Differential Geometry @ UCI

Math 162AB

Author: Zhiqin Lu
Institute: University of California, Irvine
Date: December 28, 2021
Version: 3.0
Bio: Zhiqin Lu is a professor at UCI

## Contents

0 Preliminaries
  0.1 Vector Spaces ................................................................. 1
  0.2 Inner Product ................................................................. 3
  0.3 Linear Transformation ...................................................... 5
  0.4 Orientation and Cross Product ........................................... 6
  0.5 Lines, Planes, and Spheres ............................................... 8
  0.6 Vector Calculus ............................................................... 11
  0.7 Einstein Convention ....................................................... 12

1 Calculus on Euclidean Space ................................................. 15
  1.1 Euclidean Space ............................................................. 15
  1.2 Tangent Vectors ............................................................ 15
  1.3 Directional Derivatives ................................................... 17
  1.4 Curves in \( \mathbb{R}^3 \) ...................................................... 19
  1.5 Dual Space ................................................................. 20
  1.6 1-forms ................................................................. 21
  1.7 Differential Forms ........................................................ 24
  1.8 Mappings ................................................................. 29

2 Frame Fields ........................................................................ 36
  2.1 Dot Product ................................................................. 36
  2.2 Curves ................................................................. 37
  2.3 The Frenet Formulas ....................................................... 40
  2.4 Arbitrary-Speed Curves ................................................... 43
  2.5 Covariant Derivatives ....................................................... 48
  2.6 Frame Fields ............................................................... 50
  2.7 Connection forms .......................................................... 52
  2.8 The Structure Equations .................................................. 54

3 The Fundamental Theorem of Curves ....................................... 58
  3.1 Isometry of \( \mathbb{R}^3 \) ...................................................... 58
  3.2 The Fundamental Theorem of Curves ................................... 60

4 Calculus on Surface ................................................................ 62
  4.1 Surfaces in \( \mathbb{R}^3 \) ...................................................... 62
  4.2 Patch Computations ......................................................... 67
  4.3 Differentiable Functions and Tangent Vectors ......................... 69
  4.4 Differential Forms on a Surface ......................................... 73
  4.5 Mappings of Surfaces ....................................................... 76
  4.6 Integration of Forms ......................................................... 79
  4.7 Topological properties of surfaces ...................................... 81
## CONTENTS

4.8 Manifolds (Abstract Surfaces) ......................................... 85

5 **Shape Operators** 87
  5.1 The shape operator of $M \subset \mathbb{R}^3$ ............................... 87
  5.2 Normal Curvature ...................................................... 88
  5.3 Gauss Curvature ...................................................... 90
  5.4 Computational Techniques ........................................... 92
  5.5 The Implicit Case .................................................... 94
  5.6 Special Curves in a Surface ........................................ 95
  5.7 Surface of Revolution ................................................ 96

6 **Geometry of Surfaces in $\mathbb{R}^3$** 101
  6.1 The Fundamental Equations ......................................... 101
  6.2 Form Computations .................................................. 102
  6.3 Some Global Theorems ............................................... 105
  6.4 Isometries and Local Isometries .................................... 106
  6.5 Intrinsic Geometry of Surfaces in $\mathbb{R}^3$ ....................... 108
  6.6 Orthogonal Coordinates ............................................. 109
  6.7 Integration and Orientation ........................................ 109
  6.8 Total Curvature ...................................................... 110
  6.9 Congruence of Surfaces ............................................. 112
Chapter 0  Preliminaries

Introduction

- Vector space
- Inner product on vector space
- Linear transformation
- Lines, planes, and spheres
- Einstein Convention
- Vector Calculus

0.1 Vector Spaces

Definition 0.1

A vector space $V$ is a nonempty set with two binary operations addition “+” and scalar multiplication “·” satisfying the following eight axioms: let $u, v, w \in V$ and $r, s \in \mathbb{R}$, we have

- $u + v = v + u$;
- $u + (v + w) = (u + v) + w$;
- $0 + u = u + 0 = u$ for a vector called 0;
- $(rs) \cdot u = r \cdot (s \cdot u)$;
- $(r + s) \cdot u = r \cdot u + s \cdot u$;
- $r \cdot (u + v) = r \cdot u + r \cdot v$;
- $0 \cdot u = 0$;
- $1 \cdot u = u$.

As is well-known, the set $V$ satisfying the first three axioms form an Abel semi-group. The existence of the inverse of a vector $u$ can be verified by using the scalar multiplication. Let $u$ be a vector, we claim that $(-1) \cdot u$ is the inverse of $u$ because

$$u + (-1) \cdot u = (1 + (-1)) \cdot u = 0.$$ 

Similarly, we can define the subtraction by

$$u - v = u + (-1) \cdot v.$$ 

Note We sometimes omit the - and write, for example, $ru$ for $r \cdot u$.

 External Link. Here is the video explanation of vector space (linear independence). Useful!

 In linear algebra, we restrict ourselves to finite dimensional vector space. But, a lot of results in finite dimensional case can be extended to infinite dimensional case as well as abstract vector space cases.

In the following, we give some examples of vector spaces.

Example 0.1 $\mathbb{R}^n$, the $n$-dimensional Euclidean space, is the set of all $n$-vectors

$$\mathbb{R}^n = \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\
vdots \\
x_n \end{bmatrix} \mid x_i \in \mathbb{R} \right\}.$$ 

Example 0.2 The set of all $m \times n$ matrices for an $(mn)$-dimensional vector space.
Example 0.3 The space of polynomials of degree no more than \( n \), where \( n \) is a nonnegative integer, is a vector space.

A single-variable polynomial of degree no more than \( n \) can be expressed as

\[
p(t) = a_0 + a_1 t + \cdots + a_n t^n.
\]

Let

\[
q(t) = b_0 + b_1 t + \cdots + b_n t^n
\]

be another polynomial. Define the addition to be

\[
(p + q)(t) = (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n,
\]

and the scalar multiplication to be: let \( \lambda \in \mathbb{R} \)

\[
(\lambda p)(t) = (\lambda a_0) + (\lambda a_1) t + \cdots + (\lambda a_n) t^n.
\]

With respect to the addition and scalar multiplication, the space is a vector space.

In the above two examples, the dimensions of the vector spaces are finite. Let’s show some examples of infinite dimensional vector spaces.

Note The set of all polynomials of degree equal to \( n \) is not a vector space.

Example 0.4 Moreover, the space of all real-valued functions on a set is a (infinite dimensional) vector space.

Let’s recall the definition of a function.

Definition 0.2

A function \( f : X \to Y \) is a triple \((f, X, Y)\), where \( X, Y \) are sets, and \( f \) is an assignment, or a rule, that for any element \( x \) in \( X \), there is a unique \( y = f(x) \) in \( Y \) attached to it. \( X \) is called the domain of \( f \), and \( Y \) is called the codomain of \( f \). The range is the subset of the codomain \( Y \) consisting of all \( f(x) \) when \( X \) is running through \( X \). The assignment sometimes is written as \( x \mapsto f(x) \).

Thus a complete description of a function can be given as

\[
f : X \to Y, \quad x \mapsto f(x).
\]

As above, we can define the addition and scalar multiplication as

\[
(f + g)(x) = f(x) + g(x), \quad (\lambda f)(x) = \lambda f(x).
\]

Remark The vector space we study in differential geometry are the “abstract” vector space, which is on the contrary to the vector spaces we studied in Math 3A.

The following concepts are defined in abstract vector spaces similar to those in \( \mathbb{R}^n \).

1. linear combination, span;
2. linear dependence and independence;
3. basis and dimension.

Remark In infinite dimensional space, a basis is defined by a set of vectors

\[
\mathcal{B} = \{v_1, v_2, \cdots, v_n, \cdots\}
\]

satisfying the following
0.2 Inner Product

1. any finite subset of \( \mathcal{I} \) is linearly independent;
2. any element can be expressed as a (finite) linear combination of element in \( \mathcal{I} \).

Let \( (V, +, \cdot) \) be a vector space. We can endow geometric structure onto it by defining the concept of inner product.

**Remark** Let

\[ \mathcal{S} = \{v_1, v_2, \ldots, v_n\} \]

be a finite set. Let

\[ V = \text{Span } \mathcal{S} = \text{Span } \{v_1, v_2, \ldots, v_n\}. \]

Then we say \( \mathcal{S} \) spans \( V \), and \( \mathcal{S} \) is a *spanning* set of \( V \).

### 0.2 Inner Product

The addition and scalar multiplication define the *algebraic structure* of a vector space. In order to introduce geometry to linear algebra, we can endow geometric structure onto it by defining the concept called *inner product*.

**Definition 0.3**

An inner product on a vector space \( V \) is a function \( \langle \ , \ \rangle \) \( V \times V \to \mathbb{R} \)

satisfying the following properties: let \( u, v, w \in V \) and \( r, s \in \mathbb{R} \), we have

1. \( \langle u, v \rangle = \langle v, u \rangle \)  **Symmetry**
2. \( \langle ru + sw, v \rangle = r \langle u, v \rangle + s \langle u, w \rangle \)  **Linearity**
3. \( \langle u, u \rangle \geq 0 \) and equality is true if and only if \( u = 0 \)  **Positivity**

Once we introduce the inner product, we introduce geometry into vector space. For example

**Definition 0.4**

(Length of a vector) We can define the length, or the *norm*, of a vector to be

\[ \|u\| = \sqrt{\langle u, u \rangle}. \]

(Distance) Let \( u, v \) be two vectors. Then their distance is defined to be

\[ \text{dist}(u, v) = \|u - v\| = \sqrt{\langle u - v, u - v \rangle}. \]

Note in the definition of distance, we used both the geometric structure (inner product) and algebraic structure (subtraction is defined by \( u - v = u + (-1) \cdot v \)).

**Example 0.5** In \( \mathbb{R}^3 \) (and in \( \mathbb{R}^n \)), the ordinary *dot product*

\[ \langle (a^1, a^2, a^3), (b^1, b^2, b^3) \rangle = a^1b^1 + a^2b^2 + a^3b^3 \]

is an inner product.

However, there are many other ways one can define inner product on \( \mathbb{R}^3 \). For example, we can define

\[ \langle (a^1, a^2, a^3), (b^1, b^2, b^3) \rangle = 2a^1b^1 + 3a^2b^2 + 4a^3b^3. \]

It is an inner product. In general, let

\[ R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \]
be a positive definite matrix. Then
\[
\langle (a^1, a^2, a^3), (b^1, b^2, b^3) \rangle = \sum_{i,j=1}^{3} r_{ij} a^i b^j
\]
is an inner product.

**Example 0.6** We let \( R[t] \) be the vector space of all polynomials. Define the inner product
\[
\langle p, q \rangle = \int_{-1}^{1} p(x)q(x) \, dx
\]
More general, if \( \rho(x) > 0 \) be a positive continuous function, then
\[
\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\rho(x) \, dx
\]
is an inner product.

One of the most important inequalities in mathematics is called the **Cauchy-Schwarz Inequality**.

**Theorem 0.1**

*Let \( \mathbf{u}, \mathbf{v} \) be two vectors. Then we have
\[
|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|,
\]
and the equality holds if and only if \( \mathbf{u}, \mathbf{v} \) are linearly dependent.*

**Proof.** This is the standard proof. Let \( t \) be a real number. Then by the positivity of the inner product, we have
\[
\langle \mathbf{u} + t\mathbf{v}, \mathbf{u} + t\mathbf{v} \rangle \geq 0.
\]
Using the linearity, we get
\[
t^2\langle \mathbf{v}, \mathbf{v} \rangle + 2t\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{u} \rangle \geq 0.
\]
Since the above inequality is true for any real number \( t \), the discriminant
\[
\Delta = 4|\langle \mathbf{u}, \mathbf{v} \rangle|^2 - 4\langle \mathbf{u}, \mathbf{u} \rangle \cdot \langle \mathbf{v}, \mathbf{v} \rangle \leq 0,
\]
which proves the inequality.

**Second Proof.** Assume that \( \mathbf{v} \neq 0 \). Then we have
\[
\langle \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}, \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} \rangle \geq 0.
\]
Expanding the above inequality, we obtain the Cauchy-Schwarz Inequality.

**Remark** The Cauchy-Schwarz inequality is also called the Cauchy-Bunyakovsky-Schwarz inequality.

The inequality for sums was published by Augustin-Louis Cauchy (1821), while the corresponding inequality for integrals was first proved by Viktor Bunyakovsky (1859). The modern proof of the integral version was given by Hermann Schwarz (1888).

**Definition 0.5**

*Two vectors \( \mathbf{u}, \mathbf{v} \) are called orthogonal, if \( \langle \mathbf{u}, \mathbf{v} \rangle = 0 \).*

**Example 0.7** In \( \mathbb{R}^3 \), the vector \((1, 2, 3)\) is orthogonal to \((4, -5, 2)\) with respect to the dot product, because
\[
(1, 2, 3) \cdot (4, -5, 2) = 1 \cdot 4 + 2 \cdot (-5) + 3 \cdot 2 = 0.
\]
Example 0.8 Under the inner product
\[ \langle p, q \rangle = \int_{-1}^{1} p(x)q(x) \, dx. \]
The function \( x \) and \( x^2 + 1 \) are orthogonal because \( p(x) \) is an odd function and \( q(x) \) is an even function.
An orthonormal basis of an \( n \)-dimensional vector space is a set \( \{v_1, \cdots, v_n\} \) such that
\[ \langle v_i, v_j \rangle = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}. \]
We shall introduce the Kronecker symbol \( \delta_{ij} \) as follows
\[ \delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}. \]
Under this notation, we have
\[ \langle v_i, v_j \rangle = \delta_{ij}. \]

0.3 Linear Transformation

Definition 0.6 ♣
Given two vector spaces, \( V \) and \( W \), a linear transformation \( T \) from \( V \) to \( W \) is a mapping
\[ T : V \rightarrow W \]
such that
\[ T(\alpha u + \beta v) = \alpha T(u) + \beta T(v) \]
for all \( u, v \in V \) and \( \alpha, \beta \in \mathbb{R} \).

A linear transformation of a vector space to itself is called an endomorphism. One of the most important concepts of an endomorphism is its eigenvalue and eigenvector.

Definition 0.7 ♣
Let \( T : V \rightarrow V \) be an endomorphism. Assume that there is a \( v \in V \) and \( v \neq 0 \) such that
\[ T(v) = \lambda v \]
for some complex number \( \lambda \). Then \( \lambda \) is called an eigenvalue of \( T \) and \( v \) is an eigenvector of \( \lambda \).

Note In the following, we need to prove that the definition of eigenvalue is equivalent to the definition of an eigenvalue of a matrix.

Let \( V \) be a finite dimensional space and let
\[ \mathfrak{B} = \{v_1, \cdots, v_n\} \]
be a basis of \( V \). We use the notation \( [x]_{\mathfrak{B}} \) to represent the coordinates of \( x \in V \), that is, when we write \( x \) in terms of the linear combination of the basis,
\[ x = c_1 v_1 + \cdots + c_n v_n, \]
we have
\[
[x]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.
\]

Now let \( T : V \to W \) be a linear transformation. As above, we assume that \( B \) is the basis of \( V \), and let \( C = \{ w_1, \ldots, w_m \} \) be a basis of \( W \). Then the matrix representation \( M \) of \( T \) is
\[
M = [[T(v_1)]_C, \ldots, [T(v_n)]_C]
\]
in the sense that
\[
[T(x)]_C = M \cdot [x]_B.
\]

Now we specialize the above result to the following case. Let 
\( T : V \to V \) 
be an endomorphism. Let 
\( B = \{ v_1, \ldots, v_n \} \), \( C = \{ w_1, \ldots, w_n \} \) 
be two bases. Let \( M_1, M_2 \) be the matrix representatives with respect to the two bases respectively. We thus have
\[
[T(x)]_B = M_1 \cdot [x]_B;
\]
\[
[T(x)]_C = M_2 \cdot [x]_C.
\]

Let \( A \) be the invertible matrix such that 
\[
[x]_B = A \cdot [x]_C.
\]
Such a matrix is called a transition matrix. Then
\[
[T(x)]_B = A \cdot [T(x)]_C.
\]

As a result,
\[
A \cdot M_2 \cdot [x]_C = A \cdot [T(x)]_C = [T(x)]_B = M_1 \cdot A \cdot [x]_C.
\]
Thus we have
\[
AM_2 = M_1 A,
\]
or
\[
M_2 = A^{-1}M_1 A.
\]
Thus \( M_1, M_2 \) are similar, having the same eigenvalue set.

0.4 Orientation and Cross Product

Let \( B \) and \( C \) be two bases. We say that \( B \) and \( C \) are having the same orientation, if when we write
\[
v_i = \sum_{j=1}^{n} a_{ij} w_j
\]
for \( i = 1, \ldots, n \), then we have \( \det(A) = \det(a_{ij}) > 0 \). They give the opposite orientation if \( \det(a_{ij}) < 0 \).

Example 0.9 The left hand and the right hand define two opposite orientations of \( \mathbb{R}^3 \).
Example 0.10 Let
\[ B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} \]
and
\[ C = \{(1, 1, 0), (1, 0, -1), (2, 1, 3)\} \]
are of the opposite orientation.

**Definition 0.8**
Let \( \{e_1, e_2, e_3\} \) be the standard basis of \( \mathbb{R}^3 \). If
\[ u = \sum_{i=1}^{3} a_i e_i, \quad v = \sum_{j=1}^{3} b_j e_j \]
are two vectors in \( \mathbb{R}^3 \), the cross product of \( u, v \) is given by
\[ u \times v = (a_2 b_3 - a_3 b_2)e_1 + (a_3 b_1 - a_1 b_3)e_2 + (a_1 b_2 - a_2 b_1)e_3, \]
or we can write
\[ u \times v = \det \begin{bmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}. \]

The cross product satisfies the following properties:

**Lemma 0.1**
Let \( u, v, w \in \mathbb{R}^3 \) and \( r \in \mathbb{R} \). Then
- \( u \times v = -v \times u \)
- \( (ru) \times v = r(u \times v) \)
- \( u \times v = 0 \) if and only if \( u \) and \( v \) are linearly dependent
- \( (u + v) \times w = u \times w + v \times w \)
- \( u \times v \) is perpendicular to both \( u \) and \( v \) under the usual dot product
- \( \|u \times v\| = \|u\| \cdot \|v\| \sin \theta \), where \( \theta \) is the angle between \( u \) and \( v \).
- \( \{u, v, u \times v\} \) gives a right hand orientation to \( \mathbb{R}^3 \) if \( \{u, v\} \) is linearly independent

As a result, we have the relationship between the inner product and outer product (cross product)
\[ |\langle u, v \rangle|^2 + \|u \times v\|^2 = \|u\|^2 \cdot \|v\|^2, \]
which implies the Cauchy-Schwarz Inequality in the three dimensional space.

**Definition 0.9**
The mixed (or triple) product of \( u, v, w \) is
\[ [u, v, w] = \langle u \times v, w \rangle. \]

A geometric interpretation of the norm of the cross product is that it is the area of the parallelogram spanned by \( u, v \). A geometric interpretation of the mixed scalar product is that its absolute value is the volume of the parallelepiped spanned by \( u, v, w \).

**Note** One may ask whether or not one can define cross-product-like structure one vector spaces on other dimensions. The problem is related to the Frobenius Theorem, Octonion, and the Cayley Plane. But I admit I
have never studied those.

0.5 Lines, Planes, and Spheres

In this section, we use vector notations to express some basic objects in analytic geometry.

**Definition 0.10**

The line through $x_0 \in \mathbb{R}^3$ and parallel to a vector $v \neq 0$ has the equation

$$\alpha(t) = x_0 + tv.$$

**Remark** This is a vector notation of parametrization of a line.

**Example 0.11** Let $x_1, x_2 \in \mathbb{R}^3$ be two points. Then the line through $x_1$ and $x_2$ in $\mathbb{R}^3$ has the equation.

$$\alpha(t) = x_1 + t(x_2 - x_1).$$

**Definition 0.11**

The plane through $x_0$ perpendicular to $n \neq 0$ has the equation

$$\langle x - x_0, n \rangle = 0.$$

**Lemma 0.2**

Let $\{u, v\}$ be two linearly independent vectors. Then the place through $x_0$ and parallel to the subspace spanned by $\{u, v\}$ has the equation

$$[u, v, x-x_0] = \langle x-x_0, u \times v \rangle = 0.$$

**Definition 0.12**

The sphere in $\mathbb{R}^3$ with center $m$ and radius $r > 0$ has equation

$$\langle x - m, x - m \rangle = \|x - m\|^2 = r^2.$$  \hspace{1cm} (1)

**Remark** Let

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad m = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}.$$

Then we get the usual equation of a sphere

$$(x_1 - m_1)^2 + (x_2 - m_2)^2 + (x_3 - m_3)^2 = r^2.$$

**Example 0.12 (Kelvin Transformation)**

We consider the equation of a sphere (1). Let $x_0$ be a point on the sphere, that is, we have

$$\langle x_0 - m, x_0 - m \rangle = r^2.$$

The Kelvin Transformation is a map

$$K : \mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto x_0 + \frac{x - x_0}{\|x - x_0\|^2}.$$

By a straightforward computation, we have $K^2 = id$. Assume

$$\langle K(x) - m, K(x) - m \rangle = r^2.$$
We get

$$1 + 2\langle x - x_0, x_0 - m \rangle = 0.$$  

So the Kelvin transformation maps a sphere to a plane.

_external_link_ The detailed computation can be found here.
Example 0.13 (Ptolemy Inequality) Let $u, v, w, x$ be four vectors in the Euclidean plane. Then we have

$$\|u - w\| \cdot \|v - x\| \leq \|u - v\| \cdot \|x - w\| + \|u - x\| \cdot \|v - w\|.$$ 

The equality is valid if and only if these four vectors are concyclic.

*External Link.* The Ptolemy Inequality is closely related to the Ptolemy Theorem. For details of the Ptolemy and his theorem, see Wikipedia of Ptolemy Theorem.
0.6 Vector Calculus

In differential geometry, in addition to study functions of several variables. We also need to study vector-valued functions.

We can define derivatives, indefinite integral and definite integrable in similar ways to those of multi-variable functions.

Let $V, W$ be finite dimensional vector spaces. Let $F : V \to W$ be a differentiable function, with definition as follows.

**Definition 0.13**

We fix a basis of $V$ and using that basis, we identify $V$ to $\mathbb{R}^n$. Similarly, and we fix a basis of $W$ and identify it to $\mathbb{R}^m$. Then we can identify $F : V \to W$ by $F : \mathbb{R}^n \to \mathbb{R}^m$. So $F$ is differentiable if and only if $F$ is a differentiable as a mapping of $\mathbb{R}^n \to \mathbb{R}^m$.

**External Link.** Here is a video clip for the details of the above definition.

Let $f : \mathbb{R} \to \mathbb{R}^n$ be a single variable vector-valued function. We can define

$$\frac{df}{dt} = \begin{bmatrix} \frac{df_1}{dt} \\ \vdots \\ \frac{df_n}{dt} \end{bmatrix}$$

if

$$f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

Likewise, we can define

$$\int f(t) \, dt, \quad \int_a^b f(t) \, dt$$

in similar ways.

If $f : \mathbb{R} \to V$ be an abstract vector-valued function, we can identity $V$ with $\mathbb{R}^n$ under a fixed basis, and define the derivative, integral, etc, similarly.

**Lemma 0.3**

Let $f : \mathbb{R} \to V, g : \mathbb{R} \to V$ and let $\langle \cdot, \cdot \rangle$ be an inner product on $V$. Then if $f$ and $g$ are differentiable, so is $\langle f, g \rangle$, which is a function of one variable. Moreover, we have

$$\frac{d}{dt} \langle f, g \rangle = \langle \frac{df}{dt}, g \rangle + \langle f, \frac{dg}{dt} \rangle.$$

**Lemma 0.4**

Using the notations as in the above lemma, we have

$$\frac{d}{dt} (f \times g) = \frac{df}{dt} \times g + f \times \frac{dg}{dt}.$$
Both of the above two lemmas can be proved directly. Moreover, we can generalize the above results into the following.

Let $V, W, S$ be vector spaces (probably of infinite dimensional) and let

$$K : V \times W \to S$$

be a map. We say $K$ is bilinear, if $K$ is linear with respect to each component.

Both the inner product and cross product are bilinear mappings.

Let $f : \mathbb{R} \to V, g : \mathbb{R} \to W$ be differentiable functions. Then the function

$$h(t) = K(f(t), g(t))$$

is differentiable, and

$$\frac{dh}{dt} = K\left(\frac{df}{dt}, g(t)\right) + K\left(f(t), \frac{dg}{dt}\right).$$

**Definition 0.14**

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. We say $f$ is of class $C^k$, if all derivatives up through order $k$ exist and are continuous.

$f : \mathbb{R}^n \to \mathbb{R}$ is of class $C^k$ if all its (mixed) partial derivatives of up through order $k$ exist and are continuous. A vector-valued function is of class $C^k$ if all of its components with respect to a given basis are of class $C^k$.

If $f$ is of class $C^k$ for any $k$, we say $f$ is of $C^\infty$, or we say $f$ is smooth.

We assume most of the functions we shall study in this course are smooth, or at least of $C^3$.

Finally, we review the chain rule. Let $x$ be a function of $(u_1, \cdots, u_n)$, and if each $u_i$ are functions of $(v_1, \cdots, v_m)$, say,

$$u_i = u_i(v_1, \cdots, v_m).$$

for $i = 1, \cdots, n$. Then we have the chain rule

$$\frac{\partial x}{\partial v_\alpha} = \sum_{i=1}^n \frac{\partial x}{\partial u_i} \cdot \frac{\partial u_i}{\partial v_\alpha} \quad (2)$$

for $\alpha = 1, \cdots, m$.

### 0.7 Einstein Convention

**Definition 0.15**

When an index variable appears twice in a single term and is not otherwise defined, it implies summation of that term over all the values of the index. When an index variable appears only once, it implies that the equation is valid for every value of such an index.

For example, Equation (2) can be written as

$$\frac{\partial x}{\partial v_\alpha} = \frac{\partial x}{\partial u_i} \cdot \frac{\partial u_i}{\partial v_\alpha}.$$
**Example 0.14** Let $A = (a_{ij})$, $B = (b_{ij})$, and $C = (c_{ij})$ be matrices. Then the matrix multiplication,

\[ C = AB, \]

can be written using the Einstein Convention as

\[ c_{ij} = a_{ik}b_{kj}. \]

The Einstein Convention gives another way to express and generalize linear algebra.

**Note** Let’s discuss the representation of a matrix. In linear algebra, there are three ways to represent a matrix

\[ A = [a_1, \cdots, a_n] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \]

where $a_1, \cdots, a_n$ are column vectors. The easiest way to represent a matrix is to a capital letter, say $A$. But this would contain the least amount of information about the matrix. On the other extreme, if we represent a matrix by providing all the details, it would be too clumsy to write down.

Here we give the fourth method of representing a matrix, by writing it as $(a_{ij})$, which takes care of both simplicity and information.

**Example 0.15** Prove the associativity of matrix multiplication.

**Proof.** Let $A = (a_{ij})$, $B = (b_{ij})$ be matrices. Let $D = (d_{ij})$ be the matrix $D = AB$. In terms of the Einstein Convention, $D = AB$ is equivalent to

\[ d_{ij} = a_{ik}b_{kj}. \quad (3) \]

Here the index $k$ is called a dummy index in the sense that we can replace it with other indices without changing the equations:

\[ d_{ij} = a_{ik}b_{kj} = a_{il}b_{lj} = a_{ia}b_{aj}. \quad (4) \]

Now let $C = (c_{ij})$, $E = BC = (e_{ij})$, $F = (AB)C = (f_{ij})$ and $G = A(BC) = (g_{ij})$. Then the entries for $(AB)C = DC$ would be

\[ f_{ij} = d_{ik}c_{kj} = d_{it}c_{tj}. \]

From (4), we know that $d_{it} = a_{ik}b_{kt}$. Thus

\[ f_{ij} = a_{ik}b_{kt}c_{tj}. \]

The reason we use $t$ as the dummy index in (4) is because $k$ has been used in (3) so we need to use a different one, keeping indices repeated at most twice.

Similarly, we have

\[ g_{ij} = a_{it}b_{tk}c_{kj}. \]

Thus $f_{ij} = g_{ij}$ and hence

\[ (AB)C = A(BC), \]

proving the associativity.

**Example 0.16** Using the Einstein Convention to prove the following version of the Cauchy inequality. Let $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$. Then

\[ |x \cdot y|^2 \leq \|x\|^2 \cdot \|y\|^2. \]
Proof. Using the Einstein Convention, we can write

\[ \mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i = x_i y_i. \]

Thus

\[ |\mathbf{x} \cdot \mathbf{y}|^2 = \left( \sum_{i=1}^{n} x_i y_i \right)^2 = \left( \sum_{i=1}^{n} x_i y_i \right) \cdot \left( \sum_{j=1}^{n} x_j y_j \right) = x_i y_i x_j y_j. \]

On the other hands, we can write

\[ \|\mathbf{x}\|^2 \cdot \|\mathbf{y}\|^2 = \left( \sum_{i=1}^{n} x_i^2 \right) \cdot \left( \sum_{j=1}^{n} y_j^2 \right) = x_i^2 y_j^2. \]

Thus the Cauchy inequality, written under the Einstein Convention, is

\[ x_i^2 y_j^2 - x_i y_i x_j y_j = \frac{1}{2} (x_i^2 y_j^2 + x_j^2 y_i^2) - x_i y_i x_j y_j = \frac{1}{2} (x_i y_j - x_j y_i)^2 \geq 0. \]

This completes the proof.

Note If \( n = 3 \), then we can define

\[ (\mathbf{x}, \mathbf{y}) \mapsto (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1) = \mathbf{x} \times \mathbf{y} \]

which is the cross product. If \( n \neq 3 \), then the vector \((x_i y_j - x_j y_i)\) for \( i < j \) is of dimension \( n(n - 1)/2 \neq 3 \). This explains why we can only define the cross product in 3 dimensional vector space. A more general algebraic product, called wedge product, will be used in any dimensional vector spaces to catch in the excess of the Cauchy inequality.

External Link. As fun reading, you can find the Shoelace Formula in the Wikipedia, which is related to both the cross product and wedge product.
Chapter 1  Calculus on Euclidean Space

1.1 Euclidean Space

**Definition 1.1**

Euclidean 3-space $\mathbb{R}^3$ is the set of all ordered triples of real numbers. Such a triple $\mathbf{p} = (p_1, p_2, p_3)$ is called a point of $\mathbb{R}^3$.

By last chapter, $\mathbb{R}^3$ is a vector space.

**Definition 1.2**

On $\mathbb{R}^3$, there are three natural real-valued functions $x, y, z$, defined by

$$x(\mathbf{p}) = p_1, \quad y(\mathbf{p}) = p_2, \quad z(\mathbf{p}) = p_3.$$  

These functions are called natural coordinate functions of $\mathbb{R}^3$.

**Remark** We shall also use index notation for these functions, writing

$$x_1 = x, \quad x_2 = y, \quad x_3 = z.$$  

**Definition 1.3**

A real-valued function $f$ on $\mathbb{R}^3$ is differentiable (or infinitely differentiable, or smooth, or of class $C^\infty$) provided all partial derivatives of $f$, of all orders, exist and are continuous.

As we know from the previous chapter, the space of smooth functions form a vector space, that is, let $f, g$ be two smooth functions of $\mathbb{R}^3$ and let $\lambda \in \mathbb{R}$, we have

$$(f + g)(\mathbf{p}) = f(\mathbf{p}) + g(\mathbf{p}), \quad (\lambda f)(\mathbf{p}) = \lambda f(\mathbf{p}).$$

In addition, we have

$$(fg)(\mathbf{p}) = f(\mathbf{p})g(\mathbf{p}).$$

The space of smooth functions, with respect to the three operations: addition, scalar multiplication, and the multiplication forms an algebra.

1.2 Tangent Vectors

**Definition 1.4**

A tangent vector, or a vector $\mathbf{v}_p$ to $\mathbb{R}^3$ consists of two points of $\mathbb{R}^3$: its vector part $\mathbf{v}$ and its point of application $\mathbf{p}$.
1.2 Tangent Vectors

Definition 1.5

Let \( p \) be a point of \( \mathbb{R}^3 \). The set \( T_p(\mathbb{R}^3) \) consisting of all tangent vectors that have \( p \) as point of application is called the tangent space of \( \mathbb{R}^3 \) at \( p \).

\[ T_p(\mathbb{R}^3) \]

Note: Tangent space is a vector space.

Definition 1.6

A tangent vector field, or a vector field \( V \) on \( \mathbb{R}^3 \) is a function that assigns to each point \( p \) of \( \mathbb{R}^3 \) a tangent vector \( V(p) \) to \( \mathbb{R}^3 \) at \( p \). So it is a function \( \mathbb{R}^3 \rightarrow \bigcup_p T_p(\mathbb{R}^3) \).

Note: Vector field is one of the most important concepts in differential geometry. By the above definition, a vector field is just a vector-valued function. This is because \( \mathbb{R}^3 \) is a flat space, and hence there are global basis under which all tangent spaces can be identified as \( \mathbb{R}^3 \). In general, a vector field defines a different type of "functions" comparing to the traditional one.

Remark: By definition, a vector field doesn’t have to be smooth. However, in this course, we always assume it is smooth (or at least of \( C^3 \)) when regarding it as a vector-valued function.

The domain of a vector field doesn’t have to be on the whole \( \mathbb{R}^3 \); it could be an open set of \( \mathbb{R}^3 \), or a curve or a surface in \( \mathbb{R}^3 \). In the latter to cases, we say that the vector field is along the curve or surface.

The space of vector fields is obviously a vector space. However, it has finer structure than that. It is a module over the algebra of smooth functions.

There are two operations on the space of vector fields: addition and scalar multiplication. Let \( V, W \) be two vector fields such that \( V = v_i U_i \), \( W = w_i U_i \). Let \( \lambda \in \mathbb{R} \). Then we can define
\[
V + W = \sum_i (v_i + w_i) U_i \\
\lambda V = \sum_i (\lambda v_i) U_i. 
\]

Moreover, let \( f \) be a smooth function of \( \mathbb{R}^3 \). Then we can define
\[
(fV)(p) = f(p)V(p)
\]
for all \( p \). Of course, such kind of multiplication can be localized to the case when \( V \) and \( f \) are only defined on a subset of \( \mathbb{R}^3 \).

Note: The scalar multiplication coincides with the above multiplication by regarding a scalar as a constant function.

Note: We can also say that the space of vector fields is a module over the ring of smooth functions. But the algebra of smooth functions carries more structure than the ring structure of smooth functions: it has the additional scalar multiple structure. Therefore it is better to say the module over the algebra of smooth functions than that over the ring of smooth functions.

Definition 1.7

Let \( U_1, U_2 \) and \( U_3 \) be the vector fields on \( \mathbb{R}^3 \) such that
\[
U_1(p) = (1, 0, 0)_p \\
U_2(p) = (0, 1, 0)_p \\
U_3(p) = (0, 0, 1)_p
\]
for each \( p \) of \( \mathbb{R}^3 \). We call \( \{U_1, U_2, U_3\} \) the natural frame field of \( \mathbb{R}^3 \).
1.3 Directional Derivatives

Remark For fixed point, \( \{U_1, U_2, U_3\} \) provides the standard basis of \( \mathbb{R}^3 \), usually expressed as \( \{e_1, e_2, e_3\} \).

The following result is a generalization of what we have learned in linear algebra.

**Lemma 1.1**

If \( V \) is a vector field of \( \mathbb{R}^3 \), then there are three uniquely determined real-valued functions \( v_1, v_2, v_3 \) on \( \mathbb{R}^3 \) such that

\[
V = v_1 U_1 + v_2 U_2 + v_3 U_3.
\]

These three functions are called **Euclidean coordinate functions** of \( V \).

**Proof.** For fixed \( p \in \mathbb{R}^3 \), \( V(p) \) defines a vector in \( \mathbb{R}^3 \), therefore there is unique numbers \( v_1(p), v_2(p), v_3(p) \) such that

\[
V(p) = v_1(p)U_1(p) + v_2(p)U_2(p) + v_3(p)U_3(p).
\]

Thus

\[
V = v_i U_i
\]

by definition.

1.3 Directional Derivatives

**Definition 1.8**

Let \( f \) be a differentiable real-valued function on \( \mathbb{R}^3 \), and let \( v_p \in T_p(\mathbb{R}^3) \) be a tangent vector to \( \mathbb{R}^3 \). Then the number

\[
v_p[f] = \left. \frac{d}{dt}(f(p + tv)) \right|_{t=0}
\]

is called the **derivative** of \( f \) with respect to \( v_p \).

**Note** It is called the directional derivative because \( p + tv \) for non-negative real number \( t \) represents a ray starting from \( p \) in the direction \( v \). We have encountered directional derivative in Calculus. Here the emphasis is that “vector” (which is an algebraic concept) can be identified as a “derivative” (which is a calculus concept).

**Lemma 1.2**

1. If \( v_p = (v_1, v_2, v_3) \) is a tangent vector to \( \mathbb{R}^3 \), then

\[
v_p[f] = \sum_{i=1}^{3} v_i \frac{\partial f}{\partial x_i}(p).
\]

**Proof.** Let \( p = (p_1, p_2, p_3) \); then

\[
p + tv = (p_1 + tv_1, p_2 + tv_2, p_3 + tv_3).
\]

We then

\[
v_p[f] = \left. \frac{d}{dt}(f(p + tv)) \right|_{t=0} = \sum_{i} \frac{\partial f}{\partial x_i}(p)v_i.
\]
1.3 Directional Derivatives

Example 1.1 Let \( f(x, y, z) = x^2yz \). Let \( p = (1, 1, 0) \) and \( v = (1, 0, -3) \). Then

\[
\frac{\partial f}{\partial x} = 2xyz, \quad \frac{\partial f}{\partial y} = x^2z, \quad \frac{\partial f}{\partial z} = x^2y.
\]

Thus

\[
\frac{\partial f}{\partial x}(p) = 0, \quad \frac{\partial f}{\partial y}(p) = 0, \quad \frac{\partial f}{\partial z}(p) = 1.
\]

Then using Lemma ??, we have

\[
v_p[f] = 0 + 0 + 1 \cdot (-3) = -3.
\]

Theorem 1.1

2. Let \( f \) and \( g \) be functions on \( \mathbb{R}^3 \), \( v_p \) and \( w_p \) tangent vectors, \( a \) and \( b \) numbers. Then

- \((av_p + bw_p)[f] = av_p[f] + bw_p[f]\);
- \(v_p(af + bg) = av_p[f] + bv_p[g]\);
- \(v_p[fg] = v_p[f] \cdot g(p) + f(p) \cdot v_p[g]\).

Proof. Only the 3rd equation is new, which can be proved using Lemma ??: We have

\[
v_p[fg] = v_i \frac{\partial (fg)}{\partial x_i}(p) = v_i f(p) \frac{\partial g}{\partial x_i}(p) + v_i g(p) \frac{\partial f}{\partial x_i}(p).
\]

By definition, we have

\[
v_p[fg] = v_p[f] \cdot g(p) + f(p) \cdot v_p[g].
\]

Note The space of tangent vectors is a 3-dimensional space.

Note It would be interesting to define the germ of a function at a point \( p \). Let \( U, V \) be two neighborhoods of \( p \). Let \( f \in C^\infty(U) \), and let \( g \in C^\infty(V) \). We write \( f \sim g \), or say \( f \) is equivalent to \( g \), if there exists a neighborhood \( W \) of \( p \) such that \( W \subset U \cap V \), and

\[
f|_W = g|_W.
\]

The equivalence class of \( f \) is called the germ of \( f \) at \( p \), and is denoted by \( f_p \).

The algebra of all germs of functions defined near \( p \) is called the stalk at \( p \), denoted as \( F_p \).
1.4 Curves in \( \mathbb{R}^3 \)

One of the fundamental questions in curve theory is: how to define a curve? In Euclidean geometry, only two kinds of curves are studied: straight line and circle. In analytic geometry, we study parabola, ellipse, and hyperbola. These curves have quite explicit geometric meanings. For example, an ellipse is the the set of all points in a plane such that the sum of the distances from two fixed points (foci) is constant. If we want to study more general curves, we should not expect them have clear geometric meanings.

In differential geometry, we define a curve in \( \mathbb{R}^3 \) by a function

\[
\alpha : I \rightarrow \mathbb{R}^3, \quad \alpha(t) = (\alpha_1(t), \alpha_2(t), \alpha_3(t)),
\]

where \( I \) is an open interval. In order to use calculus, we usually assume that such a function is smooth.

**Definition 1.9**

A curve in \( \mathbb{R}^3 \) is a differentiable function \( \alpha : I \rightarrow \mathbb{R}^3 \) from an open interval into \( \mathbb{R}^3 \).

We shall give a couple of examples of curves.

**Example 1.2** (Straight Line) A straight line can be expressed best using the vector notations. Let \( p, q \) be two vectors and let \( q \neq 0 \). Then we can use

\[
\alpha(t) = p + tq
\]

to represent a curve with direction \( q \).

**Example 1.3** (Helix) The parameter equations for a circle (in \( \mathbb{R}^3 \)) can be expressed by

\[
t \mapsto (a \cos t, a \sin t, 0).
\]

If we allow this curve to rise , then we obtain a helix \( \alpha : \mathbb{R} \rightarrow \mathbb{R}^3 \), given by the formula

\[
\alpha(t) = (a \cos t, a \sin t, bt),
\]

where \( a > 0 \) and \( b \neq 0 \).

**Definition 1.10**

Let \( \alpha : I \rightarrow \mathbb{R}^3 \) be a curve in \( \mathbb{R}^3 \) with \( \alpha = (\alpha_1, \alpha_2, \alpha_3) \). For each number \( t \in I \), the velocity vector of \( \alpha \) at \( t \) is the tangent vector

\[
\alpha'(t) = \left( \frac{d\alpha_1}{dt}(t), \frac{d\alpha_2}{dt}(t), \frac{d\alpha_3}{dt}(t) \right)_{\alpha(t)}
\]

at the point \( \alpha(t) \) in \( \mathbb{R}^3 \).
Example 1.4 For the helix, \[ \alpha(t) = (a \cos t, a \sin t, bt), \]
the velocity vector is \[ \alpha'(t) = (-a \sin t, a \cos t, b)_{\alpha(t)}. \]

Definition 1.11

Let \( \alpha : I \to \mathbb{R}^3 \) be a curve. If \( h : J \to I \) is an invertible differentiable function on an open interval \( J \), then the composition function \( \beta = \alpha(h) : J \to \mathbb{R}^3 \) is a curve called a reparametrization of \( \alpha \) by \( h \).

Note

The above definition is a key concept. See the next lemma.

Lemma 1.3

If \( \beta \) is the reparametrization of \( \alpha \) by \( h \), then
\[ \beta'(s) = \frac{dh}{ds} \cdot \alpha'(h(s)). \]

Proof. This is a straightforward application of the chain rule.

Lemma 1.4

Let \( \alpha \) be a curve in \( \mathbb{R}^3 \) and let \( f \) be a differentiable function on \( \mathbb{R}^3 \). Then
\[ \alpha'(t)[f] = \frac{d(f(\alpha))}{dt}(t). \]

We end up this section by a discussion of dual space.

1.5 Dual Space

Definition 1.12

Given any vector space \( V \), the dual space \( V^* \) is defined as the set of all linear transformations \( \varphi : V \to \mathbb{R} \).
The dual space is a vector space by the following definition of addition and scalar multiplication.

\[(\varphi + \psi)(x) = \varphi(x) + \psi(x)\]

\[(a\varphi)(x) = a(\varphi(x))\]

for all \(\varphi, \psi \in V^*\) and \(a \in \mathbb{R}, x \in V\). Elements of \(V^*\) is called a covector, or linear functional.

**Example 1.5** On \(\mathbb{R}^n\), any linear function

\[\ell(x) = c_1x_1 + \cdots + c_nx_n,\]

where \(x = (x_1, \cdots, x_n)\) and \(c_1, \cdots, c_n\) being real numbers, is a linear functional.

**Example 1.6** Let \(C([0, 1])\) be the vector space of continuous functions over \([0, 1]\). Then

\[f \mapsto \int_0^1 f(x)dx\]

defines a linear functional.

**Example 1.7** Let \(p \in \mathbb{R}^3\). Let \(v_p\) be a vector on \(T_p(\mathbb{R}^3)\). Then the directional derivative

\[f \mapsto v_p[f]\]

is a linear functional on the vector space of differentiable functions.

**Remark** In order words, every linear functional can be represented, through a fixed inner product, as an element of the vector space.

**Remark** We elaborate the Riesz Representation Theorem in the context of the vector space \(\mathbb{R}^n\) with the dot product. Let \(x \in \mathbb{R}^n\), and let \(\ell\) be a linear functional. By the above theorem, there is a vector \(c\) such that

\[\ell(x) = c \cdot x = c_1x_1 + \cdots + c_nx_n.\]

In this way, the dual space of \(\mathbb{R}^n\) can be identified to \(\mathbb{R}^n\).

**Note** There is an infinite dimensional version of the Riesz Representation Theorem on normed vector space, but the linear functional in question needs to be replaced by bounded linear functional.

**External Link.** The linear algebra over infinite dimensional vector spaces is called Functional Analysis.

### 1.6 1-forms

**Definition 1.13**

A 1-form \(\phi\) on \(\mathbb{R}^3\) is a smooth real-valued function on the set of all tangent vectors to \(\mathbb{R}^3\) such that \(\phi\) is linear at each point, that is,

\[\phi(a\mathbf{v} + b\mathbf{w}) = a\phi(\mathbf{v}) + b\phi(\mathbf{w})\]

for any number \(a, b\) and tangent vectors \(\mathbf{v}, \mathbf{w}\) at the same point of \(\mathbb{R}^3\).
Recall in Definition 1.6, a vector on $\mathbb{R}^3$ is a pair $(p, v)$, where $p \in \mathbb{R}^3$ and $v$ is the vector part.

As before, the space of 1-forms is a module over the algebra of smooth functions. Let $\varphi, \psi$ be two 1-forms; let $v$ be a vector on $\mathbb{R}^3$; let $\lambda \in \mathbb{R}$. Then we can define
\[
(\varphi + \psi)(v) = \varphi(v) + \psi(v); \quad (\lambda \varphi)(v) = \lambda \varphi(v).
\]
Note $\varphi(v)$ is a smooth function on $\mathbb{R}^3$. Let $f$ be a smooth function. Let $v_p$ be the vector part of $v$.
\[
(f \varphi)(v_p) = f(p) \varphi(v_p).
\]
In fact, there is a natural way to extend a 1-form as a function over vector fields. A 1-form is a linear functional in two ways: first, it is a linear functional over the vector space of vector fields, that is, if $\varphi$ is a 1-form, for any vector field $v$, $(\varphi(v))(p) = \varphi(v_p)$ is a smooth function; second, for any fixed point $p$, $\varphi$ is a linear functional over $T_p(\mathbb{R}^3)$.

**Note** At the risk to make the definition more confusing, I would try to elaborate the definition of 1-form a little more. A 1-form is a function $p \mapsto \phi_p$ whose domain is $\mathbb{R}^3$, and whose codomain is the union of all linear functionals over all tangent spaces $T_p(\mathbb{R}^3)$. This is a new kind of function which we need to define the smoothness of it. Our definition here is that $p \mapsto \phi_p$ is smooth if and only if for any vector field $v$, the function
\[
\mathbb{R}^3 \to \mathbb{R}, \quad p \mapsto \phi_p(v)
\]
is a smooth function.

**Definition 1.14**

If $f$ is a differentiable function on $\mathbb{R}^3$. Then $df$ is a 1-form defined by
\[
df(v_p) = v_p[f]
\]
for any $p \in \mathbb{R}^3$.

**Example 1.8** 1-forms on $\mathbb{R}^3$: by the above definition, we can define 1-forms $dx_1, dx_2, dx_3$. Let
\[
v_p = v_i U_i.
\]
Then by definition,
\[
dx_i[v_p] = v_p[x_i] = v_i.
\]
Let’s consider the 1-form
\[
\psi = f_i dx_i,
\]
where $f_i$ are functions. Then
\[
\psi[v_p] = f_i dx_i[v_p] = f_i(p) v_i.
\]

**Definition 1.15**

Let $V$ be a vector space and let $(e_1, \cdots, e_n)$ be a basis of $V$. $(f_1, \cdots, f_n)$ is called the dual basis of $(e_1, \cdots, e_n)$, if $f_i \in V^*$, and if
\[
f_i(e_j) = \delta_{ij}.
\]

**Theorem 1.3**

Using the above notations, then $(dx_1, dx_2, dx_3)$ is the dual basis of $(U_1, U_2, U_3)$.
1.6 1-forms

**Proof.** We have
\[ dx_i(U_j) = U_j[x_i] = \frac{\partial x_i}{\partial x_j} = \delta_{ij}. \]

**Corollary 1.1**

If \( f \) is a differentiable function on \( \mathbb{R}^3 \), then
\[ df = \frac{\partial f}{\partial x_i} dx_i = f_i dx_i. \]

**Note** As we have seen, we use \( f_i \) to represent \( \frac{\partial f}{\partial x_i} \). In differential geometry, this would greatly simplify complicated computations. In general, whether \( f_i \) is \( \frac{\partial f}{\partial x_i} \) or an arbitrary function depends on the context.

**Proof.** By definition, \( df[v_p] = v_p[f] \). But \( v_p[f] = v_i f_i(p) = (f_i dx_i)[v_p] \).

From the definition of \( df \), we observed that we can regard \( d \) as an operator, that would send a function \( f \) to a 1-form \( df \). Such an operator is called a differential operator, which plays one of the center role in differential geometry.

**Lemma 1.5**

1.5 Let \( fg \) be the product of differentiable functions \( f \) and \( g \) on \( \mathbb{R}^3 \). Then
\[ d(fg) = g df + f dg. \]

**Proof.** We have
\[ d(fg) = (fg) dx_i = (gf_i + fg_i) dx_i = g f_i dx_i + f g_i dx_i = g df + f dg. \]

**Lemma 1.6**

Let \( f \) be a function on \( \mathbb{R}^3 \) and let \( h : \mathbb{R} \rightarrow \mathbb{R} \) be a function of single variable. Then
\[ d(h(f)) = h'(f) df. \]
Proof.

\[ d(h(f)) = (h(f))_i dx_i = h'(f) f_i dx_i = h'(f) df. \]

Example 1.9 Let \( f \) be the function

\[ f(x, y) = (x^2 - 1)y + (y^2 + 2)z. \]

Then

\[ df = 2xy dx + (x^2 + 2yz - 1) dy + (y^2 + 2) dz. \]

As a result, we have

\[ df[v] = 2xy v_1 + (x^2 + 2yz - 1) v_2 + (y^2 + 2) v_3. \]

Thus

\[ df[v_p] = 2p_1 p_2 v_1 + (p_1^2 + 2p_2 p_3 - 1) v_2 + (p_3^2 + 2) v_3. \]

We also have

\[ v_p[f] = 2p_1 p_2 v_1 + (p_1^2 + 2p_2 p_3 - 1) v_2 + (p_3^2 + 2) v_3. \]

This verifies \( df[v_p] = v_p[f] \).

1.7 Differential Forms

The space of differential 1-forms is a vector space, or more precisely, it is a module over the algebra of smooth functions. To get more information from tangent spaces where the space of differential 1-forms are dual spaces of them at each point, we shall define multiplication of differential 1-forms. Since all differential 1-forms are generated by \( dx_1, dx_2, dx_3 \), we just need to define their multiplications.

What is \( dx_i dx_j \), or we called the wedge product \( dx_i \wedge dx_j \) of them? We don’t know at this moment. But we shall assume that

\[ dx_i dx_j = -dx_j dx_i \]

for \( 1 \leq i, j \leq 3 \). Obviously, this would create a new kind of algebra. For multiplication of real numbers, we have commutativity, which means, for any two real numbers \( a, b \), we have \( ab = ba \). On the other hand, for two \( n \times n \) matrices \( A, B \), in general, we have \( AB \neq BA \). The property for the multiplication of 1-forms are different from both of the above two. It is called skew commutativity. The algebra defined by the skew commutativity leads to the so-called exterior algebra.
A first observation on the definition of the wedge product reveals that, since \( dx_i \wedge dx_i = -dx_i \wedge dx_i \), we must have \( dx_i \wedge dx_i = 0 \). A quick counting shows that the only non-zero independent products would be \( dx_1 dx_2, dx_1 dx_3 \) and \( dx_2 dx_3 \).

In general, we can define the whole system of \( p \)-forms. We have already encountered 0-forms, which are smooth functions, and 1-forms. Taking multiplication of \( dx_i \) with \( dx_j \), we can define the space of two forms to be generated by \( dx_1 dx_2, dx_1 dx_3 \) and \( dx_2 dx_3 \) over smooth functions, that is, all two forms can be expressed by

\[
fx_1 dx_2 + gdx_1 dx_3 + hdx_2 dx_3,
\]

where \( f, g, h \) are functions.

We can define the 3-forms in an obvious way: all three forms have the expressions

\[
fx_1 dx_2 dx_3,
\]

where \( f \) is a function.

In the high dimensional case, we can define the \( p \)-forms for \( p > 3 \). However, on \( \mathbb{R}^3 \), all higher differential forms would be zero: consider, for example, a 4-form \( dx_i dx_j dx_k dx_l \). Since the space is of 3 dimensional, at least two of the indices must be the same. By skew commutativity, all 4-forms must be zero.

**Example 1.10** Compute the Wedge products

1. Let \( \phi = x\,dx - y\,dy, \quad \psi = z\,dx + x\,dz. \)

Then

\[
\phi \wedge \psi = (x\,dx - y\,dy) \wedge (z\,dx + x\,dz) = xzdx\,dx + x^2 dx\,dz - yzdy\,dx - xydy\,dz = yzdx\,dy + x^2 dx\,dz - xydy\,dz.
\]

2. Let \( \theta = z\,dy. \) Then

\[
\theta \wedge \phi \wedge \psi = -x^2 z\,dx\,dy\,dz.
\]

3. Let \( \eta = y\,dx\,dz + x\,dy\,dz. \) Then

\[
\phi \wedge \eta = (x\,dx - y\,dy) \wedge (y\,dx\,dz + x\,dy\,dz) = (x^2 + y^2) dx\,dy\,dz.
\]

Note It should be clear from these examples that the wedge product of a \( p \)-form and a \( q \)-form is a \((p + q)\)-form. Thus such a product is automatically zero whenever \( p + q > 3 \).

**Lemma 1.7**

Let \( \phi, \psi \) be 1-forms. Then

\[
\phi \wedge \psi = -\psi \wedge \phi.
\]

**Proof.** Let \( \phi = f_i dx_i, \psi = g_i dx_i. \) Then

\[
\phi \wedge \psi = f_i dx_i g_j dx_j = f_i g_j dx_i dx_j = -f_i g_j dx_j dx_i = -\psi \wedge \phi.
\]

**Remark** The space of any \( p \)-forms forms a module over smooth functions\(^1\). However, given that a \( p \)-form wedge

\(^1\)For any \( p \), even if \( p > 3 \) or \( p < 0 \), where we defined the module to be 0.
a $q$-form to be a $(p + q)$-form, we can take the direct sum of the modules of all $p$-forms. Obviously, this would give us a module over functions where the wedge product is well defined.

In what follows we will define arguably the most important concept in differential geometry.

**Definition 1.16**

If $\phi = f_i dx_i$ is a 1-form on $\mathbb{R}^3$. The exterior derivative, or differential, of $\phi$ is the 2-form 

$$d\phi = df_i \wedge dx_i.$$
Example 1.11 If
\[ \phi = f_1 dx_1 = f_1 dx_1 + f_2 dx_2 + f_3 dx_3, \]
then we have
\[ d\phi = \left( \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) dx_1 dx_2 + \left( \frac{\partial f_3}{\partial x_1} - \frac{\partial f_1}{\partial x_3} \right) dx_1 dx_3 + \left( \frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right) dx_2 dx_3. \]
Thus if we identify \( \phi \) to a vector-valued function \( E = (f_1, f_2, f_3) \), then \( d\phi \) can be identified as \( \text{curl}(E) \). In this sense, \( d\phi \) generalize the \text{curl} operator.

Theorem 1.4

Let \( f, g \) be functions and \( \phi, \psi \) be 1-forms. Then

1. \( d(fg) = df \cdot g + f \cdot dg \);
2. \( d(f\phi) = df \wedge \phi + f \cdot d\phi \);
3. \( d(\phi \wedge \psi) = d\phi \wedge \psi - \phi \wedge d\psi \).

This is more a definition than a property of the differential operator.

Proof. Property (1) is just the product rule. To prove (2), we let \( \phi = f_1 dx_1 \). Then
\[ d(f\phi) = d(ff_1) \wedge dx_1 = df \wedge f_1 dx_1 + f df_1 \wedge dx_1 = df \wedge \phi + f \cdot d\phi. \]
Property (3) is, straightly speaking, a definition rather than a property, since we have never defined the differential of 2-forms before. Nevertheless, let’s work on it. First,
\[ d(\phi \wedge \psi) = d(f_1 g_1 dx_1 dx_2). \]
As in the case of 1-forms, we define
\[ d(f_1 g_1 dx_1 dx_2) = d(f_1 g_1) \wedge dx_1 \wedge x_2. \]
We then have
\[ d(\phi \wedge \psi) = df_1 g_1 \wedge dx_1 \wedge dx_2 + f_1 dg_1 \wedge dx_1 \wedge dx_2 = df_1 \wedge dx_1 \wedge g_1 dx_2 - f_1 \wedge dx_1 \wedge dg_1 \wedge dx_2 = d\phi \wedge \psi - \phi \wedge d\psi. \]

Example 1.12 Let
\[ \phi = f_1 dx_2 dx_3 + f_2 dx_3 dx_1 + f_3 dx_1 dx_2. \]
be a 2-form. Then

\[ d\phi = \left( \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3} \right) dx_1 dx_2 dx_3. \]

Thus if we identify \( E = (f_1, f_2, f_3) \). Then \( d\phi \) can be identify to \( \text{div}(E) \).

**Example 1.13** Let \( f \) be a function, then

\[ df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 \]

can be identified as \( \nabla f \).

**Example 1.14** If we identify \( \phi = f_i dx_i \) to \( u \) and \( \psi = g_i dx_i \) to \( v \), then \( \phi \wedge \psi \) can be identified to \( u \times v \).

**Exercise 1.1** Can you use exterior algebra to define the dot product?
1.8 Mappings

In this section we discuss functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. If $n = 3$ and $m = 1$, this is just a function on $\mathbb{R}^3$. In the other extreme, if $n = 1$ and $m = 3$, then this is a single variable $\mathbb{R}^3$-valued function, and by the previous sections, they can be used to represent curves in $\mathbb{R}^3$. All of these functions have been studied in Calculus, but in this section, we shall study them using the idea of linearization.
Recall that a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$.

**Definition 1.17**

Given a function $F : \mathbb{R}^n \to \mathbb{R}^m$, let $f_1, \cdots, f_m$ denote the real-valued function on $\mathbb{R}^n$ such that

$$F(p) = (f_1(p), f_2(p), \cdots, f_m(p))$$

for all points $p \in \mathbb{R}^n$. These functions are called the **Euclidean coordinate functions** of $F$, and we can write $F = (f_1, \cdots, f_m)$.

The functions $F$ is **differentiable** provided its coordinate functions are differentiable in the usual sense. A differentiable function $F : \mathbb{R}^n \to \mathbb{R}^m$ is called a **mapping** from $\mathbb{R}^n$ to $\mathbb{R}^m$.

**Definition 1.18**

Let $I$ be an interval. If $\alpha : I \to \mathbb{R}^n$ is a curve in $\mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^m$ is a mapping, then the composition function $\beta = F(\alpha) : I \to \mathbb{R}^m$ is a curve in $\mathbb{R}^m$ called the **image** of $\alpha$ under $F$.

**Note** Let $B$ be the set of all curves in $\mathbb{R}^n$ and let $C$ be the set of all curves in $\mathbb{R}^m$. Then a mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ induces a map $B \to C$. 
Definition 1.19

Let $F: \mathbb{R}^n \to \mathbb{R}^m$ be a mapping. If $v$ is tangent vector to $\mathbb{R}^n$ at $p$. Let $F_\ast (v)$ be the initial velocity of the curve $t \mapsto F(p + tv)$. The resulting function $F_\ast$ sends a tangent vectors to $\mathbb{R}^n$ to tangent vectors to $\mathbb{R}^m$, and is called the tangent map of $F$.

"If $v = 0$, the straight line $p + tv$ is degenerated to a point $p$. But the definition is still valid."

Proposition 1.1

Let $F = (f_1, \cdots, f_m)$ be a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. If $v$ is a tangent vector to $\mathbb{R}^n$ at $p$, then

$$F_\ast (v) = (v[f_1], \cdots, v[f_m])$$

at $F(p)$.

Proof. We take $m = 3n = 1$ for simplicity. By definition, the curve $t \mapsto F(p + tv)$ can be written as

$$\beta(t) = F(p + tv) = (f_1(p + tv), f_2(p + tv), f_3(p + tv)).$$

By definition, we have $F_\ast(v) = \beta'(0)$. To get $\beta'(0)$, we take the derivatives, at $t = 0$, of the coordinate functions of $\beta$. But

$$\frac{d}{dt}(f_i(p + tv))|_{t=0} = v[f_i].$$

Thus

$$F_\ast (v) = (v[f_1], v[f_2], v[f_3])|_{\beta(0)},$$

where $\beta(0) = F(p)$.
Let \( p \in \mathbb{R}^n \). Then we have the linear transformation

\[ F_*p : T_p(\mathbb{R}^n) \to T_{F(p)}(\mathbb{R}^m) \]

called the tangent map of \( F \) at \( p \).

**Corollary 1.2**

If \( F : \mathbb{R}^n \to \mathbb{R}^m \) is a mapping, then at each point \( p \) of \( \mathbb{R}^n \), the tangent map \( F_*p : T_p(\mathbb{R}^n) \to T_{F(p)}(\mathbb{R}^m) \) is a linear transformation.
Note For any nonlinear function $F$, we can define a semi-linear function $F_\ast p$, where for fixed $p$, the function is a linear transformation. But the function $F_\ast : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$, $(p, v) \mapsto F_\ast p(v)$ is nonlinear with respect to $p$.

Note Let $f(t)$ be a function of single variable. Then $f_\ast : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(t, s) \mapsto sf'(t)$ is the tangent map of $f$. Such a tangent map can be identified to the derivative $f'(t)$.

Corollary 1.3

Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a mapping. If $\beta = F(\alpha)$ is the image of a curve $\alpha$ in $\mathbb{R}^n$, then $\beta' = F_\ast (\alpha')$.

Proposition 1.2

If $F = (f_1, \cdots, f_m)$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$, then

$$F_\ast (U_j(p)) = \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(p) \bar{U}_i(F(p)),$$

where $\{\bar{U}_i\}$, for $i = 1, \cdots, m$ are natural frame fields of $\mathbb{R}^m$. 


The matrix
\[ J = \left( \frac{\partial f_i}{\partial x_j} (p) \right)_{1 \leq i \leq m, 1 \leq j \leq n} \]
is called the Jacobian matrix of $F$ at $p$.

In terms of matrix notations, we have
\[ F_* \begin{bmatrix} U_1(p), \ldots, U_n(p) \end{bmatrix} = \begin{bmatrix} \bar{U}_1(F(p)), \ldots, \bar{U}_m(F(p)) \end{bmatrix} \cdot J. \]
Definition 1.21

A mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ is regular provided that at every point $p$ of $\mathbb{R}^n$ the tangent map $F_p$ is one-to-one.

Remark

By linear algebra, the following are equivalent

1. $F_p$ is one-to-one.
2. $F_p(v_p) = 0$ implies $v_p = 0$.
3. The Jacobian matrix of $F$ at $p$ has rank $n$, the dimension of the domain $\mathbb{R}^n$ of $F$.

Remark

If $m = n$, then we know that, by the Invertible Matrix Theorem, that $F_p$ is one-to-one if and only if it is onto.

Definition 1.22

Let $\mathcal{U}, \mathcal{V}$ be two open sets of $\mathbb{R}^n$. We say that $\mathcal{U}$ and $\mathcal{V}$ are diffeomorphic, if there is a differentiable map $F : \mathcal{U} \to \mathcal{V}$ which is one-to-one and onto. Moreover, the inverse mapping: $F^{-1} : \mathcal{V} \to \mathcal{U}$ is also differentiable. We also say that $F$ is a diffeomorphism of $\mathcal{U}$ to $\mathcal{V}$.

Theorem 1.5

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a mapping between Euclidean spaces of the same dimension. If $F_p$ is one-to-one at a point $p$, there is an open set $\mathcal{U}$ containing $p$ such that $F$ restricted to $\mathcal{U}$ is a diffeomorphism of $\mathcal{U}$ onto an open set $\mathcal{V}$.

External Link.

In the video here, I further elaborate the Chain Rule using the tangent map.
Chapter 2  Frame Fields

2.1 Dot Product

We have discussed inner product in Chapter 0. The dot product is a special case of inner product. So we shall very quickly go through it.

**Definition 2.1**

Let \( \mathbf{p} = (p_1, p_2, p_3) \) and \( \mathbf{q} = (q_1, q_2, q_3) \) in \( \mathbb{R}^3 \). The **dot product** is defined by

\[
\mathbf{p} \cdot \mathbf{q} = p_1 q_1 + p_2 q_2 + p_3 q_3.
\]

The **norm** of a point \( \mathbf{p} \) is defined by

\[
\| \mathbf{p} \| = \sqrt{\mathbf{p} \cdot \mathbf{p}}.
\]

The **Euclidean distance** from \( \mathbf{p} \) to \( \mathbf{q} \) is the number

\[
d(\mathbf{p}, \mathbf{q}) = \| \mathbf{p} - \mathbf{q} \|.
\]

Vectors are called **orthogonal**, if

\[
\mathbf{p} \cdot \mathbf{q} = 0.
\]

More generally, the **angle** \( \theta \) between vectors \( \mathbf{p}, \mathbf{q} \) is defined by the equation

\[
\mathbf{p} \cdot \mathbf{q} = \| \mathbf{p} \| \cdot \| \mathbf{q} \| \cdot \cos \theta.
\]

A vector \( \mathbf{p} \) is called a **unit vector**, if \( \| \mathbf{p} \| = 1 \).

**Definition 2.2**

A set \( \{ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \} \) of three mutually orthogonal unit vectors tangent to \( \mathbb{R}^3 \) at \( \mathbf{p} \) is called a **frame**\(^a\) at the point \( \mathbf{p} \).

\(^a\)It is also called an orthonormal basis.

Thus \( \{ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \} \) is a frame if and only if

\[
\mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{e}_2 \cdot \mathbf{e}_2 = \mathbf{e}_3 \cdot \mathbf{e}_3 = 1; \quad \mathbf{e}_1 \cdot \mathbf{e}_2 = \mathbf{e}_1 \cdot \mathbf{e}_3 = \mathbf{e}_2 \cdot \mathbf{e}_3 = 0.
\]

Using the Einstein Convention, we have \( \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \).

**Definition 2.3**

Let \( \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \) be a frame at a point \( \mathbf{p} \) of \( \mathbb{R}^3 \). The \( 3 \times 3 \) matrix \( A \) whose rows are the Euclidean coordinates of these three vectors is called the **attitude matrix** of the frame.
2.2 Curves

Explicitly, if

\[ e_1 = (a_{11}, a_{12}, a_{13})_p, \]
\[ e_2 = (a_{21}, a_{22}, a_{23})_p, \]
\[ e_3 = (a_{31}, a_{32}, a_{33})_p, \]

then

\[
A = \begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{bmatrix}
\]

We shall prove that \( A \) is an orthogonal matrix: we have

\[ \delta_{ij} = e_i \cdot e_j = a_{ik} a_{jk}. \]

The matrix expression of the above equation is

\[ A A^T = I, \]

where \( A^T \) is the transpose of \( A \).

\[ \medbox{Note} \] If \( A A^T = I \), then \( A^T = A^{-1} \). Therefore

\[ A^T A = A^{-1} A = I. \]

Note that this is not a trivial result. Take a \( 2 \times 2 \) matrix, for example, Let

\[
A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}
\]

Then \( A A^T = I \) is equivalent to

\[ a^2 + b^2 = 1, \quad ac + bd = 0, \quad c^2 + d^2 = 1, \]

and

\[ A^T A = I \]

is equivalent to

\[ a^2 + c^2 = 1, \quad ab + cd = 0, \quad b^2 + d^2 = 1. \]

Can you find an elementary proof of the above fact? Hint: we can prove the identity

\[
(a^2 + c^2 - 1)^2 + (b^2 + d^2 - 1)^2 + 2(ab + cd)^2
= (a^2 + b^2 - 1)^2 + (c^2 + d^2 - 1)^2 + 2(ac + bd)^2.
\]

2.2 Curves

Let \( \alpha(t) = (\alpha_1(t), \alpha_2(t), \alpha_3(t)) \) be a curve. It is well known that

\[ \alpha'(t) = (\alpha'_1(t), \alpha'_2(t), \alpha'_3(t)) \]

is the velocity of the curve. The velocity is a vector field. Its norm, \( ||\alpha'(t)|| \), is called the speed of the curve.

The speed of a curve is a function along the curve.

In terms of the components, we can write

\[ \nu = ||\alpha'|| = \left( \frac{d\alpha_1}{dt} \right)^2 + \left( \frac{d\alpha_2}{dt} \right)^2 + \left( \frac{d\alpha_3}{dt} \right)^2 \right)^{1/2}. \]
2.2 Curves

The length of the curve, from \( t = a \) to \( t = b \), can be expressed by

\[
\int_a^b \| \alpha'(t) \| \, dt.
\]

**Definition 2.4**

A curve \( \alpha : I \to \mathbb{R}^3 \) is called regular, if the velocity vector field \( \alpha'(t) \neq 0 \) for any \( t \in I \). This is equivalently to say that the speed function \( \nu = \| \alpha'(t) \| \) is not zero at any point \( t \in I \). If \( \| \alpha'(t) \| = 1 \), then the curve is called a unit speed curve.

**Definition 2.5**

A reparametrization is a map \( t : [c, d] \to [a, b] \) such that the function is one-to-one and onto, and that \( t'(s) \neq 0 \) for any \( s \in [c, d] \).

**Proposition 2.1**

Let \( t(s) \) be a reparametrization. Then we have the following two cases:

1. \( t'(s) > 0 \) and \( t(c) = a, t(d) = b \);
2. \( t'(s) < 0 \) and \( t(c) = b, t(d) = a \).

**Proof.** Since \( t'(s) \neq 0 \), we must have either \( t'(s) > 0 \) or \( t'(s) < 0 \). In the first case, \( t(s) \) is monotonically increasing. Thus \( t(c) \) must be minimal and hence equal to \( a \), and \( t(d) \) must be maximum, hence equal to \( b \). This proves the first case.

**Theorem 2.1**

The length of a curve is an invariant.

**Proof.** Let \( t = t(s) \) be a reparametrization, that is \( t : [c, d] \to [a, b] \) such that the function \( t \) is one-to-one and onto and we assume that \( t'(s) \neq 0 \). Let \( s = s(t) : [a, b] \to [c, d] \) be the inverse function. Let \( \beta(s) = \alpha(t(s)) \). Then the length of the curve \( \beta \), from \( c \) to \( d \), is

\[
\int_c^d \| \beta'(s) \| \, ds.
\]

By the chain rule, \( \beta'(s) = t'(s)\alpha'(t(s)) \). We then have

\[
\int_c^d \| \beta'(s) \| \, ds = \int_c^d |t'(s)| \cdot \| \alpha'(t(s)) \| \, ds = \int_a^b \| \alpha'(t) \| \, dt.
\]

**Theorem 2.2**

If \( \alpha \) is a regular curve, then there is a parametrization \( \beta \) such that \( \beta \) has unit speed.

**Proof.** Let \( t = t(s) \) be a reparametrization and let \( \beta(s) = \alpha(t(s)) \). The requirement is that

\[
1 = \| \beta'(s) \| = |t'(s)| \cdot \| \alpha'(t(s)) \|
\]
for any \( s \). Thus in order to find the unit speed parametrization, we need to solve the differential equation

\[
t'(s) = \frac{1}{\|\alpha'(t(s))\|}
\]

with the initial value \( t(0) = a \). This is a separable equation. We write \( s' = ds/dt \). Let \( s = s(t) \) be the inverse function of \( t(s) \). Then

\[
s(t) = \int_a^t \|\alpha'(u)\| \, du
\]
defines the unit speed reparametrization.

**Note** Note that \( s(a) = 0 \), and \( s(b) \) is the length of the curve from \( a \) to \( b \).

**Example 2.1** Consider the helix

\[
\alpha(t) = (a \cos t, a \sin t, bt).
\]

The velocity of \( \alpha \) is

\[
\alpha'(t) = (-a \sin t, a \cos t, b).
\]

The speed of \( \alpha \) is

\[
\|\alpha'(t)\| = \sqrt{a^2 \sin^2 t + a^2 \cos^2 t + b^2} = \sqrt{a^2 + b^2}.
\]

Therefore \( \alpha \) has constant speed \( c = \sqrt{a^2 + b^2} \).

Let

\[
s(t) = \int_0^t c \, du = ct.
\]

Then \( t = s/c \). We thus have

\[
\beta(s) = \alpha(s/c) = \left( a \cos \frac{s}{c}, a \sin \frac{s}{c}, \frac{bs}{c} \right)
\]
is the unit speed reparametrization.

**Definition 2.6** Let \( Y = y_i U_i \) be a vector field along a curve \( \alpha \). Thus each function \( y_i \) can be expressed as a function of \( t \) via \( \alpha \). The differentiation of \( Y \) is simply the differentiation on its Euclidean coordinate functions.

**Example 2.2** Let

\[
Y(t) = t^2 U_1 - tU_3.
\]

Then

\[
Y'(t) = 2tU_1 - U_3.
\]

In particular, we can define the acceleration \( \alpha''(t) \) of the curve \( \alpha(t) \).

**Lemma 2.1**

(1) A curve \( \alpha \) is constant if and only if its velocity is zero, \( \alpha' = 0 \).

(2) A non-constant curve \( \alpha \) is a straight line if and only if its acceleration is zero, \( \alpha'' = 0 \).

(3) A vector field \( Y \) on a curve is parallel if and only if its derivative is zero, \( Y' = 0 \).

*This is the definition of the parallelism rather than a statement.*
2.3 The Frenet Formulas

Let $\beta : I \to \mathbb{R}^3$ be a unit-speed curve. Let $T = \beta'(s)$ be the velocity vector field. Then we have $||T|| = 1$. We consider $T' = \beta''(s)$. Since $||T|| = 1$, we have $T \cdot T = 1$. Taking derivative on both sides, we have $T' \cdot T + T \cdot T' = 0$. Thus $T \cdot T' = 0$, and $T'$ is always orthogonal to $T$.

**Definition 2.7**

The curvature of a curve $\beta$ is defined by

$$\kappa(s) = ||T'(s)|| = ||\beta''(s)||.$$

By the above definition, we know that $k(s) \geq 0$. In order to introduce the Frenet formulas, we further assume that $\kappa > 0$.

When $\kappa > 0$, we have $N = T'/\kappa$. By definition, we have $||N|| = 1$. By the similar argument as above, we have $N \cdot N' = 0$.

**Definition 2.8**

Assume that $\kappa(s) > 0$. Define $N = T'/\kappa$, and $B = T \times N$. Then $(T, N, B)$ is an orthonormal basis of the tangent space at point $\beta(s)$. We call $(T, N, B)$ the Frenet frame field, or TNB frame field on $\beta$. The collection $\{T, N, B, \kappa, \tau\}$ is called the Frenet Apparatus.

**Remark** If $\kappa \equiv 0$, then the curve is a straight line. On the other hand, if $\kappa$ is nowhere zero, then we are able to define the Frenet frame field. It is beyond the scope of this book to discuss curves with vanishing curvature at isolated points.

By definition, we have $||N|| = 1$. By the similar argument as above, we have $N \cdot N' = 0$. Thus

**Definition 2.9**

We can define

$$\tau = N' \cdot B,$$

where $\tau$ is called the torsion of the curve.

**Theorem 2.3 (Frenet Formulas)**

If $\beta : I \to \mathbb{R}^3$ is a unit-speed curve with curvature $\kappa > 0$ and torsion $\tau$, then we have the following

---

1. This technique will be used repeatedly throughout the rest of the book.
2.3 The Frenet Formulas

System of ordinary differential equations:

\[
\begin{align*}
T' &= \kappa N \\
N' &= -\kappa T + \tau B \\
B' &= -\tau N
\end{align*}
\]

In matrix notation, we have

\[
(T, N, B)' = (T, N, B) \begin{bmatrix} 0 & -\kappa & 0 \\ \kappa & 0 & -\tau \\ 0 & \tau & 0 \end{bmatrix}.
\]

(2.1)

Proof. The proof is an application of the orthogonality of the frame. The first equation follows by definition. Since \(N \cdot N' = 0\), we can write

\[N' = aT + bB.\]

Thus \(a = (N' - bB) \cdot T = N' \cdot T\). Since \(N \cdot T = 0\), we have \(N' \cdot T + N \cdot T' = 0\). Thus

\[N' \cdot T = -N \cdot T' = -N \cdot \kappa N = -\kappa.\]

By definition, we have

\[b = N' \cdot B = \tau.\]

This proves the second equation.

In order to obtain the third equation, we used the similar method. Write

\[B' = pT + qN + rB.\]

We have

\[p = B' \cdot T = -B \cdot T' = -B \cdot \kappa N = 0,\]

and

\[q = B' \cdot N = -B \cdot N' = -B \cdot (-\kappa T + \tau B) = -\tau,\]

and

\[r = B' \cdot B = 0.\]

This proves the third formula.

Example 2.3 We compute the Frenet frame \(T, N, B\) and the curvature and torsion functions of the unit-speed helix

\[\beta(s) = \left( a \cos \frac{s}{c}, a \sin \frac{s}{c}, \frac{bs}{c} \right),\]

where \(c = (a^2 + b^2)^{1/2}\) and \(a > 0\). Now

\[T(s) = \beta'(s) = \left( -\frac{a}{c} \sin \frac{s}{c}, \frac{a}{c} \cos \frac{s}{c}, \frac{b}{c} \right).\]

Hence

\[T'(s) = \left( -\frac{a}{c^2} \cos \frac{s}{c}, -\frac{a}{c^2} \sin \frac{s}{c}, 0 \right).\]
Thus
\[ \kappa(s) = ||T'(s)|| = \frac{a}{c^2} = \frac{a}{a^2 + b^2} > 0. \]

Since \( T' = \kappa N \), we get
\[ N(s) = \left( -\cos \frac{s}{c}, -\sin \frac{s}{c}, 0 \right). \]

Therefore, we have
\[ B = T \times N = \left( \frac{b}{c} \sin \frac{s}{c}, -\frac{b}{c} \cos \frac{s}{c}, \frac{a}{c} \right). \]

The torsion is given by
\[ \tau = N' \cdot B = \frac{b}{c^2}. \]

In what follows, we use Frenet formulas to study the properties of curves.

**Definition 2.10**

A plane curve in \( \mathbb{R}^3 \) is a curve that lies in a single plane of \( \mathbb{R}^3 \).

**Theorem 2.4**

Let \( \beta \) be a unit-speed curve in \( \mathbb{R}^3 \) with \( \kappa > 0 \). Then \( \beta \) is a plane curve if and only if \( \tau = 0 \).

**Proof.** If a curve \( \beta(s) = (\beta_1(s), \beta_2(s), \beta_3(s)) \) lies in a plane, then there are constants \( a, b, c, d \) such that
\[ a\beta_1 + b\beta_2 + c\beta_3 = d. \]

Using the vector notation, if \( v = (a, b, c) \), then \( v \cdot \beta = d \). As a result, we have
\[ v \cdot \beta'(s) = v \cdot \beta''(s) = v \cdot \beta'''(s) = 0. \]

Since \( \beta' \perp \beta'' \), they are linearly independent. Thus we can write \( \beta''' \) as a linear combination of \( \beta' \) and \( \beta'' \).
\[ \beta''' = p\beta' + q\beta''. \]

where \( p, q \) are constants. Since
\[ B = T \times N = \kappa^{-1} \beta' \times \beta'', \]
we have
\[ \tau = N' \cdot B = (\kappa^{-1} \beta''')' \cdot B = 0. \]

On the other hand, if \( \tau \equiv 0 \), then \( B' = 0 \). Thus \( B \) is a constant vector. Let
\[ f(s) = \beta(s) \cdot B - \beta(0) \cdot B. \]

Then \( f' = 0 \) and since \( f(0) = 0 \), we know that \( f \equiv 0 \). Thus \( \beta(s) \) is on the plane
\[ x \cdot B - \beta(0) \cdot B = 0. \]

In the following, we shall use the Taylor’s formula to study the curve \( \beta \). We have
\[ \beta(s) = \beta(0) + \beta'(0)s + \frac{1}{2} \beta''(0)s^2 + \frac{1}{6} \beta'''(0)s^3 + o(s^3). \]
Apparently, we have
\[ \beta'(0) = T(0), \]
\[ \beta''(0) = \kappa(0)N(0), \]
\[ \beta'''(0) = (\kappa N)' = \kappa'(0)N + \kappa(0)(-\kappa(0)T(0) + \tau(0)B(0)). \]

Therefore, we have
\[ \beta(s) - \beta(0) = sT(0) + \frac{1}{2}s^2\kappa(0)N(0) \]
\[ + \frac{1}{6}s^3(\kappa'(0)N + \kappa(0)(-\kappa(0)T(0) + \tau(0)B(0))) + o(s^3). \]

From the above formula, we know that the Frenet Apparatus completely determined the curve, at least for the first three terms in the Taylor’s expansion.

**Remark** We have emphasized all along the distinction between a tangent vector and a point of \( \mathbb{R}^3 \). However, Euclidean space has, as we have seen, the remarkable property that given a point \( p \), there is a natural one-to-one correspondence between points \((v_1, v_2, v_3)\) and tangent vectors \((v_1, v_2, v_3)_p\) at \( p \). Thus one can transform points into tangent vectors (and vice versa) by means of this canonical isomorphism. In the next two sections particularly, it will often be convenient to switch quietly from one to the other without change of notation. Since corresponding objects have the same Euclidean coordinates, this switching can have no effect on scalar multiplication, addition, dot products, differentiation, or any other operation defined in terms of Euclidean coordinates.

So in (2.2), the left side \( \beta(s) - \beta(0) \) is a vector on \( \mathbb{R}^3 \), and the right hand side, which is a linear combination of \( T(0), N(0) \) and \( B(0) \), is a vector in the tangent space of \( \beta(0) \). We identify \( T_{\beta(0)}\mathbb{R}^3 \) with \( \mathbb{R}^3 \) so the equality makes sense.

### 2.4 Arbitrary-Speed Curves

It is a simple matter to adapt the results of the previous section to the study of a regular curve \( \alpha : I \to \mathbb{R}^3 \) that does not necessarily have unit speed. We merely transfer to \( \alpha \) the Frenet apparatus of a unit-speed reparametrization \( \bar{\alpha} \) of \( \alpha \). Explicitly, if \( s \) is an arc length function for \( \alpha \), then
\[ \alpha(t) = \bar{\alpha}(s(t)) \]
for all \( t \). Or, we can write \( \alpha(t) = \bar{\alpha}(s). \)

Assume that \( \bar{T}, \bar{N}, \bar{B}, \bar{\kappa} > 0, \bar{\tau} \) are the Frenet Apparatus with respect to the unit speed curve \( \bar{\alpha}(s) \). Then we define
- **curvature** function: \( \kappa = \bar{\kappa}(s) \),
- **torsion** function: \( \tau = \bar{\tau}(s) \),
- **unit tangent** vector field: \( T = \bar{T}(s) \),
- **principal** vector field: \( N = \bar{N}(s) \),
- **binomial** vector field: \( B = \bar{B}(s) \).

The speed of the curve is defined by \( \nu = \|\alpha'(t)\| \). We can regard \( \nu \) as a function of \( t \) but through \( t = t(s) \), it can be regarded as a function of \( s \) as well. Since \( \alpha(t) = \bar{\alpha}(s) \), we have \( \alpha'(t) = \bar{\alpha}'(s) \frac{ds}{dt} \). Thus we have
\[ \frac{ds}{dt} = \|\alpha'(t)\| = \nu. \]
2.4 Arbitrary-Speed Curves

Lemma 2.2

Assume that \( ds/dt > 0 \). If \( \alpha \) is a regular curve in \( \mathbb{R}^3 \) with \( \kappa > 0 \), then we have

\[
\begin{align*}
T' &= \kappa \nu N \\
N' &= -\kappa \nu T + \tau \nu B \\
B' &= -\tau \nu N
\end{align*}
\]

*We shall always assume this for the rest of the lecture notes. Thus we have \( \nu = ds/dt \).*

Proof. We shall use the Frenet formulas for unit speed curves. Since \( T(t) = \bar{T}(s) \), \( N(t) = \bar{N}(s) \) and \( B(t) = \bar{B}(s) \), we have \( T'(t) = \bar{T}'(s) \nu, N'(t) = \bar{N}'(s) \nu \), and \( B'(t) = \bar{B}(s) \nu \). The lemma then follows from Theorem 2.3.

Note There is a commonly used notation for the calculus that completely ignores change of parametrization. For example, the same letter would designate both a curve \( \alpha \) and its unit-speed parametrization \( \bar{\alpha} \), and similarly with the Frenet apparatus of these two curves. Differences in derivatives are handled by writing, say, \( dT/dt \) for \( T'(t) \) and \( dT/ds \) for \( \bar{T}'(s) \).

Lemma 2.3

If \( \alpha \) is a regular curve with speed function \( \nu \), then the velocity and acceleration of \( \alpha \) are given by

\[
\begin{align*}
\alpha' &= \nu T, \\
\alpha'' &= \frac{d\nu}{dt} T + \kappa \nu^2 N.
\end{align*}
\]

Proof. The proof is not difficult but the notations are confusing. By the context, \( \alpha' = \alpha'(t) \) and \( \alpha'' = \alpha''(t) \). We thus have

\[
\alpha' = T \frac{ds}{dt} = T \nu,
\]

and

\[
\alpha'' = \frac{d\nu}{dt} T + \nu T' = \frac{d\nu}{dt} T + \kappa \nu^2 N.
\]

Remark The formula \( \alpha' = \nu T \) is to be expected since \( \alpha' \) and \( T \) are each tangent to the curve and \( T \) has a unit length, while \( \|\alpha'\| = \nu \).

The formula for acceleration is more interesting. By definition, \( \alpha'' \) is the rate of change of the and in general both the length and the direction of \( \alpha' \) are changing. The tangential component \( \frac{d\nu}{dt} T \) of \( \alpha'' \) measures the rate
of change of the length of $\alpha'$ (that is, of the speed of $\alpha$). The normal component $\kappa \nu^2 N$ measures the rate of change of the direction of $\alpha'$. Newton’s laws of motion show that these components may be experienced as forces. For example, in a car that is speeding up or slowing down on a straight road, the only force one feels is due to $\frac{d\nu}{dt} T$. If one takes an unbanked curve at speed $\nu$, the resulting sideways force is due to $\kappa \nu^2 N$. Here $\kappa$ measures how sharply the road turns; the effect of speed is given by $\nu^2$, so 60 miles per hour is four times as unsettling as 30.

**Note** Assume that we live in a 1-dimensional space defined by the above curve $\alpha$. We then can only measure the tangential component of the acceleration $\alpha''$. As a 1-dimensional creature, it is not possible for the creature to understand $\kappa$, the curvature of the curve. It would think it lives in a straight line.

### Theorem 2.5

Let $\alpha$ be a regular curve in $\mathbb{R}^3$. Then

- $T = \alpha' / \| \alpha' \|$,  
- $N = B \times T$,  
- $B = \alpha' \times \alpha'' / \| \alpha' \times \alpha'' \|$,  
- $\kappa = \| \alpha' \times \alpha'' / \| \alpha' \| \|^3$,  
- $\tau = (\alpha' \times \alpha'') \cdot \alpha'' / \| \alpha' \times \alpha'' \|^2$. 

**Proof.** The proof is just a matter of applications of the chain rule and the Frenet formulas. But it contains several basic techniques in differential geometry.

First, we have $T = \alpha' / \| \alpha' \|$. Next, using the above Lemma 2.3, we have

$$\alpha' \times \alpha'' = \| \alpha' \| T \times (\nu' T + \kappa \nu^2 N) = \kappa \nu^3 B. \tag{2.3}$$

Taking the norm of the above equation, we prove the formula for $\kappa$ is proved.

That $N = B \times T$ follows from the definition. From Lemma 2.3, we have

$$\alpha''' = (\nu' T)' + (\kappa \nu^2 N)'.$$

From the above, we know that the $B$ components of $\alpha'''$ is $\kappa \tau \nu^3$. Thus we have

$$(\alpha' \times \alpha'') \cdot \alpha''' = \kappa^2 \tau \nu^6.$$

Since $\| \alpha' \times \alpha'' \|^2 = \kappa^2 \nu^6$, the formula for $\tau$ is proved.

**External Link.** As we know, the definition of the curvature is $\kappa = \nu^{-1} \| T'(t) \|$. Can we obtain the above formula by a straightforward computation. Yes, it is complicated, but such kind of computation contains useful techniques in differential geometry. See here for the method.

**Example 2.4** We compute the Frenet apparatus of the 3-curve

$$\alpha(t) = (3t - t^3, 3t^2, 3t + t^3).$$

We have

$$\alpha'(t) = 3(1 - t^2, 2t, 1 + t^2),$$
$$\alpha''(t) = 6(-t, 1, t),$$
$$\alpha'''(t) = 6(-1, 0, 1).$$

First, we have

$$\nu = \| \alpha' \| = \sqrt{\alpha' \cdot \alpha'} = 3\sqrt{2(1 + t^2)}.$$
We also have
\[
\alpha' \times \alpha'' = 18 \begin{vmatrix}
U_1 & U_2 & U_3 \\
1 - t^2 & 2t & 1 + t^2 \\
-t & 1 & t
\end{vmatrix} = 18 (-1 + t^2, -2t, 1 + t^2).
\]

Hence
\[
\|\alpha' \times \alpha''\| = 18 \sqrt{2(1 + t^2)}.
\]

We compute
\[
(\alpha' \times \alpha'') \cdot \alpha''' = 6 \cdot 18 \cdot 2.
\]

It remains only to substitute this data into the formulas in Theorem 2.5 with \(N\) being computed by another cross product. The final results are
\[
T = \left(\frac{1 - t^2, 2t, 1 + t^2}{\sqrt{2(1 + t^2)}}\right),
\]
\[
N = \left(\frac{-2t, 1 - t^2, 0}{1 + t^2}\right),
\]
\[
B = \left(\frac{-1 + t^2, -2t, 1 + t^2}{\sqrt{2(1 + t^2)}}\right),
\]
\[
\kappa = \tau = \frac{1}{3(1 + t^2)^2}.
\]

For the rest of the section, we shall do some applications of the above formulas.

**Definition 2.11**

*The spherical image of a unit-speed curve \(\beta(s)\) is the curve \(\sigma(s) = T(s) = \beta'(s)\).*

Let \(\sigma\) be the spherical image of \(\beta\). Then \(\sigma' = \beta'' = \kappa N\), where \(\kappa\) is the curvature of \(\beta\). That \(\kappa > 0\) ensures that \(\sigma\) is a regular curve. In order to compute the curvature \(\kappa_\sigma\) of \(\sigma\), we compute
\[
\sigma' = \beta'' = \kappa N,
\]
\[
\sigma''' = \beta''' = \kappa' N + \kappa N' = \kappa' N + \kappa (-\kappa T + \tau B).
\]

Thus
\[
\sigma' \times \sigma''' = \kappa^2 (\kappa B + \tau T).
\]

Using the formula in Theorem 2.5, we have
\[
\kappa_\sigma = \frac{\|\sigma' \times \sigma''\|}{\nu^3} = (1 + (\tau/\kappa)^2)^{1/2} \geq 1.
\]

**Definition 2.12**

*A regular curve \(\alpha\) in \(\mathbb{R}^3\) is a cylindrical helix provided the unit tangent vector field \(T\) of \(\alpha\) has constant angle \(\theta\) with some fixed unit vector \(u\); that is, \(T(t) \cdot u = \cos \theta\) for all \(t\).*

**Theorem 2.6**

*A regular curve \(\alpha\) with \(\kappa > 0\) is a cylindrical helix if and only if the ratio \(\tau/\kappa\) is constant.*

\(^2\)If \(\kappa > 0\), then \(\sigma'' \neq 0\). Thus the Frenet Apparatus always exists for \(\sigma\).
2.4 Arbitrary-Speed Curves

**Proof.** First assume that a unit speed curve is a cylindrical helix curve. Then there is a constant vector \( u \) such that \( T \cdot u = c \), a constant. Taking derivative on both sides, we get \( \kappa N \cdot u = 0 \). Since \( \kappa > 0 \), we get \( N \cdot u = 0 \). Therefore, if we write \( u \) as a linear combination of \( T, N, B \), there would be no \( N \) component.

By the assumption, we would get

\[
\mathbf{u} = \cos \theta \mathbf{T} + \sin \theta \mathbf{B}.
\]

Taking derivative of the above equation again, we get

\[
0 = \cos \theta \kappa N - \sin \theta \tau N.
\]

We thus have \( \kappa \cos \theta - \tau \sin \theta = 0 \). Hence \( \tau / \kappa = \cot \theta \) is a constant.

Conversely, if \( \tau / \kappa \) is a constant, we write \( \tau / \kappa = \cot \theta \). Let

\[
\mathbf{u} = \cos \theta \mathbf{T} + \sin \theta \mathbf{B}.
\]

Then

\[
\mathbf{u}' = \cos \theta \kappa N - \sin \theta \tau N = 0.
\]

Therefore \( u \) is a constant vector field. Obviously, \( \|u\| = 1 \). Then \( T \cdot u = \cos \theta \) which means that the curve is a cylindrical helix.

Can we prove the above result by solving the Frenet differential equations? Yes, in the following, we give another proof of the fact that if \( \tau / \kappa = c \) be a constant, then there must be a unit vector \( u \) such that \( T \cdot u \) is a constant.

We use (2.1) to obtain

\[
(T, N, B)' = (T, N, B) \begin{bmatrix}
0 & -\kappa & 0 \\
\kappa & 0 & -\tau \\
0 & \tau & 0
\end{bmatrix}
\begin{bmatrix}
0 & -1 & 0 \\
1 & 0 & -c \\
0 & c & 0
\end{bmatrix}.
\]

Thus we have (See the remark below)

\[
(T, N, B) = (T(0), N(0), B(0))e^{A \int_0^t \kappa(u) \, du},
\]

where

\[
A = \begin{bmatrix}
0 & -1 & 0 \\
1 & 0 & -c \\
0 & c & 0
\end{bmatrix}.
\]

Since \( A \) is a skew-symmetric \( 3 \times 3 \) matrix, it is singular\(^3\). Let \( \mathbf{v} \) be a unit eigenvector of \( A \) with respect to the zero eigenvalue. Let \( \mathbf{u} \) be the vector such that

\[
\mathbf{u} \cdot (T(0), N(0), B(0)) = \mathbf{v}.
\]

Then

\[
T \cdot \mathbf{u} = \mathbf{u} \cdot T = \mathbf{u} \cdot (T(0), N(0), B(0))e^{A \int_0^t \kappa(u) \, du} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.
\]

\(^3\)One can prove this fact by a straightforward computation. Alternatively, we observe that for skew-symmetric matrices, all eigenvalues must be purely imaginary unless they are zero. Since all purely imaginary eigenvalues must be in pairs (conjugate eigenvalues), there must be at least one zero eigenvalue for odd dimensional matrices.
Note that \( v, A = 0 \), we thus have
\[
v e^{\int_0^t \kappa(u) \, du} = v \sum_{k=0}^{\infty} \frac{1}{k!} A^k \left( \int_0^t \kappa(u) \, du \right)^k = v.
\]
Thus
\[
T \cdot u = v \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}
\]
is a constant.

**Remark** Let \( A(t) \) be a matrix-valued function. We consider the system of differential equations
\[
A'(t) = A(t) B(t), \quad A(0) = A_0,
\]
where \( B(t) \) is a matrix-valued function. Then in general, we don’t have a formula for the (unique) solution. If for any \( t, t' \), we have
\[
B(t) B(t') = B(t') B(t),
\]
then we have the solution
\[
A(t) = A_0 e^{\int_0^t B(u) \, du}.
\]
For further explanation, see [here](#).

### 2.5 Covariant Derivatives

In this section, we shall define probably one of the most important concepts in differential geometry, or may be even one of the most important concepts in mathematics: covariant derivatives on vector fields.

**Definition 2.13**

Let \( W \) be a vector field on \( \mathbb{R}^3 \), and let \( v \) be a tangent vector to \( \mathbb{R}^3 \) at the point \( p \). Then the covariant derivative of \( W \) with respect to \( v \) is the tangent vector
\[
\nabla_v W = W(p + tv)'(0)
\]
at the point \( p \).

On surface, the covariant derivative is just another version of directional derivative, especially since we have already defined such kind of derivative along a curve before. But we shall systematically use this kind of derivative under the frame of differential operators.

**Example 2.5** Let \( W = x^2 U_1 + yz U_3 \), and let
\[
v = (-1, 0, 2),
\]
at \( p = (2, 1, 0) \). Then
\[
p + tv = (2 - t, 1, 2t).
\]
So
\[
W(p + tv) = (2 - t)^2 U_1 + 2t U_3.
\]
Hence
\[
\nabla_v W = W(p + tv)'(0) = -4 U_1(p) + 2 U_3(p).
\]
Lemma 2.4

If $W = w_i U_i$ is a vector field on $\mathbb{R}^3$, and $\mathbf{v}$ is a tangent vector at $p$, then

$$\nabla_{\mathbf{v}} W = \mathbf{v}[w_i] U_i(p).$$

Proof. We have

$$W(p + t\mathbf{v}) = w_i(p + t\mathbf{v}) U_i(p + t\mathbf{v}).$$

Differentiation of the above at $t = 0$, we obtain

$$\nabla_{\mathbf{v}} W = W(p + t\mathbf{v})'(0) = \mathbf{v}[w_i] U_i(p).$$

This equation is from the book which is, of course, correct. But in all previous formulas, we use $U_i$ instead of $U_i(p + t\mathbf{v})$.

In short, to apply $\nabla_{\mathbf{v}}$ to a vector field, apply $\mathbf{v}$ to its Euclidean coordinates.

Theorem 2.7

Let $\mathbf{v}$ and $\mathbf{w}$ be tangent vectors $\mathbb{R}^3$ at $p$, and let $Y$ and $Z$ be vector fields on $\mathbb{R}^3$. Then for numbers $a, b$ and function $f$

1. $\nabla_{a\mathbf{v} + b\mathbf{w}} Y = a\nabla_{\mathbf{v}} Y + b\nabla_{\mathbf{w}} Y$,
2. $\nabla_{\mathbf{v}}(aY + bZ) = a\nabla_{\mathbf{v}} Y + b\nabla_{\mathbf{v}} Z$,
3. $\nabla_{\mathbf{v}}(fY) = \mathbf{v}[f] Y(p) + f(p)\nabla_{\mathbf{v}} Y$,
4. $\mathbf{v}[Y \cdot Z] = \nabla_{\mathbf{v}} Y \cdot Z(p) + Y(p) \cdot \nabla_{\mathbf{v}} Z$.

Proof. The proof is straightforward. For example, to prove (3), let $Y = y_i U_i$ and then $fY = (f y_i) U_i$.

Thus

$$\nabla_{\mathbf{v}}(fY) = \mathbf{v}[f y_i] U_i = \mathbf{v}[f] y_i U_i + f \mathbf{v}[y_i] U_i = \mathbf{v}[f] Y(p) + f(p)\nabla_{\mathbf{v}} Y.$$ 

If we regard the set of vector fields as a module over the algebra of smooth functions, then the operator $\nabla_{\mathbf{v}}$ is linear with respect to the addition and scalar multiplication. It is a derivative on the multiplication of smooth functions to vector fields.

Corollary 2.1

Let $V, W, Y$ and $Z$ be vector fields on $\mathbb{R}^3; f, g$ be smooth functions. Then

1. $\nabla_{fV + gW} Y = f\nabla_{V} Y + g\nabla_{W} Y$,
2. $\nabla_{V}(aY + bZ) = a\nabla_{V} Y + b\nabla_{V} Z$,
3. $\nabla_{V}(fY) = V[f] Y + f\nabla_{V} Y$,
4. $V[Y \cdot Z] = \nabla_{V} Y \cdot Z + Y \cdot \nabla_{V} Z$.

Example 2.6 Suppose $V = (y - x)U_1 + xyU_3$ and $W = x^2 U_1 + yz U + 3$. Then

$$V[x^2] = (y - x) U_1[x^2] = 2x(y - x), \quad V[yz] = xy U_3[yz] = xy^2.$$ 

Hence

$$\nabla_{V} W = 2x(y - x) U_1 + xy^2 U_3.$$ 

49
2.6 Frame Fields

When the Frenet formulas were discovered (by Frenet in 1847, and independently by Serret in 1851), the theory of surfaces in $\mathbb{R}^3$ was already a richly developed branch of geometry. The success of the Frenet approach to curves led Darboux (around 1880) to adapt this “method of moving frames” to the study of surfaces. Then it was Cartan who brought the method to full generality. His essential idea was very simple: To each point of the object under study (a curve, a surface, Euclidean space itself, . . . ) assign a frame; then using orthonormal expansion express the rate of change of the frame in terms of the frame itself. This, of course, is just what the Frenet formulas do in the case of a curve.

In the next three sections we shall carry out this scheme for the Euclidean space $\mathbb{R}^3$. We shall see that geometry of curves and surfaces in $\mathbb{R}^3$ is not merely an analogue, but actually a corollary, of these basic results.

**Definition 2.14**

Vector fields $\{E_1, E_2, E_3\}$ on $\mathbb{R}^3$ constitute a frame field on $\mathbb{R}^3$ provided

$$E_i \cdot E_j = \delta_{ij}.$$  

Apparently, $U_1, U_2, U_3$ is a frame field. We call such a frame field **Euclidean frame field**. In the following, we shall introduce two more important frame fields.

**Example 2.7** *(The cylindrical frame field)* Let $(r, \theta, z)$ be the usual cylindrical coordinate functions on $\mathbb{R}^3$. We shall pick a unit vector field in the direction in which each coordinate increases (when the other two are held constant). For $r$, this is evidently

$$E_1 = \cos \theta U_1 + \sin \theta U_2.$$  

Then

$$E_2 = -\sin \theta U_1 + \cos \theta U_2$$

points in the direction of increasing $\theta$ as in Fig. 2.19. Finally, the direction of increase of $z$ is, of course, straight up, so

$$E_3 = U_3.$$  

**Remark** I don’t think in the above example, deducing the cylindrical frame field from the cylindrical coordinates by the picture, is mathematically rigid, or even correct. In fact, as long as the above $\{E_1, E_2, E_3\}$ is an orthonormal basis at each point, it defines a frame field. We can call it the cylindrical frame field. There is no proof needed here.

A better way to show the relationship between the cylindrical coordinates and the cylindrical frame field is to use the chain rule. We have $x = r \cos \theta, y = r \sin \theta, z = z$. Then by the chain rule, for any function $f$, we
2.6 Frame Fields

have
\[
\frac{\partial f}{\partial r} = \frac{\partial x}{\partial r} \frac{\partial f}{\partial x} + \frac{\partial y}{\partial r} \frac{\partial f}{\partial y} + \frac{\partial z}{\partial r} \frac{\partial f}{\partial z}.
\]

In terms of vector notation, this implies that
\[
\frac{\partial}{\partial r} = \frac{\partial x}{\partial r} U_1 + \frac{\partial y}{\partial r} U_2 + \frac{\partial z}{\partial r} U_3 = \cos \theta U_1 + \sin \theta U_2.
\]

Using the same method, we have
\[
\frac{\partial}{\partial z} = U_3.
\]

However,
\[
\frac{\partial}{\partial \theta} = \frac{\partial x}{\partial \theta} U_1 + \frac{\partial y}{\partial \theta} U_2 + \frac{\partial z}{\partial \theta} U_3 = -r \sin \theta U_1 + r \cos \theta U_2
\]

which is not \( E_2 \) but \( r E_2 \).

In general, if \( \{x_1, x_2, x_3\} \) are another set of coordinate functions, then the vector fields
\[
\left\{ \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right\}
\]
do not define a frame field (they may not be orthonormal).

Example 2.8 (The spherical frame field) Now we define the spherical frame field \( \{F_1, F_2, F_3\} \). Let \( \rho, \theta, \varphi \) be the spherical coordinates. We then have
\[
\begin{align*}
x &= \rho \cos \varphi \cos \theta, \\
y &= \rho \cos \varphi \sin \theta, \\
z &= \rho \sin \varphi.
\end{align*}
\]

The spherical frame field is defined by
\[
F_1 = \cos \varphi \left( \cos \theta U_1 + \sin \theta U_2 \right) + \sin \varphi U_3
\]
\[
F_2 = -\sin \theta U_1 + \cos \theta U_2,
\]
\[
F_3 = -\sin \varphi \left( \cos \theta U_1 + \sin \theta U_2 \right) + \cos \varphi U_3.
\]

By a straightforward computation, we have
\[
\begin{align*}
\frac{\partial}{\partial \rho} &= F_1, \\
\frac{\partial}{\partial \theta} &= -\rho \cos \varphi \sin \theta U_1 + \rho \cos \varphi \cos \theta U_2 = \rho \cos \varphi F_2, \\
\frac{\partial}{\partial \varphi} &= -\rho \sin \varphi \cos \theta U_1 - \rho \sin \varphi \sin \theta U_2 + \rho \cos \varphi U_3 = \rho F_3.
\end{align*}
\]
2.7 Connection forms

Lemma 2.5

Let \( \{E_1, E_2, E_3\} \) be a frame field on \( \mathbb{R}^3 \).

1. If \( V \) is a vector field on \( \mathbb{R}^3 \), then \( V = f_i E_i \), where the functions \( f_i = V \cdot E_i \) are called coordinate functions of \( V \) with respect to \( \{E_1, E_2, E_3\} \).

2. If \( V = f_i E_i \) and \( W = g_i E_i \), then \( V \cdot W = f_i g_i \). In particular, \( ||V|| = (\sum f_i^2)^{1/2} \).

Note In differential geometry, we study arbitrary frame field, not concrete ones like the cylindrical or spherical frame fields. We will do so from the next section.

2.7 Connection forms

Let \( \{E_1, E_2, E_3\} \) be a frame field, and let \( v \) be a vector at \( p \in \mathbb{R}^3 \). Then we are able to write

\[
\nabla_v E_1 = c_{11} E_1(p) + c_{12} E_2(p) + c_{13} E_3(p),
\]

\[
\nabla_v E_2 = c_{21} E_1(p) + c_{22} E_2(p) + c_{23} E_3(p),
\]

\[
\nabla_v E_3 = c_{31} E_1(p) + c_{32} E_2(p) + c_{33} E_3(p).
\]

Using the Einstein Convention, we have

\[
\nabla_v E_i = c_{ij} E_j(p).
\]

Since \( \{E_1, E_2, E_3\} \) is a frame, we then have\(^4\)

\[
\]

\[c_{ij} = \nabla_v E_i \cdot E_j(p).\]

We then can define a 1-form \( \omega_{ij} \) such that

\[
\omega_{ij}(v) = c_{ij} = \nabla_v E_i \cdot E_j(p).
\]

Lemma 2.6

Using the above notations, then \( \omega_{ij} \) are 1-forms satisfying

\[
\omega_{ij} = -\omega_{ji}.
\]

These 1-forms are called connection forms of the frame field \( \{E_1, E_2, E_3\} \).

Proof. Let \( v, w \) be vectors and let \( a, b \in \mathbb{R} \). In order to prove \( \omega_{ij} \) is a 1-form, we just need to prove that

\[
\omega_{ij}(av + bw) = a \omega_{ij}(v) + b \omega_{ij}(w),
\]

where \( v, w \) are vectors and \( a, b \) are real numbers. But this follows from

\[
\nabla_{av+bw} E_i = a \nabla_v E_i + b \nabla_w E_i.
\]

To prove \( \omega_{ij} = -\omega_{ji} \), we observe that since \( E_i \cdot E_j = \delta_{ij} \). Then

\[
0 = v [E_i \cdot E_j] = \nabla_v E_i \cdot E_j + E_i \cdot \nabla_v E_j = \omega_{ij}(v) + \omega_{ji}(v).
\]

\(^4\)We know that \( 1 \leq i, j \leq 3 \). But this is implied by the Einstein convention so can be omitted.
2.7 Connection forms

**Theorem 2.8**

Let $\omega_{ij}$ be the connection 1-forms of a frame field $\{E_1, E_2, E_3\}$ on $\mathbb{R}^3$. Then for any vector field $V$ on $\mathbb{R}^3$, we have

$$\nabla_V E_i = \omega_{ij}(V)E_j.$$ 

By $\omega_{ij} = -\omega_{ji}$, we have $\omega_{ii} = 0$. We write

$$\omega = \begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{32} & \omega_{33} \end{bmatrix} = \begin{bmatrix} 0 & \omega_{12} & \omega_{13} \\ -\omega_{12} & 0 & \omega_{23} \\ -\omega_{13} & -\omega_{23} & 0 \end{bmatrix}.$$ 

If we write

$$E = (E_1, E_2, E_3),$$

then in matrix notations, we have

$$\nabla_V E = E \cdot \omega^T.$$ 

This can be used to compare with the Frenet formula

$$(T, N, B)' = (T, N, B) \begin{bmatrix} 0 & -\kappa & 0 \\ \kappa & 0 & -\tau \\ 0 & \tau & 0 \end{bmatrix}.$$ 

Given an arbitrary frame field $\{E_1, E_2, E_3\}$ on $\mathbb{R}^3$, it is fairly easy to find an explicit formula for its connection forms. First use orthonormal expansion to express the vector fields $\{E_1, E_2, E_3\}$ in terms of the natural frame field $\{U_1, U_2, U_3\}$ on $\mathbb{R}^3$. We assume that

$$E_i = a_{ij}U_j.$$ 

Then we have the matrix

$$A = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$ 

The above matrix is called an attitude matrix.

**Theorem 2.9**

If $A = (a_{ij})$ is the attitude matrix of the frame field $\{E_1, E_2, E_3\}$ and if $\omega = (\omega_{ij})$ be the connection matrix of the frame field, then

$$\omega = dA A^T,$$

or equivalently

$$\omega_{ij} = a_{jk} da_{ik}.$$ 

**Proof.** Let $V$ be a vector field, and we have

$$\omega_{ij}[V] = \nabla_V E_i \cdot E_j = (V[a_{ik}]U_k) \cdot (a_{ji}U_l) = V[a_{ik}]a_{ji}\delta_{kl} = da_{ik}[V]a_{jk}.$$ 

The theorem is proved.

It would be nice to also include a proof using the matrix notations.
Example 2.9 For cylindrical frame field, the attitude matrix can be written as

\[
A = \begin{bmatrix}
\cos \theta & \sin \theta & 0 \\
-\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{bmatrix}.
\]

Thus

\[
\omega = dA A^T = \begin{bmatrix}
-\sin \theta \, d\theta & \cos \theta \, d\theta & 0 \\
-\cos \theta \, d\theta & -\sin \theta \, d\theta & 0 \\
0 & 0 & 0
\end{bmatrix} \cdot \begin{bmatrix}
\cos \theta & -\sin \theta & 0 \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

= \begin{bmatrix}
0 & d\theta & 0 \\
-\,d\theta & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

Thus \(\omega_{12} = d\theta\), and except \(\omega_{21} = -\omega_{12}\), all other connection forms are zero. Thus we have

\[
\nabla_V E_1 = \omega_{12}[V] E_2 = V[\theta] E_2,
\]
\[
\nabla_V E_2 = -V[\theta] E_1,
\]
\[
\nabla E_3 = 0.
\]

These equations have immediate geometrical significance. Because \(V\) is arbitrary, the third equation says that the vector field \(E_3\) is parallel. We knew this already since in the cylindrical frame field, \(E_3\) is just \(U_3\). The first two equations tell us that the covariant derivatives of \(E_1\) and \(E_2\) with respect to a vector field \(V\) depend only on the rate of change of the angle \(\theta\) in the \(V\) direction.

For example, the definition of \(\theta\) shows that \(V[\theta] = 0\) whenever \(V\) is a vector field that at each point is tangent to a plane through the \(z\)-axis. Thus for a vector field of this type the connection equations above predict that \(\nabla_V E_1 = \nabla_V E_2 = 0\). In fact, it is clear from Fig. 2.19 that \(E_1\) and \(E_2\) do remain parallel on any plane through the \(z\)-axis.

### 2.8 The Structure Equations

**Definition 2.15**

If \(\{E_1, E_2, E_3\}\) is a frame field on \(\mathbb{R}^3\), then the dual 1-forms \(\{\theta_1, \theta_2, \theta_3\}\) of the frame field are the 1-forms such that

\[
\theta_i[v] = v \cdot E_i(p)
\]

for each tangent vector \(v\) to \(\mathbb{R}^3\) at \(p\).

From the above definition, we have

\[
\theta_i[E_j] = \delta_{ij}.
\]

**Example 2.10** The dual basis for the natural frame field \(\{U_1, U_2, U_3\}\) is \(\{dx_1, dx_2, dx_3\}\).

**Lemma 2.7**

Let \(\{\theta_1, \theta_2, \theta_3\}\) be the dual 1-forms of a frame field \(\{E_1, E_2, E_3\}\). Then any 1-form \(\phi\) on \(\mathbb{R}^3\) has a unique
\[ \phi = \phi(E_i) \theta_i. \]

**Proof.** Let \( V \) be any vector field. Then
\[
(\phi(E_i) \theta_i)[V] = \phi(E_i) (\theta_i[V]) = \phi(\theta_i[V]E_i) = \phi(V).
\]
Thus
\[ \phi = \phi(E_i) \theta_i. \]

Assume that \( \{E_1, E_2, E_3\} \) is a framed field. Assume that
\[ E_i = a_{ij} U_j. \]
Then we have, using the above lemma, that
\[ \theta_i = \theta_i(U_j) dx_j = (U_j \cdot E_i) dx_j = a_{ij} dx_j. \]

**Theorem 2.10**

Let \( \{E_1, E_2, E_3\} \) be a frame field on \( \mathbb{R}^3 \) with dual forms \( \{\theta_1, \theta_2, \theta_3\} \) and connection forms \( \omega_{ij} \). Then we have

1. **the first structural equations**
   \[ d\theta_i = \omega_{ij} \wedge \theta_j; \]
2. **the second structural equations**
   \[ d\omega_{ij} = \omega_{ik} \wedge \omega_{kj}. \]

**Proof.** Proof of the first structural equations. We have
\[ d\theta_i = d(a_{ij} dx_j) = da_{ij} \wedge dx_j. \]

Since \( \omega_{ij} = a_{jk} da_{ik} \), we have
\[ \omega_{ij} \wedge \theta_j = a_{jk} da_{ik} \wedge \theta_j. \]

Because \( (a_{ij}) \) is an orthogonal matrix-valued function, we have \( a_{jk} a_{js} = \delta_{ks} \). Since \( a_{jk} \theta_j = a_{jk} a_{js} dx_s = \delta_{ks} dx_s = dx_k \), we have
\[ \omega_{ij} \wedge \theta_j = da_{ik} \wedge dx_k = d\theta_i. \]

Proof of the second structural equations. We have
\[ d\omega_{ij} = d(a_{jk} da_{ik}) = da_{jk} \wedge da_{ik}. \]

On the other hand, we have
\[ \omega_{ik} \wedge \omega_{jk} = a_{ks} da_{is} \wedge a_{jt} da_{kt} = a_{ks} a_{jt} da_{is} \wedge da_{kt}. \]

Since \( a_{ks} a_{kt} = \delta_{st} \), we must have
\[ a_{ks} da_{kt} + a_{kt} da_{ks} = 0. \]
Thus
\[ \omega_{ik} \wedge \omega_{kj} = -a_{kt} a_{jt} \, da_{is} \wedge da_{ks} = -\delta_{kj} \, da_{is} \wedge da_{ks} = da_{js} \wedge da_{ts}. \]

**Example 2.11** In this example, we use verify the structure equations of the spherical frame field. The spherical coordinate system is given by
\[
\begin{align*}
x &= \rho \cos \varphi \cos \theta, \\
y &= \rho \cos \varphi \sin \theta, \\
z &= \rho \sin \varphi.
\end{align*}
\]

The spherical frame field is defined by
\[
\begin{align*}
F_1 &= \cos \varphi (\cos \theta \, U_1 + \sin \theta \, U_2) + \sin \varphi \, U_3 \\
F_2 &= -\sin \theta \, U_1 + \cos \theta \, U_2, \\
F_3 &= -\sin \varphi (\cos \theta \, U_1 + \sin \theta \, U_2) + \cos \varphi \, U_3.
\end{align*}
\]

From the above, the dual frame field is
\[
\begin{align*}
\theta_1 &= d\rho \\
\theta_2 &= \rho \cos \varphi \, d\theta \\
\theta_3 &= \rho \, d\varphi
\end{align*}
\]

One can verify the above formulas for the dual basis directly, or they follow from the fact that
\[
\begin{align*}
\frac{\partial}{\partial \rho} &= F_1, \\
\frac{\partial}{\partial \theta} &= \rho \cos \varphi \, F_2, \\
\frac{\partial}{\partial \varphi} &= \rho \, F_3.
\end{align*}
\]

Let
\[
\omega_{12} = \cos \varphi \, d\theta, \quad \omega_{13} = d\varphi, \quad \omega_{23} = \sin \varphi \, d\theta.
\]

In what follows, we check all the structural equations. By the skew-symmetry, there are 6 independent equations.
\[
\begin{align*}
d\theta_1 &= \omega_{12} \wedge \theta_2 + \omega_{13} \wedge \theta_3, \\
d\theta_2 &= -\omega_{12} \wedge \theta_1 + \omega_{23} \wedge \theta_3, \\
d\theta_3 &= -\omega_{13} \wedge \theta_1 - \omega_{23} \wedge \theta_2, \\
d\omega_{12} &= -\omega_{13} \wedge \omega_{23}, \\
d\omega_{12} &= -\omega_{13} \wedge \omega_{23}, \\
d\omega_{12} &= -\omega_{13} \wedge \omega_{23}.
\end{align*}
\]

It would be nice to write the structural equations in terms of matrix notations.

**Theorem 2.11**

Let \( \theta = (\theta_1, \theta_2, \theta_3) \). Then we have
\[
d\theta = -\theta \wedge \omega^T, \quad d\omega = \omega \wedge \omega.
\]
Proof of the matrix version of the structural equations. Let \( A = (a_{ij}) \) be the attitude matrix. Since \( \theta_i = a_{ij} \, dx_j \), we have
\[
d\theta = dx \, A^T,
\]
where \( dx = (dx_1, dx_2, dx_3) \). Thus
\[
d\theta = d(dx \, A^T) = ddx \, A^T - dx \wedge dA^T = -dx \wedge dA^T.
\]
We have
\[
-\theta \wedge \omega^T = -\theta \wedge (dA \, A^T)^T = -\theta \wedge A \, dA^T = -dx \, A \wedge dA^T = -dx \wedge dA^T.
\]
This proves the first structural equations.

To prove the second structural equations, we compute
\[
d\omega = d(dA \, A^T) = -dA \wedge dA^T.
\]
On the other hand, we have
\[
\omega \wedge \omega = -\omega \wedge \omega^T = -(dA \, A^T) \wedge (dA \, A^T)^T = -dA \, A^T \wedge A \, dA^T = -dA \wedge dA^T.
\]
This proves the second structural equations.

External Link. In this video, we further discuss the structural equations.
Chapter 3 The Fundamental Theorem of Curves

3.1 Isometry of $\mathbb{R}^3$

**Definition 3.1**

An isometry of $\mathbb{R}^3$ is a mapping $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that

$$d(F(p), F(q)) = d(p, q)$$

for any points $p, q$ in $\mathbb{R}^3$.

**Example 3.1 (Translation and Rotation)** Let $b$ be a fixed vector. Then the mapping

$$F(p) = p + b,$$

which is called a translation, is an isometry.

Let $A$ be a $3 \times 3$ orthogonal matrix. Then the mapping

$$F(p) = Ap,$$

which is called a rotation, is an isometry: let’s elaborate.

The isometry is equivalent to

$$\|Ap - Aq\| = \|p - q\|$$

which is equivalent to

$$\|A(p - q)\|^2 = \|p - q\|^2.$$

We have

$$\|A(p - q)\|^2 = (A(p - q)) \cdot (A(p - q)) = (A(p - q))^T A(p - q) = (p - q)^T A^T A(p - q).$$

Since $A$ is an orthogonal matrix, we have $A^T A = I$. Thus

$$\|A(p - q)\|^2 = (p - q)^T (p - q) = \|p - q\|^2.$$

**Example 3.2 (Reflection)** Let $v \neq 0$ be a vector. Let

$$F(p) = -p + 2 \frac{v \cdot p}{\|v\|^2} v.$$

Then $F$ is called reflection of a vector with respect to $v$. Let

$$A = -I + \frac{2v v^T}{\|v\|^2}.$$

Then it is easy to verify that $A$ is an orthogonal matrix, and $\det(A) = -1$. Thus in our terminology, a reflection is also a rotation.
3.1 Isometry of \( \mathbb{R}^3 \)

**Proposition 3.1**

If \( F, G \) are isometries of \( \mathbb{R}^3 \). Then the composition map \( GF \) is also an isometry.

**Proof.** Let \( p, q \in \mathbb{R}^3 \). We then have

\[
d(GF(p), GF(q)) = d(G(F(p)), G(F(q))) = d(F(p), F(q)) = d(p, q).
\]

Thus \( GF \) is also an isometry.

**Remark** By the above proposition, the set of all isometries of \( \mathbb{R}^3 \) forms a group. Such a group is called the isometry group of \( \mathbb{R}^3 \).

We now prove the main result of this section.

**Theorem 3.1**

Let \( F \) be an isometry of \( \mathbb{R}^3 \). Then there exist an orthogonal matrix \( A \) and a vector \( b \) such that

\[
F(p) = Ap + b.
\]

**Proof.** Let \( G(p) = F(p) - F(0) \). Then by Proposition 3.1, \( G \) is an isometry. By the definition of isometry, we have

\[
\|G(p) - G(q)\| = \|p - q\|
\]

for any \( p, q \in \mathbb{R}^3 \). In particular, since \( G(0) = 0 \), we have \( \|G(p)\| = \|p\| \).

Expanding

\[
\|G(p) - G(q)\|^2 = \|G(p)\|^2 - 2 G(p) \cdot G(q) + \|G(q)\|^2,
\]

we have

\[
\|G(p) - G(q)\|^2 = \|p\|^2 - 2 G(p) \cdot G(q) + \|q\|^2 = \|p - q\|^2.
\]

Simplifying, we get

\[
G(p) \cdot G(q) = p \cdot q.
\]

We now prove that \( G \) is a linear transformation: let \( a, b \in \mathbb{R} \). We consider

\[
\|G(ap + bq) - aG(p) - bG(q)\|^2.
\]

By a straightforward computation, we have

\[
\|G(ap + bq) - aG(p) - bG(q)\|^2
\]

\[
= \|G(ap + bq)\|^2 + a^2\|G(p)\|^2 + b^2\|G(q)\|^2 - 2a G(ap + bq) \cdot G(p) - 2b G(ap + bq) \cdot G(q) + 2ab G(p) \cdot G(q)
\]

\[
= \|ap + bq\|^2 + a^2\|p\|^2 + b^2\|q\|^2 - 2a (ap + bq) \cdot p - 2b (ap + bq) \cdot q + 2ab p \cdot q = 0.
\]

Therefore \( G \) is a linear transformation. Thus there is a \( 3 \times 3 \) matrix \( A \), we have \( G(p) = Ap \). From \( (Ap) \cdot (Aq) = p \cdot q \), we know that \( A \) is an orthogonal matrix. This completes the proof.
In this section, we shall prove the Fundamental Theorem of Curves. In a nutshell, it states that a curve is completely determined by its curvature and torsion functions.

**Definition 3.2**

Two curves \( \alpha, \beta : I \to \mathbb{R}^3 \) are congruent provided there exists an isometry \( F \) of \( \mathbb{R}^3 \) such that \( \beta = F(\alpha) \); that is, \( \beta(t) = F(\alpha(t)) \) for all \( t \) in \( I \).

**Theorem 3.2 (The Fundamental Theorem of Curves)**

Let \( \kappa(s) > 0 \) and \( \tau(s) \) be two smooth functions defined on an open interval \( I \). Assume that \( 0 \in I \). Then for any given vector \( p \in \mathbb{R}^3 \) and any frame \( \{ e_1, e_2, e_3 \} \) of the tangent space \( T_p(\mathbb{R}^3) \), there is a unique unit speed curve \( \beta(s) \) such that \( \beta(0) = p, \beta'(0) = e_1, \beta''(0) = \kappa(0)e_2 \), and moreover, the curvature and torsion of the curve are given by the functions \( \kappa(s) \) and \( \tau(s) \) respectively.

In order to prove the above result, we shall use the Picard–Lindelöf Theorem, also called the Fundamental Theorem of Ordinary Differential Equations. We cite the theorem in the version of single and multi-variables, respectively.

**Theorem 3.3 (Picard–Lindelöf Theorem, single variable version)**

Consider the initial value problem

\[
y'(t) = f(t, y(t)), \quad y(t_0) = y_0
\]

for the unknown function \( y = y(t) \), where \( y_0 \) is a real number. Suppose \( f = f(x, y) \) is uniformly Lipschitz continuous in \( y \) (meaning the Lipschitz constant can be taken independent of \( x \)) and continuous in \( x \), then for some value \( \varepsilon > 0 \), there exists a unique solution \( y(t) \) to the initial value problem on the interval \( (t_0 - \varepsilon, t_0 + \varepsilon) \).

In application, the function \( f(x, y) \) is usually smooth. Thus the conditions of the above theorem is often easily satisfied. On the other hand, the existence of the solution of the initial value problem is only guaranteed on a smooth interval \( (t_0 - \varepsilon, t_0 + \varepsilon) \), where \( \varepsilon > 0 \) depends on the function \( f \) and number \( y_0 \).

The multivariable version of the Picard–Lindelöf Theorem is as follows.

**Theorem 3.4 (Picard–Lindelöf Theorem, multivariable version)**

Let \( y = y(t) \) be an unknown \( \mathbb{R}^n \)-valued function of \( t \). Let \( f(t, y) \) be a function of \( (n + 1) \) variables. Consider the initial value problem

\[
y'(t) = f(t, y(t)), \quad y(0) = y_0,
\]

where \( y_0 \) is a vector. Suppose \( f = f(x, y) \) is uniformly Lipschitz continuous in \( y \) (meaning the Lipschitz constant can be taken independent of \( x \)) and continuous in \( x \), then for some value \( \varepsilon > 0 \), there exists a unique solution \( y(t) \) to the initial value problem on the interval \( (t_0 - \varepsilon, t_0 + \varepsilon) \).

Again, when the function \( f \) is smooth, then all the conditions in the above theorem are satisfied. Nevertheless, as in the single variable case, \( \varepsilon \) would be small. Therefore these two theorems are sometimes called local...
existence theorems, or short-time existence theorems.

Under certain conditions, we can extend the short-time existence to long time existence.

**Theorem 3.5 (Extension of the Solution)**

We use the same notations as in the above theorem. Assume that \( T \) is the maximum existence time. That is, the solution \( y(t) \) can be extended to \((t_0 - \epsilon, t_0 + T)\) but not beyond time \( T \). Then we have

\[
\lim_{t \to T} \|y(t)\| = \infty.
\]

In other words, if \( y \) is a bounded function on any open interval \((t_0 - \epsilon, t_0 + T)\), then there is an \( \epsilon > 0 \) such that the solution can be extended to \((t_0 - \epsilon, t_0 + T + \epsilon)\).

In particular, we can prove that for system of linear equations, the solution can be extended throughout the whole real number \( \mathbb{R} \).

We shall now use the above theorems to prove the main result.

**Proof of Theorem 3.2.** As is well-known, the Frenet formulas \((2.1)\)

\[
(T, N, B)' = (T, N, B) \begin{pmatrix} 0 & -\kappa & 0 \\ \kappa & 0 & -\tau \\ 0 & \tau & 0 \end{pmatrix}
\]

can be regarded as system of linear first order equations with respect to the unknown matrix-valued function\( (T, N, B) \). By Theorem 3.5, the solution \((T, N, B)\) exists and is unique with the initial values \((T(0), N(0), B(0)) = (e_1, e_2, e_1 \times e_2)\). Once \( T(s) \) is solved, it uniquely determines the curve function \( \beta(s) \) since \( \beta'(s) = T(s) \) and \( \beta(0) = p \).

---

Since a matrix can always be identified as a vector, the equation is of the form in Theorem 3.4.

**Corollary 3.1**

Let \( \alpha, \beta \) be two curves defined on an open interval. Assume that they share the curvature function and the torsion function \( \kappa(s) > 0 \) and \( \tau(s) \). Then these two curves are congruent.

**Proof.** Let \( \{e_1, e_2, e_3\} \) and \( \{f_1, f_2, f_3\} \) be the two initial frames of \( \alpha \) and \( \beta \), respectively. Let \( \alpha(0) = p \), \( \beta(0) = q \). Then there is an isometry \( F \) such that \( F(p) = q \) and \( F_e(e_i) = f_i \). By the uniqueness of the above theorem, we know that \( \beta \) and \( F(\alpha) \) are the same curve.

**Corollary 3.2**

Then \( \alpha \) is a helix if and only if both its curvature and torsion are nonzero constants.
Chapter 4  Calculus on Surface

Introduction

- Regular Surface
- (Proper) Coordinate Patch
- Differentiable Map
- Differential Forms

4.1 Surfaces in \( \mathbb{R}^3 \)

What is a surface? The definition of the book is somewhat not parallel with respect to that of a curve. A regular curve is defined by a function from an open interval to \( \mathbb{R}^3 \). However, a surface is first defined as a subset of \( \mathbb{R}^3 \).

In this section, we will define a surface rigidly, and bridge calculus to the subset defined as a surface.

Definition 4.1

An open set \( D \) of \( \mathbb{R}^2 \) is a subset of \( \mathbb{R}^2 \) such that every point of \( D \) contains a small open disk of \( \mathbb{R}^2 \).

Example 4.1 The (open) unit disk in \( \mathbb{R}^2 \)

\[ D = \{(x, y) \mid x^2 + y^2 < 1\} \]

is an open set.

Proof. To prove this, we let \((a, b) \in D\). By definition, \( a^2 + b^2 < 1 \). Let

\[ r = \frac{1}{2}(1 - \sqrt{a^2 + b^2}) > 0. \]

Let \((x, y)\) be a pair of real numbers such that

\[ (x - a)^2 + (y - b)^2 < r^2. \]

Let \(d(\cdot, \cdot)\) be the distance function. Then

\[ \sqrt{x^2 + y^2} = d((x, y), (0, 0)) \]

\[ \leq d((x, y), (a, b)) + d((a, b), (0, 0)) < r + \sqrt{a^2 + b^2} < 1. \]

Hence \((x, y) \in D\).

Definition 4.2

A coordinate patch \( \mathbf{x} : D \to \mathbb{R}^3 \) is a one-to-one regular mapping of an open set \( D \) of \( \mathbb{R}^2 \) to \( \mathbb{R}^3 \).
4.1 Surfaces in $\mathbb{R}^3$

**Lemma 4.1**

A coordinate patch $\mathbf{x}$ is a mapping $\mathbf{x} : D \rightarrow \mathbb{R}^3$ such that

1. $\mathbf{x}$ is a one-to-one mapping;
2. $\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \neq 0$.

**Proof.** The only thing that needs to prove is that $\mathbf{x}$ is regular if and only if

$$\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \neq 0.$$

By the definition of a regular map, we know that $\mathbf{x}$ is regular if and only if its Jacobian matrix

$$\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \end{bmatrix}$$

is of rank 2, which is equivalent to that the vectors $\frac{\partial \mathbf{x}}{\partial u}$ and $\frac{\partial \mathbf{x}}{\partial v}$ are linearly independent, hence the cross-product is non-zero.

**Example 4.2** Let $\mathcal{U}$ be an open set of $\mathbb{R}^2$, and let $f$ be a $C^k$ function on $\mathcal{U}$ which is one-to-one. Define

$$\mathbf{x} : \mathcal{U} \rightarrow \mathbb{R}^3, \quad (u, v) \mapsto (u, v, f(u, v)).$$

Then $\mathbf{x}$ defines a coordinate patch.

**Proof.** To prove this, we compute

$$\frac{\partial \mathbf{x}}{\partial u} = \begin{bmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial u} \end{bmatrix}, \quad \frac{\partial \mathbf{x}}{\partial v} = \begin{bmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial v} \end{bmatrix}$$

It follows that

$$\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} = \begin{bmatrix} -\frac{\partial f}{\partial u} \\ -\frac{\partial f}{\partial v} \\ 1 \end{bmatrix}$$

which is never zero (the third component is never zero). This proves that $\mathbf{x}$ is a coordinate patch.

**Remark** This can be compared to the definition of regular curve. A regular curve is a mapping $\alpha : (a, b) \rightarrow \mathbb{R}^3$, where $(a, b)$ is an open interval. In the definition of curve, we could require $\alpha$ to be one-to-one but we allowed slight generalization. On the other side, a non-one-to-one surface looks bizarre.

**Example 4.3** Let

$$\mathcal{W} = \{ (w^1, w^2) \in \mathbb{R} \mid -\frac{\pi}{2} < w^1 < \frac{\pi}{2}, -\pi < w^2 < \pi \},$$

and define a mapping by

$$z(w^1, w^2) = (\cos w^1 \cos w^2, \cos w^1 \sin w^2, \sin w^1).$$

Then $z$ is a coordinate patch.
4.1 Surfaces in $\mathbb{R}^3$

**Proof.** It is not hard to prove that the Jacobian is of rank 2, hence the mapping is regular. We compute

$$\frac{\partial z}{\partial w_1} = (-\sin w_1 \cos w^2, -\sin w_1 \sin w^2, \cos w^1);$$

$$\frac{\partial z}{\partial w_2} = (-\cos w_1 \cos w^2, \cos w_1 \sin w^2, -\sin w_1 \cos w^1).$$

It follows that

$$\left\| \frac{\partial z}{\partial w_1} \times \frac{\partial z}{\partial w_2} \right\|^2 = \cos^4(w^1) + \sin^2(w^1) \cos^2(w^1) = \cos^2(w^1).$$

Since $-\pi/2 < w^1 < \pi/2$, the above right side is not zero, this proves that the surface is simple.

*It is a little bit complicated to prove that $z$ is one-to-one.* We assume that

$$z(w^1, w^2) = z(u^1, u^2).$$

Therefore we have

$$\cos w^1 \cos w^2 = \cos u^1 \cos u^2;$$

$$\cos w^1 \sin w^2 = \cos u^1 \sin u^2;$$

$$\sin w^1 = \sin u^1.$$  

From the third equation, we conclude that

$$w^1 = u^1.$$  

Moreover, $\cos w^1, \cos u^1 \neq 0$. Thus we have

$$\cos w^2 = \cos u^2, \quad \sin w^2 = \sin u^2.$$  

Hence we have $w^2 = u^2$.

**Example 4.4** Here is an example of a non-regular surface. Let

$$\mathbf{x}(u, v) = (u^2, v^2, uv)$$

Then

$$\frac{\partial \mathbf{x}}{\partial u} = (2u, 0, v), \quad \frac{\partial \mathbf{x}}{\partial v} = (0, 2v, u).$$

At the point $(u, v) = (0, 0), \mathbf{x}_1 \times \mathbf{x}_2 = 0$, and hence the surface is not regular.

**Definition 4.3**

A coordinate patch is called **proper**, if the inverse function

$$\mathbf{x}^{-1} : \mathbf{x}(D) \to D$$

is continuous.

**External Link.** By definition, $\mathbf{x}(D)$ is only a subset of $\mathbb{R}^3$. what does it mean by the mapping $\mathbf{x}^{-1}$ is a continuous function on $\mathbf{x}(D)$? We shall elaborate this problem here.

**Definition 4.4**

A **surface** in $\mathbb{R}^3$ is a subset $M \subset \mathbb{R}^3$ such that for each point $\mathbf{p}$ of $M$ there exists a proper coordinate patch in $M$ whose image contains a neighborhood of $\mathbf{p}$ in $M$. 
4.1 Surfaces in $\mathbb{R}^3$

In particular, a single proper coordinate patch defines a surface, such a surface is defined as a **simple** surface.

**Example 4.5** The classical example of a surface is the unit sphere. Let $M$ be the subset of $\mathbb{R}^3$ defined by

\[ p = (p_1, p_2, p_3), \text{ where} \]

\[ p_1^2 + p_2^2 + p_3^2 = 1. \]

Then $M$ is a surface.

**Proof.** We shall prove that at each point $p \in M$ there is a proper coordinate patch. Let $p = (p_1, p_2, p_3)$. Since $p_1^2 + p_2^2 + p_3^2 = 1$, there is at least one of $p_1$, say $p_3$, such that $|p_3| \geq |p_1|$ and $|p_3| \geq |p_2|$. Without loss of generality, we assume that $p_3 > 1/5$. Thus $p_3$ is uniquely determined by the equation

\[ p_3 = \sqrt{1 - p_1^2 - p_2^2}. \]

Since $p_3 > 1/5$, we must have $p_1^2 + p_2^2 < 24/25$.

We thus construct a coordinate patch $x_1$ as follows. Let

\[ D = \{(u, v) \mid u^2 + v^2 < 24/25\}. \]

Let

\[ x_1 : D \to \mathbb{R}^3, \quad (u, v) \mapsto (u, v, \sqrt{1 - u^2 - v^2}). \]

Similarly, we can define the other five coordinate patches by

\[ x_2 : D \to \mathbb{R}^3, \quad (u, v) \mapsto (u, v, -\sqrt{1 - u^2 - v^2}), \]

\[ x_3 : D \to \mathbb{R}^3, \quad (u, v) \mapsto (u, \sqrt{1 - u^2 - v^2}, v), \]

\[ x_4 : D \to \mathbb{R}^3, \quad (u, v) \mapsto (u, -\sqrt{1 - u^2 - v^2}, v), \]

\[ x_5 : D \to \mathbb{R}^3, \quad (u, v) \mapsto (\sqrt{1 - u^2 - v^2}, u, v), \]

\[ x_6 : D \to \mathbb{R}^3, \quad (u, v) \mapsto (-\sqrt{1 - u^2 - v^2}, u, v). \]

These patches cover the whole $M$.

We shall prove these patches are the proper coordinate patches. We prove that for $x_1$ for example.

First, the map $x_1$ is apparently one-to-one. Next, the map has to be regular by Example 4.2. The inverse maps is given by

\[ x_1^{-1} : x_1(D) \to D, \quad (p_1, p_2, p_3) \mapsto (p_1, p_2). \]

The continuity of the mapping is ensured by the following obvious fact: if $||p - q|| < \delta$, then we must have both $|p_1 - q_1| < \delta, |p_2 - q_2| < \delta$. 

65
4.1 Surfaces in \( \mathbb{R}^3 \)

This proves that \( x_i^{-1} \) is proper. Similarly we can prove that \( x_i \) for \( 2 \leq i \leq 6 \) are all proper coordinate patches. Thus \( M \) is a surface in \( \mathbb{R}^3 \).

\[ \text{Note} \]

Let \( z = f(x, y) \) be a function on 2-variables defined on \( D \). Then by the above proof, the mapping

\[ F : D \rightarrow \mathbb{R}^3, \quad (u, v) \mapsto (u, v, f(u, v)) \]

is a proper coordinate patch. Such a coordinate patch is called the Monge Patch.

The following is a useful result to provide a lot of examples of surfaces.

**Theorem 4.1**

Let \( g \) be a differentiable real-valued function on \( \mathbb{R}^3 \), and \( c \) a number. The non-empty subset \( M : g(x, y, z) = c \) of \( \mathbb{R}^3 \) is a surface if the differential \( dg \) is not zero at any point of \( M \).

**Proof.** All we do is give geometric content to a famous result of advanced calculus – the implicit function theorem. If \( p \) is a point of \( M \), we must find a proper patch covering a neighborhood of \( p \) in \( M \). Since the hypothesis on \( dg \) is equivalent to assuming that at least one of these partial derivatives is not zero at \( p \), say \( \partial g/\partial z(p) \neq 0 \). In this case, the implicit function theorem says that near \( p \) the equation \( g(x, y, z) = c \) can be solved for \( z \). More precisely, it asserts that there is a differentiable real-valued function \( h \) defined on a neighborhood \( D \) of \( (u, v) \) such that

1. For each point \((u, v) \in D\) the point \((u, v, h(u, v))\) lies in \( M \); that is,

   \[ g(u, v, h(u, v)) = c. \]

2. Points of the form \((u, v, h(u, v))\), with \((u, v) \in D\), fill a neighborhood of \( p \in M \).

It follows immediately that the maps \( x : D \rightarrow \mathbb{R}^3 \) such that

\[ x(u, v) = (u, v, h(u, v)) \]

is a Monge Patch, which is proper and regular. Since \( p \) was an arbitrary point of \( M \), we conclude that \( M \) is a surface.

**Corollary 4.1**

The sphere \( M \) defined by

\[ \{ (x, y, z) \mid x^2 + y^2 + z^2 = r^2 \} \]

where \( r > 0 \) is a surface.

**Example 4.6** (Surfaces of revolution). Let \( C \) be a curve in a plane \( P \subset \mathbb{R}^3 \), and let \( A \) be a line in \( P \) that does not meet \( C \). When this profile curve \( C \) is revolved around the axis \( P \), it sweeps out a surface of revolution \( M \) in \( \mathbb{R}^3 \). Prove that it is a surface.

**Proof.** Without loss of generality, we assume that line \( A \) is the \( x \)-axis. Let the plane curve be defined by

\[ x = x(t), \quad y = y(t). \]

Since the curve doesn’t meet the \( x \)-axis, we can assume that \( y(t) > 0 \). We also assume that the curve
4.2 Patch Computations

is regular, that is,

\[(x'(t))^2 + (y'(t))^2 > 0\]

for all \(t\). It is therefore not hard to see that the surface of revolution can be parametrized by

\[x = x(t), \quad y = y(t)\cos \theta, \quad z = y(t)\sin \theta\]

where \(0 \leq \theta \leq 2\pi\).

Assume that \(t\) is defined on an open interval \(I\). Since the curve is regular, we can write \(I = \bigcup_i I_i\) for open intervals \(I_i\) such that the mappings \(t \mapsto (x(t), y(t))\) is one-to-one. We can also split the interval

\([0, 2\pi] = (-0.1, \pi + 0.1) \cup (\pi, 2\pi) = V_1 \cup V_2.\]

Then on each \(I_i \times V_j\), the map \((t, \theta) \mapsto (x(t), y(t)\cos \theta, y(t)\sin \theta)\) is one-to-one.

The proof of regularity of the mapping is straightforward and I will omit it.

We end this section by the following two pictures of surfaces.

4.2 Patch Computations

Let \(x : D \to \mathbb{R}^3\) be a coordinate patch. In this section, we shall first define two important vector fields \(x_u, x_v\), as follows.

**Definition 4.5**

Let \(x : D \to \mathbb{R}^3\) be a coordinate patch. Holding \(u\) or \(v\) constant in the function \((u, v) \mapsto x(u, v)\) produces curves. Explicitly, for each point \((u_0, v_0)\) in \(D\) the curve

\[u \mapsto x(u, v_0)\]

is called the \(u\)-parameter curve, \(v = v_0\), of \(x\); and the curve

\[v \mapsto x(u_0, v)\]

is the \(v\)-parameter curve, \(u = u_0\), of \(x\). We define

1. The velocity vector at \(u_0\) of the \(u\)-parameter curve, \(v = v_0\), is denoted by \(x_u(u_0, v_0)\).
2. The velocity vector at \(v_0\) of the \(v\)-parameter curve, \(u = u_0\), is denoted by \(x_v(u_0, v_0)\).

The vectors \(x_u(u_0, v_0)\) and \(x_v(u_0, v_0)\) are called the partial velocities of \(x\) at \((u_0, v_0)\).

**Remark** No matter what the wordings are, the vector fields \(x_u, x_v\) are important, as they are the partial derivative of the mapping

\[x(u, v) = (x_1(u, v), x_2(u, v), x_3(u, v)).\]
In terms of components, we can write
\[
x_u = \left( \frac{\partial x_1}{\partial u}, \frac{\partial x_2}{\partial u}, \frac{\partial x_3}{\partial u} \right),
x_v = \left( \frac{\partial x_1}{\partial v}, \frac{\partial x_2}{\partial v}, \frac{\partial x_3}{\partial v} \right).
\]

**Note** The vector fields \( x_u, x_v \) have point of application of \( x(u, v) \).

**Definition 4.6**
A regular mapping \( x : D \to \mathbb{R}^3 \) whose image lies in a surface \( M \) is called a parametrization of the region \( x(D) \) in \( M \).

**Example 4.7** The geographical patch in the sphere. Let \( \Sigma \) be the sphere of radius \( r > 0 \) centered at the origin of \( \mathbb{R}^3 \). Longitude and latitude on the earth suggest a patch in \( \Sigma \) quite different from the Monge patch used before. The point \( x(u, v) \) of \( \Sigma \) with longitude \( u (-\pi < u < \pi) \) and latitude \( v (-\pi/2 < v < \pi/2) \) has Euclidean coordinates
\[
x(u, v) = (r \cos v \cos u, r \cos v \sin u, r \sin v)
\]
We can prove that the mapping is regular.

**Remark** A coordinate patch is more restrictive: in addition to be required regularity, it has to be one-to-one.

**Example 4.8** (Parametrization of a surface of revolution). Suppose that the surface \( M \) is obtained by revolving a curve \( C \) in the upper half of the \( xy \) plane about the \( x \) axis. Now let
\[
\alpha(u) = (g(u), h(u), 0)
\]
be a parametrization of \( C \) (with \( h(u) > 0 \)). Define
\[
x(u, v) = (g(u), h(u) \cos v, h(u) \sin v).
\]

**Example 4.9** (Torus of Revolution \( T \)). This is the surface of revolution obtained when the profile curve \( C \) is a circle. Suppose that \( C \) is the circle in the \( xz \)-plane with radius \( r > 0 \) and center \((R, 0, 0)\). We shall rotate about the \( z \) axis; hence we must require \( R > r \) to keep \( C \) from meeting the axis of revolution. A natural parametrization for \( C \) is
\[
\alpha(u) = (R + r \cos u, r \sin u).
\]
We then have
\[
x(u, v) = ((R + r \cos u) \cos v, (R + r \cos u) \sin v, r \sin u).
\]

**Definition 4.7**
A ruled surface is a surface swept out by a straight line \( L \) moving along a curve \( \beta \). The various positions of the generating line \( L \) are called the rulings of the surface. Such a surface always has a ruled...
parametrization,

\[ \mathbf{x}(u, v) = \beta(u) + v\delta(u). \]

We call \( \beta \) the base curve and \( \delta \) the director curve, although \( \delta \) is usually pictured as a vector field on \( \beta \) pointing along the line \( L \).

**Example 4.10** The equation

\[ \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \]

where \( a, b, c > 0 \), represents the **Hyperboloid of one sheet**. It contains two families of straight lines \( x = \pm a, bz = \pm cy \).

### 4.3 Differentiable Functions and Tangent Vectors

We now begin an exposition of the calculus on a surface \( M \) in \( \mathbb{R}^3 \). The space \( \mathbb{R}^3 \) will gradually fade out of the picture, since our ultimate goal is a calculus for \( M \) alone.

In order to do that, the first step is to define *differentiable* function on \( M \). Let \( f : M \to \mathbb{R} \) be a function. As we have discussed before, we are able to define the continuity of function \( f \), since \( M \) is a metric space. It is somewhat indirect to defined the differentiability of \( f \), which we shall do in the following.

**Definition 4.8**

Let \( f : M \to \mathbb{R} \) be a function. Let \( \mathbf{x} : D \to \mathbb{R}^3 \) be a coordinate patch. Then \( f(\mathbf{x}) \) would be a function on \( D \). Since \( D \) is an open set of \( \mathbb{R}^2 \), we can define it differentiability as we did in Calculus. We say that \( f \) is differentiable at a point \( \mathbf{p} \), if \( \mathbf{p} = \mathbf{x}(u, v) \) for \( (u, v) \in D \), and \( f(\mathbf{x}(u, v)) \) is differentiable at \( (u, v) \). If \( f \) is differentiable at any point \( \mathbf{p} \in M \), we say that \( f \) is differentiable on \( M \).

Next, we define a differentiable map from \( \mathbb{R}^n \) to \( M \).

**Definition 4.9**

Let \( O \) be an open set of \( \mathbb{R}^n \). Let \( F : O \to M \) be a continuous mapping. Let \( \mathbf{p} \in \mathbb{R}^n \). Let \( \mathbf{x} : D \to M \) be a proper coordinate patch such that \( F(\mathbf{p}) \in \mathbf{x}(D) \). We say \( F \) is differentiable at \( \mathbf{p} \), if the function \( F(\mathbf{x}^{-1}) : O \to D \) is differentiable. If \( F \) is differentiable at any point \( \mathbf{p} \in \mathbb{R}^n \), we say that \( F \) is
4.3 Differentiable Functions and Tangent Vectors

We have studied curves on $\mathbb{R}^3$. But curve on surface is also very important. In what follows, we define curve on a surface.

**Definition 4.10**

A curve $\alpha : I \to M$ in a surface $M$ is a differentiable function from an open interval $I$ into $M$.

**Remark** A plane curve is a special case of a curve on a surface. Since a surface is a subset of $\mathbb{R}^3$, a curve on a surface is also a space curve.

The following lemma is simple but important.

**Lemma 4.2**

If $\alpha$ is a curve $\alpha : I \to M$ whose route lies in the image $x(D)$ of a single patch $x(D)$, then there exist unique differentiable functions $(a_1, a_2)$ on $I$ such that

$$\alpha(t) = x(a_1(t), a_2(t))$$

for all $t$.

**Proof.** By definition, the coordinate expression $x^{-1} \alpha : I \to D$ is differentiable. It is just a curve in $\mathbb{R}^2$ whose route lies in the domain $D$ of $x$. If $(a_1, a_2)$ are the Euclidean coordinate functions of $x^{-1} \alpha$, then

$$\alpha = x x^{-1} \alpha = x(a_1, a_2).$$

We omit the proof of the uniqueness.

**Theorem 4.2**

Let $M$ be a surface in $\mathbb{R}^3$. If $F : \mathbb{R}^n \to \mathbb{R}^3$ is a (differentiable) mapping whose image lies in $M$, then considered as a function $F : \mathbb{R}^n \to M$ into $M$, $F$ is differentiable.

**Note** We omit the proof as the proof needs some advance knowledge of calculus.

**Corollary 4.2**

If $x$ and $y$ are patches in a surface $M$ in $\mathbb{R}^3$ whose images overlap, then the composite functions $x^{-1} y$ and $y^{-1} x$ are (differentiable) mappings defined on open sets of $\mathbb{R}^2$. 
4.3 Differentiable Functions and Tangent Vectors

Corollary 4.3
If \( x \) and \( y \) are overlapping patches in \( M \), then there exist unique differentiable functions \( \bar{u} \) and \( \bar{v} \) such that
\[
y(u, v) = x(\bar{u}(u, v), \bar{v}(u, v))
\]
for all \((u, v)\) in the domain of \( x^{-1}y \). In functional notations: \( y = x(\bar{u}, \bar{v}) \).

For the rest of this section, we discuss the second topic: tangent vectors.

Definition 4.11
Let \( p \) be a point of a surface \( M \) in \( \mathbb{R}^3 \). A tangent vector \( v \) to \( \mathbb{R}^3 \) at \( p \) is tangent to \( M \) at \( p \) provided \( v \) is a velocity of some curve in \( M \) passing through \( p \).

Note
This definition provides a relationship between the “tangent vector” in analytic geometry to the directional derivative.

Lemma 4.3
Let \( p \) be a point of a surface \( M \) in \( \mathbb{R}^3 \), and let \( x \) be a patch in \( M \) such that \( x(u_0, v_0) = p \). A tangent vector \( v \) to \( \mathbb{R}^3 \) at \( p \) is tangent to \( M \) if and only if \( v \) can be written as a linear combination of \( x_u(u_0, v_0) \) and \( x_v(u_0, v_0) \).

Proof. Note that the parameter curves of \( x \) are curves in \( M \), so \( x_u \) and \( x_v \) are always tangent to \( M \) at \( p \).

First suppose that \( v \) is tangent to \( M \) at \( p \); thus there is a curve \( \alpha \) in \( M \) such that \( \alpha(0) = p \) and \( \alpha'(0) = v \). Now by Lemma 4.2, \( \alpha \) may be written
\[
\alpha = x(a_1, a_2);
\]
hence by the chain rule,

\[ \alpha' = x_u \frac{da_1}{dt} + x_v \frac{da_2}{dt}. \]

Therefore,

\[ \mathbf{v} = \alpha'(0) = x_u \frac{da_1}{dt}(0) + x_v \frac{da_2}{dt}(0), \]

which can be written as a linear combination of \( x_u, x_v \).

Conversely, assume that

\[ \mathbf{v} = c_1 x_u + c_2 x_v. \]

Then the curve

\[ t \mapsto \mathbf{x}(u_0 + tc_1, v_0 + tc_2) \]

would have its initial velocity vector \( \mathbf{v} \).

**Definition 4.12**

A Euclidean vector field \( Z \) on a surface \( M \) in \( \mathbb{R}^3 \) is a function that assigns to each point \( p \) of \( M \) a tangent vector \( Z(p) \) to \( \mathbb{R}^3 \) at \( p \).

**Definition 4.13**

A Euclidean vector field \( V \) for which each vector \( V(p) \) is tangent to \( M \) at \( p \) is called a tangent vector field on \( M \). Frequently these vector fields are defined, not on all of \( M \), but only on some region in \( M \). As usual, we always assume differentiability. A Euclidean vector \( \mathbf{z} \) at a point \( p \) of \( M \) is normal to \( M \) if it is orthogonal to the tangent plane \( T_p(M) \) - that is, to every tangent vector to \( M \) at \( p \). And a Euclidean vector field \( Z \) on \( M \) is a normal vector field on \( M \) provided each vector \( Z(p) \) is normal to \( M \).

**Lemma 4.4**

If \( M : g = c \) is a surface in \( \mathbb{R}^3 \), then the gradient vector field

\[ \nabla g = \frac{\partial g}{\partial x_i} U_i \]

(considered only at points of \( M \)) is a nonvanishing normal vector field on the entire surface \( M \).

**Proof.** The gradient is nonvanishing (that is, never zero) on \( M \) since in the implicit case we require that the partial derivatives cannot simultaneously be zero at any point of \( M \).

We must show that \( \nabla g \cdot \mathbf{v} = 0 \) for every tangent vector \( \mathbf{v} \) to \( M \) at \( p \). First note that if \( \alpha \) is a curve in \( M \), then \( g(\alpha) = g(a_1, a_2, a_3) \) has constant value \( c \). Thus by the chain rule,

\[ \frac{\partial g}{\partial x_i}(\alpha) \frac{d\alpha_i}{dt} = 0. \]

Since \( \alpha \) is arbitrary, for any \( \mathbf{v} \), there is a curve \( \alpha \) so that

\[ \mathbf{v} = \left( \frac{d\alpha_1}{dt}, \frac{d\alpha_2}{dt}, \frac{d\alpha_3}{dt} \right). \]

Thus \( \nabla g \cdot \mathbf{v} = 0. \)
Example 4.11 Vector fields on the sphere $\Sigma : g = \sum x_i^2 = r^2$. The above lemma shows that

$$X = \frac{1}{2} \nabla g = x_i U_i$$

is a normal vector.

We end up with this definition with the following directional derivative on a surface.

**Definition 4.14**

Let $v$ be a tangent vector to $M$ at $p$, and let $f$ be a differentiable real-valued function on $M$. The derivative $v[f]$ of $f$ with respect to $v$ is the common value of $(d/dt)(f \circ \alpha)(0)$ for all curves $\alpha$ in $M$ with initial velocity $v$.

4.4 Differential Forms on a Surface

On $\mathbb{R}^3$, we have defined 0-forms, 1-forms, 2-forms, and 3-forms. Since surface is of dimension 2, we shall only define 0-forms, 1-forms and 2-forms. All other order of forms are zero.

We begin with the following definition of 1-form

**Definition 4.15**

A 1-form $\omega$ on a surface $M$ is a real-valued function on tangent vector $v$.

**Remark** For each $p \in M$, $\omega$ defines a linear functional $\omega_p : T_p M \to \mathbb{R}$ by

$$v_p \mapsto \omega(v_p).$$

The mapping $p \mapsto \omega(v_p)$ is a differentiable function on $M$.

2-forms are defined in the similar way.

**Definition 4.16**

A 2-form $\eta$ on a surface $M$ is a real-valued function on all ordered pairs of tangent vectors $v, w$ to $M$ such that

1. $\eta(v, w)$ is linear in $v$ and in $w$;
2. $\eta(v, w) = -\eta(w, v)$.

**Remark** By the skew-symmetry in the above definition, we have

$$\eta(v, v) = 0.$$

**Lemma 4.5**

Let $\eta$ be a 2-form on a surface $M$, and let $v$ and $w$ be tangent vectors at some point of $M$. Then

$$\eta(av + bw, cv + dw) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \eta(v, w).$$

**Proof.** We have

$$\eta(av + bw, cv + dw) = a \eta(v, cv + dw) + b \eta(w, cv + dw)$$
$$= a \eta(v, v) + ad \eta(v, w) + bc \eta(w, v) + bd \eta(w, w)$$
$$= (ad - bc) \eta(v, w).$$
Wherever they appear, differential forms satisfy certain general properties, established (at least partially) in Chapter 1 for forms on $\mathbb{R}^3$. To begin with, the wedge product of a $p$-form and a $q$-form is always a $(p+q)$-form. If $p$ or $q$ is zero, the wedge product is just the usual multiplication by a function. On a surface, the wedge product is always zero if $p+q > 2$. So we need a definition only for the case $p = q = 1$.

**Definition 4.17**

If $\phi$ and $\psi$ are 1-forms on a surface $M$, the wedge product $\phi \wedge \psi$ is the 2-form on $M$ such that

$$(\phi \wedge \psi)(\mathbf{v}, \mathbf{w}) = \phi(\mathbf{v}) \psi(\mathbf{w}) - \phi(\mathbf{w}) \psi(\mathbf{v})$$

for all pairs $\mathbf{v}, \mathbf{w}$ of tangent vectors to $M$.

**Remark** Let $\xi$ be a $p$-form, $\eta$ be a $q$-form. Then

$$\xi \wedge \eta = (-1)^{pq} \eta \wedge \xi.$$ 

**Definition 4.18**

Let $\phi$ be a 1-form on a surface $M$. Then the exterior derivative $d\phi$ of $\phi$ is the 2-form such that for any patch $\mathbf{x}$ in $M$,

$$d\phi(\mathbf{x}_u, \mathbf{x}_v) = \frac{\partial}{\partial u}(\phi(\mathbf{x}_u)) - \frac{\partial}{\partial v}(\phi(\mathbf{x}_u)).$$

We need to prove that the above definition is well-defined.

**Lemma 4.6**

Let $\phi$ be a 1-form on $M$. If $\mathbf{x}$ and $\mathbf{y}$ are patches in $M$, then $d_x \phi = d_y \phi$ on the overlaps of $\mathbf{x}(D)$ and $\mathbf{y}(E)$.

**Proof.** It suffices to prove that

$$(d_y \phi)(\mathbf{y}_u, \mathbf{y}_v) = (d_x \phi)(\mathbf{x}_u, \mathbf{x}_v). \tag{4.1}$$

We have the relation

$$\mathbf{y}(u,v) = \mathbf{x}(\bar{u}, \bar{v}).$$

Then by the chain rule

$$\mathbf{y}_u = \frac{\partial \bar{u}}{\partial u} \mathbf{x}_u + \frac{\partial \bar{v}}{\partial u} \mathbf{x}_\bar{v},$$

$$\mathbf{y}_v = \frac{\partial \bar{u}}{\partial v} \mathbf{x}_u + \frac{\partial \bar{v}}{\partial v} \mathbf{x}_\bar{v}.$$ 

In matrix notation, we have

$$\begin{bmatrix} \mathbf{y}_u \\ \mathbf{y}_v \end{bmatrix} = A \begin{bmatrix} \mathbf{x}_\bar{u} \\ \mathbf{x}_\bar{v} \end{bmatrix} = \begin{bmatrix} \frac{\partial \bar{u}}{\partial u} & \frac{\partial \bar{v}}{\partial u} \\ \frac{\partial \bar{u}}{\partial v} & \frac{\partial \bar{v}}{\partial v} \end{bmatrix} \begin{bmatrix} \mathbf{x}_\bar{u} \\ \mathbf{x}_\bar{v} \end{bmatrix}.$$ 

By Lemma 4.5, in order to prove (4.1), we need to prove

$$(d_y \phi)(\mathbf{y}_u, \mathbf{y}_v) = \det(A)(d_x \phi)(\mathbf{x}_\bar{u}, \mathbf{x}_\bar{v}). \tag{4.2}$$

where

$$\det A = \frac{\partial \bar{u}}{\partial u} \frac{\partial \bar{v}}{\partial v} - \frac{\partial \bar{u}}{\partial v} \frac{\partial \bar{v}}{\partial u}.$$
Using Definition 4.18, we can write (4.2) in matrix notations as

\[
\begin{bmatrix}
\frac{\partial}{\partial u}, & \frac{\partial}{\partial v}
\end{bmatrix}
J
\begin{bmatrix}
y_u \\
y_v
\end{bmatrix}
= \det(A)
\begin{bmatrix}
\frac{\partial}{\partial \bar{u}}, & \frac{\partial}{\partial \bar{v}}
\end{bmatrix}
J
\begin{bmatrix}
\bar{x}_u \\
\bar{x}_v
\end{bmatrix},
\]

where

\[J = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.
\]

Using the chain rule, we have

\[
\begin{bmatrix}
\frac{\partial}{\partial u}, & \frac{\partial}{\partial v}
\end{bmatrix}
= \begin{bmatrix}
\frac{\partial}{\partial \bar{u}}, & \frac{\partial}{\partial \bar{v}}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial \bar{u}}{\partial u} & \frac{\partial \bar{u}}{\partial v} \\
\frac{\partial \bar{v}}{\partial u} & \frac{\partial \bar{v}}{\partial v}
\end{bmatrix} = \begin{bmatrix}
\frac{\partial}{\partial \bar{u}}, & \frac{\partial}{\partial \bar{v}}
\end{bmatrix} A^T.
\]

So in order to prove the result, we just need to prove that

\[A^T J A = \det(A) J.\]

But this is true for any 2 \times 2 matrices.

\[\text{Theorem 4.3}\]

\[\text{If } f \text{ is a real-valued function on } M, \text{ then } d(df) = 0.\]

\[\text{Proof.} \quad \text{Let } \psi = df. \text{ We need to prove } d\psi = 0. \text{ It then suffices to prove that}
\]

\[(d\psi)(x_u, x_v) = 0.
\]

But

\[(d\psi)(x_u, x_v) = \frac{\partial}{\partial u}(\psi(x_v)) - \frac{\partial}{\partial v}(\psi(x_u)) = \frac{\partial f}{\partial u} - \frac{\partial f}{\partial v} = 0.
\]

\[\text{Example 4.12}\]

Let \( f \) is a function, \( \phi \) is a 1-form, and \( \eta \) a 2-form, then

(1) \( \phi = f_1 du_1 + f_2 du_2 \), where \( f_i = \phi(U_i) \).

(2) \( \eta = g du_1 \wedge du_2 \), where \( g = \eta(U_1, U_2) \).

(3) for \( \psi = g_1 du_1 + g_2 du_2 \) and \( \phi \) as above, then

\[\phi \wedge \psi = (f_1 g_2 - f_2 g_1) du_1 \wedge du_2.
\]

(4) \( df = \frac{\partial f}{\partial u_1} du_1 + \frac{\partial f}{\partial u_2} du_2 \).

(5) \( d\psi = \left( \frac{\partial g}{\partial u_1} - \frac{\partial g}{\partial u_2} \right) du_1 \wedge du_2 \).

\[\text{Definition 4.19}\]

A differential form \( \psi \) is closed if its exterior derivative is zero, \( d\psi = 0 \); and \( \psi \) is exact if it is the exterior derivative of some form, \( \psi = d\xi \).
4.5 Mappings of Surfaces

In this section, we shall define a function from a surface to another surface. Since surfaces are metric spaces, it is not hard to define the continuity. In order to define the differentiability, we need to use coordinate patches.

**Definition 4.20**

A function $F : M \to N$ from one surface to another is **differentiable** provided that for each patch $x$ in $M$ and $y$ in $N$ the composite function $y^{-1}F x$ is Euclidean differentiable (and defined on an open set of $\mathbb{R}^2$). $F$ is then called a mapping of surfaces.

**Example 4.13** Let $\Sigma$ be the unit sphere in $\mathbb{R}^3$ (center at 0) but with north and south poles removed, and let $C$ be the cylinder based on the unit circle in the $xy$-plane. So $C$ is in contact with the sphere along the equator. We define a mapping $F : \Sigma \to C$ as follows: If $p$ is a point of $\Sigma$, draw the line orthogonally out from the $z$-axis through $p$, and let $F(p)$ be the point at which this line first meets $C$, as in Fig. 4.28. To prove that $F$ is a mapping, we use the geographical patch $x$ in $\Sigma$, and for $C$ the patch $y(u, v) = (\cos u, \sin u, v)$. Now $x(u, v) = (\cos v \cos u, \cos v \sin u, \sin v)$, so from the definition of $F$ we get $F(x(u, v)) = (\cos u, \sin u, \sin v)$.

But this point of $C$ is $y(u, \sin v)$; hence

$$F(x(u, v)) = y(u, \sin v).$$

Applying $y^{-1}$ to both sides of this equation gives

$$(y^{-1}F x)(u, v) = (u, \sin v).$$

So the function is differentiable.

**Example 4.14** Stereographic projection of the punctured sphere $\Sigma$ onto the plane. Let $\Sigma$ be a unit sphere resting on the $xy$-plane at the origin, so the center of $\Sigma$ is at $(0, 0, 1)$. Delete the north pole $n = (0, 0, 2)$ from $\Sigma$. Now imagine that there is a light source at the north pole, and for each point $p$ of $\Sigma$, let $P(p)$ be the shadow of $p$ in the $xy$-plane. As usual, we identify the $xy$-plane with $\mathbb{R}^2$ by $(p_1, p_2, 0) \leftrightarrow (p_1, p_2)$. Thus we have defined a function $P$ from $\Sigma$ onto $\mathbb{R}^2$. Evidently $P$ has the form

$$P(p_1, p_2, p_3) = \left(\frac{R p_1}{r}, \frac{R p_2}{r}\right),$$

where $r$ and $R$ are the distances from $p$ and $P(p)$, respectively, to the $z$-axis. We see that $R/2 = r/(2 - p_3)$; hence

$$P(p_1, p_2, p_3) = \left(\frac{2p_1}{2 - p_3}, \frac{2p_2}{2 - p_3}\right).$$
Now if \( x \) is any patch in \( \Sigma \), the composite function \( P(x) \) is Euclidean differentiable, so \( P : S \to \mathbb{R}^2 \) is a (differentiable) mapping.

**Definition 4.21**

Let \( F : M \to N \) be a mapping of surfaces. The tangent map \( F_\ast \) of \( F \) assigns to each tangent vector \( v \) to \( M \) the tangent vector \( F_\ast(v) \) to \( N \) such that if \( v \) is the initial velocity of a curve \( \alpha \) in \( M \), then \( F_\ast(v) \) is the initial velocity of the image curve \( F(\alpha) \) in \( N \) (Fig. 4.31).

The tangent map of a mapping \( F : M \to N \) may be computed in terms of partial velocities as follows. If \( x : D \to M \) is a parametrization in \( M \), let \( y \) be the composite mapping \( F(x) : D \to N \) (which need not be a parametrization). Obviously, \( F \) carries the parameter curves of \( x \) to the corresponding parameter curves of \( y \). Since \( F_\ast \) preserves velocities of curves, it follows at once that

\[
F_\ast(x_u) = y_u, \quad F_\ast(x_v) = y_v.
\]

Since \( x_u \) and \( x_v \) give a basis for the tangent space of \( M \) at each point of \( x(D) \), these readily computable formulas completely determine \( F_\ast \).

The discussion of regular mappings in Section 7 of Chapter 1 translates easily to the case of a mapping of surfaces \( F : M \to N \). \( F \) is regular provided all of its derivative maps \( F_\ast_p : T_p(M) \to T_{F(p)}(N) \) are one-to-one. Since these tangent planes have the same dimension, we know from linear algebra that the one-to-one requirement is equivalent to \( F_\ast \) being a linear isomorphism. A mapping \( F : M \to N \) that has an inverse mapping \( F^{-1} : N \to M \) is called a **diffeomorphism**. We may think of a diffeomorphism \( F \) as smoothly distorting \( M \) to produce \( N \). By applying the Euclidean formulation of the inverse function theorem to a coordinate expression \( y^{-1}Fx \) for \( F \), we can deduce this extension of the inverse function theorem (7.10 of Chapter 1) as follows

**Theorem 4.4**

Let \( F : M \to N \) be a mapping of surfaces, and suppose that \( F_\ast_p : T_p(M) \to T_{F(p)}(N) \) is a linear isomorphism at some one point \( p \) of \( M \). Then there exists a neighborhood \( \mathcal{U} \) of \( p \) in \( M \) such that the restriction of \( F \) to \( \mathcal{U} \) is a diffeomorphism onto a neighborhood \( V \) of \( F(p) \) in \( N \).

An immediate consequence is this useful result: A one-to-one regular mapping \( F \) of \( M \) onto \( N \) is a
diffeomorphism.

Diffeomorphisms have little respect for size or shape; here are some examples.

**Example 4.15** Any open rectangle in the plane \( \mathbb{R}^2 \) is diffeomorphic to the entire plane. Take \( R : -\pi/2 < u, v < \pi/2 \) for simplicity. Then \( F(u, v) = (\tan u, \tan v) \) is a mapping of \( R \) onto \( \mathbb{R}^2 \). Using a branch of the inverse tangent function, the mapping \( F^{-1}(u_1, v_1) = (\tan^{-1}(u_1), \tan^{-1}(v_1)) \) is a differentiable inverse of \( F \), so \( F \) is a diffeomorphism.

**Example 4.16** The sphere \( \Sigma \) minus one point is also diffeomorphic to the entire plane. Stereographic projection \( P \), as in Example 5.2(2), is a one-to-one mapping of the punctured sphere \( \Sigma_0 \) onto \( \mathbb{R}^2 \). A variant \( x(u, v) = (\cos v \cos u, \cos v \sin u, 1 + \sin v) \) of the usual geographical parametrization is a parametrization of \( \Sigma \setminus \{0\} \). The formula for \( P \) in Example 5.2 gives \( y(u, v) = P(x(u, v)) = \frac{2 \cos v}{1 - \sin v} (\cos u, \sin u) \).

**Example 4.17** A cylinder \( C \) over a closed curve is diffeomorphic to the plane minus one point. For simplicity, take \( C : x^2 + y^2 = 1 \), and define a mapping \( F : C \to \mathbb{R}^2 \) by \( F(x, y, z) = e^z (x, y) \). Since \( e^z \) takes on all values \( r > 0 \), \( F \) maps \( C \) onto \( \mathbb{R}^2 \setminus \{0\} \). For the inverse of \( F \), we have \( G(u, v) = \left( \frac{u}{\sqrt{u^2 + v^2}}, \frac{v}{\sqrt{u^2 + v^2}}, \log \sqrt{u^2 + v^2} \right) \).

We can check \( G = F^{-1} \).

**Definition 4.22** Let \( F : M \to N \) be a mapping of surfaces.

1. If \( \phi \) is a 1-form on \( N \), let \( F^*\phi \) be the 1-form on \( M \) such that \((F^*\phi)(v) = \phi(F_*v)\)

   for all tangent vectors \( v \) to \( M \).

2. If \( \eta \) is a 2-form on \( N \), let \( F^*\eta \) be the 2-form on \( M \) such that \((F^*(\eta))(v, w) = \eta(F_*v, F_*w)\)

   for all pairs of tangent vectors \( v, w \) on \( M \).

**Theorem 4.5** Let \( F : M \to N \) be a mapping of surfaces, and let \( \xi \) and \( \eta \) be forms on \( N \). Then

1. \( F^*(\xi + \eta) = F^*\xi + F^*\eta \),
2. \( F^*(\xi \wedge \eta) = F^*\xi \wedge F^*\eta \),
3. \( F^*(d\xi) = d(F^*\xi) \).
4.6 Integration of Forms

The three fundamental properties for definite integral: what is the geometric point of view to them?
1. integration by substitution
\[ \int_a^b F(\phi(t))\phi'(t)dt = \int_{\phi(a)}^{\phi(b)} F(s)ds \]
2. integration by parts
\[ \int_a^b f'(t)g(t)dt = f(t)g(t)|_a^b - \int_a^b f(t)g'(t)dt. \]
3. the Fundamental Theorem of Calculus
\[ \int_a^b F'(t)dt = F(b) - F(a). \]

First, we define the integration on curves.

One of the purposes of this section is to

**Definition 4.23**

Let \( M \) be a surface, and let \( \phi \) be a 1-form on \( M \). Let \( \alpha : [a, b] \rightarrow M \) be a curve segment of \( M \). Then the integral of \( \phi \) over \( \alpha \) is
\[ \int_{\alpha} \phi = \int_{[a,b]} \alpha^* \phi = \int_a^b \phi(\alpha'(t))dt. \]
\( \int_{\alpha} \phi \) is called the line integral of \( \phi \) over the line \( \alpha \).

**Theorem 4.6**

Let \( f \) be a function on \( M \), and let \( \alpha : [a, b] \rightarrow M \) be a curve segment in \( M \) from \( p = \alpha(a) \) to \( q = \alpha(b) \). Then
\[ \int_{\alpha} df = f(q) - f(p). \]

**Proof.** Let’s recall the notation \( \alpha^* \phi \), where \( \phi = df \). Since \( \phi \) is a 1-form on \( M \). Then by definition, \( \alpha^* \phi \) is a 1-form on the interval \([a, b]\). Now, \([a, b]\) is a 1-dimensional space. The tangent space of \([a, b] \subset \mathbb{R} \) is spanned by \( \frac{\partial}{\partial t} \) and therefore the cotangent space is spanned by \( \{dt\} \). Question, what is \( \alpha^* \phi \)?
\[ \alpha^* \phi(\frac{\partial}{\partial t}) = \phi(\alpha^* \frac{\partial}{\partial t}) = \phi(\alpha'(t)) \]
So \( \alpha^* \phi = \phi(\alpha'(t))dt. \)

By the above point of view, the definite integral can also be viewed as the line integral
\[ \int_a^b f(t)dt = \int_{[a,b]} \phi, \]
where
\[ \phi = f(t)dt \]
is a one form. Thus
\[ \int_{\alpha} df = \int_{a}^{b} \frac{d}{dt}(f(\alpha(t))) dt = f(\alpha(b)) - f(\alpha(a)) = f(q) - f(p). \]

Remark The integral \( \int_{\alpha} df \) is path independent, which will lead to the important result, the Stokes’ Theorem below.

Definition 4.24

A 2-segment is a smooth mapping \( x \) from a closed rectangle \( R \) of \( \mathbb{R}^2 \) to \( M \)
\[ x : [a, b] \times [c, d] \to M \]
In particular, even we don’t assume that the 2-segment is a regular mapping, the vector fields \( x_u, x_v \) are still well-defined.

Definition 4.25

Let \( \eta \) be a 2-form on \( M \), and let \( x : R \to M \) be a 2-segment. Then we define
\[ \int_{R} \eta = \int_{R} x^{*} \eta = \int_{a}^{b} \int_{c}^{d} \eta(x_u, x_v) du dv. \]

Definition 4.26

Let \( x : R \to M \) be a 2-segment in \( M \), with \( R \) the closed rectangle \( [a, b] \times [c, d] \) (Fig. 4.37). The edge curves of \( x \) are the curve segments \( \alpha, \beta, \gamma, \delta \) such that
\[ \alpha(u) = x(u, c), \]
\[ \beta(u) = x(b, v), \]
\[ \gamma(u) = x(u, d), \]
\[ \delta(u) = x(a, v). \]
Then the boundary \( \partial x \) of the 2-segment \( x \) is the formal expression
\[ \partial x = \alpha + \beta - \gamma - \delta. \]

We then define
\[ \int_{\partial x} \phi = \int_{\alpha} \phi + \int_{\beta} \phi - \int_{\gamma} \phi - \int_{\delta} \phi. \]
Theorem 4.7

If $\phi$ is a 1-form on $M$, and $x : R \to M$ is a 2-segment, then
\[
\int_x d\phi = \int_{\partial x} \phi.
\]

Proof. By definition, we have
\[
\int_x d\phi = \int_a^b \int_c^d d\phi(x_u, x_v) dudv.
\]
By the definition of $d\phi$, we have
\[
\int_a^b \int_c^d d\phi(x_u, x_v) dudv = \int_a^b \int_c^d \left( \frac{\partial}{\partial u}(\phi(x_v)) - \frac{\partial}{\partial v}(\phi(x_u)) \right) dudv
\]
We then have
\[
\int_a^b \int_c^d \frac{\partial}{\partial u}(\phi(x_v)) dudv = \int_c^d (\phi(x_v)|_{u=b} - \phi(x_v)|_{u=a}) dv = \int_\beta \phi - \int_\alpha \phi.
\]
Similarly, we have
\[
\int_a^b \int_c^d \frac{\partial}{\partial v}(\phi(x_u)) dudv = \int_\gamma \phi - \int_\alpha \phi.
\]
Thus the Stokes’ Theorem is proved.

Lemma 4.7

Let $\alpha(h) : [a, b] \to M$ be a reparametrization of a curve segment $\alpha : [c, d] \to M$ by $h : [a, b] \to [c, d]$.

For any 1-form $\phi$ on $M$,

1. If $h$ is orientation-preserving, that is, if $h' > 0$, then
\[
\int_{\alpha(h)} \phi = \int_{\alpha} \phi.
\]

2. If $h$ is orientation-reversing, that is, if $h' < 0$, then
\[
\int_{\alpha(h)} \phi = -\int_{\alpha} \phi.
\]

4.7 Topological properties of surfaces

Topological property is a property of a topological space which is invariant under homeomorphisms. Here we first need to define what a topological space is. For the most general definition, see here in the section of “Definition via open sets”.

For the purpose of this lecture, we only need to study a special kind of topological space: a metric space.

Definition 4.27

A metric space is a pair $(X, d)$, where $X$ is a non-empty set, and $d$ is a function $d : X \times X \to \mathbb{R}$ satisfying

1. (positivity) $d(x, y) \geq 0$ for any $x, y \in X$, and $d(x, y) = 0$ if and only if $x = y$;
2. (symmetry) $d(x, y) = d(y, x)$ for any $x, y \in X$;
4.7 Topological properties of surfaces

Remark Abusing of notation, we sometimes use \( X \) to represent the metric space \( (X, d) \).

Let \( x \in X \) and let \( r > 0 \) be a positive number. Let

\[
B_x(r) = \{ y \in X \mid d(y, x) < r \}
\]

be the ball of radius \( r \) centered at \( x \).

Definition 4.28

Let \( \mathcal{U} \subset X \). \( \mathcal{U} \) is called an open set, if for any \( x \in \mathcal{U} \), there is a positive number \( r > 0 \) such that \( B_x(r) \subset \mathcal{U} \).

Example 4.18

An open ball is an open set. Let \( B_x(r) \) be an open ball. Let \( y \in B_x(r) \). Then by the triangular inequality, we know that

\[
B_y(r - d(x, y)) \subset B_x(r).
\]

Thus \( B_x(r) \) is an open set.

Definition 4.29

Let \( X, Y \) be two metric spaces, and let \( f : X \to Y \) be a continuous mapping, that is, for any sequence \( \{x_n\} \) of \( X \) with \( d_X(x_n, x) \to 0 \), we have \( d_Y(f(x_n), f(x)) \to 0 \).

Proposition 4.1

Let \( F : X \to Y \) be a mapping. Then \( f \) is continuous if and only if for any open set \( \mathcal{V} \) in \( Y \), \( f^{-1}(\mathcal{V}) \) is open in \( X \).

Proof. We first prove the “only if” part: let \( \mathcal{V} \) be an open set and let \( x \in f^{-1}(\mathcal{V}) \). Assume that for any \( r_n \to 0 \), the balls \( B_x(r_n) \) do not completely belong to \( f^{-1}(\mathcal{V}) \). Then we can find a sequence \( y_n \in B_x(r_n) \) but \( f(y_n) \notin \mathcal{V} \). Since \( d_X(x_n, x) \to 0 \), we should have \( d(f(y_n), f(x)) \to 0 \). On the other hand, since \( \mathcal{V} \) is an open set, there must be a \( \delta > 0 \) such that \( B_{f(x)}(\delta) \subset \mathcal{V} \). This is a contradiction.

We omit the “if” part as the proof is similar.

Definition 4.30

If \( f \) is a bijection, and \( f^{-1} \) is also continuous, then we say that \( X \) and \( Y \) are homeomorphic.

Topology has been called rubber-sheet geometry. In a topology of two dimensions there is no difference between a circle and a square. A circle made out of a rubber band can be stretched into a square. There is a difference between a circle and a figure eight. A figure eight cannot be stretched into a circle without tearing.

Remark In this section, we study the following 4 topological properties.

1. Connectedness;
2. Compactness;
3. Orientability;
4. Simply-connectedness.
4.7 Topological properties of surfaces

**Definition 4.31**

A surface is connected provided that for any two points \( p, q \) of \( M \), there is a curve segment in \( M \) from \( p \) to \( q \).

*This is sometimes called path-connected*

**Example 4.19** Let \( f \) be a function on a connected surface \( M \) such that \( df = 0 \). Prove that \( f \) is a constant.

**Definition 4.32**

A surface is compact if and only if any open cover admit a finite subcover.

**Example 4.20** (Heine-Borel Theorem) Any open cover of the closed interval \([0, 1]\) by open intervals admits a finite open cover.

**Lemma 4.8**

A surface \( M \) is compact if and only if it can be covered by the images of a finite number of 2-segments in \( M \).

**Theorem 4.8**

A continuous function on a compact surface \( M \) is bounded, and moreover, the maximum is reached at some point.

**Proof.** Let \( f \) be a continuous function. For any \( x \in M \), there exists an \( r = r_x > 0 \) such that for any \( y \in B_x(r_x) \), we have

\[
|f(x) - f(y)| < 1.
\]

By definition, the collection \( \{B_x(r_x)\}_{x \in M} \) gives an open cover of \( M \).

Since \( M \) is compact, there is a finite subcover \( \{B_{x_i}(r_{x_i})\}_{1 \leq i \leq N} \) for some \( N > 0 \):

\[
\bigcup_{i=1}^{N} B_{x_i}(r_{x_i}) \supset M.
\]

Therefore,

\[
|f(x)| \leq \max(|f(x_1)|, \cdots, |f(x_N)|) + 1,
\]

which is bounded.

We claim that the maximum must exist. If not, let

\[
\alpha = \sup f(x).
\]

If \( \alpha \) is not achieved, then \( f(x) < \alpha \) is true for any \( x \in M \). As a result, the function

\[
g(x) = (\alpha - f(x))^{-1}
\]

is continuous on \( M \). Let \( \beta \) be an upper bound. Then

\[
(\alpha - f(x))^{-1} < \beta.
\]

Thus

\[
f(x) \leq \alpha - \beta^{-1}.
\]

This is a contradiction of the definition that \( \alpha \) is the least upper bound.
4.7 Topological properties of surfaces

**Definition 4.33**

A surface $M$ is orientable if there exists a differentiable (or merely continuous) 2-form $\mu$ on $M$ that is nonzero at every point of $M$.

_external_link_

The most famous non-orientable surface is the Möbius Strip. See [here](#).

**Theorem 4.9**

A surface $M \subset \mathbb{R}^3$ is orientable if and only if there exists a unit normal vector field on $M$. If $M$ is connected as well as orientable, there are exactly two unit normals, $\pm U$.

**Definition 4.34**

A closed curve $\alpha$ in $M$ is homotopic to a constant provided there is a 2-segment $x : \mathbb{R} \to M$ (called a homotopy) defined on

$$ R : a \leq u \leq b, 0 \leq v \leq 1 $$

such that $\alpha$ is the base curve of $x$ and the other three edge curves are constant at $p = \alpha(a) = \alpha(b)$.
A surface \( M \) is simply connected provided it is connected and every loop in \( M \) is homotopic to a constant.

Let \( \phi \) be a closed 1-form on a surface \( M \). If a loop \( \alpha \) in \( M \) is homotopic to a constant, then \( \int_\alpha \phi = 0 \).

On a simply connected surface, every closed 1-form is exact.

Consider a 1-form
\[
\psi = \frac{xdy - ydx}{x^2 + y^2}
\]
on \( \mathbb{R}^2 - \{(0, 0)\} \). We can check that \( d\psi = 0 \), but \( \psi \) is not exact.

A compact surface in \( \mathbb{R}^3 \) is orientable.

The proof of this theorem is beyond the scope of this course.

A simply connected surface is orientable.

4.8 Manifolds (Abstract Surfaces)

In mathematics, a manifold is a topological space that locally resembles Euclidean space near each point. More precisely, each point of an \( n \)-dimensional manifold has a neighborhood that is homeomorphic to the Euclidean space of dimension \( n \) – from Wikipedia, see here.

An abstract patch in \( M \) is just a one-to-one function \( x : D \to M \) from any open set \( D \) of \( \mathbb{R}^2 \) into the set \( M \).

A topological surface is a set \( M \) furnished with a collection \( \mathcal{P} \) of abstract patches in \( M \) satisfying
(1) The covering axiom: The images of the patches in the collection \( \mathcal{P} \) cover \( M \).
(2) The Hausdorff axiom: For any points \( p \neq q \) in \( M \) there exist disjoint (that is, nonoverlapping) patches \( x \) and \( y \) with \( p \) in \( x(D) \) and \( q \) in \( y(E) \).

On a topological surface, one can define continuous functions but not smooth functions. In order to put the smooth structure on, we make the following

A surface is a set \( M \) furnished with a collection \( \mathcal{P} \) of abstract patches in \( M \) satisfying
(1) The covering axiom: The images of the patches in the collection \( \mathcal{P} \) cover \( M \).
(2) The smooth overlap axiom: For any patches \( x, y \in \mathcal{P} \), the composite functions \( y^{-1}x \) and \( x^{-1}y \) are Euclidean differentiable—and defined on open sets of \( \mathbb{R}^2 \).
(3) The Hausdorff axiom: For any points \( p \neq q \) in \( M \) there exist disjoint (that is, non-overlapping)
patches $x$ and $y$ with $p$ in $x(D)$ and $q$ in $y(E)$.

**Example 4.22 (The real projective space)**

- First, as a set $RP^2 = S^2 / \sim$, where $\sim$ is an equivalence relation $p \sim -p$.
- Let $U_{upper}, U_{lower}, U_{front}, U_{rear}, U_{left}, U_{right}$ be the upper, lower, front, rear, left, right open hemisphere of $S^2$, and let $D$ be the open unit disk of $\mathbb{R}^2$. Let $\chi_i : D \rightarrow U_i \rightarrow S^2$

be the coordinate patch, where $i \in \{upper, lower, front, rear, left, right\}$.

- Let $\tilde{\chi}_i : D \rightarrow U_i \rightarrow S^2 \rightarrow RP^2 = S^2 / \sim$.

Then $\tilde{\chi}_i$ are one-to-one and are abstract patches.

1. Let $E$ be any open subset of the open disk $D$. Then $\tilde{\chi}_{i,E} : E \rightarrow D \rightarrow S^2 \rightarrow RP^2 = S^2 / \sim$ is an abstract coordinate patch.

2. Using the above, we can prove that $RP^2$ is a topological surface. It is a surface because the transition functions are smooth.

**Remark** One can generalize the above concept to $n$-dimensional manifold.

**Definition 4.39**

Let $\alpha : I \rightarrow M$ be a curve in an abstract surface $M$. $\alpha(0) = p$ for a fixed point $p \in M$. For each $t$ in $I$ the velocity vector $\alpha'(t)$ is the function such that

$$\alpha'(t)[f] = \frac{d(f\alpha)}{dt}(t)$$

for every differentiable real-valued function $f$ on $M$. The set of all initial velocity vectors $\alpha'(0)$ form the tangent space at $p$.

Alternatively, we can define tangent space as the dual space of the cotangent space.

**Definition 4.40**

Let $p \in M$. Let $\mathcal{F}_p$ be the space of smooth functions which vanishes at $p$. Then we are above to define the cotangent space by

$$T^*_pM = \frac{\mathcal{F}_p}{\mathcal{F}_p^2}$$
Chapter 5 Shape Operators

5.1 The shape operator of \( M \subset \mathbb{R}^3 \)

In \( \mathbb{R}^3 \), let \( X, Y \) be two vector fields. Then in § 2.7, we defined the covariant derivatives. Namely, we have

\[
\nabla_X Y = XY.
\]

To recap, what we really mean is that, if \( Y = (a, b, c) \), or more precisely, if

\[
Y = a \frac{\partial}{\partial x} + b \frac{\partial}{\partial y} + c \frac{\partial}{\partial z},
\]

then we have

\[
\nabla_X Y = X(a) \frac{\partial}{\partial x} + X(b) \frac{\partial}{\partial y} + X(c) \frac{\partial}{\partial z}.
\]

Now let’s assume that \( M \subset \mathbb{R}^3 \) be a surface. Let \( U \) be a unit normal vector field of \( M \).

**Definition 5.1**

Let \( p \in M \). Then the shape operator is defined by

\[
S_p(v) = -\nabla v U,
\]

where \( v \in T_p M \).

**Lemma 5.1**

The shape operator is a linear operator

\[
S_p : T_p M \to T_p M.
\]

**Proof.** Since \( U \cdot U = 1 \), we have

\[
0 = v[U \cdot U] = 2\nabla v U \cdot U = -2S_p(v) \cdot U.
\]

Thus the image of \( S_p \) is in \( T_p M \). The linearity of the operator is obvious.

**Lemma 5.2**

The shape operator is a symmetric operator

\[
S_p : T_p M \to T_p M.
\]

**Proof.** An operator is symmetric, if for any \( v, w \), we have

\[
S_p(v) \cdot w = v \cdot S_p(w).
\]
5.2 Normal Curvature

Since \( U \cdot v = U \cdot w = 0 \), we have

\[
0 = w[U \cdot v] = -S_p(w) \cdot v + U \cdot \nabla_w v,
\]

\[
0 = v[U \cdot w] = -S_p(v) \cdot w + U \cdot \nabla_v w.
\]

Later we shall prove that \( U \cdot (\nabla_w v - \nabla_v w) = 0 \). The lemma is proved.

**Example 5.1** What is the shape operator of \( \mathbb{R}^2 \subset \mathbb{R}^3 \)? It is zero, because \( U \) is parallel.

**Example 5.2** The shape operator of the sphere of radius \( r: x^2 + y^2 + z^2 = r^2 \).

The unit normal vector field is

\[
U = \frac{1}{r} \sum_i x_i U_i.
\]

Let

\[
v = v_i U_i.
\]

Then

\[
\nabla_v U = \frac{1}{r} v.
\]

Thus \( S_p(v) = -v/r \).

**Example 5.3** The shape operator of the cylinder of radius \( r: x^2 + y^2 = r^2 \).

### 5.2 Normal Curvature

**Lemma 5.3**

543 Let \( \alpha \) be a curve in \( M \subset \mathbb{R}^3 \). Then

\[
\alpha'' \cdot U = S(\alpha') \cdot \alpha'.
\]

**Proof.** We have \( \alpha' \cdot U = 0 \). Then taking derivative on both sides, we have

\[
0 = \alpha'' \cdot U + \alpha' \cdot U' = \alpha'' \cdot U - \alpha' \cdot S(\alpha').
\]

The lemma is proved.

**Definition 5.2**

Let \( v \) be a unit vector field on a surface \( M \). Then the normal curvature is defined to be

\[
k(v) = S(v) \cdot v.
\]

88
One obvious observation is that

\[ k(-v) = k(v). \]

Let \( \alpha \) be a unit-speed curve on \( M \). When regarding it as a space curve, its Frenet apparatus is given by \( (T, N, B) \), where

\[ T = \alpha'(t). \]

We therefore have

\[ k(\alpha'(t)) = S(\alpha'(t)) \cdot \alpha'(t) = \alpha'' \cdot U = \kappa(t) \cdot U = k(t) \cos \theta, \]

where \( \kappa \) is the curvature and \( \theta \) is the angle between the normal vector of the surface and the principal normal of the space curve \( \alpha \).

**Theorem 5.1**

There is a one-to-one correspondence between a symmetric operator and its corresponding quadratic form.

**Definition 5.3**

The eigenvalues of the shape operator is called the **principal curvatures** of the surface. The corresponding eigenvectors are called **principal vectors**.

This definition is different from the textbook.

**Theorem 5.2**

The eigenvalues are the maximum and minimum of the normal curvature \( k(v) \).

**Proof.** Let \( k_1, k_2 \) be the principal curvatures, and let \( v, w \) be the corresponding principal vectors. Without loss of generality, we assume that \( v, w \) forms a frame. Let

\[ u = av + bw \]

for real numbers \( a, b \) with \( a^2 + b^2 = 1 \). We compute

\[ k(u) = S(u) \cdot u = k_1a^2 + k_2b^2. \]

By the Lagrange multiplier’s method, we know that \( k_1, k_2 \) are the maximum and minimum values of the normal curvature.

**Definition 5.4**

A point \( p \) is called **umbilic**, if the two principal curvatures are the same.

**Theorem 5.3**

A point is umbilic, if and only if the shape operator is a scalar multiplication.

**Proof.** The only if part is trivial. In order to prove the if part, we shall observe that if a symmetric matrix having the same eigenvalue, then it is a scalar multiplication.
Corollary 5.1

Let $k_1, k_2$ and $e_1, e_2$ be the principal curvatures and vectors of $M \subset \mathbb{R}^3$ at $p$. Then if $u = \cos \theta \, e_1 + \sin \theta \, e_2$, the normal curvature of $M$ in the $u$ direction is

$$k(u) = k_1 \cos^2 \theta + k_2 \sin^2 \theta.$$  

5.3 Gauss Curvature

Let $T$ be a symmetric linear transformation from $\mathbb{R}^2$ to itself. What are the invariants of $T$?

Answer: $\text{Tr}(T)$ and $\det(T)$.

Justification: Let $A, B$ be the matrix representations of $T$ under different bases. Then $A$ and $B$ are similar matrices, that is,

$$A = PBP^{-1}.$$  

Therefore $\det(A) = \det(B)$ and hence $\det(T) = \det(A)$ is an invariant. The trace $\text{Tr}(A)$ is also an invariant, defined by

$$\text{Tr}(A) = a + b,$$

where the matrix $A$ can be expressed by

$$A = \begin{bmatrix} a & d \\ c & b \end{bmatrix}.$$  

Definition 5.5

Let $S$ be the shape operator. The Gauss curvature $K$ is defined by

$$K = \det(S) = k_1 k_2,$$

where $\lambda_1, \lambda_2$ are the principal curvatures. The mean curvature is defined as

$$H = \frac{1}{2} \text{Tr}(S) = \frac{1}{2} (k_1 + k_2).$$  

Remark The Gauss curvature is independent of the choice of orientation. The mean curvature does depend on the orientation.

Remark The sign of Gaussian curvature at a point $p$.

Case 1. $K(p) > 0$. If the Gauss curvature at one point is positive, then both of the principal curvatures are positive or negative. By changing the normal vector $U$ to $-U$ if necessary, we may assume that both principal curvatures are negative. Let $\alpha_1$ be a curve passing through $p$ with the direction as the principal direction. Then by the relationship between the normal curvature and the principal normal vector of $\alpha_1$ as a curve in the 3-dimensional space, we know that the angle between them is $< \pi/2$. Thus in both eigenvector directions, the surface is bending away from the normal vector. Thus the surface locally looks like a paraboloid (See Fig. 5.15).

$$z = \frac{1}{2} k_1 x^2 + \frac{1}{2} k_2 y^2.$$
Case 2. $K(p) < 0$ — similar discussion. See Fig. 5.16.

Case 3. $K(p) = 0$. There are two cases, either one principal curvature is zero and the other one is non-zero, in which the shape looks like Fig. 5.17, or both principal curvatures are zero, in which it looks close to a flat plane $\mathbb{R}^2$ in $\mathbb{R}^3$.

**Example 5.4** Compute the Gauss curvature (at $(0, 0)$) of the surface

$$z = \frac{1}{2} \lambda_1 u^2 + \frac{1}{2} \lambda_2 v^2.$$  

We have

$$\mathbf{x}_u = (1, 0, \lambda_1 u), \quad \mathbf{x}_v = (1, 0, \lambda_2 v)$$

at $(0, 0)$. The unit normal vector is given by

$$U = \left( \frac{-\lambda_1 u}{\sqrt{1 + \lambda_1^2 u^2 + \lambda_2^2 v^2}}, \frac{-\lambda_2 v}{\sqrt{1 + \lambda_1^2 u^2 + \lambda_2^2 v^2}}, \frac{1}{\sqrt{1 + \lambda_1^2 u^2 + \lambda_2^2 v^2}} \right).$$

At the origin, $\mathbf{x}_u = (1, 0, 0)$ and $\mathbf{x}_v = (0, 1, 0)$, and we have

$$S(\mathbf{x}_u) = (\lambda_1, 0, 0) = \lambda_1 \mathbf{x}_u, \quad S(\mathbf{x}_v) = (0, \lambda_2, 0) = \lambda_2 \mathbf{x}_v$$

So the principal curvatures are given by $\lambda_1, \lambda_2$, and the Gauss curvature $K$ at the origin is $K = \lambda_1 \lambda_2$.

For the rest of the sections, we introduce some formulae to compute the Gauss curvature, using the shape operator. We begin with

**Lemma 5.4**

*If $\mathbf{v}$ and $\mathbf{w}$ are linearly independent tangent vectors at a point $p$ of $M \subset \mathbb{R}^3$, then*

$$S(\mathbf{v}) \times S(\mathbf{w}) = K(p) \mathbf{v} \times \mathbf{w},$$

$$S(\mathbf{v}) \times \mathbf{w} + \mathbf{v} \times S(\mathbf{w}) = 2H(p) \mathbf{v} \times \mathbf{w}.$$
5.4 Computational Techniques

**Proof.** This is a general fact about the $2 \times 2$ linear operators.

**Corollary 5.2**

*On an oriented region $\mathcal{O}$ in $M$, the principal curvature functions are given by*

$$k_1, k_2 = H \pm \sqrt{H^2 - K}.$$

**Definition 5.6**

*A surface $M$ in $\mathbb{R}^3$ is flat provided its Gaussian curvature is zero, and minimal provided its mean curvature is zero.*

### 5.4 Computational Techniques

Let $M$ be a surface with coordinate patch $\mathbf{x} : \mathcal{O} \to M$.

Let $U$ be the unit norm of $M$. Let $p \in M$. Then the shape operator $S$ is defined by

$$S : T_p M \to T_p M, \quad S(v) = -\nabla_v U.$$

The shape operator is a symmetric operator. The corresponding quadratic form is called the normal curvature of the surface. The eigenvalues of the shape operator are called principal curvatures, the average of the principal curvatures is called the mean curvature, and the multiplication of the principal curvatures is called the Gauss curvature.

Define

$$E = \mathbf{x}_u \cdot \mathbf{x}_u, \quad F = \mathbf{x}_u \cdot \mathbf{x}_v, \quad G = \mathbf{x}_v \cdot \mathbf{x}_v.$$

Then $\{E, F, G\}$ determines the dot product, namely, if we write

$$\mathbf{v} = v_1 \mathbf{x}_u + v_2 \mathbf{x}_v, \quad \mathbf{w} = w_1 \mathbf{x}_u + w_2 \mathbf{x}_v,$$

then

$$\mathbf{v} \cdot \mathbf{w} = Ev_1 w_1 + F(v_1 w_2 + v_2 w_1) + Gv_2 w_2.$$

In particular, by the Cauchy inequality, we have $EG - F^2 > 0$. A even more refined version of the Cauchy inequality reveals that

$$\|\mathbf{x}_u \times \mathbf{x}_v\|^2 = EG - F^2 > 0.$$  

As is well-known, the unit normal vector field $U$ is given by

$$U = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|}.$$

Now we define the second derivatives. Let

$$\mathbf{x}_{uu} = \frac{\partial \mathbf{x}_u}{\partial u}, \quad \mathbf{x}_{uv} = \frac{\partial \mathbf{x}_u}{\partial v}, \quad \mathbf{x}_{vv} = \frac{\partial \mathbf{x}_v}{\partial u}, \quad \mathbf{x}_{vv} = \frac{\partial \mathbf{x}_v}{\partial v}.$$  

Let

$$L = S(\mathbf{x}_u) \cdot \mathbf{x}_u, \quad M = S(\mathbf{x}_u) \cdot \mathbf{x}_v = S(\mathbf{x}_v) \cdot \mathbf{x}_u, \quad N = S(\mathbf{x}_v) \cdot \mathbf{x}_v.$$  

Then
Lemma 5.5

We have

\[ L = U \cdot x_{uu}, \quad M = U \cdot x_{uv}, \quad N = U \cdot x_{vv}. \]

Proof. From \( x_u \cdot U = 0 \) and taking derivative with respect to \( u \), we have

\[ x_{uu} \cdot U + x_u \cdot \nabla x_u U = 0. \]

We therefore have \( L = U \cdot x_{uu} \). The other two equations can be proved similarly.
Theorem 5.4

The principal curvatures are the solutions of the following quadratic equation

$$\det \begin{bmatrix} L & M \\ M & N \end{bmatrix} - \lambda \begin{bmatrix} E & F \\ F & G \end{bmatrix} = 0$$

Proof. Let $\lambda$ be a principal curvature and let $v = ax_u + bx_v$ be a principal vector. Then

$$La + Mb = S(x_u)(ax_u + bx_v) = S(x_u) \cdot v = x_u \cdot \lambda v = \lambda(Ea + Fb).$$

Similarly, we have

$$Ma + Nb = \lambda(Fa + Gb).$$

Thus

$$\left( \begin{bmatrix} L & M \\ M & N \end{bmatrix} - \lambda \begin{bmatrix} E & F \\ F & G \end{bmatrix} \right) \begin{bmatrix} a \\ b \end{bmatrix} = 0.$$

The theorem is proved.

Theorem 5.5

We have

$$K = \frac{LN - M^2}{EG - F^2}, \quad H = \frac{GL + EN - 2FM}{2(EG - F^2)}.$$

Example 5.5 (Helicoid) The helicoid is defined by

$$x(u, v) = (u \cos v, u \sin v, bv), \quad b \neq 0.$$

Its Gauss curvature is

$$K = -\frac{b^2}{(b^2 + v^2)^2} < 0$$

and the mean curvature $H = 0$.

Example 5.6 (Saddle Surface) The saddle surface $M : z = xy$. We have

$$K = -\frac{1}{(1 + u^2 + v^2)^2}, \quad H = -\frac{uv}{(1 + u^2 + v^2)^{3/2}}.$$

5.5 The Implicit Case

Let $V, W$ be two linearly independent vector fields. Let

$$S(V) = aV + bW, \quad S(W) = cV + dW$$

Then the Gauss curvature is given by $ad - bc$, and the mean curvature is given by $\frac{1}{2}(a + d)$.

Let $M$ be defined by a single equation $g = g(x, y, z) = 0$. By analytic geometry, the norm vector field is defined by

$$Z = \nabla g,$$

where $\nabla g$ is called gradient, given by

$$\nabla g = \sum_{i=1}^{3} \frac{\partial g}{\partial x_i} U_i = g_i U_i.$$
Lemma 5.6
Assume that $V \times W = Z$. Then
\[
K = \frac{Z \cdot \nabla_V Z \times \nabla_W Z}{\|Z\|^4}
\]
\[
H = -Z \cdot \frac{\nabla_V Z \times W + V \times \nabla_W Z}{2\|Z\|^3}
\]

Example 5.7 (Ellipsoid)
\[
M : g = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.
\]
We have
\[
K = \frac{1}{a^2 b^2 c^2 \|Z\|^3} > 0.
\]

Example 5.8 (Scherk minimal surface, Problem 5, §5.5)
\[
M : e^x \cos x = \cos y
\]
It is a minimal surface. So $H = 0$.

Note If a surface is defined by $F(x, y, z) = 0$. Then its Gauss Curvature is given by
\[
K = -\frac{\text{det} \begin{bmatrix}
F_{xx} & F_{xy} & F_{xz} & F_x \\
F_{xy} & F_{yy} & F_{yz} & F_y \\
F_{xz} & F_{yz} & F_{zz} & F_z \\
F_x & F_y & F_z & 0
\end{bmatrix}}{|\nabla F|^4}
\]

See here, the section of alternative formula.

5.6 Special Curves in a Surface

Definition 5.7
A regular curve $\alpha$ in $M \subset \mathbb{R}^3$ is a principal curve provided that the velocity $\alpha'$ of $\alpha$ always points in a principal vector.

Lemma 5.7
Let $\alpha$ be a regular curve in $M \subset \mathbb{R}^3$, and let $U$ be a unit normal vector field restricted to $\alpha$. Then
1. the curve $\alpha$ is principal if and only if $U'$ and $\alpha'$ are collinear at each point.
2. if $\alpha$ is a principal curve, then the principal curvature of $M$ in the direction of $\alpha'$ is $\alpha'' \cdot U / \alpha' \cdot \alpha'$.

Lemma 5.8
Let $\alpha$ be a curve cut from a surface $M \subset \mathbb{R}^3$ by a plane $P$. If the angle between $M$ and $P$ is constant along $\alpha$, then $\alpha$ is a principal curve of $M$.

Proof. Let $U$ and $V$ be unit normal vector fields to $M$ and $P$ (respectively) along the curve $\alpha$, as shown in Fig. 5.29. Since $P$ is a plane, $V$ is parallel, that is, $V' = 0$. The constant-angle assumption
means that \( U \cdot V \) is constant; thus

\[
0 = (U \cdot V)' = U' \cdot V.
\]

Since \( U \) is a unit vector, \( U' \) is orthogonal to \( U \) as well as to \( V \). The same is of course true of \( \alpha' \) since \( \alpha \) lies in both \( M \) and \( P \). If \( U \) and \( V \) are linearly independent (as in Fig. 5.29) we conclude that \( U' \) and \( \alpha' \) are collinear; hence \( \alpha \) is principal.

However, linear independence fails only when \( U = \pm V \). But then \( U' = 0 \), so \( \alpha \) is (trivially) principal in this case as well.

Directions tangent to \( M \subset \mathbb{R}^3 \) in which the normal curvature is zero are called asymptotic directions.

**Lemma 5.9**

Let \( p \in M \).

1. if \( K(p) > 0 \), there is no asymptotic directions at \( p \);
2. if \( K(p) < 0 \), there are exactly two asymptotic directions at \( p \), and these are bisected by the principal directions at angle \( \theta \) such that

\[
\tan^2 \theta = \frac{-\lambda_1(p)}{\lambda_2(p)}.
\]

3. if \( K(p) = 0 \), then if two principal curvatures are zero, then every direction is asymptotic; otherwise, only one direction is asymptotic.

**Definition 5.8**

A ruled surface is swept out by a line moving through \( \mathbb{R}^3 \).

**Lemma 5.10**

A ruled surface \( M \) has Gauss curvature \( K \leq 0 \). Moreover, \( K = 0 \) if and only if the unit normal \( U \) is parallel along each ruling of \( M \).

### 5.7 Surface of Revolution

Surface of revolution provide rich examples of surfaces for us to test general theorems. On the other hand, many surfaces appeared naturally have some kind of symmetry, including the case of rotational symmetry. Moreover, since the surface of revolution depends on the profile curve, which can be described by function of one variable, prescribing curvature would result in ODE.

Let \( \alpha \) be a curve defined by the parametric equations

\[
x = g(u), \quad y = h(u) > 0,
\]

or in other word, \( \alpha(u) = (g(u), h(u), 0) \). Define a surface \( M \) of revolution

\[
x(u, v) = (g(u), h(u) \cos v, h(u) \sin v).
\]

We then have

\[
x_u = (g', h' \cos v, h' \sin v);
\]

\[
x_v = (0, -h \sin v, h \cos v).
\]
Hence
\[ E = (g')^2 + (h')^2, \quad F = 0, \quad G = h^2. \]

Thus
\[ x_u \times x_v = (hh', -hg' \cos v, -hg' \sin v); \]
\[ \| x_u \times x_v \| = \sqrt{EG - F^2} = h\sqrt{(g')^2 + (h')^2}; \]
\[ U = \frac{1}{\sqrt{(g')^2 + (h')^2}}(h', -g' \cos v, -g' \sin v). \]

To continue, we have
\[ x_{uu} = (g''', h''' \cos v, h''' \sin v); \]
\[ x_{uv} = (0, -h'' \sin v, h'' \cos v); \]
\[ x_{vv} = (0, -h \cos v, -h \sin v). \]

Then
\[ L = -g'' h' + g' h'' \sqrt{(g')^2 + (h')^2}, \quad M = 0, \quad N = \frac{g' h}{\sqrt{(g')^2 + (h')^2}}. \]

The fact that \( F = M = 0 \) greatly simplifies our computation. We then have
\[ S(x_u) = \frac{L}{E} x_u, \quad S(x_v) = \frac{N}{G} x_v. \]

Thus the two principal curvatures are given by
\[ k_\mu = \frac{N}{G} = \frac{g'}{h((g')^2 + (h')^2)^{1/2}}, \]
\[ k_\pi = \frac{L}{E} = \frac{-g'' h' - g' h''}{((g')^2 + (h')^2)^{3/2}}. \]

In particular, if we assume that \( g(u) = u \), then
\[ k_\mu = \frac{-h''}{(1 + (h')^2)^{3/2}}, \quad k_\pi = \frac{1}{h(1 + (h')^2)^{1/2}}, \quad K = \frac{-h''}{h(1 + (h')^2)^2}. \]

**Example 5.9 (Torus of revolution)**

We rotate the the circle
\[ g(u) = r \sin u, \quad h(u) = R + r \cos u \]
for constants \( 0 < r < R \) to obtain
\[ x(u, v) = (r \sin u, (R + r \cos u) \cos v, (R + r \cos u) \sin v) \]

We then have
\[ E = r^2, \quad F = 0, \quad G = (R + r \cos u)^2 \]
\[ L = r, \quad M = 0, \quad N = (R + r \cos u) \cos u \]

Thus
\[ k_\mu = \frac{1}{r}, \quad k_\pi = \frac{\cos u}{R + r \cos u}, \quad K = \frac{\cos u}{r(R + r \cos u)}. \]

**Example 5.10 (Catenoid)**

We rotate the Catenary
\[ y = c \cosh(x/c) \]
to obtain the Catenoid

\[ \mathbf{x}(u, v) = (u, c \cosh(u/c) \cos v, c \cosh(u/c) \sin v). \]

We can compute

\[ \mathbf{x}_u = (1, \sinh(u/c) \cos v, \sinh(u/c) \sin v); \]
\[ \mathbf{x}_v = (0, -c \cosh(u/c) \sin v, c \cosh(u/c) \cos v). \]

So

\[ E = \cosh^2(u/c), \quad F = 0, \quad G = c^2 \cosh^2(u/c). \]

For the second derivatives, we have

\[ \mathbf{x}_{uu} = \left( 0, \frac{1}{c} \cosh(u/c) \cos v, \frac{1}{c} \cosh(u/c) \sin v \right); \]
\[ \mathbf{x}_{uv} = \left( 0, -\sinh(u/c) \sin v, \sinh(u/c) \cos v \right); \]
\[ \mathbf{x}_{vv} = \left( 0, -c \cosh(u/c) \cos v, -c \cosh(u/c) \sin v \right). \]

We also have

\[ \mathbf{U} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|} = (\tanh(u/c), -\text{sech}(u/c) \cos v, -\text{sech}(u/c) \sin v). \]

Thus we have

\[ L = -\frac{1}{c}, \quad M = 0, \quad N = c. \]

We then have \( H = 0 \) and

\[ K = -\frac{1}{c^2 \cosh^4(u/c)}. \]

The following result characterizes Catenoid.

**Theorem 5.6**

*If a surface of revolution \( M \) is a minimal surface, then \( M \) is part of a plane or a catenoid.*

**Proof.** We shall renormalizing the profile curve

\[ \alpha(u) = (g(u), h(u)) \]

to be of the unit speed. That is

\[ (g')^2 + (h')^2 = 1. \]

If \( g'(u) \) is identically zero, then \( g(u) \equiv c \). Thus the service is a plane.

If \( g'(u) \) is not zero, then we can prove that \( g'(u) \) is never zero. We then reparametrilize the surface.
such that the profile curve is \( \alpha(u) = (u, h(u)) \).

Thus the surface is defined by
\[
\mathbf{x}(u, v) = (u, h(u) \cos v, h(u) \sin v).
\]

If the mean curvature is zero, then we have
\[
-\frac{h''}{(1 + (h')^2)^{3/2}} + \frac{1}{h(1 + (h')^2)^{1/2}} = 0,
\]
which can be simplified as
\[
hh'' = 1 + (h')^2.
\]

We take derivative on both sides,
\[
hh''' + hh''h' = 2hh''.
\]

Thus we have
\[
(\log hh'/h)' = 0.
\]

So we have
\[
h'' = ch.
\]

Thus
\[
h(u) = a \cosh(\sqrt{c}u + b)
\]
is the solution, where \( a, b \) are constants.

**Lemma 5.11**

For a canonical parametrization of a surface of revolution,
\[
E = 1, \quad F = 0, \quad G = h^2,
\]
and the Gaussian curvature is
\[
K = \frac{h''}{h}.
\]

**Proof.** For canonical parametrization, we have \((g')^2 + (h')^2 = 1\). Since
\[
K = -\frac{g'(g'h'' - h'g'')}{h},
\]
Since \(g'g'' + h'h'' = 0\), the above formula can be simplified by \( K = -h''/h \).

**Example 5.11** (Canonical parametrization of the catenoid \((c = 1)\))
An arc length function for the catenary \( \alpha(u) = (u, \cosh u) \) is \( s(u) = \sinh u \). Hence a unit-speed reparametrization is
\[
\beta(s) = (g(s), h(s)) = (\sinh^{-1} s, \sqrt{1 + s^2}).
\]

**Example 5.12** (Surfaces of revolution with constant positive curvature)
We shall classify surfaces of revolution with constant positive curvature \( 1/c^2 \).
Atractrix (from the Latin verb trahere "pull, drag"; plural: tractrices) is the curve along which an object moves, under the influence of friction, when pulled on a horizontal plane by a line segment attached to a tractor (pulling) point that moves at a right angle to the initial line between the object and the puller at an infinitesimal speed.

The differential equation of tractrix is

$$h' = \frac{-h}{\sqrt{c^2 - h^2}}.$$

Parametric expression is given by (when $c = 1$)

$$x = t - \tanh t, \quad y = \text{sech } t.$$

Rotating a tractrix we would get Pseudo sphere, or bugle surface, or tractroid. The curvature of the surface is $-1/c^2$. 
Chapter 6  Geometry of Surfaces in $\mathbb{R}^3$

Recall that a Euclidean frame field or a frame field, constitutes vector fields $\{E_1, E_2, E_3\}$ such that

$$\mathbf{E}_i \cdot \mathbf{E}_j = \delta_{ij}.$$  

Let $\{\theta_1, \theta_2, \theta_3\}$ be the dual frame, that is

$$\theta_i(\mathbf{E}_j) = \delta_{ij}.$$  

The connection 1-forms $\omega_{ij}$ are defined by

$$\nabla_v \mathbf{E}_i = \omega_{ij}(v) \mathbf{E}_j$$

for any vector $v$. We then have the two structural equations:

$$d\theta_i = \omega_{ij} \wedge \theta_j,$$

$$d\omega_{ij} = \omega_{ik} \wedge \omega_{kj}.$$  

6.1 The Fundamental Equations

**Definition 6.1**

A adapted frame field $\{E_1, E_2, E_3\}$ on a region $\mathcal{O}$ in $M \subset \mathbb{R}^3$ is a Euclidean frame field such that $E_3$ is always normal to $M$ (hence $E_1, E_2$ are tangent to $M$).

**Lemma 6.1**

There is an adapted frame field on a region $\mathcal{O}$ in $M \subset \mathbb{R}^3$ if and only if $\mathcal{O}$ is orientable and there exists a nonvanishing tangent vector field on $\mathcal{O}$.

**Example 6.1** (Cylinder) $M : x^2 + y^2 = r^2$.

We can take

$$E_1 = U_3,$$

$$E_2 = \frac{-yU_1 + xU_2}{r},$$

$$E_3 = \frac{xU_1 + yU_2}{r},$$

as a adapted frame field.
Example 6.2 (Sphere) $M : x^2 + y^2 + z^2 = r^2$.

We can take

$$E_3 = \frac{xU_1 + yU_2 + zU_3}{r}.$$ 

Let $V = -yU_1 + xU_2$ be a vector field. We thus have

$$E_1 = \frac{V}{\|V\|},$$

$$E_2 = E_3 \times E_1,$$

$$E_3 = \frac{xU_1 + yU_2 + zU_3}{r}.$$ 

Corollary 6.1

Let $\{E_1, E_2, E_3\}$ be an adapted frame. Then

$$S(v) = \omega_{13}(v)E_1 + \omega_{23}(v)E_2$$

for any vector $v$.

Definition 6.2

Let $\theta_1, \theta_2$ be the dual 1-forms of the frame $\{E_1, E_2\}$.

Theorem 6.1

Let $\{E_1, E_2, E_3\}$ be an adapted frame field. Then its dual forms and connection forms satisfy

$$\begin{cases} d\theta_1 = \omega_{12} \wedge \theta_2 \\ d\theta_2 = \omega_{21} \wedge \theta_1 \end{cases}$$

First structural equations

$$\omega_{31} \wedge \theta_1 + \omega_{32} \wedge \theta_2 = 0$$

Symmetry equation

$$d\omega_{12} = \omega_{13} \wedge \omega_{32}$$

Gauss equation

$$\begin{cases} d\omega_{13} = \omega_{12} \wedge \omega_{23} \\ d\omega_{23} = \omega_{21} \wedge \omega_{13} \end{cases}$$

Codazzi equations

Note This is difficult and important.

6.2 Form Computations

We first have the following linear algebraic results.
6.2 Form Computations

**Lemma 6.2**

Let $\phi$ be a 1-form and let $\mu$ be a 2-form. Then

$$
\phi = \phi(E_1) \theta_1 + \phi(E_2) \theta_2; \\
\mu = \mu(E_1, E_2) \theta_1 \wedge \theta_2.
$$

Using the definition of connection 1-forms $\omega_{ij}$, we have

**Lemma 6.3**

We have

$$
\omega_{13} \wedge \omega_{23} = K \theta_1 \wedge \theta_2; \\
\omega_{13} \wedge \theta_2 + \theta_1 \wedge \omega_{23} = 2H \theta_1 \wedge \theta_2.
$$

**Proof.** We can write

$$
S(E_1) = -\nabla_{E_1} E_3 = -(\nabla_{E_1} E_3 \cdot E_1) E_1 - (\nabla_{E_1} E_3 \cdot E_2) E_2.
$$

By the definition of the connection 1-form,

$$
S(E_1) = -\omega_{31}(E_1) E_1 - \omega_{32}(E_1) E_2.
$$

Similarly, we have

$$
S(E_2) = -\omega_{31}(E_2) E_1 - \omega_{32}(E_2) E_2.
$$

We can check

$$
\omega_{13} \wedge \omega_{23}(E_1, E_2) = \omega_{13}(E_1) \omega_{23}(E_2) - \omega_{13}(E_2) \omega_{23}(E_1) = K.
$$

On the other hand,

$$
K \theta_1 \wedge \theta_2(E_1, E_2) = K.
$$

This completes the proof.
Corollary 6.2

Gauss proved that
\[ d\omega_{12} = -K \theta_1 \land \theta_2. \]

Remark Gauss’s Theorema Egregium is a major result of differential geometry (proved by Carl Friedrich Gauss in 1827) that concerns the curvature of surfaces. It marked the birth of Differential Geometry. See the Wikipedia for details.

Definition 6.3

A principal frame field on \( M \subset \mathbb{R}^3 \) is an adapted frame field \( \{E_1, E_2, E_3\} \) such that at each point \( E_1, E_2 \) are the principal vectors of \( M \).

Lemma 6.4

If \( p \) is a non-umbilic of \( M \subset \mathbb{R}^3 \), then there exists a principal frame field on some neighborhood of \( p \) in \( M \).

Let \( k_1, k_2 \) be the principal curvatures. If \( \{E_1, E_2, E_3\} \) is a principal frame field on \( M \), then we have
\[
\begin{align*}
\omega_{13}(E_1) &= k_1, & \omega_{13}(E_2) &= 0; \\
\omega_{23}(E_1) &= 0, & \omega_{23}(E_2) &= k_2.
\end{align*}
\]

Therefore,
\[
\begin{align*}
\omega_{13} &= k_1 \theta_1, & \omega_{23} &= k_2 \theta_2,
\end{align*}
\]

Theorem 6.2

Assume that \( k_1, k_2 \) are smooth functions. In particular, if \( k_1 \neq k_2 \), then \( k_1, k_2 \) are smooth functions. Then
\[
\begin{align*}
E_1(k_2) &= (k_1 - k_2) \omega_{12}(E_2); \\
E_2(k_1) &= (k_1 - k_2) \omega_{12}(E_1).
\end{align*}
\]

Proof. We have two ways to compute \( d(k_1 \theta_1) \): first
\[
d(k_1 \theta_1) = d\omega_{13} = \omega_{12} \land \omega_{23} = k_2 \omega_{12} \land \theta_2.
\]

On the other hand,
\[
d(k_1 \theta_1) = dk_1 \land \theta_1 + k_1 d\theta_1 = dk_1 \theta_1 + k_1 \omega_{12} \land \theta_2.
\]
6.3 Some Global Theorems

The first 3 results are not global results.

**Theorem 6.3**

If the shape operator of a surface \( M \) is identically zero, then \( M \) is part of a plane in \( \mathbb{R}^3 \).

**Proof.** Since \( \omega_{13} = k_1 \theta_1 \), we know that \( k_1 \) is a smooth function. By Theorem ??, \( dk_1 = 0 \), and hence \( k_1 \) is a constant. Thus the surface has constant nonnegative Gauss curvature since \( K = k_1^2 \geq 0 \).

**Theorem 6.4**

If \( M \subset \mathbb{R}^3 \) is all-umbilic and \( K > 0 \), then \( M \) is part of a sphere in \( \mathbb{R}^3 \) of radius \( 1/\sqrt{K} \).

**Proof.** Pick at random a point \( p \) in \( M \) and a unit normal vector \( E_3(p) \) to \( M \) at \( p \). Define

\[
    c = p + \frac{1}{k(p)} E_3(p).
\]

Then

\[
    \nabla_{E_i} c = \nabla_{E_i} \left( p + \frac{1}{k(p)} E_3(p) \right) = 0
\]

for \( i = 1, 2 \). Thus \( c \) is a constant vector. We thus have

\[
    \|p - c\| = \frac{1}{k(p)} = \frac{1}{\sqrt{K}}.
\]

Thus \( M \) is part of the sphere.

The following is a global result.

**Corollary 6.3**

A compact all-umbilic surface \( M \) in \( \mathbb{R}^3 \) is an entire sphere.

**Theorem 6.5**

On every compact surface \( M \) in \( \mathbb{R}^3 \) there is a point which the Gaussian curvature \( K \) is strictly positive.

**Lemma 6.6**

Let \( m \) be a point of \( M \subset \mathbb{R}^3 \) such that

1. \( \lambda_1 \) is a local maximum at \( m \);
2. \( \lambda_2 \) is a local minimum at \( m \);
3. \( \lambda_1 > \lambda_2 \).

Then \( K \leq 0 \).
**Proof.** Since \( \lambda_1 \neq \lambda_2 \), both are smooth functions. From Theorem ??, we know that

\[
E_1(\lambda_2) = (\lambda_1 - \lambda_2) \omega_{12}(E_2).
\]

It follows that

\[
E_1(E_1(\lambda_2)) = E_1(\lambda_1 - \lambda_2) \omega_{12}(E_2) + (\lambda_1 - \lambda_2) E_1(\omega_{12}(E_2)).
\]

At the point \( m \), we have \( E_1(E_1(\lambda_2)) \geq 0 \) and \( E_1(\lambda_1 - \lambda_2) = 0 \). Thus we have

\[
E_1(\omega_{12}(E_2)) \geq 0.
\]

Similarly, we would have

\[
E_2(\omega_{12}(E_1)) \leq 0.
\]

We claim that

\[
K = E_2(\omega_{12}(E_1)) - E_1(\omega_{12}(E_2)).
\]

In order to prove that, we write

\[
\omega_{12} = \omega_{12}(E_1) \theta_1 + \omega_{12}(E_2) \theta_2.
\]

Since \( \omega_{12} = 0 \) at \( m \), \( d\theta_i = 0 \) at the point. Thus

\[
d\omega_{12} = d(\omega_{12}(E_1)) \wedge \theta_1 + d(\omega_{12}(E_2)) \wedge \theta_2,
\]

which completes the proof.

---

**Theorem 6.6**

*If \( M \) is a compact surface in \( \mathbb{R}^3 \) with constant Gaussian curvature \( K \), then \( M \) is a sphere of radius \( 1/\sqrt{K} \).*

---

### 6.4 Isometries and Local Isometries

**Definition 6.4**

If \( p, q \) are points of \( M \subset \mathbb{R}^3 \), consider the collection of all curve segments \( \alpha \) in \( M \) from \( p \) to \( q \). The **intrinsic distance** \( \rho(p, q) \) from \( p \) to \( q \) in \( M \) is the greatest lower bound of the lengths \( L(\alpha) \) of these curve segments.

**Remark** The following picture showed that difference between the intrinsic distance and the Euclidean distance in \( \mathbb{R}^3 \).
6.4 Isometries and Local Isometries

**Proposition 6.1**

With respect to the intrinsic distance, \( M \) is a metric space.

**Definition 6.5**

An isometry \( F : M \to \overline{M} \) of surfaces in \( \mathbb{R}^3 \) is a one-to-one mapping of \( M \) onto \( \overline{M} \) that preserves the dot products of tangent vectors. That is,

\[
F_*(v) \cdot F_*(w) = v \cdot w
\]

for any vectors \( v, w \).

**Theorem 6.7**

Isometries preserve intrinsic distance.

**Definition 6.6**

A local isometry \( F : M \to N \) of surfaces is a mapping that preserves dot products of tangent vectors (that is, \( F_* \) does).

**Lemma 6.7**

Let \( M \to N \) be a mapping. For each patch \( \mathbf{x} : D \to M \), consider the composite mapping

\[
\bar{x} = F(\mathbf{x}) : D \to N.
\]

Then \( F \) is a local isometry if and only if for each patch \( \mathbf{x} \) we have

\[
E = \bar{E}, \quad F = \bar{F}, \quad G = \bar{G}.
\]

**Example 6.3** (Local isometry of a plane onto a cylinder.) Define \( \mathbf{x} : \mathbb{R}^2 \to M \)

\[
\mathbf{x}(u, v) = \left( r \cos \frac{u}{r}, r \sin \frac{u}{r}, v \right),
\]

where \( M : x^2 + y^2 = r^2 \) is a cylinder. Then \( \mathbf{x} \) is a local isometry.

**Example 6.4** (Local isometry of a helicoid onto a catenoid.) Let \( H \) be the helicoid that is the image of the patch

\[
\mathbf{x}(u, v) = (u \cos v, u \sin v, v).
\]

Furnish the catenoid \( C \) with the canonical parametrization \( \mathbf{y} : \mathbb{R}^2 \to C \) discussed in Example 7.4 of Chapter 5.

Thus

\[
\mathbf{y}(u, v) = (g(u), h(u) \cos v, h(u) \sin v),
\]

\[
g(u) = \sinh^{-1} u, \quad h(u) = \sqrt{1 + u^2}.
\]

Let \( F : H \to C \) be the mapping such that

\[
F(\mathbf{x}(u, v)) = \mathbf{y}(u, v).
\]
A mapping of surfaces $F : M \to N$ is conformal provided there exists a real-valued function $\lambda > 0$ on $M$ such that

$$
\|F_*(v_p)\| = \lambda(p)\|v_p\|
$$

for all tangent vectors to $M$. The function $\lambda$ is called the scale factor of $F$.

**Lemma 6.8**

The connection form $\omega_{12} = -\omega_{21}$ is the only 1-form that satisfies the first structural equations

$$
d\theta_1 = \omega_{12} \wedge \theta_2, \quad d\theta_2 = \omega_{21} \wedge \theta_1.
$$

**Example 6.5** Consider a surface whose first fundamental form is given by

$$
\begin{bmatrix}
E & F \\
F & G
\end{bmatrix} = \begin{bmatrix}
\frac{1}{(1 + \frac{K}{4}(x^2 + y^2))^2} & 0 \\
0 & \frac{1}{(1 + \frac{K}{4}(x^2 + y^2))^2}
\end{bmatrix}.
$$

Define

$$
\theta_1 = \frac{dx}{1 + \frac{K}{4}(x^2 + y^2)}, \quad \theta_2 = \frac{dy}{1 + \frac{K}{4}(x^2 + y^2)}.
$$

Let $W = 1 + \frac{K}{4}(x^2 + y^2)$. Then we have

$$
d\theta_1 = \frac{K y}{2W^2} dx \wedge dy, \quad d\theta_2 = -\frac{K x}{2W^2} dx \wedge dy.
$$

Let $\omega_{12} = adx + b dy$. Using the first structural equations, we have

$$
\omega_{12} = \frac{K y}{2W} dx - \frac{K x}{2W} dy.
$$

A straightforward computation gives

$$
d\omega_{12} = -\frac{K}{W^2} x \wedge dy = -K \theta_1 \wedge \theta_2.
$$
6.6 Orthogonal Coordinates

**Definition 6.8**

The associated frame field $E_1, E_2$ of an orthogonal patch\(^a\) $x : D \rightarrow M$ consists of the orthogonal unit vector fields $E_1$ and $E_2$ whose values at each point $x(u, v)$ of $x(D)$ are

$$\frac{x_u(u, v)}{\sqrt{E(u, v)}}, \quad \frac{x_v(u, v)}{\sqrt{G(u, v)}}$$

\(^a\)An orthogonal patch is the one such that $F = x_u \cdot x_v = 0$.

**Lemma 6.9**

The dual frame is given by

$$\theta_1 = \sqrt{E} \, du, \quad \theta_2 = \sqrt{G} \, dv.$$ 

Moreover,

$$\omega_{12} = -\frac{(\sqrt{E})_v}{\sqrt{G}} \, du + \frac{(\sqrt{G})_u}{\sqrt{E}} \, dv.$$ 

The Gauss curvature is given by

$$K = \frac{-1}{\sqrt{EG}} \left\{ \left( \frac{(\sqrt{G})_u}{\sqrt{E}} \right)_u + \left( \frac{(\sqrt{E})_v}{\sqrt{G}} \right)_v \right\}.$$ 

**Example 6.6** For the unit sphere, we have

$$E = r^2 \cos^2 v, \quad F = 0, \quad G = r^2.$$ 

Then the Gauss curvature is $1$.

6.7 Integration and Orientation

We have already talked about the integration over a differential form in §4.6. In this section, we shall extend the definition and we shall introduce the area form, which is important and is related to the orientation.

**Definition 6.9**

Define the matrix

$$\begin{bmatrix} E & F \\ F & G \end{bmatrix}$$

to be the first fundamental form. Define

$$\begin{bmatrix} L & M \\ M & N \end{bmatrix}$$

to be the second fundamental form.

**Definition 6.10**

Define

$$\sqrt{EG - F^2} \, du \wedge dv$$

to be the area form.
Recall that a surface is orientable, if and only if there is a non-zero 2-form on the surface.

**Theorem 6.8**

*M is orientable if and only if the area form is well-defined.*

**Example 6.7** (Area form of sphere)

\[ E = r^2 \cos^2 v, \quad F = 0, \quad G = r^2. \]

**Example 6.8** (Area form of torus)

\[ E = r^2, \quad F = 0, \quad G = (R + r \cos u)^2. \]

**Example 6.9** (Area form of Bugle surface)

\[ E = 1, \quad F = 0, \quad G = h^2 = e^{-2a/c}. \]

**Definition 6.11**

An *alternating n-form* is a n-form that depends on the order of its arguments.

**Definition 6.12**

Let \( \nu \) be a 2-form on a orientable region \( \mathcal{P} \) in a surface. The integral of \( \nu \) over \( \mathcal{P} \) is

\[
\int_{\mathcal{P}} \nu = \sum_{i} \int_{\mathcal{P}_i} \nu.
\]

### 6.8 Total Curvature

**Definition 6.13**

Let \( M \) be a compact orientable surface. The *total curvature* is defined by

\[
\int_{M} K \, dM,
\]

where \( K \) is the Gauss curvature.

**Example 6.10** (Constant curvature) If the Gauss curvature is a constant, then

\[
\int_{M} K \, dM = K \int_{M} \, dM = K \text{Area}(M).
\]

For example, the total curvature of a sphere of radius \( r \) is given by

\[
\frac{1}{r^2} 4\pi r^2 = 4\pi.
\]

**Example 6.11** (Bugle Surface) The total curvature of the bugle surface is

\[
-\frac{1}{c^2} \cdot 2\pi c^2 = -2\pi.
\]

**Example 6.12** (Torus \( T \)) For the torus, we know that the Gauss curvature is given by

\[
K = \frac{\cos u}{r(R + r \cos u)}.
\]

The volume form is given by

\[
dM = r(R + r \cos u)du dv.
\]
Therefore
\[ \int_T K dT = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \cos u dudv = 0. \]

**Example 6.13** (Catenoid) The Gauss curvature is
\[ K = \frac{-1}{c^2 \cosh^2(u/c)}, \]
and
\[ dM = \sqrt{EG} = c \cosh^2(u/c). \]
Thus the total
\[ \int_{M(a)} K dM = -\int_{-a}^{a} \int_{0}^{2\pi} \frac{1}{c \cosh^2(u/c)} dudv = -4\pi \tanh(a/c) \]
As \( a \to \infty \), we have
\[ \int_M K dM = -4\pi. \]

**Definition 6.14**
Let \( M \) and \( N \) be surfaces oriented by area forms \( dM \) and \( dN \). The Jacobian of a mapping \( F : M \to N \) is the real-valued function \( J_F \) such that
\[ F^*(dN) = J_F dM. \]

**Lemma 6.10**
The Jacobian is the determinant of the Jacobian matrix.

**Definition 6.15**
The Gauss map is defined by the map of a point to its unit normal vector field.

**Theorem 6.9**
The Gauss curvature \( K \) of an oriented surface \( M \subset \mathbb{R}^3 \) is the Jacobian of its Gauss map.

**Definition 6.16**
Let \( v, w \) be unit tangent vectors at a point of an oriented surface \( M \). A number \( \phi \) is an oriented angle from \( v \) to \( w \) provided
\[ w = \cos \phi v + \sin \phi J(v), \]
where \( J(v) = U \times v \).

**Lemma 6.11**
Let \( \alpha : I \to M \) be a curve in an oriented surface \( M \). If \( V \) and \( W \) are nonvanishing tangent vector fields on \( \alpha \), there is a differentiable function \( \phi \) on \( I \) such that for each \( t \) in \( I \), \( \phi(t) \) is an oriented angle from \( V(t) \) to \( W(t) \).

**Theorem 6.10**
A simple closed curve \( \alpha \) in \( \mathbb{R}^3 \) has total curvature
\[ \int_{\alpha} \kappa ds \geq 2\pi. \]
6.9 Congruence of Surfaces

**Theorem 6.11**

If $F$ is a Euclidean isometry such that $F(M) = \overline{M}$, then the restriction of $F$ to $M$ is an isometry of surfaces. Moreover, if $M$ and $\overline{M}$ are suitably oriented, then $F$ preserves shape operators, that is,

$$F_*(S(v)) = \overline{S}(F_*(v)).$$

**Theorem 6.12**

Let $M$ and $\overline{M}$ be oriented surfaces in $\mathbb{R}^3$. Let $F : M \to \overline{M}$ be an isometry of oriented surfaces in $\mathbb{R}^3$ that preserves shape operators, so

$$F_*(S(v)) = \overline{S}(F_*(v))$$

for any tangent vectors to $M$. Then $M$ and $\overline{M}$ are congruent; in fact, there is a Euclidean isometry $F$ such that $F|_M = F$. 