alpha(X) \otimes \theta^\alpha,
\end{aligned}$$

which implies that,

$$\nabla_X \theta^\beta = -\omega^\beta{}_\alpha(X)\theta^\alpha. \tag{6.50}$$

Thus as before, since we have formulas for the covariant derivative of basis vectors and forms, we are led by induction to a general formula for the covariant derivative of an $\binom{r}{s}$-tensor given mutatis mutandis by the formula 3.32. In other words, the covariant derivative of a tensor acquires a term with a multiplicative connection factor for each contravariant index and a negative term with a multiplicative connection factor for each covariant index.

6.4.1 Definition A connection ∇ on the vector bundle E is a map

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

which satisfies the following conditions

a) $\nabla(T_1 + T_2) = \nabla T_1 + \nabla T_2, \quad T_1, T_2 \in \Gamma(E)),$
b) $\nabla(fT) = df \otimes T + f\nabla T,$
c) $\nabla_X T = i_X \nabla T.$

As a reminder of the definition of the inner product i_X, condition (c) is equivalent to the equation,

$$\nabla T(\theta^1, \ldots, \theta^r, X, X_1, \ldots, X_s) = (\nabla_X T)(\theta^1, \ldots, \theta^r, X_1, \ldots, X_s).$$

In particular, if X is vector field, then, as expected

$$\nabla X(Y) = \nabla_X Y,$$

The operator ∇ is called the *covariant differential.* Again, for a general vector bundles, we denote the sections by $s \in \Gamma(E)$ and the covariant differential by ∇s.

6.4.2 Affine Connections

A connection on the tangent bundle $T(M)$ is called an *affine connection.* In a local frame field e, we may assume that the connection is represented by a matrix of one-forms ω

$$\begin{aligned}
\nabla e_\beta &= e_\alpha \otimes \omega^\alpha{}_\beta, \\
\nabla e &= e \otimes \omega.
\end{aligned} \tag{6.51}$$

The tensor multiplication symbol is often omitted when it is clear in context. Thus, for example, the connection equation is sometimes written as $\nabla e = e\omega$.

In a local coordinate system $\{x_1,\ldots,x^n\}$, with basis vectors $\partial_\mu = \frac{\partial}{\partial x^\mu}$ and dual forms dx^μ, we have,

$$\omega^\alpha{}_{\beta\mu} = \Gamma^\alpha_{\beta\mu}\, dx^\mu.$$

From equation 6.50, it follows that

$$\nabla\theta^\alpha = -\omega^\alpha{}_\beta \otimes \theta^\beta. \tag{6.52}$$

We need to be a bit careful with the dual forms θ^α. We can view them as a vector-valued 1-form

$$\theta = e_\theta \otimes \theta^\alpha,$$

which has the same components as the identity $\binom{1}{1}$ tensor. This is a kind of a odd creature. About the closest analog to this entity in classical terms is the differential of arc-length

$$d\mathbf{x} = \mathbf{i}\,dx + \mathbf{j}\,dy + \mathbf{k}\,dz,$$

which is sort of a mixture of a vector and a form. The vector of differential forms would then be written as a column vector.
In a frame $\{e_\alpha\}$, the covariant differential of tensor-valued 0-form T is given by

$$\nabla T = \nabla_{e_\alpha} T \otimes \theta^\alpha \equiv \nabla_\alpha T \otimes \theta^\alpha.$$

In particular, if $X = v^\alpha e_\alpha$, we get,

$$\begin{aligned}
\nabla X &= \nabla_\beta X \otimes \theta^\beta = \nabla_\beta(v^\alpha e_\alpha) \otimes \theta^\beta,\\
&= (\nabla_\beta(v^\alpha)\; e_\alpha + v^\alpha \Gamma^\gamma_{\beta\alpha}\; e_\gamma) \otimes \theta^\beta,\\
&= (v^\alpha_{,\beta} + v^\gamma \Gamma^\alpha_{\beta\gamma}) e_\alpha \otimes \theta^\beta\\
&= v^\alpha_{\|\beta}\, e_\alpha \otimes \theta^\beta,
\end{aligned}$$

where we have used the classical symbols

$$v^\alpha{}_{\|\beta} = v^\alpha{}_{,\beta} + \Gamma^\alpha_{\beta\gamma}\, v^\gamma, \tag{6.53}$$

for the covariant derivative components $v^\alpha_{\|\beta}$ and the comma to abbreviate the directional derivative $\nabla_\beta(v^\alpha)$. Of course, the formula is in agreement with equation 3.25. ∇X is a $\binom{1}{1}$-tensor.
Similarly, for a covariant vector field $\alpha = v_\alpha \theta^\alpha$, we have

$$\begin{aligned}
\nabla\alpha &= \nabla(v_\alpha \otimes \theta^\alpha)\\
&= \nabla v_\alpha \otimes \theta^\alpha - v_\beta\, \omega^\beta{}_\alpha \otimes \theta^\alpha,\\
&= (\nabla_\gamma v_\alpha\; \theta^\gamma - v_\beta \Gamma^\beta_{\alpha\gamma} \theta^\gamma) \otimes \theta^\alpha,\\
&= (v_{\alpha,\gamma} - \Gamma^\alpha_{\beta\gamma}\, v_\alpha)\; \theta^\gamma \otimes \theta^\alpha,
\end{aligned}$$

hence,

$$v_{\alpha\|\beta} = v_{\alpha,\gamma} - \Gamma^\alpha_{\beta\gamma}\, v_\alpha. \tag{6.54}$$

As promised earlier, we now prove the Ricci identities for contravariant and covariant vectors when the torsion is not zero. Ricci Identities with torsion. The results are,

$$
\begin{aligned}
X^{\mu}{}_{\|\alpha\beta} - X^{\mu}{}_{\|\beta\alpha} &= \quad X^{\nu} R^{\mu}{}_{\nu\alpha\beta} - X^{\mu}{}_{,\nu} T^{\nu}{}_{\alpha\beta}, \\
X_{\mu\|\alpha\beta} - X_{\mu\|\beta\alpha} &= -X_{\nu} R^{\nu}{}_{\mu\alpha\beta} - X_{\mu,\nu} T^{\nu}{}_{\alpha\beta},
\end{aligned}
\tag{6.55}
$$

We prove the first one. Let $X = X^{\mu} e_{\mu}$. We have

$$
\begin{aligned}
\nabla X &= \nabla_{\beta} X \otimes \theta^{\beta}, \\
\nabla^2 X &= \nabla(\nabla_{\beta} X \otimes \theta^{\beta}), \\
&= \nabla(\nabla_{\beta} X) \otimes \theta^{\beta} + \nabla_{\beta} X \otimes \nabla\theta^{\beta}, \\
&= \nabla_{\alpha}\nabla_{\beta} X \otimes \theta^{\beta} \otimes \theta^{\alpha} - \nabla_{\beta} \otimes \omega^{\beta}{}_{\alpha} \otimes \theta^{\alpha}, \\
&= \nabla_{\alpha}\nabla_{\beta} X \otimes \theta^{\beta} \otimes \theta^{\alpha} - \nabla_{\mu} X \otimes \Gamma^{\mu}_{\alpha\beta}\theta^{\beta} \otimes \theta^{\alpha}, \\
\nabla^2 X &= (\nabla_{\alpha}\nabla_{\beta} X - \nabla_{\mu} X\, \Gamma^{\mu}_{\alpha\beta})\, \theta^{\beta} \otimes \theta^{\alpha}.
\end{aligned}
$$

On the other hand, we also have $\nabla X = \nabla_{\alpha} X \otimes \theta^{\alpha}$, so we can compute ∇^2 by differentiating in the reverse order to get the equivalent expression,

$$
\nabla^2 X = (\nabla_{\beta}\nabla_{\alpha} X - \nabla_{\mu} X\, \Gamma^{\mu}_{\beta\alpha})\, \theta^{\alpha} \otimes \theta^{\beta}.
$$

Subtracting the last two equations we get an alternating tensor, or a two-form that we can set equal to zero. For lack of a better notation we call this form $[\nabla, \nabla]$. The notations $Alt(\nabla^2)$ and $\nabla \wedge \nabla$ also appear in the literature. We get

$$
\begin{aligned}
[\nabla, \nabla] &= [\nabla_{\alpha}\nabla_{\beta} - \nabla_{\beta}\nabla_{\alpha})X - \nabla_{\mu} X(\Gamma^{\mu}_{\alpha\beta} - \Gamma^{\mu}_{\beta\alpha})]\theta^{\beta} \wedge \theta^{\alpha}, \\
&= [\nabla_{\alpha}\nabla_{\beta} - \nabla_{\beta}\nabla_{\alpha} - \nabla_{[e_{\alpha},\beta]})X + \nabla_{[e_{\alpha},\beta]} X - \nabla_{\mu} X(\Gamma^{\mu}_{\alpha\beta} - \Gamma^{\mu}_{\beta\alpha})]\theta^{\beta} \wedge \theta^{\alpha}, \\
&= [R(e_{\alpha}, e_{\beta})X + C^{\mu}_{\alpha\beta}\nabla_{\mu} X - \nabla_{\mu} X(\Gamma^{\mu}_{\alpha\beta} - \Gamma^{\mu}_{\beta\alpha}]\theta^{\beta} \wedge \theta^{\alpha}, \\
&= [R(e_{\alpha}, e_{\beta})X + C^{\mu}_{\alpha\beta} - \nabla_{\mu} X(\Gamma^{\mu}_{\alpha\beta} - \Gamma^{\mu}_{\beta\alpha} - C^{\mu}_{\alpha\beta}]\theta^{\beta} \wedge \theta^{\alpha}, \\
&= \tfrac{1}{2}(X^{\nu} R^{\mu}{}_{\nu\alpha\beta} - \nabla_{\mu} X\, T^{\mu}{}_{\alpha\beta})\theta^{\beta} \wedge \theta^{\alpha}.
\end{aligned}
$$

6.4.3 Exterior Covariant Derivative

Since we know how to take the covariant differential of the basis vectors, covectors, and tensor products thereof, an affine connection on the tangent bundle induces a covariant differential on the tensor bundle. It is easy to get a formula by induction for the covariant differential of a tensor-valued 0-form. A given connection can be extended in a unique way to to tensor-valued p-forms. Just as with the wedge product of a 0-form f with a p-form α for which identify $f\alpha$ with $f \otimes \alpha = f \wedge \alpha$, we write a tensor-valued p form as $T \otimes \alpha = T \wedge \alpha$, where T is a type $\binom{r}{s}$ tensor. We define the *exterior covariant derivative*

$$
D : \Omega^p(M, E) \to \Omega^{p+1}(M, E)
$$

by requiring that,

$$\begin{aligned} D(T\otimes\alpha) &= D(T\wedge\alpha),\\ &= \nabla T\wedge\alpha + (-1)^p T\wedge d\alpha. \end{aligned} \tag{6.56}$$

it is instructive to show the details of the computation of the exterior covariant derivative of the vector valued one forms θ and Θ, and the $\binom{1}{1}$ tensor-valued 2-form Ω. The results are

$$\begin{aligned} D\theta^\alpha &= d\theta^\alpha + \omega^\alpha{}_\beta\wedge\theta^\beta,\\ D\Theta^\alpha &= d\Theta^\alpha + \omega^\alpha{}_\beta\wedge\Theta^\beta,\\ D\Omega^\alpha{}_\beta &= d\Omega^\alpha{}_\beta + \omega^\alpha{}_\gamma\wedge\Omega^\alpha{}_\gamma - \Omega^\alpha{}_\gamma\wedge\omega^\gamma{}_\beta \end{aligned} \tag{6.57}$$

The first two follow immediately, we compute the third. We start by writing

$$\begin{aligned} \Omega &= \Omega^\alpha{}_\beta\, e_\alpha\otimes\theta^\beta,\\ &= (e_\alpha\otimes\theta^\beta)\wedge\Omega^\alpha{}_\beta. \end{aligned}$$

Then,

$$\begin{aligned} D\Omega &= D(e_\alpha\otimes\theta^\beta)\wedge\Omega^\alpha{}_\beta + (-1)^2(e_\alpha\otimes\theta^\beta)\wedge d\Omega^\alpha{}_\beta,\\ &= (De_\alpha\otimes\theta^\beta + e_\alpha\otimes D\theta^\beta)\wedge\Omega^\alpha{}_\beta + (e_\alpha\otimes\theta^\beta)\wedge d\Omega^\alpha{}_\beta,\\ &= (e_\gamma\otimes\omega^\gamma{}_\theta\otimes\theta^\beta + e_\alpha\otimes\omega^\beta{}_\gamma\otimes\theta^\gamma)\wedge)\Omega^\alpha{}_\beta + (e_\alpha\otimes\theta^\beta)\wedge d\Omega^\alpha{}_\beta,\\ &= (e_\alpha\otimes\theta^\beta)\wedge(d\Omega^\alpha{}_\beta + \omega^\alpha{}_\gamma\wedge\Omega^\gamma{}_\beta - \omega^\alpha{}_\gamma\wedge\Omega^\gamma{}_\beta). \end{aligned}$$

In the last step we had to relabel a couple of indices so that we could factor out $(e_\alpha\otimes\theta^\beta)$. The pattern should be clear. We get an exterior derivative for the forms, an $\omega\wedge\Omega$ term for the contravariant index and an $\Omega\wedge\omega$ term with the appropriate sign, for the covariant index. Here the computation gives

$$\begin{aligned} D\Omega^\alpha{}_\beta &= d\Omega^\alpha{}_\beta + \omega^\alpha{}_\gamma\wedge\Omega^\gamma{}_\beta - \omega^\alpha{}_\gamma\wedge\Omega^\gamma{}_\beta, \quad \text{or}\\ D\Omega &= d\omega + \omega\wedge\Omega - \Omega\wedge\omega. \end{aligned} \tag{6.58}$$

This means that we can write the equations of structure as

$$\begin{aligned} \Theta &= D\theta,\\ \Omega &= d\omega + \omega\wedge\omega, \end{aligned} \tag{6.59}$$

and the Bianchi's identities as

$$\begin{aligned} D\Theta &= \Omega\wedge\theta,\\ D\Omega &= 0 \end{aligned} \tag{6.60}$$

With apologies for the redundancy, we reproduce the change of basis formula 3.49. Let $e' = eB$ be an orthogonal change of basis. Then

$$\begin{aligned} De' &= e\otimes dB + DeB,\\ &= e\otimes dB + (e\otimes\omega)B,\\ &= e'\otimes(B^{-1}dB + B^{-1}\omega B),\\ &= e'\otimes\omega', \end{aligned}$$

where,

$$\omega' = B^{-1}dB + B^{-1}\omega B. \tag{6.61}$$

Multiply the last equation by B and take the exterior derivative d. We get.

$$\begin{aligned}
B\omega' &= dB + \omega B,\\
Bd\omega' + dB \wedge \omega' &= d\omega B - \omega \wedge dB,\\
Bd\omega' + (B\omega' - \omega B) \wedge \omega' &= d\omega B - \omega \wedge (\omega' B - \omega B),\\
B(d\omega' + \omega' \wedge \omega') &= (d\omega + \omega \wedge \omega)B,
\end{aligned}$$

Setting $\Omega = d\omega + \omega \wedge \omega$, and $\Omega' = d\omega' + \omega' \wedge \omega'$, the last equation reads,

$$\Omega' = B^{-1}\Omega B. \tag{6.62}$$

As pointed out after equation 3.49, the curvature is a tensorial form of adjoint type. The transformation law above for the connection has an extra term, so it is not tensorial. It is easy to obtain the classical transformation law for the Christoffel symbols from equation 6.61. Let $\{x^\alpha\}$ be coordinates in a patch (ϕ_α, U_α), and $\{y^\beta\}$ be coordinates on a overlapping patch (ϕ_β, U_β). The transition functions $\phi_{\alpha\beta}$ are given by the Jacobian of the change of coordinates,

$$\begin{aligned}
\frac{\partial}{\partial y^\beta} &= \frac{\partial x^\alpha}{\partial y^\beta}\frac{\partial}{\partial x^\alpha},\\
\phi_{\alpha\beta} &= \frac{\partial x^\alpha}{\partial y^\beta}.
\end{aligned}$$

Inserting the connection components $\omega'^\alpha{}_\beta = \Gamma'^\alpha_{\beta\gamma} dy^\gamma$, into the change of basis formula 6.61, with $B = \phi_{\alpha\beta}$, we get[1],

$$\begin{aligned}
\omega'^\alpha{}_\beta &= (B^{-1})^\alpha{}_\kappa dB^\kappa{}_\beta + (B^{-1})^\alpha{}_\kappa\, \omega^\kappa{}_\lambda B^\lambda{}_\beta,\\
&= \frac{\partial y^\alpha}{\partial x^\kappa} d\left(\frac{\partial x^\kappa}{\partial y^\beta}\right) + \frac{\partial y^\alpha}{\partial x^\kappa}\,\omega^\kappa{}_\lambda\,\frac{\partial x^\lambda}{\partial y^\beta},\\
\Gamma'^\alpha_{\beta\gamma} dy^\gamma &= \frac{\partial y^\alpha}{\partial x^\kappa}\frac{\partial^2 x^\kappa}{\partial y^\sigma \partial y^\beta} dy^\sigma + \frac{\partial y^\alpha}{\partial x^\kappa}\,\Gamma^\kappa_{\lambda\sigma} dx^\sigma\,\frac{\partial x^\lambda}{\partial y^\beta},\\
\Gamma'^\alpha_{\beta\gamma} &= \frac{\partial y^\alpha}{\partial x^\kappa}\frac{\partial^2 x^\kappa}{\partial y^\gamma \partial y^\beta} + \frac{\partial y^\alpha}{\partial x^\kappa}\,\Gamma^\kappa_{\lambda\sigma}\frac{\partial x^\sigma}{\partial y^\gamma}\,\frac{\partial x^\lambda}{\partial y^\beta}.
\end{aligned}$$

Thus, we retrieve the classical transformation law for Christoffel symbols that one finds in texts on general relativity.

$$\Gamma'^\alpha_{\beta\gamma} = \Gamma^\kappa_{\lambda\sigma}\frac{\partial y^\alpha}{\partial x^\kappa}\,\frac{\partial x^\sigma}{\partial y^\gamma}\,\frac{\partial x^\lambda}{\partial y^\beta} + \frac{\partial y^\alpha}{\partial x^\kappa}\frac{\partial^2 x^\kappa}{\partial y^\gamma \partial y^\beta}. \tag{6.63}$$

[1] We use this notation reluctantly, to be consistent with most literature. The notation results in violation of the index notation. We really should be writing $\phi^\alpha{}_\beta$, since in this case, the transition functions are matrix-valued.

6.4.4 Parallelism

When first introduced to vectors in elementary calculus and physics courses, vectors are often described as entities characterized by a direction and a length. This primitive notion, that two such entities in $\mathbf{R}^n$ with the same direction and length represent the same vector, regardless of location, is not erroneous in the sense that parallel translation of a vector in $\mathbf{R}^n$ does not change the attributes of a vector as described. In elementary linear algebra, vectors are described as n-tuples in $\mathbf{R}^n$ equipped with the operations of addition and multiplication by scalar, and subject to eight vector space properties. Again, those vectors can be represented by arrows which can be located anywhere in $\mathbf{R}^n$ as long as they have the same components. This is another indication that parallel transport of a vector in $\mathbf{R}^n$ is trivial, a manifestation of the fact the $\mathbf{R}^n$ is a flat space. However, in a space that is not flat, such as s sphere, parallel transport of vectors is intimately connected with the curvature of the space. To elucidate this connection, we first describe parallel transport for a surface in $\mathbf{R}^3$.

6.4.2 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 1.25. A vector field Y is called *parallel* along α if

$$\nabla_V Y = 0,$$

as illustrated in figure 6.2. The notation

$$\frac{DY}{dt} = \nabla_V Y$$

is also common in the literature. 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 to the speed by the equation $v = ds/dt$.

6.4.3 Definition A curve $\alpha(t)$ with velocity vector $V = \alpha'(t)$ is called a *geodesic* or *self-parallel* if $\nabla_V V = 0$.

6.4.4 Theorem A curve $\alpha(t)$ is geodesic iff

a) $v = \|V\|$ is constant along the curve and,

b) either $\nabla_T T = 0$, or $\kappa_g = 0$.

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 $< T, T >= 1$, so $< \nabla_T T, T >= 0$ 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 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 α is geodesic, we have $\|\alpha'\|^2 =< \alpha', \alpha' >=$constant. Differentiation gives $< \alpha', \alpha'' >= 0$, so that the acceleration α'' is orthogonal to α'. Comparing with equation 4.34 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, 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.

6.4.5 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

$$
\begin{aligned}
Y &= y^\beta \frac{\partial}{\partial u^\beta}, \\
V &= \frac{du^\alpha}{dt} \frac{\partial}{\partial u^\alpha}. \quad \text{then,} \\
\nabla_T V &= \nabla_{\dot{u^\alpha} e_\alpha} (y^\beta e_\beta), \\
&= \dot{u^\alpha} \nabla_{e^\alpha} (y^\beta e_\beta), \\
&= \frac{du^\alpha}{dt} \frac{\partial y^\beta}{\partial u^\alpha} + \dot{u^\alpha} y^\beta \Gamma^\gamma{}_{\alpha\beta} e_\gamma, \\
&= [\frac{dy^\gamma}{dt} + y^\beta \frac{du^\alpha}{dt} \Gamma^\gamma{}_{\alpha\beta}] e_\gamma.
\end{aligned}
$$

So, Y is parallel along the curve iff,

$$
\frac{dy^\gamma}{dt} + y^\beta \frac{du^\alpha}{dt} \Gamma^\gamma{}_{\alpha\beta} = 0. \tag{6.64}
$$

The existence and uniqueness of the coefficients y^β that define Y are guaranteed by the theorem on existence and uniqueness of differential equations with appropriate initial conditions.

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

$$\ddot{u}^\sigma + \Gamma^\sigma_{\alpha\beta} \dot{u}^\alpha \dot{u}^\beta = 0. \tag{6.65}$$

The existence and uniqueness theorem for solutions of differential equations leads to the following theorem

6.4.6 Theorem Let p be a point in M and V a vector T_pM. 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.

For a general vector bundles E over a manifold M, a section $s \in \Gamma(E)$ of a vector bundle is called a *parallel section* if

$$\nabla s = 0. \tag{6.66}$$

We discuss the length minimizing properties geodesics in section 6.6 and provide a number of examples for surfaces in $\mathbf{R}^3$ and for Lorentzian manifolds. Since geodesic curves have zero acceleration, in Euclidean space they are straight lines. In Einstein's theory of relativity, gravitation is a fictitious force caused by the curvature of space time, so geodesics represent the trajectory of free particles.

6.5 Lorentzian Manifolds

The formalism above refers to Riemannian manifolds, for which the metric is positive definite, but it applies just as well to pseudo-Riemannian manifolds. A 4-dimensional manifold $\{M, g\}$ 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 2.2.3. Some authors use signature $(- + ++)$.

For the purposes of general relativity, we introduce the symmetric tensor *Ricci tensor* $R_{\beta\delta}$ by the contraction

$$R_{\beta\delta} = R^\alpha{}_{\beta\alpha\delta}, \tag{6.67}$$

and the *scalar curvature* R by

$$R = R^{\alpha}{}_{\beta}. \tag{6.68}$$

The traceless part of the Ricci tensor

$$G_{\alpha\beta} = R_{\alpha\beta} - \tfrac{1}{2} R g_{\alpha\beta}, \tag{6.69}$$

is called the *Einstein tensor.* The Einstein field equations (without a cosmological constant) are

$$G_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta}, \tag{6.70}$$

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 matter tells the space how to curve. Einstein equations with cosmological constant Λ are,

$$R_{\alpha\beta} - \tfrac{1}{2} R g_{\alpha\beta} + \Lambda g_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta} \tag{6.71}$$

Fig. 6.3: Gravity

A space time which satisfies

$$R_{\alpha\beta} = 0 \tag{6.72}$$

is called *Ricci-flat.* A space which the Ricci tensor is proportional to the metric,

$$R_{\alpha\beta} = k g_{\alpha\beta} \tag{6.73}$$

is called an *Einstein manifold*

6.5.1 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 most general, spherically symmetric, static solution of the Einstein vacuum equations is the Schwarzschild metric [2]

$$ds^2 = \left[1 - \tfrac{2GM}{r}\right] dt^2 - \frac{1}{\left[1 - \tfrac{2GM}{r}\right]} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta \, d\phi^2 \tag{6.74}$$

It is convenient to set $m = GM$ and introduce the retarded coordinate transformation

$$t = u + r + 2m \ln(\tfrac{r}{2m} - 1),$$

so that,

$$dt = du + \frac{1}{\left[1 - \tfrac{2m}{r}\right]} dr.$$

Substitution for dt above gives the metric in outgoing *Eddington-Finkelstein* coordinates,

$$ds^2 = 2drdu + [1 - \tfrac{2m}{r}]\, du^2 - r^2 d\theta^2 - r^2 \sin^2\theta \, d\phi^2. \tag{6.75}$$

In these coordinates it is evident that the event horizon $r = 2m$ is not a real singularity. The Vaidya metric is the generalization

$$ds^2 = 2drdu + [1 - \tfrac{2m(u)}{r}]\, du^2 - r^2 d\theta^2 - r^2 \sin^2\theta \, d\phi^2, \tag{6.76}$$

where $m(u)$ is now 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. In all our previous curvature computations by differential forms, the metric has been diagonal; this is an instructive example of one with a non-diagonal metric. The first step in the curvature 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

$$\begin{aligned}
\theta^0 &= du,\\
\theta^1 &= dr + \tfrac{1}{2}[1 - \tfrac{2m(u)}{r}]\, du,\\
\theta^2 &= r\, d\theta,\\
\theta^3 &= r \sin\theta \, d\phi.
\end{aligned} \tag{6.77}$$

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 the coframe, the metric has components:

$$g_{\alpha\beta} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{6.78}$$

[2]The Schwarzschild radius is $r = \frac{2GM}{c^2}$, but here we follow the common convention of setting $c = 1$.

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

$$\begin{aligned}
d\theta^0 &= 0,\\
d\theta^1 &= -d[\tfrac{m}{r}du] = \frac{m}{r^2}\,dr\wedge du = \frac{m}{r^2}\,\theta^1\wedge\theta^0,\\
d\theta^2 &= dr\wedge d\theta = \frac{1}{r}\theta^1\wedge\theta^2 - \frac{1}{2r}[1-\tfrac{2m}{r}]\,\theta^0\wedge\theta^2,\\
d\theta^3 &= \sin\theta\,dr\wedge d\phi + r\cos\theta\,d\theta\wedge d\phi,\\
&= \frac{1}{r}\theta^1\wedge\theta^3 - \frac{1}{2}[1-\tfrac{2m}{r}]\,\theta^0\wedge\theta^3 + \frac{1}{r}\cot\theta\,\theta^2\wedge\theta^3. \qquad (6.79)
\end{aligned}$$

For convenience, we write below the first equation of structure [6.24] in complete detail.

$$\begin{aligned}
d\theta^0 &= {\omega^0}_0\wedge\theta^0 + {\omega^0}_1\wedge\theta^1 + {\omega^0}_2\wedge\theta^2 + {\omega^0}_3\wedge\theta^3,\\
d\theta^1 &= {\omega^1}_0\wedge\theta^0 + {\omega^1}_1\wedge\theta^1 + {\omega^1}_2\wedge\theta^2 + {\omega^1}_3\wedge\theta^3,\\
d\theta^2 &= {\omega^2}_0\wedge\theta^0 + {\omega^2}_1\wedge\theta^1 + {\omega^2}_2\wedge\theta^2 + {\omega^2}_3\wedge\theta^3,\\
d\theta^3 &= {\omega^3}_0\wedge\theta^0 + {\omega^3}_1\wedge\theta^1 + {\omega^3}_2\wedge\theta^2 + {\omega^3}_3\wedge\theta^3. \qquad (6.80)
\end{aligned}$$

Since 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:

$$\begin{aligned}
&{\omega^1}_0 = 0, && {\omega^1}_1 = \frac{m}{r^2}\,\theta^0, && {\omega^1}_2 = A\,\theta^2, && {\omega^1}_3 = B\,\theta^3,\\
&{\omega^2}_0 = -\frac{1}{2}[1-\frac{2m}{r}]\,\theta^2, && {\omega^2}_1 = \frac{1}{r}\,\theta^2, && {\omega^2}_2 = 0, && {\omega^2}_3 = C\,\theta^3,\\
&{\omega^3}_0 = -\frac{1}{2}[1-\frac{2m}{r}]\,\theta^3, && {\omega^3}_1 = \frac{1}{r}\,\theta^3, && {\omega^3}_2 = \frac{1}{r}\cot\theta\,\theta^3, && {\omega^3}_3 = 0.
\end{aligned}$$

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 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

$$\omega^1{}_2 = g^{10}\omega_{02} = -\omega_{20} = \omega^2{}_0,$$
$$\omega^1{}_3 = g^{10}\omega_{03} = -\omega_{30} = \omega^3{}_0,$$
$$\omega^2{}_3 = g^{22}\omega_{23} = \omega_{32} = -\omega^3{}_2.$$

Comparing the structure equations 6.80 with the expressions for the connection coefficients above, we find that

$$A = -\frac{1}{2}[1 - \frac{2m}{r}], \quad B = -\frac{1}{2}[1 - \frac{2m}{r}], \quad C = -\frac{1}{r}\cot\theta. \tag{6.81}$$

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,

$$\begin{aligned} \omega^0{}_0 &= -\frac{m}{r^2}\theta^0, \\ \omega^0{}_2 &= -\frac{1}{r}\theta^2, \\ \omega^0{}_3 &= \frac{1}{r}\theta^3. \end{aligned}$$

It is easy to verify 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 [6.25]. 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 straightforward and we just present the results.

$$\Omega^1{}_2 = -\frac{1}{r^2}\frac{dm}{du}\theta^2 \wedge \theta^0 - \frac{m}{r^3}\theta^1 \wedge \theta^2,$$

$$\Omega^1{}_3 = -\frac{1}{r^2}\frac{dm}{du}\theta^3 \wedge \theta^0 - \frac{m}{r^3}\theta^1 \wedge \theta^3,$$
$$\Omega^2{}_1 = \frac{m}{r^3}\theta^2 \wedge \theta^0,$$
$$\Omega^3{}_1 = \frac{m}{r^3}\theta^3 \wedge \theta^0,$$
$$\Omega^2{}_3 = \frac{2m}{r^3}\theta^2 \wedge \theta^3.$$

By antisymmetry, these are the only independent components. We can also read the components of the full Riemann curvature tensor from the definition

$$\Omega^\alpha{}_\beta = \frac{1}{2}R^\alpha{}_{\beta\gamma\delta}\theta^\gamma \wedge \theta^\delta. \tag{6.82}$$

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, we find that

$$R_{00} = 2\frac{1}{r^2}\frac{dm}{du} \tag{6.83}$$

while all the other components of the Ricci tensor vanish. As stated earlier, if m is constant, we get the Ricci flat Schwarzschild metric.

6.6 Geodesics

Geodesics were introduced in the section on parallelism. The equation of geodesics on a manifold given by equation 6.65 involves the Christoffel symbols. Whereas it is possible to compute all the Christoffel symbols starting with the metric as in equation 4.76, 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

$$\delta \int L(u^\alpha, \dot{u}^\alpha, s)\, ds = 0, \tag{6.84}$$

to minimize the arc length. Then we can pick out the non-vanishing Christoffel symbols from the geodesic equation. Following the standard methods of Lagrangian mechanics, we let u^α and $\dot{u}^\alpha$ be treated as independent (canonical) coordinates and choose the Lagrangian in this case to be

$$L = g_{\alpha\beta}\dot{u}^\alpha \dot{u}^\beta. \tag{6.85}$$

The choice will actually result in minimizing the square of the arc length, 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.

$$\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{u}^\gamma}\right) - \frac{\partial L}{\partial u^\gamma} = 0. \tag{6.86}$$

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 4.76, we get

$$\ddot{u}^\sigma + \Gamma^\sigma_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta = 0$$

which are exactly the equations of geodesics 6.65.

6.6.1 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.

6.6.2 Example Geodesics in orthogonal coordinates.
In a parametrization of a surface in which the coordinate lines are orthogonal, $F = 0$. Then first fundamental form is,

$$ds^2 = E\,du^2 + G\,dv^2,$$

and we have the Lagrangian,

$$L = E\dot{u}^2 + G\dot{v}^2.$$

The Euler-Lagrange equations for the variable u are:

$$\begin{aligned}
\frac{d}{ds}(2E\dot{u}) - E_u\dot{u}^2 - G_u\dot{v}^2 &= 0,\\
2E\ddot{u} + (2E_u\dot{u} + 2E_v\dot{v})\dot{u} - E_u\dot{u}^2 - G_u\dot{v}^2 &= 0,\\
2E\ddot{u} + E_u\dot{u}^2 + 2E_v\dot{u}\dot{v} - G_u\dot{v}^2 &= 0.
\end{aligned}$$

Similarly for the variable v,

$$\begin{aligned}
\frac{d}{ds}(2G\dot{u}) - E_v\dot{u}^2 - G_v\dot{v}^2 &= 0,\\
2G\ddot{v} + (2G_u\dot{u} + 2G_v\dot{v})\dot{v} - E_v\dot{u}^2 - G_v\dot{v}^2 &= 0,\\
2G\ddot{v} - E_v\dot{u}^2 + 2G_u\dot{u}\dot{v} + G_v\dot{v}^2 &= 0.
\end{aligned}$$

So, the equations of geodesics can be written neatly as,

$$\begin{aligned}
\ddot{u} + \frac{1}{2E}[E_u\dot{u}^2 + 2E_v\dot{u}\dot{v} - G_u\dot{v}^2] &= 0,\\
\ddot{v} + \frac{1}{2G}[G_v\dot{v}^2 + 2G_u\dot{u}\dot{v} - E_v\dot{u}^2] &= 0. \qquad (6.87)
\end{aligned}$$

6.6.3 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

$$ds^2 = (1 + f'^2)\,dr^2 + r^2\,d\phi^2, \qquad (6.88)$$

Of course, we could use the expressions for the equations of geodesics we just derived above, but since the coefficients are functions of r only, it is just a easy to start from the Lagrangian,

$$L = (1 + f'^2)\,\dot{r}^2 + r^2\dot{\phi}^2.$$

Since there is no dependance on ϕ, the Euler-Lagrange equation on ϕ gives rise to a conserved quantity.

$$\begin{aligned}
\frac{d}{ds}(2r^2\dot{\phi}) &= 0,\\
r^2\dot{\phi} &= c \qquad (6.89)
\end{aligned}$$

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 on 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

$$\begin{aligned}\cos\sigma &= \frac{< \alpha', \mathbf{x}_\phi >}{\|\alpha'\| \cdot \|\mathbf{x}_\phi\|} = \frac{G \frac{d\phi}{ds}}{k\sqrt{G}}, \\ &= \frac{1}{k}\sqrt{G}\frac{d\phi}{ds} = \frac{1}{k} r\dot{\phi}.\end{aligned}$$

We conclude from 6.89, that for a surface of revolution, the geodesics make an angle σ with meridians that satisfies the equation

$$r\cos\sigma = \text{constant}. \tag{6.90}$$

This result is called *Clairaut's relation.* Writing equation 6.89 in terms of differentials, and reusing the metric as we did in the computation of the geodesics for a sphere, we get

$$\begin{aligned} r^2\, d\phi &= c\, ds, \\ r^4\, d\phi^2 &= c^2\, ds^2, \\ &= c^2[(1+f'^2)\, dr^2 + r^2\, d\phi^2], \\ (r^4 - c^2 r^2)\, d\phi^2 &= c^2[(1+f'^2)\, dr^2, \\ r\sqrt{r^2-c^2}\, d\phi &= c\sqrt{1+f'^2}\, dr, \end{aligned}$$

so

$$\phi = \pm c \int \frac{\sqrt{1+f'^2}}{r\sqrt{r^2-c^2}}\, dr. \tag{6.91}$$

If $c = 0$, then the first equation above gives $\phi =$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 6.91 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

$$\phi = \pm \csc\alpha \sec^{-1}(r/c) \tag{6.92}$$

As shown in figure 6.4, 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.

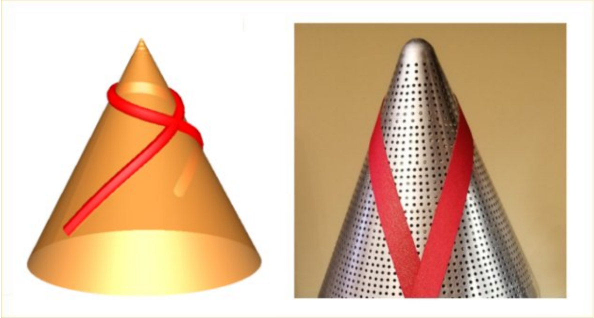


Fig. 6.4: Geodesics on a Cone.

6.7 Geodesics in GR

6.7.1 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 traversable 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

$$ds^2 = -c^2dt^2 + dl^2 + (b_0^2 + l^2)\,(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{6.93}$$

where b_0 is a constant. The obvious choice for a coframe is

$$\begin{aligned} \theta^0 &= c\,dt, & \theta^2 &= \sqrt{b_o^2 + l^2}\, d\theta, \\ \theta^1 &= dl, & \theta^3 &= \sqrt{b_o^2 + l^2}\, \sin\theta\, d\phi. \end{aligned}$$

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. That is, 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. A straightforward 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 6.36, 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

$$ds_2{}^2 = dl^2 + (b_o^2 + l^2)\,d\phi^2. \tag{6.94}$$

Let $r^2 = b_o^2 + l^2$. Then $dl^2 = (r^2/l^2)\,dr^2$ and the metric becomes

$$\begin{aligned}
ds_2{}^2 &= \frac{r^2}{r^2 - b_o^2}\,dr^2 + r^2\,d\phi^2, &(6.95)\\
&= \frac{1}{1 - \frac{b_o^2}{r^2}}\,dr^2 + r^2\,d\phi^2. &(6.96)
\end{aligned}$$

Comparing to 4.26 we recognize this to be a catenoid of revolution, so the equations of geodesics are given by 6.91 with $f(r) = b_0 \cosh^{-1}(r/b_0)$. Substituting this value of f into the geodesic equation, we get

$$\phi = \pm c \int \frac{1}{\sqrt{r^2 - b_o^2}\sqrt{r^2 - c^2}} dr. \tag{6.97}$$

There are three cases. If $c = b_o$, the integral gives immediately

$$\phi = \pm(c/b_0)\tanh^{-1}(r/b_0).$$

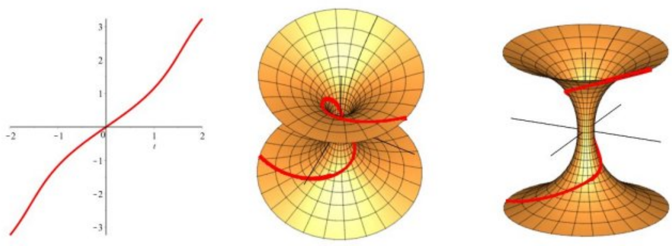


Fig. 6.5: Geodesics on Catenoid.

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 substitution leads to the integral

$$\begin{aligned}
\phi &= \pm c \int_0^s \frac{1}{r\cos\beta\sqrt{\frac{c^2}{\sin^2\beta} - b_o^2}} \frac{(-r\cos\beta)}{\sin\beta}\, d\beta, \\
&= \pm c \int_0^s \frac{1}{\sqrt{c^2 - b_o^2\sin^2\beta}}\, d\beta, \\
&= \pm \int_0^s \frac{1}{\sqrt{1 - k^2\sin^2\beta}}\, d\beta, \quad (k = b_o/c) && (6.98) \\
&= F(s, k), && (6.99)
\end{aligned}$$

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 6.5. 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 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.

6.7.2 **Example** Schwarzschild Metric

In this section we look at the geodesic equations in a Schwarzschild gravitational field, with particular emphasis on the bounded orbits. We write the metric in the form

$$ds^2 = -h(r)\,dt^2 + \frac{1}{h(r)}\,dr^2 + r^2(d\theta^2 + \sin\theta\,d\phi^2), \tag{6.100}$$

where

$$h(r) = 1 - \frac{2GM}{r}. \tag{6.101}$$

Thus, the Lagrangian is

$$\mathscr{L} = -h\,\dot{t}^2 + \frac{1}{h}\,\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin\theta\dot{\phi}^2. \tag{6.102}$$

The Euler-Lagrange equations for g_{00}, g_{22} and g_{33} yield

$$\begin{aligned} \frac{d}{ds}\left[-2h\frac{dt}{ds}\right] &= 0,\\ \frac{d}{ds}\left[r^2\frac{d\theta}{ds}\right] - r^2\sin\theta\cos\theta\left[\frac{d\phi}{ds}\right]^2 &= 0,\\ \frac{d}{ds}\left[2r^2\frac{d\phi}{ds}\right] &= 0 \end{aligned}$$

If in the equation for g_{22}, one chooses initial conditions $\theta(0) = \pi/2$, $\dot{\theta}(0) = 0$, we get $\theta(s) = \pi/2$ along the geodesic. We infer from rotation invariance that the motion takes place on a plane. Hereafter, we assume we have taken these initial conditions. From the other two equations we obtain

$$\begin{aligned} h\frac{dt}{ds} &= E,\\ r^2\frac{d\phi}{ds} &= L. \end{aligned}$$

for some constants E and L. We recognize the conserved quantities as the "energy" and the angular momentum. Along the geodesic of a massive particle, with unit time-like tangent vector, we have

$$-1 = g_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} \tag{6.103}$$

The equations of motion then reduce to

$$\begin{aligned} -1 &= -h\left[\frac{dt}{ds}\right]^2 + \frac{1}{h}\left[\frac{dr}{ds}\right]^2 + r^2\left[\frac{d\phi}{ds}\right]^2,\\ -1 &= -\frac{E^2}{h} + \frac{1}{h}\left[\frac{dr}{ds}\right]^2 + \frac{L^2}{r^2},\\ E^2 &= \left[\frac{dr}{ds}\right]^2 + h\left[1 + \frac{L^2}{r^2}\right]. \end{aligned}$$

Hence, we obtain the neat equation,

$$E^2 = \left[\frac{dr}{ds}\right]^2 + V(r), \tag{6.104}$$

where $V(r)$ represents the effective potential.

$$\begin{aligned} V(r) &= \left[1 - \frac{2GM}{r}\right]\left[1 + \frac{L^2}{r^2}\right], \\ &= 1 - \frac{2GM}{r} + \frac{L^2}{r^2} - \frac{2MGL^2}{r^3}. \end{aligned} \tag{6.105}$$

If we let $\hat{V} = V/2$ in this expression we recognize the classical $1/r$ potential, and the $1/r^2$ term corresponding to the Coriolis contribution associated with the angular momentum. The $1/r^3$ term is a new term arising from general relativity. Clearly we must have $E^2 < V(r)$. There are multiple cases depending

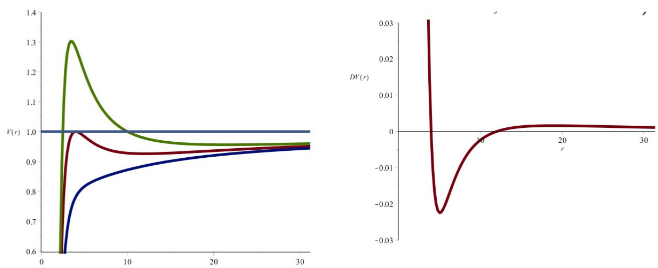


Fig. 6.6: Effective Potential for $L = 3, 4, 5$

on the values of E and L and the nature of the equilibrium points. Here we are primarily concerned with bounded orbits, so we seek conditions for the particle to be in a potential well. This presents us with a nice calculus problem. We compute $V'(r)$ and set equal to zero to find the critical points

$$V'(r) = \frac{2}{r^4}(GMr^2 - L^2 r + 3GML^2) = 0.$$

The discriminant of the quadratic is

$$D = L^2 - 12G^2M^2.$$

If $D < 0$ there are no critical points. In this case, $V(r)$ is a monotonically increasing function on the interval $(2MG, \infty)$, as shown in the bottom left graph in figure 6.6. The maple plots in this figure are in units with $GM = 1$. In the case $D < 0$, all trajectories either fall toward the event horizon or escape to infinity.

If $D > 0$, there are two critical points

$$r_1 = \frac{L^2 - L\sqrt{L^2 - 12G^2M^2}}{2GM},$$
$$r_2 = \frac{L^2 + L\sqrt{L^2 - 12G^2M^2}}{2GM}.$$

The critical point r_1 is a local maximum associated with an unstable circular orbit. The critical point $r_2 > r_1$ gives a stable circular orbit. Using the standard calculus trick of multiplying by the conjugate of the radical in the first term, we see that

$$r_1 \to 3GM,$$
$$r_2 \to \frac{L^2}{GM},$$

as $L \to \infty$. For any L, the properties of the roots of the quadratic imply that $r_1 r_2 = 3L^2$. As shown in the graph 6.6, as L gets larger, the inner radius approaches $3GM$ and the height of the bump increases, whereas the outer radius recedes to infinity. As the value of D approaches 0, the two orbits coalesce at $L^2 = 12G^2M^2$, which corresponds to $r = 6GM$, so this is the smallest value of r at which a stable circular orbit can exist. Since $V(r) \to 1$ as $r \to \infty$, to get bounded orbits we want a potential well with $V(r_1) < 1$. We can easily verify that when $L = 4GM$ the local maximum occurs at $r_1 = 4GM$, which results in a value of $V(r_1) = 1$. This case is the one depicted in the middle graph in figure 6.6, with the graph of $V'(r)$ on the right showing the two critical points at $r_1 = 4GM, \;\; r_2 = 12GM$. Hence the condition to get a bounded orbit is

$$2\sqrt{3}\,GM < L < 4GM,$$
$$E^2 < V(r_1), \;\; r > r_1,$$

so that the energy results in the particle trapped in the potential well to the right of r_1. This is the case that applies to the modification of the Kepler orbits of planets. If we rewrite

$$\frac{dr}{ds} = \frac{dr}{d\phi}\frac{d\phi}{ds} = \frac{L}{r^2}\frac{dr}{d\phi}$$

and substitute into equation 6.104, we get

$$\frac{L^2}{r^4}\left[\frac{dr}{d\phi}\right]^2 = E^2 - \left[1 + \frac{L^2}{r^2}\right]\left[1 - \frac{2GM}{r}\right].$$

If now we change variables to $u = 1/r$, we obtain

$$\frac{du}{d\phi} = -\frac{1}{r^2}\frac{dr}{d\phi} = -u^2\frac{dr}{d\phi},$$

and the orbit equation becomes

$$\left[\frac{du}{d\phi}\right]^2 = \frac{1}{L^2}[E^2 - (1 + L^2u^2)(1 - 2GMu)],$$
$$\phi = \int \frac{L du}{\sqrt{E^2 - (1 + L^2u^2)(1 - 2GML^2u)}} + \phi_0.$$

The solution of the orbit equation is therefore reduced to an elliptic integral. If we expand the denominator

$$\phi = \int \frac{Ldu}{\sqrt{(E^2-1)+2GMu-L^2u^2+2GML^2u^3}} + \phi_0,$$

and neglect the cubic term, we can complete the squares of the remaining quadratic. The integral becomes one of standard inverse cosine type; hence, the solution gives the equation of an ellipse in polar coordinates

$$u = \frac{1}{r} = C(1 + e\cos(\phi - \phi_0)),$$

for appropriate constants C, shift ϕ_0 and eccentricity e. The solution is automatically expressed in terms of the energy and the angular momentum of the system. More careful analysis of the integral shows that the inclusion of the cubic term perturbs the orbit by a precession of the ellipse. While this approach is slicker, we prefer to use the more elementary procedure of differential equations. Differentiating with respect to ϕ the equation

$$L^2\left[\frac{du}{d\phi}\right]^2 = (E^2-1)+2GMu-L^2u^2+2GML^2u^3,$$

and cancelling out the common chain rule factor $du/d\phi$, we get

$$\frac{d^2u}{d\phi^2} = \frac{GM}{L^2} - u + 3GMu^2$$

Introducing a dimensionless parameter

$$\epsilon = \frac{3G^2M^2}{L^2},$$

we can rewrite the equation of motion as

$$\frac{d^2u}{d\phi^2} + u = \frac{GM}{L^2} + \frac{L^2}{GM}u^2\epsilon. \tag{6.106}$$

The linear part of the equation corresponds precisely to Newtonian motion, and ϵ is small, so we can treat the quadratic term as a perturbation

$$u = u_0 + u_1\epsilon + u_2\epsilon^2 + \ldots.$$

Substituting u into equation 6.106, the first approximation is the linear approximation given by

$$u_0'' + u = \frac{GM}{L^2}.$$

The homogenous solution is of the form $u = A\cos(\phi - \phi_0)$, where A and ϕ_0 are the arbitrary constants, and the particular solution is a constant. So the general solution is

$$\begin{aligned} u_0 &= \frac{GM}{L^2} + A\cos(\phi-\phi_0), \\ &= \frac{GM}{L^2}[1 + e\cos(\phi-\phi_0)], \quad e = \frac{AL^2}{GM}. \end{aligned}$$

Without loss of generality, we can align the axes and set $\phi_0 = 0$. In the Newtonian orbit, we would write $u_0 = 1/r$, thus getting the equation of a polar conic.

$$u_0 = \frac{GM}{L^2}(1 + e\cos\phi) \tag{6.107}$$

In the case of the planets, the eccentricity $e < 1$, so the conics are ellipses. Having found u_0 we reinsert u into the differential equation 6.106 and keeping only the terms of order ϵ. We get

$$\begin{aligned}(u_0 + u_1\epsilon)'' + (u_0 + u_1\epsilon) &= \frac{GM}{L^2} + \frac{L^2}{GM}\epsilon(u_0 + u_1\epsilon)^2,\\ (u_0'' + u_0 - \frac{GM}{L^2}) + (u_1'' + u_1)\epsilon &= \frac{L^2}{GM}u_0^2\epsilon.\end{aligned}$$

Thus, the result is a new differential equation for u_1,

$$\begin{aligned}u_1'' + u_1' &= \frac{L^2}{GM}u_0^2,\\ &= \frac{L^2}{GM}[(1 + \tfrac{1}{2}e^2) + 2e\cos\phi + \tfrac{1}{2}e^2\cos 2\phi].\end{aligned}$$

The equation is again a linear inhomogeneous equation with constant coefficients, so it is easily solved by elementary methods. We do have to be a bit careful since we have a resonant term on the right hand side. The solution is

$$u_1 = \frac{L^2}{GM}[(1 + \tfrac{1}{2}e^2) + 2e\phi\cos\phi - \tfrac{1}{6}e^2\cos 2\phi].$$

The resonant term $\phi\cos\phi$ makes the solution non-periodic, so this is the term responsible for the precession of the elliptical orbits. The precession is obtained by looking at the perihelion, that is, the point in the elliptical orbit at which the planet is closest to the sun. This happens when

$$\begin{aligned}\frac{du}{d\phi} \approx \frac{d}{d\phi}(u_0 + u_1) &= 0,\\ -\sin\phi + (\sin\phi + e\phi\cos\phi + \tfrac{1}{3}e\sin\phi) &= 0.\end{aligned}$$

Starting with the solution $\phi = 0$, after on revolution, the perihelion drifts to $\phi = 2\pi + \delta$. By the perturbation assumptions, we assume δ is small, so to lowest order, the perihelion advance in one revolution is

$$\delta = 2\pi\epsilon = \frac{6\pi G^2M^2}{L^2}. \tag{6.108}$$

From equation 6.107 for the Newtonian elliptical orbit, the mean distance a to the sun is given by the average of the aphelion and perihelion distances, that is

$$a = \frac{1}{2}\left[\frac{L^2/GM}{1+e} + \frac{L^2/GM}{1-e}\right] = \frac{L^2}{GM}\frac{1}{1-e^2}.$$

Thus, if we divide by the period T, the rate of perihelion advance can be written in more geometric terms as

$$\delta = \frac{6\pi GM}{a(1-e^2)T}.$$

The famous computation by Einstein of a precession of 43.1" of an arc per century for the perihelion advance of the orbit of Mercury, still stands as one of the major achievements in modern physics.

For null geodesics, equation 6.103 is replaced by

$$0 = g_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds},$$

so the orbit given by the simpler equation

$$E^2 = \left[\frac{dr}{ds}\right]^2 + \frac{L^2}{r^2}h.$$

Performing the change of variables $u = 1/r$, we get

$$\frac{d^2u}{d\phi^2} + u = 3GMu^2.$$

Consider the problem of light rays from a distant star grazing the sun as they approach the earth. Since the space is asymptotically flat, we expect the geodesics to be asymptotically straight. The quantity $3GM$ is of the order of 2km, so it is very small compared to the radius of the sun, so again we can use perturbation methods. We let $\epsilon = 3GM$ and consider solutions of equation

$$u'' + u = \epsilon u^2,$$

of the form

$$u = u_0 + u_1\epsilon.$$

To lowest order the solutions are indeed straight lines

$$\begin{aligned} u_0 &= A\cos\phi + B\sin\phi, \\ 1 &= Ar\cos\phi + Br\sin\phi, \\ 1 &= Ax + By \end{aligned}$$

Without loss of generality, we can align the vertical axis parallel to the incoming light with impact parameter b (distance of closest approach)

$$u_0 = \frac{1}{b}\cos\phi.$$

As above, we reinsert the u into the differential equation and compare the coefficients of terms of order ϵ. We get an equation for u_1,

$$u_1'' + u_1 = \frac{1}{b^2}\cos^2\phi = \frac{1}{2b^2}(1+\cos 2\phi).$$

We solve the differential equation by the method of undetermined coefficients and thus we arrive at the perturbation solution to order ϵ

$$u = \frac{1}{b}\cos\phi + \frac{2\epsilon}{3b^2} - \frac{\epsilon}{3b^2}\cos^2\phi.$$

To find the the asymptotic angle of the outgoing photons, we let $r \to \infty$ or $u \to 0$. Thus we get a quadratic equation for $\cos\phi$.

$$\cos\phi = -\frac{2\epsilon}{3b} = -\frac{2GM}{b}$$

Set $\phi = \frac{\pi}{2} + \delta$. Since δ is small, we have $\sin\delta \approx \delta$, and we see that $\delta = 2GM/b$ is the approximation of the deflection angle of one of the asymptotes. The total deflection is twice that angle

$$2\delta = \frac{4GM}{b}.$$

The computation results in a deflection by the sun of light rays from a distant star of about 1.75". This was corroborated in an experiment lead by Eddington during the total solar eclipse of 1919. The part of the expedition in Brazil was featured in the 2005 movie, The House of Sand. For more details and more careful analysis of the geodesics, see for example, Misner Thorne and Wheeler [21].

6.8 Gauss-Bonnet Theorem

This section is dedicated to the memory of Professor S.-S. Chern. I prelude the section with a short anecdote that I often narrate to my students. In June 1979, an international symposium on differential geometry was held at the Berkeley campus in honor of the retirement of Professor Chern. The invited speakers included an impressive list of the most famous differential geometers at the time, At the end of the symposium, Chern walked on the stage of the packed auditorium to give thanks and to answer some questions. After a few short remarks, a member of the audience asked Chern what he thought was the most important theorem in differential geometry. Without any hesitation he answered, "there is only one theorem in differential geometry, and that is Stokes' theorem." This was followed immediately by a question about the most important theorem in analysis. Chern gave the same answer: "there is only one theorem in analysis, Stokes' theorem. A third person then asked Chern what was the most important theorem in Complex Variables. To the amusement of the crowd, Chern responded, "There is only one theorem in complex variables, and that that is Cauchy's theorem. But if one assumes the derivative of the function is continuous, then this is just Stokes' theorem." Now, of course it is well known that Goursat proved that the hypothesis of continuity of the derivative is automatically satisfied when the function is holomorphic. But the genius of Chern was always his uncanny ability to extract the essential of what makes things work, in the simplest terms.

The Gauss-Bonnet theorem is rooted on the theorem of Gauss (4.72), which combined with Stokes' theorem, provides a beautiful geometrical interpretation of the equation. This is undoubtedly part of what Chern had in mind at the symposium, and also when wrote in his Euclidean Differential Geometry Notes (Berkeley 1975) [4] that the theorem has "profound consequences and is perhaps one of the most important theorems in mathematics."

Let $\beta(s)$ by a unit speed curve on an orientable surface M, and let T be the unit tangent vector. There is Frenet frame formalism for M, but if we think of the surface intrinsically as 2-dimensional manifold, then there is no binormal. However, we can define a "geodesic normal" taking $G = J(T)$, where J is the symplectic form **??**, Then the geodesic curvature is given by the Frenet formula

$$T' = \kappa_g G. \tag{6.109}$$

6.8.1 Proposition Let $\{e_1, e_2\}$ be an orthonormal on M, and let $\beta(s)$ be a unit speed curve as above, with unit tangent T. If ϕ is the angle that T makes with e_1, then

$$\kappa_g = \frac{\partial \phi}{\partial s} - \omega^1{}_2(T). \tag{6.110}$$

Proof Since $\{T, G\}$ and $\{e_1, e_2\}$ are both orthonormal basis of the tangent space, they must be related by a rotation by an angle ϕ, that is

$$\begin{bmatrix} T \\ G \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}, \tag{6.111}$$

that is,

$$\begin{aligned} T &= \quad (\cos\phi)e_1 + (\sin\phi)e_2, \\ G &= -(\sin\phi)e_1 + (\cos\phi)e_2. \end{aligned} \tag{6.112}$$

Since $T = \beta'$, and $\beta'' = \nabla_t T$ we have

$$\begin{aligned} \beta'' &= -(\sin\phi)\frac{\partial\phi}{\partial s}e_1 + \cos\phi\nabla_T e_1 + (\cos\phi)\frac{\partial\phi}{\partial s}e_2 + \sin\phi\nabla_T e_2, \\ &= -(\sin\phi)\frac{\partial\phi}{\partial s}e_1 + (\cos\phi)\omega^2{}_1(T)e_2 + (\cos\phi)\frac{\partial\phi}{\partial s}e_2 + (\sin\phi)\omega^1{}_2(T)e_1, \\ &= [\frac{\partial\phi}{\partial s} - \omega^1{}_2(T)][-(\sin\phi)e_1] + [\frac{\partial\phi}{\partial s} - \omega^1{}_2(T)][(\cos\phi)e_2], \\ &= [\frac{\partial\phi}{\partial s} - \omega^1{}_2(T)][-(\sin\phi)e_1 + (\cos\phi)e_2], \\ &= [\frac{\partial\phi}{\partial s} - \omega^1{}_2(T)]G, \\ &= \kappa_g G. \end{aligned}$$

comparing the last two equations, we get the desired result.

This theorem is related to the notion discussed in figure 6.2 to the effect that in a space with curvature, the parallel transport of a tangent vector around a closed curve, does not necessarily result on the same vector with which one started. The difference in angle $\Delta\phi$ between a vector and the parallel transport of the vector around a closed curve C is called the *holonomy* of the curve. The holonomy of the curve is given by the integral

$$\Delta\phi = \int_C \omega^1{}_2(T)\, ds. \tag{6.113}$$

6.8.2 Definition Let C be a smooth closed curve on M parametrized by arc length with geodesic curvature κ_g. The line integral $\oint_C \kappa_g\, ds$ is called the *total geodesic curvature.* If the curve is piecewise smooth, the total geodesic curvature is the sum of the integrals of each piece.

A circle of radius R gives an elementary example. The geodesic curvature is the constant $1/R$, so the total geodesic curvature is $(1/R)2\pi R = 2\pi$.

If we integrate formula 6.110 around a smooth simple closed curve C which is the boundary of a region R and use Stokes' Theorem, we get

$$\begin{aligned}\oint_C \kappa_g\, ds &= \oint_C d\phi - \oint_C \omega^1{}_2\, ds,\\ &= \oint_C d\phi - \int\int_R d\omega^1{}_2.\end{aligned}$$

For a smooth simple closed curve, $\int_C d\phi = 2\pi$. Using the Cartan-form version of the theorema egregium 4.106 we get immediately

$$\int\int_R K\, dS + \int_C \kappa_g\, ds = 2\pi. \tag{6.114}$$

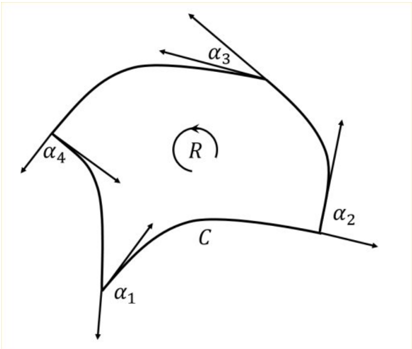


Fig. 6.7: Turning Angles

If the boundary of the region consists of k piecewise continuous functions as illustrated in figure 6.7, the change of the angle ϕ along C is still 2π, but the total change needs to be modified by adding the exterior angles α_k. Thus, we obtain a fundamental result called the *Gauss-Bonnet formula*,

6.8.3 Theorem

$$\int\int_R K\, dS + \int_C \kappa_g\, ds + \sum_k \alpha_k = 2\pi. \tag{6.115}$$

Every interior ι_k angle is the supplement of the corresponding exterior α_k angle, so the Gauss Bonnet formula can also be written as

$$\int\int_R K\, dS + \int_C \kappa_g\, ds + \sum_k (\pi - \iota_k) = 2\pi. \tag{6.116}$$

The simplest manifestation of the Gauss-Bonnet formula is for a triangle in the plane. Planes are flat surfaces, so $K = 0$ and the straight edges are geodesics, so $\kappa_g = 0$ on each of the three edges. The interior angle version of the formula then just reads $3\pi - \iota_1 - \iota_2 - \iota_3 = 2\pi$, which just says that the interior angles of a flat triangle add up to π. Since a sphere has constant positive curvature, the sum of the interior angles of a spherical triangle is larger than π. That amount of this sum over 2π is called the spherical excess. For example, the sum of the interior angles of a spherical triangle that is the boundary of one octant of a sphere is $3\pi/2$, so the spherical excess is $\pi/2$.

6.8.4 Definition The quantity $\int\int K\, dS$ is called the *total curvature*

6.8.5 Example A sphere of radius R has constant Gaussian Curvature $1/R^2$. The surface area of the sphere is $4\pi R^2$, so the total Gaussian curvature for the sphere is 4π.

6.8.6 Example For a torus generated by a circle of radius a rotating about an axis with radius b as in example (4.40), the differential of surface is $dS = a(b + a\cos\theta)\, d\theta d\phi$, and the Gaussian curvature is $K = \cos\theta/[a(b + a\cos\theta)]$, so the total Gaussian curvature is

$$\int_0^{2\pi}\int_0^{2\pi} \cos\theta\, d\theta d\phi = 0.$$

We now relate the Gauss-Bonnet formula to a topological entity.

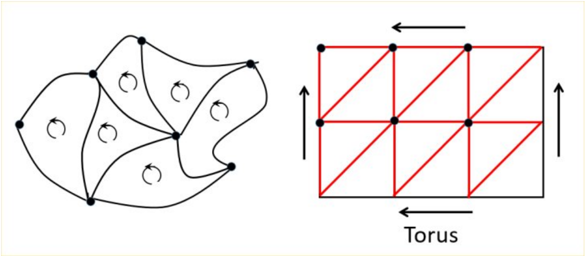


Fig. 6.8: Triangulation

6.8.7 Definition Let M be a 2-dimensional manifold. A *triangulation* of the surface is subdivision of the surface into triangular regions $\{\triangle_k\}$ which are the images of regular triangles under a coordinate patch, such that:

1) $M = \bigcup_k \triangle_k$.
2) $\triangle_i \bigcap \triangle_j$ is either empty, or a single vertex or an entire edge.
3) All the triangles are oriented in the same direction,

For an intuitive visualization of the triangulation of a sphere, think of inflating a tetrahedron or an octahedron into a spherical balloon. We state without proof:

6.8.8 Theorem Any compact surface can be triangulated.

6.8.9 Theorem Given a triangulation of a compact surface M, let V be the number of vertices, E the number of edges and F the number of faces. Then the quantity

$$\chi(M) = V - E + F, \tag{6.117}$$

is independent of the triangulation. In fact the quantity is independent of any "polyhedral" subdivision. This quantity is a topological invariant called the *Euler characteristic*.

6.8.10 Example

1. A balloon-inflated tetrahedron has $V = 4,\ E = 6,\ F = 4$, so the Euler characteristic of a sphere is 2.

2. A balloon-inflated octahedron has $V = 6,\ E = 12,\ F = 8$, so we get the same number 2.

3. The diagram on the right of gigure 6.8 represents a topological torus. In the given rectangle, opposites sides are identified in the same direction. The number of edges without double counting are shown in red, and the number of vertices not double counted are shown in black dots. We have $V = 6,\ E = 18\ F = 12$. So the Euler characteristic of a torus is 0.

4. In one has a compact surface, one can add a "handle", that is, a torus, by the following procedure. We excise a triangle in each of the two surfaces and glue the edges. We lose two faces and the number of edges and vertices cancel out, so the Euler characteristic of the new surface decreases by 2. The Euler characteristic of a pretzel is -4.

5. The Euler characteristic of an orientable surface of genus g, that is, a surface with g holes is given by $\chi(M) = 2 - 2g$.

6.8.11 Theorem Gauss-Bonnet
Let M be a compact, orientable surface. Then

$$\frac{1}{2\pi}\int_M K\, dS = \chi(M). \tag{6.118}$$

Proof Triangulate the surface so that $M = \bigcup_{k=1}^{F} \triangle_k$. We start the Gauss-Bonnet formula

$$\int\int_M K\,dS = \sum_{k=1}^{F} \int\int_{\triangle_k} K\,dS = -\sum_{k=1}^{F}\left[\oint_{\delta\triangle_k} \kappa_g\,ds + \pi + (\iota_{k1} + \iota_{k2} + \iota_{k2}))\right],$$

where F is the number of triangles and the ι_k's are the interior angles of triangle $\triangle_k$. The line integrals of the geodesic curvatures all cancel out since each edge in every triangle is traversed twice, each in opposite directions. Rewriting the equation, we get

$$\int\int_M K\,dS = -\pi F + \mathcal{S}$$

where $\mathcal{S}$ is the sum of all interior angles. Since the manifold is locally Euclidean, the sum of all interior angles at a vertex is 2π, so we have

$$\int\int_M K\,dS = -\pi F + 2\pi V$$

There are F faces. Each face has three edges, but each edge is counted twice, so $3F = 2E$, and we have $F = 2E - 2F$ Substituting in the equation above, we get,

$$\int\int_M K\,dS = -\pi(2E - 2F) + 2\pi V = 2\pi(V - E + F) = \chi(M).$$

This is a remarkable theorem because it relates the bending invariant Gaussian curvature to a topological invariant. Theorems such as this one which cut across disciplines, are the most significant in mathematics. Not surprisingly, it was Chern who proved a generalization of the Gauss-Bonnet theorem to general orientable Riemannian manifolds of even dimensions [5].

References

[1] H. Abrams 1971 *The World of M. C. Escher*, Meulenhoff International, Netherlands (1971)

[2] Atiyah, M., Hitchin, N., and Singer, I. 1978 *Self-duality in four-dimensional Riemannian geometry*, Proc. R. Soc. Lond. A. 362, (1978) pp 425-461.

[3] Belavin, A., Polyakov, A., Schwartz, A., and Tyupkin, Y. 1975 *Pseudoparticle Solutions of the Yang-Mills Equations*, Physics Letters B, Vol. 59, No, 1, (1975) pp 85-87.

[4] Chern, S. S. 1975: *Introduction to Differential Geometry.* Notes, University of California, Berkeley, CA.

[5] Chern, S. S. 1944 *A Simple Intrinsic Proof of the Gauss-Bonnet Formula for Closed Riemannian Manifolds*, Ann. Math. 45 (4), 1944, 747-752.

[6] Chern, S. S. 1955 *An Elementary Proof of the Existence of Isoterhmal Parameters on a Surface*, Proceedings of AMS, Vol. 6, No, 4, (1955), 771-782.

[7] Dray, T. 1984 1966 *The Relationship between Monopole Harmonics and Spin-weighted Spherical Harmonics*, Journal of Mathematical Physics, 26, 1030 (1985). https://doi.org/10.1063/1.526533

[8] Eguchi, T., Gilkey, P., and Hanson, A. 1980 *Gravitation, Gauge Theories and Differential Geometry*, Physics Reports, 66, No, 6 (1980). pp 213-293

[9] Georgi, H. 1999: *Lie Algebras in Particle Physics.* Frontiers in Physics V54, Westview (1999).

[10] Goldberg, J., Mcfarland, A., Newman, E., Rohrlich, F. and Sudarshan, C. 1966 *Spin-s Spherical Harmonics and* $\eth$, Proceedings of AMS, Vol. 6, No, 4, (1955), 771-782.

[11] Goldstein, Herbert, 1950: *Classical Mechanics.* Addison-Wesley, 398pp.

[12] Grey, Alfred, 2006: *Modern Differential Geometry of Curves and Surfaces with Mathematica.* 3rd ed. Chapman & Hall, CRC, 977pp.

[13] Hall, B. 2015: *Lie Groups, Lie Algebras, and Representations.* 2nd ed. Springer, Graduate Texts in Mathematics 222, Switzerland, (2015).

[14] Hicks, Noel, 1965: *Notes on Differential Geometry.* Van Nostrand Reinhold, Princeton, NJ.

[15] Hirsh, A. 2002 *Extension of 'Villarceau-Section' to Surfaces of Revolution with a Generating Conic* , J. of Geometry and Graphics. Vol 6, (2002) No, 2, pp121-132.

[16] Hoffman, K. and Kunze, R., 1971: *Linear Algebra.* 2nd ed. Prentice-Hall, 407pp.

[17] J. D. Jackson, 1962: *Classical Electrodynamics.*

[18] Kobayashi, S and Nomizu, K 1963: *Foundations of Differential Geometry.* Wiley & Sons, New York, London, 1963.

[19] Eisenhart, L. P. 1960: *A Treatise on the Differential Geometry of Curves and Surfaces.* Ginn and Company, Boston, 1909. Reprint by Dover Publications, New York, 1960.

[20] Abraham, R. and Marsden, J. 1978: *Foundations of Mechanics.* 2nd ed. Addison-Wesley, 838pp.

[21] Misner, C. W., Thorne, K. S., and Wheeler, J. A. 1973: *Gravitation.* W.H. Freeman and Company, 1279pp.

[22] Morris, M.S. and Thorne, K.S., 1987: Wormholes in Spacetime and their use for Interstellar Travel: A Tool for Teaching General Relativity. *Am. J. Phys.*, **56**(5), pp395-412.

[23] Müller, Thomas, 2008: Exact geometric optics in a Morris-Thorne wormhole spacetime. *Phys. Rev. D*, **77**, 044043-1-044043-11.

[24] Newman, E. and Penrose, R. 1962. *An Approach to Gravitational Radiation by a Method of Spin Coefficients*, Journal of Mathematical Physics 3, 566 (1962), pp 556-578. https://doi.org/10.1063/1.1724257

[25] Newman, E. and Janis, R. 1962. *Note on the Kerr Spinning-Particle Metric*, Journal of Mathematical Physics 6, 915 (1965), pp 915-917. https://doi.org/10.1063/1.1704350

[26] O'Neill, Barrett, 2006: *Elementary Differential Geometry.* 2nd ed. Academic Press, 503pp.

[27] Oprea, J., 1997: *Differential Geometry and Its Applications.* Prentice-Hall, Englewood Cliffs, NJ.

[28] Oprea, J., 2000: *The Mathematics of Soap Films.* Student Mathematical, Library, Volume 10, AMS, Providence, RI.

[29] Penrose, R. 1987. *On the Origins of Twistor Theory.* Gravitation and Geometry, a volume on honour of I. Robinson, Biblipolis, Naples, 25pp.

[30] Penrose, R. 1976. *Nonlinear Gravitons and Curved Twistor Theory*, General Relativity and Gravitation 7 (1976), No. 1 pp 31-52. and Geometry, a volume on honour of I. Robinson, Biblipolis, Naples, 25pp.

[31] Rogers, C and Schief, W. K., 2002. *On Regular Solutions of Euclidean Yang-Mills Equations.* Physic Cambridge University Press.

[32] Schwarz, A, 1977. *Bäcklund and Darboux Transformations : Geometry and Modern Applications in Soliton Theory.* Physics Letters B, Vol. 67, No, 2, (1977) pp 172-174.

[33] Spivak, M. 1965: *Calculus on Manifolds.* Addison Wesley (1965), 159pp

[34] Spivak, M. 1979: *A comprehensive introduction to Differential Geometry 5 volume set.* 2nd ed. Publish or Perish, Texas.

[35] Struik, Dirk J., 1961: *Lectures on Classical Differential Geometry.* 2nd ed. Dover Publications, Inc., 232pp.

[36] Taub, A. H. 1939: *Tensor Equations Equivalent to the Dirac Equations.* Annals of Mathematics, Vol 40, Number 4, pp.937-947.

[37] Trautman, A. 1977: *Solutions of the Maxwell and Yang-Mills Equations Associated with Hopf Fibrings.* International Journal of Theoretical Physics, Vol 16, Number 8 (1977), pp.561-565.

[38] Terng,C and Uhlenbeck, K, 2000: *Geometry of Solitons.* Notices of the AMS, Vol 47, Number 1.

[39] Ward, R. S., 1977: *On Self-dual Gauge Fields.* Physics Letters A, Vol. 61, No, 2, (1977) pp 81-82.

[40] Weisstein, E. W. From Mathworld- A Wolfram Web Resource: [Available online at `http://mathworld.wolfram.com/CostaMinimalSurface.html`.]

Index