Differentiable Manifolds

Lecture Notes for MATH 4033 (Spring 2018)

FREDERICK TSZ-HO FONG

Hong Kong University of Science and Technology
(Version: January 23, 2018)

Contents

Preface

This lecture note is written for the course MATH 4033 (Calculus on Manifolds) taught by the author in the Hong Kong University of Science and Technology. The main goal of the course is to introduce advanced undergraduates and first-year graduates the basic language of differentiable manifolds and tensor calculus. The topics covered in the course is essential for further studies on Riemannian geometry, general relativity, string theory, and related fields. The prerequisite of the course is a solid conceptual background of linear algebra and multivariable calculus.

The course MATH 4033 covers Chapters 1 to 5 in this lecture note. These chapters are about the analytic, algebraic, and topological aspects of differentiable manifolds. The appendix in this lecture note forms a crush course on differential geometry of curves and surfaces. They are not the essential parts of the course, but is strongly recommended for readers who want to acquire some workable knowledge in differential geometry (such as for the purpose of my UROP)

The author would like to thank the following students for their diligent readings of the earlier version of this lecture notes and for pointing out many typographical errors: Chow Ka-Wing, Alex Chan Yan-Long, Aaron Chow Tsz-Kiu, Jimmy Choy Ka-Hei, Toby Cheung Hin-Wa, Poon Wai-Tung, Cheng Chun-Kit, Chu Shek-Kit, Wan Jingbo, and Nicholas Chin Cheng-Hoong.

Chapter 1

Regular Surfaces

"God made solids, but surfaces were the work of the devil."

Wolfgang Pauli

A manifold is a space which locally resembles an Euclidean space. Before we learn about manifolds in the next chapter, we first introduce the notion of regular surfaces in $\mathbb{R}^3$ which motivates the definition of abstract manifolds and related concepts in the next chapter.

1.1. Local Parametrizations

In Multivariable Calculus, we expressed a surface in $\mathbb{R}^3$ in two ways, namely using a parametrization $\mathsf{F}(u,v)$ or by a level set $f(x,y,z) = 0$. In this section, let us first focus on the former.

In MATH 2023, we used a parametrization $\mathsf{F}(u,v)$ to describe a surface in $\mathbb{R}^3$ and to calculate various geometric and physical quantities such as surface areas, surface integrals and surface flux. To start the course, we first look into several technical and analytical aspects concerning $\mathsf{F}(u,v)$, such as their domains and images, their differentiability, etc. In the past, we can usually cover (or almost cover) a surface by a single parametrization $\mathsf{F}(u,v)$. Take the unit sphere as an example. We learned that it can be parametrized with the help of spherical coordinates:

$$\mathsf{F}(\theta,\varphi) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$$

where $0 < \theta < 2\pi$ and $0 < \varphi < \pi$. This parametrization covers almost every part of the sphere (except the north and south poles, and a half great circle connecting them). In order to cover the whole sphere, we need more parametrizations, such as $\mathsf{G}(\theta,\varphi) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$ with domain $-\pi < \theta < \pi$ and $0 < \varphi < \pi$.

Since the image of either F or G does not cover the whole sphere (although almost), from now on we call them *local parametrizations*.

Definition 1.1 (Local Parametrizations of Class C^k)**.** Consider a subset $M \subset \mathbb{R}^3$. A function $\mathsf{F}(u,v) : \mathcal{U} \to \mathcal{O}$ from an open subset $\mathcal{U} \subset \mathbb{R}^2$ onto an open subset $\mathcal{O} \subset M$ is called a C^k *local parametrization* (or a C^k *local coordinate chart*) of M (where $k \geq 1$) if all of the following holds:

(1) $\mathsf{F} : \mathcal{U} \to \mathbb{R}^3$ is C^k when the codomain is regarded as $\mathbb{R}^3$.

(2) $\mathsf{F} : \mathcal{U} \to \mathcal{O}$ is a *homeomorphism*, meaning that $\mathsf{F} : \mathcal{U} \to \mathcal{O}$ is bijective, and both F and F^{-1} are continuous.

(3) For all $(u,v) \in \mathcal{U}$, the cross product:

$$\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v} \neq 0.$$

The coordinates (u,v) are called the *local coordinates* of M.

If $\mathsf{F} : \mathcal{U} \to M$ is of class C^k for any integer k, then F is said to be a C^∞ (or smooth) local parametrization.

Definition 1.2 (Surfaces of Class C^k)**.** A subset $M \subset \mathbb{R}^3$ is called a C^k *surface*, where $k \in \mathbb{N} \cup \{\infty\}$, in $\mathbb{R}^3$ if at every point $p \in M$, there exists an open subset $\mathcal{U} \subset \mathbb{R}^2$, an open subset $\mathcal{O} \subset M$ containing p, and a C^k local parametrization $\mathsf{F} : \mathcal{U} \to \mathcal{O}$ which satisfies all three conditions stated in Definition 1.1.

We say M is a *regular surface* in $\mathbb{R}^3$ if it is a C^∞ surface.

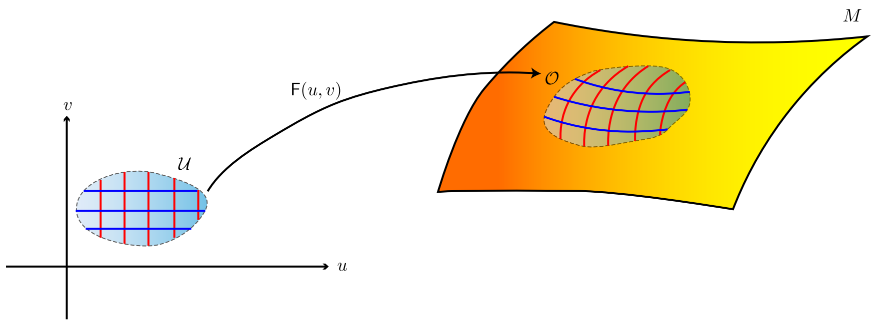
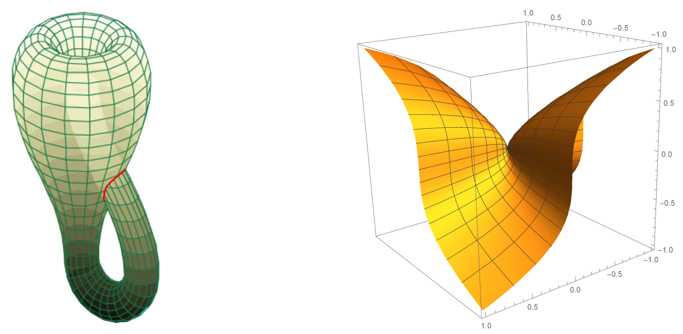


Figure 1.1. smooth local parametrization

To many students (myself included), the definition of regular surfaces looks obnoxious at the first glance. One way to make sense of it is to look at some examples and understand why each of the three conditions is needed in the definition.

The motivation behind condition (1) in the definition is that we are studying *differential* topology/geometry and so we want the parametrization to be differentiable as many times as we like. Condition (2) rules out surfaces that have self-intersection such as the Klein bottle (see Figure 1.2a). Finally, condition (3) guarantees the existence of a unique tangent plane at every point on M (see Figure 1.2b for a non-example).



(a) Klein Bottle has a self-intersection. **(b)** $\mathsf{F}(u,v) = (u^3, v^3, uv)$ fails condition (3).

Figure 1.2. Examples of non-smooth parametrizations

Example 1.3 (Graph of a Function)**.** Consider a smooth function $f(u,v) : \mathcal{U} \to \mathbb{R}$ defined on an open subset $\mathcal{U} \subset \mathbb{R}^2$. The graph of f, denoted by Γ_f, is the subset $\{(u,v,f(u,v)) : (u,v) \in \mathcal{U}\}$ of $\mathbb{R}^3$. One can parametrize Γ_f by a global parametrization:

$$\mathsf{F}(u,v) = (u,\ v,\ f(u,v)).$$

Condition (1) holds because f is given to be smooth. For condition (2), F is clearly one-to-one, and the image of F is the whole graph Γ_f. Regarding it as a map $\mathsf{F} : \mathcal{U} \to \Gamma_f$, the inverse map

$$\mathsf{F}^{-1}(x,y,z) = (x,y)$$

is clearly continuous. Therefore, $\mathsf{F} : \mathcal{U} \to \Gamma_f$ is a homeomorphism. To verify condition (3), we compute the cross product:

$$\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v} = \left(-\frac{\partial f}{\partial u},\ -\frac{\partial f}{\partial v},\ 1 \right) \neq 0$$

for all $(u,v) \in \mathcal{U}$. Therefore, F is a smooth local parametrization of Γ_f. Since the image of this single smooth local parametrization covers all of Γ_f, we have proved that Γ_f is a regular surface. □

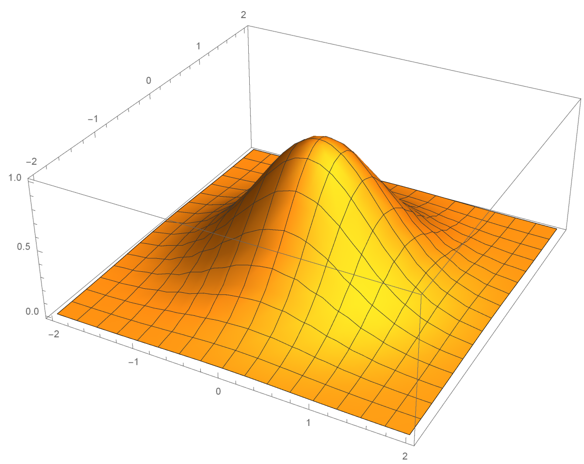


Figure 1.3. The graph of any smooth function is a regular surface.

Exercise 1.1. Show that $\mathsf{F}(u,v) : (0,2\pi) \times (0,1) \to \mathbb{R}^3$ defined by:

$$\mathsf{F}(u,v) = (\sin u, \sin 2u, v)$$

satisfies conditions (1) and (3) in Definition 1.1, but not condition (2). [Hint: Try to show F^{-1} is not continuous by finding a diverging sequence $\{(u_n,v_n)\}$ such that $\{\mathsf{F}(u_n,v_n)\}$ converges. See Figure 1.4 for reference.]

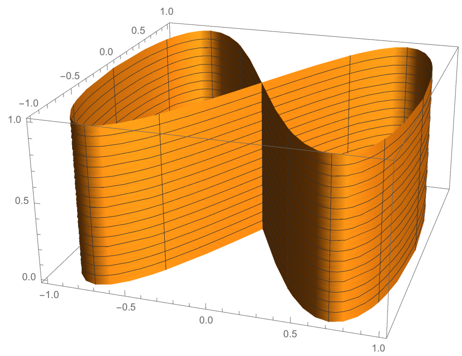


Figure 1.4. Plot of $\mathsf{F}(u,v)$ in Exercise 1.1

In Figure 1.3, one can observe that there are two families of curves on the surface. These curves are obtained by varying one of the (u,v)-variables while keeping the other constant. Precisely, they are the curves represented by $\mathsf{F}(u,v_0)$ and $\mathsf{F}(u_0,v)$ where u_0 and v_0 are fixed. As such, the partial derivatives $\frac{\partial \mathsf{F}}{\partial u}(p)$ and $\frac{\partial \mathsf{F}}{\partial v}(p)$ give a pair of tangent vectors on the surface at point p. Therefore, their cross product $\frac{\partial \mathsf{F}}{\partial u}(p) \times \frac{\partial \mathsf{F}}{\partial v}(p)$ is a normal vector to the surface at point p (see Figure 1.5). Here we have abused the notations for simplicity: $\frac{\partial \mathsf{F}}{\partial u}(p)$ means $\frac{\partial \mathsf{F}}{\partial u}$ evaluated at $(u,v) = \mathsf{F}^{-1}(p)$. Similarly for $\frac{\partial \mathsf{F}}{\partial v}(p)$.

Condition (3) requires that $\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v}$ is everywhere non-zero in the domain of F. An equivalent statement is that the vectors $\left\{\frac{\partial \mathsf{F}}{\partial u}(p), \frac{\partial \mathsf{F}}{\partial v}(p)\right\}$ are linearly independent for any $p \in \mathsf{F}(\mathcal{U})$.

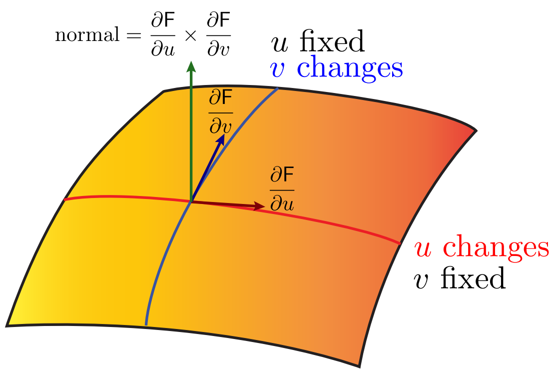


Figure 1.5. Tangent and normal vectors to a surface in $\mathbb{R}^3$

Example 1.4 (Sphere)**.** In $\mathbb{R}^3$, the unit sphere $\mathbb{S}^2$ centered at the origin can be represented by the equation $x^2 + y^2 + z^2 = 1$, or in other words, $z = \pm\sqrt{1 - x^2 - y^2}$. We can parametrize the upper and lower hemisphere by two separate local maps:

$$\mathsf{F}_1(u,v) = (u,\ v,\ \sqrt{1-u^2-v^2}) : B_1(0) \subset \mathbb{R}^2 \to \mathbb{S}^2_+$$
$$\mathsf{F}_2(u,v) = (u,\ v,\ -\sqrt{1-u^2-v^2}) : B_1(0) \subset \mathbb{R}^2 \to \mathbb{S}^2_-$$

where $B_1(0) = \{(u,v) : u^2 + v^2 < 1\}$ is the *open* unit disk in $\mathbb{R}^2$ centered at the origin, and $\mathbb{S}^2_+$ and $\mathbb{S}^2_-$ are the upper and lower hemispheres of $\mathbb{S}^2$ respectively. Since $B_1(0)$ is open, the functions $\pm\sqrt{1-u^2-v^2}$ are smooth, and according to the previous example both F_1 and F_2 are smooth local parametrizations.

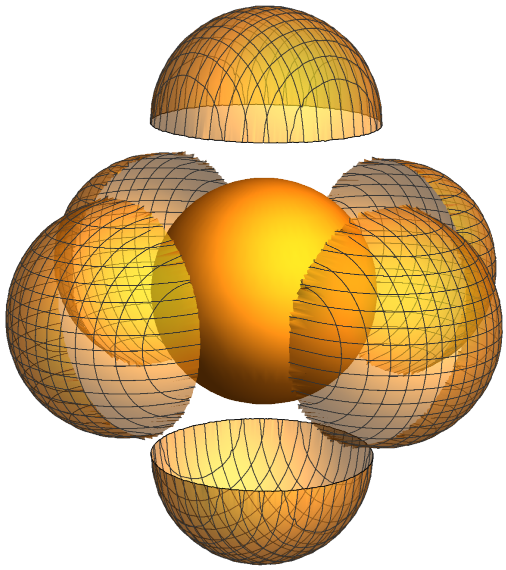


Figure 1.6. A unit sphere covered by six parametrization charts

However, not all points on the sphere are covered by $\mathbb{S}^2_+$ and $\mathbb{S}^2_-$, since points on the equator are not. In order for show that $\mathbb{S}^2$ is a regular surface, we need to write down more smooth local parametrization(s) so that each point on the sphere can be covered by at least one smooth local parametrization chart. One can construct four more smooth local parametrizations (left, right, front and back) similar to F_1 and F_2 (see Figure 1.6). It is left as an exercise for readers to write down the other four parametrizations. These six parametrizations are all smooth and they cover the whole sphere. Therefore, it shows the sphere is a regular surface. □

Exercise 1.2. Write down the left, right, front and back parametrizations F_i's ($i = 3, 4, 5, 6$) of the sphere as shown in Figure 1.6. Indicate clearly the domain and range of each F_i.

Example 1.5 (Sphere: revisited)**.** We can in fact cover the sphere by two smooth local parametrization described below. Define $\mathsf{F}_+(u,v) : \mathbb{R}^2 \to \mathbb{S}^2 \backslash \{(0,0,1)\}$ where:

$$\mathsf{F}_+(u,v) = \left(\frac{2u}{u^2+v^2+1}, \frac{2v}{u^2+v^2+1}, \frac{u^2+v^2-1}{u^2+v^2+1} \right)$$

It is called the *stereographic parametrization* of the sphere (see Figure 1.7) , which assigns each point $(u,v,0)$ on the xy-plane of $\mathbb{R}^3$ to a point where the line segment joining $(u,v,0)$ and the north pole $(0,0,1)$ intersects the sphere. Clearly F_+ is a smooth function. We leave it as exercise for readers to verify that F_+ satisfies condition (3) and that $\mathsf{F}_+^{-1} : \mathbb{S}^2 \backslash \{(0,0,1)\} \to \mathbb{R}^2$ is given by:

$$\mathsf{F}_+^{-1}(x,y,z) = \left(\frac{x}{1-z}, \frac{y}{1-z} \right).$$

As $z \neq 1$ for every (x,y,z) in the domain of F_+^{-1}, it is a continuous function. Therefore, F_+ is a smooth local parametrization. The inverse map F_+^{-1} is commonly called the stereographic projection of the sphere.

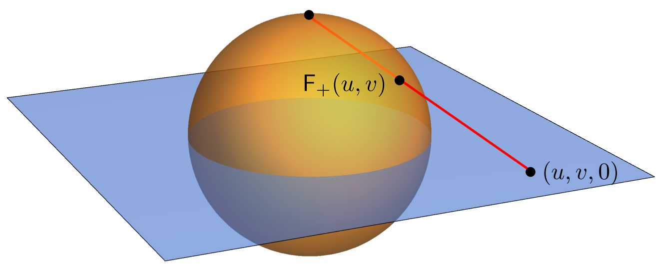


Figure 1.7. Stereographic parametrization of the sphere

Note that the range of F_+ does not include the point $(0,0,1)$. In order to show that the sphere is a regular surface, we need to cover it by another parametrization $\mathsf{F}_- : \mathbb{R}^2 \to \mathbb{S} \backslash \{(0,0,-1)\}$ which assigns each point $(u,v,0)$ on the xy-plane to a point where the line segment joining $(u,v,0)$ and the south pole $(0,0,-1)$ intersects the sphere. It is an exercise for readers to write down the explicit parametrization F_-. $\square$

Exercise 1.3. Verify that F_+ in Example 1.4 satisfies condition (3) in Definition 1.1, and that the inverse map $\mathsf{F}_+^{-1} : \mathbb{S}^2 \backslash \{(0,0,1)\} \to \mathbb{R}^2$ is given as stated. [Hint: Write down $\mathsf{F}_+(u,v) = (x,y,z)$ and solve (u,v) in terms of (x,y,z). Begin by finding u^2+v^2 in terms of z.]

Furthermore, write down explicitly the map F_- described in Example 1.4, and find its inverse map F_-^{-1}.

Exercise 1.4. Find smooth local parametrizations which together cover the whole ellipsoid:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

where a, b and c are positive constants.

Exercise 1.5. Let M be the cylinder $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1\}$. The purpose of this exercise is to construct a smooth local parametrization analogous to the stereographic parametrization in Example 1.4:

Consider the unit circle $x^2 + y^2 = 1$ on the xy-plane. For each point $(u, 0)$ on the x-axis, we construct a straight-line joining the point $(0, 1)$ and $(u, 0)$. This line intersects the unit circle at a unique point p. Denote the xy-coordinates of p by $(x(u), y(u))$.

(a) Find the coordinates $(x(u), y(u))$ in terms of u.

(b) Define:

$$\mathsf{F}_1(u, v) = (x(u), y(u), v)$$

with $\mathbb{R}^2$ as its domain. Describe the image of F_1.

(c) Denote $\mathcal{O}_1$ to be the image of F_1. Verify that $\mathsf{F}_1 : \mathbb{R}^2 \to \mathcal{O}_1$ is smooth local parametrization of M.

(d) Construct another smooth local parametrization F_2 such that the images of F_1 and F_2 cover the whole surface M (hence establish that M is a regular surface).

Let's also look at a non-example of smooth local parametrizations. Consider the map:

$$\mathsf{G}(u, v) = (u^3, v^3, 0), \quad (u, v) \in \mathbb{R} \times \mathbb{R}.$$

It is a smooth, injective map from $\mathbb{R}^2$ onto the xy-plane Π of $\mathbb{R}^3$, i.e. $\mathsf{G} : \mathbb{R}^2 \to \Pi$. However, it can be computed that

$$\frac{\partial \mathsf{G}}{\partial u}(0, 0) = \frac{\partial \mathsf{G}}{\partial v}(0, 0) = 0$$

and so condition (3) in Definition 1.1 does not hold. The map G is not a smooth local parametrization of Π. However, note that Π is a regular surface because $\mathsf{F}(u, v) = (u, v, 0)$ is a smooth global parametrization of Π, even though G is not a "good" parametrization.

In order to show M is a regular surface, what we need is to show at every point $p \in M$ there is *at least one* smooth local parametrization F near p. However, to show that M is not a regular surface, one then needs to come up with a point $p \in M$ such that there is *no* smooth local parametrization near that point p (which may not be easy).

1.2. Level Surfaces

Many surfaces are defined using an equation such as $x^2 + y^2 + z^2 = 1$, or $x^2 + y^2 = z^2 + 1$. They are level sets of a function $g(x, y, z)$. In this section, we are going to prove a theorem that allows us to show easily that some level sets $g^{-1}(c)$ are regular surfaces.

Theorem 1.6. *Let $g(x, y, z) : \mathbb{R}^3 \to \mathbb{R}$ be a smooth function of three variables. Consider a non-empty level set $g^{-1}(c)$ where c is a constant. If $\nabla g(x_0, y_0, z_0) \neq 0$ at all points $(x_0, y_0, z_0) \in g^{-1}(c)$, then the level set $g^{-1}(c)$ is a regular surface.*

Proof. The key idea of the proof is to use the Implicit Function Theorem. Given any point $p = (x_0, y_0, z_0) \in g^{-1}(c)$, since $\nabla g(x_0, y_0, z_0) \neq (0, 0, 0)$, at least one of the first partials:

$$\frac{\partial g}{\partial x}(p), \ \frac{\partial g}{\partial y}(p), \ \frac{\partial g}{\partial z}(p)$$

is non-zero. Without loss of generality, assume $\frac{\partial g}{\partial z}(p) \neq 0$, then the Implicit Function Theorem shows that locally around the point p, the level set $g^{-1}(c)$ can be regarded as a graph $z = f(x, y)$ of some smooth function f of (x, y). To be precise, there exists an open set $\mathcal{O}$ of $g^{-1}(c)$ containing p such that there is a smooth function $f(x, y) : \mathcal{U} \subset \mathbb{R}^2 \to \mathbb{R}$ from an open set $\mathcal{U}$ such that $(x, y, f(x, y)) \in \mathcal{O} \subset g^{-1}(c)$ for any $(x, y) \in \mathcal{U}$. As such, the smooth local parametrization $\mathsf{F} : \mathcal{U} \to \mathcal{O}$ defined by:

$$\mathsf{F}(u, v) = (u, v, f(u, v))$$

is a smooth local parametrization of $g^{-1}(c)$.

In the case where $\frac{\partial g}{\partial y}(p) \neq 0$, the above argument is similar as locally around p one can regard $g^{-1}(c)$ as a graph $y = h(x, z)$ for some smooth function h. Similar in the case $\frac{\partial g}{\partial x}(p) \neq 0$.

Since every point p can be covered by the image of a smooth local parametrization, the level set $g^{-1}(c)$ is a regular surface. □

Example 1.7. The unit sphere $x^2 + y^2 + z^2 = 1$ is a level surface $g^{-1}(1)$ where $g(x, y, z) := x^2 + y^2 + z^2$. The gradient vector $\nabla g = (2x,\ 2y,\ 2z)$ is zero only when $(x, y, z) = (0, 0, 0)$. Since the origin is not on the unit sphere, we have $\nabla g(x_0, y_0, z_0) \neq (0, 0, 0)$ for any $(x_0, y_0, z_0) \in g^{-1}(1)$. Therefore, the unit sphere is a regular surface.

Similarly, one can also check that the surface $x^2 + y^2 = z^2 + 1$ is a regular surface. It is a level set $h^{-1}(1)$ where $h(x, y, z) = x^2 + y^2 - z^2$. Since $\nabla h = (2x,\ 2y,\ -2z)$, the origin is the only point p at which $\nabla h(p) = (0, 0, 0)$ and it is not on the level set $h^{-1}(1)$. Therefore, $h^{-1}(1)$ is a regular surface. □

However, the cone $x^2 + y^2 = z^2$ cannot be shown to be a regular surface using Theorem 1.6. It is a level surface $h^{-1}(0)$ where $h(x, y, z) := x^2 + y^2 - z^2$. The origin $(0, 0, 0)$ is on the cone and $\nabla h(0, 0, 0) = (0, 0, 0)$. Theorem 1.6 fails to give any conclusion.

The converse of Theorem 1.6 is not true. Consider $g(x, y, z) = z^2$, then $g^{-1}(0)$ is the xy-plane which is clearly a regular surface. However, $\nabla g = (0, 0, 2z)$ is zero at the origin which is contained in the xy-plane.

Exercise 1.6. [**dC76**, P.66] Let $f(x,y,z) = (x+y+z-1)^2$. For what values of c is the set $f^{-1}(c)$ a regular surface?

Exercise 1.7. A torus is defined by the equation:

$$z^2 = R^2 - \left(\sqrt{x^2+y^2} - r\right)^2$$

where $R > r > 0$ are constants. Show that it is a regular surface.

The proof of Theorem 1.6 makes use of the Implicit Function Theorem which is an existence result. It shows a certain level set is a regular surface, but it fails to give an explicit smooth local parametrization around each point.

There is one practical use of Theorem 1.6 though. Suppose we are given $\mathsf{F}(u,v)$ which satisfies conditions (1) and (3) in Definition 1.1 and that F is continuous and F^{-1} exists. In order to verify that it is a smooth local parametrization, we need to prove continuity of F^{-1}, which is sometimes difficult. Here is one example:

$$\mathsf{F}(u,v) = (\sin u \cos v,\ \sin u \sin v,\ \cos u), \quad 0 < u < \pi,\ 0 < v < 2\pi$$

is a smooth local parametrization of a unit sphere. It is clearly a smooth map from $(0,\pi) \times (0,2\pi) \subset \mathbb{R}^2$ to $\mathbb{R}^3$, and it is quite straight-forward to verify condition (3) in Definition 1.1 and that F is one-to-one. However, it is rather difficult to write down an explicit F^{-1}, let alone to show it is continuous.

The following result tells us that if the surface is given by a level set satisfying conditions stated in Theorem 1.6, and F satisfies conditions (1) and (3), then F^{-1} is *automatically* continuous. Precisely, we have the following:

Proposition 1.8. *Assume all given conditions stated in Theorem 1.6. Furthermore, suppose* $\mathsf{F}(u,v)$ *is a bijective map from an open set* $\mathcal{U} \subset \mathbb{R}^2$ *to an open set* $\mathcal{O} \subset M := g^{-1}(c)$ *which satisfies conditions (1) and (3) in Definition 1.1. Then,* F *satisfies condition (2) as well and hence is a smooth local parametrization of* $g^{-1}(c)$.

Proof. Given any point $p \in g^{-1}(c)$, we can assume without loss of generality that $\frac{\partial g}{\partial z}(p) \neq 0$. Recall from Multivariable Calculus that $\nabla g(p)$ is a normal vector to the level surface $g^{-1}(c)$ at point p. Furthermore, if $\mathsf{F}(u,v)$ is a map satisfying conditions (1) and (3) of Definition 1.1, then $\frac{\partial \mathsf{F}}{\partial u}(p) \times \frac{\partial \mathsf{F}}{\partial v}(p)$ is also a normal vector to $g^{-1}(c)$ at p.

Now that the k-component of $\nabla g(p)$ is non-zero since $\frac{\partial g}{\partial z}(p) \neq 0$, so the k-component of the cross product $\frac{\partial \mathsf{F}}{\partial u}(p) \times \frac{\partial \mathsf{F}}{\partial v}(p)$ is also non-zero. If we express $\mathsf{F}(u,v)$ as:

$$\mathsf{F}(u,v) = (x(u,v), y(u,v), z(u,v)),$$

then the k-component of $\frac{\partial \mathsf{F}}{\partial u}(p) \times \frac{\partial \mathsf{F}}{\partial v}(p)$ is given by:

$$\left(\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}\right)(p) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}(p).$$

Define $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ by $\pi(x,y,z) = (x,y)$. The above shows that the composition $\pi \circ \mathsf{F}$ given by

$$(\pi \circ \mathsf{F})(u,v) = (x(u,v), y(u,v))$$

has non-zero Jacobian determinant at p. By the Inverse Function Theorem, $\pi \circ \mathsf{F}$ has a smooth local inverse near p. In particular, $(\pi \circ \mathsf{F})^{-1}$ is continuous near p.

Finally, by the fact that $(\pi \circ \mathsf{F}) \circ \mathsf{F}^{-1} = \pi$ and that $(\pi \circ \mathsf{F})^{-1}$ exists and is continuous locally around p, we can argue that $\mathsf{F}^{-1} = (\pi \circ \mathsf{F})^{-1} \circ \pi$ is also continuous near p. It completes the proof. $\square$

Exercise 1.8. Rewrite the proof of Proposition 1.8 by assuming $\frac{\partial g}{\partial y}(p) \neq 0$ instead.

Example 1.9. We have already shown that the unit sphere $x^2 + y^2 + z^2 = 1$ is a regular surface using Theorem 1.6 by regarding it is the level set $g^{-1}(1)$ where $g(x, y, z) = x^2 + y^2 + z^2$. We also discussed that

$$\mathsf{F}(u, v) = (\sin u \, \cos v, \, \sin u \, \sin v, \, \cos u), \quad 0 < u < \pi, \, 0 < v < 2\pi$$

is a *possible* smooth local parametrization. It is clearly smooth, and by direct computation, one can show

$$\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v} = \sin u \, (\sin u \cos v, \, \sin u \sin v, \, \cos u)$$

and so $\left| \frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v} \right| = \sin u \neq 0$ for any (u, v) in the domain $(0, \pi) \times (0, 2\pi)$. We leave it as an exercise for readers to verify that F is one-to-one (and so bijective when its codomain is taken to be its image).

Condition (2) is not easy to verify because it is difficult to write down the inverse map F^{-1} explicitly. However, thanks for Proposition 1.8, F is a smooth local parametrization since it satisfies conditions (1) and (3), and it is one-to-one. $\square$

Exercise 1.9. Consider that the *Mercator projection* of the unit sphere:

$$\mathsf{F}(u, v) = \left(\frac{\cos v}{\cosh u}, \frac{\sin v}{\cosh u}, \frac{\sinh u}{\cosh u} \right)$$

where $\sinh u := \frac{1}{2}(e^u - e^{-u})$ and $\cosh u := \frac{1}{2}(e^u + e^{-u})$.

(a) What are the domain and range of F?

(b) Show that F is a smooth local parametrization.

Exercise 1.10. Consider the following parametrization of a torus $\mathbb{T}^2$:

$$\mathsf{F}(u, v) = ((r \cos u + R) \cos v, \, (r \cos u + R) \sin v, \, r \sin u)$$

where $(u, v) \in (0, 2\pi) \times (0, 2\pi)$, and $R > r > 0$ are constants. Show that F is a smooth local parametrization.

1.3. Transition Maps

Let $M \subset \mathbb{R}^3$ be a regular surface, and $\mathsf{F}_\alpha(u_1, u_2) : \mathcal{U}_\alpha \to M$ and $\mathsf{F}_\beta(v_1, v_2) : \mathcal{U}_\beta \to M$ be two smooth local parametrizations of M with overlapping images, i.e. $\mathcal{W} := \mathsf{F}_\alpha(\mathcal{U}_\alpha) \cap \mathsf{F}_\beta(\mathcal{U}_\beta) \neq \varnothing$. Under this set-up, it makes sense to define the maps $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ and $\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta$. However, we need to shrink their domains so as to guarantee they are well-defined. Precisely:

$$(\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha) : \mathsf{F}_\alpha^{-1}(\mathcal{W}) \to \mathsf{F}_\beta^{-1}(\mathcal{W})$$
$$(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta) : \mathsf{F}_\beta^{-1}(\mathcal{W}) \to \mathsf{F}_\alpha^{-1}(\mathcal{W})$$

Note that $\mathsf{F}_\alpha^{-1}(\mathcal{W})$ and $\mathsf{F}_\beta^{-1}(\mathcal{W})$ are open subsets of $\mathcal{U}_\alpha$ and $\mathcal{U}_\beta$ respectively. The map $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ describes a relation between two sets of coordinates (u_1, u_2) and (v_1, v_2) of M. In other words, one can regard $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ as a *change-of-coordinates*, or *transition* map and we can write:

$$\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha(u_1, u_2) = (v_1(u_1, u_2),\, v_2(u_1, u_2)).$$

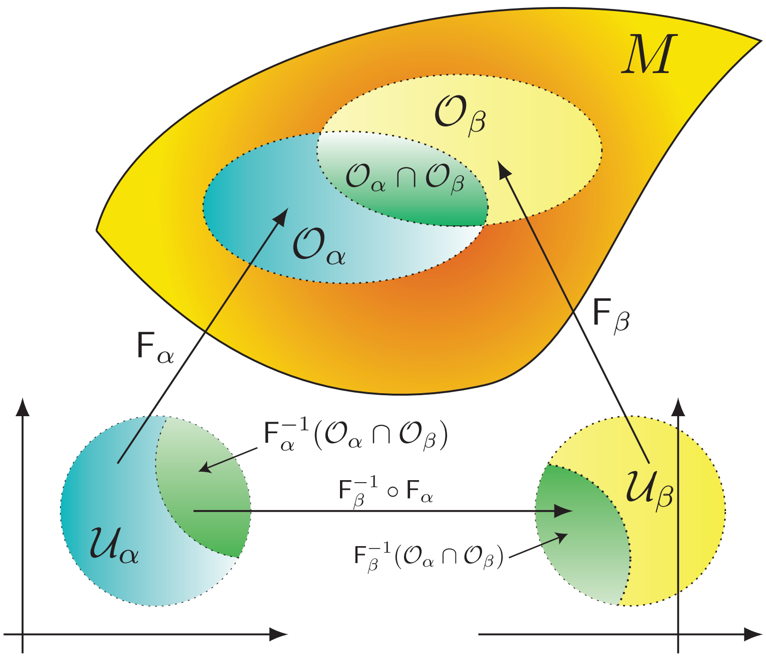


Figure 1.8. Transition maps

One goal of this section is to show that this transition map $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ is smooth provided that F_α and F_β are two overlapping smooth local parametrizations. Before we present the proof, let us look at some examples of transition maps.

Example 1.10. The xy-plane Π in $\mathbb{R}^3$ is a regular surface which admits a global smooth parametrization $\mathsf{F}_\alpha(x,y) = (x,y,0) : \mathbb{R}^2 \to \Pi$. Another way to locally parametrize Π is by polar coordinates $\mathsf{F}_\beta : (0,\infty) \times (0,2\pi) \to \Pi$

$$\mathsf{F}_\beta(r,\theta) = (r\cos\theta, r\sin\theta, 0)$$

Readers should verify that they are smooth local parametrizations. The image of F_α is the entire xy-plane Π, whereas the image of F_β is the xy-plane with the origin and positive x-axis removed. The transition map $\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta$ is given by:

$$\begin{aligned} \mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta : (0,\infty) \times (0,2\pi) &\to \mathbb{R}^2 \backslash \{(x,0) : x \geq 0\} \\ (r,\theta) &\mapsto (r\cos\theta, r\sin\theta) \end{aligned}$$

To put it in a simpler form, we can say $(x(r,\theta), y(r,\theta)) = (r\cos\theta, r\sin\theta)$. □

Exercise 1.11. Consider the stereographic parametrizations F_+ and F_- in Example 1.5. Compute the transition maps $\mathsf{F}_+^{-1} \circ \mathsf{F}_-$ and $\mathsf{F}_-^{-1} \circ \mathsf{F}_+$. State the maximum possible domain for each map. Are they smooth on their domains?

Exercise 1.12. The unit cylinder Σ^2 in $\mathbb{R}^3$ can be covered by two local parametrizations:

$$\begin{aligned} \mathsf{F} &: (0,2\pi) \times \mathbb{R} \to \Sigma^2 & \widetilde{\mathsf{F}} &: (-\pi,\pi) \times \mathbb{R} \to \Sigma^2 \\ \mathsf{F}(\theta,z) &:= (\cos\theta, \sin\theta, z) & \widetilde{\mathsf{F}}(\widetilde{\theta},\widetilde{z}) &:= (\cos\widetilde{\theta}, \sin\widetilde{\theta}, \widetilde{z}) \end{aligned}$$

Compute the transition maps $\mathsf{F}^{-1} \circ \widetilde{\mathsf{F}}$ and $\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F}$. State their maximum possible domains. Are they smooth on their domains?

Exercise 1.13. The Möbius strip Σ^2 in $\mathbb{R}^3$ can be covered by two local parametrizations:

$$\mathsf{F} : (-1,1) \times (0,2\pi) \to \Sigma^2 \qquad \widetilde{\mathsf{F}} : (-1,1) \times (-\pi,\pi) \to \Sigma^2$$

$$\mathsf{F}(u,\theta) = \begin{bmatrix} \left(3 + u\cos\frac{\theta}{2}\right)\cos\theta \\ \left(3 + u\cos\frac{\theta}{2}\right)\sin\theta \\ u\sin\frac{\theta}{2} \end{bmatrix} \qquad \widetilde{\mathsf{F}}(\widetilde{u},\widetilde{\theta}) = \begin{bmatrix} \left(3 + \widetilde{u}\cos\frac{\widetilde{\theta}}{2}\right)\cos\widetilde{\theta} \\ \left(3 + \widetilde{u}\cos\frac{\widetilde{\theta}}{2}\right)\sin\widetilde{\theta} \\ \widetilde{u}\sin\frac{\widetilde{\theta}}{2} \end{bmatrix}$$

Compute the transition maps, state their maximum possible domains and verify that they are smooth.

The proposition below shows that the transition maps between any pair of smooth local parametrizations are smooth:

Proposition 1.11. *Let $M \subset \mathbb{R}^3$ be a regular surface, and $\mathsf{F}_\alpha(u_1,u_2) : \mathcal{U}_\alpha \to M$ and $\mathsf{F}_\beta(v_1,v_2) : \mathcal{U}_\beta \to M$ be two smooth local parametrizations of M with overlapping images, i.e. $\mathcal{W} := \mathsf{F}_\alpha(\mathcal{U}_\alpha) \cap \mathsf{F}_\beta(\mathcal{U}_\beta) \neq \emptyset$. Then, the transition maps defined below are also smooth maps:*

$$\begin{aligned} (\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha) &: \mathsf{F}_\alpha^{-1}(\mathcal{W}) \to \mathsf{F}_\beta^{-1}(\mathcal{W}) \\ (\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta) &: \mathsf{F}_\beta^{-1}(\mathcal{W}) \to \mathsf{F}_\alpha^{-1}(\mathcal{W}) \end{aligned}$$

Proof. It suffices to show $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ is smooth as the other one $\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta$ can be shown by symmetry. Furthermore, since differentiability is a local property, we may fix a point $p \in \mathcal{W} \subset M$ and show that $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ is smooth at the point $\mathsf{F}_\alpha^{-1}(p)$.

By condition of (3) of smooth local parametrizations, we have:

$$\frac{\partial \mathsf{F}_\alpha}{\partial u_1}(p) \times \frac{\partial \mathsf{F}_\alpha}{\partial u_2}(p) \neq 0$$

By straight-forward computations, one can show that this cross product is given by:

$$\frac{\partial \mathsf{F}_\alpha}{\partial u_1} \times \frac{\partial \mathsf{F}_\alpha}{\partial u_2} = \left(\det \frac{\partial(y,z)}{\partial(u_1,u_2)}(p),\ \det \frac{\partial(z,x)}{\partial(u_1,u_2)}(p),\ \det \frac{\partial(x,y)}{\partial(u_1,u_2)}(p) \right).$$

Hence, at least one of the determinants is non-zero. Without loss of generality, assume that:

$$\det \frac{\partial(x,y)}{\partial(u_1,u_2)}(p) \neq 0.$$

Both $\dfrac{\partial \mathsf{F}_\alpha}{\partial u_1}(p) \times \dfrac{\partial \mathsf{F}_\alpha}{\partial u_2}(p)$ and $\dfrac{\partial \mathsf{F}_\beta}{\partial v_1}(p) \times \dfrac{\partial \mathsf{F}_\beta}{\partial v_2}(p)$ are normal vectors to the surface at p. Given that the former has non-zero k-component, then so does the latter. Therefore, we have:

$$\det \frac{\partial(x,y)}{\partial(v_1,v_2)}(p) \neq 0.$$

Then we proceed as in the proof of Proposition 1.8. Define $\pi(x,y,z) = (x,y)$, then

$$\begin{aligned} \pi \circ \mathsf{F}_\beta : \mathcal{U}_\beta &\to \mathbb{R}^2 \\ (v_1, v_2) &\mapsto (x(v_1,v_2), y(v_1,v_2)) \end{aligned}$$

has non-zero Jacobian determinant $\det \dfrac{\partial(x,y)}{\partial(v_1,v_2)}$ at p. Therefore, by the Inverse Function Theorem, $(\pi \circ \mathsf{F}_\beta)^{-1}$ exists and is smooth near p. Since $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha = (\pi \circ \mathsf{F}_\beta)^{-1} \circ (\pi \circ \mathsf{F}_\alpha)$, and all of $(\pi \circ \mathsf{F}_\beta)^{-1}$, π and F_α are smooth maps, their composition is also a smooth map. We have proved $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ is smooth near p. Since p is arbitrary, $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ is in fact smooth on the domain $\mathsf{F}_\alpha^{-1}(\mathcal{W})$. □

Exercise 1.14. Rewrite the proof of Proposition 1.11, *mutatis mutandis*, by assuming $\det \dfrac{\partial(y,z)}{\partial(u_1,u_2)}(p) \neq 0$ instead.

1.4. Maps and Functions from Surfaces

Let M be a regular surface in $\mathbb{R}^3$ with a smooth local parametrization $\mathsf{F}(u_1, u_2) : \mathcal{U} \to M$. Then, for any $p \in \mathsf{F}(\mathcal{U})$, one can define the partial derivatives for a function $f : M \to \mathbb{R}$ at p as follows. The subtle issue is that the domain of f is the surface M, but by pre-composing f with F, i.e. $f \circ \mathsf{F}$, one can regard it as a map from $\mathcal{U} \subset \mathbb{R}^2$ to $\mathbb{R}$. With a little abuse of notations, we denote:

$$\frac{\partial f}{\partial u_j}(p) := \frac{\partial (f \circ \mathsf{F})}{\partial u_j}(u_1, u_2)$$

where (u_1, u_2) is the point corresponding to p, i.e. $\mathsf{F}(u_1, u_2) = p$.

Remark 1.12. Note that $\frac{\partial f}{\partial u_j}(p)$ is defined locally on $\mathsf{F}(\mathcal{U})$, and depends on the choice of local parametrization F near p. □

Definition 1.13 (Functions of Class C^k)**.** Let M be a regular surface in $\mathbb{R}^3$, and $f : M \to \mathbb{R}$ be a function defined on M. We say f is C^k at $p \in M$ if for any smooth local parametrization $\mathsf{F} : \mathcal{U} \to M$ with $p \in \mathsf{F}(\mathcal{U})$, the composition $f \circ \mathsf{F}$ is C^k at (u_1, u_2) corresponding to p.

If f is C^k at p for any $p \in M$, then we say that f is a C^k function on M. Here k can be taken to be ∞, and in such case we call f to be a C^∞ (or smooth) function.

Remark 1.14. Although we require $f \circ \mathsf{F}$ to be C^k at $p \in M$ for *any* local parametrization F in order to say that f is C^k, by Proposition 1.11 it suffices to show that $f \circ \mathsf{F}$ is C^k at p for *at least one* F near p. It is because

$$f \circ \widetilde{\mathsf{F}} = (f \circ \mathsf{F}) \circ (\mathsf{F}^{-1} \circ \widetilde{\mathsf{F}})$$

and compositions of C^k maps (between Euclidean spaces) are C^k. □

Example 1.15. Let M be a regular surface in $\mathbb{R}^3$, then each of the x, y and z coordinates in $\mathbb{R}^3$ can be regarded as a function from M to $\mathbb{R}$. For any smooth local parametrization $\mathsf{F} : \mathcal{U} \to M$ around p given by

$$\mathsf{F}(u_1, u_2) = (x(u_1, u_2), y(u_1, u_2), z(u_1, u_2)),$$

we have $x \circ \mathsf{F}(u_1, u_2) = x(u_1, u_2)$. Since F is C^∞, we get $x \circ \mathsf{F}$ is C^∞ as well. Therefore, the coordinate functions x, y and z for any regular surface is smooth. □

Example 1.16. Let $f : M \to \mathbb{R}$ be the function from a regular surface M in $\mathbb{R}^3$ defined by:

$$f(p) := |p - p_0|^2$$

where $p_0 = (x_0, y_0, z_0)$ is a fixed point of $\mathbb{R}^3$. Suppose $\mathsf{F}(u, v)$ is a local parametrization of M. We want to compute $\frac{\partial f}{\partial u}$ and $\frac{\partial f}{\partial v}$.

Write $(x, y, z) = \mathsf{F}(u, v)$ so that x, y and z are functions of (u, v). Then

$$\begin{aligned}
\frac{\partial f}{\partial u} &:= \frac{\partial}{\partial u}(f \circ \mathsf{F}) \\
&= \frac{\partial}{\partial u} f(x(u,v), y(u,v), z(u,v)) \\
&= \frac{\partial}{\partial u}\left((x(u,v) - x_0)^2 + (y(u,v) - y_0)^2 + (z(u,v) - z_0)^2\right) \\
&= 2(x - x_0)\frac{\partial x}{\partial u} + 2(y - y_0)\frac{\partial y}{\partial u} + 2(z - z_0)\frac{\partial z}{\partial u}
\end{aligned}$$

Note that we can differentiate x, y and z by u because $\mathsf{F}(u,v)$ is smooth. Similarly, we have:

$$\frac{\partial f}{\partial v} = 2(x-x_0)\frac{\partial x}{\partial v} + 2(y-y_0)\frac{\partial y}{\partial v} + 2(z-z_0)\frac{\partial z}{\partial v}.$$

Again since $\mathsf{F}(u,v)$ (and hence x, y and z) is a smooth function of (u,v), we can differentiate $\frac{\partial f}{\partial u}$ and $\frac{\partial f}{\partial v}$ as many times as we wish. This concludes that f is a smooth funciton. □

Exercise 1.15. Let $p_0(x_0, y_0, z_0)$ be a point in $\mathbb{R}^3$ and let $f(p) = |p - p_0|$ be the Euclidean distance between p and p_0 in $\mathbb{R}^3$. Suppose M is a regular surface in $\mathbb{R}^3$, one can then restrict the domain of f to M and consider it as a function:

$$f : M \to \mathbb{R}$$
$$p \mapsto |p - p_0|$$

Under what condition is the function $f : M \to \mathbb{R}$ smooth?

Now let M and N be two regular surfaces in $\mathbb{R}^3$. Then, one can also talk about mappings $\Phi : M \to N$ between them. In this section, we will define the notion of smooth maps between two surfaces.

Suppose $\mathsf{F} : \mathcal{U}_M \to M$ and $\mathsf{G} : \mathcal{U}_N \to N$ are two smooth local parametrizations of M and N respectively. One can then consider the composition $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ after shrinking the domain. It is then a map between open subsets of $\mathbb{R}^2$.

However, in order for this composition to be well-defined, we require the image of $\Phi \circ \mathsf{F}$ to be contained in the image of G, which is not always guaranteed. Let $\mathcal{W} := \Phi(\mathcal{O}_M) \cap \mathcal{O}_N$ be the overlapping region on N of these two images. Then, provided that $\mathcal{W} \neq \emptyset$, the composition $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ becomes well-defined as a map on:

$$\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F} : (\Phi \circ \mathsf{F})^{-1}(\mathcal{W}) \to \mathcal{U}_N.$$

From now on, whenever we talk about this composition $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$, we always implicitly assume that $\mathcal{W} \neq \emptyset$ and its domain is $(\Phi \circ \mathsf{F})^{-1}(\mathcal{W})$.

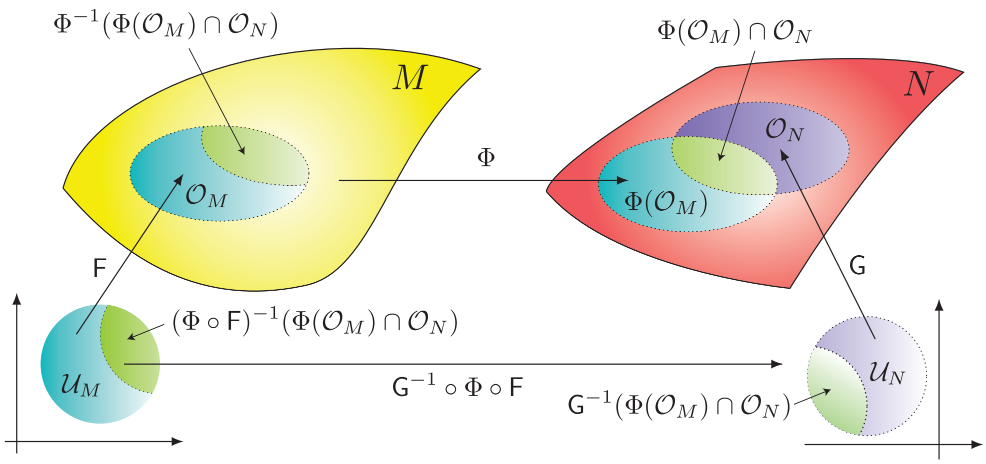


Figure 1.9. maps between regular surfaces

Definition 1.17 (Maps of Class C^k)**.** Let M and N be two regular surfaces in $\mathbb{R}^3$, and $\Phi : M \to N$ be a map between them. We say Φ is C^k at $p \in M$ if for any smooth local parametrization $\mathsf{F} : \mathcal{U}_M \to M$ with $p \in \mathsf{F}(\mathcal{U}_M)$, and $\mathsf{G} : \mathcal{U}_N \to N$ with $\Phi(p) \in \mathsf{G}(\mathcal{U}_N)$, the composition $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ is C^k at $\mathsf{F}^{-1}(p)$ as a map between subsets of $\mathbb{R}^2$.

If Φ is C^k at p for any $p \in M$, then we say that Φ is C^k on M. Here k can be taken to be ∞, and in such case we call Φ to be C^∞ (or smooth) on M.

Remark 1.18. Although we require $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ to be C^k at $p \in M$ for *any* local parametrizations F and G in order to say that Φ is C^k, by Proposition 1.11 it suffices to show that $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ is C^k at p for *at least one* pair of F and G. It is because

$$\widetilde{\mathsf{G}}^{-1} \circ \Phi \circ \widetilde{\mathsf{F}} = (\widetilde{\mathsf{G}}^{-1} \circ \mathsf{G}) \circ (\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}) \circ (\mathsf{F}^{-1} \circ \widetilde{\mathsf{F}})$$

and compositions of C^k maps (between Euclidean spaces) are C^k. □

Example 1.19. Let $\mathbb{S}^2$ be the unit sphere in $\mathbb{R}^3$. Consider the antipodal map $\Phi : \mathbb{S}^2 \to \mathbb{S}^2$ taking P to $-P$. In Example 1.4, two of the local parametrizations are given by:

$$\mathsf{F}_1(u_1, u_2) = \left(u_1,\ u_2,\ \sqrt{1 - u_1^2 - u_2^2}\right) : B_1(0) \subset \mathbb{R}^2 \to \mathbb{S}^2_+$$
$$\mathsf{F}_2(v_1, v_2) = \left(v_1,\ v_2,\ -\sqrt{1 - v_1^2 - v_2^2}\right) : B_1(0) \subset \mathbb{R}^2 \to \mathbb{S}^2_-$$

where $B_1(0)$ is the open unit disk in $\mathbb{R}^2$ centered at the origin, and $\mathbb{S}^2_+$ and $\mathbb{S}^2_-$ are the upper and lower hemispheres of $\mathbb{S}^2$ respectively. One can compute that:

$$\begin{aligned}
\mathsf{F}_2^{-1} \circ \Phi \circ \mathsf{F}_1(u_1, u_2) &= \mathsf{F}_2^{-1} \circ \Phi \left(u_1,\ u_2,\ \sqrt{1 - u_1^2 - u_2^2}\right) \\
&= \mathsf{F}_2^{-2} \left(-u_1,\ -u_2,\ -\sqrt{1 - u_1^2 - u_2^2}\right) \\
&= (-u_1,\ -u_2)
\end{aligned}$$

Clearly, the map $(u_1, u_2) \mapsto (-u_1, -u_2)$ is C^∞. It shows the antipodal map Φ is C^∞ at every point in $\mathsf{F}_1(B_1(0))$. One can show in similar way using other local parametrizations that Φ is C^∞ at points on $\mathbb{S}^2$ not covered by F_1.

Note that, for instance, the images of $\Phi \circ \mathsf{F}_1$ and F_1 are disjoint, and so $\mathsf{F}_1^{-1} \circ \Phi \circ \mathsf{F}_1$ is not well-defined. We don't need to verify whether it is smooth. □

Exercise 1.16. Let Φ be the antipodal map considered in Example 1.19, and F_+ and F_- be the two stereographic parametrizations of $\mathbb{S}^2$ defined in Example 1.5. Compute the maps $\mathsf{F}_+^{-1} \circ \Phi \circ \mathsf{F}_+$, $\mathsf{F}_-^{-1} \circ \Phi \circ \mathsf{F}_+$, $\mathsf{F}_+^{-1} \circ \Phi \circ \mathsf{F}_-$ and $\mathsf{F}_-^{-1} \circ \Phi \circ \mathsf{F}_-$. State their domains, and verify that they are smooth on their domains.

Exercise 1.17. Denote $\mathbb{S}^2$ to be the unit sphere $x^2 + y^2 + z^2 = 1$. Let $\Phi : \mathbb{S}^2 \to \mathbb{S}^2$ be the rotation map about the z-axis defined by:

$$\Phi(x, y, z) = (x \cos\alpha - y \sin\alpha, x \sin\alpha + y \cos\alpha, z)$$

where α is a fixed angle. Show that Φ is smooth.

Let M and N be two regular surfaces. If a map $\Phi : M \to N$ is C^∞, invertible, with C^∞ inverse map $\Phi^{-1} : N \to M$, then we say:

Definition 1.20 (Diffeomorphisms)**.** A map $\Phi : M \to N$ between two regular surfaces M and N in $\mathbb{R}^3$ is said to be a *diffeomorphism* if Φ is C^∞ and invertible, and also the inverse map Φ^{-1} is C^∞. If such a map Φ exists between M and N, then we say the surfaces M and N are *diffeomorphic*.

Example 1.21. The antipodal map $\Phi : \mathbb{S}^2 \to \mathbb{S}^2$ described in Example 1.19 is a diffeomorphism between $\mathbb{S}^2$ and itself. $\square$

Example 1.22. The sphere $x^2 + y^2 + z^2 = 1$ and the ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ are diffeomorphic, under the map $\Phi(x, y, z) = (ax, by, cz)$ restricted on $\mathbb{S}^2$. $\square$

Exercise 1.18. Given any pair of C^∞ functions $f, g : \mathbb{R}^2 \to \mathbb{R}$, show that the graphs Γ_f and Γ_g are diffeomorphic.

Exercise 1.19. Show that $\Phi : \mathbb{S}^2 \to \mathbb{S}^2$ defined in Exercise 1.17 is a diffeomorphism.

1.5. Tangent Planes and Tangent Maps

1.5.1. Tangent Planes of Regular Surfaces. The tangent plane is an important geometric object associated to a regular surface. Condition (3) of a smooth local parametrization $\mathsf{F}(u,v)$ requires that the cross-product $\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v}$ is non-zero for any (u,v) in the domain, or equivalently, both tangent vectors $\frac{\partial \mathsf{F}}{\partial u}$ and $\frac{\partial \mathsf{F}}{\partial v}$ must be non-zero vectors and they are non-parallel to each other.

Therefore, the two vectors $\frac{\partial \mathsf{F}}{\partial u}$ and $\frac{\partial \mathsf{F}}{\partial v}$ span a two-dimensional subspace in $\mathbb{R}^3$. We call this subspace the *tangent plane*, which is defined rigorously as follows:

Definition 1.23 (Tangent Plane)**.** Let M be a regular surface in $\mathbb{R}^3$ and p be a point on M. Suppose $\mathsf{F}(u,v) : \mathcal{U} \subset \mathbb{R}^2 \to M$ is a smooth local parametrization around p, then the *tangent plane* at p, denoted by T_pM, is defined as follows:

$$T_pM := \operatorname{span}\left\{\frac{\partial \mathsf{F}}{\partial u}(p), \frac{\partial \mathsf{F}}{\partial v}(p)\right\} = \left\{a\frac{\partial \mathsf{F}}{\partial u}(p) + b\frac{\partial \mathsf{F}}{\partial v}(p) : a, b \in \mathbb{R}\right\}.$$

Here we have abused the notations for simplicity: $\frac{\partial \mathsf{F}}{\partial u}(p)$ means $\frac{\partial \mathsf{F}}{\partial u}$ evaluated at $(u,v) = \mathsf{F}^{-1}(p)$. Similarly for $\frac{\partial \mathsf{F}}{\partial v}(p)$.

Rigorously, T_pM is a plane passing through the *origin* while $p + T_pM$ is the plane tangent to the surface at p (see Figure 1.10). The difference between T_pM and $p + T_pM$ is very subtle, and we will almost neglect this difference.

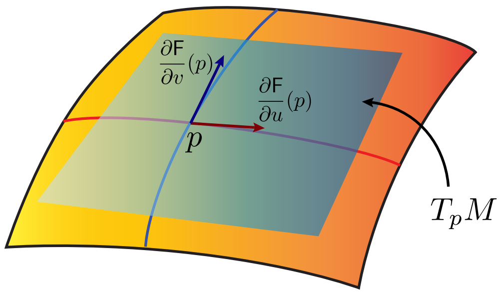


Figure 1.10. Tangent plane $p + T_pM$ at $p \in M$

Exercise 1.20. Show that the equation of the tangent plane $p + T_pM$ of the graph of a smooth function $f(x,y)$ at $p = (x_0, y_0, f(x_0,y_0))$ is given by:

$$z = f(x_0,y_0) + \left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)}(x - x_0) + \left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)}(y - y_0)$$

Exercise 1.21. [**dC76**, P.88] Consider the surface M given by $z = xf(y/x)$, where $x \neq 0$ and f is a smooth function. Show that the tangent planes $p + T_pM$ must pass through the origin $(0,0,0)$.

1.5.2. Tangent Maps between Regular Surfaces. Given a smooth map $\Phi : M \to N$ between two regular surfaces M and N, there is a naturally defined map called the *tangent map*, denoted by Φ_* in this course, between the tangent planes T_pM and $T_{\Phi(p)}N$.

Let us consider a smooth local parametrization $\mathsf{F}(u_1, u_2) : \mathcal{U}_M \to M$. The composition $\Phi \circ \mathsf{F}$ can be regarded as a map from $\mathcal{U}_M$ to $\mathbb{R}^3$, so one can talk about its partial derivatives $\frac{\partial(\Phi\circ\mathsf{F})}{\partial u_i}$:

$$\frac{\partial \Phi}{\partial u_i}(\Phi(p)) := \left.\frac{\partial(\Phi \circ \mathsf{F})}{\partial u_i}\right|_{(u_1,u_2)} = \left.\frac{d}{dt}\right|_{t=0} \Phi \circ \mathsf{F}((u_1, u_2) + t\mathsf{e}_i)$$

where (u_1, u_2) is a point in $\mathcal{U}_M$ such that $\mathsf{F}(u_1, u_2) = p$. The curve $\mathsf{F}((u_1, u_2) + t\mathsf{e}_i)$ is a curve on M with parameter t along the u_i-direction. The curve $\Phi \circ \mathsf{F}((u_1, u_2) + t\mathsf{e}_i)$ is then the image of the u_i-curve of M under the map Φ (see Figure 1.11). It is a curve on N so $\frac{\partial\Phi}{\partial u_i}$ which is a tangent vector to the surface N.

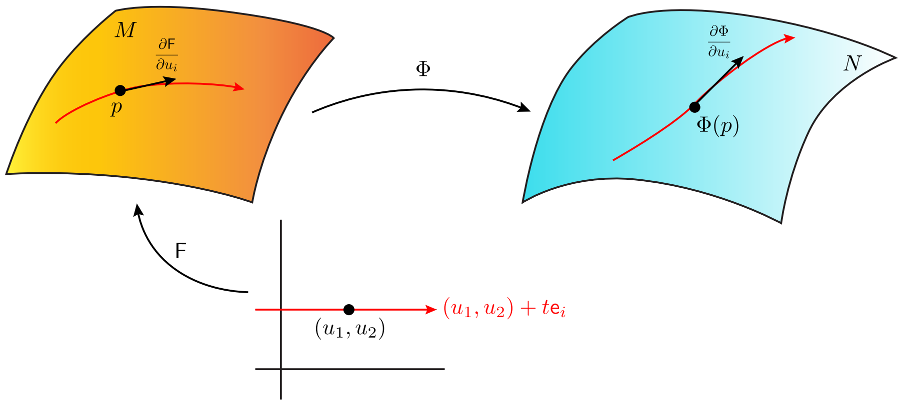


Figure 1.11. Partial derivative of the map $\Phi : M \to N$

Exercise 1.22. Denote $\mathbb{S}^2$ to be the unit sphere $x^2 + y^2 + z^2 = 1$. Let $\Phi : \mathbb{S}^2 \to \mathbb{S}^2$ be the rotation map about the z-axis defined by:

$$\Phi(x, y, z) = (x \cos\alpha - y \sin\alpha, x \sin\alpha + y \cos\alpha, z)$$

where α is a fixed angle. Calculate the following partial derivatives under the given local parametrizations:

(a) $\dfrac{\partial\Phi}{\partial\theta}$ and $\dfrac{\partial\Phi}{\partial\varphi}$ under $\mathsf{F}(\theta, \varphi) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$;

(b) $\dfrac{\partial\Phi}{\partial u}$ and $\dfrac{\partial\Phi}{\partial v}$ under F_2 in Example 1.4;

(c) $\dfrac{\partial\Phi}{\partial u}$ and $\dfrac{\partial\Phi}{\partial v}$ under F_+ in Example 1.5.

Next, we write the partial derivative $\frac{\partial\Phi}{\partial u_i}$ in a *fancy* way. Define:

$$\Phi_*\left(\frac{\partial \mathsf{F}}{\partial u_i}\right) := \frac{\partial\Phi}{\partial u_i}.$$

Then, one can regard Φ_* as a map that takes the tangent vector $\frac{\partial\mathsf{F}}{\partial u_i}$ in T_pM to another vector $\frac{\partial\Phi}{\partial u_j}$ in $T_{\Phi(p)}N$. Since $\left\{\frac{\partial\mathsf{F}}{\partial u_i}(p)\right\}$ is a basis of T_pM, one can then extend Φ_* linearly and define it as the *tangent map* of Φ. Precisely, we have:

Definition 1.24 (Tangent Maps)**.** Let $\Phi : M \to N$ be a smooth map between two regular surfaces M and N in $\mathbb{R}^3$. Let $\mathsf{F} : \mathcal{U}_M \to M$ and $\mathsf{G} : \mathcal{U}_N \to N$ be two smooth local parametrizations covering p and $\Phi(p)$ respectively. Then, the *tangent map* of Φ at $p \in M$ is denoted by $(\Phi_*)_p$ and is defined as:

$$(\Phi_*)_p : T_pM \to T_{\Phi(p)}N$$
$$(\Phi_*)_p \left(\sum_{i=1}^{2} a_i \frac{\partial \mathsf{F}}{\partial u_i}(p) \right) = \sum_{i=1}^{2} a_i \frac{\partial \Phi}{\partial u_i}(\Phi(p))$$

If the point p is clear from the context, $(\Phi_*)_p$ can be simply denoted by Φ_*.

Remark 1.25. Some textbooks may use $d\Phi_p$ to denote the tangent map of Φ at p. □

Example 1.26. Consider the unit sphere $\mathbb{S}^2$ locally parametrized by

$$\mathsf{F}(\theta, \varphi) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$$

and the rotation map:

$$\Phi(x, y, z) = (x\cos\alpha - y\sin\alpha, x\sin\alpha + y\cos\alpha, z)$$

From Exercise 1.22, one should have figured out that:

$$\frac{\partial \Phi}{\partial \theta} = (-\sin\varphi\sin(\theta+\alpha), \sin\varphi\cos(\theta+\alpha), 0)$$
$$\frac{\partial \Phi}{\partial \varphi} = (\cos\varphi\cos(\theta+\alpha), \cos\varphi\sin(\theta+\alpha), -\sin\varphi)$$

Next we want to write them in terms of the basis $\left\{ \frac{\partial \mathsf{F}}{\partial \theta}, \frac{\partial \mathsf{F}}{\partial \varphi} \right\}$. However, we should be careful about the base points of these vectors. Consider a point $p \in \mathbb{S}^2$ with local coordinates (θ, φ), the vectors $\frac{\partial \Phi}{\partial \theta}$ and $\frac{\partial \Phi}{\partial \varphi}$ computed above are based at the point $\Phi(p)$ with local coordinates $(\theta+\alpha, \varphi)$. Therefore, we should express them in terms of the basis $\left\{ \frac{\partial \mathsf{F}}{\partial \theta}(\Phi(p)), \frac{\partial \mathsf{F}}{\partial \varphi}(\Phi(p)) \right\}$, not $\left\{ \frac{\partial \mathsf{F}}{\partial \theta}(p), \frac{\partial \mathsf{F}}{\partial \varphi}(p) \right\}$!

At $\Phi(p)$, we have:

$$\frac{\partial \mathsf{F}}{\partial \theta}(\Phi(p)) = (-\sin\varphi\sin(\theta+\alpha), \sin\varphi\cos(\theta+\alpha), 0) = \frac{\partial \Phi}{\partial \theta}(\Phi(p))$$
$$\frac{\partial \mathsf{F}}{\partial \varphi}(\Phi(p)) = (\cos\varphi\cos(\theta+\alpha), \cos\varphi\sin(\theta+\alpha), -\sin\varphi) = \frac{\partial \Phi}{\partial \varphi}(\Phi(p))$$

Therefore, the tangent map $(\Phi_*)_p$ acts on the basis vectors by:

$$(\Phi_*)_p \left(\frac{\partial \mathsf{F}}{\partial \theta}(p) \right) = \frac{\partial \mathsf{F}}{\partial \theta}(\Phi(p))$$
$$(\Phi_*)_p \left(\frac{\partial \mathsf{F}}{\partial \varphi}(p) \right) = \frac{\partial \mathsf{F}}{\partial \varphi}(\Phi(p))$$

In other words, the matrix representation $[(\Phi_*)_p]$ with respect to the bases

$$\left\{ \frac{\partial \mathsf{F}}{\partial \theta}(p), \frac{\partial \mathsf{F}}{\partial \varphi}(p) \right\} \text{ for } T_p\mathbb{S}^2 \qquad \left\{ \frac{\partial \mathsf{F}}{\partial \theta}(\Phi(p)), \frac{\partial \mathsf{F}}{\partial \varphi}(\Phi(p)) \right\} \text{ for } T_{\Phi(p)}\mathbb{S}^2$$

is the identity matrix. However, it is not perfectly correct to say $(\Phi_*)_p$ is an identity map, since the domain and co-domain are different tangent planes. □

Exercise 1.23. Let Φ be as in Example 1.26. Consider the stereographic parametrization $\mathsf{F}_+(u,v)$ defined in Example 1.5. Suppose $p \in \mathbb{S}^2$, express the matrix representation $[(\Phi_*)_p]$ with respect to the bases $\left\{\frac{\partial \mathsf{F}_+}{\partial u}, \frac{\partial \mathsf{F}_+}{\partial v}\right\}_p$ and $\left\{\frac{\partial \mathsf{F}_+}{\partial u}, \frac{\partial \mathsf{F}_+}{\partial v}\right\}_{\Phi(p)}$

1.5.3. Tangent Maps and Jacobian Matrices. Let $\Phi : M \to N$ be a smooth map between two regular surfaces. Instead of computing the matrix representation of the tangent map Φ_* directly by taking partial derivatives (c.f. Example 1.26), one can also find it out by computing a Jacobian matrix.

Suppose $\mathsf{F}(u_1,u_2) : \mathcal{U}_M \to M$ and $\mathsf{G}(v_1,v_2) : \mathcal{U}_N \to N$ are local parametrizations of M and N. The composition $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ can be regarded as a map between the u_1u_2-plane to the v_1v_2-plane. As such, one can write

$$\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1,u_2) = (v_1(u_1,u_2), v_2(u_1,u_2)).$$

By considering $\Phi \circ \mathsf{F}(u_1,u_2) = \mathsf{G}(v_1(u_1,u_2), v_2(u_1,u_2))$, one can differentiate both sides with respect to u_i:

$$\frac{\partial}{\partial u_i}(\Phi \circ \mathsf{F}) = \frac{\partial}{\partial u_i}\mathsf{G}(v_1(u_1,u_2), v_2(u_1,u_2)) = \sum_{k=1}^{2} \frac{\partial \mathsf{G}}{\partial v_k}\frac{\partial v_k}{\partial u_i}. \tag{1.1}$$

Here we used the chain rule. Note that $\left\{\frac{\partial \mathsf{G}}{\partial v_k}\right\}$ is a basis for $T_{\Phi(p)}N$.

Using (1.1), one can see:

$$\Phi_*\left(\frac{\partial \mathsf{F}}{\partial u_1}\right) := \frac{\partial \Phi}{\partial u_1} = \frac{\partial}{\partial u_1}(\Phi \circ \mathsf{F}) = \frac{\partial v_1}{\partial u_1}\frac{\partial \mathsf{G}}{\partial v_1} + \frac{\partial v_2}{\partial u_1}\frac{\partial \mathsf{G}}{\partial v_2}$$
$$\Phi_*\left(\frac{\partial \mathsf{F}}{\partial u_2}\right) := \frac{\partial \Phi}{\partial u_2} = \frac{\partial}{\partial u_2}(\Phi \circ \mathsf{F}) = \frac{\partial v_1}{\partial u_2}\frac{\partial \mathsf{G}}{\partial v_1} + \frac{\partial v_2}{\partial u_2}\frac{\partial \mathsf{G}}{\partial v_2}$$

Hence the matrix representation of $(\Phi_*)_p$ with respect to the bases $\left\{\frac{\partial \mathsf{F}}{\partial u_i}(p)\right\}$ and $\left\{\frac{\partial \mathsf{G}}{\partial v_i}(\Phi(p))\right\}$ is the Jacobian matrix:

$$\left.\frac{\partial(v_1,v_2)}{\partial(u_1,u_2)}\right|_{\mathsf{F}^{-1}(p)} = \begin{bmatrix} \frac{\partial v_1}{\partial u_1} & \frac{\partial v_1}{\partial u_2} \\ \frac{\partial v_2}{\partial u_1} & \frac{\partial v_2}{\partial u_2} \end{bmatrix}_{\mathsf{F}^{-1}(p)}$$

Example 1.27. Let $\Phi : \mathbb{S}^2 \to \mathbb{S}^2$ be the rotation map as in Example 1.26. Consider again the local parametrization:

$$\mathsf{F}(\theta,\varphi) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi).$$

By standard trigonometry, one can find out that $\Phi(\mathsf{F}(\theta,\varphi)) = \mathsf{F}(\theta+\alpha,\varphi)$. Equivalently, the map $\mathsf{F}^{-1} \circ \Phi \circ \mathsf{F}$ (in a suitable domain) is the map:

$$(\theta,\varphi) \mapsto (\theta+\alpha,\varphi).$$

As α is a constant, the Jacobian matrix of $\mathsf{F}^{-1} \circ \Phi \circ \mathsf{F}$ is the identity matrix, and so the matrix $[(\Phi_*)_p]$ with respect to the bases $\left\{\frac{\partial \mathsf{F}}{\partial \theta}, \frac{\partial \mathsf{F}}{\partial \varphi}\right\}_p$ and $\left\{\frac{\partial \mathsf{F}}{\partial \theta}, \frac{\partial \mathsf{F}}{\partial \varphi}\right\}_{\Phi(p)}$ is the identity matrix (which was also obtained by somewhat tedious computations in Example 1.26). □

Exercise 1.24. Do Exercise 1.23 by considering Jacobian matrices.

Chapter 2

Abstract Manifolds

"Manifolds are a bit like pornography: hard to define, but you know one when you see one."

Shmuel Weinberger

2.1. Smooth Manifolds

Intuitively, a manifold is a space which locally resembles an Euclidean space. Regular surfaces are examples of manifolds. Being locally Euclidean, a manifold is equipped with a local coordinate system around every point so that many concepts in Calculus on Euclidean spaces can carry over to manifolds.

Unlike regular surfaces, we do *not* require a manifold to be a subset of $\mathbb{R}^n$. A manifold can just stand alone by itself like the Universe is regarded as a curved space-time sheet with nothing "outside" in General Relativity. However, we do require that a manifold satisfies certain *topological* conditions.

2.1.1. Point-Set Topology. In order to state the formal definition of a manifold, there are some topological terms (such as Hausdorff, second countable, etc.) we will briefly introduce. However, we will not take a long detour to go through every single topological concept, otherwise we will not have time to cover the more interesting material about smooth manifolds. Moreover, these topological conditions are very common as long as the space we are looking at is not "strange".

A *topological space* X is a set equipped with a collection $\mathcal{T}$ of subsets of X such that:

(a) $\varnothing, X \in \mathcal{T}$; and

(b) for any arbitrary sub-collection $\{U_\alpha\}_{\alpha\in\mathcal{A}} \subset \mathcal{T}$, we have $\bigcup_{\alpha\in\mathcal{A}} U_\alpha \in \mathcal{T}$; and

(c) for any finite sub-collection $\{U_1, \ldots, U_N\} \subset \mathcal{T}$, we have $\bigcap_{i=1}^{N} U_\alpha \in \mathcal{T}$.

If $\mathcal{T}$ is such a collection, we call $\mathcal{T}$ a *topology* of X. Elements in $\mathcal{T}$ are called *open sets* of X.

Example 2.1. The Euclidean space $\mathbb{R}^n$ equipped with the collection

$$\mathcal{T} = \text{collection of all open sets (in usual sense) in } \mathbb{R}^n$$

is an example of a topological space. The collection $\mathcal{T}$ is called the *usual topology* of $\mathbb{R}^n$. □

Example 2.2. Any subset $S \subset \mathbb{R}^n$, equipped with the collection

$$\mathcal{T}_S = \{S \cap U : U \text{ is an open set (in usual sense) in } \mathbb{R}^n\}$$

is an example of a topological space. The collection $\mathcal{T}_S$ is called the *subspace topology*. □

Given two topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$, one can talk about functions or mapping between them. A map $\Phi : X \to Y$ is said to be *continuous* with respect to $\mathcal{T}_X$ and $\mathcal{T}_Y$ if for any $U \in \mathcal{T}_Y$, we have $\Phi^{-1}(U) \in \mathcal{T}_X$. This definition is a generalization of continuous functions between Euclidean spaces equipped with the usual topologies. If the map $\Phi : X \to Y$ is one-to-one and onto, and both Φ and Φ^{-1} are continuous, then we say Φ is a *homeomorphism* and the spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ are *homeomorphic*.

A topological space $(X, \mathcal{T})$ is said to be *Hausdorff* if for any pair of distinct points $p, q \in X$, we have $U_1, U_2 \in \mathcal{T}$ such that $p \in U_1$, $q \in U_2$ and $U_1 \cap U_2 = \emptyset$. In other words, points of a Hausdorff space can be *separated* by open sets. It is intuitive that $\mathbb{R}^n$ with the usual topology is a Hausdorff space. Any subset $S \subset \mathbb{R}^n$ with subspace topology is also a Hausdorff space.

A topological space $(X, \mathcal{T}_X)$ is said to be *second countable* if there is a countable sub-collection $\{U_i\}_{i=1}^{\infty} \subset \mathcal{T}$ such that any set $U \in \mathcal{T}$ can be expressed as a union of some of these U_i's. For instance, $\mathbb{R}^n$ with usual topology is second countable since by density of rational numbers, any open set can be expressed as a countable union of open balls with rational radii and centers.

This introduction to point-set topology is intended to be short. It may not make sense to everybody, but it doesn't hurt! Point-set topology is not the main dish of the course. Many spaces we will look at are either Euclidean spaces, their subsets or sets *derived* from Euclidean spaces. Most of them are Hausdorff and second countable. Readers who want to learn more about point-set topology may consider taking MATH 4225. For more thorough treatment on point-set topology, please consult [**Mun00**]. Meanwhile, the *take-home message* of this introduction is that we don't have to worry much about point-set topology in this course!

2.1.2. Definitions and Examples. Now we are ready to learn what a manifold is. We will first introduce topological manifolds, which are objects that locally look like Euclidean space in certain *continuous* sense:

Definition 2.3 (Topological Manifolds)**.** A Hausdorff, second countable topological space M is said to be an *n-dimensional topological manifold*, or in short a *topological n-manifold*, if for any point $p \in M$, there exists a homeomorphism $\mathsf{F} : \mathcal{U} \to \mathcal{O}$ between a non-empty open subset $\mathcal{U} \subset \mathbb{R}^n$ and an open subset $\mathcal{O} \subset M$ containing p. This homeomorphism F is called a *local parametrization* (or *local coordinate chart*) around p.

Example 2.4. Any regular surface is a topological manifold since its local parametrizations are all homeomorphisms. Therefore, spheres, cylinders, torus, etc. are all topological manifolds.

However, a double cone (see Figure 2.1) is not a topological manifold since the vertex is a "bad" point. Any open set containing the vertex cannot be homeomorphic to any open set in Euclidean space.

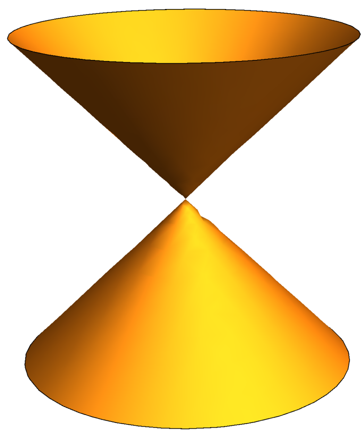


Figure 2.1. Double cone is not locally Euclidean near its vertex.

Remark 2.5. Note that around every p there may be more than one local parametrizations. If $\mathsf{F}_\alpha : \mathcal{U}_\alpha \to \mathcal{O}_\alpha$ and $\mathsf{F}_\beta : \mathcal{U}_\beta \to \mathcal{O}_\beta$ are two local parametrizations around p, then the composition:

$$(\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha) : \mathsf{F}_\alpha^{-1}(\mathcal{O}_\alpha \cap \mathcal{O}_\beta) \to \mathsf{F}_\beta^{-1}(\mathcal{O}_\alpha \cap \mathcal{O}_\beta)$$
$$(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta) : \mathsf{F}_\beta^{-1}(\mathcal{O}_\alpha \cap \mathcal{O}_\beta) \to \mathsf{F}_\alpha^{-1}(\mathcal{O}_\alpha \cap \mathcal{O}_\beta)$$

are often called the *transition maps* between these local parametrizations. We need to restrict their domains to smaller sets so as to guarantee the transition maps are well-defined (c.f. Section 1.3). □

On a topological manifold, there is a coordinate system around every point. However, many concepts in Calculus involve taking *derivatives*. In order to carry out differentiations on manifolds, it is not sufficient to be merely locally *homeomorphic* to Euclidean spaces. We need the local parametrization F to be *differentiable* in a certain sense.

For regular surfaces in $\mathbb{R}^3$, the local parametrization $\mathsf{F} : \mathcal{U} \to \mathbb{R}^3$ are maps between Euclidean spaces, so it makes sense to take derivatives of F. However, an abstract manifold may not be sitting in $\mathbb{R}^3$ or $\mathbb{R}^N$, and therefore it is difficult to make of sense of differentiability of $\mathsf{F} : \mathcal{U} \to \mathcal{O}$. To get around this issue, we will *not* talk about the differentiability of a local parametrization F, but instead talk about the differentiability of *transition maps*.

In Proposition 1.11 of Chapter 1 we showed that any two overlapping local parametrizations F_α and F_β of a regular surface M have smooth transition maps $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ and $\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta$. Now consider an abstract topological manifold. Although the local parametrizations F_α and F_β may not have a codomain sitting in Euclidean spaces, the transition maps $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ and $\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta$ are indeed maps between open subsets of Euclidean spaces!

While we cannot differentiate local parametrizations $\mathsf{F} : \mathcal{U} \to \mathcal{O} \subset M$ for abstract manifolds, we can do so for the transition maps $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ and $\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta$. This motivates the definition of a smooth manifold:

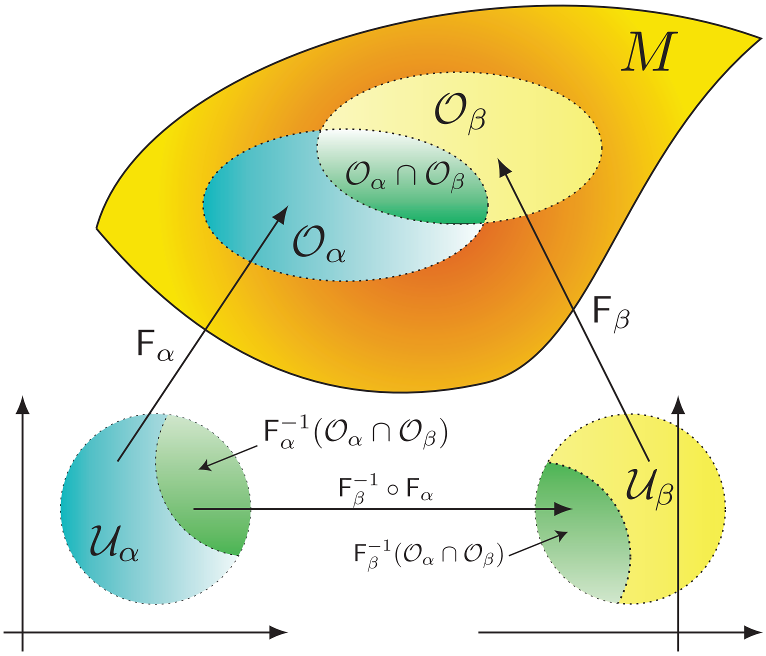


Figure 2.2. transition maps of a manifold

Definition 2.6 (Smooth Manifolds)**.** A n-dimensional topological manifold M is said to be an *n-dimensional smooth manifold*, or in short a *smooth n-manifold*, if there is a collection $\mathcal{A}$ of local parametrizations $\mathsf{F}_\alpha : \mathcal{U}_\alpha \to \mathcal{O}_\alpha$ such that

(1) $\bigcup_{\alpha \in \mathcal{A}} \mathcal{O}_\alpha = M$, i.e. these local parametrizations cover all of M; and

(2) all transition maps $\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta$ are smooth (i.e. C^∞) on their domains.

Remark 2.7. Two local parametrizations F_α and F_β with smooth transition maps $\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta$ and $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$ are said to be *compatible*. $\square$

Remark 2.8. We often use the superscript n, i.e. M^n, to mean that the manifold M is n-dimensional. $\square$

Remark 2.9. A 2-dimensional manifold is sometimes called a *surface*. In this course, we will use the term *regular surfaces* for those surfaces in $\mathbb{R}^3$ discussed in Chapter 1, while we will use the term *smooth surfaces* to describe 2-dimensional smooth manifolds in the sense of Definition 2.6. $\square$

Example 2.10. Any topological manifold which can be covered by one global parametrization (i.e. image of F is all of M) is a smooth manifold. Examples of which include $\mathbb{R}^n$ which can be covered by one parametrization $\mathrm{Id} : \mathbb{R}^n \to \mathbb{R}^n$. The graph of Γ_f of any continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is also a smooth manifold covered by one parametrization $\mathsf{F}(\mathsf{x}) = (\mathsf{x}, f(\mathsf{x})) : \mathbb{R}^n \to \Gamma_f$. Any regular curve $\mathsf{r}(t)$ is a smooth manifold of dimension 1. $\square$

Example 2.11. All regular surfaces in $\mathbb{R}^3$ are smooth manifolds by Proposition 1.11 (which we showed their transition maps are smooth). Therefore, spheres, cylinders, tori, etc. are all smooth manifolds. □

Example 2.12 (Extended complex plane)**.** Define $M = \mathbb{C} \cup \{\infty\}$. One can show (omitted here) that it is a Hausdroff, second countable topological space. Furthermore, one can cover M by two local parametrizations:

$$\begin{aligned} \mathsf{F}_1 : \mathbb{R}^2 &\to \mathbb{C} \subset M & \qquad \mathsf{F}_2 : \mathbb{R}^2 &\to (\mathbb{C}\backslash\{0\}) \cup \{\infty\} \subset M \\ (x,y) &\mapsto x + yi & (x,y) &\mapsto \frac{1}{x+yi} \end{aligned}$$

The overlap part on M is given by $\mathbb{C}\backslash\{0\}$, corresponding to $\mathbb{R}^2\backslash\{(0,0)\}$ in $\mathbb{R}^2$ under the parametrizations F_1 and F_2. One can compute that the transition maps are given by:

$$\mathsf{F}_2^{-1} \circ \mathsf{F}_1(x,y) = \left(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}\right)$$
$$\mathsf{F}_1^{-1} \circ \mathsf{F}_2(x,y) = \left(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}\right)$$

Both are smooth maps on $\mathbb{R}^2\backslash\{(0,0)\}$. Therefore, $\mathbb{C} \cup \{\infty\}$ is a smooth manifold. □

Exercise 2.1. Show that the n-dimensional sphere

$$\mathbb{S}^n := \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} : x_1^2 + \ldots + x_{n+1}^2 = 1\}$$

is a smooth n-manifold. [Hint: Generalize the stereographic projection to higher dimensions]

Exercise 2.2. Discuss: According to Example 2.10, the graph of any *continuous* function $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth manifold as there is no transition map. However, wouldn't it imply the single cone:

$$\left\{(x,y,z) \in \mathbb{R}^3 : z = \sqrt{x^2+y^2}\right\}$$

is a smooth manifold? It appears to have a "corner" point at the vertex, isn't it?

2.1.3. Product and Quotient Manifolds. Given two smooth manifolds M^m and N^n, one can form an $(m+n)$-dimensional manifold $M^m \times N^n$, which is defined by:

$$M^m \times N^n := \{(x,y) : x \in M^m \text{ and } y \in N^n\}.$$

Given a local parametrization $\mathsf{F} : \mathcal{U}_M \to \mathcal{O}_M$ for M^m, and a local parametrizaiton $\mathsf{G} : \mathcal{U}_N \to \mathcal{O}_N$ for N^n, one can define a local parametrization:

$$\begin{aligned} \mathsf{F} \times \mathsf{G} : \mathcal{U}_M \times \mathcal{U}_N &\to \mathcal{O}_M \times \mathcal{O}_N \subset M^m \times N^n \\ (\mathsf{u}, \mathsf{v}) &\mapsto (\mathsf{F}(\mathsf{u}), \mathsf{G}(\mathsf{v})) \end{aligned}$$

If $\{\mathsf{F}_\alpha\}$ is a collection of local parametrizations of M^m with smooth transition maps, and $\{\mathsf{G}_\beta\}$ is that of N^n with smooth transition maps, then one can form a collection of local parametrizations $\mathsf{F}_\alpha \times \mathsf{G}_\beta$ of the product $M^m \times N^n$. It can be shown that these local parametrizations of $M^m \times N^n$ also have smooth transition maps between open subsets of $\mathbb{R}^{m+n}$ (see Exercise 2.3).

Exercise 2.3. Show that if F_α and $\mathsf{F}_{\widetilde{\alpha}}$ are local parametrizations of M^m with smooth transition maps, and similarly for G_β and $\mathsf{G}_{\widetilde{\beta}}$ for N^n, then $\mathsf{F}_\alpha \times \mathsf{G}_\beta$ and $\mathsf{F}_{\widetilde{\alpha}} \times \mathsf{G}_{\widetilde{\beta}}$ have smooth transition maps.

The result from Exercise 2.3 showed that the product $M^m \times N^n$ of two smooth manifolds M^m and N^n is a smooth manifold with dimension $m+n$. Inductively, the product $M_1^{m_1} \times \ldots M_k^{m_k}$ of k smooth manifolds $M_1^{m_1}, \ldots, M_k^{m_k}$ is a smooth manifold with dimension $m_1 + \ldots + m_k$.

Example 2.13. The cylinder $x^2 + y^2 = 1$ in $\mathbb{R}^3$ can be regarded as $\mathbb{R} \times \mathbb{S}^1$. The torus can be regarded as $\mathbb{S}^1 \times \mathbb{S}^1$. They are both smooth manifolds. By taking products of known smooth manifolds, one can generate a great deal of new smooth manifolds. The n-dimensional cylinder can be easily seen to be a smooth manifold by regarding it as $\mathbb{R} \times \mathbb{S}^{n-1}$. The n-dimensional torus $\underbrace{\mathbb{S}^1 \times \ldots \times \mathbb{S}^1}_{n \text{ times}}$ is also a smooth manifold. □

Another common way to produce a new manifold from an old one is to take quotients. Take $\mathbb{R}$ as an example. Let us define an equivalence relation $\sim$ by declaring that $x \sim y$ if and only if $x - y$ is an integer. For instance, we have $3 \sim 5$ while $4 \not\sim \frac{9}{2}$. Then, we can talk about *equivalence classes* $[x]$ which is the following set:

$$[x] := \{y \in \mathbb{R} : y \sim x\}.$$

For instance, we have $5 \in [2]$ as $5 \sim 2$. Likewise $-3 \in [2]$ as $-3 \sim 2$ too. The set $[2]$ is the set of all integers. Similarly, one can also argue $[-1] = [0] = [1] = [2] = [3] = \ldots$ are all equal to the set of all integers.

On the contrary, $1 \notin [0.2]$ as $1 \not\sim 0.2$. Yet $-1.8, -0.8, 0.2, \ldots$ are all in the set $[0.2]$. The set $[0.2]$ is simply the set of all numbers in the form of $0.2 + N$ where N is any integer. One can also see that $[-1.8] = [-0.8] = [0.2] = [1.2] = \ldots$.

Under such notations, we see that $[1] = [2]$ while $[1] \neq [0.2]$. The notion of equivalence classes provides us with a way to "decree" what elements in the "mother" set ($\mathbb{R}$ in this case) are regarded as equal. This is how topologists and geometers interpret *gluing*. In this example, we can think of 1, 2, 3, etc. are glued together, and also $-1.8, -0.8, 0.2$, etc. are glued together. Formally, we denote

$$\mathbb{R}/\sim \; := \{[x] : x \in \mathbb{R}\}$$

which is the set of all equivalence classes under the relation $\sim$. This new set $\mathbb{R}/\sim$ is called a *quotient set* of $\mathbb{R}$ by the equivalence relation $\sim$. By sketching the set, we can see $\mathbb{R}/\sim$ is topologically a circle $\mathbb{S}^1$ (see Figure 2.3):

Exercise 2.4. Describe the set $\mathbb{R}^2/\sim$ where we declare $(x_1, y_1) \sim (x_2, y_2)$ if and only if $x_1 - x_2 \in \mathbb{Z}$ and $y_1 - y_2 \in \mathbb{Z}$.

Example 2.14 (Real Projective Space)**.** The real projective space $\mathbb{RP}^n$ is the quotient set of $\mathbb{R}^{n+1}\backslash\{0\}$ under the equivalence relation: $(x_0, x_1, \ldots, x_n) \sim (y_0, y_1, \ldots, y_n)$ if and only if there exists $\lambda \in \mathbb{R}\backslash\{0\}$ such that $(x_0, x_1, \ldots, x_n) = (\lambda y_0, \lambda y_1, \ldots, \lambda y_n)$. Each equivalence class is commonly denoted by:

$$[x_0 : x_1 : \cdots : x_n]$$

For instance, we have $[0 : 1 : -1] = [0 : -\pi : \pi]$. Under this notation, we can write:

$$\mathbb{RP}^n := \left\{[x_0 : x_1 : \cdots : x_n] : (x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1}\backslash\{0\}\right\}$$

It is important to note that $[0 : 0 : \cdots : 0] \notin \mathbb{RP}^n$.

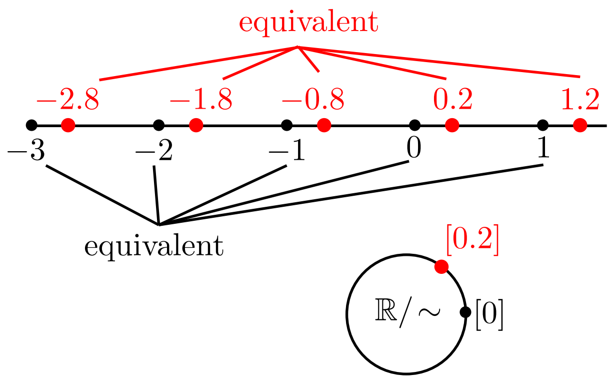


Figure 2.3. Quotient set $\mathbb{R}/\sim$

We are going to show that $\mathbb{RP}^n$ is an n-dimensional smooth manifold. For each $i = 0, 1, \ldots, n$, we denote:

$$\mathcal{O}_i := \{[x_0 : x_1 : \cdots : x_n] \in \mathbb{RP}^n : x_i \neq 0\} \subset \mathbb{RP}^n.$$

Define $\mathsf{F}_i : \mathbb{R}^n \to \mathcal{O}_i$ by:

$$\mathsf{F}_i(x_1, \ldots, x_n) = [x_1 : \cdots : \underbrace{1}_{i} : \ldots x_n].$$

For instance, $\mathsf{F}_0(x_1, \ldots, x_n) = [1 : x_1 : \cdots : x_n]$, $\mathsf{F}_1(x_1, \ldots, x_n) = [x_1 : 1 : x_2 : \cdots : x_n]$ and $\mathsf{F}_n(x_1, \ldots, x_n) = [x_1 : \cdots : x_n : 1]$.

The overlap between images of F_0 and F_1, for instance, is given by:

$$\mathcal{O}_0 \cap \mathcal{O}_1 = \{[x_0 : x_1 : x_2 : \cdots : x_n] : x_0, x_1 \neq 0\}$$

$$\mathsf{F}_0^{-1}(\mathcal{O}_0 \cap \mathcal{O}_1) = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1 \neq 0\}$$

One can compute that the transition map $\mathsf{F}_1^{-1} \circ \mathsf{F}_0$ is given by:

$$\mathsf{F}_1^{-1} \circ \mathsf{F}_0(x_1, \ldots, x_n) = \left(\frac{1}{x_1}, \frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1}\right)$$

which is smooth on the domain $\mathsf{F}_0^{-1}(\mathcal{O}_0 \cap \mathcal{O}_1)$. The smoothness of transition maps between any other pairs can be verified in a similar way. □

Exercise 2.5. Express the transition map $\mathsf{F}_3^{-1} \circ \mathsf{F}_1$ of $\mathbb{RP}^5$ and verify that it is smooth on its domain.

Example 2.15 (Complex Projective Space)**.** The complex projective space $\mathbb{CP}^n$ is an important manifold in Complex Geometry (one of my research interests) and Algebraic Geometry. It is defined similarly as $\mathbb{RP}^n$, with all $\mathbb{R}$'s replaced by $\mathbb{C}$'s. Precisely, we declare for any two elements in $(z_0, \ldots, z_n), (w_0, \ldots, w_n) \in \mathbb{C}^{n+1} \backslash \{(0, \ldots, 0)\}$, we have $(z_0, \ldots, z_n) \sim (w_0, \ldots, w_n)$ if and only if there exists $\lambda \in \mathbb{C}\backslash\{0\}$ such that $z_i = \lambda w_i$ for any $i = 0, \ldots n$. Under this equivalence relation, the equivalence classes denoted by $[z_0 : z_1 : \cdots : z_n]$ constitute the complex projective space:

$$\mathbb{CP}^n := \{[z_0 : z_1 : \cdots : z_n] : z_i \text{ not all zero}\}.$$

It can be shown to be a smooth $2n$-manifold in exactly the same way as in $\mathbb{RP}^n$. □

Exercise 2.6. Show that $\mathbb{CP}^n$ is an $2n$-dimensional smooth manifold by constructing local parametrizations in a similar way as for $\mathbb{RP}^n$. For the transition maps, express one or two of them explicitly and verify that they are smooth. What's more can you say about the transition maps apart from being smooth?

Exercise 2.7. Consider the equivalence relation of $\mathbb{R}^n$ defined as follows:

$$\mathsf{x} \sim \mathsf{y} \text{ if and only if } \mathsf{x} - \mathsf{y} \in \mathbb{Z}^n$$

Show that $\mathbb{R}^n / \sim$ is a smooth n-manifold.

Example 2.16. The Klein Bottle K (see Figure 1.2a) cannot be put inside $\mathbb{R}^3$ without self-intersection, but it can be done in $\mathbb{R}^4$. It is covered by two local parametrizations given below:

$$\mathsf{F}_1 : \mathcal{U}_1 \to \mathbb{R}^4 \quad \text{where} \quad \mathcal{U}_1 := \{(u_1, v_1) : 0 < u_1 < 2\pi \text{ and } 0 < v_1 < 2\pi\}$$

$$\mathsf{F}_1(u_1, v_1) = \begin{bmatrix} (\cos v_1 + 2)\cos u_1 \\ (\cos v_1 + 2)\sin u_1 \\ \sin v_1 \cos \frac{u_1}{2} \\ \sin v_1 \sin \frac{u_1}{2} \end{bmatrix}.$$

$$\mathsf{F}_2 : \mathcal{U}_2 \to \mathbb{R}^4 \quad \text{where} \quad \mathcal{U}_2 := \{(u_2, v_2) : 0 < u_2 < 2\pi \text{ and } 0 < v_2 < 2\pi\}$$

$$\mathsf{F}_2(u_2, v_2) = \begin{bmatrix} -(\cos v_2 + 2)\cos u_2 \\ (\cos v_2 + 2)\sin u_2 \\ \sin v_2 \cos\left(\frac{u_2}{2} + \frac{\pi}{4}\right) \\ \sin v_2 \sin\left(\frac{u_2}{2} + \frac{\pi}{4}\right) \end{bmatrix}.$$

Geometrically speaking, the Klein bottle is generated by rotating the unit circle by two independent rotations, one parallel to the xy-plane, another parallel to the zw-plane. For geometric explanations for these parametrizations, see [**dC94**, P.36].

We leave it to readers to check that F_1 and F_2 are both injective and compatible with each other. It will show that K is a 2-manifold. □

Exercise 2.8. Consider the Klein bottle K given in Example 2.16.

(a) Show that both F_1 and F_2 are injective.

(b) Let $\mathcal{W} = \mathsf{F}_1(\mathcal{U}_1) \cap \mathsf{F}_2(\mathcal{U}_2)$. Find $\mathsf{F}_1^{-1}(\mathcal{W})$ and $\mathsf{F}_2^{-1}(\mathcal{W})$.

(c) Compute the transition maps $\mathsf{F}_2^{-1} \circ \mathsf{F}_1$ and $\mathsf{F}_1^{-1} \circ \mathsf{F}_2$ defined on the overlaps.

2.1.4. Differential Structures. A smooth manifold M^n is equipped with a collection smooth local parametrizations $\mathsf{F}_\alpha : \mathcal{U}_\alpha \subset \mathbb{R}^n \to \mathcal{O}_\alpha \subset M^n$ such that the images of these F_α's cover the entire manifold, i.e.

$$M = \bigcup_{\text{all } \alpha\text{'s}} \mathcal{O}_\alpha = \bigcup_{\text{all } \alpha\text{'s}} \mathsf{F}_\alpha(\mathcal{U}_\alpha).$$

These local parametrizations need to be compatible with each other in a sense that any overlapping parametrizations F_α and F_β must have smooth transition maps $\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta$ and $\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha$. Such a collection of local parametrizations $\mathcal{A} = \{\mathsf{F}_\alpha, \mathcal{U}_\alpha, \mathcal{O}_\alpha\}_\alpha$ is called a *smooth atlas* of M.

Given a smooth atlas $\mathcal{A}$ of M, we can enlarge the atlas by including more local parametrizations $\mathsf{F}_{\text{new}} : \mathcal{U}_{\text{new}} \to \mathcal{O}_{\text{new}}$ that are compatible to all local parametrizations in $\mathcal{A}$. The *differential structure* generated by an atlas $\mathcal{A}$ is a *gigantic* atlas that contains all local parametrizations which are compatible with every local parametrizations in $\mathcal{A}$ (for more formal definition, please read [**Lee09**, Section 1.3]).

Let's take the plane $\mathbb{R}^2$ as an example. It can be parametrized by at least three different ways:

- the identity map $\mathsf{F}_1 := \mathrm{id} : \mathbb{R}^2 \to \mathbb{R}^2$.
- the map $\mathsf{F}_2 : \mathbb{R}^2 \to (0, \infty) \times (0, \infty) \subset \mathbb{R}^2$, defined as:
$$\mathsf{F}_2(u, v) := (e^u, e^v).$$
- and pathologically, by $\mathsf{F}_3 : \mathbb{R}^2 \to \mathbb{R}^2$ defined as:
$$\mathsf{F}_3(u, v) = (u, v + |u|).$$

It is clear that $\mathsf{F}_1^{-1} \circ \mathsf{F}_2(u, v) = (e^u, e^v)$ and $\mathsf{F}_2^{-1} \circ \mathsf{F}_1(u, v) = (\log u, \log v)$ are smooth on the domains at which they are defined. Therefore, we say that F_1 and F_2 are compatible, and the differential structure generated by F_1 will contain F_2.

On the other hand, $\mathsf{F}_1^{-1} \circ \mathsf{F}_3(u, v) = (u, v + |u|)$ is not smooth, and so F_1 and F_3 are not compatible. Likewise, $\mathsf{F}_2^{-1} \circ \mathsf{F}_3(u, v) = (\log u, \log(v + |u|))$ is not smooth either. Therefore, F_3 does *not* belong to the differential structure generated by F_1 and F_2.

As we can see from above, a manifold M can have many distinct differential structures. In this course, when we talk about manifolds, we usually only consider *one* differential structure of the manifold, and very often we will only deal with the most "natural" differential structure such as the one generated by F_1 or F_2 above for $\mathbb{R}^3$, but not like the pathological one such as F_3. Therefore, we usually will not specify the differential structure when we talk about a manifold, unless it is necessary in some rare occasions.

Exercise 2.9. Show that any smooth manifold has uncountably many distinct differential structures. [Hint: Let $\mathbb{B}(1) := \{\mathsf{x} \in \mathbb{R}^n : |\mathsf{x}| < 1\}$, consider maps $\Psi_s : \mathbb{B}(1) \to \mathbb{B}(1)$ defined by $\Psi_s(\mathsf{x}) = |\mathsf{x}|^s \mathsf{x}$ where $s > 0$.]

2.2. Functions and Maps on Manifolds

2.2.1. Definitions and Examples. Let's first review how smooth functions $f : M \to \mathbb{R}$ and smooth maps $\Phi : M \to N$ are defined for regular surfaces (Definitions 1.13 and 1.17). Given two and $\mathsf{G} : \mathcal{U}_N \to \mathcal{O}_N \subset N$, the compositions $f \circ \mathsf{F}$ and $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ are functions or maps between Euclidean spaces. We say the function f is smooth if $f \circ \mathsf{F}$ is smooth for any local parametrizations $\mathsf{F} : \mathcal{U}_M \to \mathcal{O}_M \subset M$. We say Φ is smooth if $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ is smooth.

The definitions of differentiable functions and maps for regular surfaces carry over to abstract manifolds in a natural way:

Definition 2.17 (Functions and Maps of Class C^k)**.** Let M^m and N^n be two smooth manifolds of dimensions m and n respectively. Then:

A scalar-valued function $f : M \to \mathbb{R}$ is said to be C^k *at* $p \in M$ if for any smooth local parametrization $\mathsf{F} : \mathcal{U} \to M$ with $p \in \mathsf{F}(\mathcal{U})$, the composition $f \circ \mathsf{F}$ is C^k at the point $\mathsf{F}^{-1}(p) \in \mathcal{U}$ as a function from subset from $\mathbb{R}^m$ to $\mathbb{R}$. Furthermore, if $f : M \to \mathbb{R}$ is C^k at every $p \in M$, then we say f is C^k *on* M.

A map $\Phi : M \to N$ is said to be C^k *at* $p \in M$ if for any smooth local parametrization $\mathsf{F} : \mathcal{U}_M \to \mathcal{O}_M \subset M$ with $p \in \mathsf{F}(\mathcal{U}_M)$, and $\mathsf{G} : \mathcal{U}_N \to \mathcal{O}_N \subset N$ with $\Phi(p) \in \mathsf{G}(\mathcal{U}_N)$, the composition $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ is C^k at $\mathsf{F}^{-1}(p)$ as a map between subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$. Furthermore, if $\Phi : M \to N$ is C^k at every $p \in M$, then Φ is said to be C^k *on* M.

When k is ∞, we can also say that the function or map is *smooth*.

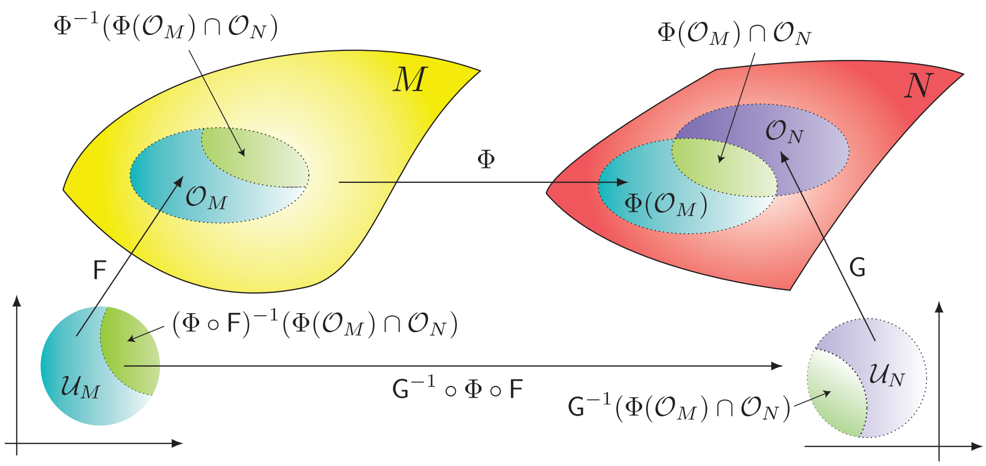


Figure 2.4. maps between two manifolds

Remark 2.18. By the definition of a smooth manifold (see condition (2) in Definition 2.6), transition maps are always smooth. Therefore, although we require $f \circ \mathsf{F}$ and $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ to be smooth for *any* local parametrizations around p, it suffices to show that they are smooth for at least *one* F covering p and at least *one* G covering $\Phi(p)$.

Exercise 2.10. Suppose $\Phi : M \to N$ and $\Psi : N \to P$ are C^k maps between smooth manifolds M, N and P. Show that the composition $\Psi \circ \Phi$ is also C^k.

Example 2.19. Consider the 3-dimensional sphere

$$\mathbb{S}^3 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1 \in \mathbb{R}\}$$

and the complex projective plane

$$\mathbb{CP}^1 = \{[z : w] : z \neq 0 \text{ or } w \neq 0\}.$$

Define a map $\Phi : \mathbb{S}^3 \to \mathbb{CP}^1$ by:

$$\Phi(x_1, x_2, x_3, x_4) = [x_1 + ix_2, x_3 + ix_4].$$

Locally parametrize $\mathbb{S}^3$ by stereographic projection:

$$\begin{aligned} \mathsf{F} &: \mathbb{R}^3 \to \mathbb{S}^3 \\ \mathsf{F}(u_1, u_2, u_3) &= \left(\frac{2u_1}{1 + \sum_k u_k^2}, \frac{2u_2}{1 + \sum_k u_k^2}, \frac{2u_3}{1 + \sum_k u_k^2}, \frac{-1 + \sum_k u_k^2}{1 + \sum_k u_k^2} \right) \end{aligned}$$

The image of F is $\mathbb{S}^3 \backslash \{(0, 0, 0, 1)\}$. As usual, we locally parametrize $\mathbb{CP}^1$ by:

$$\begin{aligned} \mathsf{G} &: \mathbb{R}^2 \to \mathbb{CP}^1 \\ \mathsf{G}(v_1, v_2) &= [1 : v_1 + iv_2] \end{aligned}$$

The domain of $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ is $(\Phi \circ \mathsf{F})^{-1} \left(\Phi \circ \mathsf{F}(\mathbb{R}^3) \cap \mathsf{G}(\mathbb{R}^2) \right)$, and the map is explicitly given by:

$$\begin{aligned} &\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1, u_2, u_3) \\ &= \mathsf{G}^{-1} \circ \Phi \left(\frac{2u_1}{1 + \sum_k u_k^2}, \frac{2u_2}{1 + \sum_k u_k^2}, \frac{2u_3}{1 + \sum_k u_k^2}, \frac{-1 + \sum_k u_k^2}{1 + \sum_k u_k^2} \right) \\ &= \mathsf{G}^{-1} \left[\frac{2u_1}{1 + \sum_k u_k^2} + i \frac{2u_2}{1 + \sum_k u_k^2} : \frac{2u_3}{1 + \sum_k u_k^2} + i \frac{-1 + \sum_k u_k^2}{1 + \sum_k u_k^2} \right] \\ &= \mathsf{G}^{-1} \left[1 : \frac{2u_3 + i \left(-1 + \sum_k u_k^2 \right)}{2u_1 + 2iu_2} \right] \\ &= G^{-1} \left[1 : \frac{2u_1 u_3 + u_2(-1 + \sum_k u_k^2)}{2(u_1^2 + u_2^2)} + i \frac{-2u_2 u_3 + u_1(-1 + \sum_k u_k^2)}{2(u_1^2 + u_2^2)} \right] \\ &= \left(\frac{2u_1 u_3 + u_2(-1 + \sum_k u_k^2)}{2(u_1^2 + u_2^2)}, \frac{-2u_2 u_3 + u_1(-1 + \sum_k u_k^2)}{2(u_1^2 + u_2^2)} \right). \end{aligned}$$

For any (u_1, u_2, u_3) in the domain of $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$, which is $(\Phi \circ \mathsf{F})^{-1} \left(\Phi \circ \mathsf{F}(\mathbb{R}^3) \cap \mathsf{G}(\mathbb{R}^2) \right)$, we have in particular $\Phi \circ \mathsf{F}(u_1, u_2, u_3) \in \mathsf{G}(\mathbb{R}^2)$, and so

$$\frac{2u_1}{1 + \sum_k u_k^2} + i \frac{2u_2}{1 + \sum_k u_k^2} \neq 0.$$

Therefore, $(u_1, u_2) \neq (0, 0)$ whenever (u_1, u_2, u_3) is in the domain of $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$. From the above computations, $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ is smooth.

One can also check similarly that $\widetilde{\mathsf{G}}^{-1} \circ \Phi \circ \widetilde{\mathsf{F}}$ is smooth for other combinations of local parametrizations, concluding Φ is a smooth map.

□

Example 2.20. Let $M \times N$ be the product of two smooth manifolds M and N. Then, the projection map π_M and π_N defined by:

$$\begin{aligned}\pi_M : M \times N &\to M \\ (p,q) &\mapsto p \\ \pi_N : M \times N &\to N \\ (p,q) &\mapsto q\end{aligned}$$

are both smooth manifolds. It can be shown by considering local parametrizations $\mathsf{F} : \mathcal{U}_M \to \mathcal{O}_M$ of M, and $\mathsf{G} : \mathcal{U}_N \to \mathcal{O}_N$ of N. Then $\mathsf{F} \times \mathsf{G} : \mathcal{U}_M \times \mathcal{U}_N \to \mathcal{O}_M \times \mathcal{O}_N$ is a local parametrization of $M \times N$. To show that π_M is smooth, we compute:

$$\begin{aligned}\mathsf{F}^{-1} \circ \pi_M \circ (\mathsf{F} \times \mathsf{G})(\mathsf{u}, \mathsf{v}) &= \mathsf{F}^{-1} \circ \pi_M (\mathsf{F}(\mathsf{u}), \mathsf{G}(\mathsf{v})) \\ &= \mathsf{F}^{-1}(\mathsf{F}(\mathsf{u})) \\ &= \mathsf{u}\end{aligned}$$

The map $(\mathsf{u}, \mathsf{v}) \mapsto \mathsf{u}$ is clearly a smooth map between Euclidean spaces. Therefore, π_M is a smooth map between $M \times N$ and M. Similarly, π_N is also a smooth map between $M \times N$ and N. □

Exercise 2.11. Suppose $\Phi : \mathbb{R}^{n+1} \backslash \{0\} \to \mathbb{R}^{m+1} \backslash \{0\}$ is a smooth map which satisfies

$$\Phi(cx_0, cx_1, \ldots, cx_n) = c^d \Phi(x_0, x_1, \ldots, x_n)$$

for any $c \in \mathbb{R}\backslash\{0\}$ and $(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1}\backslash\{0\}$. Show that the induced map $\widetilde{\Phi} : \mathbb{RP}^n \to \mathbb{RP}^m$ defined by:

$$\widetilde{\Phi}([x_0 : x_1 : \cdots : x_n]) = \Phi(x_0, x_1, \ldots, x_n)$$

is well-defined and smooth. [Hint: To check $\widetilde{\Phi}$ is well-defined means to verify that two equivalent inputs $[x_0 : x_1 : \cdots : x_n] = [y_0 : y_1 : \cdots : y_n]$ will give the same outputs $\Phi(x_0, x_1, \ldots, x_n)$ and $\Phi(y_0, y_1, \ldots, y_n)$.]

Exercise 2.12. Let $M = \{(w,z) \in \mathbb{C}^2 : |w|^2 + |z|^2 = 1\}$.

(a) Show that M is a 3-dimensional manifold.

(b) Define

$$\Phi(w,z) := \left(z\bar{w} + w\bar{z}, i(w\bar{z} - z\bar{w}), |z|^2 - |w|^2\right)$$

for any $(w,z) \in M$. Show that $\Phi(w,z) \in \mathbb{R}^3$ and it lies on the unit sphere $\mathbb{S}^2$, and then verify that $\Phi : M \to \mathbb{S}^2$ is a smooth map.

2.2.2. Diffeomorphisms. Two smooth manifolds M and N are said to be *diffeomorphic* if they are in one-to-one correspondence with each other in smooth sense. Here is the rigorous definition:

Definition 2.21 (Diffeomorphisms)**.** A smooth map $\Phi : M \to N$ between two smooth manifolds M and N is said to be a *diffeomorphism* if Φ is a one-to-one and onto (i.e. bijective), and that the inverse map $\Phi^{-1} : N \to M$ is also smooth.

If such a map exists between M and N, the two manifolds M and N are said to be *diffeomorphic*.

Example 2.22. Given a smooth function $f : \mathcal{U} \to \mathbb{R}$ from an open subset $\mathcal{U} \subset \mathbb{R}^n$. The graph Γ_f defined as:

$$\Gamma_f := \{(\mathsf{x}, f(\mathsf{x})) \in \mathbb{R}^{n+1} : \mathsf{x} \in \mathcal{U}\}$$

is a smooth manifold by Example 2.10. We claim that the projection map:

$$\begin{aligned} \pi : \Gamma_f &\to \mathcal{U} \\ (\mathsf{x}, f(\mathsf{x})) &\mapsto \mathsf{x} \end{aligned}$$

is a diffeomorphism. Both Γ_f and $\mathcal{U}$ are covered by one global parametrization. The parametrization of $\mathcal{U}$ is simply the identity map $\mathrm{id}_{\mathcal{U}}$ on $\mathcal{U}$. The parametrization of Γ_f is given by:

$$\begin{aligned} \mathsf{F} : \mathcal{U} &\to \Gamma_f \\ \mathsf{x} &\mapsto (\mathsf{x}, f(\mathsf{x})) \end{aligned}$$

To show that π is smooth, we consider $\mathrm{id}_{\mathcal{U}}^{-1} \circ \pi \circ \mathsf{F}$, which is given by:

$$\begin{aligned} \mathrm{id}_{\mathcal{U}}^{-1} \circ \pi \circ \mathsf{F}(\mathsf{x}) &= \mathrm{id}_{\mathcal{U}}^{-1} \circ \pi(\mathsf{x}, f(\mathsf{x})) \\ &= \mathrm{id}_{\mathcal{U}}^{-1}(\mathsf{x}) \\ &= \mathsf{x}. \end{aligned}$$

Therefore, the composite $\mathrm{id}_{\mathcal{U}}^{-1} \circ \pi \circ \mathsf{F}$ is simply the identity map on $\mathcal{U}$, which is clearly smooth.

π is one-to-one and onto with inverse map π^{-1} given by:

$$\begin{aligned} \pi^{-1} : \mathcal{U} &\to \Gamma_f \\ \mathsf{x} &\mapsto (\mathsf{x}, f(\mathsf{x})) \end{aligned}$$

To show π^{-1} is smooth, we consider the composite $\mathsf{F}^{-1} \circ \pi^{-1} \circ \mathrm{id}_{\mathcal{U}}$:

$$\begin{aligned} \mathsf{F}^{-1} \circ \pi^{-1} \circ \mathrm{id}_{\mathcal{U}}(\mathsf{x}) &= \mathsf{F}^{-1} \circ \pi^{-1}(\mathsf{x}) \\ &= \mathsf{F}^{-1}(\mathsf{x}, f(\mathsf{x})) \\ &= \mathsf{x}. \end{aligned}$$

Therefore, the composite $\mathsf{F}^{-1} \circ \pi^{-1} \circ \mathrm{id}_{\mathcal{U}}$ is also the identity map on $\mathcal{U}$, which is again smooth. □

Example 2.23. Let M be the cylinder $x^2 + y^2 = 1$ in $\mathbb{R}^3$. We are going to show that M is diffeomorphic to $\mathbb{R}^2 \backslash \{(0,0)\}$ via the diffeomorphism:

$$\begin{aligned} \Phi : M &\to \mathbb{R}^2 \backslash \{(0,0)\} \\ (x, y, z) &\mapsto e^z (x, y) \end{aligned}$$

We leave it for readers to verify that Φ is one-to-one and onto, and hence Φ^{-1} exists. To show it is a diffeomorphism, we first parametrize M by two local coordinate charts:

$$\begin{aligned} \mathsf{F}_1 : (0, 2\pi) \times \mathbb{R} &\to M & \mathsf{F}_2 : (-\pi, \pi) \times \mathbb{R} &\to M \\ \mathsf{F}_1(\theta, z) &= (\cos\theta, \sin\theta, z) & \mathsf{F}_2(\widetilde{\theta}, \widetilde{z}) &= (\cos\widetilde{\theta}, \sin\widetilde{\theta}, \widetilde{z}) \end{aligned}$$

The target space $\mathbb{R}^2 \backslash \{(0,0)\}$ is an open set of $\mathbb{R}^2$, and hence can be globally parametrized by $\mathrm{id} : \mathbb{R}^2 \backslash \{(0,0)\} \to \mathbb{R}^2 \backslash \{(0,0)\}$.

We need to show $\Phi \circ \mathsf{F}_i$ and $\mathsf{F}_i^{-1} \circ \Phi^{-1}$ are smooth for any $i = 1, 2$. As an example, we verify one of them only:

$$\begin{aligned} \Phi \circ \mathsf{F}_1(\theta, z) &= \Phi(\cos\theta, \sin\theta, z) \\ &= (e^z \cos\theta, e^z \sin\theta). \end{aligned}$$

To show $\mathsf{F}_1^{-1} \circ \Phi^{-1} = (\Phi \circ \mathsf{F}_1)^{-1}$ is smooth, we use Inverse Function Theorem. The Jacobian of $\Phi \circ \mathsf{F}_1$ is given by:

$$D(\Phi \circ \mathsf{F}_1) = \det \begin{bmatrix} -e^z \sin\theta & e^z \cos\theta \\ e^z \cos\theta & e^z \sin\theta \end{bmatrix} = -e^{2z} \neq 0.$$

Therefore, $\Phi \circ \mathsf{F}_1$ has a C^∞ local inverse around every point in the domain. Since $\Phi \circ \mathsf{F}_1$ is one-to-one and onto, such a local inverse is a global inverse.

Similarly, one can show $\Phi \circ \mathsf{F}_2$ and $\mathsf{F}_2^{-1} \circ \Phi^{-1}$ are smooth. All these show Φ and Φ^{-1} are smooth maps between M and $\mathbb{R}^2 \backslash \{(0,0)\}$, and hence are diffeomorphisms. □

Exercise 2.13. Show that the open square $(-1,1) \times (-1,1) \subset \mathbb{R}^2$ is diffeomorphic to $\mathbb{R}^2$. [Hint: consider the trig functions tan or $\tan^{-1}$.]

Exercise 2.14. Consider the map $\Phi : B_1(0) \to \mathbb{R}^n$ defined by:

$$\Phi(\mathsf{x}) = \frac{|\mathsf{x}|}{\sqrt{1 - |\mathsf{x}|^2}}$$

where $B_1(0)$ is the open unit ball $\{\mathsf{x} \in \mathbb{R}^n : |\mathsf{x}| < 1\}$. Show that Φ is a diffeomorphism.

Exercise 2.15. Let $M = \mathbb{R}^2/\sim$ where $\sim$ is the equivalence relation:

$$\mathsf{x} \sim \mathsf{y} \quad \text{if and only if} \quad \mathsf{x} - \mathsf{y} \in \mathbb{Z}^2.$$

From Exercise 2.7, we have already showed that M is a smooth manifold. Show that M is diffeomorphic to a $\mathbb{S}^1 \times \mathbb{S}^1$.

2.3. Tangent Spaces and Tangent Maps

At a point p on a regular surface $M \subset \mathbb{R}^3$, the tangent plane T_pM at p is spanned by the basis $\left\{\frac{\partial \mathsf{F}}{\partial u_i}\right\}_{i=1,2}$ where F is a local parametrization around p. The basis $\frac{\partial \mathsf{F}}{\partial u_i}$ are vectors in $\mathbb{R}^3$ since F has an image in $\mathbb{R}^3$. However, this definition of tangent plane can hardly be generalized to abstract manifolds, as an abstract manifold M may not sit inside any Euclidean space. Instead, we define the tangent space at a point p on a smooth manifold as the vector space spanned by partial differential operators $\left\{\frac{\partial}{\partial u_i}\right\}_{i=1}^n$. Heuristically, we generalize the concept of tangent planes of regular surfaces in $\mathbb{R}^3$ by "removing" the label F from the *geometric* vector $\frac{\partial \mathsf{F}}{\partial u_i}$, so that it becomes an *abstract* vector $\frac{\partial}{\partial u_i}$. For this generalization, we first need to define partial derivatives on abstract manifolds.

2.3.1. Partial Derivatives and Tangent Vectors. Let M^n be a smooth manifold and $\mathsf{F} : \mathcal{U} \subset \mathbb{R}^n \to \mathcal{O} \subset M^n$ be a smooth local parametrization. Then similar to regular surfaces, for any $p \in \mathcal{O}$, it makes sense to define partial derivative for a function $f : M \to \mathbb{R}$ at p by pre-composing f with F, i.e. $f \circ \mathsf{F}$, which is a map from $\mathcal{U} \subset \mathbb{R}^n$ to $\mathbb{R}$. Let $(u_1, \ldots, u_n)$ be the coordinates of $\mathcal{U} \subset \mathbb{R}^n$, then with a little abuse of notations, we denote:

$$\frac{\partial f}{\partial u_j}(p) := \frac{\partial (f \circ \mathsf{F})}{\partial u_j}(\mathsf{u})$$

where u is the point in $\mathcal{U}$ corresponding to p, i.e. $\mathsf{F}(\mathsf{u}) = p$.

Remark 2.24. Note that $\frac{\partial f}{\partial u_j}(p)$ is defined locally on $\mathcal{O}$, and depends on the choice of local parametrization F near p. □

The partial derivative $\frac{\partial}{\partial u_j}(p)$ can be thought as an operator:

$$\frac{\partial}{\partial u_j}(p) : C^1(M, \mathbb{R}) \to \mathbb{R}$$

$$f \mapsto \frac{\partial f}{\partial u_j}(p).$$

Here $C^1(M, \mathbb{R})$ denotes the set of all C^1 functions from M to $\mathbb{R}$.

On regular surfaces $\frac{\partial \mathsf{F}}{\partial u_j}(p)$ is a tangent vector at p. On an abstract manifold, $\frac{\partial \mathsf{F}}{\partial u_j}(p)$ cannot be defined since F may not be in an Euclidean space. Instead, the partial differential operator $\frac{\partial}{\partial u_j}(p)$ plays the role of $\frac{\partial \mathsf{F}}{\partial u_j}(p)$, and we will call the operator $\frac{\partial}{\partial u_j}(p)$ a *tangent vector* for an abstract manifold. It does sound less concrete and more abstract than a geometric tangent vector, but one good quote by John von Neumann is:

> "In mathematics you don't understand things. You just get used to them."

Example 2.25. Let $\mathsf{F}(x, y, z) = (x, y, z)$ be the identity parametrization of $\mathbb{R}^3$, and $\mathsf{G}(\rho, \theta, \varphi) = (\rho \sin\varphi \cos\theta, \rho \sin\varphi \sin\theta, \rho \cos\varphi)$ be local parametrization of $\mathbb{R}^3$ by spherical coordinates.

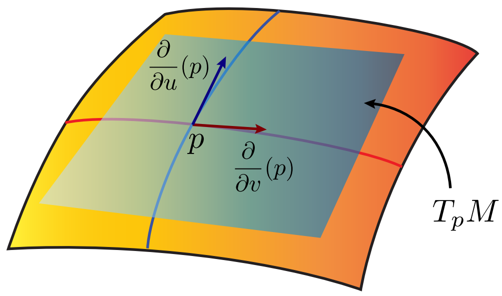


Figure 2.5. $\frac{\partial}{\partial u}$ and $\frac{\partial}{\partial v}$ are used in place of $\frac{\partial \mathsf{F}}{\partial u}$ and $\frac{\partial \mathsf{F}}{\partial v}$ on abstract manifolds.

Then at any point $p \in \mathbb{R}^3$, the vectors $\dfrac{\partial}{\partial x}(p), \dfrac{\partial}{\partial y}(p), \dfrac{\partial}{\partial z}(p)$ are regarded as the *abstract form* of the *geometric* vectors $\dfrac{\partial \mathsf{F}}{\partial x}(p), \dfrac{\partial \mathsf{F}}{\partial y}(p), \dfrac{\partial \mathsf{F}}{\partial z}(p)$, which are respectively i, j and k in standard notations.

Also, the vectors $\dfrac{\partial}{\partial \rho}(p), \dfrac{\partial}{\partial \theta}(p), \dfrac{\partial}{\partial \varphi}(p)$ are regarded as the *abstract form* of the *geometric* vectors $\dfrac{\partial \mathsf{G}}{\partial \rho}(p), \dfrac{\partial \mathsf{G}}{\partial \theta}(p), \dfrac{\partial \mathsf{G}}{\partial \varphi}(p)$, which are respectively the vectors at p tangent to the ρ-, θ- and φ-directions on the sphere. □

Example 2.26. Take another example:

$$\mathbb{RP}^2 = \{[x_0 : x_1 : x_2] : \text{at least one } x_i \neq 0\}.$$

According to Example 2.14, one of its local parametrizations is given by:

$$\begin{aligned} \mathsf{F} : \mathbb{R}^2 &\to \mathbb{RP}^2 \\ (x_1, x_2) &\mapsto [1 : x_1 : x_2] \end{aligned}$$

Such a manifold is not assumed to be in $\mathbb{R}^N$, so we can't define $\dfrac{\partial \mathsf{F}}{\partial x_1}, \dfrac{\partial \mathsf{F}}{\partial x_2}$ as *geometric vectors* in $\mathbb{R}^N$. However, as a substitute, we will regard the operators $\dfrac{\partial}{\partial x_1}, \dfrac{\partial}{\partial x_2}$ as *abstract* tangent vectors along the directions of x_1 and x_2 respectively. □

2.3.2. Tangent Spaces. Having generalized the concept of partial derivatives to abstract manifolds, we now ready to state the definition of tangent vectors for abstract manifolds.

Definition 2.27 (Tangent Spaces)**.** Let M be a smooth n-manifold, $p \in M$ and $\mathsf{F} : \mathcal{U} \subset \mathbb{R}^n \to \mathcal{O} \subset M$ be a smooth local parametrization around p. The *tangent space at p of M*, denoted by T_pM, is defined as:

$$T_pM = \operatorname{span}\left\{\frac{\partial}{\partial u_1}(p), \ldots, \frac{\partial}{\partial u_n}(p)\right\},$$

where $\dfrac{\partial}{\partial u_i}$'s are partial differential operators with respect to the local parametrization $\mathsf{F}(u_1, \ldots, u_n)$.

It *seems* that the definition of T_pM depends on the choice of local parametrization F. However, we can show that it does not. We first show that $\left\{\frac{\partial}{\partial u_i}(p)\right\}_{i=1}^n$ are linearly independent, then we have $\dim T_pM = n = \dim M$.

Given a local parametrization $\mathsf{F} : \mathcal{U} \to \mathcal{O} \subset M$ with local coordinates denoted by $(u_1, \ldots, u_n)$, then each coordinate u_i can be regarded as a locally defined function $u_i : \mathcal{O} \to \mathbb{R}$. Then we have:

$$\frac{\partial u_k}{\partial u_j}(p) = \delta_{kj} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Next we want to show $\left\{\frac{\partial}{\partial u_i}\right\}_{i=1}^n$ are linearly independent. Suppose a_i's are real numbers such that

$$\sum_{i=1}^n a_i \frac{\partial}{\partial u_i} = 0,$$

meaning that $\sum_{i=1}^n a_i \frac{\partial f}{\partial u_i} = 0$ for any differential function f (including the coordinate functions u_k's). Therefore, we have:

$$0 = \sum_{i=1}^n a_i \frac{\partial u_k}{\partial u_i} = \sum_{i=1}^n a_i \delta_{ki} = a_k$$

for any k. This shows $\left\{\frac{\partial}{\partial u_i}\right\}_{i=1}^n$ are linearly independent, show that $\dim T_pM = \dim M$.

Now we show T_pM does not depend on the choice of local coordinates. Suppose $\mathsf{F} : \mathcal{U} \subset \mathbb{R}^n \to \mathcal{O} \subset M$ and $\widetilde{\mathsf{F}} : \widetilde{\mathcal{U}} \subset \mathbb{R}^n \to \widetilde{\mathcal{O}}$ be two local parametrizations. We use $(u_1, \ldots, u_n)$ to denote the Euclidean coordinates on $\mathcal{U}$, and use $(v_1, \ldots, v_n)$ for $\widetilde{\mathcal{U}}$.

The partial derivatives $\frac{\partial f}{\partial u_j}$ are $\frac{\partial f}{\partial v_i}$ are different. Via the transition map $\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F}$, $(v_1, \ldots, v_n)$ can be regarded as functions of $(u_1, \ldots, u_n)$, and therefore it makes sense of defining $\frac{\partial v_j}{\partial u_i}$.

Given a smooth function $f : M \to \mathbb{R}$, by the chain rule, one can write the partial derivative $\frac{\partial f}{\partial u_j}$ in terms of $\frac{\partial f}{\partial v_i}$ as follows:

$$\begin{aligned}
\frac{\partial f}{\partial u_i}(p) &:= \left.\frac{\partial (f \circ \mathsf{F})}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)} \\
&= \left.\frac{\partial}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)} (f \circ \widetilde{\mathsf{F}}) \circ (\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F}) \\
&= \sum_{j=1}^n \left.\frac{\partial (f \circ \widetilde{\mathsf{F}})}{\partial v_j}\right|_{\widetilde{\mathsf{F}}^{-1}(\mathsf{p})} \left.\frac{\partial v_j}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)} \\
&= \sum_{j=1}^n \left.\frac{\partial v_j}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)} \frac{\partial f}{\partial v_j}(p)
\end{aligned} \tag{2.1}$$

In short, we can write:

$$\frac{\partial}{\partial u_i} = \sum_{j=1}^{n} \frac{\partial v_j}{\partial u_i} \frac{\partial}{\partial v_j}.$$

In other words, $\frac{\partial}{\partial u_i}$ can be expressed as a linear combination of $\frac{\partial}{\partial v_j}'$s.

Therefore, $\operatorname{span}\left\{\frac{\partial}{\partial u_i}(p)\right\}_{i=1}^{n} \subset \operatorname{span}\left\{\frac{\partial}{\partial v_i}(p)\right\}_{i=1}^{n}$. Since both spans of vectors have equal dimension, their span must be equal. This shows T_pM is independent of choice of local parametrizations. However, it is important to note that each individual basis vector $\frac{\partial}{\partial u_i}(p)$ *does* depend on local parametrizations.

Example 2.28. Consider again the real projective plane:

$$\mathbb{RP}^2 = \{[x_0 : x_1 : x_2] : \text{at least one } x_i \neq 0\}.$$

Consider the two local parametrizations:

$$\begin{aligned} &\mathsf{F} : \mathbb{R}^2 \to \mathbb{RP}^2 & &\mathsf{G} : \mathbb{R}^2 \to \mathbb{RP}^2 \\ &\mathsf{F}(x_1, x_2) = [1 : x_1 : x_2] & &\mathsf{G}(y_0, y_2) = [y_0 : 1 : y_2] \end{aligned}$$

Then, (y_0, y_2) can be regarded as a function of (x_1, x_2) via the transition map $\mathsf{G}^{-1} \circ \mathsf{F}$, which is explicitly given by:

$$\begin{aligned} (y_0, y_2) &= \mathsf{G}^{-1} \circ \mathsf{F}(x_1, x_2) = \mathsf{G}^{-1}([1 : x_1 : x_2]) \\ &= \mathsf{G}^{-1}([x_1^{-1} : 1 : x_1^{-1}x_2]) = \left(\frac{1}{x_1}, \frac{x_2}{x_1}\right). \end{aligned}$$

Using the chain rule, we can then express $\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}$ in terms of $\frac{\partial}{\partial y_0}, \frac{\partial}{\partial y_2}$:

$$\begin{aligned} \frac{\partial}{\partial x_1} &= \frac{\partial y_0}{\partial x_1}\frac{\partial}{\partial y_0} + \frac{\partial y_2}{\partial x_1}\frac{\partial}{\partial y_2} \\ &= -\frac{1}{x_1^2}\frac{\partial}{\partial y_0} - \frac{x_2}{x_1^2}\frac{\partial}{\partial y_2} \\ &= -y_0^2\frac{\partial}{\partial y_0} - y_0 y_2\frac{\partial}{\partial y_2}. \end{aligned}$$

We leave $\frac{\partial}{\partial x_2}$ as an exercise. □

Exercise 2.16. Express $\frac{\partial}{\partial x_2}$ as a linear combination $\frac{\partial}{\partial y_0}, \frac{\partial}{\partial y_2}$ in Example 2.28. Leave the final answer in terms of y_0 and y_2 only.

Exercise 2.17. Consider the extended complex plane $M := \mathbb{C} \cup \{\infty\}$ (discussed in Example 2.12) with local parametrizations:

$$\begin{aligned} &\mathsf{F}_1 : \mathbb{R}^2 \to \mathbb{C} \subset M & &\mathsf{F}_2 : \mathbb{R}^2 \to (\mathbb{C}\backslash\{0\}) \cup \{\infty\} \subset M \\ &(x_1, x_2) \mapsto x_1 + x_2 i & &(y_1, y_2) \mapsto \frac{1}{y_1 + y_2 i} \end{aligned}$$

Express the tangent space basis $\left\{\frac{\partial}{\partial x_i}\right\}$ in terms of the basis $\left\{\frac{\partial}{\partial y_j}\right\}$.

Exercise 2.18. Given two smooth manifolds M^m and N^n, and a point $(p,q) \in M \times N$, show that the tangent plane $T_{(p,q)}(M \times N)$ is isomorphic to $T_pM \oplus T_qN$. Recall that $V \oplus W$ is the *direct sum* of two vector spaces V and W, defined as:

$$V \oplus W = \{(v,w) : v \in V \text{ and } w \in W\}.$$

2.3.3. Tangent Maps. Given a smooth map Φ between two regular surfaces in $\mathbb{R}^3$, we discussed in Section 1.5.2 on how to define its partial derivatives using local parametrizations. To recap, suppose $\Phi : M \to N$ and $\mathsf{F}(u_1,u_2) : \mathcal{U}_M \to \mathcal{O}_M \subset M$ and $\mathsf{G}(v_1,v_2) : \mathcal{U}_N \to \mathcal{O}_N \subset N$ are local parametrizations of M and N respectively. Via Φ, the local coordinates (v_1,v_2) of N can be regarded as functions of (u_1,u_2), i.e.

$$(v_1,v_2) = \mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1,u_2).$$

Then, according to (1.1), the partial derivative $\dfrac{\partial \Phi}{\partial u_i}$ of the map Φ is given by:

$$\begin{aligned}\frac{\partial \Phi}{\partial u_i}(\Phi(p)) := \left.\frac{\partial(\Phi \circ \mathsf{F})}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)} &= \left.\frac{\partial(\mathsf{G} \circ (\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}))}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)} \\ &= \sum_{j=1}^{2} \left.\frac{\partial v_j}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)} \frac{\partial \mathsf{G}}{\partial v_j}(\Phi(p))\end{aligned}$$

which is a vector on the tangent plane $T_{\Phi(p)}N$.

Now, if we are given a smooth map $\Phi : M^m \to N^n$ between two smooth abstract manifolds M^m and N^n with local parametrizations $\mathsf{F}(u_1,\ldots,u_m) : \mathcal{U}_M \subset \mathbb{R}^m \to \mathcal{O}_M \subset M^m$ around $p \in M$, and $\mathsf{G}(v_1,\ldots,v_n) : \mathcal{U}_N \subset \mathbb{R}^n \to \mathcal{O}_N \subset N^n$ around $\Phi(p) \in N$, then the tangent space $T_{\Phi(p)}N$ is spanned by $\left\{\dfrac{\partial}{\partial v_j}(\Phi(p))\right\}_{j=1}^{n}$. In view of (1.1), a natural generalization of partial derivatives $\dfrac{\partial \Phi}{\partial u_i}(p)$ to smooth maps between manifolds is:

Definition 2.29 (Partial Derivatives of Maps between Manifolds)**.** Let $\Phi : M^m \to N^n$ be a smooth map between two smooth manifolds M and N. Let $\mathsf{F}(u_1,\ldots,u_m) : \mathcal{U}_M \to \mathcal{O}_M \subset M$ be a smooth local parametrization around p and $\mathsf{G}(v_1,\ldots,v_n) : \mathcal{U}_N \to \mathcal{O}_N \subset N$ be a smooth local parametrization around $\Phi(p)$. Then, the *partial derivative of Φ with respect to u_i at p* is defined to be:

$$\frac{\partial \Phi}{\partial u_i}(p) := \sum_{j=1}^{n} \left.\frac{\partial v_j}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)} \frac{\partial}{\partial v_j}(\Phi(p)). \tag{2.2}$$

Here $(v_1,\ldots,v_n)$ are regarded as functions of $(u_1,\ldots,u_m)$ in a sense that:

$$(v_1,\ldots,v_n) = \mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1,\ldots,u_m).$$

Note that the partial derivative $\dfrac{\partial \Phi}{\partial u_i}$ defined in (2.2) depends on the local parametrization F. However, one can show that it does not depend on the choice of the local parametrization G in the target space.

Suppose $\widetilde{\mathsf{G}}(w_1, \ldots, w_n)$ is another local parametrization around $\Phi(p)$. Then by the chain rule:

$$\begin{aligned}\sum_{j=1}^{n} \frac{\partial w_j}{\partial u_i} \frac{\partial}{\partial w_j} &= \sum_{j=1}^{n} \frac{\partial w_j}{\partial u_i} \left(\sum_{k=1}^{n} \frac{\partial v_k}{\partial w_j} \frac{\partial}{\partial v_k} \right) \\ &= \sum_{j,k=1}^{n} \frac{\partial w_j}{\partial u_i} \frac{\partial v_k}{\partial w_j} \frac{\partial}{\partial v_k} \\ &= \sum_{k=1}^{n} \frac{\partial v_k}{\partial u_i} \frac{\partial}{\partial v_k}\end{aligned}$$

Therefore, the way to define $\dfrac{\partial \Phi}{\partial u_i}$ in (2.2) is independent of choice of local parametrization G for the target manifold N.

Example 2.30. Consider the map $\Phi : \mathbb{RP}^1 \times \mathbb{RP}^2 \to \mathbb{RP}^5$ defined by:

$$\Phi([x_0 : x_1], [y_0 : y_1 : y_2]) = [x_0 y_0 : x_0 y_1 : x_0 y_2 : x_1 y_0 : x_1 y_1 : x_1 y_2].$$

Under the standard local parametrizations $\mathsf{F}(u) = [1 : u]$ for $\mathbb{RP}^1$, $\mathsf{G}(v_1, v_2) = [1 : v_1 : v_2]$ for $\mathbb{RP}^2$, and $\mathsf{H}(w_1, \ldots, w_5) = [1 : w_1 : \cdots : w_5]$ for $\mathbb{RP}^5$, the local expression of Φ is given by:

$$\begin{aligned}&\mathsf{H}^{-1} \circ \Phi \circ (\mathsf{F} \times \mathsf{G})(u, v_1, v_2) \\ &= \mathsf{H}^{-1} \circ \Phi([1 : u], [1 : v_1 : v_2]) \\ &= \mathsf{H}^{-1} \left([1 : v_1 : v_2 : u : u v_1 : u v_2] \right) \\ &= (v_1, v_2, u, u v_1, u v_2).\end{aligned}$$

Via the map Φ, we can regard $(w_1, w_2, w_3, w_4, w_5) = (v_1, v_2, u, u v_1, u v_2)$, and the partial derivatives of Φ are given by:

$$\begin{aligned}\frac{\partial \Phi}{\partial u} &= \frac{\partial w_1}{\partial u} \frac{\partial}{\partial w_1} + \ldots + \frac{\partial w_5}{\partial u} \frac{\partial}{\partial w_5} = \frac{\partial}{\partial w_3} + v_1 \frac{\partial}{\partial w_4} + v_2 \frac{\partial}{\partial w_5} \\ \frac{\partial \Phi}{\partial v_1} &= \frac{\partial w_1}{\partial v_1} \frac{\partial}{\partial w_1} + \ldots + \frac{\partial w_5}{\partial v_1} \frac{\partial}{\partial w_5} = \frac{\partial}{\partial w_1} + u \frac{\partial}{\partial w_4} \\ \frac{\partial \Phi}{\partial v_2} &= \frac{\partial w_1}{\partial v_2} \frac{\partial}{\partial w_1} + \ldots + \frac{\partial w_5}{\partial v_2} \frac{\partial}{\partial w_5} = \frac{\partial}{\partial w_2} + u \frac{\partial}{\partial w_5}\end{aligned}$$

□

Similar to tangent maps between regular surfaces in $\mathbb{R}^3$, we define:

$$(\Phi_*)_p \left(\frac{\partial}{\partial u_i}(p) \right) := \frac{\partial \Phi}{\partial u_i}(p)$$

and extend the map linearly to all vectors in $T_p M$. This is then a linear map between $T_p M$ and $T_{\Phi(p)} N$, and we call this map the *tangent map*.

Definition 2.31 (Tangent Maps)**.** Under the same assumption stated in Definition 2.29, the *tangent map* of Φ at $p \in M$ denoted by $(\Phi_*)_p$ is defined as:

$$(\Phi_*)_p : T_p M \to T_{\Phi(p)} N$$

$$(\Phi_*)_p \left(\sum_{i=1}^{n} a_i \frac{\partial}{\partial u_i}(p) \right) = \sum_{i=1}^{n} a_i \frac{\partial \Phi}{\partial u_i}(p)$$

If the point p is clear from the context, $(\Phi_*)_p$ can be simply denoted by Φ_*.

For brevity, we will from now on say "$(u_1, \ldots, u_m)$ are local coordinates of M around p" instead of saying in a clumsy way that "$\mathsf{F} : \mathcal{U} \to M$ is a local parametrization of M around p and that $(u_1, \ldots, u_m)$ are coordinates on $\mathcal{U}$".

Given a local coordinates $(u_1, \ldots, u_m)$ around p, and local coordinates $(v_1, \ldots, v_n)$ around $\Phi(p)$, then from (2.2), the matrix representation of $(\Phi_*)_p$ with respect to bases $\left\{ \frac{\partial}{\partial u_i}(p) \right\}_{i=1}^m$ and $\left\{ \frac{\partial}{\partial v_j}(\Phi(p)) \right\}_{j=1}^n$ is given by $\left[\frac{\partial v_j}{\partial u_i} \right]_{i=1,\ldots,m}^{j=1,\ldots,n}$ where i stands for the column, and j stands for the row. The matrix is nothing but the Jacobian matrix:

$$(2.3) \qquad [(\Phi_*)_p] = \left[\frac{\partial(v_1, \ldots, v_n)}{\partial(u_1, \ldots, u_m)} \right]_{\mathsf{F}^{-1}(p)} = \left[D(\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}) \right]_{\mathsf{F}^{-1}(p)}.$$

Example 2.32. Consider again the map $\Phi : \mathbb{RP}^1 \times \mathbb{RP}^2 \to \mathbb{RP}^5$ in Example 2.30. Under the local parametrizations considered in that example, we then have (for instance):

$$\Phi_* \left(\frac{\partial}{\partial u} \right) = \frac{\partial \Phi}{\partial u} = \frac{\partial}{\partial w_3} + v_1 \frac{\partial}{\partial w_4} + v_2 \frac{\partial}{\partial w_5}.$$

Using the results computed in Example 2.30, the matrix representation of Φ_* is given by:

$$[\Phi_*] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ v_1 & u & 0 \\ v_2 & 0 & u \end{bmatrix}$$

Hence, Φ_* is injective. Remark: To be rigorous, we have only shown $(\Phi_*)_p$ is injective at any p covered by the local coordinate charts we picked. The matrix $[\Phi_*]$ using other local coordinate charts can be computed in a similar way (left as an exercise). □

Exercise 2.19. Consider the map $\Phi : \mathbb{RP}^1 \times \mathbb{RP}^2 \to \mathbb{RP}^5$ defined as in Example 2.30. This time, we use the local parametrizations

$$\mathsf{F}(u) = [u : 1]$$
$$\mathsf{G}(v_0, v_2) = [v_0 : 1 : v_2]$$
$$\mathsf{H}(w_0, w_1, w_3, w_4, w_5) = [w_0 : w_1 : 1 : w_3 : w_4 : w_5]$$

for $\mathbb{RP}^1$, $\mathbb{RP}^2$ and $\mathbb{RP}^5$ respectively. Compute matrix representation of Φ_* using these local parametrizations.

Exercise 2.20. Note that in Definition 2.31 we defined Φ_* using local coordinates. Show that Φ_* is independent of local coordinates. Precisely, show that if:

$$\sum_i a_i \frac{\partial}{\partial u_i} = \sum_i b_i \frac{\partial}{\partial w_i}$$

where $\{u_i\}$ and $\{w_i\}$ are two local coordinates of M, then we have:

$$\sum_i a_i \frac{\partial \Phi}{\partial u_i} = \sum_i b_i \frac{\partial \Phi}{\partial w_i}, \quad \text{which implies} \quad \Phi_* \left(\sum_i a_i \frac{\partial}{\partial u_i} \right) = \Phi_* \left(\sum_i b_i \frac{\partial}{\partial w_i} \right).$$

Exercise 2.21. The identity map id_M of a smooth manifolds M takes any point $p \in M$ to itself, i.e. $\mathrm{id}_M(p) = p$. Show that its tangent map $(\mathrm{id}_M)_*$ at p is the identity map on the tangent space T_pM.

Exercise 2.22. Consider two smooth manifolds M^m and N^n, and their product $M^m \times N^n$. Find the tangent maps $(\pi_M)_*$ and $(\pi_N)_*$ of projection maps:

$$\begin{aligned} \pi_M : M \times N &\to M \\ (x, y) &\mapsto x \\ \pi_N : M \times N &\to N \\ (x, y) &\mapsto y \end{aligned}$$

2.4. Inverse Function Theorem

2.4.1. Chain Rule. Consider a smooth function $\Psi(v_1, \ldots, v_k) : \mathbb{R}^k \to \mathbb{R}^m$, another smooth function $\Phi(u_1, \ldots, u_n) : \mathbb{R}^n \to \mathbb{R}^k$ and the composition $\Psi \circ \Phi$. Under this composition, $(v_1, \ldots, v_k)$ can be regarded as a function of $(u_1, \ldots, u_n)$, and the output $(w_1, \ldots, w_m) = \Psi(v_1, \ldots, v_k)$ is ultimately a function of $(u_1, \ldots, u_n)$. In Multivariable Calculus, the chain rule is usually stated as:

$$\frac{\partial w_j}{\partial u_i} = \sum_l \frac{\partial w_j}{\partial v_l}\frac{\partial v_l}{\partial u_i}$$

or equivalently in an elegant way using Jacobian matrices:

$$\frac{\partial(w_1, \ldots, w_m)}{\partial(u_1, \ldots, u_n)} = \frac{\partial(w_1, \ldots, w_m)}{\partial(v_1, \ldots, v_k)}\frac{\partial(v_1, \ldots, v_k)}{\partial(u_1, \ldots, u_n)}.$$

Our goal here is to show that the chain rule can be generalized to maps between smooth manifolds, and can be rewritten using tangent maps:

Theorem 2.33 (Chain Rule: smooth manifolds). *Let $\Phi : M^m \to N^n$ and $\Psi : N^n \to P^k$ be two smooth maps between smooth manifolds M, N and P, then we have:*

$$(\Psi \circ \Phi)_* = \Psi_* \circ \Phi_*$$

Proof. Suppose $\mathsf{F}(u_1, \ldots, u_m)$ is a smooth local parametrization of M, $\mathsf{G}(v_1, \ldots, v_n)$ is a smooth local parametrization of N and $\mathsf{H}(w_1, \ldots, w_k)$ is a smooth parametrization of P. Locally, $(w_1, \ldots, w_k)$ are then functions of $(v_1, \ldots, v_n)$ via Ψ; and $(v_1, \ldots, v_n)$ are functions of $(u_1, \ldots, u_m)$ via Φ, i.e.

$$\begin{aligned}(w_1, \ldots, w_k) &= \mathsf{H}^{-1} \circ \Psi \circ \mathsf{G}(v_1, \ldots, v_n)\\ (v_1, \ldots, v_n) &= \mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1, \ldots, u_m)\end{aligned}$$

Ultimately, we can regard $(w_1, \ldots, w_k)$ as functions of $(u_1, \ldots, u_m)$ via the composition $\Psi \circ \Phi$:

$$\begin{aligned}(w_1, \ldots, w_k) &= (\mathsf{H}^{-1} \circ \Psi \circ \mathsf{G}) \circ (\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F})(u_1, \ldots, u_m)\\ &= \mathsf{H}^{-1} \circ (\Psi \circ \Phi) \circ \mathsf{F}(u_1, \ldots, u_m)\end{aligned}$$

To find the tangent map $(\Psi \circ \Phi)_*$, we need to figure out how it acts on the basis vectors $\dfrac{\partial}{\partial u_i}$, and recall that it is defined (see (2.2)) as follows:

$$(\Psi \circ \Phi)_*\left(\frac{\partial}{\partial u_i}\right) = \frac{\partial(\Psi \circ \Phi)}{\partial u_i} = \sum_{j=1}^{k} \frac{\partial w_j}{\partial u_i}\frac{\partial}{\partial w_j}.$$

Next, we use the (standard) chain rule for maps between Euclidean spaces:

$$\sum_{j=1}^{k} \frac{\partial w_j}{\partial u_i}\frac{\partial}{\partial w_j} = \sum_{j=1}^{k}\sum_{l=1}^{n} \frac{\partial w_j}{\partial v_l}\frac{\partial v_l}{\partial u_i}\frac{\partial}{\partial w_j}.$$

Therefore, we get:

$$(\Psi \circ \Phi)_*\left(\frac{\partial}{\partial u_i}\right) = \sum_{j=1}^{k}\sum_{l=1}^{n} \frac{\partial w_j}{\partial v_l}\frac{\partial v_l}{\partial u_i}\frac{\partial}{\partial w_j}.$$

Next, we verify that $\Psi_* \circ \Phi_*\left(\dfrac{\partial}{\partial u_i}\right)$ will give the same output:

$$\begin{aligned}
\Phi_*\left(\frac{\partial}{\partial u_i}\right) &= \frac{\partial \Phi}{\partial u_i} = \sum_{l=1}^{n} \frac{\partial v_l}{\partial u_i}\frac{\partial}{\partial v_l} \\
\Psi_* \circ \Phi_*\left(\frac{\partial}{\partial u_i}\right) &= \Psi_*\left(\sum_{l=1}^{n} \frac{\partial v_l}{\partial u_i}\frac{\partial}{\partial v_l}\right) = \sum_{l=1}^{n} \frac{\partial v_l}{\partial u_i}\Psi_*\left(\frac{\partial}{\partial v_l}\right) \\
&= \sum_{l=1}^{n} \frac{\partial v_l}{\partial u_i}\frac{\partial \Psi}{\partial v_l} = \sum_{l=1}^{n} \frac{\partial v_l}{\partial u_i}\left(\sum_{j=1}^{k} \frac{\partial w_j}{\partial v_l}\frac{\partial}{\partial w_j}\right) \\
&= \sum_{j=1}^{k}\sum_{l=1}^{n} \frac{\partial w_j}{\partial v_l}\frac{\partial v_l}{\partial u_i}\frac{\partial}{\partial w_j}.
\end{aligned}$$

Therefore, we have:

$$(\Psi \circ \Phi)_*\left(\frac{\partial}{\partial u_i}\right) = \Psi_* \circ \Phi_*\left(\frac{\partial}{\partial u_i}\right)$$

for any i, and hence $(\Psi \circ \Phi)_* = \Psi_* \circ \Phi_*$. □

Here is one immediate corollary of the chain rule:

Corollary 2.34. *If $\Phi : M \to N$ is a diffeomorphism between two smooth manifolds M and N, then at each point $p \in M$ the tangent map $\Phi_* : T_pM \to T_{\Phi(p)}N$ is invertible.*

Proof. Given that Φ is a diffeomorphism, the inverse map $\Phi^{-1} : N \to M$ exists. Since $\Phi^{-1} \circ \Phi = \mathrm{id}_M$, using the chain rule and Exercise 2.21, we get:

$$\mathrm{id}_{TM} = (\mathrm{id}_M)_* = (\Phi^{-1} \circ \Phi)_* = (\Phi^{-1})_* \circ \Phi_*.$$

Similarly, one can also show $\Phi_* \circ \Phi_*^{-1} = \mathrm{id}_{TN}$. Therefore, Φ_*, and $(\Phi^{-1})_*$ as well, are invertible. □

Exercise 2.23. Given two diffeomorphic smooth manifolds M and N, what can you say about $\dim M$ and $\dim N$?

Exercise 2.24. Let $\mathbb{S}^2 = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ be the unit sphere. Consider the maps $\pi : \mathbb{S}^2 \to \mathbb{RP}^2$ defined by

$$\pi(x,y,z) := [x : y : z]$$

and $\Phi : \mathbb{RP}^2 \to \mathbb{R}^4$ defined by:

$$\Phi([x : y : z]) = (x^2 - y^2, xy, xz, yz).$$

Locally parametrize $\mathbb{S}^2$ stereographically:

$$\mathsf{F}(u,v) = \left(\frac{2u}{u^2+v^2+1}, \frac{2v}{u^2+v^2+1}, \frac{u^2+v^2-1}{u^2+v^2+1}\right)$$

and $\mathbb{RP}^2$ by a standard parametrization:

$$\mathsf{G}(w_1, w_2) = [1 : w_1 : w_2].$$

Compute $[\Phi_*]$, $[\pi_*]$ and $[(\Phi \circ \pi)_*]$ directly, and verify that $[(\Phi \circ \pi)_*] = [\Phi_*][\pi_*]$.

2.4.2. Inverse Function Theorem. Given a diffeomorphism $\Phi : M \to N$, it is necessary that Φ_* is invertible. One natural question to ask is that if we are given Φ_* is invertible, can we conclude that Φ is a diffeomorphism?

Unfortunately, it is too good to be true. One easy counter-example is the map $\Phi : \mathbb{R} \to \mathbb{S}^1$, defined as:

$$\Phi(t) = (\cos t, \sin t).$$

As both $\mathbb{R}$ and $\mathbb{S}^1$ are one dimensional manifolds, to show that Φ_* is invertible it suffices to show that $\Phi_* \neq 0$, which can be verified by considering:

$$\Phi_*\left(\frac{\partial}{\partial t}\right) = \frac{\partial \Phi}{\partial t} = (-\sin t, \cos t) \neq 0.$$

However, it is clear that Φ is not even one-to-one, and hence Φ^{-1} does not exist.

Fortunately, the Inverse Function Theorem tells us that Φ is *locally* invertible near p whenever $(\Phi_*)_p$ is invertible. In Multivariable Calculus/Analysis, the Inverse Function Theorem asserts that if the Jacobian matrix of a smooth map $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ at p is invertible, then there exists an open set $\mathcal{U} \subset \mathbb{R}^n$ containing p, and an open set $\mathcal{V} \subset \mathbb{R}^n$ containing $\Phi(p)$ such that $\Phi|_{\mathcal{U}} : \mathcal{U} \to \mathcal{V}$ is a diffeomorphism.

Now suppose $\Phi : M \to N$ is a smooth map between two smooth manifolds M and N. According to (2.2), the matrix representation of the tangent map Φ_* is a Jacobian matrix. Therefore, one can generalize the Inverse Function Theorem to smooth manifolds. To start, we first define:

Definition 2.35 (Local Diffeomorphisms)**.** Let $\Phi : M \to N$ be a smooth map between two smooth manifolds M and N. We say Φ is a *local diffeomorphism* near p if there exists an open set $\mathcal{O}_M \subset M$ containing p, and an open set $\mathcal{O}_N \subset N$ containing $\Phi(p)$ such that $\Phi|_{\mathcal{O}_M} : \mathcal{O}_M \to \mathcal{O}_N$ is a diffeomorphism.

If such a smooth map exists, we say M is *locally diffeomorphic to N near p*, or equivalently, N *is locally diffeomorphic to M near* $\Phi(p)$. If M is locally diffeomorphic to N near every point $p \in M$, then we say M is *locally diffeomorphic* to N.

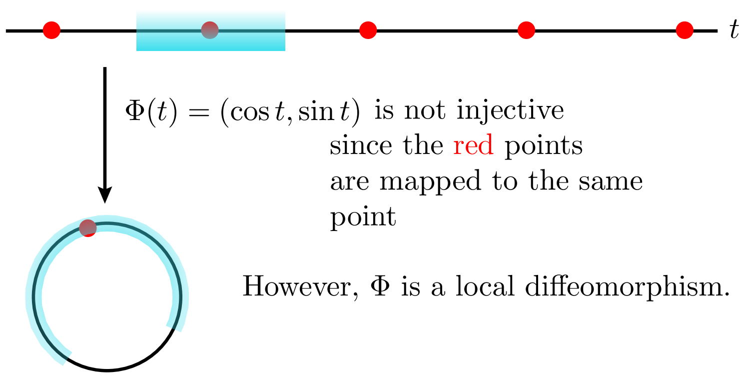


Figure 2.6. A local diffeomorphism which is not injective.

Theorem 2.36 (Inverse Function Theorem)**.** *Let $\Phi : M \to N$ be a smooth map between two smooth manifolds M and N. If $(\Phi_*)_p : T_pM \to T_{\Phi(p)}N$ is invertible, then M is locally diffeomorphic to N near p.*

Proof. The proof to be presented uses the Inverse Function Theorem for Euclidean spaces and then extends it to smooth manifolds. For the proof of the Euclidean case, readers may consult the lecture notes of MATH 3033/3043.

Let F be a local parametrization of M near p, and G be a local parametrization of N near $\Phi(p)$. Given that $(\Phi_*)_p$ is invertible, by (2.3) we know that the following Jacobian matrix $D(\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F})$ is invertible at $\mathsf{F}^{-1}(p)$. By Inverse Function Theorem for Euclidean spaces, there exist an open set $\mathcal{U}_M \subset \mathbb{R}^{\dim M}$ containing $\mathsf{F}^{-1}(p)$, and an open set $\mathcal{U}_N \subset \mathbb{R}^{\dim N}$ containing $\mathsf{G}^{-1}(\Phi(p))$ such that:

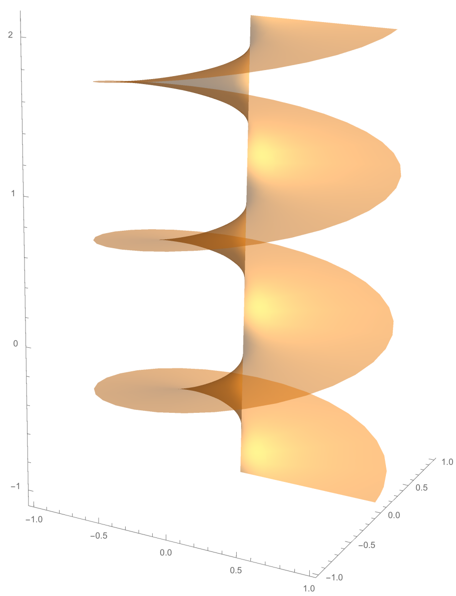
$$G^{-1} \circ \Phi \circ \mathsf{F}\Big|_{\mathcal{U}_M} : \mathcal{U}_M \to \mathcal{U}_N$$

is a diffeomorphism, i.e. the inverse $\mathsf{F}^{-1} \circ \Phi^{-1} \circ \mathsf{G}$ exists when restricted to $\mathcal{U}_N$ and is smooth.

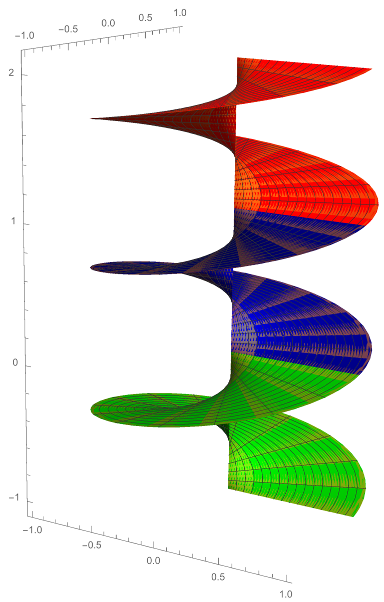
Denote $\mathcal{O}_M = \mathsf{F}(\mathcal{U}_M)$ and $\mathcal{O}_N = \mathsf{G}(\mathcal{U}_N)$. By the definition of smooth maps, this shows $\Phi|_{\mathcal{O}_M}$ and $\Phi^{-1}|_{\mathcal{O}_N}$ are smooth. Hence $\Phi|_{\mathcal{O}_M}$ is a local diffeomorphism near p. $\square$

Example 2.37. The helicoid Σ is defined to be the following surface in $\mathbb{R}^3$:

$$\Sigma := \{(r\cos\theta, r\sin\theta, \theta) \in \mathbb{R}^3 : r > 0 \text{ and } \theta \in \mathbb{R}\}.$$



(a) A helicoid



(b) Locally diffeomorphic to $\mathbb{R}^2\backslash\{0\}$

Figure 2.7. a helicoid is not globally diffeomorphic to $\mathbb{R}^2\backslash\{0\}$, but is locally diffeomorphic to $\mathbb{R}^2\backslash\{0\}$.

It can be parametrized by:

$$\mathsf{F} : (0, \infty) \times \mathbb{R} \to \Sigma$$
$$\mathsf{F}(r, \theta) = (r\cos\theta, r\sin\theta, \theta)$$

Consider the map $\Phi : \Sigma \to \mathbb{R}^2\backslash\{0\}$ defined as:

$$\Phi(r\cos\theta, r\sin\theta, \theta) = (r\cos\theta, r\sin\theta).$$

It is clear that Φ is not injective: for instance, $\Phi(\cos 2\pi, \sin 2\pi, 2\pi) = \Phi(\cos 0, \sin 0, 0)$. However, we can show that $(\Phi_*)_p$ is injective at each point $p \in \Sigma$.

The set $\mathbb{R}^2 \backslash \{0\}$ is open in $\mathbb{R}^2$. The matrix $[\Phi_*]$ is the Jacobian matrix of $\Phi \circ \mathsf{F}$:

$$\begin{aligned} \Phi \circ \mathsf{F}(r, \theta) &= \Phi(r\cos\theta, r\sin\theta, \theta) \\ &= (r\cos\theta, r\sin\theta) \end{aligned}$$

$$[\Phi_*] = D(\Phi \circ \mathsf{F}) = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

As $\det[\Phi_*] = r \neq 0$, the linear map $[\Phi_*]$ is invertible. By Inverse Function Theorem, Φ is a local diffeomorphism.

Exercise 2.25. Show that $\mathbb{S}^n$ and $\mathbb{RP}^n$ are locally diffeomorphic via the map:

$$\Phi(x_0, \ldots, x_n) = [x_0 : \cdots : x_n].$$

2.5. Immersions and Submersions

2.5.1. Review of Linear Algebra: injectivity and surjectivity. Given a linear map $T : V \to W$ between two finite dimensional vector spaces V and W, the following are equivalent:

(a) T is injective;

(b) $\ker T = \{0\}$;

(c) The row reduced echelon form (RREF) of the matrix of T has no *free column*.

In each RREF of a matrix, we call the first non-zero entry (if exists) of each row to be a pivot. A free column of an RREF is a column which does *not* have a pivot. For instance, the following RREF:

$$R = \begin{bmatrix} 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

has three pivots, and two free columns (namely the second and fourth columns). Any map with a matrix which can be row reduced to this R is not injective.

Surjectivity of a linear map $T : V \to W$ can also be stated in several equivalent ways:

(a) T is surjective;

(b) $\operatorname{rank}(T) = \dim W$;

(c) All rows in the RREF of the matrix of T are non-zero.

For instance, all rows of the matrix R above are non-zero. Hence any map with a matrix which can be row reduced to R is surjective.

Exercise 2.26. Let $T : V \to W$ be a linear map between two finite dimensional vector spaces V and W. Given that T is injective, what can you say about $\dim V$ and $\dim W$? Explain. Now given that T is surjective, what can you say about $\dim V$ and $\dim W$? Explain.

2.5.2. Immersions. Loosely speaking, an immersion from one smooth manifold to another is a map that is "locally injective". Here is the rigorous definition:

Definition 2.38 (Immersions)**.** Let $\Phi : M \to N$ be a smooth map between two smooth manifolds M and N. We say Φ is an *immersion at* $p \in M$ if the tangent map $(\Phi_*)_p : T_pM \to T_{\Phi(p)}N$ is injective. If Φ is an immersion at every point on M, then we simply say Φ is an *immersion*.

Remark 2.39. As a linear map $T : V \to W$ between any two finite dimensional vector spaces cannot be injective if $\dim V > \dim W$, an immersion $\Phi : M \to N$ can only exist when $\dim M \le \dim N$. □

Example 2.40. The map $\Phi : \mathbb{R} \to \mathbb{S}^1$ defined by:

$$\Phi(t) = (\cos t, \sin t)$$

is an immersion. The tangent space of $\mathbb{R}$ at any point t_0 is simply $\operatorname{span}\left\{ \left.\frac{\partial}{\partial t}\right|_{t=t_0} \right\}$. The tangent map $(\Phi_*)_{t_0}$ is given by:

$$(\Phi_*)_{t_0}\left(\frac{\partial}{\partial t}\right) = \left.\frac{\partial \Phi}{\partial t}\right|_{t=t_0} = (-\sin t_0, \cos t_0) \neq 0.$$

Therefore, the "matrix" of Φ_* is a one-by-one matrix with a non-zero entry. Clearly, there is no free column and so Φ_* is injective at every $t_0 \in \mathbb{R}$. This shows Φ is an immersion.

This example tells us that an immersion Φ is not necessary injective. $\square$

Example 2.41. Let M^2 be a regular surface in $\mathbb{R}^3$, then the inclusion map $\iota : M^2 \to \mathbb{R}^3$, defined as $\iota(p) = p \in \mathbb{R}^3$, is a smooth map, since for any local parametrization $\mathsf{F}(u_1, u_2)$ of M^2, we have $\iota \circ \mathsf{F} = \mathsf{F}$, which is smooth by definition (see p.2). We now show that ι is an immersion:

$$(\iota_*)_p \left(\frac{\partial \mathsf{F}}{\partial u_i}\right) = \left.\frac{\partial(\iota \circ \mathsf{F})}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)} = \left.\frac{\partial \mathsf{F}}{\partial u_i}\right|_{\mathsf{F}^{-1}(p)}.$$

Let $\mathsf{F}(u_1, u_2) = (x_1(u_1, u_2),\, x_2(u_1, u_2),\, x_3(u_1, u_2))$, then

$$\frac{\partial \mathsf{F}}{\partial u_i} = \sum_{j=1}^{3} \frac{\partial x_j}{\partial u_i} \mathsf{e}_i$$

where $\{\mathsf{e}_i\}$ is the standard basis of $\mathbb{R}^3$. Therefore, the matrix of ι_* is given by:

$$[\iota_*] = \begin{bmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} \\ \frac{\partial x_3}{\partial u_1} & \frac{\partial x_3}{\partial u_2} \end{bmatrix}.$$

By the condition $0 \neq \dfrac{\partial \mathsf{F}}{\partial u_1} \times \dfrac{\partial \mathsf{F}}{\partial u_2} = \dfrac{\partial(x_2, x_3)}{\partial(u_1, u_2)}\mathsf{e}_1 + \dfrac{\partial(x_3, x_1)}{\partial(u_1, u_2)}\mathsf{e}_2 + \dfrac{\partial(x_1, x_2)}{\partial(u_1, u_2)}\mathsf{e}_3$, at each $p \in M$ one least one of the following is invertible:

$$\frac{\partial(x_2, x_3)}{\partial(u_1, u_2)}, \quad \frac{\partial(x_3, x_1)}{\partial(u_1, u_2)}, \quad \frac{\partial(x_1, x_2)}{\partial(u_1, u_2)}.$$

and hence has the 2×2 identity as its RREF. Using this fact, one can row reduce $[\iota_*]$ so that it becomes:

$$[\iota_*] \to \ldots \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ * & * \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

which has no free column. Therefore, $[\iota_*]$ is an injective linear map at every $p \in M$. This shows ι is an immersion. $\square$

Exercise 2.27. Define a map $\Phi : \mathbb{R}^2 \to \mathbb{R}^4$ by:

$$\Phi(x, y) = (x^3, x^2 y, x y^2, y^3).$$

Show that Φ is an immersion at any $(x, y) \neq (0, 0)$.

Exercise 2.28. Given two immersions $\Phi : M \to N$ and $\Psi : N \to P$ between smooth manifolds M, N and P, show that $\Psi \circ \Phi : M \to P$ is also an immersion.

Exercise 2.29. Consider two smooth maps $\Phi_1 : M_1 \to N_1$ and $\Phi_2 : M_2 \to N_2$ between smooth manifolds. Show that:

(a) If both Φ_1 and Φ_2 are immersions, then so does $\Phi_1 \times \Phi_2 : M_1 \times M_2 \to N_1 \times N_2$.

(b) If Φ_1 is not an immersion, then $\Phi_1 \times \Phi_2$ cannot be an immersion.

A nice property of an immersion $\Phi : M^m \to N^n$ is that for every $\Phi(p) \in N$, one can find a *special* local parametrization G of N such that $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ is an inclusion map from $\mathbb{R}^m$ to $\mathbb{R}^n$, which is a map which takes $(x_1, \ldots, x_m)$ to $(x_1, \ldots, x_m, 0, \ldots, 0)$. Let's state the result in a precise way:

Theorem 2.42 (Immersion Theorem). *Let $\Phi : M^m \to N^{m+k}$ be an immersion at $p \in M$ between two smooth manifolds M^m and N^{n+k} with $k \geq 1$. Given any local parametrization $\mathsf{F} : \mathcal{U}_M \to \mathcal{O}_M$ of M near $p \in M$, and any local parametrization $\mathsf{G} : \mathcal{U}_N \to \mathcal{O}_N$ of N near $\Phi(p) \in N$, there exists a smooth reparametrization map $\psi : \widetilde{\mathcal{U}}_N \to \mathcal{U}_N$ such that:*

$$(\mathsf{G} \circ \psi)^{-1} \circ \Phi \circ \mathsf{F}(u_1, \ldots, u_m) = (u_1, \ldots, u_m, \underbrace{0, \ldots, 0}_{k}).$$

See Figure 2.8 for an illustration.

Proof. The proof uses the Inverse Function Theorem. By translation, we may assume that $\mathsf{F}(0) = p$ and $\mathsf{G}(0) = \Phi(p)$. Given that $(\Phi_*)_p$ is injective, there are n linearly independent rows in the matrix $[(\Phi_*)_p]$. WLOG we may assume that the first m rows of $[(\Phi_*)_p]$ are linearly independent. As such, the matrix can be decomposed into the form:

$$[(\Phi_p)_*] = \begin{bmatrix} A \\ * \end{bmatrix}$$

where A is an invertible $m \times m$ matrix, and $*$ denotes any $k \times m$ matrix.

Now define $\psi : \mathbb{R}^{m+k} \to \mathbb{R}^{m+k}$ as:

(*)

$$\psi(u_1, \ldots, u_m, u_{m+1}, \ldots, u_{m+k}) = \mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1, \ldots, u_m) + (0, \ldots, 0, u_{m+1}, \ldots, u_{m+k}).$$

We claim that this is the map ψ that we want. First note that $\psi(0) = \mathsf{G}^{-1} \circ \Phi(p) = 0$ by our earlier assumption. Next we show that ψ has a smooth inverse near 0. The Jacobian matrix of this map at 0 is given by:

$$[(D\psi)_0] = \begin{bmatrix} A & 0 \\ * & I_k \end{bmatrix}.$$

As rows of A are linearly independent, it is easy to see then all rows of $[(D\psi)_0]$ are linearly independent, and hence $[(D\psi)_0]$ is invertible. By Inverse Function Theorem, ψ is locally invertible near 0, i.e. there exists an open set $\widetilde{\mathcal{U}}_N \subset \mathbb{R}^{m+k}$ containing 0 such that the restricted map:

$$\psi|_{\widetilde{\mathcal{U}}_N} : \widetilde{\mathcal{U}}_N \to \psi(\widetilde{\mathcal{U}}_N) \subset \mathcal{U}_N.$$

has a smooth inverse.

Finally, we verify that this is the map ψ that we want. We compute:

$$(\mathsf{G} \circ \psi)^{-1} \circ \Phi \circ \mathsf{F}(u_1, \ldots, u_m) = \psi^{-1}\left((\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F})(u_1, \ldots, u_m)\right).$$

By (*), we have $\psi(u_1, \ldots, u_m, 0, \ldots, 0) = \mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1, \ldots, u_m)$, and hence:

$$\psi^{-1}\left((\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F})(u_1, \ldots, u_m)\right) = (u_1, \ldots, u_m, 0, \ldots, 0).$$

It completes our proof. □

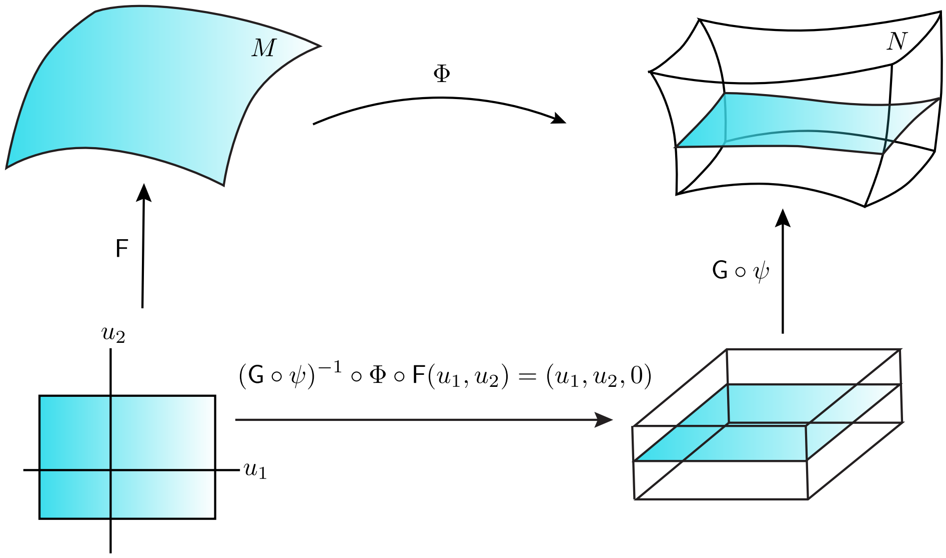


Figure 2.8. Geometric illustration of the Immersion Theorem.

Example 2.43. Consider the map $\Phi : \mathbb{R} \to \mathbb{R}^2$ defined by:

$$\Phi(\theta) = (\cos\theta, \sin\theta).$$

It is easy to see that $[\Phi_*] = (-\sin\theta, \cos\theta) \neq (0,0)$ for any θ. Hence Φ is an immersion. We can locally parametrize $\mathbb{R}^2$ near image of Φ by:

$$\widetilde{\mathsf{G}}(\theta, r) := ((1-r)\cos\theta, (1-r)\sin\theta),$$

then $\widetilde{\mathsf{G}}^{-1} \circ \Phi(\theta) = \widetilde{\mathsf{G}}^{-1}(\cos\theta, \sin\theta) = (\theta, 0)$. Note that the Immersion Theorem (Theorem 2.42) asserts that such $\widetilde{\mathsf{G}}$ exists, it fails to give an explicit form of such a $\widetilde{\mathsf{G}}$. □

Exercise 2.30. Consider the sphere $\mathbb{S}^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ in $\mathbb{R}^3$. Find local parametrizations F for $\mathbb{S}^2$, and G for $\mathbb{R}^3$ such that the composition

$$\mathsf{G}^{-1} \circ \iota \circ \mathsf{F}$$

takes (u_1, u_2) to $(u_1, u_2, 0)$. Here $\iota : \mathbb{S}^2 \to \mathbb{R}^3$ is the inclusion map.

2.5.3. Submersions. Another important type of smooth maps are submersions. Loosely speaking, a submersion is a map that is "locally surjective". Here is the rigorous definition:

Definition 2.44 (Submersions)**.** Let $\Phi : M \to N$ be a smooth map between smooth manifolds M and N. We say Φ is a *submersion at* $p \in M$ if the tangent map $(\Phi_*)_p : T_pM \to T_{\Phi(p)}N$ is surjective. If Φ is a submersion at every point on M, then we simply say Φ is a *submersion*.

Remark 2.45. Clearly, in order for $\Phi : M \to N$ to be a submersion at any point $p \in M$, it is necessary that $\dim M \geq \dim N$. □

Example 2.46. Given two smooth manifolds M and N, the projection maps $\pi_M : M \times N \to M$ and $\pi_N : M \times N \to N$ are both submersions. To verify this, recall that

$T_{(p,q)}(M \times N) = T_pM \oplus T_qN$, and from Exercise 2.22 that $(\pi_M)_* = \pi_{TM}$ where π_{TM} is the projection map of the tangent space:

$$\pi_{TM}(v, w) = v \quad \text{for any } v \in T_pM \text{ and } w \in T_qN.$$

The matrix $[\pi_{TM}]$ is then of the form: $\begin{bmatrix} I & 0 \end{bmatrix}$ where I is the identity matrix of dimension $\dim M$ and 0 is the $\dim M \times \dim N$ zero matrix. There are pivots in every row so π_{TM} is surjective. Similarly we can also show $(\pi_N)_* = \pi_{TN}$ is also surjective. □

Example 2.47. Given a smooth function $f : \mathbb{R}^n \to \mathbb{R}$, and at the point $p \in \mathbb{R}^n$ such that $\nabla f(p) \neq 0$, one can show f is a submersion at p. To show this, let $\{\mathsf{e}_i\}_{i=1}^n$ be the standard basis of $\mathbb{R}^n$, then

$$f_*(\mathsf{e}_i) = f_*\left(\frac{\partial}{\partial x_i}\right) = \frac{\partial f}{\partial x_i}$$

and so the matrix of $[f_*]$ is given by $\begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$. At the point p, we have $\nabla f(p) \neq 0$ which is equivalent to show $[f_*]$ at p is a non-zero $1 \times n$ matrix, which always have 1 pivot in its RREF. Therefore, $(f_*)_p$ is surjective and f is a submersion at p. □

Exercise 2.31. Show that if M and N are two smooth manifolds of equal dimension, then the following are equivalent:

(i) $\Phi : M \to N$ is a local diffeomorphism.

(ii) $\Phi : M \to N$ is an immersion.

(iii) $\Phi : M \to N$ is a submersion.

Exercise 2.32. Find a smooth map $\Phi : \mathbb{R} \to \mathbb{R}$ which is a submersion but is not surjective.

Exercise 2.33. Show that the map $\Phi : \mathbb{R}^{n+1}\backslash\{0\} \to \mathbb{RP}^n$ defined as:

$$\Phi(x_0, \ldots, x_n) = [x_0 : \cdots : x_n]$$

is a submersion.

One nice property of a submersion $\Phi : M^m \to N^n$ that locally around every $p \in M$, one can find *special* local parametrizations F of M near p, and G of N near $\Phi(p)$ such that $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}$ is a projection map. We will see later that this result will show any level-set of Φ, if non-empty, must be a smooth manifold. Let's state this result in a precise way:

Theorem 2.48 (Submersion Theorem). *Let $\Phi : M^{n+k} \to N^n$ be a submersion at $p \in M$ between two smooth manifolds M^{n+k} and N^n with $k \geq 1$. Given any local parametrization $\mathsf{F} : \mathcal{U}_M \to \mathcal{O}_M$ of M near $p \in M$, and any local parametrization $\mathsf{G} : \mathcal{U}_N \to \mathcal{O}_N$ of N near $\Phi(p) \in N$, there exists a smooth reparametrization map $\psi : \widetilde{\mathcal{U}}_M \to \mathcal{U}_M$ such that:*

$$\mathsf{G}^{-1} \circ \Phi \circ (\mathsf{F} \circ \psi)(u_1, \ldots, u_n, u_{n+1}, \ldots, u_{n+k}) = (u_1, \ldots, u_n).$$

See Figure 2.9 for illustration.

Proof. The proof uses again the Inverse Function Theorem. First by translation we may assume that $\mathsf{F}(0) = p$ and $\mathsf{G}(0) = \Phi(p)$. Given that $(\Phi_*)_p$ is surjective, there are n linearly independent columns in the matrix $[(\Phi_*)_p]$. WLOG assume that they are the first n columns, then $[(\Phi_*)_p]$ is of the form:

$$[(\Phi_*)_p] = \left[D(\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F})_0\right] = \begin{bmatrix} A & * \end{bmatrix}$$

where A is an $n \times n$ invertible matrix, and $*$ is any $n \times k$ matrix.

Now define $\phi : \mathcal{U}_M \to \mathbb{R}^{n+k}$ as:

$$(*) \qquad \phi(u_1, \ldots, u_n, u_{n+1}, \ldots, u_{n+k}) = (\underbrace{\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1, \ldots, u_{n+k})}_{\in \mathbb{R}^n}, u_{n+1}, \ldots, u_{n+k}).$$

This map is has an invertible Jacobian matrix at $\mathsf{F}^{-1}(p)$ since:

$$[(D\phi)_0] = \begin{bmatrix} A & * \\ 0 & I \end{bmatrix}.$$

By Inverse Function Theorem, there exists a local inverse $\phi^{-1} : \widetilde{\mathcal{U}}_M \to \phi^{-1}(\widetilde{\mathcal{U}}_M) \subset \mathcal{U}_M$. Finally, we verify that:

$$\begin{aligned} &\mathsf{G}^{-1} \circ \Phi \circ (\mathsf{F} \circ \phi^{-1})(u_1, \ldots, u_n, u_{n+1}, \ldots, u_{n+k}) \\ &= (\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}) \left(\phi^{-1}(u_1, \ldots, u_{n+k})\right) \end{aligned}$$

Let $\phi^{-1}(u_1, \ldots, u_{n+k}) = (v_1, \ldots, v_{n+k})$, then $\phi(v_1, \ldots, v_{n+k}) = (u_1, \ldots, u_{n+k})$. From (*), we get:

$$\begin{aligned} (\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(v_1, \ldots, v_{n+k}), v_{n+1}, \ldots, v_{n+k}) &= \phi(v_1, \ldots, v_{n+k}) \\ &= (u_1, \ldots, u_n, u_{n+1}, \ldots, u_{n+k}) \end{aligned}$$

which implies $\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(v_1, \ldots, v_{n+k}) = (u_1, \ldots, u_n)$. Combine with previous result, we get:

$$\begin{aligned} &(\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}) \left(\phi^{-1}(u_1, \ldots, u_{n+k})\right) \\ &= (\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F})(v_1, \ldots, v_{n+k}) \\ &= (u_1, \ldots, u_n). \end{aligned}$$

Hence, $\mathsf{G}^{-1} \circ \Phi \circ (\mathsf{F} \circ \phi^{-1})$ is the projection from $\mathbb{R}^{n+k}$ onto $\mathbb{R}^n$. It completes the proof by taking $\psi = \phi^{-1}$. $\square$

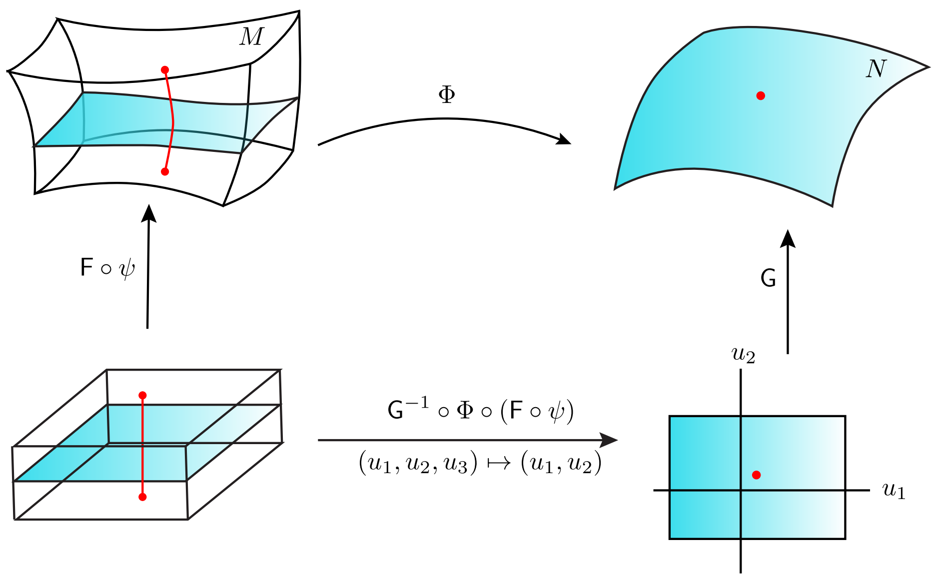


Figure 2.9. Geometric illustration of the Submersion Theorem

2.6. Submanifolds

In this section we talk about submanifolds. A subspace W of a vector space V is a subset of V and is itself a vector space. A subgroup H of a group G is a subset of G and is itself a group. It *seems* that a smooth submanifold N of a smooth manifold M *might* be defined as a subset of M and is itself a smooth manifold. However, it is just one side of the full story – we need more than that because we hope that the local coordinates of a submanifold is in some sense *compatible* with the local coordinates of the manifold M.

Definition 2.49 (Submanifolds)**.** Let M be a smooth n-manifold. A subset $N \subset M$ is said to be a *smooth k-submanifold of M* if N is a smooth k-manifold and the inclusion map $\iota : N \to M$ is an smooth immersion.

Example 2.50. Let $\Phi : M^m \to N^n$ be a smooth map. Define Γ_Φ to be the graph of Φ. Precisely:

$$\Gamma_\Phi = \{(p, \Phi(p)) \in M \times N : p \in M\}.$$

We are going to show that the graph Γ_Φ is a submanifold of $M \times N$, with $\dim \Gamma_\Phi = \dim M$. To show this, consider an arbitrary point $(p, \Phi(p)) \in \Gamma_\Phi$ where $p \in M$. The product manifold $M \times N$ can be locally parametrized by $\mathsf{F} \times \mathsf{G}$ where F is a smooth local parametrization of M near p and G is a smooth local parametrization of N around $\Phi(p)$.

Γ_Φ is locally parametrized around $(p, \Phi(p))$ by:

$$\widetilde{\mathsf{F}}(\mathsf{u}) := (\mathsf{F}(\mathsf{u}), \Phi \circ \mathsf{F}(\mathsf{u})).$$

Here for simplicity, we denote $\mathsf{u} := (u_1, \ldots, u_m)$ where $m = \dim M$. It can be verified that if F_1 and F_2 are compatible (i.e. with smooth transition maps) parametrizations of M around p, then the induced parametrizations $\widetilde{\mathsf{F}}_1$ and $\widetilde{\mathsf{F}}_2$ are also compatible (see exercise below).

Recall that for any $\mathsf{u} = (u_1, \ldots, u_m)$ and $\mathsf{v} = (v_1, \ldots, v_n)$, the product map $\mathsf{F} \times \mathsf{G}$ defined as:

$$(\mathsf{F} \times \mathsf{G})(\mathsf{u}, \mathsf{v}) = (\mathsf{F}(\mathsf{u}), \mathsf{G}(\mathsf{v}))$$

is a local parametrization of $M \times N$. Now we show that the inclusion $\iota : \Gamma_\Phi \to M \times N$ is an immersion:

$$\begin{aligned}(\mathsf{F} \times \mathsf{G})^{-1} \circ \iota \circ \widetilde{\mathsf{F}}(\mathsf{u}) &= (\mathsf{F} \times \mathsf{G})^{-1} \circ \iota\,(\mathsf{F}(\mathsf{u}), \Phi(\mathsf{F}(\mathsf{u}))) \\ &= (\mathsf{F} \times \mathsf{G})^{-1}\,(\mathsf{F}(\mathsf{u}), \Phi(\mathsf{F}(\mathsf{u}))) \\ &= \left(\mathsf{u}, \mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(\mathsf{u})\right).\end{aligned}$$

Its Jacobian matrix has the form:

$$[\iota_*] = \begin{bmatrix} I \\ [\Phi_*] \end{bmatrix}$$

which is injective since its RREF does not have any free column. This completes the proof that ι is an immersion and so Γ_Φ is a submanifold of $M \times N$.

□

Exercise 2.34. Complete the *exercise* stated in Example 2.50 that if F_1 and F_2 are compatible parametrizations of M around p, then the induced parametrizations $\widetilde{\mathsf{F}}_1$ and $\widetilde{\mathsf{F}}_2$ of Γ_Φ are also compatible.

Exercise 2.35. Let M^m be a smooth manifold. Consider the *diagonal* subset $\Delta_M \subset M \times M$ defined as:

$$\Delta_M := \{(x,x) \in M \times M : x \in M\}.$$

Show that Δ_M is a submanifold of $M \times M$.

Exercise 2.36. Show that if N is a submanifold of M, and P is a submanifold of N, then P is also a submanifold of M.

Exercise 2.37. Show that any non-empty open subset N of a smooth manifold M is a submanifold of M with $\dim N = \dim M$.

We require a submanifold to have the inclusion map being smooth because we want to rule out some pathological cases. Consider the graph of an absolute function, i.e. $\Gamma = \{(x,|x|) : x \in \mathbb{R}\}$, and $\mathbb{R}^2$. The graph Γ can be parametrized by a single parametrization $\mathsf{F} : \mathbb{R} \to \Gamma$ defined by:

$$\mathsf{F}(t) = (t, |t|).$$

Then, since Γ equipped with this *single* parametrization, it is considered as a *smooth* manifold (although quite difficult to accept) since there is essentially no transition map. However, we (fortunately) can show that Γ is not a submanifold of $\mathbb{R}^2$ (with usual differential structure, parametrized by the identity map). It is because the inclusion map is not smooth:

$$\mathrm{id}_{\mathbb{R}^2}^{-1} \circ \iota \circ \mathsf{F}(t) = (t, |t|)$$

Exercise 2.38. Show that if $\mathbb{R}^2$ is (pathologically) parametrized by

$$\begin{aligned} \mathsf{G} : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x,y) &\mapsto (x, y + |x|) \end{aligned}$$

and $\Gamma = \{(x,|x|) : x \in \mathbb{R}\}$ is parametrized by $\mathsf{F}(t) = (t,|t|)$, then with differential structures generated by these parametrizations, Γ becomes a submanifold of $\mathbb{R}^2$.

That says: the "pathologically" smooth manifold Γ is a submanifold of this "pathological" $\mathbb{R}^2$.

We require the inclusion map is an immersion because we want a submanifold N of M to be equipped with local coordinates "compatible" with that of M in the following sense:

Proposition 2.51. *If N^n is a submanifold of M^m, then near every point $p \in N$, there exists a smooth local parametrization $\mathsf{G}(u_1, \ldots, u_m) : \mathcal{U} \to \mathcal{O}$ of M near p such that $\mathsf{G}(0) = p$ and*

$$N \cap \mathcal{O} = \{\mathsf{G}(u_1, \ldots, u_n, 0, \ldots, 0) : (u_1, \ldots, u_n, 0, \ldots, 0) \in \mathcal{U}\}.$$

Proof. By Theorem 2.42 (Immersion Theorem), given that $\iota : N \to M$ is an immersion, then around every point $p \in N$ one can find a local parametrization $\mathsf{F} : \mathcal{U}_N \to \mathcal{O}_N$ of N near p, and another local parametrization $\mathsf{G}(u_1, \ldots, u_m) : \mathcal{U}_M \to \mathcal{O}_M$ of M near $\iota(p) = p$ such that:

$$\mathsf{G}^{-1} \circ \iota \circ \mathsf{F}(u_1, \ldots, u_n) = (u_1, \ldots, u_n, \underbrace{0, \ldots, 0}_{m-n})$$

and so $\mathsf{F}(u_1, \ldots, u_n) = \mathsf{G}(u_1, \ldots, u_n, 0, \ldots, 0)$. Note that in order for $\mathsf{G}^{-1} \circ \iota \circ \mathsf{F}$ to be well-defined, we assume (by shrinking the domains if necessary) that $\mathcal{O}_N = N \cap \mathcal{O}_M$.

Therefore,

$$\begin{aligned}&\{\mathsf{G}(u_1,\ldots,u_n,0,\ldots,0) : (u_1,\ldots,u_n,0,\ldots,0) \in \mathcal{U}\}\\&= \{\mathsf{F}(u_1,\ldots,u_n) : (u_1,\ldots,u_n) \in \mathcal{U}_N\}\\&= \mathcal{O}_N = N \cap \mathcal{O}_M\end{aligned}$$

It completes our proof. □

We introduced submersions because the level-set of a submersion, if non-empty, can in fact shown to be a submanifold! We will state and prove this result. Using this fact, one can show a lot of sets are in fact manifolds.

Proposition 2.52. *Let $\Phi : M^m \to N^n$ be a smooth map between two smooth manifolds M and N. Suppose $q \in N$ such that $\Phi^{-1}(q)$ is non-empty, and that Φ is a submersion at any $p \in \Phi^{-1}(q)$, then the level-set $\Phi^{-1}(q)$ is a submanifold of M with* $\dim \Phi^{-1}(q) = m - n$.

Proof. Using Theorem 2.48 (Submersion Theorem), given any point $p \in \Phi^{-1}(q) \subset M$, there exist a local parametrization $\mathsf{F} : \mathcal{U}_M \to \mathcal{O}_M$ of M near p, and a local parametrization G of N near $\Phi(p) = q$, such that:

$$\mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1,\ldots,u_n,u_{n+1},\ldots,u_m) = (u_1,\ldots,u_n)$$

and that $\mathsf{F}(0) = p$, $\mathsf{G}(0) = q$.

We first show that $\Phi^{-1}(q)$ is a smooth manifold. Note that we have:

$$\Phi\left(\mathsf{F}(0,\ldots,0,u_{n+1},\ldots,u_m)\right) = \mathsf{G}(0,\ldots,0) = q.$$

Therefore, $\mathsf{F}(0,\ldots,0,u_{n+1},\ldots,u_m) \in \Phi^{-1}(q)$. Hence, $\Phi^{-1}(q)$ can be locally parametrized by $\widetilde{\mathsf{F}}(u_{n+1},\ldots,u_m) := \mathsf{F}(0,\ldots,0,u_{n+1},\ldots,u_m)$. One can also verify that compatible F's gives compatible $\widetilde{\mathsf{F}}$'s. This shows $\Phi^{-1}(q)$ is a smooth manifold of dimension $m - n$.

To show it is a submanifold of M, we need to compute the tangent map ι_*. First consider the composition:

$$\mathsf{F}^{-1} \circ \iota \circ \widetilde{\mathsf{F}}(u_{n+1},\ldots,u_m) = \mathsf{F}^{-1}(\mathsf{F}(0,\ldots,0,u_{n+1},\ldots,u_m)) = (0,\ldots,0,u_{n+1},\ldots,u_m).$$

The matrix $[\iota_*]$ with respect to local parametrizations $\widetilde{\mathsf{F}}$ of $\Phi^{-1}(q)$, and F of M is given by the Jacobian:

$$[\iota_*] = [D(\mathsf{F}^{-1} \circ \iota \circ \widetilde{\mathsf{F}})] = \begin{bmatrix} 0 \\ I \end{bmatrix}$$

which shows ι_* is injective. Therefore, $\Phi^{-1}(q)$ is a submanifold of M. □

Using Proposition 2.52, one can produce a lot of examples of manifolds which are level-sets of smooth functions.

Example 2.53. In $\mathbb{R}^4$, the set $\Sigma := \{x^3 + y^3 + z^3 + w^3 = 1\}$ is a smooth 3-manifold. It can be shown by consider $\Phi : \mathbb{R}^4 \to \mathbb{R}$ defined by:

$$\Phi(x,y,z,w) = x^3 + y^3 + z^3 + w^3.$$

Then, $\Sigma = \Phi^{-1}(1)$. To show it is a manifold, we show Φ is a submersion at every $p \in \Sigma$. By direct computation, we get:

$$[\Phi_*] = [3x^2 \quad 3y^2 \quad 3z^2 \quad 3w^2]$$

Since $[\Phi_*] = 0$ only when $(x,y,z,w) = (0,0,0,0)$ which is not contained in Σ, we have shown $(\Phi_*)_p$ is injective for any $p \in \Sigma$. By Proposition 2.52, we have proved $\Sigma = \Phi^{-1}(1)$ is a smooth manifold of dimension $4 - 1 = 3$. □

Example 2.54. The set $M_{n\times n}(\mathbb{R})$ of all $n \times n$ real matrices can be regarded as $\mathbb{R}^{2n}$ equipped with the usual differential structure. Consider these subsets of $M_{n\times n}(\mathbb{R})$:

(a) $\mathrm{GL}(n,\mathbb{R})$ = the set of all invertible $n \times n$ matrices;

(b) $\mathrm{Sym}(n,\mathbb{R})$ = the set of all symmetric $n \times n$ matrices;

(c) $\mathrm{O}(n,\mathbb{R})$ = the set of all orthogonal matrices;

We are going to show that they are all submanifolds of $M_{n\times n}(\mathbb{R})$. Consider the determinant function $f : M_{n\times n}(\mathbb{R}) \to \mathbb{R}$ defined as:

$$f(A) := \det(A).$$

Since f is a continuous function, the set $\mathrm{GL}(n,\mathbb{R}) = f^{-1}(\mathbb{R}\backslash\{0\})$ is an open subset of $M_{n\times n}(\mathbb{R})$. Any (non-empty) open subset N of a smooth manifold M is a submanifold of M with $\dim N = \dim M$ (see Exercise 2.37).

For $\mathrm{Sym}(n,\mathbb{R})$, we first label the coordinates of $\mathbb{R}^{\frac{n(n+1)}{2}}$ by (x_{ij}) where and $1 \leq i \leq j \leq n$. Then one can parametrize $\mathrm{Sym}(n,\mathbb{R})$ by $\mathsf{F} : \mathbb{R}^{\frac{n(n+1)}{2}} \to \mathrm{Sym}(n,\mathbb{R})$ taking $(x_{ij})_{i\leq j} \in \mathbb{R}^{\frac{n(n+1)}{2}}$ to the matrix A with (i,j)-th entry x_{ij} when $i \leq j$, and x_{ji} when $i > j$. For instance, when $n = 2$, $\mathbb{R}^{\frac{n(n+1)}{2}}$ becomes $\mathbb{R}^3$ with coordinates labelled by (x_{11}, x_{12}, x_{22}). The parametrization F will take the point $(x_{11}, x_{12}, x_{22}) = (a,b,c) \in \mathbb{R}^3$ to the matrix:

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathrm{Sym}(n,\mathbb{R}).$$

Back to the general n, this F is a global parametrization and it makes $\mathrm{Sym}(n,\mathbb{R})$ a smooth $\frac{n(n+1)}{2}$-manifold. To show that it is a submanifold of $M_{n\times n}(\mathbb{R})$, we computed the tangent map ι_* of the inclusion map $\iota : \mathrm{Sym}(n,\mathbb{R}) \to M_{n\times n}(\mathbb{R})$:

$$\begin{aligned} \iota_*\left(\frac{\partial}{\partial x_{ij}}\right) &= \frac{\partial \iota}{\partial x_{ij}} = \frac{\partial}{\partial x_{ij}}(\iota \circ \mathsf{F}) \\ &= \frac{1}{2}(E_{ij} + E_{ji}) \end{aligned}$$

where E_{ij} is the $n \times n$ matrix with 1 in the (i,j)-th entry, and 0 elsewhere. The tangent space $T\mathrm{Sym}(n,\mathbb{R})$ at each point is spanned by the basis $\left\{\frac{\partial}{\partial x_{ij}}\right\}_{1\leq i\leq j\leq n}$. Its image $\left\{\frac{1}{2}(E_{ij} + E_{ji})\right\}_{1\leq i\leq j\leq n}$ under the map ι_* is linearly independent (why?). This shows ι_* is injective, and hence $\mathrm{Sym}(n,\mathbb{R})$ is a submanifold of $M_{n\times n}(\mathbb{R})$. The image of the inclusion map is the set of all symmetric matrices in $M_{n\times n}(\mathbb{R})$, hence we conclude that $T_{A_0}\mathrm{Sym}(n,\mathbb{R}) \cong T\mathrm{Sym}(n,\mathbb{R})$ for any $A_0 \in \mathrm{Sym}(n,\mathbb{R})$.

The set of all orthogonal matrices $\mathrm{O}(n)$ can be regarded as the level-set $\Phi^{-1}(I)$ of the following map:

$$\begin{aligned} \Phi : M_{n\times n}(\mathbb{R}) &\to \mathrm{Sym}(n,\mathbb{R}) \\ A &\mapsto A^T A \end{aligned}$$

We are going to show that Φ is a submersion at every $A_0 \in \Phi^{-1}(I)$, we compute its tangent map:

$$(\Phi_*)\left(\frac{\partial}{\partial x_{ij}}\right) = \frac{\partial}{\partial x_{ij}} A^T A = E_{ij}^T A + A^T E_{ij} = (A^T E_{ij})^T + A^T E_{ij}.$$

From now on we denote $[A]_{ij}$ to be the (i,j)-th entry of any matrix A (Be cautious that E_{ij} without the square brackets is a matrix, not the (i,j)-th entry of E). In fact, any matrix A can be written as:

$$A = \sum_{i,j=1}^{n} [A]_{ij} E_{ij}.$$

At $A_0 \in \Phi^{-1}(I)$, we have $A_0^T A_0 = I$ and so for any symmetric matrix B, we have:

$$(\Phi_*)_{A_0}\left(\frac{\partial}{\partial x_{ij}}\right) = (A_0^T E_{ij})^T + A_0^T E_{ij}$$

$$\begin{aligned}
(\Phi_*)_{A_0}\left(\frac{1}{2}\sum_{i,j=1}^{n}[A_0B]_{ij}\frac{\partial}{\partial x_{ij}}\right) &= \frac{1}{2}\sum_{i,j=1}^{n}[A_0B]_{ij}\left((A_0^T E_{ij})^T + A_0^T E_{ij}\right) \\
&= \frac{1}{2}\left(A_0^T \sum_{i,j=1}^{n}[A_0B]_{ij}E_{ij}\right)^T + \frac{1}{2}\left(A_0^T \sum_{i,j=1}^{n}[A_0B]_{ij}E_{ij}\right) \\
&= \frac{1}{2}(A_0^T A_0 B)^T + \frac{1}{2}(A_0^T A_0 B) \\
&= \frac{1}{2}B^T + \frac{1}{2}B = B.
\end{aligned}$$

Therefore, $(\Phi_*)_{A_0}$ is surjective. This shows Φ_* is a submersion at every point $A_0 \in \Phi^{-1}(I)$. This shows $\mathrm{Sym}(n,\mathbb{R}) = \Phi^{-1}(I)$ is a submanifold of $M_{n\times n}(\mathbb{R})$ of dimension $\dim M_{n\times n}(\mathbb{R}) - \dim \mathrm{Sym}(n,\mathbb{R})$, which is $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$.

$\square$

Exercise 2.39. Show that the subset Σ of $\mathbb{R}^3$ defined by the two equations below is a 1-dimensional manifold:

$$\begin{aligned}
x^3 + y^3 + z^3 &= 1 \\
x + y + z &= 0
\end{aligned}$$

Exercise 2.40. Define

- $\mathrm{SL}(n,\mathbb{R})$ = the set of all $n \times n$ matrices with determinant 1
- $\mathfrak{sl}(n,\mathbb{R})$ = the set of all $n \times n$ skew-symmetric matrices (i.e. the set of matrices $A \in M_{n\times n}(\mathbb{R})$ such that $A^T = -A$).

Show that they are both submanifolds of $M_{n\times n}(\mathbb{R})$, and find their dimensions.

Exercise 2.41. Consider the map $\Phi : \mathbb{S}^3 \backslash \{(0,0)\} \to \mathbb{CP}^1$ defined by:

$$\Phi(x_1, x_2, x_3, x_4) := [x_1 + ix_2 : x_3 + ix_4].$$

Show that $\Phi^{-1}([1:0])$ is a smooth manifold of (real) dimension 1, and show that $\Phi^{-1}([1:0])$ is diffeomorphic to a circle.

Chapter 3

Tensors and Differential Forms

> "In the beginning, God said, the four-dimensional divergence of an antisymmetric, second-rank tensor equals zero, and there was light."
>
> *Michio Kaku*

In Multivariable Calculus, we learned about gradient, curl and divergence of a vector field, and three important theorems associated to them, namely Green's, Stokes' and Divergence Theorems. In this and the next chapters, we will generalize these theorems to higher dimensional manifolds, and unify them into one single theorem (called the *Generalized Stokes' Theorem*). In order to carry out this generalization and unification, we need to introduce tensors and differential forms. The reasons of doing so are *many-folded*. We will explain it in detail. Meanwhile, one obvious reason is that the curl of a vector field is only defined in $\mathbb{R}^3$ since it uses the cross product. In this chapter, we will develop the language of *differential forms* which will be used in place of gradient, curl, divergence and all that in Multivariable Calculus.

3.1. Cotangent Spaces

3.1.1. Review of Linear Algebra: dual spaces. Let V be an n-dimensional real vector space, and $\mathcal{B} = \{\mathsf{e}_1, \ldots, \mathsf{e}_n\}$ be a basis for V. The set of all linear maps $T : V \to \mathbb{R}$ from V to the scalar field $\mathbb{R}$ (they are commonly called *linear functionals*) forms a vector space with dimension n. This space is called the *dual space* of V, denoted by V^*.

Associated to the basis $\mathcal{B} = \{\mathsf{e}_i\}_{i=1}^n$ for V, there is a basis $\mathcal{B}^* = \{\mathsf{e}_i^*\}_{i=1}^n$ for V^*:

$$\mathsf{e}_i^*(\mathsf{e}_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The basis $\mathcal{B}^*$ for V^* (do Exericse 3.1 to verify it is indeed a basis) is called the *dual basis* of V^* with respect to $\mathcal{B}$.

Exercise 3.1. Given that V is a finite-dimensional real vector space, show that:

(a) V^* is a vector space

(b) $\dim V^* = \dim V$

(c) If $\mathcal{B} = \{\mathsf{e}_i\}_{i=1}^n$ is a basis for V, then $\mathcal{B}^* := \{\mathsf{e}_i^*\}_{i=1}^n$ is a basis for V^*.

Given $T \in V^*$ and that $T(\mathsf{e}_i) = a_i$, verify that:

$$T = \sum_{i=1}^n a_i \mathsf{e}_i^*.$$

3.1.2. Cotangent Spaces of Smooth Manifolds. Let M^n be a smooth manifold. Around $p \in M$, suppose there is a local parametrization $\mathsf{F}(u_1, \ldots, u_n)$. Recall that the tangent space T_pM at p is defined as the span of partial differential operators $\left\{\frac{\partial}{\partial u_i}(p)\right\}_{i=1}^n$. The *cotangent space* denoted by T_p^*M is defined as follows:

Definition 3.1 (Cotangent Spaces)**.** Let M^n be a smooth manifold. At every $p \in M$, the *cotangent space of M at p* is the dual space of the tangent space T_pM, i.e.:

$$T_p^*M = \left(T_pM\right)^*.$$

The elements in T_p^*M are called *cotangent vectors* of M at p.

Remark 3.2. Some authors use T_pM^* to denote the cotangent space. □

Associated to the basis $\mathcal{B}_p = \left\{\frac{\partial}{\partial u_i}(p)\right\}_{i=1}^n$ of T_pM, there is a dual basis $\mathcal{B}_p^* = \left\{(du^1)_p, \ldots, (du^n)_p\right\}$ for T_p^*M, which is defined as follows:

$$(du^i)_p\left(\frac{\partial}{\partial u_j}(p)\right) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

As $(du^i)_p$ is a linear map from T_pM to $\mathbb{R}$, from the above definition we have:

$$(du^i)_p\left(\sum_{j=1}^n a_j \frac{\partial}{\partial u_j}(p)\right) = \sum_{j=1}^n a_j \delta_{ij} = a_i.$$

Occasionally (just for aesthetic purpose), $\left(du^i\right)_p$ can be denoted as $du^i\big|_p$. Moreover, whenever p is clear from the context (or not significant), we may simply write du^i and $\frac{\partial}{\partial u_i}$.

Note that both $\mathcal{B}_p$ and $\mathcal{B}_p^*$ depend on the choice of local coordinates. Suppose $(v_1, \ldots, v_n)$ is another local coordinates around p, then by chain rule we have:

$$\frac{\partial}{\partial v_j} = \sum_{k=1}^n \frac{\partial u_k}{\partial v_j}\frac{\partial}{\partial u_k}$$

$$\frac{\partial}{\partial u_j} = \sum_{k=1}^n \frac{\partial v_k}{\partial u_j}\frac{\partial}{\partial v_k}.$$

We are going to express dv^i in terms of du^j's:

$$\begin{aligned} dv^i\left(\frac{\partial}{\partial u_j}\right) &= dv^i\left(\sum_{k=1}^{n}\frac{\partial v_k}{\partial u_j}\frac{\partial}{\partial v_k}\right) \\ &= \sum_{k=1}^{n}\frac{\partial v_k}{\partial u_j}dv^i\left(\frac{\partial}{\partial v_k}\right) \\ &= \sum_{k=1}^{n}\frac{\partial v_k}{\partial u_j}\delta_{ik} \\ &= \frac{\partial v_i}{\partial u_j}. \end{aligned}$$

This proves the transition formula for the cotangent basis:

$$dv^i = \sum_{k=1}^{n}\frac{\partial v_i}{\partial u_k}du^k. \tag{3.1}$$

Example 3.3. Consider $M = \mathbb{R}^2$ which can be parametrized by

$$\begin{aligned} \mathsf{F}_1(x,y) &= (x,y) \\ \mathsf{F}_2(r,\theta) &= (r\cos\theta, r\sin\theta). \end{aligned}$$

From (3.1), the conversion between $\{dr, d\theta\}$ and $\{dx, dy\}$ is given by:

$$\begin{aligned} dx &= \frac{\partial x}{\partial r}dr + \frac{\partial x}{\partial \theta}d\theta \\ &= (\cos\theta)\ dr - (r\sin\theta)\ d\theta \\ dy &= \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \theta}d\theta \\ &= (\sin\theta)\ dr + (r\cos\theta)\ d\theta \end{aligned}$$

□

Exercise 3.2. Consider $M = \mathbb{R}^3$ which can be parametrized by:

$$\begin{aligned} \mathsf{F}_1(x,y,z) &= (x,y,z) \\ \mathsf{F}_2(r,\theta,z) &= (r\cos\theta, r\sin\theta, z) \\ \mathsf{F}_3(\rho,\phi,\theta) &= (\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi) \end{aligned}$$

Express $\{dx, dy, dz\}$ in terms of $\{dr, d\theta, dz\}$ and $\{d\rho, d\phi, d\theta\}$.

Exercise 3.3. Suppose $\mathsf{F}(u_1, \ldots, u_n)$ and $\mathsf{G}(v_1, \ldots, v_n)$ are two local parametrizations of a smooth manifold M. Let $\omega : M \to TM$ be a smooth differential 1-form such that on the overlap of local coordinates we have:

$$\omega = \sum_j a_j du^j = \sum_i b_i dv^i.$$

Find a conversion formula between a_j's and b_i's.

3.2. Tangent and Cotangent Bundles

3.2.1. Definitions. Let M be a smooth manifold. Loosely speaking, the *tangent bundle* (resp. *cotangent bundle*) are defined as the disjoint union of all tangent (resp. cotangent) spaces over the whole M. Precisely:

Definition 3.4 (Tangent and Cotangent Bundles)**.** Let M be a smooth manifold. The *tangent bundle*, denoted by TM, is defined to be:

$$TM = \bigcup_{p \in M} \left(\{p\} \times T_pM\right).$$

Elements in TM can be written as (p, V) where $V \in T_pM$.

Similarly, the *cotangent bundle*, denoted by T^*M, is defined to be:

$$T^*M = \bigcup_{p \in M} \left(\{p\} \times T_p^*M\right).$$

Elements in T^*M can be written as (p, ω) where $\omega \in T_p^*M$.

Suppose $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \to M$ is a local parametrization of M, then a general element of TM can be written as:

$$\left(p, \sum_{i=1}^n V^i \frac{\partial}{\partial u_i}(p)\right)$$

and a general element of T^*M can be written as:

$$\left(p, \sum_{i=1}^n a_i \, du^i\Big|_p\right).$$

We are going to explain why both TM and T^*M are smooth manifolds. The local parametrization $\mathsf{F}(u_1, \ldots, u_n)$ of M induces a local parametrization $\widetilde{\mathsf{F}}$ of TM defined by:

$$\begin{aligned} \widetilde{\mathsf{F}} &: \mathcal{U} \times \mathbb{R}^n \to TM \\ (u_1, \ldots, u_n; V^1, \ldots, V^n) &\mapsto \left(\mathsf{F}(u_1, \ldots, u_n), \sum_{i=1}^n V^i \frac{\partial}{\partial u_i}\Big|_{\mathsf{F}(u_1,\ldots,u_n)}\right). \end{aligned} \tag{3.2}$$

Likewise, it also induces a local parametrization $\widetilde{\mathsf{F}}^*$ of T^*M defined by:

$$\begin{aligned} \widetilde{\mathsf{F}}^* &: \mathcal{U} \times \mathbb{R}^n \to T^*M \\ (u_1, \ldots, u_n; a_1, \ldots, a_n) &\mapsto \left(\mathsf{F}(u_1, \ldots, u_n), \sum_{i=1}^n a_i \, du^i\Big|_{\mathsf{F}(u_1,\ldots,u_n)}\right). \end{aligned} \tag{3.3}$$

It suggests that TM and T^*M are both smooth manifolds of dimension $2\dim M$. To do so, we need to verify compatible F's induce compatible $\widetilde{\mathsf{F}}$ and $\widetilde{\mathsf{F}}^*$. Let's state this as a proposition and we leave the proof as an exercise for readers:

Proposition 3.5. *Let M^n be a smooth manifold. Suppose F and G are two overlapping smooth local parametrizations of M, then their induced local parametrizations $\widetilde{\mathsf{F}}$ and $\widetilde{\mathsf{G}}$ defined as in* (3.2) *on the tangent bundle TM are compatible, and also that $\widetilde{\mathsf{F}}^*$ and $\widetilde{\mathsf{G}}^*$ defined as in* (3.3) *on the cotangent bundle T^*M are also compatible.*

Corollary 3.6. *The tangent bundle TM and the cotangent bundle T^*M of a smooth manifold M are both smooth manifolds of dimension* $2\dim M$.

Exercise 3.4. Prove Proposition 3.5.

Exercise 3.5. Show that the bundle map $\pi : TM \to M$ taking $(p, V) \in TM$ to $p \in M$ is a submersion. Show also that the set:

$$\Sigma_0 := \{(p, 0) \in TM : p \in M\}$$

is a submanifold of TM.

3.2.2. Vector Fields. Intuitively, a vector field V on a manifold M is an assignment of a vector to each point on M. Therefore, it can be regarded as a map $V : M \to TM$ such that $V(p) \in \{p\} \times T_pM$. Since we have shown that the tangent bundle TM is also a smooth manifold, one can also talk about C^k and *smooth* vector fields.

Definition 3.7 (Vector Fields of Class C^k)**.** Let M be a smooth manifold. A map $V : M \to TM$ is said to be a *vector field* if for each $p \in M$, we have $V(p) = (p, V_p) \in \{p\} \times T_pM$.

If V is of class C^k as a map between M and TM, then we say V is a C^k *vector field*. If V is of class C^∞, then we say V is a *smooth vector field*.

Remark 3.8. In the above definition, we used $V(p)$ to be denote the element (p, V_p) in TM, and V_p to denote the vector in T_pM. We will distinguish between them for a short while. After getting used to the notations, we will abuse the notations and use V_p and $V(p)$ interchangeably. □

Remark 3.9. Note that a vector field can also be defined *locally* on an open set $\mathcal{O} \subset M$. In such case we say V *is a* C^k *on* $\mathcal{O}$ if the map $V : \mathcal{O} \to TM$ is C^k. □

Under a local parametrization $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \to M$ of M, a vector field $V : M \to TM$ can be expressed in terms of local coordinates as:

$$V(p) = \left(p, \sum_{i=1}^{n} V^i(p) \frac{\partial}{\partial u_i}(p)\right).$$

The functions $V^i : \mathsf{F}(\mathcal{U}) \subset M \to \mathbb{R}$ are all locally defined and are commonly called the *components* of V with respect to local coordinates $(u_1, \ldots, u_n)$.

Let $\widetilde{\mathsf{F}}(u_1, \ldots, u_n; V^1, \ldots, V^n)$ be the induced local parametrization of TM defined as in (3.2). Then, one can verify that:

$$\begin{aligned}\widetilde{\mathsf{F}}^{-1} \circ V \circ \mathsf{F}(u_1, \ldots, u_n) &= \widetilde{\mathsf{F}}^{-1}\left(\mathsf{F}(u_1, \ldots, u_n), \sum_{i=1}^{n} V^i(\mathsf{F}(u_1, \ldots, u_n)) \left.\frac{\partial}{\partial u_i}\right|_{\mathsf{F}(u_1, \ldots, u_n)}\right) \\ &= \left(u_1, \ldots, u_n; V^1(\mathsf{F}(u_1, \ldots, u_n)), \ldots, V^n(\mathsf{F}(u_1, \ldots, u_n))\right).\end{aligned}$$

Therefore, $\widetilde{\mathsf{F}}^{-1} \circ V \circ \mathsf{F}(u_1, \ldots, u_n)$ is smooth if and only if the components V^i's are all smooth. Similarly for class C^k. In short, a vector field V is smooth if and only if the components V^i in every its local expression:

$$V(p) = \left(p, \sum_{i=1}^{n} V^i(p) \frac{\partial}{\partial u_i}(p)\right)$$

are all smooth.

3.2.3. Differential 1-Forms. Differential 1-forms are the *dual* counterpart of vector fields. It is essentially an assignment of a cotangent vector to each point on M. Precisely:

Definition 3.10 (Differential 1-Forms of Class C^k). Let M be a smooth manifold. A map $\omega : M \to T^*M$ is said to be a *differential 1-form* if for each $p \in M$, we have $\omega(p) = (p, \omega_p) \in \{p\} \times T_p^*M$.

If ω is of class C^k as a map between M and T^*M, then we say ω is a C^k *differential 1-form*. If ω is of class C^∞, then we say ω is a *smooth differential 1-form*.

Remark 3.11. At this moment we use $\omega(p)$ to denote an element in $\{p\} \times T_pM$, and ω_p to denote an element in T_p^*M. We will abuse the notations later on and use them interchangeably, since such a distinction is unnecessary for many practical purposes. □

Remark 3.12. If a differential 1-form ω is locally defined on an open set $\mathcal{O} \subset M$, we may say ω is C^k on $\mathcal{O}$ to mean the map $\omega : \mathcal{O} \to T^*M$ is of class C^k. □

Under a local parametrization $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \to M$ of M, a differential 1-form $\omega : M \to T^*M$ has a local coordinate expression given by:

$$\omega(p) = \left(p, \sum_{i=1}^{n} \omega_i(p)\, du^i\Big|_p\right)$$

where $\omega_i : \mathsf{F}(\mathcal{U}) \subset M \to \mathbb{R}$ are locally defined functions and are commonly called the *components* of ω with respect to local coordinates $(u_1, \ldots, u_n)$. Similarly to vector fields, one can show that ω is a C^∞ differential 1-form if and only if all ω_i's are smooth under any local coordinates in the atlas of M (see Exercise 3.6).

Exercise 3.6. Show that a differential 1-form ω is C^k on M if and only if all components ω_i's are C^k under any local coordinates in the atlas of M.

Example 3.13. The differential 1-form:

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

is smooth on $\mathbb{R}^2 \backslash \{0\}$, but is not smooth on $\mathbb{R}^2$. □

3.2.4. Push-Forward and Pull-Back. Consider a smooth map $\Phi : M \to N$ between two smooth manifolds M and N. The tangent map at p denoted by $(\Phi_*)_p$ is the induced map between tangent spaces T_pM and $T_{\Phi(p)}N$. Apart from calling it the tangent map, we often call Φ_* to be the *push-forward by* Φ, since Φ and Φ_* are both from the space M to the space N.

The push-forward map Φ_* takes tangent vectors on M to tangent vectors on N. There is another induced map Φ^*, called the *pull-back by* Φ, which is *loosely* defined as follows:

$$(\Phi^*\omega)(V) = \omega(\Phi_* V)$$

where ω is a cotangent vector and V is a tangent vector. In order for the above to make sense, V has to be a tangent vector on M (say at p). Then, $\Phi_* V$ is a tangent vector in $T_{\Phi(p)}N$. Therefore, $\Phi^*\omega$ needs to act on V and hence is a cotangent vector in T_p^*M; whereas ω acts on $\Phi_* V$ and so it should be a cotangent vector in $T_{\Phi(p)}^*N$. It is precisely defined as follows:

Definition 3.14 (Pull-Back of Cotangent Vectors)**.** Let $\Phi : M \to N$ be a smooth map between two smooth manifolds M and N. Given any cotangent vector $\omega_{\Phi(p)} \in T^*_{\Phi(p)}N$, the *pull-back of ω by Φ at p* denoted by $(\Phi^*\omega)_p$ is an element in T^*_pM and is defined to be the following linear functional on T_pM:

$$(\Phi^*\omega)_p : T_pM \to \mathbb{R}$$
$$(\Phi^*\omega)_p (V_p) := \omega_{\Phi(p)} \big((\Phi_*)_p(V_p)\big)$$

Therefore, one can think of Φ^* is a map which takes a *cotangent vector* $\omega_{\Phi(p)} \in T^*_{\Phi(p)}N$ to a cotangent vector $(\Phi^*\omega)_p$ on T^*_pM. As it is in the opposite direction to $\Phi : M \to N$, we call Φ^* the *pull-back* whereas Φ_* is called the *push-forward*.

Remark 3.15. In many situations, the points p and $\Phi(p)$ are clear from the context. Therefore, we often omit the subscripts p and $\Phi(p)$ when dealing with pull-backs and push-forwards. □

Example 3.16. Consider the map $\Phi : \mathbb{R} \to \mathbb{R}^2\backslash\{0\}$ defined by:

$$\Phi(\theta) = (\cos\theta, \sin\theta).$$

Let ω be the following 1-form on $\mathbb{R}^2\backslash\{0\}$:

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

First note that

$$\Phi_*\left(\frac{\partial}{\partial\theta}\right) = \frac{\partial\Phi}{\partial\theta} = \frac{\partial\overbrace{(\cos\theta)}^{x}}{\partial\theta}\frac{\partial}{\partial x} + \frac{\partial\overbrace{(\sin\theta)}^{y}}{\partial\theta}\frac{\partial}{\partial y} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}.$$

Therefore, one can compute:

$$\begin{aligned}(\Phi^*\omega)\left(\frac{\partial}{\partial\theta}\right) &= \omega\left(\Phi_*\left(\frac{\partial}{\partial\theta}\right)\right) = \omega\left(-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}\right)\\ &= -y\left(\frac{-y}{x^2+y^2}\right) + x\left(\frac{x}{x^2+y^2}\right)\\ &= 1.\end{aligned}$$

Therefore, $\Phi^*\omega = d\theta$. □

Example 3.17. Let $M := \mathbb{R}^2\backslash\{(0,0)\}$ (equipped with polar (r,θ)-coordinates) and $N = \mathbb{R}^2$ (with (x,y)-coordinates), and define:

$$\Phi : M \to N$$
$$\Phi(r,\theta) := (r\cos\theta, r\sin\theta)$$

One can verify that:

$$\begin{aligned}\Phi_*\left(\frac{\partial}{\partial r}\right) &= \frac{\partial\Phi}{\partial r} = (\cos\theta)\frac{\partial}{\partial x} + (\sin\theta)\frac{\partial}{\partial y}\\ \Phi_*\left(\frac{\partial}{\partial\theta}\right) &= \frac{\partial\Phi}{\partial\theta} = (-r\sin\theta)\frac{\partial}{\partial x} + (r\cos\theta)\frac{\partial}{\partial y}\\ &= -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}\end{aligned}$$

Hence, we have:

$$\begin{aligned}
(\Phi^* dx)\left(\frac{\partial}{\partial r}\right) &= dx\left(\Phi_*\left(\frac{\partial}{\partial r}\right)\right) \\
&= dx\left((\cos\theta)\frac{\partial}{\partial x} + (\sin\theta)\frac{\partial}{\partial y}\right) \\
&= \cos\theta \\
(\Phi^* dx)\left(\frac{\partial}{\partial \theta}\right) &= dx\left(\Phi_*\left(\frac{\partial}{\partial \theta}\right)\right) \\
&= dx\left(-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}\right) \\
&= -y = -r\sin\theta
\end{aligned}$$

We conclude:

$$\Phi^* dx = \cos\theta\, dr - r\sin\theta\, d\theta.$$

□

Given a smooth map $\Phi : M^m \to N^n$, and local coordinates $(u_1, \ldots, u_m)$ of M around p and local coordinates $(v_1, \ldots, v_n)$ of N around $\Phi(p)$. One can compute a local expression for Φ^*:

$$\Phi^* dv^i = \sum_{j=1}^{n} \frac{\partial v_i}{\partial u_j} du^j \tag{3.4}$$

where $(v_1, \ldots, v_n)$ is regarded as a function of $(u_1, \ldots, u_m)$ via the map $\Phi : M \to N$.

Exercise 3.7. Prove (3.4).

Exercise 3.8. Express $\Phi^* dy$ in terms of dr and $d\theta$ in Example 3.17. Try computing it directly and then verify that (3.4) gives the same result.

Exercise 3.9. Denote (x_1, x_2) the coordinates for $\mathbb{R}^2$ and (y_1, y_2, y_3) the coordinates for $\mathbb{R}^3$. Define the map $\Phi : \mathbb{R}^2 \to \mathbb{R}^3$ by:

$$\Phi(x_1, x_2) = (x_1 x_2, x_2 x_3, x_3 x_1).$$

Compute $\Phi^*(dy^1)$, $\Phi^*(dy^2)$ and $\Phi^*(dy^3)$.

Exercise 3.10. Consider the map $\Phi : \mathbb{R}^3 \backslash \{0\} \to \mathbb{RP}^2$ defined by:

$$\Phi(x, y, z) = [x : y : z].$$

Consider the local parametrization $\mathsf{F}(u_1, u_2) = [1 : u_1 : u_2]$ of $\mathbb{RP}^2$. Compute $\Phi^*(du^1)$ and $\Phi^*(du^2)$.

3.2.5. Lie Derivatives. Derivatives of a function f or a vector field Y in Euclidean spaces along a curve $\gamma(t) : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ are defined as follows:

$$\begin{aligned}
D_{\gamma'(t)} f &:= \frac{d}{dt}(f \circ \gamma)(t) = \lim_{\delta \to 0} \frac{f(\gamma(t+\delta)) - f(\gamma(t))}{\delta} \\
D_{\gamma'(t)} Y &:= \frac{d}{dt}(Y \circ \gamma)(t) = \lim_{\delta \to 0} \frac{Y(\gamma(t+\delta)) - Y(\gamma(t))}{\delta}
\end{aligned}$$

Now given any vector field X and any point $p \in \mathbb{R}^n$, if one can find a curve $\gamma(t) : (-\varepsilon, \varepsilon) \to M$ such that $\gamma'(t) = X(\gamma(t))$ for $t \in (-\varepsilon, \varepsilon)$ and $\gamma(0) = 0$, then it is well-defined to denote:

$$(D_X Y)_p := D_{\gamma'(t)} Y, \quad (D_X f)_p := D_{\gamma'(t)} f \qquad \text{at } t = 0$$

By the existence and uniqueness theorems of ODE, such a curve $\gamma(t)$ exists uniquely provided that the vector field X is C^1.

Furthermore, it can also be checked that if $\gamma_1, \gamma_2 : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ are two curves with $\gamma_1(0) = \gamma_2(0) = p$ and with the same velocity vectors $\gamma_1'(0) = \gamma_2'(0)$ at p, then it is necessarily that $D_{\gamma_1'} f = D_{\gamma_2'} f$ and $D_{\gamma_1'} Y = D_{\gamma_2'} Y$ at p. Therefore, just the existence theorem of ODE is sufficient to argue that $D_X Y$ and $D_X f$ are well-defined.

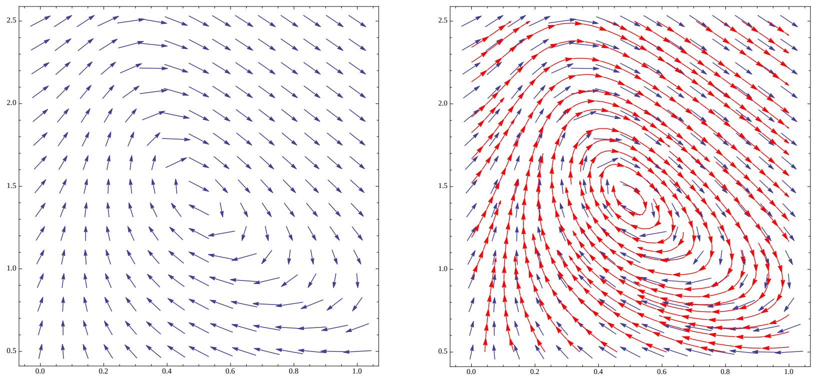
Exercise 3.11. Prove the above claim that $D_{\gamma_1'} f = D_{\gamma_2'} f$ and $D_{\gamma_1'} Y = D_{\gamma_2'} Y$ at p.

Remark 3.18. Consult any standard textbook about theory of ODEs for a proof of existence and uniqueness of the curve $\gamma(t)$ given any vector field X. Most standard textbook uses contraction mapping to prove existence, and Gronwall's inequality to prove uniqueness. □

Now let M be a smooth manifold, and X be a smooth vector field on M. Then, one can also extend the existence and uniqueness theorem of ODE to manifolds to prove that for any point $p \in M$, there exists a smooth curve $\gamma(t) : (-\varepsilon, \varepsilon) \to M$ on M with $\gamma(0) = p$ such that:

$$\frac{d}{dt}\gamma(t) = X(\gamma(t)).$$

Recall that $\frac{d}{dt}\gamma(t)$ is defined as $\gamma_*\left(\frac{\partial}{\partial t}\right)$. This curve γ is called the *integral curve* of X passing through p. This extension can be justified by applying standard ODE theorems on the local coordinate chart covering p. Then one solves for the integral curve within this chart until the curve approaches the boundary of the chart (say at point q). Since the boundary of one chart must be the interior of another local coordinate chart, one can then continue solving for the integral curve starting from q.



(a) vector field X **(b)** integral curves of X

Figure 3.1. a vector field and its integral curves

Now one can still talk about integral curves $\gamma(t)$ given a vector field X on a manifold, so one can define $D_{\gamma'(t)} f$ and $D_X f$ in the same way as in $\mathbb{R}^n$ (as it makes perfect sense to talk about $f(\gamma(t+\delta)) - f(\gamma(t))$. However, it is not straight-forward how to generalize the definitions of $D_{\gamma'(t)} Y$ and $D_X Y$ where Y is a vector field on a manifold. The vectors $Y(\gamma(t+\delta))$ and $Y(\gamma(t))$ are at different based points, so one cannot make sense of $Y(\gamma(t+\delta)) - Y(\gamma(t))$.

One notion of differentiating a vector field by another one is called the *Lie derivatives*. The key idea is to push-forward tangent vectors in a natural way so that they become vectors at the same based point. Then, it makes sense to consider subtraction of vectors and also derivatives.

To begin with, we denote the integral curves using a map. First fix a vector field X on a manifold M. Then, given any point $p \in M$, as discussed before, one can find an integral curve $\gamma(t)$ so that $\gamma(0) = p$ and $\gamma'(t) = X(\gamma(t))$. We denote this curve $\gamma(t)$ by $\Phi_t(p)$, indicating that it depends on p. Now, for any fixed t, we can view $\Phi_t : M \to M$ as a map. One nice way to interpret this map is to regard $\Phi_t(p)$ as the point on M reached by flowing p along the vector field X for t unit time. As such, this map Φ_t is often called the *flow map*.

Exercise 3.12. Consider the unit sphere S^2 parametrized by spherical coordinates. (θ, φ), and the vector field $X = \dfrac{\partial}{\partial \theta}$. Describe the flow map Φ_t for this vector field, i.e. state how Φ_t maps the point with coordinates (θ, φ).

There are many meaningful purposes of this interpretation of integral curves. Standard theory of ODE shows Φ_t is smooth as long as X is a smooth vector field. Moreover, given $s, t \in \mathbb{R}$ and $p \in M$, we can regard $\Phi_s(\Phi_t(p))$ as the point obtained by flowing p along X first for t unit time, then for s unit time. Naturally, one would expect that the point obtained is exactly $\Phi_{s+t}(p)$. It is indeed true provided that X is independent of t.

Proposition 3.19. *Given any smooth vector field X on a smooth manifold M, and denote its flow map by $\Phi_t : M \to M$. Then, given any $t, s \in \mathbb{R}$ and $p \in M$, we have:*

(3.5) $$\Phi_t(\Phi_s(p)) = \Phi_{t+s}(p), \qquad \text{or equivalently} \quad \Phi_t \circ \Phi_s = \Phi_{t+s}.$$

Consequently, for each fixed $t \in \mathbb{R}$, the flow map Φ_t is a diffeomorphism with inverse Φ_{-t}.

Proof. The proof is a direct consequence of the uniqueness theorem of ODE. Consider s as fixed and t as the variable, then $\Phi_t(\Phi_s(p))$ and $\Phi_{t+s}(p)$ can be regarded as curves on M. When $t = 0$, both curves pass through the point $\Phi_s(p)$. It remains to show that both curves satisfy the same ODE, then uniqueness theorem of ODE guarantee that the two curves must be the same. We leave the detail as an exercise for readers. □

Exercise 3.13. Complete the detail of the above proof that the curves $\{\Phi_t(\Phi_s(p))\}_{t\in\mathbb{R}}$ and $\{\Phi_{t+s}(p)\}_{t\in\mathbb{R}}$ both satisfy the same ODE.

Now we are ready to introduce Lie derivatives of vector fields. Given two vector fields X and Y, we want to develop a notion of differentiating Y along X, i.e. the rate of change of Y when moving along integral curves of X. Denote the flow map of X by Φ_t. Fix a point $p \in M$, we want to compare $Y_{\Phi_t(p)}$ with Y_p. However, they are at different base points, so we push-forward $Y_{\Phi_t(p)}$ so that it becomes a vector based at p. To do so, the natural way is to push it forward by the map Φ_{-t} as it maps tangent vectors at $\Phi_t(p)$ to tangent vectors at $\Phi_{-t}(\Phi_t(p)) = p$.

Definition 3.20 (Lie Derivatives of Vector Fields)**.** Let X and Y be smooth vector fields on a manifold M. We define the *Lie derivative* of X along Y by:

$$(\mathcal{L}_X Y)_p := \left.\frac{d}{dt}\right|_{t=0} (\Phi_{-t})_* \left(Y_{\Phi_t(p)}\right)$$

where Φ_t denotes the flow map of X.

It sounds like a very technical definition that is very difficult to compute! Fortunately, we will prove that $\mathcal{L}_X Y$ is simply the commutator $[X, Y]$, to be defined below. First recall that a vector field on a manifold is a differential operator acting on scalar functions f. After differentiating f by a vector field Y, we get another scalar function $Y(f)$. Then, we can differentiate $Y(f)$ by another vector field X and obtaining $X(Y(f))$ (which for simplicity we denote it by XYf. The commutator, or the Lie brackets, measure the difference between XYf and YXf:

Definition 3.21 (Lie Brackets)**.** Given two vector fields X and Y on a manifold M, we define the *Lie brackets* $[X, Y]$ to be the vector field such that for any smooth function $f : M \to \mathbb{R}$, we have:

$$[X, Y]f := XYf - YXf.$$

Remark 3.22. Suppose under local coordinates $(u_1, \ldots, u_n)$, the vector fields X and Y can be written as:

$$X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial u_i} \qquad\qquad Y = \sum_{i=1}^{n} Y^i \frac{\partial}{\partial u_i},$$

then $[X, Y]$ has following the local expression:

$$[X, Y] = \sum_{i,j=1}^{n} \left(X^i \frac{\partial Y^j}{\partial u_i} - Y^i \frac{\partial X^j}{\partial u_i} \right) \frac{\partial}{\partial u_j} \tag{3.6}$$

□

Exercise 3.14. Verify (3.6), i.e. show that for any smooth function $f : M \to \mathbb{R}$, we have:

$$XYf - YXf = \sum_{i,j=1}^{n} \left(X^i \frac{\partial Y^j}{\partial u_i} - Y^i \frac{\partial X^j}{\partial u_i} \right) \frac{\partial f}{\partial u_j}$$

Exercise 3.15. Let $(u_1, \ldots, u_n)$ be a local coordinate of a manifold M, and define

$$X = \frac{\partial}{\partial u_i} \qquad\qquad Y = \frac{\partial}{\partial u_j}$$

Then, what is $[X, Y]$?

Exercise 3.16. Let X, Y, Z be vector fields on a manifold M, and $\varphi : M \to \mathbb{R}$ be a smooth scalar function. Show that:

(i) $[X + Y, Z] = [X, Z] + [Y, Z]$
(ii) $[X, \varphi Y] = (X\varphi)Y + \varphi[X, Y]$

It appears that $[X, Y]$ is more like an algebraic operation whereas Lie derivative $\mathcal{L}_X Y$ is a differential operation. Amazingly, they are indeed equal!

Proposition 3.23. *Let X and Y be smooth vector fields on a manifold M. Then, we have:*
$$\mathcal{L}_X Y = [X, Y].$$

Proof. Denote Φ_t to be the flow map of the vector field X. Fix a point $p \in M$ and let $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \to M$ be a local parametrization covering p. In order to compute $\mathcal{L}_X Y$ at p, we may assume that t is sufficiently small so that $\Phi_t(p)$ is also covered by F. Denote that coordinate representation of Φ_t by:
$$\mathsf{F}^{-1} \circ \Phi_t \circ \mathsf{F}(u_1, \ldots, u_n) = (v_t^1(u_1, \ldots, u_n), \ldots, v_t^n(u_1, \ldots, u_n)).$$
In local coordinates, the flow map Φ_t is then related to X under the relation:
$$\underbrace{\sum_{i=1}^n \frac{\partial v_t^i}{\partial t} \frac{\partial}{\partial u^i}\bigg|_{\Phi_t(p)}}_{\frac{\partial \Phi_t}{\partial t}\big|_p} = \underbrace{\sum_{i=1}^n X^i(\Phi_t(p)) \frac{\partial}{\partial u^i}\bigg|_{\Phi_t(p)}}_{X_{\Phi_t(p)}}.$$
Equating the coefficients, we have
$$\frac{\partial v_t^i}{\partial t} = X^i(\Phi_t(p)) \tag{3.7}$$
for $i = 1, \ldots, n$. Recall that the Lie derivative $(\mathcal{L}_X Y)_p$ is the time derivative at $t = 0$ of $(\Phi_{-t})_*\big(Y(\Phi_t(p))\big)$, which is given by:
$$\begin{aligned}(\Phi_{-t})_*\big(Y_{\Phi_t(p)}\big) &= (\Phi_t)_*^{-1}\left(\sum_{i=1}^n Y^i(\Phi_t(p)) \frac{\partial}{\partial u_i}(\Phi_t(p))\right) \\ &= \sum_{i=1}^n Y^i(\Phi_t(p)) \cdot (\Phi_t)_*^{-1}\left(\frac{\partial}{\partial u_i}(\Phi_t(p))\right).\end{aligned}$$
It then follows that:

(3.8)
$$\begin{aligned}(\mathcal{L}_X Y)_p &= \frac{\partial}{\partial t}\bigg|_{t=0} (\Phi_{-t})_*\big(Y_{\Phi_t(p)}\big) \\ &= \sum_{i=1}^n \frac{\partial}{\partial t}\bigg|_{t=0} Y^i(\Phi_t(p)) \cdot (\Phi_t)_*^{-1}\left(\frac{\partial}{\partial u_i}(\Phi_t(p))\right)\bigg|_{t=0} \\ &\quad + \sum_{i=1}^n Y^i(\Phi_t(p))\bigg|_{t=0} \cdot \frac{\partial}{\partial t}\bigg|_{t=0} (\Phi_t)_*^{-1}\left(\frac{\partial}{\partial u_i}(\Phi_t(p))\right) \\ &= \sum_{i=1}^n \frac{\partial}{\partial t}\bigg|_{t=0} Y^i(\Phi_t(p)) \cdot \frac{\partial}{\partial u_i}(p) + \sum_{i=1}^n Y^i(p) \cdot \frac{\partial}{\partial t}\bigg|_{t=0} (\Phi_t)_*^{-1}\left(\frac{\partial}{\partial u_i}(\Phi_t(p))\right).\end{aligned}$$
To compute $\frac{\partial}{\partial t} Y^i(\Phi_t(p))$, we use the chain rule:
$$\begin{aligned}\frac{\partial}{\partial t} Y^i(\Phi_t(p)) &= \sum_{j=1}^n \frac{\partial Y^i}{\partial u_j} \frac{\partial v_t^j}{\partial t} \\ &= \sum_{j=1}^n \frac{\partial Y^i}{\partial u_j} X^j(\Phi_t(p)) \qquad \text{(from (3.7))}.\end{aligned} \tag{3.9}$$

The term $\frac{\partial}{\partial t}\big|_{t=0}(\Phi_t)_*^{-1}\left(\frac{\partial}{\partial u_i}(\Phi_t(p))\right)$ is a bit more tricky. We first define Z_i^k by the coefficients of:

$$\sum_{k=1}^n Z_i^k \frac{\partial}{\partial u_k}(p) = (\Phi_t)_*^{-1}\left(\frac{\partial}{\partial u_i}(\Phi_t(p))\right) \quad \Longrightarrow \quad \sum_{k=1}^n Z_i^k \cdot (\Phi_t)_*\left(\frac{\partial}{\partial u_k}\bigg|_p\right) = \frac{\partial}{\partial u_i}\bigg|_{\Phi_t(p)}.$$

Recall that the coordinate representation of Φ_t is given by v_t^k's, so we get:

$$\sum_{j,k=1}^n Z_i^k \frac{\partial v_t^j}{\partial u_k}\frac{\partial}{\partial u_j}\bigg|_{\Phi_t(p)} = \frac{\partial}{\partial u_i}\bigg|_{\Phi_t(p)}.$$

By the linear independence of coordinate vectors, the coefficients Z_i^k satisfy:

$$\sum_{k=1}^n Z_i^k \frac{\partial v_t^j}{\partial u_k} = \delta_{ij}.$$

Differentiate both sides with respect to t, we get:

$$\sum_{k=1}^n \left(\frac{\partial Z_i^k}{\partial t}\frac{\partial v_t^j}{\partial u_k} + Z_i^k \frac{\partial X^j}{\partial u_k}\right)\bigg|_{t=0} = 0$$

where we have used (3.7). At $t = 0$, we have $v_t^j = u_j$, hence

$$\frac{\partial v_t^j}{\partial u_k}\bigg|_{t=0} = \delta_{jk} \quad \Longrightarrow \quad \underbrace{Z_i^j(p) = \delta_{ij}}_{\text{from definition}} \quad \text{and} \quad \sum_{k=1}^n \left(\frac{\partial Z_i^k}{\partial t}\delta_{jk} + \delta_{ik}\frac{\partial X^j}{\partial u_k}\right)\bigg|_{t=0} = 0.$$

It implies that:

$$\frac{\partial Z_i^j}{\partial t}\bigg|_{t=0} = -\frac{\partial X^j}{\partial u_i}.$$

Then, we can compute that:

$$\begin{aligned} \frac{\partial}{\partial t}\bigg|_{t=0}(\Phi_t)_*^{-1}\left(\frac{\partial}{\partial u_i}(\Phi_t(p))\right) &= \frac{\partial}{\partial t}\bigg|_{t=0}\sum_{k=1}^n Z_i^k \frac{\partial}{\partial u_k}(p) \\ &= -\sum_{k=1}^n \frac{\partial X^k}{\partial u_i}\frac{\partial}{\partial u_k}(p). \end{aligned} \tag{3.10}$$

Finally, substitute (3.9) and (3.10) back into (3.8), we get:

$$(\mathcal{L}_X Y)_p = \sum_{i=1}^n\sum_{j=1}^n \frac{\partial Y^i}{\partial u_j}X^j(p)\frac{\partial}{\partial u_i}(p) - \sum_{i=1}^n\sum_{k=1}^n Y^i(p)\frac{\partial X^k}{\partial u_i}\frac{\partial}{\partial u_k}(p),$$

which is exactly $[X, Y]$ at p according to (3.6). □

Now we know the geometric meaning of two commuting vector fields X and Y, i.e. $[X, Y] = 0$. According to Proposition 3.23, it is equivalent to saying $\mathcal{L}_X Y = 0$ at any $p \in M$. Then by the definition of Lie derivatives, we can conclude that:

$$(\Phi_{-t})_*(Y_{\Phi_t(p)}) = (\Phi_{-0})_*(Y_{\Phi_0(p)}) = Y_p \quad \text{for any } t.$$

In other words, we have $Y_{\Phi_t(p)} = (\Phi_t)_*(Y_p)$ for any t, meaning that pushing Y at p forward by the flow map Φ_t of X will yield the vector field Y at the point $\Phi_t(p)$. This result can further extends to show the flow maps of X and Y commute:

Exercise 3.17. Let X and Y be two vector fields on M such that $[X, Y] = 0$. Denote Φ_t and Ψ_t be the flow maps of X and Y respectively, show that for any $s, t \in \mathbb{R}$, we have:

$$\Phi_s \circ \Psi_t = \Psi_t \circ \Phi_s.$$

Sketch a diagram to illustrate its geometric meaning.

Lie derivatives on 1-forms can be defined similarly as on vector fields, expect that we uses *pull-backs* instead of *push-forwards* this time.

Definition 3.24 (Lie Derivatives of Differential 1-Forms)**.** Let X and a smooth vector field and α be a smooth 1-form on a manifold M. Denote the flow map of X by Φ_t, then we define the *Lie derivative* of α at $p \in M$ along X by:

$$(\mathcal{L}_X \alpha)_p := \left.\frac{d}{dt}\right|_{t=0} (\Phi_t)^* \alpha_{\Phi_t(p)}.$$

One can compute similarly as in Proposition 3.23 that the Lie derivative of a 1-form $\alpha = \sum_{i=1}^{n} \alpha_i \, du^i$ can be locally expressed as:

$$(\mathcal{L}_X \alpha)_p = \sum_{i,j=1}^{n} \left(X^i \frac{\partial \alpha_j}{\partial u_i} + \alpha_i \frac{\partial X^i}{\partial u_j} \right) du^j \tag{3.11}$$

Exercise 3.18. Verify (3.11).

Exercise 3.19. Let X and Y be two vector fields on M, and α be a 1-form on M. Show that:

$$X(\alpha(Y)) = (\mathcal{L}_X \alpha)(Y) + \alpha(\mathcal{L}_X Y).$$

3.3. Tensor Products

In Differential Geometry, tensor products are often used to produce bilinear, or in general multilinear, maps between tangent and cotangent spaces. The first and second fundamental forms of a regular surface, the Riemann curvature, etc. can all be expressed using tensor notations.

3.3.1. Tensor Products in Vector Spaces. Given two vector spaces V and W, their dual spaces V^* and W^* are vector spaces of all linear functionals $T : V \to \mathbb{R}$ and $S : W \to \mathbb{R}$ respectively. Pick two linear functionals $T \in V^*$ and $S \in W^*$, their tensor product $T \otimes S$ is a map from $V \times W$ to $\mathbb{R}$ defined by:

$$\begin{aligned} T \otimes S &: V \times W \to \mathbb{R} \\ (T \otimes S)(X, Y) &:= T(X)\, S(Y) \end{aligned}$$

It is easy to verify that $T \otimes S$ is bilinear, meaning that it is linear at each slot:

$$\begin{aligned} &(T \otimes S)\,(a_1X_1 + a_2X_2,\, b_1Y_1 + b_2Y_2) \\ &= a_1b_1(T \otimes S)(X_1, Y_1) + a_2b_1(T \otimes S)(X_2, Y_1) \\ &\quad + a_1b_2(T \otimes S)(X_1, Y_2) + a_2b_2(T \otimes S)(X_1, Y_2) \end{aligned}$$

Given three vector spaces U, V, W, and linear functionals $T_U \in U^*$, $T_V \in V^*$ and $T_W \in W^*$, one can define a triple tensor product $T_U \otimes (T_V \otimes T_W)$ by:

$$\begin{aligned} T_U \otimes (T_V \otimes T_W) &: U \times (V \times W) \to \mathbb{R} \\ (T_U \otimes (T_V \otimes T_W))(X, Y, Z) &:= T_U(X)\,(T_V \otimes T_W)(Y, Z) \\ &= T_U(X)\, T_V(Y)\, T_W(Z) \end{aligned}$$

One check easily that $(T_U \otimes T_V) \otimes T_W = T_U \otimes (T_V \otimes T_W)$. Since there is no ambiguity, we may simply write $T_U \otimes T_V \otimes T_W$. Inductively, given finitely many vector spaces $V_1, \ldots, V_k$, and linear functions $T_i \in V_i^*$, we can define the tensor product $T_1 \otimes \cdots \otimes T_k$ as a k-linear map by:

$$\begin{aligned} T_1 \otimes \cdots \otimes T_k &: V_1 \times \cdots \times V_k \to \mathbb{R} \\ (T_1 \otimes \cdots \otimes T_k)(X_1, \ldots, X_k) &:= T_1(X_1) \cdots T_k(X_k) \end{aligned}$$

Given two tensor products $T_1 \otimes S_1 : V \times W \to \mathbb{R}$ and $T_2 \otimes S_2 : V \times W \to \mathbb{R}$, one can form a linear combination of them:

$$\begin{aligned} \alpha_1(T_1 \otimes S_1) + \alpha_2(T_2 \otimes S_2) &: V \times W \to \mathbb{R} \\ (\alpha_1(T_1 \otimes S_1) + \alpha_2(T_2 \otimes S_2))(X, Y) &:= \alpha_1(T_1 \otimes S_1)(X, Y) + \alpha_2(T_2 \otimes S_2)(X, Y) \end{aligned}$$

The tensor products $T \otimes S$ with $T \in V^*$ and $S \in W^*$ generate a vector space. We denote this vector space by:

$$V^* \otimes W^* := \operatorname{span}\{T \otimes S : T \in V^* \text{ and } S \in W^*\}.$$

Exercise 3.20. Verify that $\alpha\,(T \otimes S) = (\alpha T) \otimes S = T \otimes (\alpha S)$. Therefore, we can simply write $\alpha T \otimes S$.

Exercise 3.21. Show that the tensor product is bilinear in a sense that:

$$T \otimes (\alpha_1 S_1 + \alpha_2 S_2) = \alpha_1 T \otimes S_1 + \alpha_2 T \otimes S_2$$

and similar for the T slot.

Let's take the dual basis as an example to showcase the use of tensor products. Consider a vector space V with a basis $\{\mathsf{e}_i\}_{i=1}^n$. Let $\{\mathsf{e}_i^*\}_{i=1}^n$ be its dual basis for V^*. Then, one can check that:

$$\begin{aligned}(\mathsf{e}_i^* \otimes \mathsf{e}_j^*)(\mathsf{e}_k, \mathsf{e}_l) &= \mathsf{e}_i^*(\mathsf{e}_k)\, \mathsf{e}_k^*(\mathsf{e}_l) \\ &= \delta_{ik}\,\delta_{jl} \\ &= \begin{cases} 1 & \text{if } i = k \text{ and } j = l \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

Generally, the sum $\sum_{i,j=1}^n A_{ij}\mathsf{e}_i^* \otimes \mathsf{e}_j^*$ will act on vectors in V by:

$$\begin{aligned}&\left(\sum_{i,j=1}^n A_{ij}\mathsf{e}_i^* \otimes \mathsf{e}_j^*\right)\left(\sum_{k=1}^n \alpha_k \mathsf{e}_k, \sum_{l=1}^n \beta_l \mathsf{e}_l\right) \\ &= \sum_{i,j,k,l=1}^n A_{ij}\alpha_k\beta_l(\mathsf{e}_i^* \otimes \mathsf{e}_j^*)(\mathsf{e}_k, \mathsf{e}_l) = \sum_{i,j,k,l=1}^n A_{ij}\alpha_k\beta_l\delta_{ik}\delta_{jl} = \sum_{k,l=1}^n A_{kl}\alpha_k\beta_l\end{aligned}$$

In other words, the sum of tensor products $\sum_{i,j=1}^n A_{ij}\mathsf{e}_i^* \otimes \mathsf{e}_j^*$ is the inner product on V represented by the matrix $[A_{kl}]$ with respect to the basis $\{\mathsf{e}_i\}_{i=1}^n$ of V. For example, when $A_{kl} = \delta_{kl}$, then $\sum_{i,j=1}^n A_{ij}\mathsf{e}_i^* \otimes \mathsf{e}_j^* = \sum_{i=1}^n \mathsf{e}_i^* \otimes \mathsf{e}_i^*$. It is the usual dot product on V.

Exercise 3.22. Show that $\{\mathsf{e}_i^* \otimes \mathsf{e}_j^*\}_{i,j=1}^n$ is a basis for $V^* \otimes V^*$. What is the dimension of $V^* \otimes V^*$?

Exercise 3.23. Suppose $\dim V = 2$. Let $\omega \in V^* \otimes V^*$ satisfy:

$$\begin{aligned}\omega(\mathsf{e}_1, \mathsf{e}_1) &= 0 \qquad & \omega(\mathsf{e}_1, \mathsf{e}_2) &= 3 \\ \omega(\mathsf{e}_2, \mathsf{e}_1) &= -3 \qquad & \omega(\mathsf{e}_2, \mathsf{e}_2) &= 0\end{aligned}$$

Express ω in terms of e_i^*'s.

To describe linear or multilinear map between two vector spaces V and W (where W is not necessarily the one-dimensional space $\mathbb{R}$), one can also use tensor products. Given a linear functional $f \in V^*$ and a vector $w \in W$, we can form a tensor $f \otimes w$, which is regarded as a linear map $f \otimes w : V \to W$ defined by:

$$(f \otimes w)(v) := f(v)w.$$

Let $\{\mathsf{e}_i\}$ be a basis for V, and $\{\mathsf{f}_j\}$ be a basis for W. Any linear map $T : V \to W$ can be expressed in terms of these bases. Suppose:

$$T(\mathsf{e}_i) = \sum_j A_i^j \mathsf{f}_j.$$

Then, we claim that T can be expressed using the following tensor notations:

$$T = \sum_{i,j} A_i^j \mathsf{e}_i^* \otimes \mathsf{f}_j$$

Let's verify this. Note that a linear map is determined by its action on the basis $\{\mathsf{e}_i\}$ for V. It suffices to show:

$$\left(\sum_{i,j} A_i^j \mathsf{e}_i^* \otimes \mathsf{f}_j\right)(\mathsf{e}_k) = T(\mathsf{e}_k).$$

Using the fact that:

$$(\mathsf{e}_i^* \otimes \mathsf{f}_j)(\mathsf{e}_k) = \mathsf{e}_i^*(\mathsf{e}_k)\mathsf{f}_j = \delta_{ik}\mathsf{f}_j,$$

one can compute:

$$\begin{aligned}\left(\sum_{i,j} A_i^j \mathsf{e}_i^* \otimes \mathsf{f}_j\right)(\mathsf{e}_k) &= \sum_{i,j} A_i^j (\mathsf{e}_i^* \otimes \mathsf{f}_j)(\mathsf{e}_k)\\ &= \sum_{i,j} A_i^j \delta_{ik}\mathsf{f}_j = \sum_j A_k^j \mathsf{f}_j = T(\mathsf{e}_k)\end{aligned}$$

as desired.

Generally, if $T_1, \ldots, T_k \in V^*$ and $X \in V$, then

$$T_1 \otimes \cdots \otimes T_k \otimes X$$

is regarded to be a k-linear map from $V \times \ldots \times V$ to V, defined by:

$$T_1 \otimes \cdots \otimes T_k \otimes X : \underbrace{V \times \ldots \times V}_{k} \to V$$
$$(T_1 \otimes \cdots \otimes T_k \otimes X)(Y_1, \ldots, Y_k) := T_1(Y_1) \cdots T_k(Y_k)\, X$$

Example 3.25. One can write the cross-product in $\mathbb{R}^3$ using tensor notations. Think of the cross product as a bilinear map $\omega : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ that takes two input vectors u and v, and outputs the vector $\mathsf{u} \times \mathsf{v}$. Let $\{\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3\}$ be the standard basis in $\mathbb{R}^3$ (i.e. $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$). Then one can write:

$$\begin{aligned}\omega = \;& \mathsf{e}_1^* \otimes \mathsf{e}_2^* \otimes \mathsf{e}_3 - \mathsf{e}_2^* \otimes \mathsf{e}_1^* \otimes \mathsf{e}_3\\ &+ \mathsf{e}_2^* \otimes \mathsf{e}_3^* \otimes \mathsf{e}_1 - \mathsf{e}_3^* \otimes \mathsf{e}_2^* \otimes \mathsf{e}_1\\ &+ \mathsf{e}_3^* \otimes \mathsf{e}_1^* \otimes \mathsf{e}_2 - \mathsf{e}_1^* \otimes \mathsf{e}_3^* \otimes \mathsf{e}_2\end{aligned}$$

One can check that, for instance, $\omega(\mathsf{e}_1, \mathsf{e}_2) = \mathsf{e}_3$, which is exactly $\mathsf{e}_1 \times \mathsf{e}_2 = \mathsf{e}_3$. □

3.3.2. Tensor Products on Smooth Manifolds.

In the previous subsection we take tensor products on a general abstract vector space V. In this course, we will mostly deal with the case when V is the tangent or cotangent space of a smooth manifold M.

Recall that if $\mathsf{F}(u_1, \ldots, u_n)$ is a local parametrization of M, then there is a local coordinate basis $\left\{\frac{\partial}{\partial u_i}(p)\right\}_{j=1}^n$ for the tangent space T_pM at every $p \in M$ covered by F. The cotangent space T_p^*M has a dual basis $\left\{du^j\big|_p\right\}_{j=1}^n$ defined by $du_j\left(\frac{\partial}{\partial u_i}\right) = \delta_{ij}$ at every $p \in M$.

Then, one can take tensor products of du^i's and $\frac{\partial}{\partial u_i}$'s to express multilinear maps between tangent and cotangent spaces. For instance, the tensor product $g = \sum_{i,j=1}^n g_{ij} du^i \otimes du^j$, where g_{ij}'s are scalar functions, means that it is a bilinear map at each point $p \in M$ such that:

$$g(X, Y) = \sum_{i,j=1}^n g_{ij}(du^i \otimes du^j)(X, Y) = \sum_{i,j=1}^n g_{ij} du^i(X)\, du^j(Y)$$

for any vector field $X, Y \in TM$. In particular, we have:

$$g\left(\frac{\partial}{\partial u_i}, \frac{\partial}{\partial u_j}\right) = g_{ij}.$$

We can also express multilinear maps from $T_pM \times T_pM \times T_pM$ to T_pM. For instance, we let:

$$\mathrm{Rm} = \sum_{i,j,k,l=1}^{n} R_{ijk}^{l} du^i \otimes du^j \otimes du^k \otimes \frac{\partial}{\partial u_l}.$$

Then, Rm is a mutlilinear map at each $p \in M$ such that:

$$\mathrm{Rm}(X,Y,Z) = \sum_{i,j,k,l=1}^{n} R_{ijk}^{l} du^i(X)\, du^j(Y)\, du^k(Z)\, \frac{\partial}{\partial u_l}.$$

It is a trilinear map such that:

$$\mathrm{Rm}\left(\frac{\partial}{\partial u_i}, \frac{\partial}{\partial u_j}, \frac{\partial}{\partial u_k}\right) = \sum_{l=1}^{n} R_{ijk}^{l} \frac{\partial}{\partial u_l}.$$

We call g a $(2,0)$-tensor (meaning that it maps two vectors to a scalar), and Rm a $(3,1)$-tensor (meaning that it maps three vectors to one vector). In general, we can also define $(k,0)$-tensor ω on M which has the general form:

$$\omega_p = \sum_{i_1,\ldots,i_k=1}^{n} \omega_{i_1 i_2 \cdots i_k}(p)\, du^{i_1}\Big|_p \otimes \cdots \otimes du^{i_k}\Big|_p$$

Here $\omega_{i_1 i_2 \cdots i_k}$'s are scalar functions. This tensor maps the tangent vectors $\left(\frac{\partial}{\partial u_{i_1}}, \ldots, \frac{\partial}{\partial u_{i_k}}\right)$ to the scalar $\omega_{i_1 i_2 \ldots i_k}$ at the corresponding point.

Like the Rm-tensor, we can also generally define $(k,1)$-tensor Ω on M which has the general form:

$$\Omega_p = \sum_{i_1,\ldots,i_k,j=1}^{n} \Omega_{i_1 i_2 \cdots i_k}^{j}(p)\, du^{i_1}\Big|_p \otimes \cdots \otimes du^{i_k}\Big|_p \otimes \frac{\partial}{\partial u_j}(p)$$

where $\Omega_{i_1 i_2 \ldots i_k}^{j}$'s are scalar functions. This tensor maps the tangent vectors $\left(\frac{\partial}{\partial u_{i_1}}, \ldots, \frac{\partial}{\partial u_{i_k}}\right)$ to the tangent vector $\sum_j \Omega_{i_1 i_2 \ldots i_k}^{j} \frac{\partial}{\partial u_j}$ at the corresponding point.

Note that these g_{ij}, R_{ijk}^{l}, $\omega_{i_1 i_2 \cdots i_k}$ and $\Omega_{i_1 i_2 \ldots i_k}^{j}$ are scalar functions locally defined on the open set covered by the local parametrization F, so we can talk about whether they are smooth or not:

Definition 3.26 (Smooth Tensors on Manifolds). A *smooth* $(k,0)$-*tensor* ω on M is a k-linear map $\omega_p : \underbrace{T_pM \times \ldots \times T_pM}_{k} \to \mathbb{R}$ at each $p \in M$ such that under any local parametrization $\mathsf{F}(u_1,\ldots,u_n) : \mathcal{U} \to M$, it can be written in the form:

$$\omega_p = \sum_{i_1,\ldots,i_k=1}^{n} \omega_{i_1 i_2 \cdots i_k}(p)\, du^{i_1}\Big|_p \otimes \cdots \otimes du^{i_k}\Big|_p$$

where $\omega_{i_1 i_2 \ldots i_k}$'s are smooth scalar functions locally defined on $\mathsf{F}(\mathcal{U})$.

A *smooth* $(k,1)$-*tensor* Ω on M is a k-linear map $\Omega_p : \underbrace{T_pM \times \ldots \times T_pM}_{k} \to T_pM$ at each $p \in M$ such that under any local parametrization $\mathsf{F}(u_1,\ldots,u_n) : \mathcal{U} \to M$, it can be written in the form:

$$\Omega_p = \sum_{i_1,\ldots,i_k,j=1}^{n} \Omega^j_{i_1 i_2 \cdots i_k}(p)\, du^{i_1}\Big|_p \otimes \cdots \otimes du^{i_k}\Big|_p \otimes \frac{\partial}{\partial u_j}(p)$$

where $\Omega^j_{i_1 i_2 \ldots i_k}$'s are smooth scalar functions locally defined on $\mathsf{F}(\mathcal{U})$.

Remark 3.27. Since T_pM is finite dimensional, from Linear Algebra we know $(T_pM)^{**}$ is isomorphic to T_pM. Therefore, a tangent vector $\frac{\partial}{\partial u_i}(p)$ can be regarded as a linear functional on cotangent vectors in T_p^*M, meaning that:

$$\frac{\partial}{\partial u_i}\bigg|_p \left(du^j\Big|_p \right) = \delta_{ij}.$$

Under this interpretation, one can also view a $(k,1)$-tensor Ω as a $(k+1)$-linear map $\Omega_p : \underbrace{T_pM \times \ldots \times T_pM}_{k} \times T_p^*M \to \mathbb{R}$, which maps $\left(du^{i_1},\ldots,du^{i_k},\frac{\partial}{\partial u_j}\right)$ to $\Omega^j_{i_1 i_2 \ldots i_k}$. However, we will not view a $(k,1)$-tensor this way in this course.

Generally, we can also talk about (k,s)-tensors, which is a $(k+s)$-linear map $\Omega_p : \underbrace{T_pM \times \ldots \times T_pM}_{k} \times \underbrace{T_p^*M \times \ldots \times T_p^*M}_{s} \to \mathbb{R}$ taking $\left(du^{i_1},\ldots,du^{i_k},\frac{\partial}{\partial u_{j_1}},\ldots,\frac{\partial}{\partial u_{j_s}}\right)$ to a scalar. However, we seldom deal with these tensors in this course. □

Exercise 3.24. Let M be a smooth manifold with local coordinates (u_1,u_2). Consider the tensor products:

$$T_1 = du^1 \otimes du^2 \quad \text{and} \quad T_2 = du^1 \otimes \frac{\partial}{\partial u_2}.$$

Which of the following is well-defined?

(a) $T_1\left(\frac{\partial}{\partial u_1}\right)$

(b) $T_2\left(\frac{\partial}{\partial u_1}\right)$

(c) $T_1\left(\frac{\partial}{\partial u_1},\frac{\partial}{\partial u_2}\right)$

(d) $T_2\left(\frac{\partial}{\partial u_1},\frac{\partial}{\partial u_2}\right)$

Exercise 3.25. let M be a smooth manifold with local coordinates (u_1, u_2). The linear map $T : T_pM \to T_pM$ satisfies:

$$T\left(\frac{\partial}{\partial u_1}\right) = \frac{\partial}{\partial u_1} + \frac{\partial}{\partial u_2}$$
$$T\left(\frac{\partial}{\partial u_2}\right) = \frac{\partial}{\partial u_1} - \frac{\partial}{\partial u_2}.$$

Express T using tensor products.

One advantage of using tensor notations, instead of using matrices, to denote a multilinear map between tangent or cotangent spaces is that one can figure out the conversion rule between local coordinate systems easily (when compared to using matrices)

Example 3.28. Consider the extended complex plane $M := \mathbb{C} \cup \{\infty\}$ defined in Example 2.12. We cover M by two local parametrizations:

$$\begin{aligned} \mathsf{F}_1 : \mathbb{R}^2 &\to \mathbb{C} \subset M & \qquad \mathsf{F}_2 : \mathbb{R}^2 &\to (\mathbb{C}\backslash\{0\}) \cup \{\infty\} \subset M \\ (x, y) &\mapsto x + yi & (u, v) &\mapsto \frac{1}{u + vi} \end{aligned}$$

The transition maps on the overlap are given by:

$$(u, v) = \mathsf{F}_2^{-1} \circ \mathsf{F}_1(x, y) = \left(\frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2}\right)$$
$$(x, y) = \mathsf{F}_1^{-1} \circ \mathsf{F}_2(u, v) = \left(\frac{u}{u^2 + v^2}, -\frac{v}{u^2 + v^2}\right)$$

Consider the $(2, 0)$-tensor ω defined using local coordinates (x, y) by:

$$\omega = e^{-(x^2+y^2)}\, dx \otimes dy.$$

Using the chain rule, we can express dx and dy in terms of du and dv:

$$\begin{aligned} dx &= d\left(\frac{u}{u^2 + v^2}\right) = \frac{(u^2 + v^2)\, du - u(2u\, du + 2v\, dv)}{(u^2 + v^2)^2} \\ &= \frac{v^2 - u^2}{(u^2 + v^2)^2} du - \frac{2uv}{(u^2 + v^2)^2} dv \\ dy &= -d\left(\frac{v}{u^2 + v^2}\right) = -\frac{(u^2 + v^2) dv - v(2u\, du + 2v\, dv)}{(u^2 + v^2)^2} \\ &= -\frac{2uv}{(u^2 + v^2)^2} du + \frac{v^2 - u^2}{(u^2 + v^2)^2} dv \end{aligned}$$

Therefore, we get:

$$\begin{aligned} dx \otimes dy = &\frac{2uv(u^2 - v^2)}{(u^2 + v^2)^4} du \otimes du + \frac{(u^2 - v^2)^2}{(u^2 + v^2)^4} du \otimes dv \\ &+ \frac{4u^2v^2}{(u^2 + v^2)^4} dv \otimes du + \frac{2uv(u^2 - v^2)}{(u^2 + v^2)^4} dv \otimes dv \end{aligned}$$

Recall that $\omega = e^{-(x^2+y^2)}\, dx \otimes dy$, and in terms of (u, v), we have:

$$e^{-(x^2+y^2)} = e^{-\frac{1}{u^2+v^2}}.$$

Hence, in terms of (u,v), ω is expressed as:

$$\begin{aligned}\omega = e^{-\frac{1}{u^2+v^2}} \Big\{ &\frac{2uv(u^2-v^2)}{(u^2+v^2)^4} du \otimes du + \frac{(u^2-v^2)^2}{(u^2+v^2)^4} du \otimes dv \\ &+ \frac{4u^2v^2}{(u^2+v^2)^4} dv \otimes du + \frac{2uv(u^2-v^2)}{(u^2+v^2)^4} dv \otimes dv \Big\}\end{aligned}$$

□

Exercise 3.26. Consider the extended complex plane $\mathbb{C} \cup \{\infty\}$ as in Example 3.28, and the $(1,1)$-tensor of the form:

$$\Omega = e^{-(x^2+y^2)}\, dx \otimes \frac{\partial}{\partial y}.$$

Express Ω in terms of (u,v).

Generally, if $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ are two overlapping local coordinates on a smooth manifold M, then given a $(2,0)$-tensor:

$$g = \sum_{i,j} g_{ij} du^i \otimes du^j$$

written using the u_i's coordinates, one can convert it to v_α's coordinates by the chain rule:

$$\begin{aligned} g = \sum_{i,j} g_{ij} du^i \otimes du^j &= \sum_{i,j} g_{ij} \left(\sum_\alpha \frac{\partial u_i}{\partial v_\alpha} dv^\alpha \right) \otimes \left(\sum_\beta \frac{\partial u_j}{\partial v_\beta} dv^\beta \right) \\ &= \sum_{\alpha,\beta} \left(\sum_{i,j} g_{ij} \frac{\partial u_i}{\partial v_\alpha} \frac{\partial u_j}{\partial v_\beta} \right) dv^\alpha \otimes dv^\beta \end{aligned}$$

Exercise 3.27. Given that u_i's and v_α's are overlapping local coordinates of a smooth manifold M. Using these coordinates, one can express the following $(3,1)$-tensor in two ways:

$$\mathrm{Rm} = \sum_{i,j,k,l} R_{ijk}^l du^i \otimes du^j \otimes du^k \otimes \frac{\partial}{\partial u_l} = \sum_{\alpha,\beta,\gamma,\eta} \widetilde{R}_{\alpha\beta\gamma}^\eta dv^\alpha \otimes dv^\beta \otimes dv^\gamma \otimes \frac{\partial}{\partial v_\eta}$$

Express R_{ijk}^l in terms of $R_{\alpha\beta\gamma}^\eta$'s.

Exercise 3.28. Given that u_i's and v_α's are overlapping local coordinates of a smooth manifold M. Suppose g and h are two $(2,0)$-tensors expressed in terms of local coordinates as:

$$\begin{aligned} g &= \sum_{i,j} g_{ij}\, du^i \otimes du^j = \sum_{\alpha,\beta} \widetilde{g}_{\alpha\beta} dv^\alpha \otimes dv^\beta \\ h &= \sum_{i,j} h_{ij}\, du^i \otimes du^j = \sum_{\alpha,\beta} \widetilde{h}_{\alpha\beta} dv^\alpha \otimes dv^\beta. \end{aligned}$$

Let G be the matrix with g_{ij} as its (i,j)-th entry, and let g^{ij} be the (i,j)-th entry of G^{-1}. Similarly, define $\widetilde{g}^{\alpha\beta}$ to be the inverse of $\widetilde{g}_{\alpha\beta}$. Show that:

$$\sum_{i,j} g^{ik} h_{kj}\, du^i \otimes du^j = \sum_{\alpha,\beta} \widetilde{g}^{\alpha\gamma} \widetilde{h}_{\gamma\beta}\, dv^\alpha \otimes dv^\beta.$$

3.4. Wedge Products

Recall that in Multivariable Calculus, the cross product plays a crucial role in many aspects. It is a bilinear map which takes two vectors to one vectors, and so it is a $(2, 1)$-tensor on $\mathbb{R}^3$.

The determinant is another important quantity in Multivariable Calculus and Linear Algebra. Using tensor languages, an $n \times n$ determinant can be regarded as a n-linear map taking n vectors in $\mathbb{R}^n$ to a scalar. For instance, for the 2×2 case, one can view:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

as a bilinear map taking column vectors $(a, c)^T$ and $(b, d)^T$ in $\mathbb{R}^2$ to a number $ad - bc$. Therefore, it is a $(2, 0)$-tensor on $\mathbb{R}^2$; and generally for $n \times n$, the determinant is an $(n, 0)$-tensor on $\mathbb{R}^n$.

Both the cross product in $\mathbb{R}^3$ and determinant ($n \times n$ in general) are *alternating*, in a sense that interchanging any pair of inputs will give a negative sign for the output. For the cross product, we have $\mathsf{a} \times \mathsf{b} = -\mathsf{b} \times \mathsf{a}$; and for the determinant, switching any pair of columns will give a negative sign:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = - \begin{vmatrix} b & a \\ d & c \end{vmatrix}.$$

In the previous section we have seen how to express k-linear maps over tangent vectors using tensor notations. To deal with the above alternating tensors, it is more elegant and concise to use *alternating tensors*, or *wedge products* that we are going to learn in this section.

3.4.1. Wedge Product on Vector Spaces.

Let's start from the easiest case. Suppose V is a finite dimensional vector space and V^* is the dual space of V. Given any two elements $T, S \in V^*$, the tensor product $T \otimes S$ is a map given by:

$$(T \otimes S)(X, Y) = T(X)\, S(Y)$$

for any $X, Y \in V$. The *wedge product* $T \wedge S$, where $T, S \in V^*$, is a bilinear map defined by:

$$T \wedge S := T \otimes S - S \otimes T$$

meaning that for any $X, Y \in V$, we have:

$$\begin{aligned} (T \wedge S)(X, Y) &= (T \otimes S)(X, Y) - (S \otimes T)(X, Y) \\ &= T(X)\, S(Y) - S(X)\, T(Y) \end{aligned}$$

It is easy to note that $T \wedge S = -S \wedge T$.

Take the cross product in $\mathbb{R}^3$ as an example. Write the cross product as a bilinear map $\omega(\mathsf{a}, \mathsf{b}) := \mathsf{a} \times \mathsf{b}$. It is a $(2, 1)$-tensor on $\mathbb{R}^3$ which can be represented as:

$$\begin{aligned} \omega = {} & \mathsf{e}_1^* \otimes \mathsf{e}_2^* \otimes \mathsf{e}_3 - \mathsf{e}_2^* \otimes \mathsf{e}_1^* \otimes \mathsf{e}_3 \\ & + \mathsf{e}_2^* \otimes \mathsf{e}_3^* \otimes \mathsf{e}_1 - \mathsf{e}_3^* \otimes \mathsf{e}_2^* \otimes \mathsf{e}_1 \\ & + \mathsf{e}_3^* \otimes \mathsf{e}_1^* \otimes \mathsf{e}_2 - \mathsf{e}_1^* \otimes \mathsf{e}_3^* \otimes \mathsf{e}_2 \end{aligned}$$

Now using the wedge product notations, we can express ω as:

$$\omega = (\mathsf{e}_1^* \wedge \mathsf{e}_2^*) \otimes \mathsf{e}_3 + (\mathsf{e}_2^* \wedge \mathsf{e}_3^*) \otimes \mathsf{e}_1 + (\mathsf{e}_3^* \wedge \mathsf{e}_1^*) \otimes \mathsf{e}_2$$

which is a half shorter than using tensor products alone.

Now given three elements $T_1, T_2, T_3 \in V^*$, one can also form a triple wedge product $T_1 \wedge T_2 \wedge T_3$ which is a $(3,0)$-tensor so that switching any pair of T_i and T_j (with $i \neq j$) will give a negative sign. For instance:

$$T_1 \wedge T_2 \wedge T_3 = -T_2 \wedge T_1 \wedge T_3 \quad \text{and} \quad T_1 \wedge T_2 \wedge T_3 = -T_3 \wedge T_2 \wedge T_1.$$

It can be defined in a precise way as:

$$\begin{aligned} T_1 \wedge T_2 \wedge T_3 := {} & T_1 \otimes T_2 \otimes T_3 - T_1 \otimes T_3 \otimes T_2 \\ & + T_2 \otimes T_3 \otimes T_1 - T_2 \otimes T_1 \otimes T_3 \\ & + T_3 \otimes T_1 \otimes T_2 - T_3 \otimes T_2 \otimes T_1 \end{aligned}$$

Exercise 3.29. Verify that the above definition of triple wedge product will result in $T_1 \wedge T_2 \wedge T_3 = -T_3 \wedge T_2 \wedge T_1$.

Exercise 3.30. Propose the definition of $T_1 \wedge T_2 \wedge T_3 \wedge T_4$. Do this exercise before reading ahead.

We can also define $T_1 \wedge T_2 \wedge T_3$ in a more systematic (yet equivalent) way using symmetric groups. Let S_3 be the permutation group of $\{1,2,3\}$. An element $\sigma \in S_3$ is a bijective map $\sigma : \{1,2,3\} \to \{1,2,3\}$. For instance, a map satisfying $\sigma(1) = 2$, $\sigma(2) = 3$ and $\sigma(3) = 1$ is an example of an element in S_3. We can express this σ by:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{or simply:} \quad (123)$$

A map $\tau \in S_3$ given by $\tau(1) = 2$, $\tau(2) = 1$ and $\tau(3) = 3$ can be expressed as:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \text{or simply:} \quad (12)$$

This element, which switches two of the elements in $\{1,2,3\}$ and fixes the other one, is called a *transposition*.

Multiplication of two elements $\sigma_1, \sigma_2 \in S_3$ is defined by composition. Precisely, $\sigma_1\sigma_2$ is the composition $\sigma_1 \circ \sigma_2$. Note that this means the elements $\{1,2,3\}$ are input into σ_2 first, and then into σ_1. In general, $\sigma_1\sigma_2 \neq \sigma_2\sigma_1$. One can check easily that, for instance, we have:

$$\begin{aligned} (12)(23) &= (123) \\ (23)(12) &= (132) \end{aligned}$$

Elements in the permutation group S_n of n elements (usually denoted by $\{1,2,\ldots,n\}$) can be represented and mutliplied in a similar way.

Exercise 3.31. Convince yourself that in S_5, we have:

$$(12345)(31) = (32)(145) = (32)(15)(14)$$

The above exercise shows that we can decompose $(12345)(31)$ into a product of three transpositions (32), (15) and (14). In fact, any element in S_n can be decomposed this way. Here we state a standard theorem in elementary group theory:

Theorem 3.29. *Every element $\sigma \in S_n$ can be expressed as a product of transpositions: $\sigma = \tau_1\tau_2 \ldots \tau_r$. Such a decomposition is not unique. However, if $\sigma = \widetilde{\tau}_1\widetilde{\tau}_2 \ldots \widetilde{\tau}_k$ is another decomposition of σ into transpositions, then we have $(-1)^k = (-1)^r$.*

Proof. Consult any standard textbook on Abstract Algebra. □

In view of Theorem 3.29, given an element $\sigma \in S_n$ which can be decomposed into the product of r transpositions, we define:

$$\operatorname{sgn}(\sigma) := (-1)^r.$$

For instance, $\operatorname{sgn}(12345) = (-1)^3 = -1$, and $\operatorname{sgn}(123) = (-1)^2 = 1$. Certainly, if τ is a transposition, we have $\operatorname{sgn}(\sigma\tau) = -\operatorname{sgn}(\sigma)$.

Now we are ready to state an equivalent way to define triple wedge product using the above notations:

$$T_1 \wedge T_2 \wedge T_3 := \sum_{\sigma \in S_3} \operatorname{sgn}(\sigma) T_{\sigma(1)} \otimes T_{\sigma(2)} \otimes T_{\sigma(3)}.$$

We can verify that it gives the same expression as before:

$$\begin{aligned}
&\sum_{\sigma \in S_3} \operatorname{sgn}(\sigma) T_{\sigma(1)} \otimes T_{\sigma(2)} \otimes T_{\sigma(3)} \\
&= T_1 \otimes T_2 \otimes T_3 && \sigma = \mathrm{id} \\
&\quad - T_2 \otimes T_1 \otimes T_3 && \sigma = (12) \\
&\quad - T_3 \otimes T_2 \otimes T_1 && \sigma = (13) \\
&\quad - T_1 \otimes T_3 \otimes T_2 && \sigma = (23) \\
&\quad + T_2 \otimes T_3 \otimes T_1 && \sigma = (123) = (13)(12) \\
&\quad + T_3 \otimes T_1 \otimes T_2 && \sigma = (132) = (12)(13)
\end{aligned}$$

In general, we define:

Definition 3.30 (Wedge Product)**.** Let V be a finite dimensional vector space, and V^* be the dual space of V. Then, given any $T_1, \ldots, T_k \in V^*$, we define their *k-th wedge product* by:

$$T_1 \wedge \cdots \wedge T_k := \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, T_{\sigma(1)} \otimes \ldots \otimes T_{\sigma(k)}$$

where S_k is the permutation group of $\{1, \ldots, k\}$. The vector space spanned by $T_1 \wedge \cdots \wedge T_k$'s (where $T_1, \ldots, T_k \in V^*$) is denoted by $\wedge^k V^*$.

Remark 3.31. It is a convention to define $\wedge^0 V^* := \mathbb{R}$. □

If we switch any pair of the T_i's, then the wedge product differs by a minus sign. To show this, let $\tau \in S_k$ be a transposition, then for any $\sigma \in S_k$, we have $\operatorname{sgn}(\sigma \circ \tau) = -\operatorname{sgn}(\sigma)$. Therefore, we get:

$$\begin{aligned}
T_{\tau(1)} \wedge \cdots \wedge T_{\tau(k)} &= \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) T_{\sigma(\tau(1))} \otimes \ldots \otimes T_{\sigma(\tau(k))} \\
&= -\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma \circ \tau) T_{\sigma \circ \tau(1)} \otimes \ldots \otimes T_{\sigma \circ \tau(k)} \\
&= -\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma') T_{\sigma'(1)} \otimes \ldots \otimes T_{\sigma' \tau(k)} && (\text{where } \sigma' := \sigma \circ \tau) \\
&= -T_1 \wedge \cdots \wedge T_k.
\end{aligned}$$

The last step follows from the fact that $\sigma \mapsto \sigma \circ \tau$ is a bijection between S_k and itself.

Exercise 3.32. Write down $T_1 \wedge T_2 \wedge T_3 \wedge T_4$ explicitly in terms of tensor products (with no wedge and summation sign).

Exercise 3.33. Show that $\dim \wedge^k V^* = C_k^n$, when $n = \dim V$ and $0 \leq k \leq n$, by writing a basis for $\wedge^k V^*$. Show also that $\wedge^k V^* = \{0\}$ if $k > \dim V$.

Exercise 3.34. Let $\{\mathsf{e}_i\}_{i=1}^n$ be a basis for a vector space V, and $\{\mathsf{e}_i^*\}_{i=1}^n$ be the corresponding dual basis for V^*. Show that:

$$\left(\mathsf{e}_{i_1}^* \wedge \cdots \wedge \mathsf{e}_{i_k}^*\right)(\mathsf{e}_{j_1}, \ldots, \mathsf{e}_{j_k}) = \delta_{i_1 j_1} \cdots \delta_{i_k j_k}.$$

Remark 3.32. The vector space $\wedge^k V^*$ is *spanned* by $T_1 \wedge \cdots \wedge T_k$'s where $T_1, \ldots, T_k \in V^*$. Note that not all elements in V^* can be expressed in the form of $T_1 \wedge \cdots \wedge T_k$. For instance when $V = \mathbb{R}^4$ with standard basis $\{\mathsf{e}_i\}_{i=1}^4$, the element $\sigma = \mathsf{e}_1^* \wedge \mathsf{e}_2^* + \mathsf{e}_3^* \wedge \mathsf{e}_4^* \in \wedge^2 V^*$ cannot be written in the form of $T_1 \wedge T_2$ where $T_1, T_2 \in V^*$. It is because $(T_1 \wedge T_2) \wedge (T_1 \wedge T_2) = 0$ for any $T_1, T_2 \in V^*$, while $\sigma \wedge \sigma = 2\mathsf{e}_1^* \wedge \mathsf{e}_2^* \wedge \mathsf{e}_3^* \wedge \mathsf{e}_4^* \neq 0$. □

In the above remark, we take the wedge product between elements in $\wedge^2 V^*$. It is defined in a natural way that for any $T_1, \ldots T_k, S_1, \ldots, S_r \in V^*$, we have:

$$\underbrace{(T_1 \wedge \cdots \wedge T_k)}_{\in \wedge^k V^*} \wedge \underbrace{(S_1 \wedge \cdots \wedge S_r)}_{\in \wedge^r V^*} = \underbrace{T_1 \wedge \cdots \wedge T_k \wedge S_1 \wedge \cdots \wedge S_r}_{\in \wedge^{k+r} V^*}$$

and extended linearly to other elements in $\wedge^k V^*$ and $\wedge^r V^*$. For instance, we have:

$$\underbrace{(T_1 \wedge T_2 + S_1 \wedge S_2)}_{\in \wedge^2 V^*} \wedge \underbrace{\sigma}_{\in \wedge^k V^*} = \underbrace{T_1 \wedge T_2 \wedge \sigma + S_1 \wedge S_2 \wedge \sigma}_{\in \wedge^{k+2} V^*}.$$

While it is true that $T_1 \wedge T_2 = -T_2 \wedge T_1$ for any $T_1, T_2 \in V^*$, it is generally *not* true that $\sigma \wedge \eta = -\eta \wedge \sigma$ where $\sigma \in \wedge^k V^*$ and $\eta \in \wedge^r V^*$. For instance, let $T_1, \ldots, T_5 \in V^*$ and consider $\sigma = T_1 \wedge T_2$ and $\eta = T_3 \wedge T_4 \wedge T_5$. Then we can see that:

$$\begin{aligned}
\sigma \wedge \eta &= T_1 \wedge T_2 \wedge T_3 \wedge T_4 \wedge T_5 \\
&= -T_1 \wedge T_3 \wedge T_4 \wedge T_5 \wedge T_2 && \text{(switching } T_2 \text{ subsequently with } T_3, T_4, T_5\text{)} \\
&= T_3 \wedge T_4 \wedge T_5 \wedge T_1 \wedge T_2 && \text{(switching } T_1 \text{ subsequently with } T_3, T_4, T_5\text{)} \\
&= \eta \wedge \sigma.
\end{aligned}$$

Proposition 3.33. *Let V be a finite dimensional vector space, and V^* be the dual space of V. Given any $\sigma \in \wedge^k V^*$ and $\eta \in \wedge^r V^*$, we have:*

$$\sigma \wedge \eta = (-1)^{kr} \eta \wedge \sigma. \tag{3.12}$$

Clearly from (3.12), *any $\omega \in \wedge^{\text{even}} V^*$ commutes with any $\sigma \in \wedge^k V^*$.*

Proof. By linearity, it suffices to prove that case $\sigma = T_1 \wedge \cdots \wedge T_k$ and $\eta = S_1 \wedge \cdots \wedge S_r$ where $T_i, S_j \in V^*$, in which we have:

$$\sigma \wedge \eta = T_1 \wedge \cdots \wedge T_k \wedge S_1 \wedge \cdots \wedge S_r$$

In order to switch all the T_i's with the S_j's, we can first switch T_k subsequently with each of $S_1, \ldots, S_r$ and each switching contributes to a factor of (-1). Precisely, we have:

$$T_1 \wedge \cdots \wedge T_k \wedge S_1 \wedge \cdots \wedge S_r = (-1)^r T_1 \wedge \cdots \wedge T_{k-1} \wedge S_1 \wedge \cdots \wedge S_r \wedge T_k.$$

By repeating this sequence of switching on each of T_{k-1}, T_{k-2}, etc., we get a factor of $(-1)^r$ for each set of switching, and so we finally get the following as desired:

$$T_1 \wedge \cdots \wedge T_k \wedge S_1 \wedge \cdots \wedge S_r = [(-1)^r]^k \, S_1 \wedge \cdots \wedge S_r \wedge T_1 \wedge \cdots \wedge T_k$$

□

From Exercise 3.33, we know that $\dim \wedge^n V^* = 1$ if $n = \dim V$. In fact, every element $\sigma \in \dim \wedge^n V^*$ is a constant multiple of $\mathsf{e}_1^* \wedge \cdots \wedge \mathsf{e}_n^*$, and it is interesting (and important) to note that this constant multiple is related to a determinant! Precisely, for each $i = 1, \ldots, n$, we consider the elements:

$$\omega_i = \sum_{j=1}^n a_{ij} \mathsf{e}_j^* \in V^*$$

where a_{ij} are real constants. Then, the wedge product of all ω_i's are given by:

$$\begin{aligned}
\omega_1 \wedge \cdots \wedge \omega_n &= \left(\sum_{j_1=1}^n a_{1j_1} \mathsf{e}_{j_1}^* \right) \wedge \left(\sum_{j_2=1}^n a_{2j_2} \mathsf{e}_{j_2}^* \right) \wedge \cdots \wedge \left(\sum_{j_n=1}^n a_{nj_n} \mathsf{e}_{j_n}^* \right) \\
&= \sum_{j_1, \ldots, j_n \text{ distinct}} a_{1j_1} a_{2j_2} \ldots a_{nj_n} \mathsf{e}_{j_1}^* \wedge \cdots \wedge \mathsf{e}_{j_n}^* \\
&= \sum_{\sigma \in S_n} a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)} \mathsf{e}_{\sigma(1)}^* \wedge \cdots \wedge \mathsf{e}_{\sigma(n)}^*
\end{aligned}$$

Next we want to find a relation between $\mathsf{e}_{\sigma(1)}^* \wedge \cdots \wedge \mathsf{e}_{\sigma(n)}^*$ and $\mathsf{e}_1^* \wedge \cdots \wedge \mathsf{e}_n^*$. $\sigma \in S_n$, by decomposing it into transpositions $\sigma = \tau_1 \circ \cdots \circ \tau_k$, then we have:

$$\begin{aligned}
\mathsf{e}_{\sigma(1)}^* \wedge \cdots \wedge \mathsf{e}_{\sigma(n)}^* &= \mathsf{e}_{\tau_1 \circ \cdots \circ \tau_k(1)}^* \wedge \cdots \wedge \mathsf{e}_{\tau_1 \circ \cdots \circ \tau_k(n)}^* \\
&= (-1) \mathsf{e}_{\tau_2 \circ \cdots \circ \tau_k(1)}^* \wedge \cdots \wedge \mathsf{e}_{\tau_2 \circ \cdots \circ \tau_k(n)}^* \\
&= (-1)^2 \mathsf{e}_{\tau_3 \circ \cdots \circ \tau_k(1)}^* \wedge \cdots \wedge \mathsf{e}_{\tau_3 \circ \cdots \circ \tau_k(n)}^* \\
&= \ldots \\
&= (-1)^{k-1} \mathsf{e}_{\tau_k(1)}^* \wedge \cdots \wedge \mathsf{e}_{\tau_k(n)}^* \\
&= (-1)^k \mathsf{e}_1^* \wedge \cdots \wedge \mathsf{e}_n^* \\
&= \operatorname{sgn}(\sigma) \mathsf{e}_1^* \wedge \cdots \wedge \mathsf{e}_n^*.
\end{aligned}$$

Therefore, we have:

$$\omega_1 \wedge \cdots \wedge \omega_n = \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)} \right) \mathsf{e}_1^* \wedge \cdots \wedge \mathsf{e}_n^*.$$

Note that the sum:

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)}$$

is exactly the determinant of the matrix A whose (i, j)-th entry is a_{ij}. To summarize, let's state it as a proposition:

Proposition 3.34. *Let V^* be the dual space of a vector space V of dimension n, and let $\{\mathsf{e}_i\}_{i=1}^n$ be a basis for V, and $\{\mathsf{e}_i^*\}_{i=1}^n$ be the corresponding dual basis for V^*. Given any n elements $\omega_i = \sum_{j=1}^n a_{ij} \mathsf{e}_j^* \in V^*$, we have:*

$$\omega_1 \wedge \cdots \wedge \omega_n = (\det A)\ \mathsf{e}_1^* \wedge \cdots \wedge \mathsf{e}_n^*,$$

where A is the $n \times n$ matrix whose (i, j)-th entry is a_{ij}.

Exercise 3.35. Given an n-dimensional vector space V. Show that $\omega_1, \ldots, \omega_n \in V^*$ are linearly independent if and only if $\omega_1 \wedge \cdots \wedge \omega_n \neq 0$.

Exercise 3.36. Generalize Proposition 3.34. Precisely, now given

$$\omega_i = \sum_{j=1}^{n} a_{ij} \mathsf{e}_j^* \in V^*$$

where $1 \leq i \leq k < \dim V$, express $\omega_1 \wedge \cdots \wedge \omega_k$ in terms of e_i^*'s.

Exercise 3.37. Regard $\det : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as a multilinear map:

$$\det(\mathsf{v}_1, \ldots, \mathsf{v}_n) := \begin{vmatrix} | & & | \\ \mathsf{v}_1 & \cdots & \mathsf{v}_n \\ | & & | \end{vmatrix}.$$

Denote $\{\mathsf{e}_i\}$ the standard basis for $\mathbb{R}^n$. Show that:

$$\det = \mathsf{e}_1^* \wedge \cdots \wedge \mathsf{e}_n^*.$$

3.4.2. Differential Forms on Smooth Manifolds. In the simplest term, differential forms on a smooth manifold are wedge products of cotangent vectors in T^*M. At each point $p \in M$, let $(u_1, \ldots, u_n)$ be the local coordinates near p, then the cotangent space T_p^*M is spanned by $\left\{ du^1\big|_p, \ldots, du^n\big|_p \right\}$, and a smooth differential 1-form α is a map from M to T^*M such that it can be locally expressed as:

$$\alpha(p) = \left(p, \sum_{i=1}^{n} \alpha_i(p)\, du^i\Big|_p \right)$$

where α_i are smooth functions locally defined near p. Since the based point p can usually be understood from the context, we usually denote α by simply:

$$\alpha = \sum_{i=1}^{n} \alpha_i\, du^i.$$

Since T_p^*M is a finite dimensional vector space, we can consider the wedge products of its elements. A *differential k-form* ω on a smooth manifold M is a map which assigns each point $p \in M$ to an element in $\wedge^k T_p^*M$. Precisely:

Definition 3.35 (Smooth Differential k-Forms)**.** Let M be a smooth manifold. A *smooth differential k-form* ω on M is a map $\omega_p : \underbrace{T_pM \times \ldots \times T_pM}_{k \text{ times}} \to \mathbb{R}$ at each $p \in M$ such that under any local parametrization $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \to M$, it can be written in the form:

$$\omega = \sum_{i_1, \ldots, i_k = 1}^{n} \omega_{i_1 i_2 \cdots i_k}\, du^{i_1} \wedge \cdots \wedge du^{i_k}$$

where $\omega_{i_1 i_2 \ldots i_k}$'s are smooth scalar functions locally defined in $\mathsf{F}(\mathcal{U})$, and they are commonly called the *local components* of ω. The vector space of all smooth differential k-forms on M is denoted by $\wedge^k T^*M$.

Remark 3.36. It is a convention to denote $\wedge^0 T^*M := C^\infty(M, \mathbb{R})$, the vector space of all smooth scalar functions defined on M. $\square$

We will mostly deal with differential k-forms that are smooth. Therefore, we will very often call a *smooth differential k-form* simply by a *differential k-form*, or even simpler, a *k-form*. As we will see in the next section, the language of differential forms will unify and generalize the curl, grad and div in Multivariable Calculus and Physics courses.

From algebraic viewpoint, the manipulations of differential k-forms on a manifold are similar to those for wedge products of a finite-dimensional vector space. The major difference is a manifold is usually covered by more than one local parametrizations, hence there are conversion rules for differential k-forms from one local coordinate system to another.

Example 3.37. Consider $\mathbb{R}^2$ with (x, y) and (r, θ) as its two local coordinates. Given a 2-form $\omega = dx \wedge dy$, for instance, we can express it in terms of the polar coordinates (r, θ):

$$\begin{aligned} dx &= \frac{\partial x}{\partial r}\, dr + \frac{\partial x}{\partial \theta}\, d\theta \\ &= (\cos\theta)\, dr - (r\sin\theta)\, d\theta \\ dy &= \frac{\partial y}{\partial r}\, dr + \frac{\partial y}{\partial \theta}\, d\theta \\ &= (\sin\theta)\, dr + (r\cos\theta)\, d\theta \end{aligned}$$

Therefore, using $dr \wedge dr = 0$ and $d\theta \wedge d\theta = 0$, we get:

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

□

Exercise 3.38. Define a 2-form on $\mathbb{R}^3$ by:

$$\omega = x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy.$$

Express ω in terms of spherical coordinates (ρ, θ, φ), defined by:

$$(x, y, z) = (\rho\sin\varphi\cos\theta, \rho\sin\varphi\sin\theta, \rho\cos\varphi).$$

Exercise 3.39. Let ω be the 2-form on $\mathbb{R}^{2n}$ given by:

$$\omega = dx^1 \wedge dx^2 + dx^3 \wedge dx^4 + \ldots + dx^{2n-1} \wedge dx^{2n}.$$

Compute $\underbrace{\omega \wedge \cdots \wedge \omega}_{n \text{ times}}$.

Exercise 3.40. Let $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ be two local coordinates of a smooth manifold M. Show that:

$$du^1 \wedge \cdots \wedge du^n = \det \frac{\partial(u_1, \ldots, u_n)}{\partial(v_1, \ldots, v_n)}\, dv^1 \wedge \cdots \wedge dv^n.$$

Exercise 3.41. Show that on $\mathbb{R}^3$, a $(2, 0)$-tensor T is in $\wedge^2(\mathbb{R}^2)^*$ if and only if $T(\mathsf{v}, \mathsf{v}) = 0$ for any $\mathsf{v} \in \mathbb{R}^3$.

3.5. Exterior Derivatives

Exterior differentiation is an important operations on differential forms. It not only generalizes and unifies the curl, grad, div operators in Multivariable Calculus and Physics, but also leads to the development of de Rham cohomology to be discussed in Chapter 5.

3.5.1. Definition of Exterior Derivatives. Exterior differentiation, commonly denoted by the symbol d, takes a k-form to a $(k+1)$-form. To begin, let's define it on scalar functions first. Suppose $(u_1, \ldots, u_n)$ are local coordinates of M^n, then given any smooth scalar function $f \in C^\infty(M, \mathbb{R})$, we define:

$$df := \sum_{i=1}^{n} \frac{\partial f}{\partial u_i}\, du^i \tag{3.13}$$

Although (3.13) involves local coordinates, it can be easily shown that df is independent of local coordinates. Suppose $(v_1, \ldots, v_n)$ is another local coordinates of M which overlap with $(u_1, \ldots, u_n)$. By the chain rule, we have:

$$\begin{aligned} \frac{\partial f}{\partial u_i} &= \sum_{k=1}^{n} \frac{\partial f}{\partial v_k} \frac{\partial v_k}{\partial u_i} \\ dv^k &= \sum_{i=1}^{n} \frac{\partial v_k}{\partial u_i}\, du^i \end{aligned}$$

which combine to give:

$$\sum_{i=1}^{n} \frac{\partial f}{\partial u_i}\, du^i = \sum_{i=1}^{n} \sum_{k=1}^{n} \frac{\partial f}{\partial v_k} \frac{\partial v_k}{\partial u_i}\, du^i = \sum_{k=1}^{n} \frac{\partial f}{\partial v_k}\, dv^k.$$

Therefore, if f is smooth on M then df is a smooth 1-form on M. The components of df are $\frac{\partial f}{\partial u_i}$'s, and so df is analogous to ∇f in Multivariable Calculus. Note that as long as f is C^∞ just in an open set $\mathcal{U} \subset M$, we can also define df locally on $\mathcal{U}$ since (3.13) is a local expression.

Exterior derivatives can also be defined on differential forms of higher degrees. Let $\alpha \in \wedge^1 T^*M$, which can be locally written as:

$$\alpha = \sum_{i=1}^{n} \alpha_i\, du^i$$

where α_i's are smooth functions locally defined in a local coordinate chart. Then, we define:

$$d\alpha := \sum_{i=1}^{n} d\alpha_i \wedge du^i = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial \alpha_i}{\partial u_j}\, du^j \wedge du^i. \tag{3.14}$$

Using the fact that $du^j \wedge du^i = -du^i \wedge du^j$ and $du^i \wedge du^i = 0$, we can also express $d\alpha$ as:

$$d\alpha = \sum_{1 \le j < i \le n} \left(\frac{\partial \alpha_i}{\partial u_j} - \frac{\partial \alpha_j}{\partial u_i} \right) du^j \wedge du^i.$$

Example 3.38. Take $M = \mathbb{R}^3$ as an example, and let (x, y, z) be the (usual) coordinates of $\mathbb{R}^3$, then given any 1-form $\alpha = P\,dx + Q\,dy + R\,dz$ (which is analogous to the vector field $P\mathsf{i} + Q\mathsf{j} + R\mathsf{k}$), we have:

$$\begin{aligned}
d\alpha &= dP \wedge dx + dQ \wedge dy + dR \wedge dz \\
&= \left(\frac{\partial P}{\partial x}\,dx + \frac{\partial P}{\partial y}\,dy + \frac{\partial P}{\partial z}\,dz\right) \wedge dx + \left(\frac{\partial Q}{\partial x}\,dx + \frac{\partial Q}{\partial y}\,dy + \frac{\partial Q}{\partial z}\,dz\right) \wedge dy \\
&\quad + \left(\frac{\partial R}{\partial x}\,dx + \frac{\partial R}{\partial y}\,dy + \frac{\partial R}{\partial z}\,dz\right) \wedge dz \\
&= \frac{\partial P}{\partial y}\,dy \wedge dx + \frac{\partial P}{\partial z}\,dz \wedge dx + \frac{\partial Q}{\partial x}\,dx \wedge dy + \frac{\partial Q}{\partial z}\,dz \wedge dy \\
&\quad + \frac{\partial R}{\partial x}\,dx \wedge dz + \frac{\partial R}{\partial y}\,dy \wedge dz \\
&= \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy - \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right) dz \wedge dx + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dy \wedge dz
\end{aligned}$$

which is analogous to $\nabla \times (P\mathsf{i} + Q\mathsf{j} + R\mathsf{k})$ by declaring the correspondence $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$ with $\{dy \wedge dz, dz \wedge dx, dx \wedge dy\}$. □

One can check that the definition of $d\alpha$ stated in (3.14) is independent of local coordinates. On general k-forms, the exterior derivatives are defined in a similar way as:

Definition 3.39 (Exterior Derivatives)**.** Let M^n be a smooth manifold and $(u_1, \ldots, u_n)$ be local coordinates on M. Given any (smooth) k-form

$$\omega = \sum_{j_1,\ldots,j_k=1}^{n} \omega_{j_1\cdots j_k}\,du^{j_1} \wedge \cdots \wedge du^{j_k},$$

we define:

$$\begin{aligned}
d\omega &:= \sum_{j_1,\cdots,j_k=1}^{n} d\omega_{j_1\cdots j_k} \wedge du^{j_1} \wedge \cdots \wedge du^{j_k} \\
&= \sum_{j_1,\cdots,j_k=1}^{n} \sum_{i=1}^{n} \frac{\partial \omega_{j_1\cdots j_k}}{\partial u_i} du^i \wedge du^{j_1} \wedge \cdots \wedge du^{j_k}
\end{aligned} \tag{3.15}$$

In particular, if ω is an n-form (where $n = \dim V$), we have $d\omega = 0$.

Exercise 3.42. Show that $d\omega$ defined as in (3.15) does not depend on the choice of local coordinates.

Example 3.40. Consider $\mathbb{R}^2$ equipped with polar coordinates (r, θ). Consider the 1-form:

$$\omega = (r \sin\theta)\,dr.$$

Then, we have

$$\begin{aligned}
d\omega &= \frac{\partial(r\sin\theta)}{\partial r}\,dr \wedge dr + \frac{\partial(r\sin\theta)}{\partial \theta}\,d\theta \wedge dr \\
&= 0 + (r\cos\theta)\,d\theta \wedge dr \\
&= -(r\cos\theta)\,dr \wedge d\theta.
\end{aligned}$$

□

Exercise 3.43. Let $\omega = F_1\, dy \wedge dz + F_2\, dz \wedge dx + F_3\, dx \wedge dy$ be a smooth 2-form on $\mathbb{R}^3$. Compute $d\omega$. What operator in Multivariable Calculus is the d analogous to in this case?

Exercise 3.44. Let ω, η, θ be the following differential forms on $\mathbb{R}^3$:

$$\begin{aligned} \omega &= x\, dx - y, dy \\ \eta &= z\, dx \wedge dy + x\, dy \wedge dz \\ \theta &= z\, dy \end{aligned}$$

Compute: $\omega \wedge \eta$, $\omega \wedge \eta \wedge \theta$, $d\omega$, $d\eta$ and $d\theta$.

3.5.2. Properties of Exterior Derivatives. The exterior differentiation d can hence be regarded as a chain of maps:

$$\wedge^0 T^*M \xrightarrow{d} \wedge^1 T^*M \xrightarrow{d} \wedge^2 T^*M \xrightarrow{d} \cdots \xrightarrow{d} \wedge^{n-1} T^*M \xrightarrow{d} \wedge^n T^*M.$$

Here we abuse the use of the symbol d a little bit – we use the same symbol d for all the maps $\wedge^k T^*M \xrightarrow{d} \wedge^{k+1} T^*M$ in the chain. The following properties about exterior differentiation are not difficult to prove:

Proposition 3.41. *For any k-forms ω and η, and any smooth scalar function f, we have the following:*

(1) $d(\omega + \eta) = d\omega + d\eta$

(2) $d(f\omega) = df \wedge \omega + f\, d\omega$

Proof. (1) is easy to prove (left as an exercise for readers). To prove (2), we consider local coordinates $(u_1, \ldots, u_n)$ and let $\omega = \sum_{j_1,\ldots,j_k=1}^{n} \omega_{j_1\cdots j_k}\, du^{j_1} \wedge \cdots \wedge du^{j_k}$. Then, we have:

$$\begin{aligned} d(f\omega) &= \sum_{j_1,\ldots,j_k=1}^{n} \sum_{i=1}^{n} \frac{\partial}{\partial u_i}(f\omega_{j_1\cdots j_k})\, du^i \wedge du^{j_1} \wedge \cdots \wedge du^{j_k} \\ &= \sum_{j_1,\ldots,j_k=1}^{n} \sum_{i=1}^{n} \left(\frac{\partial f}{\partial u_i}\, \omega_{j_1\cdots j_k} + f\, \frac{\partial \omega_{j_1\cdots j_k}}{\partial u_i} \right) du^i \wedge du^{j_1} \wedge \cdots \wedge du^{j_k} \\ &= \left(\sum_{i=1}^{n} \frac{\partial f}{\partial u_i}\, du^i \right) \wedge \left(\sum_{j_1,\ldots,j_k=1}^{n} \omega_{j_1\cdots j_k}\, du^{j_1} \wedge \cdots \wedge du^{j_k} \right) \\ &\quad + f \sum_{j_1,\cdots,j_k=1}^{n} \sum_{i=1}^{n} \frac{\partial \omega_{j_1\cdots j_k}}{\partial u_i} du^i \wedge du^{j_1} \wedge \cdots \wedge du^{j_k} \end{aligned}$$

as desired. □

Identity (2) in Proposition 3.41 can be regarded as a kind of product rule. Given a k-form α and a r-form β, the general product rule for exterior derivative is stated as:

Proposition 3.42. *Let $\alpha \in \wedge^k T^*M$ and $\beta \in \wedge^r T^*M$ be smooth differential forms on M, then we have:*

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k\, \alpha \wedge d\beta.$$

Exercise 3.45. Prove Proposition 3.42. Based on your proof, explain briefly why the product rule does not involve any factor of $(-1)^r$.

Exercise 3.46. Given three differential forms α, β and γ such that $d\alpha = 0$, $d\beta = 0$ and $d\gamma = 0$. Show that:

$$d(\alpha \wedge \beta \wedge \gamma) = 0.$$

An crucial property of exterior derivatives is that the *composition* is zero. For instance, given a smooth scalar function $f(x, y, z)$ defined on $\mathbb{R}^3$, we have:

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz.$$

Taking exterior derivative one more time, we get:

$$\begin{aligned}
d(df) &= \left(\frac{\partial}{\partial x}\frac{\partial f}{\partial x}\,dx + \frac{\partial}{\partial y}\frac{\partial f}{\partial x}\,dy + \frac{\partial}{\partial z}\frac{\partial f}{\partial x}\,dz\right) \wedge dx \\
&\quad + \left(\frac{\partial}{\partial x}\frac{\partial f}{\partial y}\,dx + \frac{\partial}{\partial y}\frac{\partial f}{\partial y}\,dy + \frac{\partial}{\partial z}\frac{\partial f}{\partial y}\,dz\right) \wedge dy \\
&\quad + \left(\frac{\partial}{\partial x}\frac{\partial f}{\partial z}\,dx + \frac{\partial}{\partial y}\frac{\partial f}{\partial z}\,dy + \frac{\partial}{\partial z}\frac{\partial f}{\partial z}\,dz\right) \wedge dz \\
&= \left(\frac{\partial}{\partial x}\frac{\partial f}{\partial y} - \frac{\partial}{\partial y}\frac{\partial f}{\partial x}\right) dx \wedge dy + \left(\frac{\partial}{\partial z}\frac{\partial f}{\partial x} - \frac{\partial}{\partial x}\frac{\partial f}{\partial z}\right) dz \wedge dx \\
&\quad + \left(\frac{\partial}{\partial y}\frac{\partial f}{\partial z} - \frac{\partial}{\partial z}\frac{\partial f}{\partial y}\right) dy \wedge dz
\end{aligned}$$

Since partial derivatives commute, we get $d(df) = 0$, or in short $d^2 f = 0$, for any scalar function f. The fact that $d^2 = 0$ is generally true on smooth differential forms, not only for scalar functions. Precisely, we have:

Proposition 3.43. *Let ω be a smooth k-form defined on a smooth manifold M, then we have:*

$$d^2\omega := d(d\omega) = 0.$$

Proof. Let $\omega = \sum_{j_1,\ldots,j_k=1}^{n} \omega_{j_1\cdots j_k}\,du^{j_1} \wedge \cdots \wedge du^{j_k}$, then:

$$d\omega = \sum_{j_1,\cdots,j_k=1}^{n} \sum_{i=1}^{n} \frac{\partial \omega_{j_1\cdots j_k}}{\partial u_i}\,du^i \wedge du^{j_1} \wedge \cdots \wedge du^{j_k}.$$

$$\begin{aligned}
d^2\omega &= d\left(\sum_{j_1,\cdots,j_k=1}^{n} \sum_{i=1}^{n} \frac{\partial \omega_{j_1\cdots j_k}}{\partial u_i}\,du^i \wedge du^{j_1} \wedge \cdots \wedge du^{j_k}\right) \\
&= \sum_{j_1,\cdots,j_k=1}^{n} \sum_{i=1}^{n} \sum_{l=1}^{n} \frac{\partial^2 \omega_{j_1\ldots j_k}}{\partial u_l \partial u_i}\,du^l \wedge du^i \wedge du^{j_1} \wedge \cdots \wedge du^{j_k}
\end{aligned}$$

For each fixed k-tuple $(j_1, \ldots, j_k)$, the term $\sum_{i,l=1}^{n} \frac{\partial^2 \omega_{j_1\ldots j_k}}{\partial u_l \partial u_i}\,du^l \wedge du^i$ can be rewritten as:

$$\sum_{1 \leq i < l \leq n} \left(\frac{\partial^2 \omega_{j_1\ldots j_k}}{\partial u_l \partial u_i} - \frac{\partial^2 \omega_{j_1\ldots j_k}}{\partial u_i \partial u_l}\right) du^l \wedge du^i$$

which is zero since partial derivatives commute. It concludes that $d^2\omega = 0$. □

Proposition 3.43 is a important fact that leads to the development of de Rham cohomology in Chapter 5.

In Multivariable Calculus, we learned that given a vector field $\mathsf{F} = P\mathsf{i} + Q\mathsf{j} + R\mathsf{k}$ and a scalar function f, we have:

$$\nabla \times \nabla f = 0$$
$$\nabla \cdot (\nabla \times \mathsf{F}) = 0$$

These two formulae can be unified using the language of differential forms. The one-form df corresponds to the vector field ∇f:

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz$$
$$\nabla f = \frac{\partial f}{\partial x}\mathsf{i} + \frac{\partial f}{\partial y}\mathsf{j} + \frac{\partial f}{\partial z}\mathsf{k}$$

Define a one-form $\omega = P\,dx + Q\,dy + R\,dz$ on $\mathbb{R}^3$, which corresponds to the vector field F, then we have discussed that $d\omega$ corresponds to taking curl of F:

$$d\omega = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy - \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right) dz \wedge dx + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dy \wedge dz$$
$$\nabla \times \mathsf{F} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathsf{k} - \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right)\mathsf{j} + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathsf{i}$$

If one takes $\omega = df$, and $\mathsf{F} = \nabla f$, then we have $d\omega = d(df) = 0$, which corresponds to the fact that $\nabla \times \mathsf{G} = \nabla \times \nabla f = 0$ in Multivariable Calculus.

Taking exterior derivative on a two-form $\beta = A\,dy \wedge dz + B\,dz \wedge dx + C\,dx \wedge dy$ corresponds to taking the divergence on the vector field $\mathsf{G} = A\mathsf{i} + B\mathsf{j} + C\mathsf{k}$ according to Exercise 3.43:

$$d\beta = \left(\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z}\right) dx \wedge dy \wedge dz$$
$$\nabla \cdot \mathsf{G} = \left(\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z}\right)$$

By taking $\beta = d\omega$, and $\mathsf{G} = \nabla \times \mathsf{F}$, then we have $d\beta = d(d\omega) = 0$ corresponding to $\nabla \cdot \mathsf{G} = \nabla \cdot (\nabla \times \mathsf{F}) = 0$ in Multivariable Calculus.

Here is a summary of the correspondences:

Differential Form on $\mathbb{R}^3$	Multivariable Calculus
$f(x,y,z)$	$f(x,y,z)$
$\omega = P\,dx + Q\,dy + R\,dz$	$\mathsf{F} = P\mathsf{i} + Q\mathsf{j} + R\mathsf{k}$
$\beta = A\,dy \wedge dz + B\,dz \wedge dx + C\,dx \wedge dy$	$\mathsf{G} = A\mathsf{i} + B\mathsf{j} + C\mathsf{k}$
df	∇f
$d\omega$	$\nabla \times \mathsf{F}$
$d\beta$	$\nabla \cdot \mathsf{G}$
$d^2 f = 0$	$\nabla \times \nabla f = 0$
$d^2\omega = 0$	$\nabla \cdot (\nabla \times \mathsf{F}) = 0$

3.5.3. Exact and Closed Forms. In Multivariable Calculus, we discussed various concepts of vector fields including potential functions, conservative vector fields, solenoidal vector fields, curl-less and divergence-less vector fields, etc. All these concepts can be unified using the language of differential forms.

As a reminder, a conservative vector field F is one that can be expressed as $\mathsf{F} = \nabla f$ where f is a scalar function. It is equivalent to saying that the 1-form ω can be expressed as $\omega = df$. Moreover, a solenoidal vector field G is one that can be expressed as $\mathsf{G} = \nabla \times \mathsf{F}$ for some vector field F. It is equivalent to saying that the 2-form β can be expressed as $\beta = d\omega$ for some 1-form ω.

Likewise, a curl-less vector field F (i.e. $\nabla \times \mathsf{F} = 0$) corresponds to a 1-form ω satisfying $d\omega = 0$; and a divergence-less vector field G (i.e. $\nabla \cdot \mathsf{G} = 0$) corresponds to a 2-form β satisfying $d\beta = 0$.

In view of the above correspondence, we introduce two terminologies for differential forms, namely *exact-ness* and *closed-ness*:

Definition 3.44 (Exact and Closed Forms)**.** Let ω be a smooth k-form defined on a smooth manifold M, then we say:

- ω is *exact* if there exists a $(k-1)$-form η defined on M such that $\omega = d\eta$;
- ω is *closed* if $d\omega = 0$.

Remark 3.45. By the fact that $d^2 = 0$ (Proposition 3.43), it is clear that every exact form is a closed form (but *not* vice versa). □

The list below showcases the corresponding concepts of exact/closed forms in Multivariable Calculus.

Differential Form on $\mathbb{R}^3$	Multivariable Calculus
exact 1-form	conservative vector field
closed 1-form	curl-less vector field
exact 2-form	solenoidal vector field
closed 2-form	divergence-less vector field

Example 3.46. On $\mathbb{R}^3$, the 1-form:

$$\alpha = yz\,dx + zx\,dy + xy\,dz$$

is exact since $\alpha = df$ where $f(x,y,z) = xyz$. By Proposition 3.43, we immediately get $d\alpha = d(df) = 0$, so α is a closed form. One can also verify this directly:

$$\begin{aligned} d\alpha &= (z\,dy + y\,dz) \wedge dx + (z\,dx + x\,dz) \wedge dy + (y\,dx + x\,dy) \wedge dz \\ &= (z-z)\,dx \wedge dy + (y-y)\,dz \wedge dx + (x-x)\,dy \wedge dz = 0. \end{aligned}$$

□

Example 3.47. The 1-form:

$$\alpha := -\frac{y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy$$

defined on $\mathbb{R}^2 \backslash \{(0,0)\}$ is closed:

$$\begin{aligned} d\alpha &= \frac{\partial}{\partial y}\left(-\frac{y}{x^2+y^2}\right) dy \wedge dx + \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) dx \wedge dy \\ &= \frac{y^2-x^2}{(x^2+y^2)^2}\,dy \wedge dx + \frac{y^2-x^2}{(x^2+y^2)^2}\,dx \wedge dy \\ &= 0 \end{aligned}$$

as $dx \wedge dy = -dy \wedge dx$. However, we will later see that α is not exact.

Note that even though we have $\alpha = df$ where $f(x,y) = \tan^{-1}\frac{y}{x}$, such an f is NOT smooth on $\mathbb{R}^2 \backslash \{(0,0)\}$. In order to claim α is exact, we require such an f to be smooth on the domain of α. □

Exercise 3.47. Consider the forms ω, η and θ on $\mathbb{R}^3$ defined in Exercise 3.44. Determine whether each of them is closed and/or exact on $\mathbb{R}^3$.

Exercise 3.48. The purpose of this exercise is to show that any closed 1-form ω on $\mathbb{R}^3$ must be exact. Let

$$\omega = P(x,y,z)\,dx + Q(x,y,z)\,dy + R(x,y,z)\,dz$$

be a closed 1-form on $\mathbb{R}^3$. Define $f : \mathbb{R}^3 \to \mathbb{R}$ by:

$$f(x,y,z) = \int_{t=0}^{t=1} \left(xP(tx,ty,tz) + yQ(tx,ty,tz) + zR(tx,ty,tz)\right)\,dt$$

Show that $\omega = df$. Point out exactly where you have used the fact that $d\omega = 0$.

3.5.4. Pull-Back of Tensors. Let's first begin by reviewing the *push-forward* and *pull-back* of tangent and cotangent vectors. Given a smooth map $\Phi : M \to N$ between two smooth manifolds M^m and N^n, its tangent map Φ_* takes a tangent vector in T_pM to a tangent vector in $T_{\Phi(p)}N$. If we let $\mathsf{F}(u_1, \ldots, u_m)$ be local coordinates of M, $\mathsf{G}(v_1, \ldots, v_n)$ be local coordinates of N and express the map Φ locally as:

$$(v_1, \ldots, v_n) = \mathsf{G}^{-1} \circ \Phi \circ \mathsf{F}(u_1, \ldots, u_m),$$

then Φ_* acts on the basis vectors $\left\{\dfrac{\partial}{\partial u_i}\right\}$ by:

$$\Phi_*\left(\frac{\partial}{\partial u_i}\right) = \frac{\partial \Phi}{\partial u_i} = \sum_j \frac{\partial v_j}{\partial u_i}\frac{\partial}{\partial v_j}.$$

The tangent map Φ_* is also commonly called the *push-forward* map. It is important to note that the v_j's in the partial derivatves $\frac{\partial v_j}{\partial u_i}$ can sometimes cause confusion if we talk about the push-forwards of two different smooth maps $\Phi : M \to N$ and $\Psi : M \to N$. Even with the same input $(u_1, \ldots, u_m)$, the output $\Phi(u_1, \ldots, u_m)$ and $\Psi(u_1, \ldots, u_m)$ are generally different and have different v_j-coordinates. To avoid this confusion, it is best to write:

$$\Phi_*\left(\frac{\partial}{\partial u_i}\right) = \sum_j \frac{\partial(v_j \circ \Phi)}{\partial u_i}\frac{\partial}{\partial v_j}$$

$$\Psi_*\left(\frac{\partial}{\partial u_i}\right) = \sum_j \frac{\partial(v_j \circ \Psi)}{\partial u_i}\frac{\partial}{\partial v_j}$$

Here each v_j in the partial derivatives $\frac{\partial v_j}{\partial u_i}$ are considered to be a locally defined function taking a point $p \in N$ to its v_j-coordinate.

For cotangent vectors (i.e. 1-forms), we talk about *pull-back* instead. According to Definition 3.14, Φ^* takes a cotangent vector in $T^*_{\Phi(p)}N$ to a cotangent vector in T^*_pM, defined as follows:

$$\Phi^*(dv^i)(X) = dv^i\,(\Phi_* X) \qquad \text{for any } X \in T_pM.$$

In terms of local coordinates, it is given by:

$$\Phi^*(dv^i) = \sum_j \frac{\partial(v_i \circ \Phi)}{\partial u_j}\,du^j.$$

The pull-back action by a smooth $\Phi : M \to N$ between manifolds can be extended to $(k, 0)$-tensors (and hence to differential forms):

Definition 3.48 (Pull-Back on $(k,0)$-Tensors)**.** Let $\Phi : M \to N$ be a smooth map between two smooth manifolds. Given T a smooth $(k,0)$-tensor on N, then we define:

$$(\Phi^*T)_p\,(X_1,\ldots,X_k) = T_{\Phi(p)}\,(\Phi_*(X_1),\ldots,\Phi_*(X_k)) \qquad \text{for any } X_1,\ldots,X_k \in T_pM$$

Remark 3.49. An equivalent way to state the definition is as follows: let $T_1,\ldots T_k \in TN$ be 1-forms on N, then we define:

$$\Phi^*(T_1 \otimes \cdots \otimes T_k) = (\Phi^*T_1) \otimes \cdots \otimes (\Phi^*T_k).$$

□

Remark 3.50. It is easy to verify that Φ^* is linear, in a sense that:

$$\Phi^*(aT + bS) = a\Phi^*T + b\Phi^*S$$

for any $(k,0)$-tensors T and S, and scalars a and b. □

Example 3.51. Let's start with an example on $\mathbb{R}^2$. Let $\Phi : \mathbb{R}^2 \to \mathbb{R}^3$ be a map defined by:

$$\Phi(x_1,x_2) = \left(e^{x_1+x_2}, \sin(x_1^2x_2), x_1\right).$$

To avoid confusion, we use (x_1,x_2) to label the coordinates of the domain $\mathbb{R}^2$, and use (y_1,y_2,y_3) to denote the coordinates of the codomain $\mathbb{R}^3$. Then, we have:

$$\begin{aligned}
\Phi^*(dy^1)\left(\frac{\partial}{\partial x_1}\right) &= dy^1\left(\Phi_*\left(\frac{\partial}{\partial x_1}\right)\right) = dy^1\left(\frac{\partial\Phi}{\partial x_1}\right)\\
&= dy^1\left(\frac{\partial(y_1\circ\Phi)}{\partial x_1}\frac{\partial}{\partial y_1} + \frac{\partial(y_2\circ\Phi)}{\partial x_1}\frac{\partial}{\partial y_2} + \frac{\partial(y_3\circ\Phi)}{\partial x_1}\frac{\partial}{\partial y_3}\right)\\
&= \frac{\partial(y_1\circ\Phi)}{\partial x_1} = \frac{\partial}{\partial x_1}e^{x_1+x_2} = e^{x_1+x_2}.
\end{aligned}$$

Similarly, we have:

$$\Phi^*(dy^1)\left(\frac{\partial}{\partial x_2}\right) = \frac{\partial(y_1\circ\Phi)}{\partial x_2} = \frac{\partial}{\partial x_2}e^{x_1+x_2} = e^{x_1+x_2}.$$

Therefore, $\Phi^*(dy^1) = e^{x_1+x_2}dx^1 + e^{x_1+x_2}dx^2 = e^{x_1+x_2}(dx^1+dx^2)$. We leave it as an exercise for readers to verify that:

$$\begin{aligned}
\Phi^*(dy^2) &= 2x_1x_2\cos(x_1^2x_2)\,dx^1 + x_1^2\cos(x_1^2x_2)\,dx^2\\
\Phi^*(dy^3) &= dx^1
\end{aligned}$$

Let $f(y_1,y_2,y_3)$ be a scalar function on $\mathbb{R}^3$, and consider the $(2,0)$-tensor on $\mathbb{R}^3$:

$$T = f(y_1,y_2,y_3)\,dy^1\otimes dy^2$$

The pull-back of T by Φ is given by:

$$\begin{aligned}
\Phi^*T &= f(y_1,y_2,y_3)\,\Phi^*(dy^1)\otimes\Phi^*(dy^2)\\
&= f(\Phi(x_1,x_2))\left(e^{x_1+x_2}(dx^1+dx^2)\right)\otimes\left(2x_1x_2\cos(x_1^2x_2)\,dx^1 + x_1^2\cos(x_1^2x_2)\,dx^2\right)
\end{aligned}$$

The purpose of writing $f(y_1,y_2,y_3)$ as $f(\Phi(x_1,x_2))$ is to leave the final expression in terms of functions and tensors in (x_1,x_2)-coordinates. □

Example 3.52. Let Σ be a regular surface in $\mathbb{R}^3$. The standard dot product in $\mathbb{R}^3$ is given by the following $(2,0)$-tensor:

$$\omega = dx\otimes dx + dy\otimes dy + dz\otimes dz.$$

Consider the inclusion map $\iota : \Sigma \to \mathbb{R}^3$. Although the input and output are the same under the map ι, the cotangents dx and $\iota^*(dx)$ are different! The former is a cotangent vector on $\mathbb{R}^3$, while $\iota^*(dx)$ is a cotangent vector on the surface Σ. If $(x, y, z) = \mathsf{F}(u, v)$ is a local parametrization of Σ, then $\iota^*(dx)$ should be in terms of du and dv, but not dx, dy and dz. Precisely, we have:

$$\begin{aligned}
\iota_* \left(\frac{\partial \mathsf{F}}{\partial u}\right) &= \frac{\partial \iota}{\partial u} := \frac{\partial(\iota \circ \mathsf{F})}{\partial u} = \frac{\partial \mathsf{F}}{\partial u} \\
\iota^*(dx)\left(\frac{\partial \mathsf{F}}{\partial u}\right) &= dx\left(\iota_*\left(\frac{\partial \mathsf{F}}{\partial u}\right)\right) = dx\left(\frac{\partial \mathsf{F}}{\partial u}\right) \\
&= dx\left(\frac{\partial x}{\partial u}\frac{\partial}{\partial x} + \frac{\partial y}{\partial u}\frac{\partial}{\partial y} + \frac{\partial z}{\partial u}\frac{\partial}{\partial z}\right) \\
&= \frac{\partial x}{\partial u}.
\end{aligned}$$

Similarly, we also have $\iota^*(dx)\left(\frac{\partial \mathsf{F}}{\partial v}\right) = \frac{\partial x}{\partial v}$, and hence:

$$\iota^*(dx) = \frac{\partial x}{\partial u}\, du + \frac{\partial x}{\partial v}\, dv.$$

As a result, we have:

$$\begin{aligned}
\iota^*\omega &= \iota^*(dx) \otimes \iota^*(dx) + \iota^*(dy) \otimes \iota^*(dy)\iota^*(dz) \otimes \iota^*(dz) \\
&= \left(\frac{\partial x}{\partial u}\, du + \frac{\partial x}{\partial v}\, dv\right) \otimes \left(\frac{\partial x}{\partial u}\, du + \frac{\partial x}{\partial v}\, dv\right) \\
&\quad + \left(\frac{\partial y}{\partial u}\, du + \frac{\partial y}{\partial v}\, dv\right) \otimes \left(\frac{\partial y}{\partial u}\, du + \frac{\partial y}{\partial v}\, dv\right) \\
&\quad + \left(\frac{\partial z}{\partial u}\, du + \frac{\partial z}{\partial v}\, dv\right) \otimes \left(\frac{\partial z}{\partial u}\, du + \frac{\partial z}{\partial v}\, dv\right).
\end{aligned}$$

After expansion and simplification, one will get:

$$\iota^*\omega = \frac{\partial \mathsf{F}}{\partial u} \cdot \frac{\partial \mathsf{F}}{\partial u}\, du \otimes du + \frac{\partial \mathsf{F}}{\partial u} \cdot \frac{\partial \mathsf{F}}{\partial v}\, du \otimes dv + \frac{\partial \mathsf{F}}{\partial v} \cdot \frac{\partial \mathsf{F}}{\partial u}\, dv \otimes du + \frac{\partial \mathsf{F}}{\partial v} \cdot \frac{\partial \mathsf{F}}{\partial v}\, dv \otimes dv,$$

which is the first fundamental form in Differential Geometry. □

Exercise 3.49. Let the unit sphere $\mathbb{S}^2$ be locally parametrized by spherical coordinates (θ, φ). Consider the $(2, 0)$-tensor on $\mathbb{R}^3$:

$$\omega = x\, dy \otimes dz$$

Express the pull-back $\iota^*\omega$ in terms of (θ, φ).

One can derive a general formula (which you do not need to remember in practice) for the local expression of pull-backs. Consider local coordinates $\{u_i\}$ for M and $\{v_i\}$ for N, and write $(v_1, \ldots, v_n) = \Phi(u_1, \ldots, u_m)$ and

$$T = \sum_{i_1, \ldots, i_k = 1}^{n} T_{i_1 \cdots i_k}(v_1, \ldots, v_n)\, dv^{i_1} \otimes \cdots \otimes dv^{i_k}.$$

The pull-back Φ^*T then has the following local expression:

$$\begin{aligned}
(3.16)\quad \Phi^*T &= \sum_{i_1,\ldots,i_k=1}^{n} T_{i_1\cdots i_k}(v_1,\ldots,v_n)\,\Phi^*(dv^{i_1})\otimes\cdots\otimes\Phi^*(dv^{i_k})\\
&= \sum_{i_1,\ldots,i_k=1}^{n} T_{i_1\cdots i_k}(\Phi(u_1,\ldots,u_m))\left(\sum_{j_1=1}^{m}\frac{\partial v_{i_1}}{\partial u_{j_1}}\,du^{j_1}\right)\otimes\cdots\otimes\left(\sum_{j_k=1}^{m}\frac{\partial v_{i_k}}{\partial u_{j_k}}\,du^{j_k}\right)\\
&= \sum_{i_1,\ldots,i_k=1}^{n}\sum_{j_1,\ldots,j_k=1}^{m} T_{i_1\cdots i_k}(\Phi(u_1,\ldots,u_m))\frac{\partial v_{i_1}}{\partial u_{j_1}}\cdots\frac{\partial v_{i_k}}{\partial u_{j_k}}\,du^{j_1}\otimes\cdots\otimes du^{j_k}.
\end{aligned}$$

In view of $T_{i_1\cdots i_k}(v_1,\ldots,v_n) = T_{i_1\cdots i_k}(\Phi(u_1,\ldots,u_m))$ and the above local expression, we define

$$\Phi^*f := f\circ\Phi$$

for any scalar function of f. Using this notation, we then have $\Phi^*(fT) = (\Phi^*f)\,\Phi^*T$ for any scalar function f and $(k,0)$-tensor T.

Exercise 3.50. Let $\Phi : M \to N$ be a smooth map between smooth manifolds M and N, f be a smooth scalar function defined on N. Show that

$$\Phi^*(df) = d(\Phi^*f).$$

In particular, if $(v_1,\ldots,v_n)$ are local coordinates of N, we have $\Phi^*(dv^j) = d(\Phi^*v^j)$.

Example 3.53. Using the result from Exercise 3.50, one can compute the pull-back by inclusion map $\iota : \Sigma \to \mathbb{R}^3$ for regular surfaces Σ in $\mathbb{R}^3$. Suppose $\mathsf{F}(u,v)$ is a local parametrization of Σ, then:

$$\iota^*(dx) = d(\iota^*x) = d(x\circ\iota).$$

Although $x\circ\iota$ and x (as a coordinate function) have the same output, their domains are different! Namely, $x\circ\iota : \Sigma \to \mathbb{R}$ while $x : \mathbb{R}^3 \to \mathbb{R}$. Therefore, when computing $d(x\circ\iota)$, one should express it in terms of local coordinates (u,v) of Σ:

$$d(x\circ\iota) = \frac{\partial(x\circ\iota)}{\partial u}\,du + \frac{\partial(x\circ\iota)}{\partial v}\,dv = \frac{\partial x}{\partial u}\,du + \frac{\partial x}{\partial v}\,dv.$$

□

Recall that the tangent maps (i.e. push-forwards) acting on tangent vectors satisfy the chain rule: if $\Phi : M \to N$ and $\Psi : N \to P$ are smooth maps between smooth manifolds, then we have $(\Psi\circ\Phi)_* = \Psi_*\circ\Phi_*$. It is easy to extend the chain rule to $(k,0)$-tensors:

Theorem 3.54 (Chain Rule for $(k,0)$-tensors)**.** *Let $\Phi : M \to N$ and $\Psi : N \to P$ be smooth maps between smooth manifolds M, N and P, then the pull-back maps Φ^* and Ψ^* acting on $(k,0)$-tensors for any $k \geq 1$ satisfy the following chain rule:*

$$(\Psi\circ\Phi)^* = \Phi^*\circ\Psi^*. \tag{3.17}$$

Exercise 3.51. Prove Theorem 3.54.

Exercise 3.52. Denote id_M and id_{TM} to be the identity maps of a smooth manifold M and its tangent bundle respectively. Show that $(\mathrm{id}_M)^* = \mathrm{id}_{TM}$. Hence, show that if M and N are diffeomorphic, then for $k \geq 1$ the vector spaces of $(k,0)$-tensors $\otimes^k T^*M$ and $\otimes^k T^*N$ are isomorphic.

3.5.5. Pull-Back of Differential Forms. By linearity of the pull-back map, and the fact that differential forms are linear combinations of tensors, the pull-back map acts on differential forms by the following way:

$$\Phi^*(T_1 \wedge \cdots \wedge T_k) = \Phi^* T_1 \wedge \cdots \wedge \Phi^* T_k$$

for any 1-forms $T_1, \ldots, T_k$.

Example 3.55. Consider the map $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ given by:

$$\underbrace{\Phi(x_1, x_2)}_{(y_1, y_2)} = (x_1^2 - x_2, x_2^3).$$

By straight-forward computations, we have:

$$\Phi^*(dy^1) = 2x_1 \, dx^1 - dx^2$$
$$\Phi^*(dy^2) = 3x_2 \, dx^2$$

Therefore, we have:

$$\Phi^*(dy^1 \wedge dy^2) = \Phi^*(dy^1) \wedge \Phi^*(dy^2) = 6x_1 x_2 \, dx^1 \wedge dx^2.$$

Note that $6x_1x_2$ is the Jacobian determinant $\det[\Phi_*]$. We will see soon that it is not a coincident, and it holds true in general. □

Although the computation of pull-back on differential forms is not much different from that on tensors, there are several distinctive features for pull-back on forms. One feature is that the pull-back on forms is closely related to Jacobian determinants:

Proposition 3.56. *Let $\Phi : M \to N$ be a smooth map between two smooth manifolds. Suppose $(u_1, \ldots, u_m)$ are local coordinates of M, and $(v_1, \ldots, v_n)$ are local coordinates of N, then for any $1 \leq i_1, \ldots, i_k \leq n$, we have:*

$$\Phi^*(dv^{i_1} \wedge \cdots \wedge dv^{i_k}) = \sum_{1 \leq j_1 < \cdots < j_k \leq m} \det \frac{\partial(v_{i_1}, \ldots, v_{i_k})}{\partial(u_{j_1}, \ldots, u_{j_k})} \, du^{j_1} \wedge \cdots \wedge du^{j_k}. \tag{3.18}$$

In particular, if $\dim M = \dim N = n$*, then we have:*

$$\Phi^*(dv^1 \wedge \cdots \wedge dv^n) = \det[\Phi_*] \, du^1 \wedge \cdots \wedge du^n \tag{3.19}$$

where $[\Phi_*]$ is the Jacobian matrix of Φ with respect to local coordinates $\{u_i\}$ and $\{v_i\}$, i.e.

$$[\Phi_*] = \frac{\partial(v_1, \ldots, v_n)}{\partial(u_1, \ldots, u_n)}.$$

Proof. Proceed as in the derivation of (3.16) by simply replacing all tensor products by wedge products, we get:

$$\begin{aligned} \Phi^*(dv^{i_1} \wedge \cdots \wedge dv^{i_k}) &= \sum_{j_1, \ldots, j_k = 1}^{m} \left(\frac{\partial v_{i_1}}{\partial u_{j_1}} \cdots \frac{\partial v_{i_k}}{\partial u_{j_k}} \, du^{j_1} \wedge \cdots \wedge du^{j_k} \right) \\ &= \sum_{\substack{j_1, \ldots, j_k = 1 \\ j_1, \ldots, j_k \text{ distinct}}}^{m} \left(\frac{\partial v_{i_1}}{\partial u_{j_1}} \cdots \frac{\partial v_{i_k}}{\partial u_{j_k}} \, du^{j_1} \wedge \cdots \wedge du^{j_k} \right) \end{aligned}$$

The second equality follows from the fact that $du^{j_1} \wedge \cdots \wedge du^{j_k} = 0$ if $\{j_1, \ldots, j_k\}$ are not all distinct. Each k-tuples $(j_1, \ldots, j_k)$ with distinct j_i's can be obtained by permuting a strictly increasing sequence of j's. Precisely, we have:

$$\begin{aligned}&\{(j_1,\ldots,j_k) : 1 \le j_1,\ldots,j_k \le n \text{ and } j_1,\ldots,j_k \text{ are all distinct}\}\\ &= \bigcup_{\sigma\in S_k} \{(j_{\sigma(1)},\ldots,j_{\sigma(k)}) : 1 \le j_1 < j_2 < \ldots < j_k \le n\}\end{aligned}$$

Therefore, we get:

$$\begin{aligned}&\Phi^*(dv^{i_1} \wedge \cdots \wedge dv^{i_k})\\ &= \sum_{1\le j_1<\ldots<j_k\le m} \sum_{\sigma\in S_k} \left(\frac{\partial v_{i_1}}{\partial u_{j_{\sigma(1)}}} \cdots \frac{\partial v_{i_k}}{\partial u_{j_{\sigma(k)}}} du^{j_{\sigma(1)}} \wedge \cdots \wedge du^{j_{\sigma(k)}}\right)\\ &= \sum_{1\le j_1<\ldots<j_k\le m} \sum_{\sigma\in S_k} \operatorname{sgn}(\sigma) \frac{\partial v_{i_1}}{\partial u_{j_{\sigma(1)}}} \cdots \frac{\partial v_{i_k}}{\partial u_{j_{\sigma(k)}}} du^{j_1} \wedge \cdots \wedge du^{j_k}\end{aligned}$$

By observing that $\sum_{\sigma\in S_k} \operatorname{sgn}(\sigma) \frac{\partial v_{i_1}}{\partial u_{j_{\sigma(1)}}} \cdots \frac{\partial v_{i_k}}{\partial u_{j_{\sigma(k)}}}$ is the determinant of $\left[\frac{\partial v_{i_p}}{\partial u_{j_q}}\right]_{1\le p,q\le k}$, the desired result (3.18) follows easily.

The second result (3.19) follows directly from (3.18). In case of $\dim M = \dim N = n$ and $k = n$, the only possible strictly increasing sequence $1 \le j_1 < \ldots < j_n \le n$ is $(j_1,\ldots,j_n) = (1,2,\ldots,n)$. □

Proposition 3.57. *Let $\Phi : M \to N$ be a smooth map between two smooth manifolds. For any $\omega \in \wedge^k T^*N$, we have:*

$$\Phi^*(d\omega) = d(\Phi^*\omega). \tag{3.20}$$

To be precise, we say $\Phi^*(d_N\omega) = d_M(\Phi^*\omega)$, where $d_N : \wedge^k T^*N \to \wedge^{k+1} T^*N$ and $d_M : \wedge^k T^*M \to \wedge^{k+1} T^*M$ are the exterior derivatives on N and M respectively.

Proof. Let $\{u_j\}$ and $\{v_i\}$ be local coordinates of M and N respectively. By linearity, it suffices to prove (3.20) for the case $\omega = f\, dv^{i_1} \wedge \cdots \wedge dv^{i_k}$ where f is a locally defined scalar function. The proof follows from computing both LHS and RHS of (3.20):

$$\begin{aligned}d\omega &= df \wedge dv^{i_1} \wedge \cdots \wedge dv^{i_k}\\ \Phi^*(d\omega) &= \Phi^*(df) \wedge \Phi^*(dv^{i_1}) \wedge \cdots \wedge \Phi^*(dv^{i_k})\\ &= d(\Phi^* f) \wedge d(\Phi^* v^{j_1}) \wedge \cdots \wedge d(\Phi^* v^{j_k}).\end{aligned}$$

Here we have used Exercise 3.50. On the other hand, we have:

$$\begin{aligned}\Phi^*\omega &= (\Phi^* f)\, \Phi^*(dv^{j_1}) \wedge \cdots \wedge \Phi^*(dv^{j_k})\\ &= (\Phi^* f)\, d(\Phi^* v^{i_1}) \wedge \cdots \wedge d(\Phi^* v^{i_k})\\ d(\Phi^*\omega) &= d(\Phi^* f) \wedge d(\Phi^* v^{i_1}) \wedge \cdots \wedge d(\Phi^* v^{i_k})\\ &\quad + \Phi^* f\, d\left(d(\Phi^* v^{i_1}) \wedge \cdots \wedge d(\Phi^* v^{i_k})\right)\end{aligned}$$

Since $d^2 = 0$, each of $d(\Phi^* v^{i_q})$ is a closed 1-form. By Proposition 3.42 (product rule) and induction, we can conclude that:

$$d\left(d(\Phi^* v^{i_1}) \wedge \cdots \wedge d(\Phi^* v^{i_k})\right) = 0$$

and so $d(\Phi^*\omega) = d(\Phi^* f) \wedge d(\Phi^* v^{i_1}) \wedge \cdots \wedge d(\Phi^* v^{i_k})$ as desired. □

Exercise 3.53. Show that the pull-back of any closed form is closed, and the pull-back of any exact form is exact.

Exercise 3.54. Consider the unit sphere $\mathbb{S}^2$ locally parametrized by

$$F(\theta, \varphi) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi).$$

Define a map $\Phi : \mathbb{S}^2 \to \mathbb{R}^3$ by $\Phi(x, y, z) = (xz, yz, z^2)$, and consider a 2-form $\omega = z\,dx \wedge dy$. Compute $d\omega$, $\Phi^*(d\omega)$, $\Phi^*\omega$ and $d(\Phi^*\omega)$, and verify they satisfy Proposition 3.57.

3.5.6. Unification of Green's, Stokes' and Divergence Theorems. Given a submanifold M^m in $\mathbb{R}^n$, a differential form on $\mathbb{R}^n$ induces a differential form on M^m. For example, let C be a smooth regular curve in $\mathbb{R}^3$ parametrized by $\mathsf{r}(t) = (x(t), y(t), z(t))$. The 1-form:

$$\alpha = \alpha_x\,dx + \alpha_y\,dy + \alpha_z\,dz$$

is *a priori* defined on $\mathbb{R}^3$, but we can regard the coordinates (x, y, z) as functions on the curve C parametrized by $\mathsf{r}(t)$, then we have $dx = \frac{dx}{dt}\,dt$ and similarly for dy and dz. As such, dx can now be regarded as a 1-form on C. Therefore, the 1-form α on $\mathbb{R}^3$ induces a 1-form α (abuse in notation) on C:

$$\begin{aligned}\alpha &= \alpha_x(\mathsf{r}(t))\frac{dx}{dt}\,dt + \alpha_y(\mathsf{r}(t))\frac{dy}{dt}\,dt + \alpha_z(\mathsf{r}(t))\frac{dz}{dt}\,dt \\ &= \left(\alpha_x(\mathsf{r}(t))\frac{dx}{dt} + \alpha_y(\mathsf{r}(t))\frac{dy}{dt} + \alpha_z(\mathsf{r}(t))\frac{dz}{dt}\right)dt\end{aligned}$$

In practice, there is often no issue of using α to denote both the 1-form on $\mathbb{R}^3$ and its induced 1-form on C. To be (overly) rigorous over notations, we can use the inclusion map $\iota : C \to \mathbb{R}^3$ to distinguish them. The 1-form α on $\mathbb{R}^3$ is transformed into a 1-form $\iota^*\alpha$ on C by the pull-back of ι. From the previous subsection, we learned that:

$$\iota^*(dx) = d(\iota^*x) = d(x \circ \iota).$$

Note that dx and $d(x \circ \iota)$ are different in a sense that $x \circ \iota : C \to \mathbb{R}$ has the curve C as its domain, while $x : \mathbb{R}^3 \to \mathbb{R}$ has $\mathbb{R}^3$ as its domain. Therefore, we have:

$$d(x \circ \iota) = \frac{d(x \circ \iota)}{dt}\,dt = \frac{dx}{dt}\,dt.$$

In short, we may use $\iota^*(dx) = \frac{dx}{dt}\,dt$ to distinguish it from dx if necessary. Similarly, we may use $\iota^*\alpha$ to denote the induced 1-form of α on C:

$$\iota^*\alpha = \left(\alpha_x(\mathsf{r}(t))\frac{dx}{dt} + \alpha_y(\mathsf{r}(t))\frac{dy}{dt} + \alpha_z(\mathsf{r}(t))\frac{dz}{dt}\right)dt.$$

An induced 1-form on a curve in $\mathbb{R}^3$ is related to line integrals in Multivariable Calculus. Recall that the 1-form $\alpha = \alpha_x\,dx + \alpha_y\,dy + \alpha_z\,dz$ corresponds to the vector field $\mathsf{F} = \alpha_x\mathsf{i} + \alpha_y\mathsf{j} + \alpha_z\mathsf{k}$ on $\mathbb{R}^3$. In Multivariable Calculus, we denote $d\mathsf{r} = dx\mathsf{i} + dy\mathsf{j} + dz\mathsf{k}$ and

$$\mathsf{F} \cdot d\mathsf{r} = \left(\alpha_x\mathsf{i} + \alpha_y\mathsf{j} + \alpha_z\mathsf{k}\right) \cdot (dx\mathsf{i} + dy\mathsf{j} + dz\mathsf{k}) = \alpha.$$

The line integral $\int_C \mathsf{F} \cdot d\mathsf{r}$ over the curve $C \subset \mathbb{R}^3$ can be written using differential form notations:

$$\int_C \mathsf{F} \cdot d\mathsf{r} = \int_C \alpha \quad \text{or more rigorously:} \quad \int_C \iota^*\alpha.$$

Now consider a regular surface $M \subset \mathbb{R}^3$. Suppose $\mathsf{F}(u,v) = (x(u,v), y(u,v), z(u,v))$ is a smooth local parametrization of M. Consider a vector $\mathsf{G} = \beta_x \mathsf{i} + \beta_y \mathsf{j} + \beta_z \mathsf{k}$ on $\mathbb{R}^3$ and its corresponding 2-form on $\mathbb{R}^3$:

$$\beta = \beta_x \, dy \wedge dz + \beta_y \, dz \wedge dx + \beta_z \, dx \wedge dy.$$

Denote $\iota : M \to \mathbb{R}^3$ the inclusion map. The induced 2-form $\iota^*\beta$ on M is in fact related to the surface flux of G through M. Let's explain why:

$$\begin{aligned}
\iota^*(dy \wedge dz) &= (\iota^* dy) \wedge (\iota^* dz) = d(y \circ \iota) \wedge d(z \circ \iota) \\
&= \left(\frac{\partial y}{\partial u} \, du + \frac{\partial y}{\partial v} \, dv \right) \wedge \left(\frac{\partial z}{\partial u} \, du + \frac{\partial z}{\partial v} \, dv \right) \\
&= \left(\frac{\partial y}{\partial u} \frac{\partial z}{\partial v} - \frac{\partial z}{\partial u} \frac{\partial y}{\partial v} \right) du \wedge dv \\
&= \det \frac{\partial(y,z)}{\partial(u,v)} \, du \wedge dv.
\end{aligned}$$

Similarly, we have:

$$\begin{aligned}
\iota^*(dz \wedge dx) &= \det \frac{\partial(z,x)}{\partial(u,v)} \, du \wedge dv \\
\iota^*(dx \wedge dy) &= \det \frac{\partial(x,y)}{\partial(u,v)} \, du \wedge dv
\end{aligned}$$

All these show:

$$\iota^*\beta = \left(\beta_x \det \frac{\partial(y,z)}{\partial(u,v)} + \beta_y \det \frac{\partial(z,x)}{\partial(u,v)} + \beta_z \det \frac{\partial(x,y)}{\partial(u,v)} \right) du \wedge dv$$

Compared with the flux element $\mathsf{G} \cdot \mathsf{N} \, dS$ in Multivariable Calculus:

$$\begin{aligned}
\mathsf{G} \cdot \mathsf{N} \, dS &= \underbrace{(\beta_x \mathsf{i} + \beta_y \mathsf{j} + \beta_z \mathsf{k})}_{\mathsf{G}} \cdot \underbrace{\frac{\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v}}{\left| \frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v} \right|}}_{\mathsf{N}} \underbrace{\left| \frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v} \right| dudv}_{dS} \\
&= (\beta_x \mathsf{i} + \beta_y \mathsf{j} + \beta_z \mathsf{k}) \cdot \left(\det \frac{\partial(y,z)}{\partial(u,v)} \mathsf{i} + \det \frac{\partial(z,x)}{\partial(u,v)} \mathsf{j} + \det \frac{\partial(x,y)}{\partial(u,v)} \mathsf{k} \right) \\
&= \left(\beta_x \det \frac{\partial(y,z)}{\partial(u,v)} + \beta_y \det \frac{\partial(z,x)}{\partial(u,v)} + \beta_z \det \frac{\partial(x,y)}{\partial(u,v)} \right) dudv,
\end{aligned}$$

the only difference is that $\iota^*\beta$ is in terms of the wedge product $du \wedge dv$ while the flux element $\mathsf{G} \cdot \mathsf{N} \, dS$ is in terms of $dudv$. Ignoring this minor difference (which will be addressed in the next chapter), the surface flux $\iint_M \mathsf{G} \cdot \mathsf{N} \, dS$ can be expressed in terms of differential forms in the following way:

$$\iint_M \mathsf{G} \cdot \mathsf{N} \, dS = \iint_M \beta \quad \text{or more rigorously:} \quad \iint_M \iota^*\beta$$

Recall that the classical Stokes' Theorem is related to line integrals of a curve and surface flux of a vector field. Based on the above discussion, we see that Stokes' Theorem can be restated in terms of differential forms. Consider the 1-form $\alpha = \alpha_x \, dx + \alpha_y \, dy + \alpha_z \, dz$ and its corresponding vector field $\mathsf{F} = \alpha_x \mathsf{i} + \alpha_y \mathsf{j} + \alpha_z \mathsf{k}$. We have already discussed that the 2-form $d\alpha$ corresponds to the vector field $\nabla \times \mathsf{F}$. Therefore, the surface flux of the vector field $\nabla \times \mathsf{F}$ through M can be expressed in terms of differential forms as:

$$\iint_M (\nabla \times \mathsf{F}) \cdot \mathsf{N} \, dS = \iint_M \iota^*(d\alpha) = \iint_M d(\iota^*\alpha).$$

If C is the boundary curve of M, then from our previous discussion we can write:

$$\int_C \mathsf{F} \cdot d\mathsf{r} = \int_C \iota^* \alpha.$$

The classical Stokes' Theorem asserts that:

$$\int_C \mathsf{F} \cdot \, d\mathsf{r} = \iint_M (\nabla \times \mathsf{F}) \cdot \mathsf{N} \, dS$$

which can be expressed in terms of differential form as:

$$\int_C \iota^* \alpha = \iint_M d(\iota^* \alpha) \quad \text{or simply:} \int_C \alpha = \iint_M d\alpha.$$

Due to this elegant way (although not very practical for physicists and engineers) of expressing Stokes' Theorem, we often denote the boundary of a surface M as ∂M, then the classical Stokes' Theorem can be expressed as:

$$\int_{\partial M} \alpha = \iint_M d\alpha.$$

Using differential forms, one can also express Divergence Theorem in Multivariable Calculus in a similar way as above. Let D be a solid region in $\mathbb{R}^3$ and ∂D be the boundary surface of D. Divergence Theorem in MATH 2023 asserts that:

$$\iint_{\partial D} \mathsf{G} \cdot \mathsf{N} \, dS = \iiint_D \nabla \cdot \mathsf{G} \, dV,$$

where $\mathsf{G} = \beta_x \mathsf{i} + \beta_y \mathsf{j} + \beta_z \mathsf{k}$. As discussed before, the LHS is $\iint_{\partial D} \beta$ where $\beta = \beta_x \, dy \wedge dz + \beta_y \, dz \wedge dx + \beta_z \, dx \wedge dy$. We have seen that:

$$d\beta = \left(\frac{\partial \beta_x}{\partial x} + \frac{\partial \beta_y}{\partial y} + \frac{\partial \beta_z}{\partial z} \right) dx \wedge dy \wedge dz,$$

which is (almost) the same as:

$$\nabla \cdot \mathsf{G} \, dV = \left(\frac{\partial \beta_x}{\partial x} + \frac{\partial \beta_y}{\partial y} + \frac{\partial \beta_z}{\partial z} \right) dxdydz.$$

Hence, the RHS of Divergence Theorem can be expressed as $\iiint_D d\beta$; and therefore we can rewrite Divergence Theorem as:

$$\iint_{\partial D} \beta = \iiint_D d\beta.$$

Again, the same expression! Stokes' and Divergence Theorems can therefore be unified. Green's Theorem can also be unified with Stokes' and Divergence Theorems as well. Try the exercise below:

Exercise 3.55. Let C be a simple closed smooth curve in $\mathbb{R}^2$ and R be the region enclosed by C in $\mathbb{R}^2$. Given a smooth vector field $\mathsf{F} = P\mathsf{i} + Q\mathsf{j}$ on $\mathbb{R}^2$, Green's Theorem asserts that:

$$\int_C \mathsf{F} \cdot d\mathsf{r} = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy.$$

Express Green's Theorem using the language of differential forms.

3.5.7. Differential Forms and Maxwell's Equations. The four Maxwell's equations are a set of partial differential equations that form the foundation of electromagnetism. Denote the components of the electric field E, magnetic field B, and current density J by

$$\begin{aligned}\mathsf{E} &= E_x\mathsf{i} + E_y\mathsf{j} + E_z\mathsf{k}\\ \mathsf{B} &= B_x\mathsf{i} + B_y\mathsf{j} + B_z\mathsf{k}\\ \mathsf{J} &= j_x\mathsf{i} + j_y\mathsf{j} + j_z\mathsf{k}\end{aligned}$$

All components of E, B and J are considered to be time-dependent. Denote ρ to be the charge density. The four Maxwell's equations assert that:

$$\begin{aligned}\nabla\cdot\mathsf{E} &= \rho & \nabla\cdot\mathsf{B} &= 0\\ \nabla\times\mathsf{E} &= -\frac{\partial\mathsf{B}}{\partial t} & \nabla\times\mathsf{B} &= \mathsf{J} + \frac{\partial\mathsf{E}}{\partial t}\end{aligned}$$

These four equations can be rewritten using differential forms in a very elegant way. Consider $\mathbb{R}^4$ with coordinates (t, x, y, z), which is also denoted as (x_0, x_1, x_2, x_3) in this problem. First we introduce the Minkowski Hodge-star operator $*$ on $\mathbb{R}^4$, which is a linear map taking p-forms on $\mathbb{R}^4$ to $(4-p)$-forms on $\mathbb{R}^4$. In particular, for 2-forms $\omega = dx^i \wedge dx^j$ (where $i, j = 0, 1, 2, 3$ and $i \neq j$), we define $*\omega$ to be the unique 2-form on $\mathbb{R}^4$ such that:

$$\omega\wedge *\omega = \begin{cases} dt\wedge dx\wedge dy\wedge dz & \text{if } i, j \neq 0\\ -dt\wedge dx\wedge dy\wedge dz & \text{otherwise.}\end{cases}$$

For instance, $*(dx\wedge dy) = dt\wedge dz$ since $dx\wedge dy\wedge dt\wedge dz = dt\wedge dx\wedge dy\wedge dz$ and there is no dt term in $dx\wedge dy$. On the other hand, $*(dt\wedge dx) = -dy\wedge dz$ since there is a dt term in $dt\wedge dx$. The operator $*$ then extends linearly to all 2-forms on $\mathbb{R}^4$.

Exercise 3.56. Compute each of the following:

$$\begin{array}{ccc} *(dt\wedge dx) & *(dt\wedge dy) & *(dt\wedge dz)\\ *(dx\wedge dy) & *(dy\wedge dz) & *(dz\wedge dx)\end{array}$$

To convert the Maxwell's equations using the language of differential forms, we define the following analogue of E, B, J and ρ using differential forms:

$$\begin{aligned}E &= E_x\,dx + E_y\,dy + E_z\,dz\\ B &= B_x\,dy\wedge dz + B_y\,dz\wedge dx + B_z\,dx\wedge dy\\ J &= -(j_x\,dy\wedge dz + j_y\,dz\wedge dx + j_z\,dx\wedge dy)\wedge dt + \rho\,dx\wedge dy\wedge dz\end{aligned}$$

Note that E_i's and B_j's may depend on t although there is no dt above. Define the 2-form:

$$F := B + E\wedge dt.$$

Exercise 3.57. Show that the four Maxwell's equations can be rewritten in an elegant way as:

$$\begin{aligned}dF &= 0\\ d(*F) &= J\end{aligned}$$

where d is the exterior derivative on $\mathbb{R}^4$.

3.5.8. Global Expressions of Exterior Derivatives. We defined exterior differentiation using local coordinates. In fact, using Lie derivatives, one can derive a global expression (i.e. without using local coordinates) of exterior derivatives on differential forms.

We first introduce Lie derivatives on $(p,0)$-tensors, which are similarly defined as those on 1-forms. Let T be a $(p,0)$-tensor and X be a vector field on M. Denote the flow map of X by Φ_t, then the Lie derivative of T along X at $p \in M$ is defined as:

$$(\mathcal{L}_X T)_p := \left.\frac{d}{dt}\right|_{t=0} \Phi_t^* \left(T_{\Phi_t(p)}\right).$$

Exercise 3.58. Guess the definition of Lie derivatives of a general (p,q)-tensor along a vector field X. Check any standard textbook to see if your guess is right.

Remark 3.58. On a regular surface M in $\mathbb{R}^3$ with the first fundamental form denoted by g, if X is a vector field on M such that $\mathcal{L}_X g = 0$, then we call X to be a *Killing vector field*. The geometric meaning of such an X is that g is invariant when M moves along the vector field X, or equivalently, g is symmetric in the direction of X. This concept of Killing vector fields can be generalize to Riemannian manifolds and is important in Differential Geometry and General Relativity, whenever *symmetry* plays an important role. □

Since the pull-back of a tensor product satisfies $\Phi^*(T \otimes S) = \Phi^* T \otimes \Phi^* S$, it is easy to show from definition that the Lie derivative satisfies the product rule:

$$\mathcal{L}_X(T \otimes S) = (\mathcal{L}_X T) \otimes S + T \otimes (\mathcal{L}_X S). \tag{3.21}$$

Exercise 3.59. Prove (3.21).

Since differential forms are simply linear combinations of tensor products, the definition of their Lie derivatives is the same as that for $(k,0)$-tensors. One nice fact about Lie derivatives on differential forms is so-called the *Cartan's magic formula*, which relates Lie derivatives and exterior derivatives. We first introduce the interior product:

Definition 3.59 (Interior Product)**.** Let α be a k-form (where $k \geq 2$) on a manifold M, and X be a vector field on M. Then, the *interior product* $i_X\alpha$ is a $(k-1)$-form defined as follows. For any vector fields $Y_1, \ldots, Y_{k-1}$ on M, we define:

$$(i_X\alpha)(Y_1, \ldots, Y_{k-1}) := \alpha(X, Y_1, \ldots, Y_{k-1}).$$

Example 3.60. In local coordinates, if a vector field X can be written as $X = \sum_{i=1}^n X^i \frac{\partial}{\partial u_i}$, then $i_X(du^j \wedge du^k)$ is an 1-form and we have:

$$\left(i_X(du^j \wedge du^k)\right)\left(\frac{\partial}{\partial u_l}\right) = (du^j \wedge du^k)\left(X, \frac{\partial}{\partial u_l}\right) = X^j \delta_{kl} - X^k \delta_{jl}.$$

In other words, we have:

$$i_X(du^j \wedge du^k) = \sum_{l=1}^n (X^j \delta_{kl} - X^k \delta_{jl})\, du^l = X^j\, du^k - X^k\, du^j.$$

□

Exercise 3.60. Let $X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial u_i}$ be a vector field on a manifold M with local coordinates $(u_1, \ldots, u_n)$. Derive the local expression of:

$$i_X(du^{j_1} \wedge du^{j_2} \wedge \cdots \wedge du^{j_k})$$

where $1 \leq j_1 < j_2 < \cdots < j_k \leq n$.

Now we are ready to present a beautiful and elegant formula due to Elie Cartan:

Proposition 3.61 (Cartan's Magic Formula)**.** *Let X be a smooth vector field on a manifold M, then for any differential k-form ω, we have:*

$$\mathcal{L}_X \omega = i_X(d\omega) + d(i_X\omega) \tag{3.22}$$

Proof. The proof is by induction on k, the degree of ω. We first show that (3.22) holds for 1-forms.

Consider $\omega = \sum_{j=1}^{n} \omega_j \, du^j$, we have already computed in (3.11) that:

$$\mathcal{L}_X \omega = \sum_{i,j=1}^{n} \left(X^i \frac{\partial \omega_j}{\partial u_i} + \omega_i \frac{\partial X^i}{\partial u_j} \right) du^j.$$

Next we verify it is equal to the RHS of (3.22).

$$d\omega = \sum_{i,j=1}^{n} \frac{\partial \omega_j}{\partial u_i} \, du^i \wedge du^j$$

From Example 3.60, we have

$$i_X(d\omega) = \sum_{i,j=1}^{n} \frac{\partial \omega_j}{\partial u_i} i_X(du^i \wedge du^j) = \sum_{i,j=1}^{n} \frac{\partial \omega_j}{\partial u_i} (X^i \, du^j - X^j \, du^i).$$

Moreover, we have

$$i_X \omega = \omega(X) = \sum_{j=1}^{n} X^j \omega_j$$

$$d(i_X\omega) = \sum_{i,j=1}^{n} \left(\frac{\partial X^j}{\partial u_i} \omega_j + X^j \frac{\partial \omega_j}{\partial u_i} \right) du^i$$

and it follows easily that:

$$i_X(d\omega) + d(i_X\omega) = \sum_{i,j=1}^{n} X^i \frac{\partial \omega_j}{\partial u_i} \, du^j + \sum_{i,j=1}^{n} \omega_j \frac{\partial X^j}{\partial u_i} \, du^i$$

which is exactly $\mathcal{L}_X \omega$ after relabelling indices.

Now that (3.22) holds for 1-form. To complete the inductive proof, we just need to show that if (3.22) holds for both differential forms ω and σ, then it also holds for $\omega \wedge \sigma$. It is left as an exercise for readers. □

Exercise 3.61. Complete the above inductive proof. [Note: the proof is somewhat algebraic.]

Exercise 3.62. Show that if ω is closed, then $\mathcal{L}_X \omega$ is exact for any vector field X.

The purpose of introducing Cartan's magic formula is it gives a coordinate-free expression of exterior derivatives. Consider a 1-form ω, and two vector fields X and Y. Then, from (3.22), we have:

$$(\mathcal{L}_X\omega)(Y) = \left(i_X(d\omega)\right)(Y) + \left(d(i_X\omega)\right)(Y),$$

which, from the definition of i_X and (3.19), can be simplified to:

$$X(\omega(Y)) - \omega(\mathcal{L}_X Y) = (d\omega)(X,Y) + d(\omega(X))(Y).$$

As $\omega(X)$ is a scalar function, we also have:

$$d(\omega(X))(Y) = Y\left(\omega(X)\right).$$

[Note that generally, $(df)(Y) = Y(f)$ for any scalar function f.]

Finally, we get:

$$(d\omega)(X,Y) = X\left(\omega(Y)\right) - Y\left(\omega(X)\right) - \omega([X,Y]) \tag{3.23}$$

for any vector fields X and Y. This is a global formula for $d\omega$ as it does not involve any local coordinates.

The expression (3.23) can be generalized to k-forms ω. The proof is by induction and the Cartan's magic formula again. For any k-form ω, and vector fields $X_0, X_1, \ldots, X_k$, we have:

$$\begin{aligned}
&(d\omega)(X_0, X_1, \cdots, X_k) \\
&= \sum_{i=0}^{k}(-1)^i X_i(\omega(X_0, \cdots, \hat{X}_i, \cdots, X_k)) \\
&\quad + \sum_{0\leq i<j\leq k}(-1)^{i+j}\omega\left([X_i, X_j], X_0, \cdots, \hat{X}_i, \cdots, \hat{X}_j, \cdots, X_k\right).
\end{aligned}$$

Readers interested in the proof may consult [**Lee13**, P.370, Proposition 14.32].

Chapter 4

Generalized Stokes' Theorem

> "It is very difficult for us, placed as we have been from earliest childhood in a condition of training, to say what would have been our feelings had such training never taken place."
>
> *Sir George Stokes, 1st Baronet*

4.1. Manifolds with Boundary

We have seen in the Chapter 3 that Green's, Stokes' and Divergence Theorem in Multivariable Calculus can be unified together using the language of differential forms. In this chapter, we will generalize Stokes' Theorem to higher dimensional and abstract manifolds.

These classic theorems and their generalizations concern about an integral over a manifold with an integral over its boundary. In this section, we will first rigorously define the notion of a *boundary* for abstract manifolds. Heuristically, an *interior* point of a manifold locally looks like a ball in Euclidean space, whereas a *boundary* point locally looks like an upper-half space.

4.1.1. Smooth Functions on Upper-Half Spaces. From now on, we denote $\mathbb{R}^n_+ := \{(u_1, \ldots, u_n) \in \mathbb{R}^n : u_n \geq 0\}$ which is the upper-half space of $\mathbb{R}^n$. Under the subspace topology, we say a subset $V \subset \mathbb{R}^n_+$ is *open in* $\mathbb{R}^n_+$ if there exists a set $\widetilde{V} \subset \mathbb{R}^n$ open in $\mathbb{R}^n$ such that $V = \widetilde{V} \cap \mathbb{R}^n_+$. It is intuitively clear that if $V \subset \mathbb{R}^n_+$ is disjoint from the subspace $\{u_n = 0\}$ of $\mathbb{R}^n$, then V is open in $\mathbb{R}^n_+$ if and only if V is open in $\mathbb{R}^n$.

Now consider a set $V \subset \mathbb{R}^n_+$ which is open in $\mathbb{R}^n_+$ and that $V \cap \{u_n = 0\} \neq \emptyset$. We need to first develop a notion of differentiability for functions such an V as their domain. Given a vector-valued function $\mathsf{G} : V \to \mathbb{R}^m$, then near a point $\mathsf{u} \in V \cap \{u_n = 0\}$, we can only approach u from one side only, namely from directions with positive u_n-coordinates. The usual definition of differentiability does not apply at such a point, so we define:

Definition 4.1 (Functions of Class C^k on $\mathbb{R}^n_+$)**.** Let $V \subset \mathbb{R}^n_+$ be open in $\mathbb{R}^n_+$ and that $V \cap \{u_n = 0\} \neq \varnothing$. Consider a vector-valued function $\mathsf{G} : V \to \mathbb{R}^m$. We say G is C^k (resp. smooth) at $\mathsf{u} \in V \cap \{u_n = 0\}$ if there exists a C^k (resp. smooth) local extension $\widetilde{\mathsf{G}} : B_\varepsilon(\mathsf{u}) \to \mathbb{R}^m$ such that $\widetilde{\mathsf{G}}(\mathsf{y}) = \mathsf{G}(\mathsf{y})$ for any $\mathsf{y} \in B_\varepsilon(\mathsf{u}) \cap V$. Here $B_\varepsilon(\mathsf{u}) \subset \mathbb{R}^n$ refers to an open ball in $\mathbb{R}^n$.

If G is C^k (resp. smooth) at every $\mathsf{u} \in V$ (including those points with $u_n > 0$), then we say G is C^k (resp. smooth) on V.

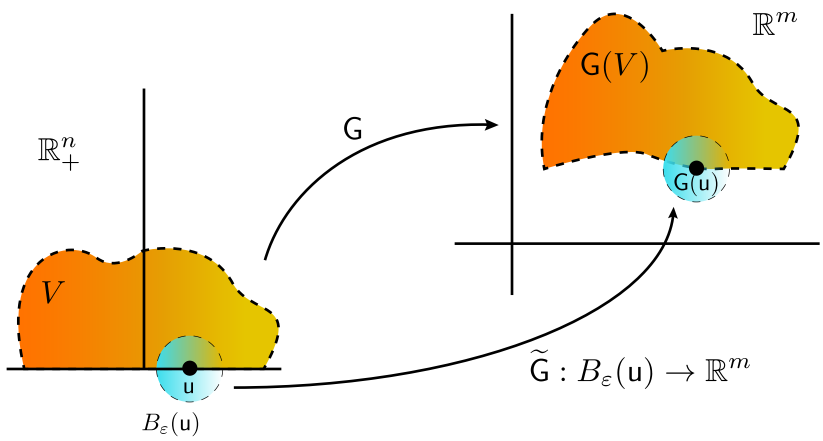


Figure 4.1. G is C^k at u if there exists a local extension $\widetilde{\mathsf{G}}$ near u.

Example 4.2. Let $V = \{(x, y) : y \geq 0 \text{ and } x^2 + y^2 < 1\}$, which is an open set in $\mathbb{R}^2_+$ since $V = \underbrace{\{(x, y) : x^2 + y^2 < 1\}}_{\text{open in } \mathbb{R}^2} \cap \mathbb{R}^2_+$. Then $f(x, y) : V \to \mathbb{R}$ defined by $f(x, y) = \sqrt{1 - x^2 - y^2}$ is a smooth function on V since $\sqrt{1 - x^2 - y^2}$ is smoothly on the whole ball $x^2 + y^2 < 1$.

However, the function $g : V \to \mathbb{R}$ defined by $g(x, y) = \sqrt{y}$ is not smooth at every point on the y-axis because $\frac{\partial g}{\partial y} \to \infty$ as $y \to 0^+$. Any extension $\widetilde{g}$ of g will agree with g on the upper-half plane, and hence will also be true that $\frac{\partial \widetilde{g}}{\partial y} \to \infty$ as $y \to 0^+$, which is sufficient to argue that such $\widetilde{g}$ is not smooth. □

Exercise 4.1. Consider $f : \mathbb{R}^2_+ \to \mathbb{R}$ defined by $f(x, y) = |x|$. Is f smooth on $\mathbb{R}^2_+$? If not, at which point(s) in $\mathbb{R}^2_+$ is f not smooth? Do the same for $g : \mathbb{R}^2_+ \to \mathbb{R}$ defined by $g(x, y) = |y|$.

4.1.2. Boundary of Manifolds. After understanding the definition of a smooth function when defined on subsets of the upper-half space, we are ready to introduce the notion of manifolds with boundary:

Definition 4.3 (Manifolds with Boundary)**.** We say M is a smooth *manifold with boundary* if there exist two families of local parametrizations $\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M$ where $\mathcal{U}_\alpha$ is open in $\mathbb{R}^n$, and $\mathsf{G}_\beta : \mathcal{V}_\beta \to M$ where $\mathcal{V}_\beta$ is open in $\mathbb{R}^n_+$ such that every F_α and G_β is a homeomorphism between its domain and image, and that the transition functions of all types:

$$\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_{\alpha'} \qquad \mathsf{F}_\alpha^{-1} \circ \mathsf{G}_\beta \qquad \mathsf{G}_\beta^{-1} \circ \mathsf{G}_{\beta'} \qquad \mathsf{G}_\beta^{-1} \circ \mathsf{F}_\alpha$$

are smooth on the overlapping domain for any α, α', β and β'.

Moreover, we denote and define the boundary of M by:

$$\partial M := \bigcup_\beta \{\mathsf{G}_\beta(u_1, \ldots, u_{n-1}, 0) : (u_1, \ldots, u_{n-1}, 0) \in \mathcal{V}_\beta\}.$$

Remark 4.4. In this course, we will call these F_α's to be local parametrizations of *interior type*, and these G_β's to be local parametrizations of *boundary type*. □

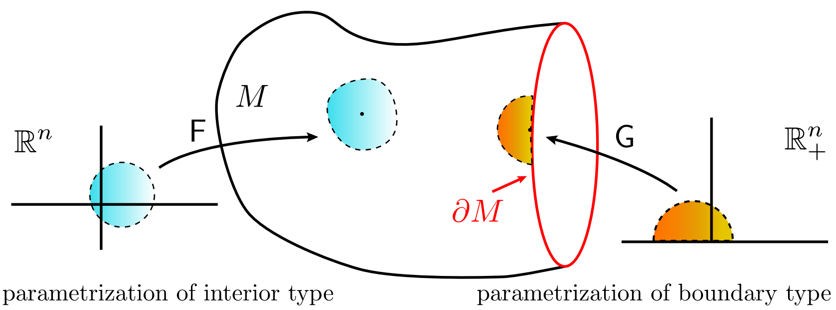


Figure 4.2. A manifold with boundary

Example 4.5. Consider the solid ball $\mathbb{B}^2 := \{\mathsf{x} \in \mathbb{R}^2 : |\mathsf{x}| \leq 1\}$. It can be locally parametrized using polar coordinates by:

$$\mathsf{G} : (0, 2\pi) \times [0, 1) \to \mathbb{B}^2$$
$$\mathsf{G}(\theta, r) := (1 - r)(\cos\theta, \sin\theta)$$

Note that the domain of G can be regarded as a subset

$$\mathcal{V} := \{(\theta, r) : \theta \in (0, 2\pi) \text{ and } 0 \leq r < 1\} \subset \mathbb{R}^2_+.$$

Here we used $1 - r$ instead of r so that the boundary of $\mathbb{B}^2$ has zero r-coordinate, and the interior of $\mathbb{B}^2$ has positive r-coordinate.

Note that the image of G does not cover the whole solid ball $\mathbb{B}^2$. Precisely, the image of G is $\mathbb{B}^2 \backslash \{\text{non-negative } x\text{-axis}\}$. In order to complete the proof that $\mathbb{B}^2$ is a manifold with boundary, we cover $\mathbb{B}^2$ by two more local parametrizations:

$$\widetilde{\mathsf{G}} : (-\pi, \pi) \times [0, 1) \to \mathbb{B}^2$$
$$\widetilde{\mathsf{G}}(\theta, r) := (1 - r)(\cos\theta, \sin\theta)$$

and also the inclusion map $\iota : \{\mathsf{u} \in \mathbb{R}^2 : |\mathsf{u}| < 1\} \to \mathbb{B}^2$. We need to show that the transition maps are smooth. There are six possible transition maps:

$$\widetilde{\mathsf{G}}^{-1} \circ \mathsf{G}, \quad \mathsf{G}^{-1} \circ \widetilde{\mathsf{G}}, \quad \iota^{-1} \circ \mathsf{G}, \quad \iota^{-1} \circ \widetilde{\mathsf{G}}, \quad \mathsf{G}^{-1} \circ \iota, \quad \text{and} \quad \widetilde{\mathsf{G}}^{-1} \circ \iota.$$

The first one is given by (we leave it as an exercise for computing these transition maps):

$$\widetilde{\mathsf{G}}^{-1} \circ \mathsf{G} : ((0,\pi) \cup (\pi, 2\pi)) \times [0,1) \to ((-\pi, 0) \cup (0,\pi)) \times [0,1)$$

$$\widetilde{\mathsf{G}}^{-1} \circ \mathsf{G}(\theta, r) = \begin{cases} (\theta, r) & \text{if } \theta \in (0,\pi) \\ (\theta - 2\pi, r) & \text{if } \theta \in (\pi, 2\pi) \end{cases}$$

which can be smoothly extended to the domain $((0,\pi) \cup (\pi, 2\pi)) \times (-1,1)$. Therefore, $\widetilde{\mathsf{G}}^{-1} \circ \mathsf{G}$ is smooth. The second transition map $\mathsf{G}^{-1} \circ \widetilde{\mathsf{G}}$ can be computed and verified to be smooth in a similar way.

For $\iota^{-1} \circ \mathsf{G}$, by examining the overlap part of ι and G on $\mathbb{B}^2$, we see that the domain of the transition map is an open set $(0, 2\pi) \times (0,1)$ in $\mathbb{R}^2$. On this domain, $\iota^{-1} \circ \mathsf{G}$ is essentially G, which is clearly smooth. Similar for $\iota^{-1} \circ \widetilde{\mathsf{G}}$.

To show $\mathsf{G}^{-1} \circ \iota$ is smooth, we use the Inverse Function Theorem. The domain of $\iota^{-1} \circ \mathsf{G}$ is $(0, 2\pi) \times (0,1)$. By writing $(x, y) = \iota^{-1} \circ \mathsf{G}(\theta, r) = (1-r)(\cos\theta, \sin\theta)$, we check that on the domain of $\iota^{-1} \circ \mathsf{G}$, we have:

$$\det \frac{\partial(x,y)}{\partial(\theta, r)} = 1 - r \neq 0.$$

Therefore, the inverse $\mathsf{G}^{-1} \circ \iota$ is smooth. Similar for $\widetilde{\mathsf{G}}^{-1} \circ \iota$.

Combining all of the above verifications, we conclude that $\mathbb{B}^2$ is a 2-dimensional manifold with boundary. The boundary $\partial\mathbb{B}^2$ is given by points with zero r-coordinates, namely the unit circle $\{|\mathsf{x}| = 1\}$. □

Exercise 4.2. Compute all transition maps

$$\widetilde{\mathsf{G}}^{-1} \circ \mathsf{G}, \quad \mathsf{G}^{-1} \circ \widetilde{\mathsf{G}}, \quad \iota^{-1} \circ \mathsf{G}, \quad \iota^{-1} \circ \widetilde{\mathsf{G}}, \quad \mathsf{G}^{-1} \circ \iota, \quad \text{and} \quad \widetilde{\mathsf{G}}^{-1} \circ \iota$$

in Example 4.5. Indicate clearly their domains, and verify that they are smooth on their domains.

Exercise 4.3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth scalar function. The region in $\mathbb{R}^{n+1}$ above the graph of f is given by:

$$\Gamma_f := \{(u_1, \ldots, u_{n+1}) \in \mathbb{R}^{n+1} : u_{n+1} \geq f(u_1, \ldots, u_n)\}.$$

Show that Γ_f is an n-dimensional manifold with boundary, and the boundary $\partial\Gamma_f$ is the graph of f in $\mathbb{R}^{n+1}$.

Exercise 4.4. Show that ∂M (assumed non-empty) of any n-dimensional manifold M is an $(n-1)$-dimensional manifold without boundary.

From the above example and exercise, we see that verifying a set is a manifold with boundary may be cumbersome. The following proposition provides us with a very efficient way to do so.

Proposition 4.6. *Let $f : M^m \to \mathbb{R}$ be a smooth function from a smooth manifold M. Suppose $c \in \mathbb{R}$ such that the set $\Sigma := f^{-1}([c, \infty))$ is non-empty and that f is a submersion at any $p \in f^{-1}(c)$, then the set Σ is an m-dimensional manifold with boundary. The boundary $\partial\Sigma$ is given by $f^{-1}(c)$.*

Proof. We need to construct local parametrizations for the set Σ. Given any point $p \in \Sigma$, then by the definition of Σ, we have $f(p) > c$ or $f(p) = c$.

For the former case $f(p) > c$, we are going to show that near p there is a local parametrization of Σ of interior type. Regarding p as a point in the manifold M, there exists a smooth local parametrization $\mathsf{F} : \mathcal{U} \subset \mathbb{R}^n \to M$ of M covering p. We argue that such a local parametrization of M induces naturally a local parametrization of Σ near p. Note that f is continuous and so $f^{-1}(c, \infty)$ is an open set of M containing p. Denote $\mathcal{O} = f^{-1}(c, \infty)$, then F restricted to $\mathcal{U} \cap \mathsf{F}^{-1}(\mathcal{O})$ will have its image in $\mathcal{O} \subset \Sigma$, and so is a local parametrization of Σ near p.

For the later case $f(p) = c$, we are going to show that near p there is a local parametrization of Σ of boundary type. Since f is a submersion at p, by the Submersion Theorem (Theorem 2.48) there exist a local parametrization $\mathsf{G} : \widetilde{\mathcal{U}} \to M$ of M near p, and a local parametrization H of $\mathbb{R}$ near c such that $\mathsf{G}(0) = p$ and $\mathsf{H}(0) = c$, and:

$$\mathsf{H}^{-1} \circ f \circ \mathsf{G}(u_1, \ldots, u_m) = u_m.$$

Without loss of generality, we assume that H is an increasing function near 0. We argue that by restricting the domain of G to $\mathcal{U} \cap \{u_m \geq 0\}$, which is an open set in $\mathbb{R}^m_+$, the restricted G is a boundary-type local parametrization of Σ near p. To argue this, we note that:

$$f(\mathsf{G}(u_1, \ldots, u_m)) = \mathsf{H}(u_m) \geq \mathsf{H}(0) = c \quad \text{whenever } u_m \geq 0.$$

Therefore, $\mathsf{G}(u_1, \ldots, u_m) \in f^{-1}([c, \infty)) = \Sigma$ whenever $u_m \geq 0$, and so G (when restricted to $\mathcal{U} \cap \{u_m \geq 0\}$) is a local parametrization of Σ.

Since all local parametrizations F and G of Σ constructed above are induced from local parametrizations of M (whether it is of interior or boundary type), their transition maps are all smooth. This shows Σ is an m-dimensional manifold with boundary. To identify the boundary, we note that for any boundary-type local parametrization G constructed above, we have:

$$\mathsf{H}^{-1} \circ f \circ \mathsf{G}(u_1, \ldots, u_{m-1}, 0) = 0$$

and so $f(\mathsf{G}(u_1, \ldots, u_{m-1})) = \mathsf{H}(0) = c$, and therefore:

$$\mathsf{G}(u_1, \ldots, u_{m-1}, 0) \in f^{-1}(c).$$

This show $\partial\Sigma \subset f^{-1}(c)$. The other inclusion $f^{-1}(c) \subset \partial\Sigma$ follows from the fact that for any $p \in f^{-1}(c)$, the boundary-type local parametrization G has the property that $\mathsf{G}(0) = p$ (and hence $p = \mathsf{G}(0, \ldots, 0, 0) \in \partial\Sigma$). □

Remark 4.7. It is worthwhile to note that the above proof only requires that f is a submersion at any $p \in f^{-1}(c)$, and we do not require that it is a submersion at any $p \in \Sigma = f^{-1}([c, \infty))$. Furthermore, the codomain of f is $\mathbb{R}$ which has dimension 1, hence f is a submersion at p if and only if the tangent map $(f_*)_p$ at p is non-zero – and so it is very easy to verify this condition. □

With the help of Proposition 4.6, one can show many sets are manifolds with boundary by picking a suitable submersion f.

Example 4.8. The n-dimensional ball $\mathbb{B}^n = \{\mathsf{x} \in \mathbb{R}^n : |\mathsf{x}| \leq 1\}$ is an n-manifold with boundary. To argue this, let $f : \mathbb{R}^n \to \mathbb{R}$ be the function:

$$f(\mathsf{x}) = 1 - |\mathsf{x}|^2.$$

Then $\mathbb{B}^n = f^{-1}([0, \infty))$.

The tangent map f_* is represented by the matrix:

$$[f_*] = \left[\frac{\partial f}{\partial x_1}, \ \cdots, \ \frac{\partial f}{\partial x_n}\right] = -2\,[x_1, \ \cdots, \ x_n]$$

which is surjective if and only if $(x_1, \ldots, x_n) \neq (0, \ldots, 0)$. For any $\mathsf{x} \in f^{-1}(0)$, we have $|\mathsf{x}|^2 = 1$ and so in particular $\mathsf{x} \neq 0$. Therefore, f is a submersion at every $\mathsf{x} \in f^{-1}(0)$. By Proposition 4.6, we proved $\mathbb{B}^n = f^{-1}([0, \infty))$ is an n-dimensional manifold with boundary, and the boundary is $f^{-1}(0) = \{\mathsf{x} \in \mathbb{R}^n : |\mathsf{x}| = 1\}$, i.e. the unit circle. $\square$

Exercise 4.5. Suppose $f : M^m \to \mathbb{R}$ is a smooth function defined on a smooth manifold M. Suppose $a, b \in \mathbb{R}$ such that $\Sigma := f^{-1}([a, b])$ is non-empty, and that f is a submersion at any $p \in f^{-1}(a)$ and any $q \in f^{-1}(b)$. Show that Σ is an m-manifold with boundary, and $\partial\Sigma = f^{-1}(a) \cup f^{-1}(b)$.

4.1.3. Tangent Spaces at Boundary Points. On a manifold M^n without boundary, the tangent space T_pM at p is the span of partial differential operators $\left\{ \frac{\partial}{\partial u_i}\bigg|_p \right\}_{i=1}^n$, where $(u_1, \ldots, u_n)$ are local coordinates of a parametrization $\mathsf{F}(u_1, \ldots, u_n)$ near p.

Now on a manifold M^n with boundary, near any boundary point $p \in \partial M^n$ there exists a local parametrization $\mathsf{G}(u_1, \ldots, u_n) : \mathcal{V} \subset \mathbb{R}^n_+ \to M$ of boundary type. Although G is only defined when $u_n \geq 0$, we *still* define T_pM to be the span of $\left\{ \frac{\partial}{\partial u_i}\bigg|_p \right\}_{i=1}^n$. Although such a definition of T_pM (when $p \in \partial M$) is a bit counter-intuitive, the perk is that T_pM is still a vector space. Given a vector $V \in T_pM$ with coefficients:

$$V = \sum_{i=1}^n V^i \frac{\partial}{\partial u_i}\bigg|_p.$$

We say that V is *inward-pointing* if $V^n > 0$; and *outward-pointing* if $V^n < 0$.

Furthermore, the tangent space $T_p(\partial M)$ of the boundary manifold ∂M at p can be regarded as a subspace of T_pM:

$$T_p(\partial M) = \operatorname{span}\left\{ \frac{\partial}{\partial u_i}\bigg|_p \right\}_{i=1}^{n-1} \subset T_pM.$$

4.2. Orientability

In Multivariable Calculus, we learned (or was told) that Stokes' Theorem requires the surface to be orientable, meaning that the unit normal vector $\hat{\mathsf{n}}$ varies continuously on the surface. The Möbius strip is an example of *non-orientable* surface.

Now we are talking about abstract manifolds which may not sit inside any Euclidean space, and so it does not make sense to define *normal vectors* to the manifold. Even when the manifold M is a subset of $\mathbb{R}^n$, if the dimension of the manifold is $\dim M \leq n-2$, the manifold does not have a unique normal vector direction. As such, in order to generalize the notion of orientability of abstract manifolds, we need to seek a reasonable definition without using normal vectors.

In this section, we first show that for hypersurfaces M^n in $\mathbb{R}^{n+1}$, the notion of orientability using normal vectors is equivalent to another notion using transition maps. Then, we extend the notion of orientability to abstract manifolds using transition maps.

4.2.1. Orientable Hypersurfaces.

To begin, we first state the definition of orientable hypersurfaces in $\mathbb{R}^{n+1}$:

Definition 4.9 (Orientable Hypersurfaces)**.** A regular hypersurface M^n in $\mathbb{R}^{n+1}$ is said to be *orientable* if there exists a continuous unit normal vector $\hat{\mathsf{n}}$ defined on the whole M^n

Let's explore the above definition a bit in the easy case $n = 2$. Given a regular surface M^2 in $\mathbb{R}^3$ with a local parametrization $(x, y, z) = \mathsf{F}(u_1, u_2) : \mathcal{U} \to M$, one can find a normal vector to the surface by taking cross product:

$$\frac{\partial \mathsf{F}}{\partial u_1} \times \frac{\partial \mathsf{F}}{\partial u_2} = \det \frac{\partial(y,z)}{\partial(u_1,u_2)}\mathsf{i} + \det \frac{\partial(z,x)}{\partial(u_1,u_2)}\mathsf{j} + \det \frac{\partial(x,y)}{\partial(u_1,u_2)}\mathsf{k}$$

and hence the unit normal along this direction is given by:

$$\hat{\mathsf{n}}_\mathsf{F} = \frac{\det \frac{\partial(y,z)}{\partial(u_1,u_2)}\mathsf{i} + \det \frac{\partial(z,x)}{\partial(u_1,u_2)}\mathsf{j} + \det \frac{\partial(x,y)}{\partial(u_1,u_2)}\mathsf{k}}{\left|\det \frac{\partial(y,z)}{\partial(u_1,u_2)}\mathsf{i} + \det \frac{\partial(z,x)}{\partial(u_1,u_2)}\mathsf{j} + \det \frac{\partial(x,y)}{\partial(u_1,u_2)}\mathsf{k}\right|} \quad \text{on } \mathsf{F}(\mathcal{U}).$$

Note that the above $\hat{\mathsf{n}}$ is defined locally on the domain $\mathsf{F}(\mathcal{U})$.

Now given another local parametrization $(x, y, z) = \mathsf{G}(v_1, v_2) : \mathcal{V} \to M$, one can find a unit normal using G as well:

$$\hat{\mathsf{n}}_\mathsf{G} = \frac{\det \frac{\partial(y,z)}{\partial(v_1,v_2)}\mathsf{i} + \det \frac{\partial(z,x)}{\partial(v_1,v_2)}\mathsf{j} + \det \frac{\partial(x,y)}{\partial(v_1,v_2)}\mathsf{k}}{\left|\det \frac{\partial(y,z)}{\partial(v_1,v_2)}\mathsf{i} + \det \frac{\partial(z,x)}{\partial(v_1,v_2)}\mathsf{j} + \det \frac{\partial(x,y)}{\partial(v_1,v_2)}\mathsf{k}\right|} \quad \text{on } \mathsf{G}(\mathcal{V}).$$

Using the chain rule, we have the following relation between the Jacobian determinants:

$$\det \frac{\partial(*,**)}{\partial(v_1,v_2)} = \det \frac{\partial(u_1,u_2)}{\partial(v_1,v_2)} \det \frac{\partial(*,**)}{\partial(u_1,u_2)}$$

(here $*$ and $**$ mean any of the x, y and z) and therefore $\hat{\mathsf{n}}_\mathsf{F}$ and $\hat{\mathsf{n}}_\mathsf{G}$ are related by:

$$\hat{\mathsf{n}}_\mathsf{G} = \frac{\det \frac{\partial(u_1,u_2)}{\partial(v_1,v_2)}}{\left|\det \frac{\partial(u_1,u_2)}{\partial(v_1,v_2)}\right|}\hat{\mathsf{n}}_\mathsf{F}.$$

Therefore, if there is an overlap between local coordinates (u_1, u_2) and (v_1, v_2), the unit normal vectors $\hat{\mathsf{n}}_\mathsf{F}$ and $\hat{\mathsf{n}}_\mathsf{G}$ agree with each other on the overlap $\mathsf{F}(\mathcal{U}) \cap \mathsf{G}(\mathcal{V})$ if and only if $\det \dfrac{\partial(u_1,u_2)}{\partial(v_1,v_2)} > 0$ (equivalently, $\det D(\mathsf{F}^{-1} \circ \mathsf{G}) > 0$).

From above, we see that consistency of unit normal vector on different local coordinate charts is closely related to the positivity of the determinants of transition maps. A consistence choice of unit normal vector $\hat{\mathsf{n}}$ exists if and only if it is possible to pick a family of local parametrizations $\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M^2$ covering the whole M such that $\det D(\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha) > 0$ on $\mathsf{F}_\alpha^{-1}\left(\mathsf{F}_\alpha(\mathcal{U}_\alpha) \cap \mathsf{F}_\beta(\mathcal{U}_\beta)\right)$ for any α and β in the family. The notion of normal vectors makes sense only for hypersurfaces in $\mathbb{R}^n$, while the notion of transition maps can extend to any abstract manifold.

Note that given two local parametrizations $\mathsf{F}(u_1, u_2)$ and $\mathsf{G}(v_1, v_2)$, it is *not* always possible to make sure $\det \dfrac{\partial(u_1, u_2)}{\partial(v_1, v_2)} > 0$ on the overlap even by switching v_1 and v_2. It is because it sometimes happens that the overlap $\mathsf{F}(\mathcal{U}) \cap \mathsf{G}(\mathcal{V})$ is a disjoint union of two open sets. If on one open set the determinant is positive, and on another one the determinant is negative, then switching v_1 and v_2 cannot make the determinant positive on both open sets. Let's illustrate this issue through two contrasting examples: the cylinder and the Möbius strip:

Example 4.10. The unit cylinder Σ^2 in $\mathbb{R}^3$ can be covered by two local parametrizations:

$$\mathsf{F} : (0, 2\pi) \times \mathbb{R} \to \Sigma^2 \qquad \widetilde{\mathsf{F}} : (-\pi, \pi) \times \mathbb{R} \to \Sigma^2$$

$$\mathsf{F}(\theta, z) := (\cos\theta, \sin\theta, z) \qquad \widetilde{\mathsf{F}}(\widetilde{\theta}, \widetilde{z}) := (\cos\widetilde{\theta}, \sin\widetilde{\theta}, \widetilde{z})$$

Then, the transition map $\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F}$ is defined on a disconnected domain $\theta \in (0, \pi) \cup (\pi, 2\pi)$ and $z \in \mathbb{R}$, and it is given by:

$$\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F}(\theta, z) = \begin{cases} (\theta, z) & \text{if } \theta \in (0, \pi) \\ (\theta - 2\pi, z) & \text{if } \theta \in (\pi, 2\pi) \end{cases}$$

By direct computations, the Jacobian of this transition map is given by:

$$D(\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F})(\theta, z) = I$$

in either case $\theta \in (0, \pi)$ or $\theta \in (\pi, 2\pi)$. Therefore, $\det D(\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F}) > 0$ on the overlap.

The unit normal vectors defined using these F and $\widetilde{\mathsf{F}}$:

$$\hat{\mathsf{n}}_{\mathsf{F}} = \frac{\frac{\partial \mathsf{F}}{\partial r} \times \frac{\partial \mathsf{F}}{\partial \theta}}{\left|\frac{\partial \mathsf{F}}{\partial r} \times \frac{\partial \mathsf{F}}{\partial \theta}\right|} \quad \text{on } \mathsf{F}((0, 2\pi) \times \mathbb{R})$$

$$\hat{\mathsf{n}}_{\widetilde{\mathsf{F}}} = \frac{\frac{\partial \widetilde{\mathsf{F}}}{\partial \widetilde{r}} \times \frac{\partial \widetilde{\mathsf{F}}}{\partial \widetilde{\theta}}}{\left|\frac{\partial \widetilde{\mathsf{F}}}{\partial \widetilde{r}} \times \frac{\partial \widetilde{\mathsf{F}}}{\partial \widetilde{\theta}}\right|} \quad \text{on } \widetilde{\mathsf{F}}((-\pi, \pi) \times \mathbb{R})$$

will agree with each other on the overlap. Therefore, it defines a *global* continuous unit normal vector across the whole cylinder. □

Example 4.11. The Möbius strip Σ^2 in $\mathbb{R}^3$ can be covered by two local parametrizations:

$$\mathsf{F} : (-1, 1) \times (0, 2\pi) \to \Sigma^2 \qquad \widetilde{\mathsf{F}} : (-1, 1) \times (-\pi, \pi) \to \Sigma^2$$

$$\mathsf{F}(u, \theta) = \begin{bmatrix} \left(3 + u\cos\frac{\theta}{2}\right)\cos\theta \\ \left(3 + u\cos\frac{\theta}{2}\right)\sin\theta \\ u\sin\frac{\theta}{2} \end{bmatrix} \qquad \widetilde{\mathsf{F}}(\widetilde{u}, \widetilde{\theta}) = \begin{bmatrix} \left(3 + \widetilde{u}\cos\frac{\widetilde{\theta}}{2}\right)\cos\widetilde{\theta} \\ \left(3 + \widetilde{u}\cos\frac{\widetilde{\theta}}{2}\right)\sin\widetilde{\theta} \\ \widetilde{u}\sin\frac{\widetilde{\theta}}{2} \end{bmatrix}$$

In order to compute the transition map $\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F}(u, \theta)$, we need to solve the system of equations, i.e. find $(\widetilde{u}, \widetilde{\theta})$ in terms of (u, θ):

$$\left(3 + u\cos\frac{\theta}{2}\right)\cos\theta = \left(3 + \widetilde{u}\cos\frac{\widetilde{\theta}}{2}\right)\cos\widetilde{\theta} \tag{4.1}$$

$$\left(3 + u\cos\frac{\theta}{2}\right)\sin\theta = \left(3 + \widetilde{u}\cos\frac{\widetilde{\theta}}{2}\right)\sin\widetilde{\theta} \tag{4.2}$$

$$u\sin\frac{\theta}{2} = \widetilde{u}\sin\frac{\widetilde{\theta}}{2} \tag{4.3}$$

By considering $(4.1)^2 + (4.2)^2$, we get:

$$u\cos\frac{\theta}{2} = \widetilde{u}\cos\frac{\widetilde{\theta}}{2} \tag{4.4}$$

We leave it as an exercise for readers to check that $\theta \neq \pi$ in order for the system to be solvable. Therefore, $\theta \in (0, \pi) \cup (\pi, 2\pi)$ and so the domain of overlap is a disjoint union of two open sets.

When $\theta \in (0, \pi)$, from (4.3) and (4.4) we can conclude that $\widetilde{u} = u$ and $\widetilde{\theta} = \theta$.

When $\theta \in (\pi, 2\pi)$, we *cannot* have $\widetilde{\theta} = \theta$ since $\widetilde{\theta} \in (-\pi, \pi)$. However, one can have $\widetilde{u} = -u$ so that (4.3) and (4.4) become:

$$\sin\frac{\theta}{2} = -\sin\frac{\widetilde{\theta}}{2} \quad \text{and} \quad \cos\frac{\theta}{2} = -\cos\frac{\widetilde{\theta}}{2}$$

which implies $\widetilde{\theta} = \theta - 2\pi$.

To conclude, we have:

$$\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F}(u, \theta) = \begin{cases} (u, \theta) & \text{if } \theta \in (0, \pi) \\ (-u, \theta - 2\pi) & \text{if } \theta \in (\pi, 2\pi) \end{cases}$$

By direct computations, we get:

$$\det D(\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F})(u, \theta) = \begin{cases} 1 & \text{if } \theta \in (0, \pi) \\ -1 & \text{if } \theta \in (\pi, 2\pi) \end{cases}$$

Therefore, no matter how we switch the order of u and θ, or $\widetilde{u}$ and $\widetilde{\theta}$, we can never allow $\det D(\widetilde{\mathsf{F}}^{-1} \circ \mathsf{F}) > 0$ everywhere on the overlap. In other words, even if the unit normal vectors $\hat{\mathsf{n}}_{\mathsf{F}}$ and $\hat{\mathsf{n}}_{\widetilde{F}}$ agree with each other when $\theta \in (0, \pi)$, it would point in opposite direction when $\theta \in (\pi, 2\pi)$. □

Next, we are back to hypersurfaces M^n in $\mathbb{R}^{n+1}$ and prove the equivalence between consistency of unit normal and positivity of transition maps. To begin, we need the following result about normal vectors (which is left as an exercise for readers):

Exercise 4.6. Let M^n be a smooth hypersurface in $\mathbb{R}^{n+1}$ whose coordinates are denoted by $(x_1, \ldots, x_{n+1})$, and the unit vector along the x_i-direction is denoted by e_i. Let $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \to M^n$ be a local parametrization of M. Show that the following vector defined on $\mathsf{F}(\mathcal{U})$ is normal to the hypersurface M^n:

$$\sum_{i=1}^{n+1} \det \frac{\partial(x_{i+1}, \ldots, x_{n+1}, x_1, \ldots, x_{i-1})}{\partial(u_1, \ldots, u_n)} \mathsf{e}_i.$$

Proposition 4.12. *Given a smooth hypersurface M^n in $\mathbb{R}^{n+1}$, the following are equivalent:*

(i) M^n is orientable;

(ii) There exists a family of local parametrizations $\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M$ covering M such that for any $\mathsf{F}_\alpha, \mathsf{F}_\beta$ in the family with $\mathsf{F}_\beta(\mathcal{U}_\beta) \cap \mathsf{F}_\alpha(\mathcal{U}_\alpha) \neq \emptyset$, we have:

$$\det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta) > 0 \quad \textit{on } \mathsf{F}_\beta^{-1}\left(\mathsf{F}_\beta(\mathcal{U}_\beta) \cap \mathsf{F}_\alpha(\mathcal{U}_\alpha)\right).$$

Proof. We first prove (ii) $\Longrightarrow$ (i). Denote $(u_1^\alpha, \ldots, u_n^\alpha)$ to be the local coordinates of M under the parametrization F_α. On every $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$, using the result from Exercise 4.6, one can construct a unit normal vector locally defined on $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$:

$$\hat{\mathsf{n}}_\alpha = \frac{\sum_{i=1}^{n+1} \det \frac{\partial(x_{i+1},\ldots,x_{n+1},x_1,\ldots,x_{i-1})}{\partial(u_1^\alpha,\ldots,u_n^\alpha)} \mathsf{e}_i}{\left|\sum_{i=1}^{n+1} \det \frac{\partial(x_{i+1},\ldots,x_{n+1},x_1,\ldots,x_{i-1})}{\partial(u_1^\alpha,\ldots,u_n^\alpha)} \mathsf{e}_i\right|}$$

Similarly, on $\mathsf{F}_\beta(\mathcal{U}_\beta)$, we have another locally defined unit normal vectors:

$$\hat{\mathsf{n}}_\beta = \frac{\sum_{i=1}^{n+1} \det \frac{\partial(x_{i+1},\ldots,x_{n+1},x_1,\ldots,x_{i-1})}{\partial(u_1^\beta,\ldots,u_n^\beta)} \mathsf{e}_i}{\left|\sum_{i=1}^{n+1} \det \frac{\partial(x_{i+1},\ldots,x_{n+1},x_1,\ldots,x_{i-1})}{\partial(u_1^\beta,\ldots,u_n^\beta)} \mathsf{e}_i\right|}$$

Then on the overlap $\mathsf{F}_\beta^{-1}\left(\mathsf{F}_\alpha(\mathcal{U}_\alpha) \cap \mathsf{F}_\beta(\mathcal{U}_\beta)\right)$, the chain rule asserts that:

$$\begin{aligned}
&\det \frac{\partial(x_{i+1},\ldots,x_{n+1},x_1,\ldots,x_{i-1})}{\partial(u_1^\beta,\ldots,u_n^\beta)} \\
&= \det \frac{\partial(u_1^\alpha,\ldots,u_n^\alpha)}{\partial(u_1^\beta,\ldots,u_n^\beta)} \;\det \frac{\partial(x_{i+1},\ldots,x_{n+1},x_1,\ldots,x_{i-1})}{\partial(u_1^\alpha,\ldots,u_n^\alpha)} \\
&= \det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta)\;\det \frac{\partial(x_{i+1},\ldots,x_{n+1},x_1,\ldots,x_{i-1})}{\partial(u_1^\alpha,\ldots,u_n^\alpha)}
\end{aligned}$$

and so the two unit normal vectors are related by:

$$\hat{\mathsf{n}}_\beta = \frac{\det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta)}{\left|\det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta)\right|}\hat{\mathsf{n}}_\alpha.$$

By the condition that $\det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta) > 0$, we have $\hat{\mathsf{n}}_\beta = \hat{\mathsf{n}}_\alpha$ on the overlap. Define $\hat{\mathsf{n}} := \hat{\mathsf{n}}_\alpha$ on every $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$, it is then a continuous unit normal vector globally defined on M. This proves (i).

Now we show (i) $\Longrightarrow$ (ii). Suppose $\hat{\mathsf{n}}$ is a continuous unit normal vector defined on the whole M. Suppose $\mathsf{F}_\alpha(u_1^\alpha, \ldots, u_n^\alpha) : \mathcal{U}_\alpha \to M$ is *any* family of local parametrizations that cover the whole M. On every $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$, we consider the locally defined unit normal vector:

$$\hat{\mathsf{n}}_\alpha = \frac{\sum_{i=1}^{n+1} \det \frac{\partial(x_{i+1},\ldots,x_{n+1},x_1,\ldots,x_{i-1})}{\partial(u_1^\alpha,\ldots,u_n^\alpha)} \mathsf{e}_i}{\left|\sum_{i=1}^{n+1} \det \frac{\partial(x_{i+1},\ldots,x_{n+1},x_1,\ldots,x_{i-1})}{\partial(u_1^\alpha,\ldots,u_n^\alpha)} \mathsf{e}_i\right|}.$$

As a hypersurface M^n in $\mathbb{R}^{n+1}$, there is only one direction of normal vectors, and so we have either $\hat{\mathsf{n}}_\alpha = \hat{\mathsf{n}}$ or $\hat{\mathsf{n}}_\alpha = -\hat{\mathsf{n}}$ on $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$. For the latter case, one can modify the parametrization F_α by switching any pair of u_i^α's such that $\hat{\mathsf{n}}_\alpha = \hat{\mathsf{n}}$.

After making suitable modification on every F_α, we can assume without loss of generality that F_α's are local parametrizations such that $\hat{\mathsf{n}}_\alpha = \hat{\mathsf{n}}$ on every $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$. In particular, on the overlap $\mathsf{F}_\beta^{-1}\left(\mathsf{F}_\alpha(\mathcal{U}_\alpha) \cap \mathsf{F}_\beta(\mathcal{U}_\beta)\right)$, we have $\hat{\mathsf{n}}_\alpha = \hat{\mathsf{n}}_\beta$.

By $\hat{\mathsf{n}}_\beta = \dfrac{\det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta)}{\left|\det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta)\right|}\hat{\mathsf{n}}_\alpha$, we conclude that $\det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta) > 0$, proving (ii).

□

Remark 4.13. According to Proposition 4.12, the cylinder in Example 4.10 is orientable, while the Möbius strip in Example 4.11 is not orientable. □

Exercise 4.7. Show that the unit sphere $\mathbb{S}^2$ in $\mathbb{R}^3$ is orientable.

Exercise 4.8. Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a smooth function. Suppose $c \in \mathbb{R}$ such that $f^{-1}(c)$ is non-empty and f is a submersion at every $p \in f^{-1}(c)$. Show that $f^{-1}(c)$ is an orientable hypersurface in $\mathbb{R}^3$.

4.2.2. Orientable Manifolds. On an abstract manifold M, it is not possible to define normal vectors on M, and so the notion of orientability cannot be defined using normal vectors. However, thanks to Proposition 4.12, the notion of orientability of hypersurfaces is equivalent to positivity of Jacobians of transition maps, which we can also talk about on abstract manifolds. Therefore, motivated by Proposition 4.12, we define:

Definition 4.14 (Orientable Manifolds)**.** A smooth manifold M is said to be *orientable* if there exists a family of local parametrizations $\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M$ covering M such that for any F_α and F_β in the family with $\mathsf{F}_\beta(\mathcal{U}_\beta) \cap \mathsf{F}_\alpha(\mathcal{U}_\alpha) \neq \emptyset$, we have:

$$\det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{F}_\beta) > 0 \quad \text{on } \mathsf{F}_\beta^{-1}\left(\mathsf{F}_\beta(\mathcal{U}_\beta) \cap \mathsf{F}_\alpha(\mathcal{U}_\alpha)\right).$$

In this case, we call the family $\mathcal{A} = \{\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M\}$ of local parametrizations to be an *oriented atlas* of M.

Example 4.15. Recall that the real projective space $\mathbb{RP}^2$ consists of homogeneous triples $[x_0 : x_1 : x_2]$ where $(x_0, x_1, x_2) \neq (0,0,0)$. The standard parametrizations are given by:

$$\begin{aligned}\mathsf{F}_0(x_1, x_2) &= [1 : x_1 : x_2]\\ \mathsf{F}_1(y_0, y_2) &= [y_0 : 1 : y_2]\\ \mathsf{F}_2(z_0, z_1) &= [z_0 : z_1 : 1]\end{aligned}$$

By the fact that $[y_0 : 1 : y_2] = [1 : y_0^{-1} : y_2 y_0^{-1}]$, the transition map $\mathsf{F}_0^{-1} \circ \mathsf{F}_1$ is defined on $\{(y_0, y_2) \in \mathbb{R}^2 : y_0 \neq 0\}$, and is given by: $(x_1, x_2) = (y_0^{-1}, y_2 y_0^{-1})$. Hence,

$$\begin{aligned}D(\mathsf{F}_0^{-1} \circ \mathsf{F}_1) &= \frac{\partial(x_1, x_2)}{\partial(y_0, y_2)} = \begin{bmatrix} -y_0^{-2} & 0 \\ -y_2 y_0^{-2} & y_0^{-1} \end{bmatrix}\\ \det D(\mathsf{F}_0^{-1} \circ \mathsf{F}_1) &= -\frac{1}{y_0^3}\end{aligned}$$

Therefore, it is impossible for $\det D(\mathsf{F}_0^{-1} \circ \mathsf{F}_1) > 0$ on the overlap domain $\{(y_0, y_2) \in \mathbb{R}^2 : y_0 \neq 0\}$.

At this stage, we have shown that this altas is not an oriented one. In order to prove $\mathbb{RP}^2$ is non-orientable, we need to show *any* altas of $\mathbb{RP}^2$ is not oriented. We will prove this using Proposition 4.25 later. □

Exercise 4.9. Show that $\mathbb{RP}^3$ is orientable. Propose a conjecture about the orientability of $\mathbb{RP}^n$.

Exercise 4.10. Show that for *any* smooth manifold M (whether or not it is orientable), the tangent bundle TM must be orientable.

Exercise 4.11. Show that for a smooth orientable manifold M with boundary, the boundary manifold ∂M must also be orientable.

4.3. Integrations of Differential Forms

Generalized Stokes' Theorem concerns about integrals of differential forms. In this section, we will give a rigorous definition of these integrals.

4.3.1. Single Parametrization. In the *simplest* case if a manifold M can be covered by a single parametrization:

$$\mathsf{F}(u_1, \ldots, u_n) : (\alpha_1, \beta_1) \times \cdots \times (\alpha_n, \beta_n) \to M,$$

then given an n-form $\varphi(u_1, \ldots, u_n)\, du^1 \wedge du^2 \wedge \cdots \wedge du^n$, the integral of ω over the manifold M is given by:

$$\underbrace{\int_M \varphi(u_1, \ldots, u_n)\, du^1 \wedge du^2 \wedge \cdots \wedge du^n}_{\text{integral of differential form}} := \underbrace{\int_{\alpha_n}^{\beta_n} \cdots \int_{\alpha_1}^{\beta_1} \varphi(u_1, \ldots, u_n)\, du^1\, du^2 \cdots du^n}_{\text{ordinary integral in Multivariable Calculus}}$$

From the definition, we see that it only makes sense to integrate an n-form on an n-dimensional manifold.

Very few manifolds can be covered by a single parametrization. Of course, $\mathbb{R}^n$ is an example. One less trivial example is the graph of a smooth function. Suppose $f(x, y) : \mathbb{R}^2 \to \mathbb{R}$ is a smooth function. Consider its graph:

$$\Gamma_f := \{(x, y, f(x, y)) \in \mathbb{R}^3 : (x, y) \in \mathbb{R}^2\}$$

which can be globally parametrized by $\mathsf{F} : \mathbb{R}^2 \to \Gamma_f$ where

$$\mathsf{F}(x, y) = (x, y, f(x, y)).$$

Let $\omega = e^{-x^2-y^2}\, dx \wedge dy$ be a 2-form on Γ_f, then its integral over Γ_f is given by:

$$\int_{\Gamma_f} \omega = \int_{\Gamma_f} e^{-x^2-y^2}\, dx \wedge dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2-y^2}\, dx\, dy = \pi.$$

Here we leave the computational detail as an exercise for readers.

It *appears* that integrating a differential form is just like "erasing the wedges", yet there are two subtle (but important) issues:

(1) In the above example, note that ω can also be written as:

$$\omega = -e^{-x^2-y^2}\, dy \wedge dx.$$

It suggests that we also have:

$$\int_{\Gamma_f} \omega = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} -e^{-x^2-y^2}\, dy\, dx = -\pi,$$

which is *not* consistent with the previous result. How shall we fix it?

(2) Even if a manifold can be covered by one single parametrization, such a parametrization may not be unique. If both $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ are global coordinates of M, then a differential form ω can be expressed in terms of either u_i's or v_i's. Is the integral independent of the chosen coordinate system?

The first issue can be resolved easily. Whenever we talk about integration of differential forms, we need to first fix the order of the coordinates. Say on $\mathbb{R}^2$ we fix the order to be (x, y), then for any given 2-form we should express it in terms of $dx \wedge dy$ before "erasing the wedges". For the 2-form ω above, we must first express it as:

$$\omega = e^{-x^2-y^2}\, dx \wedge dy$$

before integrating it.

For higher (say dim $= 4$) dimensional manifolds M^4 covered by a single parametrization $\mathsf{F}(u_1, \ldots, u_4) : \mathcal{U} \to M$, if we choose (u_1, u_2, u_3, u_4) to be the order of coordinates and given a 4-form:

$$\Omega = f(u_1, \ldots, u_4)\, du^1 \wedge du^3 \wedge du^2 \wedge du^4 + g(u_1, \ldots, u_4)\, du^4 \wedge du^3 \wedge du^2 \wedge du^1,$$

then we need to re-order the wedge product so that:

$$\Omega = -f(u_1, \ldots, u_4)\, du^1 \wedge du^2 \wedge du^3 \wedge du^4 + g(u_1, \ldots, u_4) du^1 \wedge du^2 \wedge du^3 \wedge du^4.$$

The integral of ω over M^4 with respect to the order (u_1, u_2, u_3, u_4) is given by:

$$\int_M \Omega = \int_{\mathcal{U}} \left(-f(u_1, \ldots, u_4) + g(u_1, \ldots, u_4)\right)\, du^1\, du^2\, du^3\, du^4.$$

Let's examine the second issue. Suppose M is an n-manifold with two different global parametrizations $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \to M$ and $\mathsf{G}(v_1, \ldots, v_n) : \mathcal{V} \to M$. Given an n-form ω which can be expressed as:

$$\omega = \varphi\, du^1 \wedge \cdots \wedge du^n,$$

then from Proposition 3.56, ω can be expressed in terms of v_i's by:

$$\omega = \varphi \det \frac{\partial(u_1, \ldots, u_n)}{\partial(v_1, \ldots, v_n)}\, dv^1 \wedge \cdots \wedge dv^n.$$

Recall that the change-of-variable formula in Multivariable Calculus asserts that:

$$\int_{\mathcal{U}} \varphi\, du^1 \cdots du^n = \int_{\mathcal{V}} \varphi \left| \det \frac{\partial(u_1, \ldots, u_n)}{\partial(v_1, \ldots, v_n)} \right| dv^1 \cdots dv^n.$$

Therefore, in order for $\displaystyle\int_M \omega$ to be well-defined, we need

$$\int_{\mathcal{U}} \varphi\, du^1 \wedge \cdots \wedge du^n \text{ and } \int_{\mathcal{V}} \varphi \det \frac{\partial(u_1, \ldots, u_n)}{\partial(v_1, \ldots, v_n)}\, dv^1 \wedge \cdots \wedge dv^n$$

to be equal, and so we require:

$$\det \frac{\partial(u_1, \ldots, u_n)}{\partial(v_1, \ldots, v_n)} > 0.$$

When defining an integral of a differential form, we *not only* need to choose a convention on the order of coordinates, say $(u_1, \ldots, u_n)$, but also we shall only consider those coordinate systems $(v_1, \ldots, v_n)$ such that $\det \dfrac{\partial(u_1, \ldots, u_n)}{\partial(v_1, \ldots, v_n)} > 0$. Therefore, in order to integrate a differential form, we require the manifold to be *orientable*.

4.3.2. Multiple Parametrizations. A majority of smooth manifolds are covered by more than one parametrizations. Integrating a differential form over such a manifold is not as straight-forward as previously discussed.

In case M can be "almost" covered by a single parametrization $\mathsf{F} : \mathcal{U} \to M$ (i.e. the set $M \backslash \mathsf{F}(\mathcal{U})$ has measure zero) and the n-form ω is continuous, then it is still possible to compute $\displaystyle\int_M \omega$ by computing $\displaystyle\int_{\mathsf{F}(\mathcal{U})} \omega$. Let's consider the example of a sphere:

Example 4.16. Let $\mathbb{S}^2$ be the unit sphere in $\mathbb{R}^3$ centered at the origin. Consider the 2-form ω on $\mathbb{R}^3$ defined as:

$$\omega = dx \wedge dy.$$

Let $\iota : \mathbb{S}^2 \to \mathbb{R}^3$ be the inclusion map, then $\iota^*\omega$ is a 2-form on $\mathbb{S}^2$. We are interested in the value of the integral $\displaystyle\int_{\mathbb{S}^2} \iota^*\omega$.

Note that $\mathbb{S}^2$ can be covered almost everywhere by spherical coordinate parametrization $\mathsf{F}(\varphi,\theta) : (0,\pi) \times (0,2\pi) \to \mathbb{S}^2$ given by:

$$\mathsf{F}(\varphi,\theta) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi).$$

Under the local coordinates (φ,θ), we have:

$$\begin{aligned}
\iota^*(dx) &= d(\sin\varphi\cos\theta) = \cos\varphi\cos\theta\,d\varphi - \sin\varphi\sin\theta\,d\theta \\
\iota^*(dy) &= d(\sin\varphi\sin\theta) = \cos\varphi\sin\theta\,d\varphi + \sin\varphi\cos\theta\,d\theta \\
\iota^*\omega &= \iota^*(dx) \wedge \iota^*(dy) \\
&= \sin\varphi\cos\varphi\,d\varphi \wedge d\theta.
\end{aligned}$$

Therefore,

$$\int_M \iota^*\omega = \int_M \sin\varphi\cos\varphi\,d\varphi \wedge d\theta = \int_0^{2\pi}\int_0^{\pi} \sin\varphi\cos\varphi\,d\varphi\,d\theta = 0.$$

Here we pick (φ,θ) as the order of coordinates. □

Exercise 4.12. Let $\omega = x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy$. Compute

$$\int_{\mathbb{S}^2} \iota^*\omega$$

where $\mathbb{S}^2$ is the unit sphere in $\mathbb{R}^3$ centered at the origin, and $\iota : \mathbb{S}^2 \to \mathbb{R}^3$ is the inclusion map.

Exercise 4.13. Let $\mathbb{T}^2$ be the torus in $\mathbb{R}^4$ defined as:

$$\mathbb{T}^2 := \left\{ (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1^2 + x_2^2 = x_3^2 + x_4^2 = \frac{1}{2} \right\}.$$

Let $\iota : \mathbb{T}^2 \to \mathbb{R}^4$ be the inclusion map. Compute the following integral:

$$\int_{\mathbb{T}^2} \iota^* \left(x_1 x_2 x_3\,dx^4 \wedge dx^3 \right).$$

FYI: Clifford Torus

The torus $\mathbb{T}^2$ in Exercise 4.13 is a well-known object in Differential Geometry called the *Clifford Torus*. A famous conjecture called the Hsiang-Lawson's Conjecture concerns about this torus. One of the proposers Wu-Yi Hsiang is a retired faculty of HKUST Math. This conjecture was recently solved by Simon Brendle in 2012.

Next, we will discuss how to define integrals of differential forms when M is covered by multiple parametrizations none of which can almost cover the whole manifold. The key idea is to break down the n-form into small pieces, so that each piece is completely covered by one single parametrization. It will be done using *partition of unity* to be discussed.

We first introduce the notion of *support* which appears often in the rest of the course (as well as in advanced PDE courses).

Definition 4.17 (Support)**.** Let M be a smooth manifold. Given a k-form ω (where $0 \le k \le n$) defined on M, we denote and define the *support of* ω to be:

$$\operatorname{supp}\omega := \overline{\{p \in M : \omega(p) \neq 0\}},$$

i.e. the *closure* of the set $\{p \in M : \omega(p) \neq 0\}$.

Suppose M^n is an oriented manifold with $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \to M$ as one of (many) local parametrizations. If an n-form ω on M^n only has "stuff" inside $\mathsf{F}(\mathcal{U})$, or precisely:

$$\operatorname{supp} \omega \subset \mathsf{F}(\mathcal{U}),$$

then one can define $\int_M \omega$ as in the previous subsection. Namely, if on $\mathsf{F}(\mathcal{U})$ we have $\omega = \varphi\, du^1 \wedge \cdots \wedge du^n$, then we define:

$$\int_M \omega = \int_{\mathsf{F}(\mathcal{U})} \omega = \int_{\mathcal{U}} \varphi\, du^1 \cdots du^n.$$

Here we pick the order of coordinates to be $(u_1, \ldots, u_n)$.

The following important tool called *partitions of unity* will "chop" a differential form into "little pieces" such that each piece has support covered by a single parametrization.

Definition 4.18 (Partitions of Unity)**.** Let M be a smooth manifold with an atlas $\mathcal{A} = \{\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M\}$ such that $M = \bigcup_{\text{all } \alpha} \mathsf{F}_\alpha(\mathcal{U}_\alpha)$. A *partition of unity subordinate to the atlas* $\mathcal{A}$ is a family of smooth functions $\rho_\alpha : M \to [0, 1]$ with the following properties:

(i) $\operatorname{supp} \rho_\alpha \subset \mathsf{F}_\alpha(\mathcal{U}_\alpha)$ for any α.

(ii) For any $p \in M$, there exists an open set $\mathcal{O} \subset M$ containing p such that

$$\operatorname{supp} \rho_\alpha \cap \mathcal{O} \neq \emptyset$$

for finitely many α's only.

(iii) $\sum_{\text{all } \alpha} \rho_\alpha \equiv 1$ on M.

Remark 4.19. It can be shown that given any smooth manifold with any atlas, partitions of unity subordinate to that given atlas must exist. The proof is very technical and is not in the same spirit with other parts of the course, so we omit the proof here. It is more important to know what partitions of unity are for, than to know the proof of existence. □

Remark 4.20. Note that partitions of unity subordinate to a given atlas may not be unique! □

Remark 4.21. Condition (ii) in Definition 4.18 is merely a technical analytic condition to make sure the sum $\sum_{\text{all } \alpha} \rho_\alpha(p)$ is a finite sum for each fixed $p \in M$, so that we do not need to worry about convergence issues. If the manifold can be covered by finitely many local parametrizations, then condition (ii) automatically holds (and we do not need to worry about). □

Now, take an n-form ω defined on an orientable manifold M^n, which is parametrized by an oriented atlas $\mathcal{A} = \{\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M\}$. Let $\{\rho_\alpha : M \to [0, 1]\}$ be a partition of unity subordinate to $\mathcal{A}$, then by condition (iii) in Definition 4.18, we get:

$$\omega = \underbrace{\left(\sum_{\text{all } \alpha} \rho_\alpha \right)}_{=1} \omega = \sum_{\text{all } \alpha} \rho_\alpha \omega.$$

Condition (i) says that $\operatorname{supp} \rho_\alpha \subset \mathsf{F}_\alpha(\mathcal{U}_\alpha)$, or heuristically speaking ρ_α vanishes outside $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$. Naturally, we have $\operatorname{supp}(\rho_\alpha \omega) \subset \mathsf{F}_\alpha(\mathcal{U}_\alpha)$ for each α. Therefore, as previously discussed, we can integrate $\rho_\alpha \omega$ for each *individual* α:

$$\int_M \rho_\alpha \omega := \int_{\mathsf{F}_\alpha(\mathcal{U}_\alpha)} \rho_\alpha \omega.$$

Given that we can integrate each $\rho_\alpha\omega$, we define the integral of ω as:

$$\int_M \omega := \sum_{\text{all } \alpha} \int_M \rho_\alpha\omega = \sum_{\text{all } \alpha} \int_{\mathsf{F}_\alpha(\mathcal{U}_\alpha)} \rho_\alpha\omega. \tag{4.5}$$

However, the sum involved in (4.5) is in general an infinite (possible uncountable!) sum. To avoid convergence issue, from now on we will only consider n-forms ω which have *compact support*, i.e.

$$\operatorname{supp}\omega \text{ is a compact set.}$$

Recall that every open cover of a compact set has a finite sub-cover. Together with condition (ii) in Definition 4.18, one can show that $\rho_\alpha\omega$ are identically zero for all except finitely many α's. The argument goes as follows: at each $p \in \operatorname{supp}\omega$, by condition (ii) in Definition 4.18, there exists an open set $\mathcal{O}_p \subset M$ containing p such that the set:

$$S_p := \{\alpha : \operatorname{supp}\rho_\alpha \cap \mathcal{O}_p \neq \emptyset\}$$

is finite. Evidently, we have

$$\operatorname{supp}\omega \subset \bigcup_{p\in\operatorname{supp}\omega} \mathcal{O}_p$$

and by compactness of $\operatorname{supp}\omega$, there exists $p_1, \ldots, p_N \in \operatorname{supp}\omega$ such that

$$\operatorname{supp}\omega \subset \bigcup_{i=1}^N \mathcal{O}_{p_i}.$$

Since $\{q \in M : \rho_\alpha(q)\omega(q) \neq 0\} \subset \{q \in M : \rho_\alpha(q) \neq 0\} \cap \{q \in M : \omega(q) \neq 0\}$, we have:

$$\begin{aligned}
\operatorname{supp}(\rho_\alpha\omega) &= \overline{\{q \in M : \rho_\alpha(q)\omega(q) \neq 0\}} \\
&\subset \overline{\{q \in M : \rho_\alpha(q) \neq 0\} \cap \{q \in M : \omega(q) \neq 0\}} \\
&\subset \overline{\{q \in M : \rho_\alpha(q) \neq 0\}} \cap \overline{\{q \in M : \omega(q) \neq 0\}} \\
&= \operatorname{supp}\rho_\alpha \cap \operatorname{supp}\omega \subset \bigcup_{i=1}^N \left(\operatorname{supp}\rho_\alpha \cap \mathcal{O}_{p_i}\right).
\end{aligned}$$

Therefore, if α is an index such that $\operatorname{supp}(\rho_\alpha\omega) \neq \emptyset$, then there exists $i \in \{1, \ldots, N\}$ such that $\operatorname{supp}\rho_\alpha \cap \mathcal{O}_{p_i} \neq \emptyset$, or in other words, $\alpha \in S_{p_i}$ for some i, and so:

$$\{\alpha : \operatorname{supp}(\rho_\alpha\omega) \neq \emptyset\} \subset \bigcup_{i=1}^N S_{p_i}.$$

Since each S_{p_i} is a finite set, the set $\{\alpha : \operatorname{supp}(\rho_\alpha\omega) \neq \emptyset\}$ is also finite. Therefore, there are only finitely many α's such that $\displaystyle\int_{\mathsf{F}_\alpha(\mathcal{U}_\alpha)}$ is non-zero, and so the sum stated in (4.5) is in fact a *finite sum*.

Now we have understood that there is no convergence issue for (4.5) provided that ω has compact support (which is automatically true if the manifold M is itself compact). There are still two well-definedness issues to resolve, namely whether the integral in (4.5) is independent of oriented atlas $\mathcal{A}$, and for each atlas whether the integral is independent of the choice of partitions of unity.

Proposition 4.22. *Let M^n be an orientable smooth manifold with two oriented atlas*

$$\mathcal{A} = \{\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M\} \text{ and } \mathcal{B} = \{\mathsf{G}_\beta : \mathcal{V}_\beta \to M\}$$

such that $\det D(\mathsf{F}_\alpha^{-1} \circ \mathsf{G}_\beta) > 0$ *on the overlap for any pair of α and β. Suppose $\{\rho_\alpha : M \to [0,1]\}$ and $\{\sigma_\beta : M \to [0,1]\}$ are partitions of unity subordinate to $\mathcal{A}$ and $\mathcal{B}$ respectively. Then, given any compactly supported differential n-form ω on M^n, we have:*

$$\sum_{\text{all } \alpha} \int_{\mathsf{F}_\alpha(\mathcal{U}_\alpha)} \rho_\alpha \omega = \sum_{\text{all } \beta} \int_{\mathsf{G}_\beta(\mathcal{V}_\beta)} \sigma_\beta \omega.$$

Proof. By the fact that $\sum_{\text{all } \beta} \sigma_\beta \equiv 1$ on M, we have:

$$\sum_{\text{all } \alpha} \int_{\mathsf{F}_\alpha(\mathcal{U}_\alpha)} \rho_\alpha \omega = \sum_{\text{all } \alpha} \int_{\mathsf{F}_\alpha(\mathcal{U}_\alpha)} \left(\sum_{\text{all } \beta} \sigma_\beta \right) \rho_\alpha \omega = \sum_{\text{all } \alpha} \sum_{\text{all } \beta} \int_{\mathsf{F}_\alpha(\mathcal{U}_\alpha) \cap \mathsf{G}_\beta(\mathcal{V}_\beta)} \rho_\alpha \sigma_\beta \omega.$$

The last equality follows from the fact that $\operatorname{supp} \sigma_\beta \subset \mathsf{G}_\beta(\mathcal{V}_\beta)$.

One can similarly work out that

$$\sum_{\text{all } \beta} \int_{\mathsf{G}_\beta(\mathcal{V}_\beta)} \sigma_\beta \omega = \sum_{\text{all } \beta} \sum_{\text{all } \alpha} \int_{\mathsf{F}_\alpha(\mathcal{U}_\alpha) \cap \mathsf{G}_\beta(\mathcal{V}_\beta)} \rho_\alpha \sigma_\beta \omega.$$

Note that $\sum_\alpha \sum_\beta$ is a finite double sum and so there is no issue of switching them. It completes the proof. □

By Proposition 4.22, we justified that (4.5) is independent of oriented atlas and the choice of partitions of unity. We can now define:

Definition 4.23. Let M^n be an orientable smooth manifold with an oriented atlas $\mathcal{A} = \{\mathsf{F}_\alpha(u_\alpha^1, \ldots, u_\alpha^n) : \mathcal{U}_\alpha \to M\}$ where $(u_\alpha^1, \ldots, u_\alpha^n)$ is the chosen order of local coordinates. Pick a partition of unity $\{\rho_\alpha : M \to [0,1]\}$ subordinate to the atlas $\mathcal{A}$. Then, given any n-form ω, we define its *integral over M* as:

$$\int_M \omega := \sum_{\text{all } \alpha} \int_{\mathsf{F}_\alpha(\mathcal{U}_\alpha)} \rho_\alpha \omega.$$

If $\omega = \varphi_\alpha \, du_\alpha^1 \wedge \cdots \wedge du_\alpha^n$ on each $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$, then:

$$\int_M \omega = \sum_{\text{all } \alpha} \int_{\mathcal{U}_\alpha} \rho_\alpha \varphi_\alpha \, du_\alpha^1 \cdots du_\alpha^n.$$

Remark 4.24. It is generally impossible to compute such an integral, as we know only the existence of ρ_α's but typically not the exact expressions. Even if such a partition of unity ρ_α's can be found, it often involves some terms such as e^{-1/x^2}, which is almost impossible to integrate. To conclude, we do not attempt compute such an integral, but we will study the properties of it based on the definition. □

4.3.3. Orientation of Manifolds. Partition of unity is a powerful tool to construct a smooth *global* item from *local* ones. For integrals of differential forms, we first defines integral of forms with support contained in a single parametrization chart, then we uses a partition of unity to *glue* each chart together. There are some other uses in this spirit. The following beautiful statement can be proved using partitions of unity:

Proposition 4.25. *A smooth n-dimensional manifold M is orientable if and only if there exists a non-vanishing smooth n-form globally defined on M.*

Proof. Suppose M is orientable, then by definition there exists an oriented atlas $\mathcal{A} = \{\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M\}$ such that $\det D(\mathsf{F}_\beta^{-1} \circ \mathsf{F}_\alpha) > 0$ for any α and β. For each local parametrization F_α, we denote $(u_\alpha^1, \ldots, u_\alpha^n)$ to be its local coordinates, then the n-form:

$$\eta_\alpha := du_\alpha^1 \wedge \cdots \wedge du_\alpha^n$$

is *locally defined* on $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$.

Let $\{\rho_\alpha : M \to [0,1]\}$ be a partition of unity subordinate to $\mathcal{A}$. We define:

$$\omega = \sum_{\text{all } \alpha} \rho_\alpha \eta_\alpha = \sum_{\text{all } \alpha} \rho_\alpha \, du_\alpha^1 \wedge \cdots \wedge du_\alpha^n.$$

We claim $\omega(p) \neq 0$ at every point $p \in M$. Suppose $p \in \mathsf{F}_\beta(\mathcal{U}_\beta)$ for some β in the atlas. By (3.19), for each α, locally near p we have:

$$du_\alpha^1 \wedge \cdots \wedge du_\alpha^n = \det \frac{\partial(u_\alpha^1, \ldots, u_\alpha^n)}{\partial(u_\beta^1, \ldots, u_\beta^n)} \, du_\beta^1 \wedge \cdots \wedge du_\beta^n,$$

and so:

$$\omega = \left(\sum_{\text{all } \alpha} \rho_\alpha \, \det \frac{\partial(u_\alpha^1, \ldots, u_\alpha^n)}{\partial(u_\beta^1, \ldots, u_\beta^n)} \right) du_\beta^1 \wedge \cdots \wedge du_\beta^n.$$

Since $\rho_\alpha \geq 0$, $\sum_{\text{all } \alpha} \rho_\alpha \equiv 1$ and $\det \dfrac{\partial(u_\alpha^1, \ldots, u_\alpha^n)}{\partial(u_\beta^1, \ldots, u_\beta^n)} > 0$, we must have:

$$\sum_{\text{all } \alpha} \rho_\alpha \, \det \frac{\partial(u_\alpha^1, \ldots, u_\alpha^n)}{\partial(u_\beta^1, \ldots, u_\beta^n)} > 0 \quad \text{near } p.$$

This shows ω is a non-vanishing n-form on M.

Conversely, suppose Ω is a non-vanishing n-form on M. Let $\mathcal{C} = \{\mathsf{G}_\alpha : \mathcal{V}_\alpha \to M\}$ be any atlas on M, and for each α we denote $(v_\alpha^1, \ldots, v_\alpha^n)$ to be its local coordinates. Express Ω in terms of local coordinates:

$$\Omega = \varphi_\alpha \, dv_\alpha^1 \wedge \cdots \wedge dv_\alpha^n.$$

Since Ω is non-vanishing, φ_α must be either positive on $\mathcal{V}_\alpha$, or negative on $\mathcal{V}_\alpha$. Re-define the local coordinates by:

$$(\widetilde{v}_\alpha^1, \widetilde{v}_\alpha^2, \ldots, \widetilde{v}_\alpha^n) := \begin{cases} (v_\alpha^1, v_\alpha^2, \ldots, v_\alpha^n) & \text{if } \varphi_\alpha > 0 \\ (-v_\alpha^1, v_\alpha^2, \ldots, v_\alpha^n) & \text{if } \varphi_\alpha < 0 \end{cases}$$

Then, under these new local coordinates, we have:

$$\Omega = |\varphi_\alpha| \; d\widetilde{v}_\alpha^1 \wedge \cdots \wedge d\widetilde{v}_\alpha^n.$$

From (3.19), we can deduce:

$$\Omega = |\varphi_\alpha| \; d\widetilde{v}_\alpha^1 \wedge \cdots \wedge d\widetilde{v}_\alpha^n = |\varphi_\alpha| \; \det \frac{\partial(\widetilde{v}_\alpha^1, \ldots, \widetilde{v}_\alpha^n)}{\partial(\widetilde{v}_\beta^1, \ldots, \widetilde{v}_\beta^n)} \, d\widetilde{v}_\beta^1 \wedge \cdots \wedge d\widetilde{v}_\beta^n$$

on the overlap of any two local coordinates $(\widetilde{v}_\alpha^1, \ldots, \widetilde{v}_\alpha^n)$ and $(\widetilde{v}_\beta^1, \ldots, \widetilde{v}_\beta^n)$. On the other hand, we have:

$$\Omega = \left|\varphi_\beta\right| \; d\widetilde{v}_\beta^1 \wedge \cdots \wedge d\widetilde{v}_\beta^n.$$

This shows:

$$\det \frac{\partial(\widetilde{v}_\alpha^1, \ldots, \widetilde{v}_\alpha^n)}{\partial(\widetilde{v}_\beta^1, \ldots, \widetilde{v}_\beta^n)} = \left| \frac{\varphi_\beta}{\varphi_\alpha} \right| > 0 \quad \text{for any } \alpha, \beta.$$

Therefore, M is orientable. $\square$

The significance of Proposition 4.25 is that it relates the orientability of an n-manifold (which was defined in a rather *local* way) with the existence of a non-vanishing n-form (which is a *global* object). For abstract manifolds, unit normal vectors cannot be defined. Here the non-vanishing global n-form plays a similar role as a continuous unit normal does for hypersurfaces. In the rest of the course we will call:

Definition 4.26 (Orientation of Manifolds)**.** Given an orientable manifold M^n, a non-vanishing global n-form Ω is called an *orientation of* M. A basis of tangent vectors $\{T_1, \ldots, T_n\} \in T_pM$ is said to be Ω-*oriented* if $\Omega(T_1, \ldots, T_n) > 0$. A local coordinate system $(u_1, \ldots, u_n)$ is said to be Ω-*oriented* if $\Omega\left(\frac{\partial}{\partial u_1}, \ldots, \frac{\partial}{\partial u_n}\right) > 0$.

Recall that when we integrate an n-form, we need to first pick an order of local coordinates $(u_1, \ldots, u_n)$, then express the n-form according to this order, and locally define the integral as:

$$\int_{\mathsf{F}(\mathcal{U})} \varphi \, du^1 \wedge \cdots \wedge du^n = \int_{\mathcal{U}} \varphi \, du^1 \cdots du^n.$$

Note that picking the *order of coordinates* is a *local* notion. To rephrase it using *global* terms, we can first pick an orientation Ω (which is a global object on M), then we require the order of any local coordinates $(u_1, \ldots, u_n)$ to be Ω-oriented. Any pair of local coordinate systems $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ which are both Ω-oriented will automatically satisfy $\det \dfrac{\partial(u_1, \ldots, u_n)}{\partial(v_1, \ldots, v_n)} > 0$ on the overlap.

To summarize, given an orientable manifold M^n with a chosen orientation Ω, then for *any* local coordinate system $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \to M$, we define:

$$\int_{\mathsf{F}(\mathcal{U})} \varphi \, du^1 \wedge \cdots \wedge du^n = \begin{cases} \int_{\mathcal{U}} \varphi \, du^1 \cdots du^n & \text{if } (u_1, \ldots, u_n) \text{ is } \Omega\text{-oriented} \\ -\int_{\mathcal{U}} \varphi \, du^1 \cdots du^n & \text{if } (u_1, \ldots, u_n) \text{ is not } \Omega\text{-oriented} \end{cases}$$

or to put it in a more elegant (yet equivalent) way:

$$\int_{\mathsf{F}(\mathcal{U})} \varphi \, du^1 \wedge \cdots \wedge du^n = \operatorname{sgn}\left[\Omega\left(\frac{\partial}{\partial u_1}, \ldots, \frac{\partial}{\partial u_n}\right)\right] \int_{\mathcal{U}} \varphi \, du^1 \cdots du^n.$$

Exercise 4.14. Let $\Omega := dx \wedge dy \wedge dz$ be the orientation of $\mathbb{R}^3$. Which of the following is Ω-oriented?

(a) local coordiantes (x, y, z)

(b) vectors $\{\mathsf{i}, \mathsf{k}, \mathsf{j}\}$

(c) vectors $\{\mathsf{u}, \mathsf{v}, \mathsf{u} \times \mathsf{v}\}$ where u and v are linearly independent vectors in $\mathbb{R}^3$.

Exercise 4.15. Consider three linearly independent vectors $\{\mathsf{u}, \mathsf{v}, \mathsf{w}\}$ in $\mathbb{R}^3$ such that $\mathsf{u} \perp \mathsf{w}$ and $\mathsf{v} \perp \mathsf{w}$. Show that $\{\mathsf{u}, \mathsf{v}, \mathsf{w}\}$ has the same orientation as $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$ if and only if $\mathsf{w} = c\mathsf{u} \times \mathsf{v}$ for some *positive* constant c.

Proposition 4.25 can be used to complete the proof that $\mathbb{RP}^2$ is not orientable in Example 4.15. In that example, we demonstrated that there are two local parametrizations $\mathsf{F}_0(u_1, u_2)$ and $\mathsf{F}_1(v_1, v_2)$ with the properties that:

- the domain of each of F_i is connected; while
- their overlap, i.e. domain of $\mathsf{F}_0^{-1} \circ \mathsf{F}_1$, is not connected; and

- $\det D(\mathsf{F}_0^{-1} \circ \mathsf{F}_1)$ is positive on one component U, but negative on another component V.

To show that $\mathbb{RP}^2$ is not orientable, we argue by contradiction that there exists a global non-vanishing 2-form Ω. Then, if $\Omega(\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}) > 0$, then one has $\Omega(\frac{\partial}{\partial v_1}, \frac{\partial}{\partial v_2}) > 0$ on U since $\det D(\mathsf{F}_0^{-1} \circ \mathsf{F}_1) > 0$ on U, and $\Omega(\frac{\partial}{\partial v_1}, \frac{\partial}{\partial v_2}) < 0$ on V since $\det D(\mathsf{F}_0^{-1} \circ \mathsf{F}_1) < 0$. However, since the domain of $\mathsf{F}_1(v_1, v_2)$ is connected and $\Omega(\frac{\partial}{\partial v_1}, \frac{\partial}{\partial v_2})$ is a smooth (in particular continuous) function on that domain, there must be a point p in the domain of F_1 such that $\Omega(\frac{\partial}{\partial v_1}, \frac{\partial}{\partial v_2}) = 0$ at p. It leads to a contradiction that Ω is non-vanishing. Similar for the case $\Omega(\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}) < 0$.

4.4. Generalized Stokes' Theorem

In this section, we (finally) state and give a proof of an elegant theorem, Generalized Stokes' Theorem. It not only unifies Green's, Stokes' and Divergence Theorems which we learned in Multivariable Calculus, but also generalize it to higher dimensional abstract manifolds.

4.4.1. Boundary Orientation. Since the statement of Generalized Stokes' Theorem involves integration on differential forms, we will assume all manifolds discussed in this section to be *orientable*. Let's fix an orientation Ω of M^n, which is a non-vanishing n-form, and this orientation determines how local coordinates on M are ordered as discussed in the previous section.

Now we deal with the orientation of the boundary manifold ∂M. Given a local parametrization $\mathsf{G}(u_1, \ldots, u_n) : \mathcal{V} \subset \mathbb{R}^n_+ \to M$ of boundary type. The tangent space T_pM for points $p \in \partial M$ is defined as the span of $\left\{\dfrac{\partial}{\partial u_i}\right\}_{i=1}^n$. As $\mathcal{V}$ is a subset of the upper half-space $\{u_n \geq 0\}$, the vector $\nu := -\frac{\partial}{\partial u_n}$ in T_pM is often called an *outward-pointing* "normal" vector to ∂M.

An orientation Ω of M^n is a non-vanishing n-form. The boundary manifold ∂M^n is an $(n-1)$-manifold, and so an orientation of ∂M^n should be a non-vanishing $(n-1)$-form. Using the outward-pointing normal vector ν, one can produce such an $(n-1)$-form in a natural way. Given any tangent vectors $T_1, \ldots, T_{n-1}$ on $T(\partial M)$, we consider the interior product $i_\nu\Omega$, which is defined as:

$$(i_\nu\Omega)(T_1, \ldots, T_{n-1}) := \Omega(\nu, T_1, \ldots, T_{n-1}).$$

Then $i_\nu\Omega$ is an alternating multilinear map in $\wedge^{n-1}T^*(\partial M)$.

Locally, given a local coordinate system $(u_1, \ldots, u_n)$, by recalling that $\nu = -\frac{\partial}{\partial u_n}$ we can compute:

$$\begin{aligned}
(i_\nu\Omega)\left(\frac{\partial}{\partial u_1}, \ldots, \frac{\partial}{\partial u_{n-1}}\right) &= \Omega\left(\nu, \frac{\partial}{\partial u_1}, \ldots, \frac{\partial}{\partial u_{n-1}}\right) \\
&= \Omega\left(-\frac{\partial}{\partial u_n}, \frac{\partial}{\partial u_1}, \ldots, \frac{\partial}{\partial u_{n-1}}\right) \\
&= (-1)^n\Omega\left(\frac{\partial}{\partial u_1}, \ldots, \frac{\partial}{\partial u_{n-1}}, \frac{\partial}{\partial u_n}\right)
\end{aligned}$$

which is non-zero. Therefore, $i_\nu\Omega$ is a non-vanishing $(n-1)$-form on ∂M, and we can take it as an orientation for ∂M. From now on, whenever we pick an orientation Ω for M^n, we will by-default pick $i_\nu\Omega$ to be the orientation for ∂M.

Given an Ω-oriented local coordinate system $\mathsf{G}(u_1, \ldots, u_n) : \mathcal{V} \to M$ of boundary type for M^n, then $(u_1, \ldots, u_{n-1})$ is $i_\nu\Omega$-oriented if n is *even*; and is not $i_\nu\Omega$-oriented if n is *odd*. Therefore, when integrating an $(n-1)$-form $\varphi\, du^1 \wedge \cdots \wedge du^{n-1}$ on ∂M, we need to take into account of the parity of n, i.e.

$$\int_{\mathsf{G}(\mathcal{V})\cap\partial M} \varphi\, du^1 \wedge \cdots \wedge du^{n-1} = (-1)^n \int_{\mathcal{V}\cap\{u_n=0\}} \varphi\, du^1 \cdots du^{n-1}. \tag{4.6}$$

The "extra" factor of $(-1)^n$ does not look nice at the first glance, but as we will see later, it will make Generalized Stokes' Theorem nicer. We are now ready to state Generalized Stokes' Theorem in a precise way:

Theorem 4.27 (Generalized Stokes' Theorem)**.** *Let M be an orientable smooth n-manifold, and let ω be a compactly supported smooth $(n-1)$-form on M. Then, we have:*

$$\int_M d\omega = \int_{\partial M} \omega. \tag{4.7}$$

Here if Ω is a chosen to be an orientation of M, then we will take $i_\nu\Omega$ to be the orientation of ∂M where ν is an outward-point normal vector of ∂M.

In particular, if $\partial M = \emptyset$, then $\int_M d\omega = 0$.

4.4.2. Proof of Generalized Stokes' Theorem. The proof consists of three steps:

Step 1: a special case where $\operatorname{supp}\omega$ is contained inside a single parametrization chart of interior type;

Step 2: another special case where $\operatorname{supp}\omega$ is contained inside a single parametrization chart of boundary type;

Step 3: use partitions of unity to deduce the general case.

Proof of Theorem 4.27. Throughout the proof, we will let Ω be the orientation of M, and $i_\nu\Omega$ be the orientation of ∂M with ν being an outward-point normal vector to ∂M. All local coordinate system $(u_1, \ldots, u_n)$ of M is assumed to be Ω-oriented.

Step 1: Suppose $\operatorname{supp}\omega$ *is contained in a single parametrization chart of interior type.*

Let $\mathsf{F}(u_1, \ldots, u_n) : \mathcal{U} \subset \mathbb{R}^n \to M$ be a local parametrization of interior type such that $\operatorname{supp}\omega \subset \mathsf{F}(\mathcal{U})$. Denote:

$$du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^n := du^1 \wedge \cdots \wedge du^{i-1} \wedge du^{i+1} \wedge \cdots \wedge du^n,$$

or in other words, it means the form with du^i removed.

In terms of local coordinates, the $(n-1)$-form ω can be expressed as:

$$\omega = \sum_{i=1}^n \omega_i \, du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^n.$$

Taking the exterior derivative, we get:

$$d\omega = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial \omega_i}{\partial u_j} du^j \wedge du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^n$$

For each i, the wedge product $du^j \wedge du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^n$ is zero if $j \neq i$. Therefore,

$$\begin{aligned} d\omega &= \sum_{i=1}^n \frac{\partial \omega_i}{\partial u_i} du^i \wedge du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^n \\ &= \sum_{i=1}^n (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \wedge \cdots \wedge du^i \wedge \cdots \wedge du^n \end{aligned}$$

By definition of integrals of differential forms, we get:

$$\int_M d\omega = \int_{\mathcal{U}} \sum_{i=1}^n (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \cdots du^n.$$

Since $\operatorname{supp}\omega \subset \mathsf{F}(\mathcal{U})$, the functions ω_i's are identically zero near and outside the boundary of $\mathcal{U} \subset \mathbb{R}^n$. Therefore, we can replace the domain of integration $\mathcal{U}$ of the RHS integral by a rectangle $[-R, R] \times \cdots \times [-R, R]$ in $\mathbb{R}^n$ where $R > 0$ is a sufficiently large number. The value of the integral is unchanged. Therefore, using the Fubini's Theorem, we get:

$$\begin{aligned}\int_M d\omega &= \int_{-R}^{R} \cdots \int_{-R}^{R} \sum_{i=1}^{n} (-1)^i \frac{\partial \omega_i}{\partial u_i} du^1 \cdots du^n \\ &= \sum_{i=1}^{n} (-1)^{i-1} \int_{-R}^{R} \cdots \int_{-R}^{R} \frac{\partial \omega_i}{\partial u_i} du^i \, du^1 \cdots \widehat{du^i} \cdots du^n \\ &= \sum_{i=1}^{n} (-1)^{i-1} \int_{-R}^{R} \cdots \int_{-R}^{R} [\omega_i]_{u_i=-R}^{u_i=R} \, du^1 \cdots \widehat{du^i} \cdots du^n.\end{aligned}$$

Since ω_i's vanish at the boundary of the rectangle $[-R, R]^n$, we have $\omega_i = 0$ when $u_i = \pm R$. As a result, we proved $\int_M d\omega = 0$. Since $\operatorname{supp} \omega$ is contained in a single parametrization chart of interior type, we have $\omega = 0$ on the boundary ∂M. Evidently, we have $\int_{\partial M} \omega = 0$ in this case. Hence, we proved

$$\int_M d\omega = \int_{\partial M} \omega = 0$$

in this case.

Step 2: Suppose $\operatorname{supp} \omega$ *is contained inside a single parametrization chart of boundary type.*

Let $\mathsf{G}(u_1, \ldots, u_n) : \mathcal{V} \subset \mathbb{R}^n_+ \to M$ be a local parametrization of boundary type such that $\operatorname{supp} \omega \subset \mathsf{G}(\mathcal{V})$. As in Step 1, we express

$$\omega = \sum_{i=1}^{n} \omega_i \, du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^n.$$

Proceed exactly in the same way as before, we arrive at:

$$\int_M d\omega = \int_{\mathcal{V}} \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \cdots du^n.$$

Now $\mathcal{V}$ is an open set in $\mathbb{R}^n_+$ instead of $\mathbb{R}^n$. Recall that the boundary is the set of points with $u_n = 0$. Therefore, this time we replace $\mathcal{V}$ by the half-space rectangle $[-R, R] \times \cdots \times [-R, R] \times [0, R]$ where $R > 0$ again is a sufficiently large number.

One key difference from Step 1 is that even though ω_i's has compact support inside $\mathcal{V}$, it may not vanish on the boundary of M. Therefore, we can only guarantee $\omega_i(u_1, \ldots, u_n) = 0$ when $u_n = R$, but we cannot claim $\omega_i = 0$ when $u_n = 0$. Some more work needs to be done:

$$\begin{aligned}\int_M d\omega &= \int_{\mathcal{V}} \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \cdots du^n \\ &= \int_0^R \int_{-R}^{R} \cdots \int_{-R}^{R} \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \cdots du^n \\ &= \sum_{i=1}^{n-1} (-1)^{i-1} \int_0^R \int_{-R}^{R} \cdots \int_{-R}^{R} (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \cdots du^n \\ &\quad + (-1)^{n-1} \int_0^R \int_{-R}^{R} \cdots \int_{-R}^{R} \frac{\partial \omega_n}{\partial u_n} du^1 \cdots du^n\end{aligned}$$

One can proceed as in Step 1 to show that the first term:

$$\sum_{i=1}^{n-1} (-1)^{i-1} \int_0^R \int_{-R}^{R} \cdots \int_{-R}^{R} (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \cdots du^n = 0,$$

which follows from the fact that whenever $1 \leq i \leq n-1$, we have $\omega_i = 0$ on $u_i = \pm R$.

The second term:

$$(-1)^{n-1}\int_0^R \int_{-R}^R \cdots \int_{-R}^R \frac{\partial \omega_n}{\partial u_n} du^1 \cdots du^n$$

is handled in a different way:

$$\begin{aligned}
&(-1)^{n-1}\int_0^R \int_{-R}^R \cdots \int_{-R}^R \frac{\partial \omega_n}{\partial u_n} du^1 \cdots du^n \\
&= (-1)^{n-1}\int_{-R}^R \cdots \int_{-R}^R \int_0^R \frac{\partial \omega_n}{\partial u_n}\, du^n\, du^1 \cdots du^{n-1} \\
&= (-1)^{n-1}\int_{-R}^R \cdots \int_{-R}^R [\omega_n]_{u_n=0}^{u_n=R}\, du^1 \cdots du^{n-1} \\
&= (-1)^{n}\int_{-R}^R \cdots \int_{-R}^R \omega_n(u_1, \ldots, u_{n-1}, 0)\, du^1 \cdots du^{n-1}
\end{aligned}$$

where we have used the following fact:

$$\begin{aligned}
[\omega_n(u_1, \ldots, u_n)]_{u_n=0}^{u_n=R} &= \omega_n(u_1, \ldots, u_{n-1}, R) - \omega_i(u_1, \ldots, u_{n-1}, 0) \\
&= 0 - \omega_n(u_1, \ldots, u_{n-1}, 0).
\end{aligned}$$

Combining all results proved so far, we have:

$$\int_M d\omega = (-1)^n \int_{-R}^R \cdots \int_{-R}^R \omega_n(u_1, \ldots, u_{n-1}, 0)\, du^1 \cdots du^{n-1}$$

On the other hand, we compute $\int_{\partial M} \omega$ and then compare it with $\int_M d\omega$. Note that the boundary ∂M are points with $u_n = 0$. Therefore, across the boundary ∂M, we have $du^n \equiv 0$, and so on ∂M we have:

$$\begin{aligned}
\omega &= \sum_{i=1}^{n} \omega_i(u_1, \ldots, u_{n-1}, 0)\, du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge \underbrace{du^n}_{=0} \\
&= \omega_n(u_1, \ldots, u_{n-1}, 0)\, du^1 \wedge \cdots \wedge du^{n-1} \\
\int_{\partial M} \omega &= \int_{\mathsf{G}(\mathcal{V}) \cap \partial M} \omega_n(u_1, \ldots, u_{n-1}, 0)\, du^1 \wedge \cdots \wedge du^{n-1} \\
&= (-1)^n \int_{\mathcal{V} \cap \{u_n = 0\}} \omega_n(u_1, \ldots, u_{n-1}, 0)\, du^1 \cdots du^{n-1} \\
&= (-1)^n \int_{-R}^R \cdots \int_{-R}^R \omega_n(u_1, \ldots, u_{n-1}, 0)\, du^1 \cdots du^{n-1}
\end{aligned}$$

Recall that we have a factor of $(-1)^n$ because the local coordinate system $(u_1, \ldots, u_{n-1})$ for ∂M is $i_\nu \Omega$ if and only if n is even, as discussed in the previous subsection.

Consequently, we have proved

$$\int_M d\omega = \int_{\partial M} \omega$$

in this case.

Step 3: Use partitions of unity to deduce the general case

Finally, we "glue" the previous two steps together and deduce the general case. Let $\mathcal{A} = \{\mathsf{F}_\alpha : \mathcal{U}_\alpha \to M\}$ be an atlas of M where all local coordinates are Ω-oriented. Here $\mathcal{A}$ contain both interior and boundary types of local parametrizations. Suppose $\{\rho_\alpha : M \to [0,1]\}$ is a partition of unity subordinate to $\mathcal{A}$. Then, we have:

$$\omega = \underbrace{\left(\sum_\alpha \rho_\alpha\right)}_{\equiv 1} \omega = \sum_\alpha \rho_\alpha \omega$$

$$\int_{\partial M} \omega = \int_{\partial M} \sum_\alpha \rho_\alpha \omega = \sum_\alpha \int_{\partial M} \rho_\alpha \omega.$$

For each α, the $(n-1)$-form $\rho_\alpha \omega$ is compactly supported in a single parametrization chart (either of interior or boundary type). From Step 1 and Step 2, we have already proved that Generalized Stokes' Theorem is true for each $\rho_\alpha \omega$. Therefore, we have:

$$\begin{aligned}
\sum_\alpha \int_{\partial M} \rho_\alpha \omega &= \sum_\alpha \int_M d(\rho_\alpha \omega) \\
&= \sum_\alpha \int_M (d\rho_\alpha \wedge \omega + \rho_\alpha \, d\omega) \\
&= \int_M d\left(\sum_\alpha \rho_\alpha\right) \wedge \omega + \left(\sum_\alpha \rho_\alpha\right) d\omega.
\end{aligned}$$

Since $\sum_\alpha \rho_\alpha \equiv 1$ and hence $d\left(\sum_\alpha \rho_\alpha\right) \equiv 0$, we have proved:

$$\int_{\partial M} \omega = \sum_\alpha \int_{\partial M} \rho_\alpha \omega = \int_M 0 \wedge \omega + 1 \, d\omega = \int_M d\omega.$$

It completes the proof of Generalized Stokes' Theorem. □

Remark 4.28. As we can see from that the proof (Step 2), if we simply choose an orientation for ∂M such that $(u_1, \ldots, u_{n-1})$ becomes the order of local coordinates for ∂M, then (4.7) would have a factor of $(-1)^n$ on the RHS, which does not look nice. Moreover, if we pick $i_{-\nu}\Omega$ to be the orientation of ∂M (here $-\nu$ is then an inward-pointing normal to ∂M), then the RHS of (4.7) would have a minus sign, which is not nice either. □

4.4.3. Fundamental Theorems in Vector Calculus. We briefly discussed at the end of Chapter 3 how the three fundamental theorems in Vector Calculus, namely Green's, Stokes' and Divergence Theorems, can be formulated using differential forms. Given that we now have proved Generalized Stokes' Theorem (Theorem 4.27), we are going to give a formal proof of the three Vector Calculus theorems in MATH 2023 using the Theorem 4.27.

Corollary 4.29 (Green's Theorem). *Let R be a closed and bounded smooth 2-submanifold in $\mathbb{R}^2$ with boundary ∂R. Given any smooth vector field $\mathsf{V} = (P(x,y), Q(x,y))$ defined in R, then we have:*

$$\oint_{\partial R} \mathsf{V} \cdot d\mathsf{r} = \int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \, dy,$$

The line integral on the LHS is oriented such that $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\}$ has the same orientation as $\{\nu, \mathsf{T}\}$ where ν is the outward-pointing normal of R, and T is the velocity vector of the curve ∂R. See Figure 4.3.

Proof. Consider the 1-form $\omega := P\,dx + Q\,dy$ defined on R, then we have:

$$d\omega = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy.$$

Suppose we fix an orientation $\Omega = dx \wedge dy$ for R so that the order of coordinates is (x, y), then by generalized Stokes' Theorem we get:

$$\underbrace{\oint_{\partial R} P\,dx + Q\,dy}_{\oint_{\partial R}\omega} = \underbrace{\int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy}_{\int_R d\omega} = \underbrace{\int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy}_{(x,y) \text{ is the orientation}}.$$

The only thing left to figure out is the orientation of the line integral. Locally parametrize R by local coordinates (s, t) so that $\{t = 0\}$ is the boundary ∂R and $\{t > 0\}$ is the interior of R (see Figure 4.3). By convention, the local coordinate s for ∂R must be chosen so that $\Omega(\nu, \frac{\partial}{\partial s}) > 0$ where ν is a outward-pointing normal vector to ∂R. In other words, the pair $\{\nu, \frac{\partial}{\partial s}\}$ should have the same orientation as $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\}$. According to Figure 4.3, we must choose the local coordinate s for ∂R such that for the outer boundary, s goes counter-clockwisely as it increases; whereas for each inner boundary, s goes clockwisely as it increases. $\square$

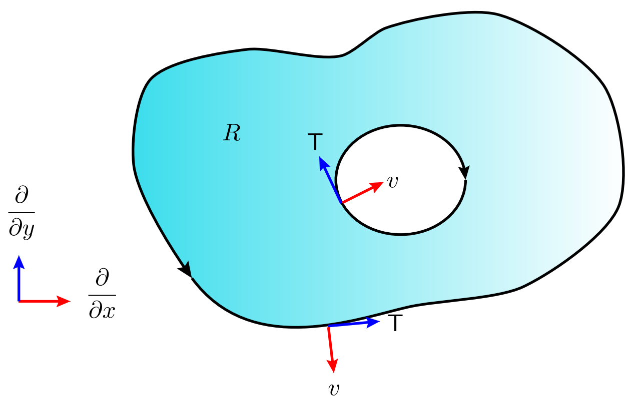


Figure 4.3. Orientation of Green's Theorem

Next we show that Stokes' Theorem in Multivariable Calculus is also a consequence of Generalized Stokes' Theorem. Recall that in MATH 2023 we learned about surface integrals. If $\mathsf{F}(u, v) : \mathcal{U} \to \Sigma \subset \mathbb{R}^3$ is a parametrization of the whole surface Σ, then we define the surface element as:

$$dS = \left|\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v}\right| du\,dv,$$

and the surface integral of a scalar function φ is defined as:

$$\int_\Sigma \varphi\,dS = \int_{\mathcal{U}} \varphi(u, v) \left|\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v}\right| du\,dv.$$

However, not every surface can be covered (or almost covered) by a single parametrization chart. Generally, if $\mathcal{A} = \{\mathsf{F}_\alpha(u_\alpha, v_\alpha) : \mathcal{U}_\alpha \to \mathbb{R}^3\}$ is an oriented atlas of Σ with a partition of unity $\{\rho_\alpha : \Sigma \to [0, 1]\}$ subordinate to $\mathcal{A}$, we then define:

$$dS := \sum_\alpha \rho_\alpha \left|\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}\right| du_\alpha\,dv_\alpha.$$

Corollary 4.30 (Stokes' Theorem). *Let Σ be a closed and bounded smooth 2-submanifold in $\mathbb{R}^3$ with boundary $\partial\Sigma$, and $\mathsf{V} = (P(x,y,z), Q(x,y,z), R(x,y,z))$ be a vector field which is smooth on Σ, then we have:*

$$\oint_{\partial\Sigma} \mathsf{V} \cdot d\mathsf{r} = \int_{\Sigma} (\nabla \times \mathsf{V}) \cdot \mathsf{N}\, dS.$$

Here $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$ has the same orientation as $\{\nu, \mathsf{T}, \mathsf{N}\}$, where ν is the outward-point normal vector of Σ at points of $\partial\Sigma$, T is the velocity vector of $\partial\Sigma$, and N is the unit normal vector to Σ in $\mathbb{R}^3$. See Figure 4.4.

Proof. Define:

$$\omega = P\,dx + Q\,dy + R\,dz$$

which is viewed as a 1-form on Σ. Then,

$$\oint_{\partial\Sigma} \omega = \oint_{\partial\Sigma} \mathsf{V} \cdot d\mathsf{r}. \tag{4.8}$$

By direct computation, the 2-form $d\omega$ is given by:

$$d\omega = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dz \wedge dx + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dy \wedge dz.$$

Now consider an oriented atlas $\mathcal{A} = \{\mathsf{F}_\alpha(u_\alpha, v_\alpha) : \mathcal{U}_\alpha \to \mathbb{R}^3\}$ of Σ with a partition of unity $\{\rho_\alpha : \Sigma \to [0,1]\}$, then according to the discussion near the end of Chapter 3, we can express each of $dx \wedge dy$, $dz \wedge dx$ and $dy \wedge dz$ in terms of $du_\alpha \wedge dv_\alpha$, and obtain:

$$\begin{aligned}
d\omega &= \sum_\alpha \rho_\alpha\, d\omega \\
&= \sum_\alpha \rho_\alpha \left[\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dz \wedge dx + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dy \wedge dz\right] \\
&= \sum_\alpha \rho_\alpha \left\{ \left(\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \det \frac{\partial(y,z)}{\partial(u_\alpha, v_\alpha)} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) \det \frac{\partial(z,x)}{\partial(u_\alpha, v_\alpha)} \right.\right. \\
&\qquad \left.\left. + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) \det \frac{\partial(x,y)}{\partial(u_\alpha, v_\alpha)} \right) \right\} du_\alpha \wedge dv_\alpha.
\end{aligned}$$

On each local coordinate chart $\mathsf{F}_\alpha(\mathcal{U}_\alpha)$, a normal vector to Σ in $\mathbb{R}^3$ can be found using cross products:

$$\begin{aligned}
\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha} &= \det \frac{\partial(y,z)}{\partial(u_\alpha, v_\alpha)}\mathsf{i} + \det \frac{\partial(z,x)}{\partial(u_\alpha, v_\alpha)}\mathsf{j} + \det \frac{\partial(x,y)}{\partial(u_\alpha, v_\alpha)}\mathsf{k} \\
\nabla \times \mathsf{V} &= \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathsf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathsf{j} + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathsf{k}.
\end{aligned}$$

Hence,

$$d\omega = \sum_\alpha (\nabla \times \mathsf{V}) \cdot \left(\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}\right) \rho_\alpha\, du_\alpha \wedge dv_\alpha,$$

and so

$$\int_\Sigma d\omega = \sum_\alpha \int_{\mathcal{U}_\alpha} (\nabla \times \mathsf{V}) \cdot \left(\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}\right) \rho_\alpha\, du_\alpha\, dv_\alpha.$$

Denote $\mathsf{N} = \dfrac{\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}}{\left|\frac{\partial \mathsf{F}}{\partial u_\alpha} \times \frac{\partial \mathsf{F}}{\partial v_\alpha}\right|}$, and recall the fact that $dS := \sum_\alpha \rho_\alpha \left|\dfrac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \dfrac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}\right| du_\alpha\, dv_\alpha$, we get:

$$\int_\Sigma d\omega = \int_\Sigma (\nabla \times \mathsf{V}) \cdot \mathsf{N}\, dS. \tag{4.9}$$

Combining the results of (4.8) and (4.9), using Generalized Stokes' Theorem (Theorem 4.7, we get:

$$\oint_{\partial\Sigma} \mathsf{V} \cdot d\mathsf{r} = \int_{\Sigma} (\nabla \times \mathsf{V}) \cdot \mathsf{N}\, dS$$

as desired.

To see the orientation of $\partial\Sigma$, we locally parametrize Σ by coordinates (s,t) such that $\{t=0\}$ are points on $\partial\Sigma$, and so $\partial\Sigma$ is locally parametrized by s. The outward-pointing normal of $\partial\Sigma$ in Σ is given by $\nu := -\frac{\partial}{\partial t}$. By convention, the orientation of $\left\{\nu, \frac{\partial}{\partial s}\right\}$ is the same as $\left\{\frac{\partial}{\partial u_\alpha}, \frac{\partial}{\partial v_\alpha}\right\}$, and hence:

$$\left\{\nu, \frac{\partial}{\partial s}, \mathsf{N}\right\} \text{ has the same orientation as } \left\{\frac{\partial}{\partial u_\alpha}, \frac{\partial}{\partial v_\alpha}, \mathsf{N}\right\}.$$

As $\mathsf{N} = \dfrac{\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}}{\left|\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}\right|}$, the set $\left\{\frac{\partial}{\partial u_\alpha}, \frac{\partial}{\partial v_\alpha}, \mathsf{N}\right\}$ has the same orientation as $\{\mathsf{i},\mathsf{j},\mathsf{k}\}$. As a result, the set $\{\nu, \frac{\partial}{\partial s}, \mathsf{N}\}$ is oriented in the way as in Figure 4.4. □

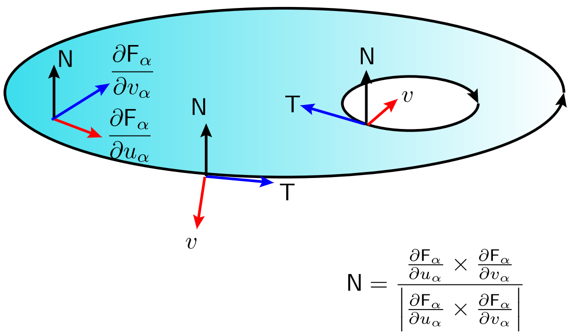


Figure 4.4. Orientation of Stokes' Theorem

Finally, we discuss how to use Generalized Stokes' Theorem to prove Divergence Theorem in Multivariable Calculus.

Corollary 4.31 (Divergence Theorem). *Let D be a closed and bounded 3-submanifold of $\mathbb{R}^3$ with boundary ∂D, and $\mathsf{V} = (P(x,y,z), Q(x,y,z), R(x,y,z))$ be a smooth vector field defined on D. Then, we have:*

$$\oint_{\partial D} \mathsf{V} \cdot \mathsf{N}\, dS = \int_D \nabla \cdot \mathsf{V}\, dx\, dy\, dz.$$

Here N is the unit normal vector of ∂D in $\mathbb{R}^3$ which points away from D.

Proof. Let $\omega := P\, dy \wedge dz + Q\, dz \wedge dx + R\, dx \wedge dy$. Then by direct computations, we get:

$$d\omega = \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dx \wedge dy \wedge dz = \nabla \cdot \mathsf{V}\, dx \wedge dy \wedge dz.$$

Using $\{\mathsf{i},\mathsf{j},\mathsf{k}\}$ as the orientation for D, then it is clear that:

$$\int_D d\omega = \int_D \nabla \cdot \mathsf{V}\, dx\, dy\, dz. \tag{4.10}$$

Consider an atlas $\mathcal{A} = \{\mathsf{F}_\alpha(u_\alpha, v_\alpha, w_\alpha) : \mathcal{U}_\alpha \to \mathbb{R}^3\}$ of D such that for the local parametrization of boundary type, the boundary points are given by $\{w_\alpha = 0\}$, and interior points are $\{w_\alpha > 0\}$. Then, ∂D is locally parametrized by (u_α, v_α).

As a convention, the orientation of (u_α, v_α) is chosen such that $\{-\frac{\partial}{\partial w_\alpha}, \frac{\partial}{\partial u_\alpha}, \frac{\partial}{\partial v_\alpha}\}$ has the same orientation as $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$, or equivalently, $\{\frac{\partial}{\partial u_\alpha}, \frac{\partial}{\partial v_\alpha}, -\frac{\partial}{\partial w_\alpha}\}$ has the same orientation as $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$.

Furthermore, let N be the unit normal of ∂D given by:

$$\mathsf{N} = \frac{\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}}{\left|\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}\right|}.$$

By the convention of cross products, $\{\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha}, \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}, \mathsf{N}\}$ must have the same orientation as $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$. Now that $\{\frac{\partial}{\partial u_\alpha}, \frac{\partial}{\partial v_\alpha}, -\frac{\partial}{\partial w_\alpha}\}$ and $\{\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha}, \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha}, \mathsf{N}\}$ have the same orientation, so N and $-\frac{\partial}{\partial w_\alpha}$ are both pointing in the same direction. In other words, N is the outward-point normal.

The rest of the proof goes by writing ω in terms of $du_\alpha \wedge dv_\alpha$ on each local coordinate chart:

$$\begin{aligned}
\omega &= \sum_\alpha \rho_\alpha \omega \\
&= \sum_\alpha \rho_\alpha \left(P \det \frac{\partial(y,z)}{\partial(u_\alpha, v_\alpha)} + Q \det \frac{\partial(z,x)}{\partial(u_\alpha, v_\alpha)} + R \det \frac{\partial(x,y)}{\partial(u_\alpha, v_\alpha)} \right) du_\alpha \wedge dv_\alpha \\
&= \sum_\alpha \mathsf{V} \cdot \left(\frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha} \right) \rho_\alpha \, du_\alpha \wedge dv_\alpha \\
&= \sum_\alpha \mathsf{V} \cdot \mathsf{N} \, \rho_\alpha \left| \frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha} \right| du_\alpha \wedge dv_\alpha
\end{aligned}$$

Therefore, we get:

$$\oint_{\partial D} \omega = \oint_{\partial D} \sum_\alpha \mathsf{V} \cdot \mathsf{N} \, \rho_\alpha \left| \frac{\partial \mathsf{F}_\alpha}{\partial u_\alpha} \times \frac{\partial \mathsf{F}_\alpha}{\partial v_\alpha} \right| du_\alpha \, dv_\alpha = \oint_{\partial D} \mathsf{V} \cdot \mathsf{N} \, dS. \tag{4.11}$$

Combining with (4.10), (4.11) and Generalized Stokes' Theorem, the proof of this corollary is completed. □

De Rham Cohomology

> "Hydrodynamics procreated complex analysis, partial differential equations, Lie groups and algebra theory, cohomology theory and scientific computing."
>
> *Vladimir Arnold*

In Chapter 3, we discussed closed and exact forms. As a reminder, a smooth k-form ω on a smooth manifold M is *closed* if $d\omega = 0$ on M, and is *exact* if $\omega = d\eta$ for some smooth $(k-1)$-form η defined on the whole M.

By the fact that $d^2 = 0$, an exact form must be closed. It is then natural to ask whether every closed form is exact. The answer is no in general. Here is a counter-example. Let $M = \mathbb{R}^2 \backslash \{(0,0)\}$, and define

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

It can be computed easily that $d\omega = 0$ on M, and so ω is closed.

However, we can show that ω is *not* exact. Consider the unit circle C parametrized by $(x,y) = (\cos t, \sin t)$ where $0 < t < 2\pi$, and also the induced 1-form $\iota^*\omega$ (where $\iota : C \to M$). By direct computation, we get:

$$\oint_C \iota^*\omega = \int_0^{2\pi} -\frac{\sin t}{\cos^2 t + \sin^2 t}\,d(\cos t) + \frac{\cos t}{\cos^2 t + \sin^2 t}\,d(\sin t) = 2\pi.$$

If ω *were* exact, then $\omega = df$ for some smooth function $f : M \to \mathbb{R}$. Then, we would have:

$$\oint_C \iota^*\omega = \oint_C \iota^*(df) = \oint_C d(\iota^* f) = \int_0^{2\pi} \frac{d(\iota^* f)}{dt}\,dt.$$

Since $t = 0$ and $t = 2\pi$ represent the same point on C, by Fundamental Theorem of Calculus, we finally get:

$$\oint_C \iota^*\omega = 0$$

which is a contradiction! Therefore, ω is *not* exact on $\mathbb{R}^2 \backslash \{(0,0)\}$.

Heuristically, de Rham cohomology studies "how many" smooth k-forms defined on a given manifold M are closed but not exact. We should refine the meaning of "how many". Certainly, if η is any $(k-1)$-form on M, then $\omega + d\eta$ is also closed but not exact. Therefore, when we "count" how many smooth k-forms on M which are closed but not exact, it is fair to group ω and $\omega + d\eta$'s together, and count them as one. In formal mathematical language, equivalence classes are used as we will discuss in detail. It turns out that the "number" of closed, not exact k-forms on a given M is a related to the *topology* of M!

In this chapter, we will learn the basics of de Rham cohomology, which is a beautiful topic to end the course MATH 4033.

5.1. De Rham Cohomology

Let M be a smooth manifold (with or without boundary). Recall that the exterior derivative d is a linear map that takes a k-form to a $(k+1)$-form, i.e. $d : \wedge^k T^*M \to \wedge^{k+1} T^*M$. We can then talk about the kernel and image of these maps. We define:

$$\begin{aligned}\ker\left(d : \wedge^k T^*M \to \wedge^{k+1} T^*M\right) &= \{\omega \in \wedge^k T^*M : d\omega = 0\} \\ &= \{\text{closed } k\text{-forms on } M\} \\ \operatorname{Im}\left(d : \wedge^{k-1} T^*M \to \wedge^k T^*M\right) &= \{\omega \in \wedge^k T^*M : \omega = d\eta \text{ for some } \eta \in \wedge^{k-1} T^*M\} \\ &= \{\text{exact } k\text{-forms on } M\}\end{aligned}$$

In many occasions, we may simply denote the above kernel and image by $\ker(d)$ and $\operatorname{Im}(d)$ whenever the value of k is clear from the context.

By $d^2 = 0$, it is easy to see that:

$$\operatorname{Im}\left(d : \wedge^{k-1} T^*M \to \wedge^k T^*M\right) \subset \ker\left(d : \wedge^k T^*M \to \wedge^{k+1} T^*M\right).$$

If all closed k-forms on a certain manifold are exact, then we have $\operatorname{Im}(d) = \ker(d)$. How "many" closed k-forms are exact is then measured by how $\operatorname{Im}(d)$ is "smaller" than $\ker(d)$, which is precisely measured by the size of the quotient vector space $\ker(d)/\operatorname{Im}(d)$. We call this quotient the de Rham cohomology group[1].

> **Definition 5.1** (de Rham Cohomology Group). Let M be a smooth manifold. For any positive integer k, we define the *k-th de Rham cohomology group of M* to be the quotient vector space:
>
> $$H^k_{\mathrm{dR}}(M) := \frac{\ker\left(d : \wedge^k T^*M \to \wedge^{k+1} T^*M\right)}{\operatorname{Im}\left(d : \wedge^{k-1} T^*M \to \wedge^k T^*M\right)}.$$

Remark 5.2. When $k = 0$, then $\wedge^k T^*M = \wedge^0 T^*M = C^\infty(M, \mathbb{R})$ and $\wedge^{k-1} T^*M$ is not defined. Instead, we *define*

$$H^0_{\mathrm{dR}}(M) := \ker\left(d : C^\infty(M, \mathbb{R}) \to \wedge^1 T^*M\right) = \{f \in C^\infty(M, \mathbb{R}) : df = 0\},$$

which is the vector space of all locally constant functions on M. If M has N connected components, then a locally constant function f is determined by its value on each of the components. The space of functions $\{f : df = 0\}$ is in one-to-one correspondence an N-tuple $(k_1, \ldots, k_N) \in \mathbb{R}^N$, where k_i is the value of f on the i-th component of M. Therefore, $H^0_{\mathrm{dR}}(M) \simeq \mathbb{R}^N$ where N is the number of connected components of M. □

---

[1] A vector space is also a group whose addition is the vector addition. Although it is more appropriate or precise to call the quotient the "de Rham cohomology space", we will follow the history to call it a group.

5.1.1. Quotient Vector Spaces. Let's first review the basics about quotient vector spaces in Linear Algebra. Given a subspace W of a vector space V, we can define an equivalence relation $\sim$ by declaring that $v_1 \sim v_2$ if and only if $v_1 - v_2 \in W$. For example, if W is the x-axis and V is the xy-plane, then two vector v_1 and v_2 are equivalent under this relation if and only if they have the same j-component.

For each element $v \in V$ (the bigger space), one can define an equivalence class:

$$[v] := \{u \in V : u \sim v\} = \{u \in V : u - v \in W\}$$

which is the set of all vectors in V that are equivalent to v. For example, if W is the x-axis and V is $\mathbb{R}^2$, then the class $[(2,3)]$ is given by:

$$[(2,3)] = \{(x,3) : x \in \mathbb{R}\}$$

which is the horizontal line $\{y = 3\}$. Similarly, one can figure out $[(1,3)] = [(2,3)] = [(3,3)] = \ldots$ as well, but $[(2,3)] \neq [(2,2)]$, and the latter is the line $\{y = 2\}$.

The quotient space V/W is defined to be the set of all equivalence classes, i.e.

$$V/W := \{[v] : v \in V\}.$$

For example, if V is $\mathbb{R}^2$ and W is the x-axis, then V/W is the set of all horizontal lines in $\mathbb{R}^2$. For finite dimensional vector spaces, one can show (see Exercise 5.1) that

$$\dim(V/W) = \dim V - \dim W,$$

and so the "size" (precisely, the dimension) of the quotient V/W measures how small W is when compared to V. In fact, if the bases of V and W are suitably chosen, we can describe the basis of V/W in a precise way (see Exercise 5.1).

Exercise 5.1. Let W be a subspace of a finite dimensional vector space V. Suppose $\{w_1, \ldots, w_k\}$ is a basis for W, and $\{w_1, \ldots, w_k, v_1, \ldots, v_l\}$ is a basis for V (Remark: given any basis $\{w_1, \ldots, w_k\}$ for the subspace W, one can always complete it to form a basis for V).

(a) Show that given any vector $\sum_{i=1}^k \alpha_i w_i + \sum_{j=1}^l \beta_j v_j \in V$, the equivalence class represented by this vector is given by:

$$\left[\sum_{i=1}^k \alpha_i w_i + \sum_{j=1}^l \beta_j v_j\right] = \left\{\sum_{i=1}^k \gamma_i w_i + \sum_{j=1}^l \beta_j v_j : \gamma_i \in \mathbb{R}\right\} = \left[\sum_{j=1}^l \beta_j v_j\right].$$

(b) Hence, show that $\{[v_1], \ldots, [v_l]\}$ is a basis for V/W, and so

$$\dim V/W = l = \dim V - \dim W.$$

Exercise 5.2. Given a subspace W of a vector space V, and define an equivalence relation $\sim$ by declaring that $v_1 \sim v_2$ if and only if $v_1 - v_2 \in W$. Show that the following are equivalent:

(1) $u \in [v]$

(2) $u - v \in W$

(3) $[u] = [v]$

5.1.2. Cohomology Classes and Betti numbers. Recall that the k-th de Rham cohomology group $H^k_{\mathrm{dR}}(M)$, where $k \geq 1$, of a smooth manifold M is defined to be the quotient vector space:

$$H^k_{\mathrm{dR}}(M) := \frac{\ker\left(d : \wedge^k T^*M \to \wedge^{k+1} T^*M\right)}{\mathrm{Im}\ \left(d : \wedge^{k-1} T^*M \to \wedge^k T^*M\right)}.$$

Given a closed k-form ω, we then define its equivalence class to be:

$$\begin{aligned}[\omega] &:= \{\omega' : \omega' - \omega \text{ is exact}\} \\ &= \{\omega' : \omega' = \omega + d\eta \text{ for some } \eta \in \wedge^{k-1}T^*M\} \\ &= \{\omega + d\eta : \eta \in \wedge^{k-1}T^*M\}.\end{aligned}$$

An equivalence class $[\omega]$ is called the *de Rham cohomology class* represented by (or containing) ω, and ω is said to be a representative of this de Rham cohomology class.

By Exercise 5.1, its dimension is given by

$$\begin{aligned}&\dim H^k_{\mathrm{dR}}(M) \\ &= \dim\ker\left(d : \wedge^k T^*M \to \wedge^{k+1}T^*M\right) - \dim\operatorname{Im}\left(d : \wedge^{k-1}T^*M \to \wedge^k T^*M\right)\end{aligned}$$

provided that both kernel and image are finite-dimensional.

Therefore, the dimension of $H^k_{\mathrm{dR}}(M)$ is a measure of "how many" closed k-forms on M are not exact. Due to the importance of this dimension, we have a special name for it:

Definition 5.3 (Betti Numbers)**.** Let M be a smooth manifold. The *k-th Betti number* of M is defined to be:

$$b_k(M) := \dim H^k_{\mathrm{dR}}(M).$$

In particular, $b_0(M) = \dim H^0_{\mathrm{dR}}(M)$ is the number of connected components of M.

In case when $M = \mathbb{R}^2\backslash\{(0,0)\}$, we discussed that there is a closed 1-form

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

defined on M which is not exact. Therefore, $\omega \in \ker\left(d : \wedge^1 T^*M \to \wedge^2 T^*M\right)$ yet $\omega \notin \operatorname{Im}\left(d : \wedge^0 T^*M \to \wedge^1 T^*M\right)$, and so in $H^1_{\mathrm{dR}}(M)$ we have $[\omega] \neq [0]$. From here we can conclude that $H^1_{\mathrm{dR}}(M) \neq \{[0]\}$ and $b_1(M) \geq 1$. We will later show that in fact $b_1(M) = 1$ using some tools in later sections.

Exercise 5.3. If $k > \dim M$, what can you say about $b_k(M)$?

5.1.3. Poincaré Lemma. A *star-shaped* open set U in $\mathbb{R}^n$ is a region containing a point $\mathsf{p} \in U$ (call it a base point) such that any line segment connecting a point $\mathsf{x} \in U$ and the base point p must be contained inside U. Examples of star-shaped open sets include convex open sets such an open ball $\{\mathsf{x} \in \mathbb{R}^n : |x| < 1\}$, and all of $\mathbb{R}^n$. The following Poincaré Lemma asserts that $H^1_{\mathrm{dR}}(U) = \{[0]\}$.

Theorem 5.4 (Poincaré Lemma for H^1_{dR})**.** *For any star-shaped open set U in $\mathbb{R}^n$, we have $H^1_{\mathrm{dR}}(U) = \{[0]\}$. In other words, any closed 1-form defined on a star-shaped open set is exact on that open set.*

Proof. Given a closed 1-form ω defined on U, given by $\omega = \sum_i \omega_i\,dx^i$, we need to find a smooth function $f : U \to \mathbb{R}$ such that $\omega = df$. In other words, we need $\dfrac{\partial f}{\partial x_i} = \omega_i$ for any i.

Let p be the base point of U, then given any $\mathsf{x} \in U$, we define:

$$f(\mathsf{x}) := \int_{L_{\mathsf{x}}} \omega$$

where L_{x} is the line segment joining p and x, which can be parametrized by:

$$\mathsf{r}(t) = (1-t)\mathsf{p} + t\mathsf{x}, \quad t \in [0,1].$$

Write $\mathsf{p} = (p_1, \ldots, p_n)$, $\mathsf{x} = (x_1, \ldots, x_n)$, then $f(x)$ can be expressed in terms of t by:

$$f(\mathsf{x}) = \int_0^1 \sum_{i=1}^n \omega_i(\mathsf{r}(t)) \cdot (x_i - p_i)\,dt.$$

Using the chain rule, we can directly verify that:

$$\begin{aligned}
\frac{\partial f}{\partial x_j}(\mathsf{x}) &= \frac{\partial}{\partial x_j}\int_0^1 \sum_{i=1}^n \omega_i(\mathsf{r}(t)) \cdot (x_i - p_i)\,dt \\
&= \sum_{i=1}^n \int_0^1 \left(\frac{\partial}{\partial x_j}\omega_i(\mathsf{r}(t)) \cdot (x_i - p_i) + \omega_i(\mathsf{r}(t)) \cdot \frac{\partial}{\partial x_j}(x_i - p_i) \right) dt \\
&= \sum_{i=1}^n \int_0^1 \left(\sum_{k=1}^n \frac{\partial \omega_i}{\partial x_k} \frac{\partial \overbrace{((1-t)p_k + tx_k)}^{x_k \circ \mathsf{r}(t)}}{\partial x_j} \cdot (x_i - p_i) + \omega_i(\mathsf{r}(t)) \cdot \delta_{ij} \right) dt \\
&= \sum_{i=1}^n \int_0^1 \left(\sum_{k=1}^n t\frac{\partial \omega_i}{\partial x_k}\delta_{jk} \cdot (x_i - p_i) + \omega_j(\mathsf{r}(t)) \right) dt \\
&= \sum_{i=1}^n \int_0^1 \left(t\frac{\partial \omega_i}{\partial x_j} \cdot (x_i - p_i) + \omega_j(\mathsf{r}(t)) \right) dt
\end{aligned}$$

Since ω is closed, we have:

$$0 = d\omega = \sum_{i<j}^n \left(\frac{\partial \omega_i}{\partial x_j} - \frac{\partial \omega_j}{\partial x_i} \right) dx^j \wedge dx^i$$

and hence $\dfrac{\partial \omega_i}{\partial x_j} = \dfrac{\partial \omega_j}{\partial x_i}$ for any i, j. Using this to proceed our calculation:

$$\begin{aligned}
\frac{\partial f}{\partial x_j}(\mathsf{x}) &= \int_0^1 \left(t\frac{\partial \omega_j}{\partial x_i} \cdot (x_i - p_i) + \omega_j(\mathsf{r}(t)) \right) dt \\
&= \int_0^1 \frac{d}{dt}\left(t\omega_j(\mathsf{r}(t)) \right) dt \\
&= \left[t\omega_j(\mathsf{r}(t)) \right]_{t=0}^{t=1} = \omega_j(\mathsf{r}(1)) = \omega_j(\mathsf{x}).
\end{aligned}$$

In the second equality above, we have used the chain rule backward:

$$\frac{d}{dt}\left(t\omega_j(\mathsf{r}(t)) \right) = t\frac{\partial \omega_j}{\partial x_i} \cdot (x_i - p_i) + \omega_j(\mathsf{r}(t)).$$

From this, we conclude that $\omega = df$ on U, and hence $[\omega] = [0]$ in $H^1_{\mathrm{dR}}(U)$. Since ω is an arbitrary closed 1-form on U, we have $H^1_{\mathrm{dR}}(U) = \{[0]\}$. □

Remark 5.5. Poincaré Lemma also holds for H^k_{dR}, meaning that if U is a star-shaped open set in $\mathbb{R}^n$, then $H^k_{\mathrm{dR}}(U) = \{[0]\}$ for any $k \geq 1$. However, the proof involves the use of Lie derivatives and a formula by Cartan, both of which are beyond the scope of this course. Note also that $H^0_{\mathrm{dR}}(U) \simeq \mathbb{R}$ since a star-shaped open set must be connected. □

Remark 5.6. We have discussed that the 1-form

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

is closed but not exact. To be precise, it is not exact *on* $\mathbb{R}^2 \backslash \{(0,0)\}$. However, if we regard the domain to be the first quadrant $U := \{(x,y) : x > 0 \text{ and } y > 0\}$, which is a star-shaped open set in $\mathbb{R}^2$, then by Poinaré Lemma (Theorem 5.4), ω is indeed an exact 1-form on U. In fact, it is not difficult to verify that

$$\omega = d\left(\tan^{-1}\frac{y}{x}\right) \quad \text{on } U.$$

Note that the scalar function $\tan^{-1}\frac{y}{x}$ is smoothly defined on U. Whether a form is exact or not depends on the choice of its domain! □

5.1.4. Diffeomorphic Invariance. By Proposition 3.57, we learned that the exterior derivative d commutes with the pull-back of a smooth map between two manifolds. An important consequence is that the de Rham cohomology group is invariant under diffeomorphism.

Let $\Phi : M \to N$ be any smooth map between two smooth manifolds. The pull-back map $\Phi^* : \wedge^k T^*N \to \wedge^k T^*M$ induces a well-defined pull-back map (which is also denoted by Φ^*) from $H^k_{\mathrm{dR}}(N)$ to $H^k_{\mathrm{dR}}(M)$. Precisely, given any closed k-form ω on N, we define:

$$\Phi^*[\omega] := [\Phi^*\omega].$$

$\Phi^*\omega$ is a k-form on M. It is closed since $d(\Phi^*\omega) = \Phi^*(d\omega) = \Phi^*(0) = 0$. To show it is well-defined, we take another k-form ω' on N such that $[\omega'] = [\omega]$ in $H^k_{\mathrm{dR}}(N)$. Then, there exists a $(k-1)$-form η on N such that:

$$\omega' - \omega = d\eta \quad \text{on } N.$$

Using again $d \circ \Phi^* = \Phi^* \circ d$, we get:

$$\Phi^*\omega' - \Phi^*\omega = \Phi^*(d\eta) = d(\Phi^*\eta) \quad \text{on } M$$

We conclude $\Phi^*\omega' - \Phi^*\omega$ is exact and so

$$[\Phi^*\omega'] = [\Phi^*\omega] \text{ in } H^k_{\mathrm{dR}}(M).$$

This shows $\Phi^* : H^k_{\mathrm{dR}}(N) \to H^k_{\mathrm{dR}}(M)$ is a well-defined map.

Theorem 5.7 (Diffeomorphism Invariance of H^k_{dR})**.** *If two smooth manifolds M and N are diffeomorphic, then $H^k_{\mathrm{dR}}(M)$ and $H^k_{\mathrm{dR}}(N)$ are isomorphic for any $k \geq 0$.*

Proof. Let $\Phi : M \to N$ be a diffeomorphism, then $\Phi^{-1} : N \to M$ exists and we have $\Phi \circ \Phi^{-1} = \mathrm{id}_N$ and $\Phi^{-1} \circ \Phi = \mathrm{id}_M$. By the chain rule for tensors (Theorem 3.54), we have:

$$(\Phi^{-1})^* \circ \Phi^* = \mathrm{id}_{\wedge^k T^*N} \quad \text{and} \quad \Phi^* \circ (\Phi^{-1})^* = \mathrm{id}_{\wedge^k T^*M}.$$

Given any closed k-form ω on M, then in $H^k_{\mathrm{dR}}(M)$ we have:

$$\Phi^* \circ (\Phi^{-1})^*[\omega] = \Phi^*[(\Phi^{-1})^*\omega] = [\Phi^* \circ (\Phi^{-1})^*\omega] = [\omega].$$

In other words, $\Phi^* \circ (\Phi^{-1})^*$ is also the identity map of $H^k_{\mathrm{dR}}(M)$. Similarly, one can also show $(\Phi^{-1})^* \circ \Phi^*$ is the identity map of $H^k_{\mathrm{dR}}(N)$. Therefore, $H^k_{\mathrm{dR}}(M)$ and $H^k_{\mathrm{dR}}(N)$ are isomorphic (as vector spaces). □

Corollary 5.8. *Given any smooth manifold M which is diffeomorphic to a star-shaped open set in $\mathbb{R}^n$, we have $H^1_{\mathrm{dR}}(M) \simeq \{[0]\}$, or in other words, every closed 1-form ω on such a manifold M is exact.*

Proof. Combine the results of the Poincaré Lemma (Theorem 5.4) and the diffeomorphism invariance of H^1_{dR} (Theorem 5.7). □

Consequently, a large class of open sets in $\mathbb{R}^n$ has trivial H^1_{dR} as long as it is diffeomorphic to a star-shaped manifold. For open sets in $\mathbb{R}^2$, there is a celebrated result called Riemann Mapping Theorem, which says any (non-empty) simply-connected open bounded subset U in $\mathbb{R}^2$ is diffeomorphic to the unit open ball in $\mathbb{R}^2$. In fact, the diffeomorphism can be chosen so that angles are preserved, but we don't need this when dealing with de Rham cohomology.

Under the assumption of Riemann Mapping Theorem (whose proof can be found in advanced Complex Analysis textbooks), we can establish that $H^1_{\mathrm{dR}}(U) = \{[0]\}$ for any (non-empty) simply-connected subset U in $\mathbb{R}^2$. Consequently, any closed 1-form on such a domain U is exact on U. Using the language in Multivariable Calculus (or Physics), this means any curl-zero vector field defined on a (non-empty) simply-connected domain U in $\mathbb{R}^2$ must be conservative on U. You might have learned this fact without proof in MATH 2023.

5.2. Deformation Retracts

In the previous section, we learned that two diffeomorphic manifolds have isomorphic de Rham cohomology groups. In short, we say de Rham cohomology is a diffeomorphic invariance. In this section, we will discuss another type of invariance: *deformation retracts*.

Let M be a smooth manifold (with or without boundary), and Σ is a submanifold of M. Note that Σ can have lower dimension than M. Roughly speaking, we say Σ is a *deformation retract* of M if one can continuously contract M onto Σ. Let's make it more precise:

Definition 5.9 (Deformation Retract)**.** Let M be a smooth manifold, and Σ is a submanifold of M. If there exists a C^1 family of smooth maps $\{\Psi_t : M \to M\}_{t\in[0,1]}$ satisfying all three conditions below:

- $\Psi_0(x) = x$ for any $x \in M$, i.e. $\Psi_0 = \mathrm{id}_M$;
- $\Psi_1(x) \in \Sigma$ for any $x \in M$, i.e. $\Psi_1 : M \to \Sigma$;
- $\Psi_t(p) = p$ for any $p \in \Sigma$, $t \in [0,1]$, i.e. $\Psi_t|_\Sigma = \mathrm{id}_\Sigma$ for any $t \in [0,1]$,

then we say Σ is a deformation retract of M. Equivalently, we can also say M deformation retracts onto Σ.

One good way to think of a deformation retract is to regard t as the time, and Ψ_t is a "movie" that demonstates how M collapses onto Σ. The condition $\Psi_0 = \mathrm{id}_M$ says initially (at $t = 0$), the "movie" starts with the image M. At the final scene (at $t = 1$), the condition $\Psi_1 : M \to \Sigma$ says that the image eventually becomes Σ . The last condition $\Psi_t(p) = p$ for any $p \in \Sigma$ means the points on Σ do not move throughout the movie. Before we talk about the relation between cohomology and deformation retract, let's first look at some examples:

Example 5.10. The unit circle $\mathbb{S}^1$ defined by $\{(x,y) : x^2 + y^2 = 1\}$ is a deformation retract of the annulus $\{(x,y) : \frac{1}{4} < x^2 + y^2 < 4\}$. To describe such a retract, it's best to use polar coordinates:

$$\Psi_t(re^{i\theta}) = (r + t(1-r))\, e^{i\theta}$$

For each $t \in [0,1]$, the map Ψ_t has image inside the annulus since $r + t(1-r) \in (\frac{1}{2}, 2)$ whenever $r \in (\frac{1}{2}, 2)$ and $t \in [0,1]$. One can easily check that $\Psi_0(re^{i\theta}) = re^{i\theta}$, $\Psi_1(re^{i\theta}) = e^{i\theta}$ and $\Psi_t(e^{i\theta}) = e^{i\theta}$ for any (r,θ) and $t \in [0,1]$. Hence Ψ_t fulfills all three conditions stated in Definition 5.9. □

Example 5.11. Intuitively, we can see the letters E, F, H, K, L, M and N all deformation retract onto the letter I. Also, the letter Q deformation retracts onto the letter O. The explicit Ψ_t for each deformation retract is not easy to write down. □

Example 5.12. A two-dimensional torus with a point removed can deformation retract onto two circles joined at one point. Try to visualize it! □

Exercise 5.4. Show that the unit circle $x^2 + y^2 = 1$ in $\mathbb{R}^2$ is a deformation retract of $\mathbb{R}^2 \backslash \{(0,0)\}$.

Exercise 5.5. Show that any star-shaped open set U in $\mathbb{R}^n$ deformation retracts onto its base point.

Exercise 5.6. Let M be a smooth manifold, and Σ_0 be the zero section of the tangent bundle, i.e. Σ_0 consists of all pairs $(p, 0_p)$ in TM where $p \in M$ and 0_p is the zero vector in T_pM. Show that the zero section Σ_0 is a deformation retract of the tangent bundle TM.

Exercise 5.7. Define a relation $\sim$ of manifolds by declaring that $M_1 \sim M_2$ if and only if M_1 is a deformation retract of M_2. Is $\sim$ an equivalence relation?

We next show an important result in de Rham theory, which asserts that deformation retracts preserve the first de Rham cohomology group.

Theorem 5.13 (Invariance under Deformation Retracts)**.** *Let M be a smooth manifold, and Σ be a submanifold of M. If Σ is a deformation retract of M, then $H^1_{\mathrm{dR}}(M)$ and $H^1_{\mathrm{dR}}(\Sigma)$ are isomorphic.*

Proof. Let $\iota : \Sigma \to M$ be the inclusion map, and $\{\Psi_t : M \to M\}_{t\in[0,1]}$ be the family of maps satisfying all conditions stated in Definition 5.9. Then, the pull-back map $\iota^* : \wedge^1 T^*M \to \wedge^1 T^*\Sigma$ induces a map $\iota^* : H^1_{\mathrm{dR}}(M) \to H^1_{\mathrm{dR}}(\Sigma)$. Also, the map $\Psi_1 : M \to \Sigma$ induces a pull-back map $\Psi_1^* : H^1_{\mathrm{dR}}(\Sigma) \to H^1_{\mathrm{dR}}(M)$. The key idea of the proof is to show that Ψ_1^* and ι^* are inverses of each other as maps between $H^1_{\mathrm{dR}}(M)$ and $H^1_{\mathrm{dR}}(\Sigma)$.

Let ω be an arbitrary closed 1-form defined on M. Similar to the proof of Poincaré Lemma (Theorem 5.4), we consider the scalar function $f : M \to \mathbb{R}$ defined by:

$$f(x) = \int_{\Psi_t(x)} \omega$$

Here, $\Psi_t(x)$ is regarded as a curve with parameter t joining $\Psi_0(x) = x$ and $\Psi_1(x) \in \Sigma$. We will show the following result:

$$\Psi_1^* \iota^* \omega - \omega = df \tag{5.1}$$

which will imply $[\omega] = \Psi_1^*\iota^*[\omega]$, or in other words, $\Psi_t^* \circ \iota^* = \mathrm{id}$ on $H^1_{\mathrm{dR}}(M)$.

To prove (5.1), we use local coordinates $(u_1, \ldots, u_n)$, and express ω in terms of local coordinates $\omega = \sum_i \omega_i du^i$. For simplicity, let's assume that such a local coordinate chart can cover the whole curve $\Psi_t(x)$ for $t \in [0,1]$. We will fix this issue later. For each $t \in [0,1]$, we write $\Psi_t^i(x)$ to be the u_i-coordinate of $\Psi_t(x)$, i.e. $\Psi_t^i = u_i \circ \Psi_t$. Then, one can calculate df using local coordinates. The calculation is similar to the one we did in the proof of Poincaré Lemma (Theorem 5.4):

$$\begin{aligned}
f(x) &= \int_{\Psi_t(x)} \omega = \int_0^1 \sum_i \omega_i(\Psi_t(x)) \frac{\partial \Psi_t^i}{\partial t}\, dt \\
(df)(x) &= \sum_j \frac{\partial f}{\partial u_j} du^j = \sum_j \left\{ \int_0^1 \frac{\partial}{\partial u_j} \left(\sum_i \omega_i(\Psi_t(x)) \frac{\partial \Psi_t^i}{\partial t} \right) dt \right\} du^j \\
&= \sum_j \left\{ \int_0^1 \left[\sum_{i,k} \left.\frac{\partial \omega_i}{\partial u_k}\right|_{\Psi_t(x)} \frac{\partial \Psi_t^k}{\partial u_j} \frac{\partial \Psi_t^i}{\partial t} + \sum_i \omega_i(\Psi_t(x)) \frac{\partial}{\partial t} \left(\frac{\partial \Psi_t^i}{\partial u_j} \right) \right] dt \right\} du^j
\end{aligned}$$

Next, recall that ω is a closed 1-form, so we have $\dfrac{\partial \omega_i}{\partial u_k} = \dfrac{\partial \omega_k}{\partial u_i}$ for any i, k. Using this on the first term, and by switching indices of the second term in the integrand, we get:

$$\begin{aligned}(df)(x) &= \sum_j \left\{ \int_0^1 \left[\sum_{i,k} \frac{\partial \omega_k}{\partial u_i}\bigg|_{\Psi_t(x)} \frac{\partial \Psi_t^k}{\partial u_j} \frac{\partial \Psi_t^i}{\partial t} + \sum_k \omega_k(\Psi_t(x)) \frac{\partial}{\partial t}\left(\frac{\partial \Psi_t^k}{\partial u_j} \right) \right] dt \right\} du^j \\ &= \sum_j \left\{ \int_0^1 \frac{\partial}{\partial t} \left(\sum_k \omega_k(\Psi_t(x)) \frac{\partial \Psi_t^k}{\partial u_j} \right) dt \right\} du^j = \sum_{j,k} \left[\omega_k(\Psi_t(x)) \frac{\partial \Psi_t^k}{\partial u_j} \right]_{t=0}^{t=1} du^j\end{aligned}$$

where the last equality follows from the (backward) chain rule.

Denote $\iota_t : \Psi_t(M) \to M$ the inclusion map at time t, then one can check that

$$\begin{aligned}\Psi_t^* \iota_t^* \omega(x) = (\iota_t \circ \Psi_t)^* \omega(x) &= (\iota_t \circ \Psi_t)^* \sum_k \omega_k du^k \\ &= \sum_k \omega_k(\iota_t \circ \Psi_t(x))\, d(u_k \circ \iota_t \circ \Psi_t(x)) \\ &= \sum_k \omega_k(\iota_t \circ \Psi_t(x))\, d\Psi_t^k \\ &= \sum_{j,k} \omega_k(\Psi_t(x))\, \frac{\partial \Psi_t^k}{\partial u_j}\, du^j.\end{aligned}$$

Therefore, we get:

$$df = \sum_{j,k} \left[\omega_k(\Psi_t(x)) \frac{\partial \Psi_t^k}{\partial u_j} \right]_{t=0}^{t=1} du^j = [\Psi_t^* \iota_t^* \omega]_{t=0}^{t=1} = \Psi_1^* \iota_1^* \omega - \Psi_0^* \iota_0^* \omega.$$

Since $\Psi_0 = \mathrm{id}_M$ and $\iota_0 = \mathrm{id}_M$, we have proved (5.1). In case $\Psi_t(x)$ cannot be covered by one single local coordinate chart, one can then modify the above proof a bit by covering the curve $\Psi_t(x)$ by finitely many local coordinate charts. It can be done because $\Psi_t(x)$ is compact. Suppose $0 = t_0 < t_1 < \ldots < t_N = 1$ is a partition of $[0, 1]$ such that for each α, the curve $\Psi_t(x)$ restricted to $t \in [t_{\alpha-1}, t_\alpha]$ can be covered by a single local coordinate chart, then we have:

$$f(x) = \sum_{\alpha=1}^N \int_{t_{\alpha-1}}^{t_\alpha} \sum_i \omega_i(\Psi_t(x)) \frac{\partial \Psi_t^i}{\partial t}\, dt.$$

Proceed as in the above proof, we can get:

$$df = \sum_{\alpha=1}^N \left(\Psi_{t_\alpha}^* \iota_{t_\alpha}^* \omega - \Psi_{t_\alpha - 1}^* \iota_{t_\alpha - 1}^* \omega \right) = \Psi_1^* \iota_1^* \omega - \Psi_0^* \iota_0^* \omega,$$

which completes the proof of (5.1) in the general case.

To complete the proof of the theorem, we consider an arbitrary 1-form η on Σ. We claim that

$$\iota^* \Psi_1^* \eta = \eta. \tag{5.2}$$

We prove by direct verification using local coordinates $(u_1, \ldots, u_n)$ on M such that:

$$(u_1, \ldots, u_k, 0, \ldots, 0) \in \Sigma.$$

Such a local coordinate system always exists near Σ by Immersion Theorem (Theorem 2.42). Locally, denote $\eta = \sum_{i=1}^{k} \eta_i \, du^i$, then

$$\begin{aligned}(\Psi_1^*\eta)(x) &= \sum_{i=1}^{k} \Psi_1^*(\eta_i(x)\, du^i) = \sum_{i=1}^{k} \eta_i(\Psi_1(x))\, d(u^i \circ \Psi_1)\\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} \eta_i(\Psi_1(x)) \frac{\partial \Psi_1^i(x)}{\partial u_j}\, du^j.\end{aligned}$$

Since $\Psi_1(x) = x$ whenever $x \in \Sigma$, we have $\Psi_1^i(x) = u_i(x)$ where $u_i(x)$ is the i-th coordinate of x. Therefore, we get $\dfrac{\partial \Psi_1^i(x)}{\partial u_j} = \dfrac{\partial u_i}{\partial u_j} = \delta_{ij}$ and so:

$$(\Psi_1^*\eta)(x) = \sum_{i,j=1}^{k} \eta_i(x)\delta_{ij}\, du^j = \sum_{i=1}^{k} \eta_i(x)\, du^i = \eta(x)$$

for any $x \in \Sigma$. In other words, $\iota^*\Psi_1^*\eta = \eta$ on Σ. This proves (5.2).

Combining (5.1) and (5.2), we get $\iota^* \circ \Psi_1^* = \mathrm{id}$ on $H^1_{\mathrm{dR}}(\Sigma)$, and $\Psi_1^* \circ \iota^* = \mathrm{id}$ on $H^1_{\mathrm{dR}}(M)$. As a result, Ψ_1^* and ι^* are inverses of each other in H^1_{dR}. It completes the proof that $H^1_{\mathrm{dR}}(M)$ and $H^1_{\mathrm{dR}}(\Sigma)$ are isomorphic. □

Using Theorem 5.13, we see that $H^1_{\mathrm{dR}}(\mathbb{R}^2 \backslash \{(0,0)\})$ and $H^1_{\mathrm{dR}}(\mathbb{S}^1)$ are isomorphic, and hence $b_1(\mathbb{R}^2 \backslash \{(0,0)\}) = b_1(\mathbb{S}^1)$. At this moment, we still don't know the exact value of $b_1(\mathbb{S}^1)$, but we will figure it out in the next section.

Note that Theorem 5.13 holds for H^k_{dR} for any $k \geq 2$ as well, but the proof again uses some Lie derivatives and Cartan's formula, which are beyond the scope of this course.

Another nice consequence of Theorem 5.13 is the 2-dimensional case of the following celebrated theorem in topology:

Theorem 5.14 (Brouwer's Fixed-Point Theorem on $\mathbb{R}^2$)**.** *Let $B_1(0)$ be the closed ball with radius 1 centered at origin in $\mathbb{R}^2$. Suppose $\Phi : B_1(0) \to B_1(0)$ is a smooth map between $B_1(0)$. Then, there exists a point $\mathsf{x} \in B_1(0)$ such that $\Phi(\mathsf{x}) = \mathsf{x}$.*

Proof. We prove by contradiction. Suppose $\Phi(\mathsf{x}) \neq \mathsf{x}$ for any $\mathsf{x} \in B_1(0)$. Then, we let $\Psi_t(\mathsf{x})$ be a point in $B_1(0)$ defined in the following way:

(1) Consider the vector $\mathsf{x} - \Phi(\mathsf{x})$ which is non-zero.
(2) Consider the straight ray emanating from x in the direction of $\mathsf{x} - \Phi(\mathsf{x})$. This ray will intersect the unit circle $\mathbb{S}^1$ at a unique point p_x.
(3) We then define $\Psi_t(\mathsf{x}) := (1-t)\mathsf{x} + t\mathsf{p}_\mathsf{x}$

We leave it as an exercise for readers to write down the explicit formula for $\Psi_t(\mathsf{x})$, and show that it is smooth for each $t \in [0,1]$.

Clearly, we have $\Psi_0(\mathsf{x}) = \mathsf{x}$ for any $\mathsf{x} \in B_1(0)$; $\Psi_1(\mathsf{x}) = \mathsf{p}_\mathsf{x} \in \mathbb{S}^1$; and if $|\mathsf{x}| = 1$, then $\mathsf{p}_\mathsf{x} = \mathsf{x}$ and so $\Psi_t(\mathsf{x}) = \mathsf{x}$.

Therefore, it shows $\mathbb{S}^1$ is a deformation retract of $B_1(0)$, and by Theorem 5.13, their H^1_{dR}'s are isomorphic. However, we know $H^1_{\mathrm{dR}}(B_1(0)) \simeq \{[0]\}$, while $H^1_{\mathrm{dR}}(\mathbb{S}^1) \simeq H^1_{\mathrm{dR}}(\mathbb{R}^2 \backslash \{(0,0)\}) \neq \{[0]\}$. It is a contradiction! It completes the proof that there is at least a point $\mathsf{x} \in B_1(0)$ such that $\Phi(\mathsf{x}) = \mathsf{x}$. □

Exercise 5.8. Write down an explicit expression of p_x in the above proof, and hence show that Ψ_t is smooth for each fixed t.

Exercise 5.9. Generalize the Brouwer's Fixed-Point Theorem in the following way: given a manifold Ω which is diffeomorphic to $B_1(0)$, and a smooth map $\Phi : \Omega \to \Omega$. Using Theorem 5.14, show that there exists a point $p \in \Omega$ such that $\Phi(p) = p$.

Exercise 5.10. What fact(s) are needed to be established in order to prove the Brouwer's Fixed-Point Theorem for general $\mathbb{R}^n$ using a similar way as in the proof of Theorem 5.14?

5.3. Mayer-Vietoris Theorem

In the previous section, we showed that if Σ is a deformation retract of M, then $H^1_{\mathrm{dR}}(\Sigma)$ and $H^1_{\mathrm{dR}}(M)$ are isomorphic. For instance, this shows $H^1_{\mathrm{dR}}(\mathbb{S}^1)$ is isomorphic to $H^1_{\mathrm{dR}}(\mathbb{R}^2\backslash\{(0,0)\})$. Although we have discussed that $H^2_{\mathrm{dR}}(\mathbb{R}^2\backslash\{(0,0)\})$ is non-trivial, we still haven't figured out what this group is. In this section, we introduce a useful tool, called Mayer-Vietoris sequence, that we can use to compute the de Rham cohomology groups of $\mathbb{R}^2\backslash\{(0,0)\}$, as well as many other spaces.

5.3.1. Exact Sequences. Consider a sequence of homomorphism between abelian groups:

$$\cdots \xrightarrow{T_{k-1}} G_{k-1} \xrightarrow{T_k} G_k \xrightarrow{T_{k+1}} G_{k+1} \xrightarrow{G_{k+1}} \cdots$$

We say it is an *exact sequence* if the image of each homomorphism is equal to the kernel of the next one, i.e.

$$\operatorname{Im} T_{i-1} = \ker T_i \quad \text{for each } i.$$

One can also talk about exact-ness for a finite sequence, say:

$$G_0 \xrightarrow{T_1} G_1 \xrightarrow{T_2} G_2 \xrightarrow{T_3} \cdots \xrightarrow{T_{n-1}} G_{n-1} \xrightarrow{T_n} G_n$$

However, such a T_1 would not have a previous map, and such an T_n would not have the next map. Therefore, whenever we talk about the exact-ness of a finite sequence of maps, we will add two trivial maps at both ends, i.e.

$$0 \xrightarrow{0} G_0 \xrightarrow{T_1} G_1 \xrightarrow{T_2} G_2 \xrightarrow{T_3} \cdots G_{n-1} \xrightarrow{T_n} G_n \xrightarrow{0} 0. \tag{5.3}$$

The first map $0 \xrightarrow{0} G_0$ is the homomorphism taking the zero in the trivial group to the zero in G_0. The last map $G_n \xrightarrow{0} 0$ is the linear map that takes every element in G_n to the zero in the trivial group. We say the finite sequence (5.3) an *exact sequence* if

$$\operatorname{Im}(0 \xrightarrow{0} G_0) = \ker T_1, \quad \operatorname{Im} T_n = \ker(G_n \xrightarrow{0} 0), \text{ and } \operatorname{Im} T_i = \ker T_{i+1} \quad \text{for any } i.$$

Note that $\operatorname{Im}(0 \xrightarrow{0} G_0) = \{0\}$ and $\ker(G_n \xrightarrow{0} 0) = G_n$, so if (5.3) is an exact sequence, it is necessary that

$$\ker T_1 = \{0\} \quad \text{and} \quad \operatorname{Im} T_n = G_n$$

or equivalently, T_1 is injective and T_n is surjective.

One classic example of a finite exact sequence is:

$$0 \to \mathbb{Z} \xrightarrow{\iota} \mathbb{C} \xrightarrow{f} \mathbb{C}\backslash\{0\} \to 0$$

where $\iota : \mathbb{Z} \to \mathbb{C}$ is the inclusion map taking $n \in \mathbb{Z}$ to itself $n \in \mathbb{C}$. The map $f : \mathbb{C} \to \mathbb{C}\backslash\{0\}$ is the map taking $z \in \mathbb{C}$ to $e^{2\pi i z} \in \mathbb{C}\backslash\{0\}$.

It is clear that ι is injective and f is surjective (from Complex Analysis). Also, we have $\operatorname{Im}\iota = \mathbb{Z}$ and $\ker f = \mathbb{Z}$ as well (note that the identity of $\mathbb{C}\backslash\{0\}$ is 1, not 0). Therefore, this is an exact sequence.

Exercise 5.11. Given an exact sequence of group homomorphisms:

$$0 \to A \xrightarrow{T} B \xrightarrow{S} C \to 0,$$

(a) If it is given that $C = \{0\}$, what can you say about A and B?

(b) If it is given that $A = \{0\}$, what can you say about B and C?

5.3.2. Mayer-Vietoris Sequences. We talk about exact sequences because there is such a sequence concerning de Rham cohomology groups. This exact sequence, called the Mayer-Vietoris sequence, is particularly useful for computing H^k_{dR} for many manifolds.

The basic setup of a Mayer-Vietoris sequence is a smooth manifold (with or without boundary) which can be expressed a union of two open sets U and V, i.e. $M = U \cup V$. Note that we do not require U and V are disjoint. The intersection $U \cap V$ is a subset of both U and V; and each of U and V is in turn a subset of M. To summarize, we have the following relations of sets:

$$\begin{tikzcd} & U \arrow[dr, "j_U"] & \\ U \cap V \arrow[ur, "i_U"] \arrow[dr, "i_V"'] & & M \\ & V \arrow[ur, "j_V"'] & \end{tikzcd}$$

where i_U, i_V, j_U and j_V are inclusion maps. Each inclusion map, say $j_U : U \to M$, induces a pull-back map $j_U^* : \wedge^k T^* M \to \wedge^k T^* U$ which takes any k-form ω on M, to the k-form $\omega|_U$ restricted on U, i.e. $j_U^*(\omega) = \omega|_U$ for any $\omega \in \wedge^k T^* M$. In terms of local expressions, there is essentially no difference between ω and $\omega|_U$ since U is open. If locally $\omega = \sum_i \omega_i \, du^i$ on M, then $\omega|_U = \sum_i \omega_i \, du^i$ as well. The only difference is the domain: $\omega(p)$ is defined for every $p \in M$, while $\omega|_U(p)$ is defined only when $p \in U$.

To summarize, we have the following diagram:

$$\begin{tikzcd} & U \arrow[dl, "i_U^*"'] & \\ U \cap V & & M \arrow[ul, "j_U^*"'] \arrow[dl, "j_V^*"] \\ & V \arrow[ul, "i_V^*"] & \end{tikzcd}$$

Using the pull-backs of these four inclusions i_U, i_V, j_U and j_V, one can form a sequence of linear maps for each integer k:

$$0 \to \wedge^k T^* M \xrightarrow{j_U^* \oplus j_V^*} \wedge^k T^* U \oplus \wedge^k T^* V \xrightarrow{i_U^* - i_V^*} \wedge^k T^*(U \cap V) \to 0 \tag{5.4}$$

Here, $\wedge^k T^* U \oplus \wedge^k T^* V$ is the direct sum of the vector spaces $\wedge^k T^* U$ and $\wedge^k T^* V$, meaning that:

$$\wedge^k T^* U \oplus \wedge^k T^* V = \{(\omega, \eta) : \omega \in \wedge^k T^* U \text{ and } \eta \in \wedge^k T^* V\}.$$

The map $j_U^* \oplus j_V^* : \wedge^k T^* M \to \wedge^k T^* U \oplus \wedge^k T^* V$ is defined by:

$$(j_U^* \oplus j_V^*)(\omega) = (j_U^* \omega, j_V^* \omega) = (\omega|_U, \omega|_V).$$

The map $\wedge^k T^* U \oplus \wedge^k T^* V \xrightarrow{i_U^* - i_V^*} \wedge^k T^*(U \cap V)$ is given by:

$$(i_U^* - i_V^*)(\omega, \eta) = i_U^* \omega - i_V^* \eta = \omega|_{U \cap V} - \eta|_{U \cap V}.$$

We next show that the sequence (5.4) is exact. Let's first try to understand the image and kernel of each map involved.

Given $(\omega, \eta) \in \ker(i_U^* - i_V^*)$, we will have $\omega|_{U \cap V} = \eta|_{U \cap V}$. Therefore, $\ker(i_U^* - i_V^*)$ consists of pairs (ω, η) where ω and η agree on the intersection $U \cap V$.

Now consider $\operatorname{Im}(j_U^* \oplus j_V^*)$, which consists of pairs of the form $(\omega|_U, \omega|_V)$. Certainly, the restrictions of both $\omega|_U$ and $\omega|_V$ on the intersection $U \cap V$ are the same, and hence the pair is inside $\ker(i_U^* - i_V^*)$. Therefore, we have $\operatorname{Im}(j_U^* \oplus j_V^*) \subset \ker(i_U^* - i_V^*)$.

In order to show (5.4) is exact, we need further that:

(1) $j_U^* \oplus j_V^*$ is injective;

(2) $i_U^* - i_V^*$ is surjective; and

(3) $\ker(i_U^* - i_V^*) \subset \operatorname{Im}(j_U^* \oplus j_V^*)$

We leave (1) as an exercises, and will give the proofs of (2) and (3).

Exercise 5.12. Show that $j_U^* \oplus j_V^*$ is injective in the sequence (5.4).

Proposition 5.15. *Let M be a smooth manifold. Suppose there are two open subsets U and V of M such that $M = U \cup V$, and $U \cap V$ is non-empty, then the sequence of maps* (5.4) *is exact.*

Proof. So far we have proved that $j_U^* \oplus j_V^*$ is injective, and $\operatorname{Im}(j_U^* \oplus j_V^*) \subset \ker(i_U^* - i_V^*)$. We next claim that $\ker(i_U^* - i_V^*) \subset \operatorname{Im}(j_U^* \oplus j_V^*)$:

Let $(\omega, \eta) \in \ker(i_U^* - i_V^*)$, meaning that ω is a k-form on U, η is a k-form on V, and that $\omega|_{U\cap V} = \eta|_{U\cap V}$. Define a k-form σ on $M = U \cup V$ by:

$$\sigma = \begin{cases} \omega & \text{on } U \\ \eta & \text{on } V \end{cases}$$

Note that σ is well-defined on $U \cap V$ since ω and η agree on $U \cap V$. Then, we have:

$$(\omega, \eta) = (\sigma|_U, \sigma|_V) = (j_U^*\sigma, j_V^*\sigma) = (j_U^* \oplus j_V^*)\sigma \in \operatorname{Im}(j_U^* \oplus j_V^*).$$

Since (ω, η) is arbitrary in $\ker(i_U^* - i_V^*)$, this shows:

$$\ker(i_U^* - i_V^*) \subset \operatorname{Im}(j_U^* \oplus j_V^*).$$

Finally, we show $i_U^* - i_V^*$ is surjective. Given any k-form $\theta \in \wedge^k T^*(U \cap V)$, we need to find a k-form ω' on U, and a k-form η' on V such that $\omega' - \eta' = \theta$ on $U \cap V$. Let $\{\rho_U, \rho_V\}$ be a partition of unity subordinate to $\{U, V\}$. We define:

$$\omega' = \begin{cases} \rho_V\theta & \text{on } U \cap V \\ 0 & \text{on } U \backslash V \end{cases}$$

Note that ω' is smooth: If $p \in \operatorname{supp} \rho_V \subset V$, then $p \in V$ (which is open) and so $\omega' = \rho_V\theta$ in an open neighborhood of p. Note that ρ_V and θ are smooth at p, so ω' is also smooth at p. On the other hand, if $p \notin \operatorname{supp} \rho_V$, then $\omega' = 0$ in an open neighborhood of p. In particular, ω' is smooth at p.

Similarly, we define:

$$\eta' = \begin{cases} -\rho_U\theta & \text{on } U \cap V \\ 0 & \text{on } V \backslash U \end{cases}$$

which can be shown to be smooth in a similar way.

Then, when restricted to $U \cap V$, we get:

$$\omega'|_{U\cap V} - \eta'|_{U\cap V} = \rho_V\theta + \rho_U\theta = (\rho_V + \rho_U)\,\theta = \theta.$$

In other words, we have $(i_U^* - i_V^*)(\omega', \eta') = \theta$. Since θ is arbitrary, we proved $i_U^* - i_V^*$ is surjective. $\square$

Recall that a pull-back map on k-forms induces a well-defined pull-back map on H^k_{dR}. The sequence of maps (5.4) between space of wedge products induces a sequence of maps between de Rham cohomology groups:

$$0 \to H^k_{\mathrm{dR}}(M) \xrightarrow{j_U^* \oplus j_V^*} H^k_{\mathrm{dR}}(U) \oplus H^k_{\mathrm{dR}}(V) \xrightarrow{i_U^* - i_V^*} H^k_{\mathrm{dR}}(U \cap V) \to 0. \tag{5.5}$$

Here, $j_U^* \oplus j_V^*$ and $i_U^* - i_V^*$ are defined by:

$$\begin{aligned}(j_U^* \oplus j_V^*)[\omega] &= (j_U^*[\omega], j_V^*[\omega]) = ([j_U^*\omega], [j_V^*\omega]) \\ (i_U^* - i_V^*)\,([\omega],[\eta]) &= i_U^*[\omega] - i_V^*[\eta] = [i_U^*\omega] - [i_V^*\eta].\end{aligned}$$

However, the sequence (5.5) is *not* exact because $j_U^* \oplus j_V^*$ may not be injective, and $i_U^* - i_V^*$ may not be surjective. For example, take $M = \mathbb{R}^2 \backslash \{(0,0)\}$, and define using polar coordinates the open sets $U = \{re^{i\theta} : r > 0, \theta \in (0, 2\pi)\}$ and $V = \{re^{i\theta} : r > 0, \theta \in (-\pi, \pi)\}$. Then, both U and V are star-shaped and hence both $H^1_{\mathrm{dR}}(U)$ and $H^1_{\mathrm{dR}}(V)$ are trivial. Nonetheless we have exhibited that $H^1_{\mathrm{dR}}(M)$ is non-trivial. The map $j_U^* \oplus j_V^*$ from a non-trivial group to the trivial group can never be injective!

Exercise 5.13. Find an example of M, U and V such that the map $i_U^* - i_V^*$ in (5.5) is not surjective.

Nonetheless, it is still true that $\ker(i_U^* - i_V^*) = \operatorname{Im}(j_U^* \oplus j_V^*)$, and we will verify it in the proof of Mayer-Vietoris Theorem (Theorem 5.16). Mayer-Vietoris Theorem asserts that although (5.5) is not exact in general, but we can connect each short sequence below:

$$\begin{aligned}
&H^0_{\mathrm{dR}}(M) \xrightarrow{j_U^* \oplus j_V^*} H^0_{\mathrm{dR}}(U) \oplus H^0_{\mathrm{dR}}(V) \xrightarrow{i_U^* - i_V^*} H^0_{\mathrm{dR}}(U \cap V) \\
&H^1_{\mathrm{dR}}(M) \xrightarrow{j_U^* \oplus j_V^*} H^1_{\mathrm{dR}}(U) \oplus H^1_{\mathrm{dR}}(V) \xrightarrow{i_U^* - i_V^*} H^1_{\mathrm{dR}}(U \cap V) \\
&H^2_{\mathrm{dR}}(M) \xrightarrow{j_U^* \oplus j_V^*} H^2_{\mathrm{dR}}(U) \oplus H^2_{\mathrm{dR}}(V) \xrightarrow{i_U^* - i_V^*} H^2_{\mathrm{dR}}(U \cap V) \\
&\qquad\qquad\qquad\qquad\vdots
\end{aligned}$$

to produce a long exact sequence.

Theorem 5.16 (Mayer-Vietoris Theorem). *Let M be a smooth manifold, and U and V be open sets of M such that $M = U \cup V$. Then, for each $k \geq 0$ there is a homomorphism $\delta : H^k_{\mathrm{dR}}(U \cap V) \to H^{k+1}_{\mathrm{dR}}(M)$ such that the following sequence is exact:*

$$\cdots \xrightarrow{\delta} H^k_{\mathrm{dR}}(M) \xrightarrow{j_U^* \oplus j_V^*} H^k_{\mathrm{dR}}(U) \oplus H^k_{\mathrm{dR}}(V) \xrightarrow{i_U^* - i_V^*} H^k_{\mathrm{dR}}(U \cap V) \xrightarrow{\delta} H^{k+1}_{\mathrm{dR}}(M) \to \cdots$$

This long exact sequence is called the Mayer-Vietoris sequence.

The proof of Theorem 5.16 is purely algebraic. We will learn the proof after looking at some examples.

5.3.3. Using Mayer-Vietoris Sequences.

The Mayer-Vietoris sequence is particularly useful for computing de Rham cohomology groups and Betti numbers using linear algebraic methods. Suppose M can be expressed as a union $U \cup V$ of two open sets, such that the H^k_{dR}'s of U, V and $U \cap V$ can be computed easily, then $H^k_{\mathrm{dR}}(M)$ can be deduced by "playing around" the kernels and images in the Mayer-Vietoris sequence. One useful result in Linear (or Abstract) Algebra is the following:

Theorem 5.17 (First Isomorphism Theorem). *Let $T : V \to W$ be a linear map between two vector spaces V and W. Then, we have:*

$$\operatorname{Im} T \cong V/\ker T.$$

In particular, if V and W are finite dimensional, we have:

$$\dim \ker T + \dim \operatorname{Im} T = \dim V.$$

Proof. Let $\Phi : \operatorname{Im} T \to V/\ker T$ be the map defined by:

$$\Phi(T(v)) = [v]$$

for any $T(v) \in \operatorname{Im} T$. This map is well-defined since if $T(v) = T(w)$ in $\operatorname{Im} T$, then $v - w \in \ker T$, which implies $[v] = [w]$ in the quotient vector space $V/\ker T$. It is easy (hence omitted) to verify that Φ is linear.

Φ is injective since whenever $T(v) \in \ker \Phi$, we have $\Phi(T(v)) = [0]$ which implies $[v] = [0]$ and hence $v \in \ker T$ (i.e. $T(v) = 0$). Also, Φ is surjective since given any $[v] \in V/\ker T$, we have $\Phi(T(v)) = [v]$ by the definition of Φ.

These show Φ is an isomorphism, hence completing the proof. $\square$

Example 5.18. In this example, we use the Mayer-Vietoris sequence to compute $H^1_{\mathrm{dR}}(\mathbb{S}^1)$. Let:

$$M = \mathbb{S}^1, \quad U = M\backslash\{\text{north pole}\}, \quad V = M\backslash\{\text{south pole}\}.$$

Then clearly $M = U \cup V$, and $U \cap V$ consists of two disjoint arcs (each of which deformation retracts to a point). Here are facts which we know and which we haven't yet known:

$$\begin{array}{llll} H^0_{\mathrm{dR}}(M) \cong \mathbb{R} & H^0_{\mathrm{dR}}(U) \cong \mathbb{R} & H^0_{\mathrm{dR}}(V) \cong \mathbb{R} & H^0_{\mathrm{dR}}(U \cap V) \cong \mathbb{R} \oplus \mathbb{R} \\ H^1_{\mathrm{dR}}(M) \text{ unknown} & H^1_{\mathrm{dR}}(U) \cong 0 & H^1_{\mathrm{dR}}(V) \cong 0 & H^1_{\mathrm{dR}}(U \cap V) \cong 0 \end{array}$$

By Theorem 5.16, we know that the following sequence is exact:

$$\cdots \to \underbrace{H^0_{\mathrm{dR}}(U) \oplus H^0_{\mathrm{dR}}(V)}_{\mathbb{R}\oplus\mathbb{R}} \xrightarrow{i_U^* - i_V^*} \underbrace{H^0_{\mathrm{dR}}(U \cap V)}_{\mathbb{R}\oplus\mathbb{R}} \xrightarrow{\delta} \underbrace{H^1_{\mathrm{dR}}(M)}_{?} \xrightarrow{j_U^* \oplus j_V^*} \underbrace{H^1_{\mathrm{dR}}(U) \oplus H^1_{\mathrm{dR}}(V)}_{0}$$

Therefore, δ is surjective.

By First Isomorphism Theorem (Theorem 5.17), we know:

$$H^1_{\mathrm{dR}}(M) = \operatorname{Im} \delta \cong \frac{H^0_{\mathrm{dR}}(U \cap V)}{\ker \delta}.$$

Elements of $H^0_{\mathrm{dR}}(U \cap V)$ are locally constant functions of the form:

$$f_{a,b} = \begin{cases} a & \text{on left arc} \\ b & \text{on right arc} \end{cases}$$

Since the Mayer-Vietoris sequence is exact, we have $\ker \delta = \operatorname{Im}(i_U^* - i_V^*)$. The space $H^0_{\mathrm{dR}}(U)$, $H^0_{\mathrm{dR}}(V)$ and $H^0_{\mathrm{dR}}(U \cap V)$ consist of locally constant functions on U, V and $U \cap V$ respectively, and the maps $i_U^* - i_V^*$ takes constant functions $(k_1, k_2) \in H^0_{\mathrm{dR}}(U) \oplus H^0_{\mathrm{dR}}(V)$ to the constant function $f_{k_1 - k_2, k_1 - k_2}$ on $U \cap V$. Therefore, the first de Rham cohomology group of M is given by:

$$H^1_{\mathrm{dR}}(M) \cong \frac{\{f_{a,b} : a, b \in \mathbb{R}\}}{\{f_{a-b,a-b} : a, b \in \mathbb{R}\}} \cong \frac{\mathbb{R}^2}{\{(x,y) : x = y\}},$$

and hence $b_1(M) = \dim H^1_{\mathrm{dR}}(M) = 2 - 1 = 1$. $\square$

Example 5.19. Let's discuss some consequences of the result proved in the previous example. Recall that $\mathbb{R}^2\backslash\{(0,0)\}$ deformation retracts to $\mathbb{S}^1$. By Theorem 5.13, we know $H^1_{\mathrm{dR}}(\mathbb{R}^2\backslash\{(0,0)\}) \cong H^1_{\mathrm{dR}}(\mathbb{S}^1)$.

This tells us $b_1(\mathbb{R}^2\backslash\{(0,0)\}) = 1$ as well. Recall that the following 1-form:

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

is closed but not exact. The class $[\omega]$ is then trivial in $H^1_{\mathrm{dR}}(\mathbb{R}^2\backslash\{(0,0)\})$. In an one-dimensional vector space, any non-zero vector spans that space. Therefore, we conclude:

$$H^1_{\mathrm{dR}}(\mathbb{R}^2\backslash\{(0,0)\} = \{c[\omega] : c \in \mathbb{R}\}.$$

where ω is defined as in above.

As a result, if ω' is a closed 1-form on $\mathbb{R}^2\backslash\{(0,0)\}$, then we must have

$$[\omega'] = c[\omega]$$

for some $c \in \mathbb{R}$, and so $\omega' = c\omega + df$ for some smooth function $f : \mathbb{R}^2\backslash\{(0,0)\} \to \mathbb{R}$.

Using the language of vector fields, if $\mathsf{V}(x,y) : \mathbb{R}^2\backslash\{(0,0)\} \to \mathbb{R}^2$ is a smooth vector field with $\nabla \times \mathsf{V} = 0$, then there is a constant $c \in \mathbb{R}$ and a smooth function $f : \mathbb{R}^2\backslash\{(0,0)\} \to \mathbb{R}$ such that:

$$\mathsf{V} = c\left(\frac{-y\mathsf{i} + x\mathsf{j}}{x^2+y^2}\right) + \nabla f.$$

□

Exercise 5.14. Let $\mathbb{T}^2$ be the two-dimensional torus. Show that $b_1(\mathbb{T}^2) = 2$.

Exercise 5.15. Show that $b_1(\mathbb{S}^2) = 0$. Based on this result, show that any curl-zero vector field defined on $\mathbb{R}^3\backslash\{(0,0,0)\}$ must be conservative.

One good technique of using the Mayer-Vietoris sequence (as demonstrated in the examples and exercises above) is to consider a segment of the sequence that starts and ends with the trivial space, i.e.

$$0 \to V_1 \to V_2 \to \cdots \to V_n \to 0.$$

If all vector spaces V_i's except one of them are known, then the remaining one (at least its dimension) can be deduced using First Isomorphism Theorem. Below is a useful lemma which is particularly useful for finding the Betti number of a manifold:

Lemma 5.20. *Let the following be an exact sequence of finite dimensional vector spaces:*

$$0 \to V_1 \xrightarrow{T_1} V_2 \xrightarrow{T_2} \cdots \xrightarrow{T_{n-1}} V_n \to 0.$$

Then, we have:

$$\dim V_1 - \dim V_2 + \dim V_3 - \cdots + (-1)^{n-1} \dim V_n = 0$$

Proof. By exact-ness, the map $T_{n-1} : V_{n-1} \to V_n$ is surjective. By First Isomorphism Theorem (Theorem 5.17), we get:

$$V_n = \operatorname{Im} T_{n-1} \cong V_{n-1}/\ker T_{n-1} = V_{n-1}/\operatorname{Im} T_{n-2}.$$

As a result, we have:

$$\dim V_n = \dim V_{n-1} - \dim \operatorname{Im} T_{n-2}.$$

Similarly, apply First Isomorphism Theorem on $T_{n-2} : V_{n-2} \to V_{n-1}$, we get:

$$\dim \operatorname{Im} T_{n-2} = \dim V_{n-2} - \dim \operatorname{Im} T_{n-3},$$

and combine with the previous result, we get:

$$\dim V_n = \dim V_{n-1} - \dim V_{n-2} + \dim \operatorname{Im} T_{n-3}.$$

Proceed similarly as the above, we finally get:

$$\dim V_n = \dim V_{n-1} - \dim V_{n-2} + \ldots + (-1)^n \dim V_1,$$

as desired. $\square$

In Example 5.18 (about computing $H^1_{\mathrm{dR}}(\mathbb{S}^1)$), the following exact sequence was used:

$$0 \to \underbrace{H^0_{\mathrm{dR}}(\mathbb{S}^1)}_{\mathbb{R}} \to \underbrace{H^0_{\mathrm{dR}}(U) \oplus H^0_{\mathrm{dR}}(V)}_{\mathbb{R}\oplus\mathbb{R}} \to \underbrace{H^0_{\mathrm{dR}}(U \cap V)}_{\mathbb{R}\oplus\mathbb{R}} \to \underbrace{H^1_{\mathrm{dR}}(\mathbb{S}^1)}_{?} \to \underbrace{H^1_{\mathrm{dR}}(U) \oplus H^1_{\mathrm{dR}}(V)}_{0}$$

Using Lemma 5.20, the dimension of $H^1_{\mathrm{dR}}(\mathbb{S}^1)$ can be computed easily:

$$\dim \mathbb{R} - \dim \mathbb{R} \oplus \mathbb{R} + \dim \mathbb{R} \oplus \mathbb{R} - \dim H^1_{\mathrm{dR}}(\mathbb{S}^1) = 0$$

which implies $\dim H^1_{\mathrm{dR}}(\mathbb{S}^1) = 1$ (or equivalently, $b_1(\mathbb{S}^1) = 1$). Although this method does not give a precise description of $H^1_{\mathrm{dR}}(\mathbb{S}^1)$ in terms of inclusion maps, it is no doubt much easier to adopt.

In the forthcoming examples, we will assume the following facts stated below (which we have only proved the case $k = 1$):

- $H^k_{\mathrm{dR}}(U) = 0$, where $k \geq 1$, for any star-shaped region $U \subset \mathbb{R}^n$.
- If Σ is a deformation retract of M, then $H^k_{\mathrm{dR}}(\Sigma) \cong H^k_{\mathrm{dR}}(M)$ for any $k \geq 1$.

Example 5.21. Consider $\mathbb{R}^2 \backslash \{p_1, \ldots, p_n\}$ where $p_1, \ldots, p_n$ are n distinct points in $\mathbb{R}^2$. We want to find b_1 of this open set.

Define $U = \mathbb{R}^2 \backslash \{p_1, \ldots, p_{n-1}\}$, $V = \mathbb{R}^2 \backslash \{p_n\}$, then $U \cup V = \mathbb{R}^2$ and $U \cap V = \mathbb{R}^2 \backslash \{p_1, \ldots, p_n\}$. Consider the Mayer-Vietoris sequence:

$$\underbrace{H^1_{\mathrm{dR}}(U \cup V)}_{0} \to H^1_{\mathrm{dR}}(U) \oplus H^1_{\mathrm{dR}}(V) \to H^1_{\mathrm{dR}}(U \cap V) \to \underbrace{H^2_{\mathrm{dR}}(U \cup V)}_{0}.$$

Using Lemma 5.20, we know:

$$\dim H^1_{\mathrm{dR}}(U) \oplus H^1_{\mathrm{dR}}(V) - \dim H^1_{\mathrm{dR}}(U \cap V) = 0$$

We have already figured out that $\dim H^1_{\mathrm{dR}}(V) = 1$. Therefore, we get:

$$\dim H^1_{\mathrm{dR}}(\mathbb{R}^2 \backslash \{p_1, \ldots, p_n\}) = \dim H^1_{\mathrm{dR}}(\mathbb{R}^2 \backslash \{p_1, \ldots, p_{n-1}\}) + 1.$$

By induction, we conclude:

$$b_1(\mathbb{R}^2 \backslash \{p_1, \ldots, p_n\}) = \dim H^1_{\mathrm{dR}}(\mathbb{R}^2 \backslash \{p_1, \ldots, p_n\}) = n.$$

$\square$

Example 5.22. Consider the n-sphere S^n (where $n \geq 2$). It can be written as $U \cup V$ where $U := S^n \backslash \{\text{north pole}\}$ and $V := S^n \backslash \{\text{south pole}\}$. Using stereographic projections, one can show both U and V are diffeomorphic to $\mathbb{R}^n$. Furthermore, $U \cap V$ is diffeomorphic to $\mathbb{R}^n \backslash \{0\}$, which deformation retracts to S^{n-1}. Hence $H^k_{\mathrm{dR}}(S^{n-1}) = H^k_{\mathrm{dR}}(U \cap V)$ for any k.

Now consider the Mayer-Vietoris sequence with these U and V, we have for each $k \geq 2$ an exact sequence:

$$\underbrace{H^{k-1}_{\mathrm{dR}}(U) \oplus H^{k-1}_{\mathrm{dR}}(V)}_{0} \to H^{k-1}_{\mathrm{dR}}(U \cap V) \to H^k_{\mathrm{dR}}(S^n) \to \underbrace{H^k_{\mathrm{dR}}(U) \oplus H^k_{\mathrm{dR}}(V)}_{0}.$$

This shows $H^{k-1}_{\mathrm{dR}}(S^{n-1}) \cong H^k_{\mathrm{dR}}(S^n)$ for any $k \geq 2$. By induction, we conclude that $H^n_{\mathrm{dR}}(S^n) \cong H^1_{\mathrm{dR}}(\mathbb{S}^1) \cong \mathbb{R}$ for any $n \geq 2$. $\square$

5.3.4. Proof of Mayer-Vietoris Theorem. To end this chapter (and this course), we present the proof of the Mayer-Vietoris's Theorem (Theorem 5.16). As mentioned before, the proof is purely algebraic. The key ingredient of the proof applies to many other kinds of cohomologies as well (de Rham cohomology is only one kind of many types of cohomology).

For simplicity, we denote:

$$X^k := \wedge^k T^*M \qquad Y^k := \wedge^k T^*U \oplus \wedge^k T^*V \qquad Z^k := \wedge^k T^*(U \cap V)$$
$$H^k(X) := H^k_{\mathrm{dR}}(M) \qquad H^k(Y) := H^k_{\mathrm{dR}}(U) \oplus H^k_{\mathrm{dR}}(V) \qquad H^k(Z) := H^k_{\mathrm{dR}}(U \cap V)$$

Furthermore, we denote the pull-back maps $i_U^* - i_V^*$ and $j_U^* \oplus j_V^*$ by simply i and j respectively. We then have the following commutative diagram between all these X, Y and Z:

$$\begin{tikzcd}
0 \arrow[r] & X^k \arrow[r, "j"] \arrow[d, "d"] & Y^k \arrow[r, "i"] \arrow[d, "d"] & Z^k \arrow[r] \arrow[d, "d"] & 0 \\
0 \arrow[r] & X^{k+1} \arrow[r, "j"] \arrow[d, "d"] & Y^{k+1} \arrow[r, "i"] \arrow[d, "d"] & Z^{k+1} \arrow[r] \arrow[d, "d"] & 0 \\
0 \arrow[r] & X^{k+2} \arrow[r, "j"] & Y^{k+2} \arrow[r, "i"] & Z^{k+2} \arrow[r] & 0
\end{tikzcd}$$

The maps in the diagram commute because the exterior derivative d commute with any pull-back map. The map $d : Y^k \to Y^{k+1}$ takes (ω, η) to $(d\omega, d\eta)$.

To give a proof of the Mayer-Vietoris Theorem, we first need to construct a linear map $\delta : H^k_{\mathrm{dR}}(Z) \to H^{k+1}_{\mathrm{dR}}(Z)$. Then, we need to check that the connected sequence:

$$\cdots \xrightarrow{i} H^k(Z) \xrightarrow{\delta} H^{k+1}(X) \xrightarrow{j} H^{k+1}(Y) \xrightarrow{i} H^{k+1}(Z) \xrightarrow{\delta} \cdots$$

is exact. Most arguments involved are done by "chasing the commutative diagram".

Step 1: Construction of $H^k(Z) \xrightarrow{\delta} H^{k+1}(X)$

Let $[\theta] \in H^k(Z)$, where $\theta \in Z^k$ is a closed k-form on $U \cap V$. Recall from Proposition 5.15 that the sequence

$$0 \to X^k \xrightarrow{j} Y^k \xrightarrow{i} Z^k \to 0$$

is exact, and in particular i is surjective. As a result, there exists $\omega \in Y^k$ such that $i(\omega) = \theta$.

From the commutative diagram, we know $id\omega = di\omega = d\theta = 0$, and hence $d\omega \in \ker i$. By exact-ness, $\operatorname{Im} j = \ker i$ and so there exists $\eta \in X^{k+1}$ such that $j(\eta) = d\omega$.

Next we argue that such η must be closed: since $j(d\eta) = d(j\eta) = d(d\omega) = 0$, and j is injective by exact-ness. We must have $d\eta = 0$, and so η represents a class in $H^{k+1}(X)$. To summarize, given $[\theta] \in H^k(Z)$, ω and η are elements such that

$$i(\omega) = \theta \quad \text{and} \quad j(\eta) = d\omega.$$

We then define $\delta[\theta] := [\eta] \in H^{k+1}(X)$.

Step 2: Verify that δ is a well-defined map

Suppose $[\theta] = [\theta']$ in $H^k_{\mathrm{dR}}(Z)$. Let $\omega' \in Y^k$ and $\eta' \in X^{k+1}$ be the corresponding elements associated with θ', i.e.

$$i(\omega') = \theta' \quad \text{and} \quad j(\eta') = d\omega'.$$

We need to show $[\eta] = [\eta']$ in $H^{k+1}(X)$.

From $[\theta] = [\theta']$, there exists a $(k-1)$-form β in Z^{k-1} such that $\theta - \theta' = d\beta$, which implies:

$$i(\omega - \omega') = \theta - \theta' = d\beta.$$

By surjectivity of $i : Y^{k-1} \to Z^{k-1}$, there exists $\alpha \in Y^{k-1}$ such that $i\alpha = \beta$. Then we get:

$$i(\omega - \omega') = d(i\alpha) = id\alpha$$

which implies $(\omega - \omega') - d\alpha \in \ker i$.

By exact-ness, $\ker i = \operatorname{Im} j$ and so there exists $\gamma \in X^k$ such that

$$j\gamma = (\omega - \omega') - d\alpha.$$

Differentiating both sides, we arrive at:

$$dj\gamma = d(\omega - \omega') - d^2\alpha = j(\eta - \eta').$$

Therefore, $jd\gamma = dj\gamma = j(\eta - \eta')$, and by injectivity of j, we get:

$$\eta - \eta' = d\gamma$$

and so $[\eta] = [\eta']$ in $H^{k+1}(X)$.

Step 3: Verify that δ is a linear map

We leave this step as an exercise for readers.

Step 4: Check that $H^k(Y) \xrightarrow{i} H^k(Z) \xrightarrow{\delta} H^{k+1}(X)$ is exact

To prove $\operatorname{Im} i \subset \ker \delta$, we take an arbitrary $[\theta] \in \operatorname{Im} i \subset H^k(Z)$, there is $[\omega] \in H^k(Y)$ such that $[\theta] = i[\omega]$, we will show $\delta[\theta] = 0$. Recall that $\delta[i\omega]$ is the element $[\eta]$ in $H^{k+1}(X)$ such that $j\eta = d\omega$. Now that ω is closed, the injectivity of j implies $\eta = 0$. Therefore, $\delta[\theta] = \delta[i\omega] = [0]$, proving $[\theta] \in \ker \delta$.

Next we show $\ker \delta \subset \operatorname{Im} i$. Suppose $[\theta] \in \ker \delta$, and let ω and η be the forms such that $i(\omega) = \theta$ and $j(\eta) = d\omega$. Then $[\eta] = \delta[\theta] = [0]$, so there exists $\gamma \in X^{k-1}$ such that $\eta = d\gamma$, which implies $j(d\gamma) = d\omega$, and so $\omega - j\gamma$ is closed. By exact-ness, $i(j\gamma) = 0$, and so:

$$\theta = i(\omega) = i(\omega - j\gamma).$$

For $\omega - j\gamma$ being closed, we conclude $[\theta] = i[\omega - j\gamma] \in \operatorname{Im} i$ in $H^k(Z)$.

Step 5: Check that $H^k(Z) \xrightarrow{\delta} H^{k+1}(X) \xrightarrow{j} H^{k+1}(Y)$ is exact

First show $\operatorname{Im} \delta \subset \ker j$. Let $[\theta] \in H^{k+1}(Z)$, then $\delta[\theta] = [\eta]$ where

$$i(\omega) = \theta \quad \text{and } j(\eta) = d\omega.$$

As a result, $j\delta[\theta] = j[\eta] = [d\omega] = [0]$. This shows $\delta[\theta] \in \ker j$.

Next we show $\ker j \subset \operatorname{Im} \delta$. Let $j[\omega] = [0]$, then $j\omega = d\alpha$ for some $\alpha \in Y^k$. Since:

$$i(\alpha) = i\alpha \quad \text{and} \quad j(\omega) = d\alpha$$

We conclude $\delta[i\alpha] = [\omega]$, or in other words, $[\omega] \in \operatorname{Im} \delta$.

Step 6: Check that $H^{k+1}(X) \xrightarrow{j} H^{k+1}(Y) \xrightarrow{i} H^{k+1}(Z)$ is exact

The inclusion $\operatorname{Im} j \subset \ker i$ follows from the fact that $i(j\eta) = 0$ for any closed $\eta \in X^{k+1}$, and hence $ij[\eta] = [0]$. Finally, we show $\ker i \subset \operatorname{Im} j$: suppose $[\omega] \in \ker i$ so that $i\omega = d\beta$ for some $\beta \in Z^k$. By surjectivity of $i : Y^k \to Z^k$, there exists $\alpha \in Y^k$ such that $\beta = i\alpha$. As a result, we get:

$$i\omega = di\alpha = id\alpha \quad \Longrightarrow \quad \omega - d\alpha \in \ker i.$$

Since $\ker i = \operatorname{Im} j$ on the level of $X^{k+1} \to Y^{k+1} \to Z^{k+1}$, there exists $\gamma \in X^{k+1}$ such that $j\gamma = \omega - d\alpha$. One can easily show γ is closed by injectivity of j:

$$jd\gamma = dj\gamma = d(\omega - d\alpha) = 0 \quad \Longrightarrow \quad d\gamma = 0$$

and so $[\gamma] \in H^{k+1}(X)$. Finally, we conclude:

$$j[\gamma] = [\omega - d\alpha] = [\omega]$$

and so $[\omega] \in \operatorname{Im} j$.

*** End of the proof of the Mayer-Vietoris Theorem ***
**** End of MATH 4033 ****
***** I hope you enjoy it. *****

Geometry of Curves

> "Arc, amplitude, and curvature sustain a similar relation to each other as time, motion and velocity, or as volume, mass and density."
>
> *Carl Friedrich Gauss*

The rest of this lecture notes is about geometry of curves and surfaces in $\mathbb{R}^2$ and $\mathbb{R}^3$. It will not be covered during lectures in MATH 4033 and is not essential to the course. However, it is recommended for readers who want to acquire *workable* knowledge on Differential Geometry.

A.1. Curvature and Torsion

A.1.1. Regular Curves. A curve in the Euclidean space $\mathbb{R}^n$ is regarded as a function $\mathsf{r}(t)$ from an interval I to $\mathbb{R}^n$. The interval I can be finite, infinite, open, closed or half-open. Denote the coordinates of $\mathbb{R}^n$ by $(x_1, x_2, \ldots, x_n)$, then a curve $\mathsf{r}(t)$ in $\mathbb{R}^n$ can be written in coordinate form as:

$$\mathsf{r}(t) = (x_1(t), x_2(t), \ldots, x_n(t)).$$

One easy way to make sense of a curve is to regard it as the trajectory of a particle. At any time t, the functions $x_1(t), x_2(t), \ldots, x_n(t)$ give the coordinates of the particle in $\mathbb{R}^n$. Assuming all $x_i(t)$, where $1 \leq i \leq n$, are at least twice differentiable, then the first derivative $\mathsf{r}'(t)$ represents the *velocity* of the particle, its magnitude $|\mathsf{r}'(t)|$ is the *speed* of the particle, and the second derivative $\mathsf{r}''(t)$ represents the *acceleration* of the particle.

As a course on *Differential* Manifolds/Geometry, we will mostly study those curves which are infinitely differentiable (i.e. C^∞). For some technical purposes as we will explain later, we only study those C^∞ curves $\mathsf{r}(t)$ whose velocity $\mathsf{r}'(t)$ is never zero. We call those curves:

Definition A.1 (Regular Curves)**.** A *regular curve* is a C^∞ function $\mathsf{r}(t) : I \to \mathbb{R}^n$ such that $\mathsf{r}'(t) \neq 0$ for any $t \in I$.

Example A.2. The curve $\mathsf{r}(t) = (\cos(e^t), \sin(e^t))$, where $t \in (-\infty, \infty)$, is a regular curve since $\mathsf{r}'(t) = \left(-e^t \sin(e^t), e^t \cos(e^t)\right)$ and $|\mathsf{r}'(t)| = e^t \neq 0$ for any t.

However, $\widetilde{r}(t) = (\cos t^2, \sin t^2)$, where $t \in (-\infty, \infty)$, is *not* a regular curve since $\widetilde{r}'(t) = (-2t \sin t^2, 2t \cos t^2)$ and so $\widetilde{r}'(0) = 0$.

Although both curves $r(t)$ and $\widetilde{r}(t)$ represent the unit circle centered at the origin in $\mathbb{R}^2$, one is regular but another is not. Therefore, the term *regular* refers to the parametrization rather than the trajectory. $\square$

A.1.2. Arc-Length Parametrization. From Calculus, the arc-length of a curve $r(t)$ from $t = t_0$ to $t = t_1$ is given by:

$$\int_{t_0}^{t_1} \left|r'(t)\right| dt.$$

Now suppose the curve $r(t)$ starts at $t = 0$ (call it the initial time). Then the following quantity:

$$s(t) := \int_0^t \left|r'(\tau)\right| d\tau$$

measures the distance traveled by the particle after t unit time since its initial time.

Given a curve $r(t) = (\cos(e^t - 1), \sin(e^t - 1))$, we have

$$\begin{aligned} r'(t) &= (-e^t \sin(e^t - 1), e^t \cos(e^t - 1)), \\ \left|r'(t)\right| &= e^t \neq 0 \quad \text{for any } t \in (-\infty, \infty). \end{aligned}$$

Therefore, $r(t)$ is a regular curve. By an easy computation, one can show $s(t) = e^t - 1$ and so, regarding t as a function of s, we have $t(s) = \log(s + 1)$. By substituting $t = \log(s + 1)$ into $r(t)$, we get:

$$r(t(s)) = r(\log(s+1)) = \Big(\cos(e^{\log(s+1)} - 1), \sin(e^{\log(s+1)} - 1)\Big) = (\cos s, \sin s).$$

The curve $r(t(s))$ is ultimately a function of s. With abuse of notations, we denote $r(t(s))$ simply by $r(s)$. Then, this $r(s)$ has the same trajectory as $r(t)$ and both curves at C^∞. The difference is that the former travels at a unit speed. The curve $r(s)$ is a *reparametrization* of $r(t)$, and is often called an *arc-length parametrization* of the curve.

However, if we *attempt* to do find a reparametrization on a non-regular curve say $\widetilde{r}(t) = (\cos(t^2), \sin(t^2))$, in a similar way as the above, we can see that such the reparametrization obtained will not be smooth. To see this, we first compute

$$s(t) = \int_0^t \left|\widetilde{r}'(\tau)\right| d\tau = \int_0^t 2|\tau| d\tau = \begin{cases} t^2 & \text{if } t \geq 0; \\ -t^2 & \text{if } t < 0. \end{cases}$$

Therefore, regarding t as a function of s, we have

$$t(s) = \begin{cases} \sqrt{s} & \text{if } s \geq 0; \\ -\sqrt{-s} & \text{if } s < 0. \end{cases}$$

Then,

$$\widetilde{r}(s) := \widetilde{r}(t(s)) = \begin{cases} (\cos(s), \sin(s)) & \text{if } s \geq 0; \\ (\cos(-s), \sin(-s)) & \text{if } s < 0, \end{cases}$$

or in short, $\widetilde{r}(s) = (\cos(s), \sin|s|)$, which is not differentiable at $s = 0$.

It turns out the reason why the reparametrization by s works well for $r(t)$ but not for $\widetilde{r}(t)$ is that the former is regular but the later is not. In general, one can always reparametrize a regular curve by its arc-length s. Let's state it as a theorem:

Theorem A.3. *Given any regular curve* $\mathsf{r}(t) : I \to \mathbb{R}^n$*, one can always reparametrize it by arc-length. Precisely, let* $t_0 \in I$ *be a fixed number and consider the following function of* t*:*

$$s(t) := \int_{t_0}^{t} \left|\mathsf{r}'(\tau)\right| d\tau.$$

Then, t *can be regarded as a* C^∞ *function of* s*, and the reparametrized curve* $\mathsf{r}(s) := \mathsf{r}(t(s))$ *is a regular curve such that* $\left|\frac{d}{ds}\mathsf{r}(s)\right| = 1$ *for any* s*.*

Proof. The Fundamental Theorem of Calculus shows

$$\frac{ds}{dt} = \frac{d}{dt}\int_{t_0}^{t} \left|\mathsf{r}'(\tau)\right| d\tau = \left|\mathsf{r}'(t)\right| > 0.$$

We have $|\mathsf{r}'(t)| > 0$ since $\mathsf{r}(t)$ is a regular curve. Now $s(t)$ is a strictly increasing function of t, so one can regard t as a function of s by the Inverse Function Theorem. Since $s(t)$ is C^∞ (because $\mathsf{r}(t)$ is C^∞ and $|\mathsf{r}'(t)| \neq 0$), by the Inverse Function Theorem $t(s)$ is C^∞ too.

To verify that $\left|\frac{d}{ds}\mathsf{r}(s)\right| = 1$, we use the chain rule:

$$\begin{aligned}
\frac{d}{ds}\mathsf{r}(s) &= \frac{d\mathsf{r}}{dt} \cdot \frac{dt}{ds} \\
&= \mathsf{r}'(t) \cdot \frac{1}{\frac{ds}{dt}} \\
\left|\frac{d}{ds}\mathsf{r}(s)\right| &= \left|\mathsf{r}'(t)\right| \cdot \frac{1}{|\mathsf{r}'(t)|} = 1.
\end{aligned}$$

□

Exercise A.1. Determine whether each of the following is a regular curve. If so, reparametrize the curve by arc-length:

(a) $\mathsf{r}(t) = (\cos t, \sin t, t), \quad t \in (-\infty, \infty)$

(b) $\mathsf{r}(t) = (t - \sin t, 1 - \cos t), \quad t \in (-\infty, \infty)$

A.1.3. Definition of Curvature. Curvature is quantity that measures the sharpness of a curve, and is closely related to the acceleration. Imagine you are driving a car along a curved road. On a sharp turn, the force exerted on your body is proportional to the acceleration according to the Newton's Second Law. Therefore, given a parametric curve $\mathsf{r}(t)$, the magnitude of the acceleration $|\mathsf{r}''(t)|$ *somewhat* reflects the sharpness of the path – the sharper the turn, the larger the $|\mathsf{r}''(t)|$.

However, the magnitude $|\mathsf{r}''(t)|$ is not *only* affected by the sharpness of the curve, but also on how *fast* you drive. In order to give a *fair* and *standardized* measurement of sharpness, we need to get an arc-length parametrization $\mathsf{r}(s)$ so that the "car" travels at unit speed.

Definition A.4 (Curvature)**.** Let $\mathsf{r}(s) : I \to \mathbb{R}^n$ be an arc-length parametrization of a path γ in $\mathbb{R}^n$. The curvature of γ is a function $\kappa : I \to \mathbb{R}$ defined by:

$$\kappa(s) = \left|\mathsf{r}''(s)\right|.$$

Remark A.5. Since an arc-length parametrization is required in the definition, we talk about curvature for only for regular curves. □

Another way (which is less *physical*) to understand curvature is to regard $\mathsf{r}''(s)$ as $\frac{d}{ds}\mathsf{T}(s)$ where $\mathsf{T}(s) := \mathsf{r}'(s)$ is the unit tangent vector at $\mathsf{r}(s)$. The curvature $\kappa(s)$ is then given by $\left|\frac{d}{ds}\mathsf{T}(s)\right|$ which measures how fast the unit tangents $\mathsf{T}(s)$ move or turn along the curve (see Figure A.1).

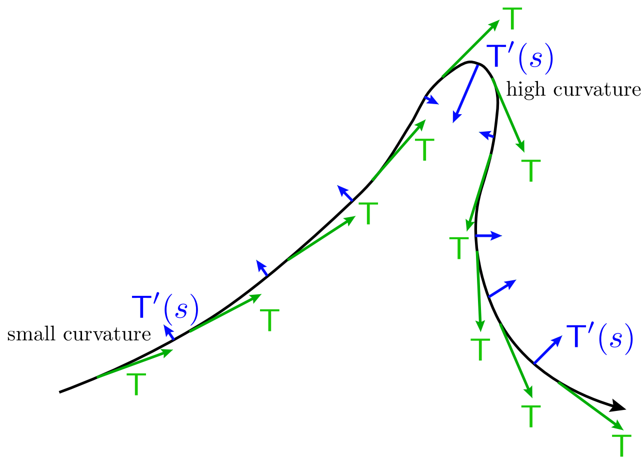


Figure A.1. curvature measures how fast the unit tangents move

Example A.6. The circle of radius R centered at the origin $(0,0)$ on the xy-plane can be parametrized by $\mathsf{r}(t) = (R\cos t,\ R\sin t)$. It can be easily verified that $|\mathsf{r}'(t)| = R$ and so $\mathsf{r}(t)$ is not an arc-length parametrization.

To find an arc-length parametrization, we let:

$$s(t) = \int_0^t \left|\mathsf{r}'(\tau)\right|\,d\tau = \int_0^t R\,d\tau = Rt.$$

Therefore, $t(s) = \frac{s}{R}$ as a function of s and so an arc-length parametrization of the circle is:

$$\mathsf{r}(s) := \mathsf{r}(t(s)) = \left(R\cos\frac{s}{R},\ R\sin\frac{s}{R}\right).$$

To find its curvature, we compute:

$$\begin{aligned}
\mathsf{r}'(s) &= \frac{d}{ds}\left(R\cos\frac{s}{R},\ R\sin\frac{s}{R}\right) \\
&= \left(-\sin\frac{s}{R},\ \cos\frac{s}{R}\right) \\
\mathsf{r}''(s) &= \left(-\frac{1}{R}\cos\frac{s}{R},\ -\frac{1}{R}\sin\frac{s}{R}\right) \\
\kappa(s) &= \left|\mathsf{r}''(s)\right| = \frac{1}{R}.
\end{aligned}$$

Thus the curvature of the circle is given by $\frac{1}{R}$, i.e. the larger the circle, the smaller the curvature. □

Exercise A.2. Find an arc-length parametrization of the helix:

$$\mathsf{r}(t) = (a\cos t,\ a\sin t,\ bt)$$

where a and b are positive constants. Hence compute its curvature.

Exercise A.3. Prove that a regular curve $\mathsf{r}(t)$ is a straight line if and only if its curvature κ is identically zero.

A.1.4. Curvature Formula. Although the curvature is defined as $\kappa(s) = |\mathsf{r}''(s)|$ where $\mathsf{r}(s)$ is an arc-length parametrization of the curve, it is very often impractical to compute the curvature this way. The main reason is that the arc-length parametrizations of many paths are very difficult to find explicitly. A "notorious" example is the ellipse:

$$\mathsf{r}(t) = (a\cos t,\ b\sin t)$$

where a and b are positive constants with $a \neq b$. The arc-length function is given by:

$$s(t) = \int_0^t \sqrt{a^2\sin^2\tau + b^2\cos^2\tau}\, d\tau.$$

While it is very easy to compute the integral when $a = b$, there is no closed form or explicit anti-derivative for the integrand if $a \neq b$. Although the arc-length parametrization *exists* theoretically speaking (Theorem A.3), it cannot be written down explicitly and so the curvature cannot be computed from the definition.

The purpose of this section is to derive a formula for computing curvature without the need of finding its arc-length parametrization. To begin, we first prove the following important observation:

Lemma A.7. *Let $\mathsf{r}(s) : I \to \mathbb{R}^n$ be a curve parametrized by arc-length, then the velocity $\mathsf{r}'(s)$ and the acceleration $\mathsf{r}''(s)$ is always orthogonal for any $s \in I$.*

Proof. Since $\mathsf{r}(s)$ is parametrized by arc-length, we have $|\mathsf{r}'(s)| = 1$ for any s, and so:

$$\begin{aligned}
\frac{d}{ds}\left|\mathsf{r}'(s)\right|^2 &= \frac{d}{ds}1 = 0 \\
\frac{d}{ds}\left(\mathsf{r}'(s)\cdot\mathsf{r}'(s)\right) &= 0 \\
\mathsf{r}''(s)\cdot\mathsf{r}'(s) + \mathsf{r}'(s)\cdot\mathsf{r}''(s) &= 0 \\
2\mathsf{r}''(s)\cdot\mathsf{r}'(s) &= 0 \\
\mathsf{r}''(s)\cdot\mathsf{r}'(s) &= 0
\end{aligned}$$

Therefore, $\mathsf{r}'(s)$ is orthogonal to $\mathsf{r}''(s)$ for any s. □

Proposition A.8. *Given any regular curve $\mathsf{r}(t)$ in $\mathbb{R}^3$, the curvature as a function of t can be computed by the following formula:*

$$\kappa(t) = \frac{|\mathsf{r}'(t) \times \mathsf{r}''(t)|}{|\mathsf{r}'(t)|^3}.$$

Proof. Since $\mathsf{r}(t)$ is a regular curve, there exists an arc-length parametrization $\mathsf{r}(t(s))$, which for simplicity we denote it by $\mathsf{r}(s)$. From now on, we denote $\mathsf{r}'(t)$ as $\frac{d\mathsf{r}(t)}{dt}$, regarding t as the parameter of the curve, and $\mathsf{r}'(s)$ as $\frac{d\mathsf{r}(s)}{ds}$ regarding s as the parameter of the curve.

By the chain rule, we have:

$$\frac{d\mathsf{r}}{dt} = \frac{d\mathsf{r}}{ds}\,\frac{ds}{dt} = \mathsf{r}'(s)\,\frac{ds}{dt} \tag{A.1}$$

$$\begin{aligned}\frac{d^2\mathsf{r}}{dt^2} &= \frac{d}{dt}\left(\frac{d\mathsf{r}}{dt}\right) = \frac{d}{dt}\left(\mathsf{r}'(s)\,\frac{ds}{dt}\right) && \text{(from (A.1))}\\ &= \frac{d\mathsf{r}'(s)}{dt}\,\frac{ds}{dt} + \mathsf{r}'(s)\,\frac{d^2s}{dt^2}\end{aligned} \tag{A.2}$$

By the chain rule again, we get:

$$\frac{d\mathsf{r}'(s)}{dt} = \frac{d\mathsf{r}'(s)}{ds}\,\frac{ds}{dt} = \mathsf{r}''(s)\,\frac{ds}{dt}$$

Substitute this back to (A.2), we obtain:

$$\frac{d^2\mathsf{r}}{dt^2} = \mathsf{r}''(s)\,\left(\frac{ds}{dt}\right)^2 + \mathsf{r}'(s)\,\frac{d^2s}{dt^2} \tag{A.3}$$

Taking the cross product of (A.1) and (A.3) yields:

$$\frac{d\mathsf{r}}{dt} \times \frac{d^2\mathsf{r}}{dt^2} = \left(\frac{ds}{dt}\right)^3\,\mathsf{r}'(s) \times \mathsf{r}''(s) + \underbrace{\frac{d^2s}{dt^2}\,\frac{ds}{dt}\,\mathsf{r}'(s) \times \mathsf{r}'(s)}_{=0} = \left(\frac{ds}{dt}\right)^3\,\mathsf{r}'(s) \times \mathsf{r}''(s). \tag{A.4}$$

Since $\mathsf{r}'(s)$ and $\mathsf{r}''(s)$ are two orthogonal vectors by Lemma A.7, we have $|\mathsf{r}'(s) \times \mathsf{r}''(s)| = |\mathsf{r}'(s)|\,|\mathsf{r}''(s)| = \kappa(s)$. Taking the magnitude on both sides of (A.4), we get:

$$\left|\frac{d\mathsf{r}}{dt} \times \frac{d^2\mathsf{r}}{dt^2}\right| = \kappa\left|\frac{ds}{dt}\right|^3.$$

Therefore, we get:

$$\kappa = \frac{|\mathsf{r}'(t) \times \mathsf{r}''(t)|}{\left|\frac{ds}{dt}\right|^3}.$$

The proof can be easily completed by the definition of $s(t)$ and the Fundamental Theorem of Calculus:

$$\begin{aligned} s &= \int_0^t \left|\mathsf{r}'(\tau)\right|\,d\tau\\ \frac{ds}{dt} &= \left|\mathsf{r}'(t)\right|\end{aligned}$$

□

Remark A.9. Since the cross product is involved, Proposition A.8 can only be used for curves in $\mathbb{R}^2$ or $\mathbb{R}^3$. To apply the result for curves in $\mathbb{R}^2$, say $\mathsf{r}(t) = (x(t),\ y(t))$, one may regard it as the curve $\mathsf{r}(t) = (x(t),\ y(t),\ 0)$ in $\mathbb{R}^3$. □

By Proposition A.8, the curvature of the ellipse can be computed easily. See the example below:

Example A.10. Let $\mathsf{r}(t) = (a\cos t,\ b\sin t,\ 0)$ be a parametrization of an ellipse on the xy-plane where a and b are positive constants, then we have:

$$\begin{aligned}\mathsf{r}'(t) &= (-a\sin t,\ b\cos t,\ 0)\\ \mathsf{r}''(t) &= (-a\cos t,\ -b\sin t,\ 0)\\ \mathsf{r}'(t) \times \mathsf{r}''(t) &= (ab\sin^2 t + ab\cos^2 t)\ \mathsf{k} = ab\ \mathsf{k}\end{aligned}$$

Therefore, by Proposition A.8, it's curvature function is given by:

$$\kappa(t) = \frac{|\mathsf{r}'(t) \times \mathsf{r}''(t)|}{|\mathsf{r}'(t)|^3} = \frac{ab}{(a^2 \sin^2 t + b^2 \cos^2 t)^{3/2}}.$$

□

Exercise A.4. Consider the graph of a smooth function $y = f(x)$. Regarding the graph as a curve in $\mathbb{R}^3$, it can be parametrized using x as the parameter by $\mathsf{r}(x) = (x,\ f(x),\ 0)$. Show that the curvature of the graph is given by:

$$\kappa(x) = \frac{|f''(x)|}{(1 + f'(x)^2)^{3/2}}.$$

Exercise A.5. For each of the following curves: (i) compute the curvature $\kappa(t)$ using Proposition A.8; (ii) If it is easy to find an explicit arc-length parametrization of the curve, compute also the curvature from the definition; (iii) find the (x, y, z)-coordinates of the point(s) on the curve at which the curvature is the maximum.

(a) $\mathsf{r}(t) = (3\cos t,\ 4\cos t,\ 5t)$.

(b) $\mathsf{r}(t) = (t^2,\ 0,\ t)$.

(c) $\mathsf{r}(t) = \left(2t,\ t^2,\ -\frac{1}{3}t^3\right)$.

A.1.5. Frenet-Serret Frame. For now on, we will concentrate on regular curves in $\mathbb{R}^3$. Furthermore, we consider mostly those curves whose curvature function κ is *nowhere vanishing*. Therefore, straight-lines in $\mathbb{R}^3$, or paths such as the graph of $y = x^3$, are excluded in our discussion.

Definition A.11 (Non-degenerate Curves)**.** A regular curve $\mathsf{r}(t) : I \to \mathbb{R}^3$ is said to be *non-degenerate* if its curvature satisfies $\kappa(t) \neq 0$ for any $t \in I$.

We now introduce an important basis of $\mathbb{R}^3$ in the studies of space curves, the Frenet-Serret Frame, or the TNB-frame. It is an orthonormal basis of $\mathbb{R}^3$ associated to each point of a regular curve in $\mathbb{R}^3$.

Definition A.12 (Frenet-Serret Frame)**.** Given a non-degenerate curve $\mathsf{r}(s) : I \to \mathbb{R}^3$ parametrized by arc-length, we define:

$$\begin{aligned} \mathsf{T}(s) &:= \mathsf{r}'(s) && \text{(tangent)} \\ \mathsf{N}(s) &:= \frac{\mathsf{r}''(s)}{|\mathsf{r}''(s)|} && \text{(normal)} \\ \mathsf{B}(s) &:= \mathsf{T}(s) \times \mathsf{N}(s) && \text{(binormal)} \end{aligned}$$

The triple $\{\mathsf{T}(s),\ \mathsf{N}(s),\ \mathsf{B}(s)\}$ is called the *Frenet-Serret Frame* of $\mathbb{R}^3$ at the point $\mathsf{r}(s)$ of the curve. See Figure A.2.

Remark A.13. Note that T is a unit vector since $\mathsf{r}(s)$ is arc-length parametrized. Recall that $\kappa(s) := |\mathsf{r}''(s)|$ and the curve $\mathsf{r}(s)$ is assumed to be non-degenerate. Therefore, N is well-defined for any $s \in I$ and is a unit vector by its definition. From Lemma A.7, T and N are orthogonal to each other for any $s \in I$. Therefore, by the definition of cross product, B is also a unit vector and is orthogonal to both T and N. To conclude, for each fixed $s \in I$, the Frenet-Serret Frame is an orthonormal basis of $\mathbb{R}^3$. □

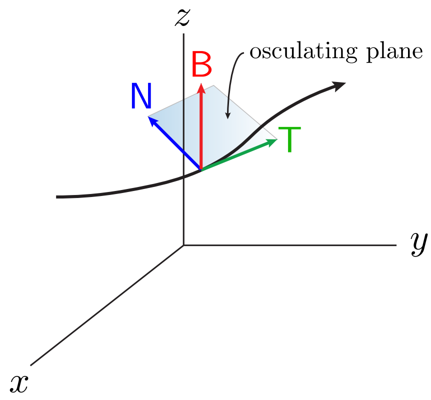


Figure A.2. Frenet-Serret frame

Example A.14. Let $\mathsf{r}(s) = \left(\cos\frac{s}{\sqrt{2}},\ \sin\frac{s}{\sqrt{2}},\ \frac{s}{\sqrt{2}}\right)$ where $s \in \mathbb{R}$. It can be verified easily that it is arc-length parametrized, i.e. $|\mathsf{r}'(s)| = 1$ for any $s \in \mathbb{R}$. The Frenet-Serret Frame of this curve is given by:

$$\begin{aligned}
\mathsf{T}(s) = \mathsf{r}'(s) &= \left(-\frac{1}{\sqrt{2}}\sin\frac{s}{\sqrt{2}},\ \frac{1}{\sqrt{2}}\cos\frac{s}{\sqrt{2}},\ \frac{1}{\sqrt{2}}\right)\\
\mathsf{r}''(s) &= \left(-\frac{1}{2}\cos\frac{s}{\sqrt{2}},\ -\frac{1}{2}\sin\frac{s}{\sqrt{2}},\ 0\right)\\
\mathsf{N}(s) &= \frac{\mathsf{r}''(s)}{|\mathsf{r}''(s)|} = \left(-\cos\frac{s}{\sqrt{2}},\ -\sin\frac{s}{\sqrt{2}},\ 0\right)\\
\mathsf{B}(s) &= \mathsf{T}(s) \times \mathsf{N}(s)\\
&= \begin{vmatrix} \mathsf{i} & \mathsf{j} & \mathsf{k} \\ -\frac{1}{\sqrt{2}}\sin\frac{s}{\sqrt{2}} & \frac{1}{\sqrt{2}}\cos\frac{s}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\cos\frac{s}{\sqrt{2}} & -\sin\frac{s}{\sqrt{2}} & 0 \end{vmatrix}\\
&= \left(\frac{1}{\sqrt{2}}\sin\frac{s}{\sqrt{2}},\ -\frac{1}{\sqrt{2}}\cos\frac{s}{\sqrt{2}},\ \frac{1}{\sqrt{2}}\right)
\end{aligned}$$

□

Definition A.15 (Osculating Plane)**.** Given a non-degenerate arc-length parametrized curve $\mathsf{r}(s) : I \to \mathbb{R}^3$, the *osculating plane* $\Pi(s)$ of the curve is a plane in $\mathbb{R}^3$ containing the point represented by $\mathsf{r}(s)$ and parallel to both $\mathsf{T}(s)$ and N, i.e.

$$\Pi(s) := \mathsf{r}(s) + \operatorname{span}\{\mathsf{T}(s),\ \mathsf{N}(s)\}.$$

(See Figure A.2)

Remark A.16. By the definition of the Frenet-Serret Frame, the binormal vector $\mathsf{B}(s)$ is a unit normal vector to the osculating plane $\Pi(s)$. □

Exercise A.6. Consider the curve $\mathsf{r}(t) = (a\cos t,\ a\sin t,\ bt)$ where a and b positive constants. First find its arc-length parametrization $\mathsf{r}(s) := \mathsf{r}(t(s))$, and then compute its Frenet-Serret Frame.

Exercise A.7. Show that if $\mathsf{r}(s) : I \to \mathbb{R}^3$ is a non-degenerate arc-length parametrized curve contained in the plane $Ax + By + Cz = D$ where A, B, C and D are constants, then $\mathsf{T}(s)$ and $\mathsf{N}(s)$ are parallel to the plane $Ax + By + Cz = 0$ for any $s \in I$, and $\mathsf{B}(s)$ is a constant vector which is normal to the plane $Ax + By + Cz = 0$.

Exercise A.8. [**dC76**, P.23] Let $\mathsf{r}(s)$ be an arc-length parametrized curve in $\mathbb{R}^3$. The normal line at $\mathsf{r}(s)$ is the infinite straight line parallel to $\mathsf{N}(s)$ passing through the point represented by $\mathsf{r}(s)$. Suppose the normal line at every $\mathsf{r}(s)$ pass through a fixed point $\mathsf{p} \in \mathbb{R}^3$. Show that $\mathsf{r}(s)$ is a part of a circle.

A.1.6. Torsion. If the curve $\mathsf{r}(s) : I \to \mathbb{R}^3$ is contained in a plane Π in $\mathbb{R}^3$, then the osculating plane $\Pi(s)$ coincides the plane Π for any $s \in I$, and hence the binormal vector $\mathsf{B}(s)$ is a unit normal vector to Π for any $s \in I$. By continuity, $\mathsf{B}(s)$ is a constant vector.

On the other hand, the helix considered in Example A.14 is not planar since $\mathsf{B}(s)$ is changing over s. As s increases, the osculating plane $\Pi(s)$ not only translates but also rotates. The magnitude of $\frac{d\mathsf{B}}{ds}$ is therefore a measurement of how much the osculating plane rotates and how non-planar the curve $\mathsf{r}(s)$ looks. It motivates the introduction of *torsion*.

However, instead of defining the torsion of a curve to be $\left|\frac{d\mathsf{B}}{ds}\right|$, we hope to give a *sign* for the torsion. Before we state the definition of torsion, we first prove the following fact:

Lemma A.17. *Given any non-degenerate, arc-length parametrized curve $\mathsf{r}(s) : I \to \mathbb{R}^3$, the vector $\frac{d\mathsf{B}}{ds}$ must be parallel to the normal $\mathsf{N}(s)$ for any $s \in I$.*

Proof. First note that $\{\mathsf{T}(s),\ \mathsf{N}(s),\ \mathsf{B}(s)\}$ is an orthonormal basis of $\mathbb{R}^3$ for any $s \in I$. Hence, we have:

$$\frac{d\mathsf{B}(s)}{ds} = a(s)\mathsf{T}(s) + b(s)\mathsf{N}(s) + c(s)\mathsf{B}(s)$$

where $a(s) = \frac{d\mathsf{B}(s)}{ds} \cdot \mathsf{T}(s)$, $b(s) = \frac{d\mathsf{B}(s)}{ds} \cdot \mathsf{N}(s)$ and $c(s) = \frac{d\mathsf{B}(s)}{ds} \cdot \mathsf{B}(s)$. It suffices to show $a(s) = c(s) = 0$ for any $s \in I$.

Since $\mathsf{B}(s)$ is unit, one can easily see that $c(s) \equiv 0$ by considering $\frac{d}{ds}|\mathsf{B}|^2$ (c.f. Lemma A.7). To show $a(s) \equiv 0$, we consider the fact that:

$$\mathsf{T}(s) \cdot \mathsf{B}(s) = 0 \quad \text{for any } s \in I.$$

Differentiate both sides with respect to s, we get:

$$\frac{d\mathsf{T}}{ds} \cdot \mathsf{B} + \mathsf{T} \cdot \frac{d\mathsf{B}}{ds} = 0. \tag{A.5}$$

Since $\frac{d\mathsf{T}}{ds} = \frac{d}{ds}\mathsf{r}'(s) = \mathsf{r}''(s) = \kappa\mathsf{N}$, we get $\frac{d\mathsf{T}}{ds} \cdot \mathsf{B} = 0$ by the definition of B.

Combining this result with (A.5), we get $a(s) = \mathsf{T} \cdot \frac{d\mathsf{B}}{ds} = 0$. Hence we have $\frac{d\mathsf{B}}{ds} = b(s)\mathsf{N}$ and it completes the proof. $\square$

Definition A.18 (Torsion)**.** Let $\mathsf{r}(s) : I \to \mathbb{R}^3$ be an arc-length parametrized, non-degenerate curve. The *torsion* of the curve is a function $\tau : I \to \mathbb{R}$ defined by:

$$\tau(s) := -\frac{d\mathsf{B}}{ds} \cdot \mathsf{N}.$$

Remark A.19. By Lemma A.17, the vector $\frac{d\mathsf{B}}{ds}$ and N are parallel. Combining with the fact that N is unit, one can see easily that:

$$|\tau(s)| = \left|\frac{d\mathsf{B}}{ds}\right| |\mathsf{N}| \cos 0 = \left|\frac{d\mathsf{B}}{ds}\right|.$$

Therefore, the torsion can be regarded as a *signed* $\left|\frac{d\mathsf{B}}{ds}\right|$ which measures the rate that the osculating plane rotates as s increases (see Figure A.3). The negative sign appeared in the definition is a historical convention. □

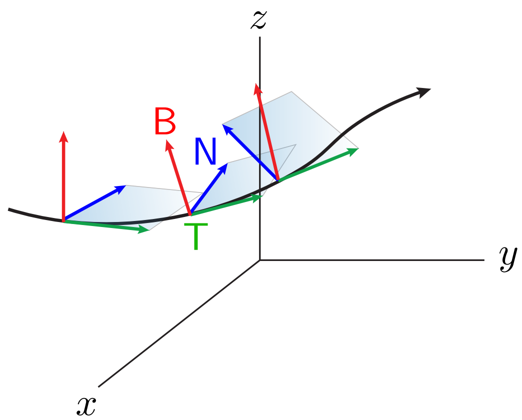


Figure A.3. Torsion measures how fast the osculating plane changes along a curve

Example A.20. Consider the curve $\mathsf{r}(s) = \left(\cos\frac{s}{\sqrt{2}}, \sin\frac{s}{\sqrt{2}}, \frac{s}{\sqrt{2}}\right)$ which is the helix appeared in Example A.14. The normal and binormal were already computed:

$$\mathsf{N}(s) = \frac{\mathsf{r}''(s)}{|\mathsf{r}''(s)|} = \left(-\cos\frac{s}{\sqrt{2}}, -\sin\frac{s}{\sqrt{2}}, 0\right)$$
$$\mathsf{B}(s) = \left(\frac{1}{\sqrt{2}}\sin\frac{s}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\cos\frac{s}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

Taking the derivative, we get:

$$\frac{d\mathsf{B}}{ds} = \left(\frac{1}{2}\cos\frac{s}{\sqrt{2}}, \frac{1}{2}\sin\frac{s}{\sqrt{2}}, 0\right).$$

Therefore, the torsion of the curve is:

$$\tau(s) = -\frac{d\mathsf{B}}{ds}\cdot\mathsf{N} = \frac{1}{2}.$$

□

Exercise A.9. Consider the curve $\mathsf{r}(t) = (a\cos t, a\sin t, bt)$ where a and b are positive constants. Find its torsion $\tau(s)$ as a function of the arc-length parameter s.

Exercise A.10. Let $\mathsf{r}(s) : I \to \mathbb{R}^3$ be a non-degenerate, arc-length parametrized curve. Prove that $\tau(s) = 0$ for any $s \in I$ if and only if $\mathsf{r}(s)$ is contained in a plane. [Hint: for the "only if" part, consider the dot product $\mathsf{B}\cdot\mathsf{T}$.]

Exercise A.11. [**dC76**, P.25] Suppose $\mathsf{r}(s) : I \to \mathbb{R}^3$ is a non-degenerate, arc-length parametrized curve such that $\tau(s) \neq 0$ and $\kappa'(s) \neq 0$ for any $s \in I$. Show that the curve lies on a sphere if and only if $\frac{1}{\kappa^2} + \left(\frac{d}{ds}\frac{1}{\kappa}\right)^2 \frac{1}{\tau^2}$ is a constant.

The torsion of a non-degenerate curve $\mathsf{r}(t)$ can be difficult to compute from the definition since it involves finding an explicit arc-length parametrization. Fortunately, just like the curvature, there is a formula for computing torsion.

Proposition A.21. *Let $\mathsf{r}(t) : I \to \mathbb{R}^3$ be a non-degenerate curve, then the torsion of the curve is given by:*

$$\tau(t) = \frac{(\mathsf{r}'(t) \times \mathsf{r}''(t)) \cdot \mathsf{r}'''(t)}{|\mathsf{r}'(t) \times \mathsf{r}''(t)|^2}.$$

Proof. See Exercise #**??**. □

Exercise A.12. The purpose of this exercise is to give a proof of Proposition A.21. As $\mathsf{r}(t)$ is a (regular) non-degenerate curve, there exist an arc-length parametrization $\mathsf{r}(s) := \mathsf{r}(t(s))$ and a Frenet-Serret Frame $\{\mathsf{T}(s),\ \mathsf{N}(s),\ \mathsf{B}(s)\}$ at every point on the curve. With a little abuse of notations, we denote $\kappa(s) := \kappa(t(s))$ and $\tau(s) := \tau(t(s))$.

(a) Show that
$$\tau(s) = \frac{(\mathsf{r}'(s) \times \mathsf{r}''(s)) \cdot \mathsf{r}'''(s)}{\kappa(s)^2}.$$

(b) Using (A.3) in the proof of Proposition A.8, show that
$$\mathsf{r}'''(t) = \left(\frac{ds}{dt}\right)^3 \mathsf{r}'''(s) + \mathsf{v}(s)$$
where $\mathsf{v}(s)$ is a linear combination of $\mathsf{r}'(s)$ and $\mathsf{r}''(s)$ for any s.

(c) Hence, show that
$$(\mathsf{r}'(s) \times \mathsf{r}''(s)) \cdot \mathsf{r}'''(s) = \frac{(\mathsf{r}'(t) \times \mathsf{r}''(t)) \cdot \mathsf{r}'''(t)}{|\mathsf{r}'(t)|^6}.$$
[Hint: use (A.4) in the proof of Proposition A.8.]

(d) Finally, complete the proof of Proposition A.21. You may use the curvature formula proved in Proposition A.8.

Exercise A.13. Compute the torsion $\tau(t)$ for the ellipsoidal helix:
$$\mathsf{r}(t) = (a\cos t,\ b\sin t,\ ct)$$
where a and b are positive and c is non-zero.

A.2. Fundamental Theorem of Space Curves

In this section, we discuss a deep result about non-degenerate curves in $\mathbb{R}^3$. Given an arc-length parametrized, non-degenerate curve $\mathsf{r}(s)$, one can define its curvature $\kappa(s)$ and torsion $\tau(s)$ as discussed in the previous section. They are scalar-valued functions of s. The former must be positive-valued, while the latter can take any real value. Both functions are smooth.

Now we ask the following questions:

Existence: If we are *given* a pair of smooth real-valued functions $\alpha(s)$ and $\beta(s)$ defined on $s \in I$ where $\alpha(s) > 0$ for any $s \in I$, does there exist a regular curve $\mathsf{r}(s) : I \to \mathbb{R}^3$ such that its curvature $\kappa(s)$ is identically equal to $\alpha(s)$, and its torsion $\tau(s)$ is identically equal to $\beta(s)$?

Uniqueness: Furthermore, if there are two curves $\mathsf{r}(s)$ and $\bar{\mathsf{r}}(s)$ in $\mathbb{R}^3$ whose curvature are both identical to $\alpha(s)$, and torsion are both identical to $\beta(s)$, then is it necessary that $\mathsf{r}(s) \equiv \bar{\mathsf{r}}(s)$?

The *Fundamental Theorem of Space Curves* answers both questions above. Using the classic existence and uniqueness theorems in Ordinary Differential Equations (ODEs), one can give an affirmative answer to the above existence question – yes, such a curve exists – and an "almost" affirmative answer to the uniqueness question – that is, although the curves $\mathsf{r}(s)$ and $\bar{\mathsf{r}}(s)$ may not be identical, one can be transformed from another by a rigid body motion in $\mathbb{R}^3$. The proof of this theorem is a good illustration of how *Differential Equations* interact with *Differential Geometry* – nowadays a field called *Geometric Analysis*.

FYI: Geometric Analysis

Geometric Analysis is a modern field in mathematics which uses Differential Equations to study Differential Geometry. In the past few decades, there are several crowning achievements in this area. Just to name a few, these include Yau's solution to the Calabi Conjecture (1976), and Hamilton–Perelman's solution to the Poincaré Conjecture (2003), and Brendle–Schoen's solution to the Differentiable Sphere Theorem (2007).

A.2.1. Existence and Uniqueness of ODEs.

A system of ODEs (or ODE system) is a set of one or more ODEs. The general form of an ODE system is:

$$\begin{aligned}
x_1'(t) &= f_1(x_1, x_2, \ldots, x_n, t) \\
x_2'(t) &= f_2(x_1, x_2, \ldots, x_n, t) \\
&\vdots \\
x_n'(t) &= f_n(x_1, x_2, \ldots, x_n, t)
\end{aligned}$$

where t is the independent variable, $x_i(t)$'s are unknown functions, and f_j's are prescribed functions of $(x_1, \ldots, x_n, t)$ from $\mathbb{R}^n \times I \to \mathbb{R}$.

An ODE system with a given initial condition, such as $(x_1(0), \ldots, x_n(0)) = (a_1, \ldots, a_n)$ where a_i's are constants, is called an initial-value problem (IVP).

We first state a fundamental existence and uniqueness theorem for ODE systems:

Theorem A.22 (Existence and Uniqueness Theorem of ODEs). *Given functions f_i's ($1 \leq i \leq n$) defined on $\mathbb{R}^n \times I$, we consider the initial-value problem:*

$$x_i'(t) = f_i(x_1, \ldots, x_n, t) \quad \text{for } 1 \leq i \leq n$$

with initial condition $(x_1(0), \ldots, x_n(0)) = (a_1, \ldots, a_n)$. Suppose for every $1 \leq i, j \leq n$, the first partial derivative $\frac{\partial f_i}{\partial x_j}$ exists and is continuous on $\mathbb{R}^n \times I$, then there exists a unique solution $(x_1(t), \ldots, x_n(t))$, defined at least on a short-time interval $t \in (-\varepsilon, \varepsilon)$, to the initial-value problem. Furthermore, as long as the solution remains bounded, the solution exists for all $t \in I$.

Proof. MATH 4051. □

A.2.2. Frenet-Serret System. Given an arc-length parametrized and non-degenerate curve $\mathsf{r}(s) : I \to \mathbb{R}^3$, recall that tangent and binormal satisfy:

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

Using the fact that $\mathsf{N} = \mathsf{B} \times \mathsf{T}$, one can also compute:

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

The *Frenet-Serret System* is an ODE system for the Frenet-Serret Frame of a non-degenerate curve $\mathsf{r}(s)$:

$$\begin{array}{rcrcr} \mathsf{T}' & = & & \kappa\mathsf{N} & \\ \mathsf{N}' & = & -\kappa\mathsf{T} & & +\tau\mathsf{B} \\ \mathsf{B}' & = & & -\tau\mathsf{N} & \end{array}$$

or equivalently in matrix form:

$$\begin{bmatrix} \mathsf{T} \\ \mathsf{N} \\ \mathsf{B} \end{bmatrix}' = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \mathsf{T} \\ \mathsf{N} \\ \mathsf{B} \end{bmatrix} \tag{A.6}$$

Since each vector of the $\{\mathsf{T}, \mathsf{N}, \mathsf{B}\}$ frame has three components, therefore the Frenet-Serret System (A.6) is an ODE system of 9 equations with 9 unknown functions.

A.2.3. Fundamental Theorem. We now state the main theorem of this section:

Theorem A.23 (Fundamental Theorem of Space Curves). *Given any smooth positive function $\alpha(s) : I \to (0, \infty)$, and a smooth real-valued function $\beta(s) : I \to \mathbb{R}$, there exists an arc-length parametrized, non-degenerate curve $\mathsf{r}(s) : I \to \mathbb{R}^3$ such that its curvature $\kappa(s) \equiv \alpha(s)$ and its torsion $\tau(s) \equiv \beta(s)$.*

Moreover, if $\bar{\mathsf{r}}(\mathsf{s}) : I \to \mathbb{R}^3$ is another arc-length parametrized, non-degenerate curve whose curvature $\bar{\kappa}(s) \equiv \alpha(s)$ and torsion $\bar{\tau}(s) \equiv \beta(s)$, then there exists a 3×3 constant matrix A with $A^T A = I$, and a constant vector p, such that $\bar{\mathsf{r}}(s) = A\mathsf{r}(s) + \mathsf{p}$ for any $s \in I$.

Proof. The existence part consists of three major steps.

Step 1: Use the existence theorem of ODEs (Theorem A.22) to show there exists a moving orthonormal frame $\{\mathsf{e}_1(s), \mathsf{e}_2(s)\ \mathsf{e}_3(s)\}$ which satisfies an ODE system (see (A.7) below) *analogous* to the Frenet-Serret System (A.6).

Step 2: Show that there exists a curve $\mathsf{r}(s)$ whose Frenet-Serret Frame is given by $\mathsf{T}(s) = \mathsf{e}_1(s)$, $\mathsf{N}(s) = \mathsf{e}_2(s)$ and $\mathsf{B}(s) = \mathsf{e}_3(s)$. Consequently, from the system (A.7), one can claim $\mathsf{r}(s)$ is a curve that satisfies the required conditions.

Step 3: Prove the uniqueness part of the theorem.

Step 1: To begin, let's consider the ODE system with unknowns e_1, e_2 and e_3:

$$\begin{bmatrix} \mathsf{e}_1(s) \\ \mathsf{e}_2(s) \\ \mathsf{e}_3(s) \end{bmatrix}' = \begin{bmatrix} 0 & \alpha(s) & 0 \\ -\alpha(s) & 0 & \beta(s) \\ 0 & -\beta(s) & 0 \end{bmatrix} \begin{bmatrix} \mathsf{e}_1(s) \\ \mathsf{e}_2(s) \\ \mathsf{e}_3(s) \end{bmatrix} \tag{A.7}$$

which is an analogous system to the Frenet-Serret System (A.6). Impose the initial conditions:

$$\mathsf{e}_1(0) = \mathsf{i},\ \mathsf{e}_2(0) = \mathsf{j},\ \mathsf{e}_3(0) = \mathsf{k}.$$

Recall that $\alpha(s)$ and $\beta(s)$ are given to be smooth (in particular, continuously differentiable). By Theorem A.22, there exists a solution $\{\mathsf{e}_1(s),\ \mathsf{e}_2(s),\ \mathsf{e}_3(s)\}$ defined on a maximal interval $s \in (T_-, T_+)$ that satisfies the system with the above initial conditions.

Note that $\{\mathsf{e}_1(s),\ \mathsf{e}_2(s),\ \mathsf{e}_3(s)\}$ is orthonormal initially at $s = 0$, we claim it remains so as long as solution exists. To prove this, we first derive (see Exercise A.14):

$$\frac{d}{ds}\begin{bmatrix} \mathsf{e}_1 \cdot \mathsf{e}_1 \\ \mathsf{e}_2 \cdot \mathsf{e}_2 \\ \mathsf{e}_3 \cdot \mathsf{e}_3 \\ \mathsf{e}_1 \cdot \mathsf{e}_2 \\ \mathsf{e}_2 \cdot \mathsf{e}_3 \\ \mathsf{e}_3 \cdot \mathsf{e}_1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 2\alpha & 0 & 0 \\ 0 & 0 & 0 & -2\alpha & 2\beta & 0 \\ 0 & 0 & 0 & 0 & -2\beta & 0 \\ -\alpha & \alpha & 0 & 0 & 0 & \beta \\ 0 & -\beta & \beta & 0 & 0 & -\alpha \\ 0 & 0 & 0 & -\beta & \alpha & 0 \end{bmatrix} \begin{bmatrix} \mathsf{e}_1 \cdot \mathsf{e}_1 \\ \mathsf{e}_2 \cdot \mathsf{e}_2 \\ \mathsf{e}_3 \cdot \mathsf{e}_3 \\ \mathsf{e}_1 \cdot \mathsf{e}_2 \\ \mathsf{e}_2 \cdot \mathsf{e}_3 \\ \mathsf{e}_3 \cdot \mathsf{e}_1 \end{bmatrix} \tag{A.8}$$

Exercise A.14. Verify (A.8).

Regarding $\mathsf{e}_i \cdot \mathsf{e}_j$'s are unknowns, (A.8) is a linear ODE system of 6 equations with initial conditions:

$$(\mathsf{e}_1 \cdot \mathsf{e}_1,\ \mathsf{e}_2 \cdot \mathsf{e}_2,\ \mathsf{e}_3 \cdot \mathsf{e}_3,\ \mathsf{e}_1 \cdot \mathsf{e}_2,\ \mathsf{e}_2 \cdot \mathsf{e}_3,\ \mathsf{e}_3 \cdot \mathsf{e}_1)_{s=0} = (1,\ 1,\ 1,\ 0,\ 0,\ 0)$$

It can be verified easily that the constant solution $(1,\ 1,\ 1,\ 0,\ 0,\ 0)$ is indeed a solution to (A.8). Therefore, by the uniqueness part of Theorem A.22, we must have

$$(\mathsf{e}_1 \cdot \mathsf{e}_1,\ \mathsf{e}_2 \cdot \mathsf{e}_2,\ \mathsf{e}_3 \cdot \mathsf{e}_3,\ \mathsf{e}_1 \cdot \mathsf{e}_2,\ \mathsf{e}_2 \cdot \mathsf{e}_3,\ \mathsf{e}_3 \cdot \mathsf{e}_1) = (1,\ 1,\ 1,\ 0,\ 0,\ 0)$$

for any $s \in (T_-, T_+)$. In other words, the frame $\{\mathsf{e}_1(s),\ \mathsf{e}_2(s),\ \mathsf{e}_3(s)\}$ is orthonormal as long as solution exists.

Consequently, each of $\{\mathsf{e}_1(s),\ \mathsf{e}_2(s),\ \mathsf{e}_3(s)\}$ remains bounded, and by last statement of Theorem A.22, this orthonormal frame can be extended so that it is defined for all $s \in I$.

Step 2: Using the frame $\mathsf{e}_1(s) : I \to \mathbb{R}^3$ obtained in Step 1, we define:

$$\mathsf{r}(s) = \int_0^s \mathsf{e}_1(s)\, ds.$$

Evidently, $\mathsf{r}(s)$ is a curve starting from the origin at $s = 0$. Since $\mathsf{e}_1(s)$ is continuous, $\mathsf{r}(s)$ is well-defined on I and by the Fundamental Theorem of Calculus, we get:

$$\mathsf{T}(s) := \mathsf{r}'(s) = \mathsf{e}_1(s)$$

which is a unit vector for any $s \in I$. Therefore, $\mathsf{r}(s)$ is arc-length parametrized.

Next we verify that $\mathsf{r}(s)$ is the curve require by computing its curvature and torsion. By (A.8),

$$\mathsf{r}''(s) = \mathsf{e}_1'(s) = \alpha(s)\mathsf{e}_2(s)$$

By the fact that $\mathsf{e}_2(s)$ is unit, we conclude that:

$$\kappa(s) = \left|\mathsf{r}''(s)\right| = \alpha(s)$$

and so $\mathsf{N}(s) = \frac{1}{\kappa(s)}\mathsf{r}''(s) = \mathsf{e}_2(s)$. For the binormal, we observe that $\mathsf{e}_3 = \mathsf{e}_1 \times \mathsf{e}_2$ initially at $s = 0$ and that the frame $\{\mathsf{e}_1(s),\ \mathsf{e}_2(s),\ \mathsf{e}_3(s)\}$ remains to be orthonormal for all $s \in I$, we must have $\mathsf{e}_3 = \mathsf{e}_1 \times \mathsf{e}_2$ for all $s \in I$ by continuity. Therefore, $\mathsf{B}(s) = \mathsf{T}(s) \times \mathsf{N}(s) = \mathsf{e}_1(s) \times \mathsf{e}_2(s) = \mathsf{e}_3(s)$ for any $s \in I$. By (A.8), we have:

$$\mathsf{B}'(s) = \mathsf{e}_3'(s) = -\beta(s)\mathsf{e}_2(s) = -\beta(s)\mathsf{N}.$$

Therefore, $\tau(s) = -\mathsf{B}'(s) \cdot \mathsf{N}(s) = \beta(s)$.

Step 3: Now suppose there exists another curve $\bar{\mathsf{r}}(s) : I \to \mathbb{R}^3$ with the same curvature and torsion as $\mathsf{r}(s)$. Let $\{\bar{\mathsf{T}}(s),\ \bar{\mathsf{N}}(s),\ \bar{\mathsf{B}}(s)\}$ be the Frenet-Serret Frame of $\bar{\mathsf{r}}(s)$. We define the matrix:

$$A = \begin{bmatrix}\bar{\mathsf{T}}(0) & \bar{\mathsf{N}}(0) & \bar{\mathsf{B}}(0)\end{bmatrix}.$$

By orthonormality, one can check that $A^T A = I$. We claim that $\bar{\mathsf{r}}(s) = A\mathsf{r}(s) + \bar{\mathsf{r}}(0)$ for any $s \in I$ using again the uniqueness theorem of ODEs (Theorem A.22).

First note that A is an orthogonal matrix, so the Frenet-Serret Frame of the transformed curve $A\mathsf{r}(s) + \bar{\mathsf{r}}(0)$ is given by $\{A\mathsf{T}(s),\ A\mathsf{N}(s),\ A\mathsf{B}(s)\}$ and the frame satisfies the ODE system:

$$\begin{bmatrix}A\mathsf{T}(s)\\ A\mathsf{N}(s)\\ A\mathsf{B}(s)\end{bmatrix}' = \begin{bmatrix}0 & \alpha(s) & 0\\ -\alpha(s) & 0 & \beta(s)\\ 0 & -\beta(s) & 0\end{bmatrix}\begin{bmatrix}A\mathsf{T}(s)\\ A\mathsf{N}(s)\\ A\mathsf{B}(s)\end{bmatrix}$$

since the Frenet-Serret Frame $\{\mathsf{T}(s),\ \mathsf{N}(s),\ \mathsf{B}(s)\}$ does.

Furthermore, the curve $\bar{\mathsf{r}}(s)$ also has curvature $\alpha(s)$ and torsion $\beta(s)$, so its Frenet-Serret Frame $\{\bar{\mathsf{T}}(s),\ \bar{\mathsf{N}}(s),\ \bar{\mathsf{B}}(s)\}$ also satisfies the ODE system:

$$\begin{bmatrix}\bar{\mathsf{T}}(s)\\ \bar{\mathsf{N}}(s)\\ \bar{\mathsf{B}}(s)\end{bmatrix}' = \begin{bmatrix}0 & \alpha(s) & 0\\ -\alpha(s) & 0 & \beta(s)\\ 0 & -\beta(s) & 0\end{bmatrix}\begin{bmatrix}\bar{\mathsf{T}}(s)\\ \bar{\mathsf{N}}(s)\\ \bar{\mathsf{B}}(s)\end{bmatrix}.$$

Initially at $s = 0$, the two Frenet-Serret Frames are equal by the definition of A and choice of $\mathsf{e}_i(0)$'s in Step 1:

$$\begin{aligned}A\mathsf{T}(0) &= \begin{bmatrix}\bar{\mathsf{T}}(0) & \bar{\mathsf{N}}(0) & \bar{\mathsf{B}}(0)\end{bmatrix}\mathsf{i} = \bar{\mathsf{T}}(0)\\ A\mathsf{N}(0) &= \begin{bmatrix}\bar{\mathsf{T}}(0) & \bar{\mathsf{N}}(0) & \bar{\mathsf{B}}(0)\end{bmatrix}\mathsf{j} = \bar{\mathsf{N}}(0)\\ A\mathsf{B}(0) &= \begin{bmatrix}\bar{\mathsf{T}}(0) & \bar{\mathsf{N}}(0) & \bar{\mathsf{B}}(0)\end{bmatrix}\mathsf{k} = \bar{\mathsf{B}}(0)\end{aligned}$$

By the uniqueness part of Theorem A.22, the two frames are equal for all $s \in I$. In particular, we have:

$$A\mathsf{T}(s) \equiv \bar{\mathsf{T}}(s).$$

Finally, to show that $\bar{\mathsf{r}}(s) \equiv A\mathsf{r}(s) + \bar{\mathsf{r}}(0)$, we consider the function

$$f(s) := \left|\bar{\mathsf{r}}(s) - (A\mathsf{r}(s) - \bar{\mathsf{r}}(0))\right|^2.$$

Taking its derivative, we get:

$$\begin{aligned} f'(s) &= 2\left(\bar{\mathsf{r}}'(s) - A\mathsf{r}'(s)\right) \cdot \left(\bar{\mathsf{r}}(s) - (A\mathsf{r}(s) - \bar{\mathsf{r}}(0))\right) \\ &= 2\underbrace{\left(\bar{\mathsf{T}}(s) - A\mathsf{T}(s)\right)}_{=0} \cdot \left(\bar{\mathsf{r}}(s) - (A\mathsf{r}(s) - \bar{\mathsf{r}}(0))\right) \\ &= 0 \end{aligned}$$

for any $s \in I$. Since $f(0) = 0$ initially by the fact that $\mathsf{r}(0) = 0$, we have $f(s) \equiv 0$ and so $\bar{\mathsf{r}}(s) \equiv A\mathsf{r}(s) + \bar{\mathsf{r}}(0)$, completing the proof of the theorem. □

The existence part of Theorem A.23 only shows a curve with prescribed curvature and torsion *exists*, but it is in general difficult to find such a curve explicitly. While the existence part does not have much practical use, the uniqueness part has some nice corollaries.

First recall that a helix is a curve of the form $\mathsf{r}_{a,b}(t) = (a\cos t,\ a\sin t,\ bt)$ where $a \neq 0$ and b can be any real number. It's arc-length parametrization is given by:

$$\mathsf{r}_{a,b}(s) = \left(a\cos\frac{s}{\sqrt{a^2+b^2}},\ a\sin\frac{s}{\sqrt{a^2+b^2}},\ \frac{bs}{\sqrt{a^2+b^2}}\right).$$

It can be computed that its curvature and torsion are both constants:

$$\begin{aligned} \kappa_{a,b}(s) &\equiv \frac{a}{a^2+b^2} \\ \tau_{a,b}(s) &\equiv \frac{b}{a^2+b^2}. \end{aligned}$$

Conversely, given two constants $\kappa_0 > 0$ and $\tau_0 \in \mathbb{R}$, by taking $a = \frac{\kappa_0}{\kappa_0^2+\tau_0^2}$ and $b = \frac{\tau_0}{\kappa_0^2+\tau_0^2}$, the helix $\mathsf{r}_{a,b}(s)$ with this pair of a and b has curvature κ_0 and torsion τ_0. Hence, the uniqueness part of Theorem A.23 asserts that:

Corollary A.24. *A non-degenerate curve $\mathsf{r}(s)$ has constant curvature and torsion if and only if $\mathsf{r}(s)$ is congruent to one of the helices $\mathsf{r}_{a,b}(s)$.*

Remark A.25. Two space curves $\mathsf{r}(s)$ and $\widetilde{\mathsf{r}}(s)$ are said to be congruent if there exists a 3×3 orthogonal matrix A and a constant vector $\mathsf{p} \in \mathbb{R}^3$ such that $\widetilde{\mathsf{r}}(s) = A\mathsf{r}(s) + \mathsf{p}$. In simpler terms, one can obtain $\widetilde{\mathsf{r}}(s)$ by rotating and translating $\mathsf{r}(s)$. □

A.3. Plane Curves

A plane curve $\mathsf{r}(s)$ is an arc-length parametrized curve in $\mathbb{R}^2$. While it can be considered as a space curve by identifying $\mathbb{R}^2$ and the xy-plane in $\mathbb{R}^3$, there are several aspects of plane curves that make them distinguished from space curves.

A.3.1. Signed Curvature. Given an arc-length parametrized curve $\mathsf{r}(s) : I \to \mathbb{R}^2$, we define the tangent frame $\mathsf{T}(s)$ as in space curves, i.e.

$$\mathsf{T}(s) = \mathsf{r}'(s).$$

However, instead of defining the normal frame $\mathsf{N}(s) = \frac{1}{\kappa(s)}\mathsf{T}'(s)$, we use the frame $\mathsf{JT}(s)$ where J is the counter-clockwise rotation by $\frac{\pi}{2}$, i.e.

$$\mathsf{J} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

One can easily check that $\{\mathsf{T}(s), \mathsf{JT}(s)\}$ is an orthonormal frame of $\mathbb{R}^2$ for any $s \in I$. Let's call this the TN-Frame of the curve. We will work with the TN-Frame in place of the Frenet-Serret Frame for plane curves. The reasons for doing so are two-folded. For one thing, the normal frame. $\mathsf{JT}(s)$ is well-defined for any $s \in I$ even though $\kappa(s)$ is zero for some $s \in I$. Hence, one can relax the non-degeneracy assumption here. For another, we can introduce the signed curvature $k(s)$:

Definition A.26 (Signed Curvature)**.** Given an arc-length parametrized plane curve $\mathsf{r}(s) : I \to \mathbb{R}^2$, the signed curvature $k(s) : I \to \mathbb{R}$ is defined as:

$$k(s) := \mathsf{T}'(s) \cdot \mathsf{JT}(s).$$

Note that $\mathsf{T}(s)$ is unit, so by Lemma A.7 we know $\mathsf{T}(s)$ and $\mathsf{T}'(s)$ are always orthogonal and hence it is either in or against the direction of $\mathsf{JT}(s)$. Therefore, we have

$$|k(s)| = \left|\mathsf{T}'(s)\right| |\mathsf{JT}(s)| = \left|\mathsf{r}''(s)\right| = \kappa(s).$$

The sign of $k(s)$ is determined by whether T' and JT are along or against each other (see Figure A.4).

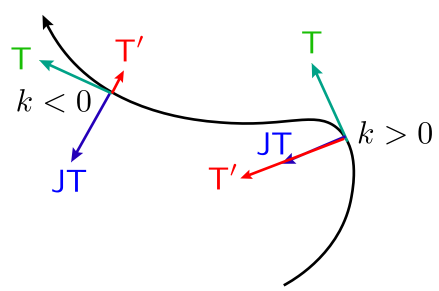


Figure A.4. Signed curvature

Example A.27. Let's compute the signed curvature of the unit circle

$$\mathsf{r}(s) = (\cos \varepsilon s, \ \sin \varepsilon s)$$

where $\varepsilon = \pm 1$. The curve is counter-clockwise orientable when $\varepsilon = 1$, and is clockwise orientable when $\varepsilon = -1$. Clearly it is arc-length parametrized, and

$$\begin{aligned} \mathsf{T}(s) &= \mathsf{r}'(s) = (-\varepsilon \sin \varepsilon s,\ \varepsilon \cos \varepsilon s) \\ \mathsf{JT}(s) &= (-\varepsilon \cos \varepsilon s,\ -\varepsilon \sin \varepsilon s) \\ k(s) &= \mathsf{T}'(s) \cdot \mathsf{JT}(s) \\ &= (-\varepsilon^2 \cos \varepsilon s,\ -\varepsilon^2 \sin \varepsilon s) \cdot (-\varepsilon \cos \varepsilon s,\ -\varepsilon \sin \varepsilon s) \\ &= \varepsilon^3 = \varepsilon. \end{aligned}$$

□

Exercise A.15. Consider a plane curve $\mathsf{r}(s)$ parametrized by arc-length. Let $\theta(s)$ be the angle between the x-axis and the unit tangent vector $\mathsf{T}(s)$. Show that:

$$\mathsf{T}'(s) = \theta'(s)\mathsf{JT}(s) \quad \text{and} \quad k(s) = \theta'(s).$$

Exercise A.16. [**dC76**, P.25] Consider a plane curve $\mathsf{r}(s)$ parametrized by arc-length. Suppose $|\mathsf{r}(s)|$ is maximum at $s = s_0$. Show that:

$$|k(s_0)| \geq \frac{1}{|\mathsf{r}(s_0)|}.$$

Given a regular plane curve $\mathsf{r}(t) : I \to \mathbb{R}^2$, not necessarily arc-length parametrized. Denote the components of the curve by $\mathsf{r}(t) = (x_1(t),\ x_2(t))$.

(a) Show that its signed curvature (as a function of t) is given by:

$$k(t) = \frac{\mathsf{r}''(t) \cdot \mathsf{Jr}'(t)}{|\mathsf{r}'(t)|^3} = \frac{x_1'(t)x_2''(t) - x_2'(t)x_1''(t)}{\left(x_1'(t)^2 + x_2'(t)^2\right)^{3/2}}.$$

(b) Hence, show that the graph of a smooth function $y = f(x)$, when considered as a curve parametrized by x, has signed curvature given by:

$$k(x) = \frac{f''(x)}{(1 + f'(x)^2)^{3/2}}.$$

The signed curvature characterizes regular plane curves, as like curvature and torsion characterize non-degenerate space curves.

Theorem A.28 (Fundamental Theorem of Plane Curves)**.** *Given any smooth real-valued function $\alpha(s) : I \to \mathbb{R}$, there exists a regular plane curve $\mathsf{r}(s) : I \to \mathbb{R}^2$ such that its signed curvature $k(s) \equiv \alpha(s)$. Moreover, if $\bar{\mathsf{r}}(s) : I \to \mathbb{R}^2$ is another regular plane curve such that its signed curvature $\bar{k}(s) \equiv \alpha(s)$, then there exists a 2×2 orthogonal matrix A and a constant vector $\mathsf{p} \in \mathbb{R}^2$ such that $\bar{\mathsf{r}}(s) \equiv A\mathsf{r}(s) + \mathsf{p}$.*

Proof. See Exercise #A.17. □

Exercise A.17. Prove Theorem A.28. Although the proof is similar to that of Theorem A.23 for non-degenerate space curves, please do not use the latter to prove the former in this exercise. Here is a hint on how to begin the proof: Consider the initial-value problem

$$\begin{aligned} \mathsf{e}'(s) &= \alpha(s)\ \mathsf{Je}(s) \\ \mathsf{e}(0) &= \mathsf{i} \end{aligned}$$

Exercise A.18. Using Theorem A.28, show that a regular plane curve has constant signed curvature if and only if it is a straight line or a circle

Exercise A.19. [**Küh05**, P.50] Find an explicit plane curve $\mathsf{r}(s)$ such that the signed curvature is given by $k(s) = \frac{1}{\sqrt{s}}$.

A.3.2. Total Curvature. In this subsection, we explore an interesting result concerning the signed curvature of a plane curve. We first introduce:

Definition A.29 (Closed Curves)**.** An arc-length parametrized plane curve $\mathsf{r}(s) : [0, L] \to \mathbb{R}^2$ is said to be *closed* if $\mathsf{r}(0) = \mathsf{r}(L)$. It is said to be *simple closed* if $\mathsf{r}(s)$ is closed and if $\mathsf{r}(s_1) = \mathsf{r}(s_2)$ for some $s_i \in [0, L]$ then one must have $s_1,\ s_2 = 0$ or L.

The following is a celebrated result that relates the local property (i.e. signed curvature) to the global property (topology) of simple closed curves:

Theorem A.30 (Hopf)**.** *For any arc-length parametrized, simple closed curve* $\mathsf{r}(s) : [0, L] \to \mathbb{R}^2$ *such that* $\mathsf{r}'(0) = \mathsf{r}'(L)$*, we must have:*

$$\int_0^L k(s)\, ds = \pm 2\pi.$$

The original proof was due to Hopf. We will not discuss Hopf's original proof in this course, but we will prove a weaker result, under the same assumption as Theorem A.30, that

$$\int_0^L k(s)\, ds = 2\pi n$$

for some integer n.

Let $\{\mathsf{T}(s),\ J\mathsf{T}(s)\}$ be the TN-frame of $\mathsf{r}(s)$. Since $\mathsf{T}(s)$ is unit for any $s \in [0, L]$, one can find a smooth function $\theta(s) : [0, L] \to \mathbb{R}$ such that

$$\mathsf{T}(s) = (\cos\theta(s),\ \sin\theta(s))$$

for any $s \in [0, L]$. Here $\theta(s)$ can be regarded as $2k\pi +$ angle between $\mathsf{T}(s)$ and i. In order to ensure continuity, we allow $\theta(s)$ to take values beyond $[0, 2\pi]$.

Then $J\mathsf{T}(s) = (-\sin\theta(s),\ \cos\theta(s))$, and so by the definition of $k(s)$, we get:

$$\begin{aligned} k(s) &= \mathsf{T}'(s) \cdot J\mathsf{T}(s) \\ &= (-\theta'(s)\sin\theta(s),\ \theta'(s)\cos\theta(s)) \cdot (-\sin\theta(s),\ \cos\theta(s)) \\ &= \theta'(s). \end{aligned}$$

Therefore, the total curvature is given by:

$$\int_0^L k(s)\, ds = \int_0^L \theta'(s)\, ds = \theta(L) - \theta(0).$$

Since it is assumed that $\mathsf{T}(0) = \mathsf{T}(L)$ in Theorem A.30, we have

$$\theta(L) \equiv \theta(0) \pmod{2\pi}$$

and so we have:

$$\int_0^L k(s)\, ds = 2\pi n$$

for some integer n.

Appendix B

Geometry of Surfaces

> "If you can't explain it *simply*, you don't understand it well enough."
>
> *Albert Einstein*

Throughout this chapter, unless otherwise is stated, M is a regular surface in $\mathbb{R}^3$, p is a point on M and $\mathsf{F}(u, v)$ is a smooth local parametrization around p.

B.1. First Fundamental Form

B.1.1. First Fundamental Form: basic notions. In this subsection, we introduce an important concept in differential geometry – the first fundamental form. Loosely speaking, it is the dot product of the tangent vectors of a regular surface. It captures and encodes intrinsic geometric information (such as curvature) about the surface M. Precisely, it is defined as follows:

Definition B.1 (First Fundamental Form)**.** The *first fundamental form* of a regular surface M is a bilinear map $g_p : T_pM \times T_pM \to \mathbb{R}$ on each T_pM defined as:

$$\begin{aligned} g_p &\left(\alpha \frac{\partial \mathsf{F}}{\partial u}(p) + \beta \frac{\partial \mathsf{F}}{\partial v}(p), \gamma \frac{\partial \mathsf{F}}{\partial u}(p) + \delta \frac{\partial \mathsf{F}}{\partial v}(p)\right) \\ &:= \left(\alpha \frac{\partial \mathsf{F}}{\partial u}(p) + \beta \frac{\partial \mathsf{F}}{\partial v}(p)\right) \cdot \left(\gamma \frac{\partial \mathsf{F}}{\partial u}(p) + \delta \frac{\partial \mathsf{F}}{\partial v}(p)\right) \end{aligned}$$

where the "dot" on the right hand side is the usual dot product on $\mathbb{R}^3$. Whenever the point p is clear from the context, we can omit the subscript p and denote the first fundamental form simply by g.

A bilinear map is completely determined by its action on basis vectors $\left\{\frac{\partial \mathsf{F}}{\partial u}, \frac{\partial \mathsf{F}}{\partial v}\right\}$. In other words, once we know these four values:

$$g\left(\frac{\partial \mathsf{F}}{\partial u}, \frac{\partial \mathsf{F}}{\partial u}\right) = \frac{\partial \mathsf{F}}{\partial u} \cdot \frac{\partial \mathsf{F}}{\partial u} \qquad g\left(\frac{\partial \mathsf{F}}{\partial u}, \frac{\partial \mathsf{F}}{\partial v}\right) = \frac{\partial \mathsf{F}}{\partial u} \cdot \frac{\partial \mathsf{F}}{\partial v}$$
$$g\left(\frac{\partial \mathsf{F}}{\partial v}, \frac{\partial \mathsf{F}}{\partial u}\right) = \frac{\partial \mathsf{F}}{\partial v} \cdot \frac{\partial \mathsf{F}}{\partial u} \qquad g\left(\frac{\partial \mathsf{F}}{\partial v}, \frac{\partial \mathsf{F}}{\partial v}\right) = \frac{\partial \mathsf{F}}{\partial v} \cdot \frac{\partial \mathsf{F}}{\partial v},$$

then we know how the bilinear map g acts on other tangent vectors in T_pM.

Example B.2. Let M be the unit sphere and F be the following smooth local parametrization:

$$\mathsf{F}(u,v) = (\sin u \cos v,\ \sin u \sin v,\ \cos u), \quad (u,v) \in (0,\pi) \times (0,2\pi)$$

By direct computations, we have:

$$\frac{\partial \mathsf{F}}{\partial u} = (\cos u \cos v,\ \cos u \sin v,\ -\sin u)$$
$$\frac{\partial \mathsf{F}}{\partial v} = (-\sin u \sin v,\ \sin u \cos v,\ 0)$$

$$g\left(\frac{\partial \mathsf{F}}{\partial u}, \frac{\partial \mathsf{F}}{\partial u}\right) = \frac{\partial \mathsf{F}}{\partial u} \cdot \frac{\partial \mathsf{F}}{\partial u} = 1 \qquad g\left(\frac{\partial \mathsf{F}}{\partial u}, \frac{\partial \mathsf{F}}{\partial v}\right) = \frac{\partial \mathsf{F}}{\partial u} \cdot \frac{\partial \mathsf{F}}{\partial v} = 0$$
$$g\left(\frac{\partial \mathsf{F}}{\partial v}, \frac{\partial \mathsf{F}}{\partial u}\right) = \frac{\partial \mathsf{F}}{\partial v} \cdot \frac{\partial \mathsf{F}}{\partial u} = 0 \qquad g\left(\frac{\partial \mathsf{F}}{\partial v}, \frac{\partial \mathsf{F}}{\partial v}\right) = \frac{\partial \mathsf{F}}{\partial v} \cdot \frac{\partial \mathsf{F}}{\partial v} = \sin^2 u$$

Therefore, the action of g on other tangent vectors in T_pM is given by:

$$\begin{aligned}
&g\left(\alpha\frac{\partial \mathsf{F}}{\partial u} + \beta\frac{\partial \mathsf{F}}{\partial v}, \gamma\frac{\partial \mathsf{F}}{\partial u} + \delta\frac{\partial \mathsf{F}}{\partial v}\right) \\
&= \left(\alpha\frac{\partial \mathsf{F}}{\partial u} + \beta\frac{\partial \mathsf{F}}{\partial v}\right) \cdot \left(\gamma\frac{\partial \mathsf{F}}{\partial u} + \delta\frac{\partial \mathsf{F}}{\partial v}\right) \\
&= \alpha\gamma + \beta\delta \sin^2 u.
\end{aligned}$$

□

B.1.2. Matrix Representation of g. One elegant way to represent the first fundamental form is by the matrix:

$$[g] := \begin{bmatrix} g\left(\frac{\partial \mathsf{F}}{\partial u}, \frac{\partial \mathsf{F}}{\partial u}\right) & g\left(\frac{\partial \mathsf{F}}{\partial u}, \frac{\partial \mathsf{F}}{\partial v}\right) \\ g\left(\frac{\partial \mathsf{F}}{\partial v}, \frac{\partial \mathsf{F}}{\partial u}\right) & g\left(\frac{\partial \mathsf{F}}{\partial v}, \frac{\partial \mathsf{F}}{\partial v}\right) \end{bmatrix}$$

It is a *symmetric* matrix since $g\left(\frac{\partial \mathsf{F}}{\partial u}, \frac{\partial \mathsf{F}}{\partial v}\right) = g\left(\frac{\partial \mathsf{F}}{\partial v}, \frac{\partial \mathsf{F}}{\partial u}\right)$ since the dot product in $\mathbb{R}^3$ is commutative.

By identifying (u,v) with (u_1,u_2), and so $\frac{\partial \mathsf{F}}{\partial u_1} = \frac{\partial \mathsf{F}}{\partial u}$ and $\frac{\partial \mathsf{F}}{\partial u_2} = \frac{\partial \mathsf{F}}{\partial v}$, we can further denote

$$g_{ij} := g\left(\frac{\partial \mathsf{F}}{\partial u_i}, \frac{\partial \mathsf{F}}{\partial u_j}\right) \quad \text{for } i = 1,2$$

so that the matrix $[g]$ can be further written as

$$[g] = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

Given two tangent vectors $\mathsf{Y} = \sum_{i=1}^{2} \alpha_i \frac{\partial \mathsf{F}}{\partial u_i}$ and $\mathsf{Z} = \sum_{j=1}^{2} \beta_j \frac{\partial \mathsf{F}}{\partial u_j}$ on T_pM, the value of $g(\mathsf{Y}, \mathsf{Z})$ is related to the matrix $[g]$ in the following way:

$$\begin{aligned} g(\mathsf{Y}, \mathsf{Z}) &= g\left(\sum_{i=1}^{2} \alpha_i \frac{\partial \mathsf{F}}{\partial u_i}, \sum_{j=1}^{2} \beta_j \frac{\partial \mathsf{F}}{\partial u_j}\right) \\ &= \sum_{i=1}^{2}\sum_{j=1}^{2} \alpha_i \beta_j g\left(\frac{\partial \mathsf{F}}{\partial u_i}, \frac{\partial \mathsf{F}}{\partial u_j}\right) \\ &= \sum_{i=1}^{2}\sum_{j=1}^{2} \alpha_i \beta_j g_{ij} \\ &= \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} [g] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \end{aligned}$$

Therefore, the matrix $[g]$ "encodes" the most crucial data of the first fundamental form g, in a sense that if you know all the components of the matrix $[g]$, you can write down $g(\mathsf{Y}, \mathsf{Z})$ for any pair of tangent vectors Y and Z on T_pM.

As computed in Example B.2, the matrix $[g]$ of the unit sphere (with respect the parametrization F used in the example) is given by:

$$[g] := \begin{bmatrix} 1 & 0 \\ 0 & \sin^2 u \end{bmatrix}$$

Evidently, it is a diagonal matrix. Try to think about the geometric significance of $[g]$ being diagonal!

B.1.3. Tensor Representation of g. The first fundamental form is sometimes (quite common in physics) represented using tensor notations. Again we identify (u, v) with (u_1, u_2). The first fundamental form g can be represented as:

$$\begin{aligned} g &= g_{11} du \otimes du + g_{12} du \otimes dv + g_{21} dv \otimes du + g_{22} dv \otimes dv \\ &= \sum_{i,j=1}^{2} g_{ij} du^i \otimes du^j \end{aligned}$$

As $g_{ij} = g_{ji}$, the tensor g is symmetric. As such the tensor notation $du^i \otimes du^j$ is often written simply as $du^i \, du^j$. For instance, the first fundamental form of the unit sphere in Example B.2 can be expressed as:

$$g = du^2 + \sin^2 u \, dv^2$$

where du^2 is interpreted as $(du)^2$ (not $d(u^2)$).

B.1.4. Geometric Significance of g. We will see in subsequent sections that g "encodes" crucial geometric information such as curvatures of the surface. There are also some familiar geometric quantities, such as length and area, which are related to the first fundamental form g.

Consider a curve γ on a regular surface M parametrized by $\mathsf{F}(u, v)$. Suppose the curve can be parametrized by $\mathsf{r}(t)$, $a < t < b$, then from calculus we know the arc-length of the curve is given by:

$$l(\gamma) = \int_a^b \left|\mathsf{r}'(t)\right| \, dt$$

In fact one can express this above in terms of g. The argument is as follows:

Since the curve γ is assumed to be on the surface M, on every point $\mathsf{r}(t)$ on the curve there is a corresponding coordinates $(u(t), v(t))$ on the uv-plane such that $\mathsf{F}(u(t), v(t)) = \mathsf{r}(t)$. Using the chain rule, we then have:

$$\mathsf{r}'(t) = \frac{\partial \mathsf{F}}{\partial u}\frac{du}{dt} + \frac{\partial \mathsf{F}}{\partial v}\frac{dv}{dt} = \sum_{i=1}^{2} u_i'(t)\frac{\partial \mathsf{F}}{\partial u_i}$$

Therefore, the tangent vector $\mathsf{r}'(t)$ of the curve γ is in the span of $\left\{\frac{\partial \mathsf{F}}{\partial u}, \frac{\partial \mathsf{F}}{\partial v}\right\}$.

Recall that $|\mathsf{r}'(t)| = \sqrt{\mathsf{r}'(t)\cdot \mathsf{r}'(t)}$ and that $\mathsf{r}'(t)$ lies on T_pM, we then have:

$$\left|\mathsf{r}'(t)\right| = \sqrt{g\left(\mathsf{r}'(t), \mathsf{r}'(t)\right)}$$

We can first express it in terms of the matrix components g_{ij}'s. We first recall that $\mathsf{r}'(t) = \sum_{i=1}^{2} u_i'(t)\frac{\partial \mathsf{F}}{\partial u_i}$, so

$$\begin{aligned} g\left(\mathsf{r}'(t), \mathsf{r}'(t)\right) &= g\left(\sum_{i=1}^{2} u_i'(t)\frac{\partial \mathsf{F}}{\partial u_i}, \sum_{j=1}^{2} u_j'(t)\frac{\partial \mathsf{F}}{\partial u_j}\right) \\ &= \sum_{i,j=1}^{2} u_i'(t)u_j'(t)\, g\left(\frac{\partial \mathsf{F}}{\partial u_i}, \frac{\partial \mathsf{F}}{\partial u_j}\right) \\ &= \sum_{i,j=1}^{2} u_i'(t)u_j'(t)\, g_{ij} \end{aligned} \tag{B.1}$$

where g_{ij}'s are evaluated at the point $\mathsf{r}(t)$. Therefore, the arc-length can be expressed in terms of the first fundamental form by:

$$l(\gamma) = \int_a^b \sqrt{g\left(\mathsf{r}'(t), \mathsf{r}'(t)\right)}\, dt = \int_a^b \sqrt{\sum_{i,j=1}^{2} u_i'(t)u_j'(t)\, g_{ij}}\, dt$$

Another familiar geometric quantity which is also related to g is the area of a surface. Given a regular surface M (almost everywhere) parametrized by $\mathsf{F}(u, v)$ with $(u, v) \in D \subset \mathbb{R}^2$ where D is a bounded domain, the area of this surface is given by:

$$A(M) = \iint_D \left|\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v}\right|\, dudv$$

It is also possible to express $\left|\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v}\right|$ in terms of the first fundamental form g. Let θ be the angle between the two vectors $\frac{\partial \mathsf{F}}{\partial u}$ and $\frac{\partial \mathsf{F}}{\partial v}$, then from elementary vector geometry, we have:

$$\begin{aligned} \left|\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v}\right|^2 &= \left|\frac{\partial \mathsf{F}}{\partial u}\right|^2 \left|\frac{\partial \mathsf{F}}{\partial v}\right|^2 \sin^2\theta \\ &= \left|\frac{\partial \mathsf{F}}{\partial u}\right|^2 \left|\frac{\partial \mathsf{F}}{\partial v}\right|^2 - \left|\frac{\partial \mathsf{F}}{\partial u}\right|^2 \left|\frac{\partial \mathsf{F}}{\partial v}\right|^2 \cos^2\theta \\ &= \left|\frac{\partial \mathsf{F}}{\partial u}\right|^2 \left|\frac{\partial \mathsf{F}}{\partial v}\right|^2 - \left(\frac{\partial \mathsf{F}}{\partial u}\cdot\frac{\partial \mathsf{F}}{\partial v}\right)^2 \\ &= \left(\frac{\partial \mathsf{F}}{\partial u}\cdot\frac{\partial \mathsf{F}}{\partial u}\right)\left(\frac{\partial \mathsf{F}}{\partial v}\cdot\frac{\partial \mathsf{F}}{\partial v}\right) - \left(\frac{\partial \mathsf{F}}{\partial u}\cdot\frac{\partial \mathsf{F}}{\partial v}\right)^2 \\ &= g_{11}g_{22} - (g_{12})^2 \\ &= \det[g]. \end{aligned}$$

Therefore,

$$A(M) = \iint_D \sqrt{\det[g]}\, dudv \tag{B.2}$$

Example B.3. Let Γ_f be the graph of a smooth function $f(u,v) : \mathcal{U} \to \mathbb{R}$ defined on an open subset $\mathcal{U}$ of the uv-plane, then Γ_f has a globally defined smooth parametrization:

$$\mathsf{F}(u,v) = (u,\, v,\, f(u,v)).$$

By straight-forward computations, we can get:

$$\begin{aligned}
\frac{\partial \mathsf{F}}{\partial u} &= \left(1,\, 0,\, \frac{\partial f}{\partial u}\right) & \frac{\partial \mathsf{F}}{\partial v} &= \left(0,\, 1,\, \frac{\partial f}{\partial v}\right) \\
g_{11} &= 1 + \frac{\partial f}{\partial u}\frac{\partial f}{\partial u} & g_{12} &= 0 + \frac{\partial f}{\partial u}\frac{\partial f}{\partial v} \\
g_{21} &= 0 + \frac{\partial f}{\partial v}\frac{\partial f}{\partial u} & g_{22} &= 1 + \frac{\partial f}{\partial v}\frac{\partial f}{\partial v}
\end{aligned}$$

Therefore, the matrix representation of g is given by:

$$[g] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \frac{\partial f}{\partial u}\frac{\partial f}{\partial u} & \frac{\partial f}{\partial u}\frac{\partial f}{\partial v} \\ \frac{\partial f}{\partial v}\frac{\partial f}{\partial u} & \frac{\partial f}{\partial v}\frac{\partial f}{\partial v} \end{bmatrix} = I_{2\times 2} + \begin{bmatrix} \frac{\partial f}{\partial u} \\ \frac{\partial f}{\partial v} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{bmatrix}$$

where $I_{2\times 2}$ is the 2×2 identity matrix. By further defining $\nabla f = \begin{bmatrix} \frac{\partial f}{\partial u} \\ \frac{\partial f}{\partial v} \end{bmatrix}$, we have:

$$[g] = I_{2\times 2} + (\nabla f)\,(\nabla f)^T.$$

By identifying $(u_1, u_2) = (u, v)$, we can write the first fundamental form in a more concise way:

$$g_{ij} = \delta_{ij} + \frac{\partial f}{\partial u_i}\frac{\partial f}{\partial u_j}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Using tensor notations, the first fundamental form can be written as:

$$\begin{aligned}
g &= \left(1 + \left(\frac{\partial f}{\partial u}\right)^2\right) du^2 + 2\frac{\partial f}{\partial u}\frac{\partial f}{\partial v} dudv + \left(1 + \left(\frac{\partial f}{\partial v}\right)^2\right) dv^2 \\
&= \sum_{i,j=1}^{2} \left(\delta_{ij} + \frac{\partial f}{\partial u_i}\frac{\partial f}{\partial u_i}\right) du_i du_j
\end{aligned}$$

Therefore, to compute the length the curve $\mathsf{r}(t) := \mathsf{F}(t, v_0)$, where v_0 is a constant and $t \in [a,b]$, on the surface Γ_f, we may first write down the uv-coordinates of this curve, namely:

$$u(t) = t,\ v(t) = v_0,$$

and so $u'(t) = 1$ and $v'(t) = 0$. Hence, according to (B.1), we have:

$$\begin{aligned}
g\left(\mathsf{r}'(t), \mathsf{r}'(t)\right) &= g_{11} u_1'(t)^2 + \underbrace{g_{12} u_1'(t) u_2'(t) + g_{21} u_2'(t) u_1'(t) + g_{22} u_2'(t)^2}_{=0 \text{ since } u_2'(t) = v'(t) = 0} \\
&= \left(1 + \left(\frac{\partial f}{\partial u}(t, v_0)\right)^2\right) \\
\text{Length of the curve} &= \int_a^b \sqrt{g\left(\mathsf{r}'(t), \mathsf{r}'(t)\right)}\, dt \\
&= \int_a^b \sqrt{1 + \left(\frac{\partial f}{\partial u}(t, v_0)\right)^2}\, dt
\end{aligned}$$

To compute the surface area of a region $\mathsf{F}(\Omega) \subset \Gamma_f$ where Ω is a bounded domain on the uv-plane, we first compute that:

$$\begin{aligned}
\det[g] &= g_{11} g_{22} - g_{12} g_{21} \\
&= \left(1 + \left(\frac{\partial f}{\partial u}\right)^2\right) \left(1 + \left(\frac{\partial f}{\partial v}\right)^2\right) - \left(\frac{\partial f}{\partial u} \frac{\partial f}{\partial v}\right)^2 \\
&= 1 + \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2
\end{aligned}$$

and according to (B.2), we have:

$$A(\mathsf{F}(\Omega)) = \iint_\Omega \sqrt{1 + \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2}\, du dv$$

which is exactly the same as what you have seen in multivariable calculus. □

B.2. Second Fundamental Form

In this section we introduce another important inner product on T_pM, the second fundamental form h. While the first fundamental form g encodes information about angle, length and area of a regular surface, the second fundamental form encodes information about various curvatures of a surface.

We will first present some preliminaries, discuss the motivation and define the second fundamental form. In the next section, we will see how the second fundamental form is related to the curvatures of a surface.

B.2.1. Gauss Map. We will see in subsequent sections that curvature of a regular surface is, roughly speaking, determined by the interaction between tangent and normal just like the case for regular curves. While the first fundamental form concerns only the tangent vectors, the second fundamental form involves both tangent and normal. Now let's talk about the normal vector – or in differential geometry jargon – the Gauss Map.

Given a regular surface M in $\mathbb{R}^3$ with $\mathsf{F}(u,v) : \mathcal{U} \subset \mathbb{R}^2 \to M$ as one of its smooth local parametrization, and let $p \in M$, the vectors: $\frac{\partial \mathsf{F}}{\partial u}(p)$ and $\frac{\partial \mathsf{F}}{\partial v}(p)$ are two linearly independent tangents of M at p. Therefore, their cross product $\frac{\partial \mathsf{F}}{\partial u}(p) \times \frac{\partial \mathsf{F}}{\partial v}(p)$ is a normal vector to M at p, and so a unit normal vector at p is given by:

$$\mathsf{N}(p) = \frac{\frac{\partial \mathsf{F}}{\partial u}(p) \times \frac{\partial \mathsf{F}}{\partial v}(p)}{\left|\frac{\partial \mathsf{F}}{\partial u}(p) \times \frac{\partial \mathsf{F}}{\partial v}(p)\right|}.$$

Naturally, the opposite vector

$$-\frac{\frac{\partial \mathsf{F}}{\partial u}(p) \times \frac{\partial \mathsf{F}}{\partial v}(p)}{\left|\frac{\partial \mathsf{F}}{\partial u}(p) \times \frac{\partial \mathsf{F}}{\partial v}(p)\right|}, \quad \text{or equivalently,} \quad \frac{\frac{\partial \mathsf{F}}{\partial v}(p) \times \frac{\partial \mathsf{F}}{\partial u}(p)}{\left|\frac{\partial \mathsf{F}}{\partial v}(p) \times \frac{\partial \mathsf{F}}{\partial u}(p)\right|}$$

is another unit normal vector at p.

Example B.4. Consider the unit sphere $\mathbb{S}^2(1)$ with smooth local parametrization:

$$\mathsf{F}(u,v) = (\sin u \cos v, \, \sin u \sin v, \, \cos u), \quad (u,v) \in (0,\pi) \times (0, 2\pi)$$

It is straight-forward to compute that:

$$\begin{aligned} \frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v} &= \left(\sin^2 u \cos v, \, \sin^2 u \sin v, \, \sin u \cos u\right) \\ \left|\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v}\right| &= \sin u \\ \mathsf{N}(u,v) &= (\sin u \cos v, \, \sin u \sin v, \, \cos u) = \mathsf{F}(u,v) \end{aligned}$$

This unit normal vector N is outward point. □

Given a smooth local parametrization F, there are always two choices of normal vector. For a sphere, once the normal vector direction is chosen, it is always *consistent* with your choice when we move the normal vector across the sphere. That is, when you draw a closed path on the sphere and see how the unit normal vector varies along the path, you will find that the unit normal remains the same when you come back to the original point. We call it an *orientable* surface.

Most surfaces we have encountered so far are orientable. A celebrated example of a non-orientable surface is the Möbius strip as discussed in Chapter 4.

When M is an orientable regular surface, the unit normal vector N can then be regarded as a map. The domain of N is M. Since N is unit, the codomain can be taken to be the *unit sphere* $\mathbb{S}^2$. Let's summarize our discussion by stating the definition of:

Definition B.5 (Gauss Map)**.** Suppose M is an orientable regular surface. The *Gauss map* of M is a smooth function $\mathsf{N} : M \to \mathbb{S}^2$ such that for any $p \in M$, the output $\mathsf{N}(p)$ is a unit normal vector of M at p. Here $\mathbb{S}^2$ is the unit sphere in $\mathbb{R}^3$:

$$\mathbb{S}^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}.$$

As computed in Example B.4, the Gauss map N for the unit sphere $\mathbb{S}^2(1)$ is given by F (assuming the outward-pointing convention is observed). It is not difficult to see that the Gauss map N for a sphere with radius R centered the origin in $\mathbb{R}^3$ is given by $\frac{1}{R}\mathsf{F}$. Readers should verify this as an exercise.

For a plane Π, the unit normal vector at each point is the same. Therefore, the Gauss map $\mathsf{N}(p)$ is a constant vector independent of p.

A unit cylinder with z-axis as its central axis can be parametrized by:

$$\mathsf{F}(u, v) = (\cos u,\ \sin u,\ v), \qquad (u, v) \in (0, 2\pi) \times \mathbb{R}.$$

By straight-forward computations, one can get:

$$\frac{\partial \mathsf{F}}{\partial u} \times \frac{\partial \mathsf{F}}{\partial v} = (\cos u,\ \sin u,\ 0)$$

which is already unit. Therefore, the Gauss map of the cylinder is given by:

$$\mathsf{N}(u, v) = (\cos u,\ \sin u,\ 0).$$

The image of N in $\mathbb{S}^2$ is the equator.

It is not difficult to see that the *image* of the Gauss map N, which is a subset of $\mathbb{S}^2$, is related to how "spherical" or "planar" the surface looks. The smaller the image, the more planar it is.

B.2.2. Normal Curvature. The curvature of a regular curve is a scalar function $\kappa(p)$. Since a curve is one dimensional, we can simply use one single value to measure the curvature at each point. However, a regular surface is two dimensional and has higher degree of freedom than curves, and hence may bend in a different way along different direction. As such, there are various notions of curvatures for regular surfaces. In this subsection, we talk about the normal curvature, which is fundamental to many other notions of curvatures.

Let M be an orientable regular surface with its Gauss map denoted by N. At each point $p \in M$, we pick a unit tangent vector T in T_pM. Heuristically, the normal curvature at p measures the curvature of the surface along a direction T. Precisely, we define:

Definition B.6 (Normal Curvature)**.** Let M be an orientable regular surface with Gauss map N. For each point $p \in M$, and any unit tangent vector $\mathsf{T} \in T_pM$, we let $\Pi(p, \mathsf{T})$ be the plane in $\mathbb{R}^3$ spanned by $\mathsf{N}(p)$ and T (see Figure B.1).

Let γ be the curve of intersection of M and $\Pi(p, \mathsf{T})$. The *normal curvature at p in the direction of* T of the surface M, denoted by $k_n(p, \mathsf{T})$, is defined to be the signed curvature $k(p)$ of the curve γ at p with normal vector taken to be the Gauss map N.

Remark B.7. We can talk about the signed curvature of γ because it is on the plane $\Pi(p, \mathsf{T})$. For curves in $\mathbb{R}^2$, the normal vector N is always taken to be $\mathsf{J}\mathsf{T}$, where J is the counter-clockwise rotation by $\frac{\pi}{2}$ in $\mathbb{R}^2$. However, as a plane in $\mathbb{R}^3$, $\Pi(p, \mathsf{T})$ does not have a natural notion of counter-clockwise rotation. Therefore, when computing the signed curvature $k(p)$ stated in the definition above, we always pick $\mathsf{N}(p)$ to be the chosen Gauss map.

Remark B.8. Since $k_n(p, \mathsf{T})$ is defined using the Gauss map N, which always comes with two possible choice for any orientable regular surface, the normal curvature *depends* on the choice of the Gauss map N. If the opposite unit normal is chosen to be the Gauss map, the normal curvature will differ by a sign.

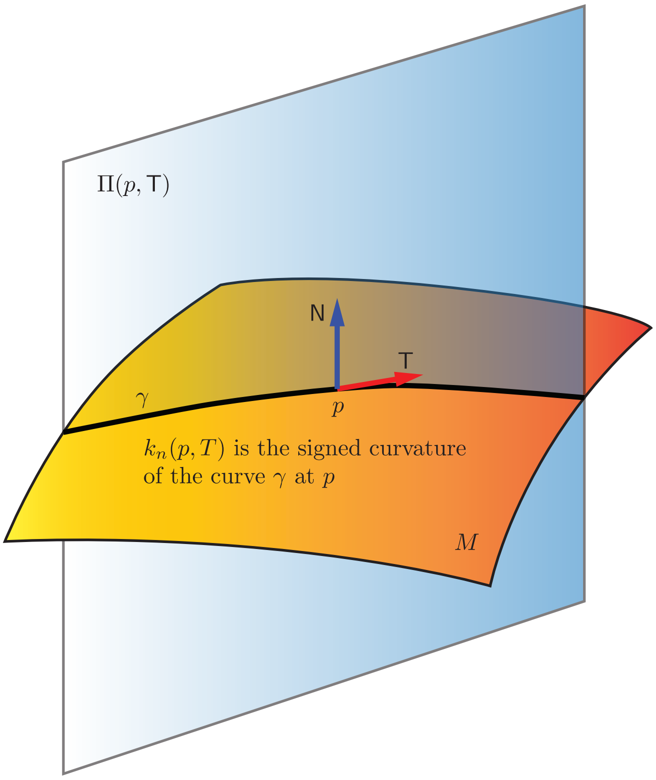


Figure B.1. normal curvature at p in a given direction T

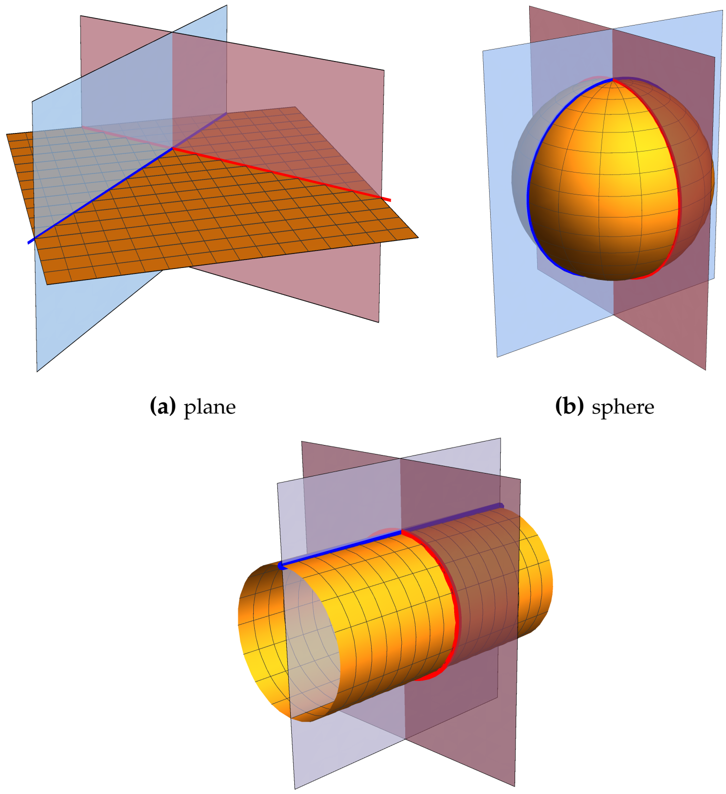
We will first make sense of normal curvatures through elementary examples, then we will prove a general formula for computing normal curvatures.

Example B.9. Let P be any plane in $\mathbb{R}^3$. For any point $p \in P$ and unit tangent $\mathsf{T} \in T_pP$, the plane $\Pi(p, \mathsf{T})$ must cut through P along a straight-line γ. Since γ has curvature 0, we have:

$$k_n(p, \mathsf{T}) = 0$$

for any $p \in P$ and $\mathsf{T} \in T_pP$. See Figure B.2a. □

Example B.10. Let $\mathbb{S}^2(R)$ be the sphere with radius R centered at the origin in $\mathbb{R}^3$ with Gauss map N taken to be inward-pointing. For any point $p \in \mathbb{S}^2$ and $\mathsf{T} \in T_p\mathbb{S}^2(R)$, the



(a) plane **(b)** sphere

(c) cylinder

Figure B.2. Normal curvatures of various surfaces

plane $\Pi(p, \mathsf{T})$ cuts $\mathbb{S}^2(R)$ along a great circle (with radius R). Since a circle with radius R has constant curvature $\frac{1}{R}$, we have:

$$k_n(p, \mathsf{T}) = \frac{1}{R}$$

for any $p \in \mathbb{S}^2(R)$ and $\mathsf{T} \in T_p\mathbb{S}^2(R)$. See Figure B.2b. $\square$

Example B.11. Let M be the (infinite) cylinder of radius R with x-axis as the central axis with outward-pointing Gauss map N. Given any $p \in M$, if T_x is the unit tangent vector at p parallel to the x-axis, then the $\Pi(p, \mathsf{T}_x)$ cuts the cylinder M along a straight-line. Therefore, we have:

$$k_n(p, \mathsf{T}_x) = 0$$

for any $p \in M$. See the blue curve in Figure B.2c.

On the other hand, if T_{yz} is a horizontal unit tangent vector at p, then $\Pi(p, \mathsf{T}_{yz})$ cuts M along a circle with radius R. Therefore, we have:

$$k_n(p, \mathsf{T}_{yz}) = -\frac{1}{R}$$

for any $p \in M$. See the red curve in Figure B.2c. Note that the tangent vector of the curve is moving away from the outward-pointing N. It explains the negative sign above.

For any other choice of unit tangent T at p, the plane $\Pi(p, \mathsf{T})$ cuts the cylinder along an ellipse, so the normal curvature along T may vary between 0 and $-\frac{1}{R}$. $\square$

In the above examples, the normal curvatures $k_n(p, \mathsf{T})$ are easy to find since the curve of intersection between $\Pi(p, \mathsf{T})$ and the surface M is either a straight line or a circle. Generally speaking, the curve of intersection may be of arbitrary shape such as an ellipse, and sometimes it is not even easy to identify what curve it is. Fortunately, it is possible to compute $k_n(p, \mathsf{T})$ for any given unit tangent $\mathsf{T} \in T_pM$ in a systematic way.

Under a smooth local parametrization $\mathsf{F}(u_1, u_2)$ of M, the tangent plane T_pM is spanned by $\left\{\frac{\partial \mathsf{F}}{\partial u_1}(p), \frac{\partial \mathsf{F}}{\partial u_2}(p)\right\}$. Then, any tangent vector $\mathsf{T} \in T_pM$ is a linear combination of $\left\{\frac{\partial \mathsf{F}}{\partial u_1}(p), \frac{\partial \mathsf{F}}{\partial u_2}(p)\right\}$. The following proposition proves a formula for computing the normal curvature along T in terms of local coordinates:

Proposition B.12. *Consider a smooth local parametrization $\mathsf{F}(u_1, u_2)$ of an orientable regular surface M with Gauss map N. Let $p \in M$, and the pick a unit tangent vector $\mathsf{T} \in T_pM$ given by:*

$$\mathsf{T} = a_1 \frac{\partial \mathsf{F}}{\partial u_1}(p) + a_2 \frac{\partial \mathsf{F}}{\partial u_2}(p) = \sum_{i=1}^{2} a_i \frac{\partial \mathsf{F}}{\partial u_i}(p).$$

Then, the normal curvature $k_n(p, \mathsf{T})$ at p along T is given by:

$$\begin{aligned} k_n(p, \mathsf{T}) &= \left(\frac{\partial^2 \mathsf{F}}{\partial u_1^2} \cdot \mathsf{N}\right) a_1^2 + 2\left(\frac{\partial^2 \mathsf{F}}{\partial u_1 \partial u_2} \cdot \mathsf{N}\right) a_1 a_2 + \left(\frac{\partial^2 \mathsf{F}}{\partial u_2^2} \cdot \mathsf{N}\right) a_2^2 \\ &= \sum_{i,j=1}^{2} \left(\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}(p) \cdot \mathsf{N}(p)\right) a_i a_j \end{aligned} \tag{B.3}$$

Proof. Recall that $k_n(p, \mathsf{T})$ is the signed curvature of the intersection curve γ between the plane $\Pi(p, \mathsf{T})$ and the surface M. Since γ is on M, one can arc-length parametrize γ by:

$$\mathsf{r}(s) = \mathsf{F}(u_1(s), u_2(s)), \quad s \in (-\varepsilon, \varepsilon). \tag{B.4}$$

We may assume $\mathsf{r}(0) = p$. Here $(u_1(s), u_2(s))$ is the pre-image of γ under the map F. It represents a curve on the (u_1, u_2)-plane. Note that although s is an arc-length parameter for $\mathsf{r}(s)$, i.e. $|\mathsf{r}'(s)| \equiv 1$, it may not be an arc-length parameter for the curve $(u_1(s), u_2(s))$.

Since the curve $\mathsf{r}(s)$ is a planar curve in $\Pi(p, \mathsf{T})$, its signed curvature at p (with respect to the Gauss map N) is given by:

$$k(p) = \mathsf{r}''(0) \cdot \mathsf{N}(p).$$

By applying the chain rule on (B.4), we get:

$$\begin{aligned} \mathsf{r}'(s) &= \frac{d}{ds} \mathsf{F}(u_1(s), u_2(s)) \\ &= \frac{\partial \mathsf{F}}{\partial u_1} \frac{du_1}{ds} + \frac{\partial \mathsf{F}}{\partial u_2} \frac{du_2}{ds} \\ &= \sum_{i=1}^{2} \frac{\partial \mathsf{F}}{\partial u_i} \frac{du_i}{ds}. \end{aligned}$$

Since the curve γ has a unit tangent T at p, we must have $\mathsf{r}'(0) = \mathsf{T}$. Therefore,

$$\underbrace{\sum_{i=1}^{2} \frac{\partial \mathsf{F}}{\partial u_i}(p) \frac{du_i}{ds}(0)}_{\mathsf{r}'(0)} = \underbrace{\sum_{i=1}^{2} a_i \frac{\partial \mathsf{F}}{\partial u_i}(p)}_{\mathsf{T}}.$$

By equating the coefficients, we get $\frac{du_i}{ds}(0) = a_i$ for $i = 1, 2$.

Then we apply the chain rule again for the second derivative:

$$\begin{aligned}
\mathsf{r}''(s) &= \sum_{i=1}^{2} \frac{d}{ds}\left(\frac{\partial \mathsf{F}}{\partial u_i}\frac{du_i}{ds}\right) \\
&= \sum_{i=1}^{2} \left\{\left(\frac{d}{ds}\frac{\partial \mathsf{F}}{\partial u_i}\right)\frac{du_i}{ds} + \frac{\partial \mathsf{F}}{\partial u_i}\frac{d^2u_i}{ds^2}\right\} \\
&= \sum_{i=1}^{2}\left(\frac{\partial^2 \mathsf{F}}{\partial u_1 \partial u_i}\frac{du_1}{ds} + \frac{\partial^2 \mathsf{F}}{\partial u_2 \partial u_i}\frac{du_2}{ds}\right)\frac{du_i}{ds} + \sum_{i=1}^{2}\frac{\partial \mathsf{F}}{\partial u_i}\frac{d^2u_i}{ds^2} \\
&= \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\partial^2 \mathsf{F}}{\partial u_j \partial u_i}\frac{du_j}{ds}\frac{du_i}{ds} + \sum_{i=1}^{2}\frac{\partial \mathsf{F}}{\partial u_i}\frac{d^2u_i}{ds^2}.
\end{aligned}$$

Therefore, using the fact that $u_i'(0) = a_i$, we get:

$$\mathsf{r}''(0) = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}\cdot a_i a_j + \sum_{i=1}^{2}\frac{\partial \mathsf{F}}{\partial u_i}\frac{d^2u_i}{ds^2}(0).$$

Note that the second term is tangential. Finally, we have:

$$\begin{aligned}
k(p) &= \mathsf{r}''(0)\cdot \mathsf{N}(p) \\
&= \left(\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}a_i a_j + \underbrace{\sum_{i=1}^{2}\frac{\partial \mathsf{F}}{\partial u_i}\frac{d^2u_i}{ds^2}(0)}_{\text{tangential}}\right)\cdot \mathsf{N}(p) \\
&= \left(\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}a_i a_j\right)\cdot \mathsf{N}(p)
\end{aligned}$$

as desired. □

B.2.3. Second Fundamental Form. When computing the normal curvature $k_n(p, \mathsf{T})$ where $\mathsf{T} = \sum_{i=1}^{2} a_i \frac{\partial \mathsf{F}}{\partial u_i}(p)$, we can make use of Proposition B.12. The dot products

$$\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}(p)\cdot \mathsf{N}(p), \quad i = 1, 2$$

are crucial quantities when applying (B.3). These quantities are determined by the smooth local parametrization F and are independent of the unit tangent T. Therefore, when we regard $k_n(p, \mathsf{T})$ as a function of T with p fixed, then $k_n(p, \mathsf{T})$ can be regarded as a function of (a_1, a_2), and the dot products $\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}(p)\cdot \mathsf{N}(p)$ can be considered as constants (when p is fixed). Due to the fundamental importance of these quantities when it comes to computing curvatures, they are coined with a name:

Definition B.13 (Second Fundamental Form)**.** Given an orientable regular surface M with Gauss map N, then for each $p \in M$ with smooth local parametrization $\mathsf{F}(u_1, u_2)$, we define:

$$h_{ij}(p) := \frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}(p) \cdot \mathsf{N}(p).$$

The *second fundamental form* is the bilinear map $h_p : T_pM \times T_pM \to \mathbb{R}$ at each $p \in M$ defined by:

$$h_p\left(\frac{\partial \mathsf{F}}{\partial u_i}, \frac{\partial \mathsf{F}}{\partial u_j}\right) := h_{ij}(p) \quad \text{for } 1 \le i, j \le 2.$$

When the point p is clear from the context, one can omit the subscript p and simply write h and h_{ij}.

Remark B.14. Now (B.3) in Proposition B.12 can be restated as:

$$k_n(p, \mathsf{T}) = \sum_{i,j=1}^{2} h_{ij}(p) \cdot a_i a_j = h_p(\mathsf{T}, \mathsf{T}) \quad \text{where } T = \sum_{i=1}^{2} a_i \frac{\partial \mathsf{F}}{\partial u_i}(p).$$

□

Example B.15. Let Γ_f be the graph of a smooth function $f(u_1, u_2) : \mathcal{U} \to \mathbb{R}$ defined on an open subset $\mathcal{U}$ of $\mathbb{R}^2$, then Γ_f has a globally defined smooth parametrization:

$$\mathsf{F}(u_1, u_2) = (u_1,\, u_2,\, f(u_1, u_2)).$$

By straight-forward computations, we can get:

$$\frac{\partial \mathsf{F}}{\partial u_1} = \left(1,\, 0,\, \frac{\partial f}{\partial u_1}\right) \qquad \frac{\partial \mathsf{F}}{\partial u_2} = \left(0,\, 1,\, \frac{\partial f}{\partial u_2}\right)$$

$$\frac{\partial^2 \mathsf{F}}{\partial u_1^2} = \left(0,\, 0,\, \frac{\partial^2 f}{\partial u_1^2}\right) \qquad \frac{\partial^2 \mathsf{F}}{\partial u_1 \partial u_2} = \left(0,\, 0,\, \frac{\partial^2 f}{\partial u_1 \partial u_2}\right)$$

$$\frac{\partial^2 \mathsf{F}}{\partial u_2 \partial u_1} = \left(0,\, 0,\, \frac{\partial^2 f}{\partial u_2 \partial u_1}\right) \qquad \frac{\partial^2 \mathsf{F}}{\partial u_2^2} = \left(0,\, 0,\, \frac{\partial^2 f}{\partial u_2^2}\right)$$

In short, we have

$$\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j} = \left(0,\, 0,\, \frac{\partial^2 f}{\partial u_i \partial u_j}\right).$$

Let's take the Gauss map N to be:

$$\mathsf{N} = \frac{\frac{\partial \mathsf{F}}{\partial u_1} \times \frac{\partial \mathsf{F}}{\partial u_2}}{\left|\frac{\partial \mathsf{F}}{\partial u_1} \times \frac{\partial \mathsf{F}}{\partial u_2}\right|} = \frac{\left(-\frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_2}, 1\right)}{\sqrt{1 + \left(\frac{\partial f}{\partial u_1}\right)^2 + \left(\frac{\partial f}{\partial u_2}\right)^2}}.$$

Then, the second fundamental form is given by:

$$h_{ij} = \frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j} \cdot \mathsf{N} = \left(0,\, 0,\, \frac{\partial^2 f}{\partial u_i \partial u_j}\right) \cdot \frac{\left(-\frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_2}, 1\right)}{\sqrt{1 + \left(\frac{\partial f}{\partial u_1}\right)^2 + \left(\frac{\partial f}{\partial u_2}\right)^2}} = \frac{\frac{\partial^2 f}{\partial u_i \partial u_j}}{\sqrt{1 + |\nabla f|^2}}.$$

The matrix whose (i, j)-th entry given by $\frac{\partial^2 f}{\partial u_i \partial u_j}$ is commonly called the *Hessian* of f, denoted by $\nabla\nabla f$ or $\mathrm{Hess}(f)$. Using this notation, the matrix of second fundamental form of Γ_f is given by:

$$[h] = \frac{\text{Hess}(f)}{\sqrt{1+|\nabla f|^2}}.$$

$\square$

B.2.4. Principal Curvatures. Given an orientable regular surface M with Gauss map N and again pick a fixed $p \in M$, the normal curvature $k_n(p, \mathsf{T})$ measures the how curved the surface is at p in the direction T. As T is unit, we can regard $k_n(p, \cdot)$ as a function defined on the unit circle $\mathbb{S}^1(1)$. Since $\mathbb{S}^1(1)$ is closed and bounded, Extreme Value Theorem asserts that $k_n(p, \mathsf{T})$ must achieve its maximum and minimum at certain unit tangents T. In this subsection, we are going to investigate the maximum and minimum normal curvatures $k_n(p, \mathsf{T})$.

Definition B.16 (Principal Curvatures)**.** Given an orientable regular surface M with Gauss map N, then for each $p \in M$ we define the *principal curvatures* at p to be:

$$k_1(p) := \min\{k_n(p, \mathsf{T}) : \mathsf{T} \in T_pM \text{ and } |\mathsf{T}| = 1\};$$
$$k_2(p) := \max\{k_n(p, \mathsf{T}) : \mathsf{T} \in T_pM \text{ and } |\mathsf{T}| = 1\}.$$

Furthermore, the unit tangents $\mathsf{T} \in T_pM$ which realize the principal curvatures are called *principal directions*.

Recall that when given a smooth local parametrization $\mathsf{F}(u_1, u_2)$, the normal curvatures are given by $k_n(p, \mathsf{T}) = \sum_{i,j=1}^2 h_{ij}a_ia_j$ where $\mathsf{T} = \sum_{i=1}^2 a_i \frac{\partial \mathsf{F}}{\partial u_i}$. Moreover, the tangent T is unit and so we have:

$$1 = |\mathsf{T}|^2 = g(\mathsf{T}, \mathsf{T}) = g\left(\sum_{i=1}^2 a_i \frac{\partial \mathsf{F}}{\partial u_i}, \sum_{j=1}^2 a_j \frac{\partial \mathsf{F}}{\partial u_j}\right) = \sum_{i,j=1}^2 a_i a_j g\left(\frac{\partial \mathsf{F}}{\partial u_i}, \frac{\partial \mathsf{F}}{\partial u_j}\right) = \sum_{i,j=1}^2 g_{ij}a_ia_j.$$

Therefore, the problem of maximizing and minimizing $k_n(p, \mathsf{T})$ with a fixed p can be reformulated as:

maximize and minimize: $\sum_{i,j=1}^2 h_{ij}(p)a_ia_j$

subject to constraint: $\sum_{i,j=1}^2 g_{ij}(p)a_ia_j = 1$

The unknown variables in this optimization problem are (a_1, a_2). To solve this optimization problem, one can use the Lagrange's Multiplier from MATH 2023, which will result in the following:

Proposition B.17. *Let M be an orientable regular surface with Gauss map N. At a point $p \in M$ with smooth local parametrization $\mathsf{F}(u_1, u_2)$, if a unit vector $\mathsf{T} = \sum_{i=1}^2 a_i \frac{\partial \mathsf{F}}{\partial u_i} \in T_pM$ is a principal direction at p, then the vector $[a_1 \ \ a_2]^T$ is an eigenvector of the matrix $[g]^{-1}[h]$ with eigenvalue exactly equal to $k_n(p, \mathsf{T})$.*

Proof. The unit tangent vector $\mathsf{T} = \sum_{i=1}^2 a_i \frac{\partial \mathsf{F}}{\partial u_i}$ is one that maximizes or minimizes $\sum_{i,j=1}^2 h_{ij}(p)a_ia_j$ subject to constraint $\sum_{i,j=1}^2 g_{ij}(p)a_ia_j = 1$. The method of Lagrange's

Multiplier asserts that the pair (a_1, a_2) is a solution to the following system:

$$\text{(*)} \qquad \frac{\partial}{\partial a_1}\sum_{i,j=1}^{2} h_{ij}(p)a_i a_j = \lambda\frac{\partial}{\partial a_1}\sum_{i,j=1}^{2} g_{ij}(p)a_i a_j$$

$$\text{(**)} \qquad \frac{\partial}{\partial a_2}\sum_{i,j=1}^{2} h_{ij}(p)a_i a_j = \lambda\frac{\partial}{\partial a_2}\sum_{i,j=1}^{2} g_{ij}(p)a_i a_j$$

$$\text{(***)} \qquad \sum_{i,j=1}^{2} g_{ij}(p)a_i a_j = 1$$

Note that both $g_{ij}(p)$ and $h_{ij}(p)$ are independent of the choice of unit tangents T, and that:

$$\frac{\partial}{\partial a_k}a_i = \delta_{ik} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases}.$$

Therefore, (*) in the above system can be simplified as:

$$\begin{aligned}
\sum_{i,j=1}^{2} h_{ij}(p)(\delta_{1i}a_j + a_i\delta_{1j}) &= \lambda\sum_{i,j=1}^{2} g_{ij}(p)(\delta_{1i}a_j + a_i\delta_{1j}) \\
\implies \underbrace{\sum_{j=1}^{2} h_{1j}(p)a_j + \sum_{i=1}^{2} h_{i1}(p)a_i}_{\text{same terms!}} &= \lambda\left(\sum_{j=1}^{2} g_{1j}(p)a_j + \sum_{i=1}^{2} g_{i1}(p)a_i\right) \\
\implies \quad 2\sum_{j=1}^{2} h_{1j}(p)a_j &= 2\lambda\sum_{j=1}^{2} g_{1j}(p)a_j \\
\implies \quad h_{11}a_1 + h_{12}a_2 &= \lambda(g_{11}a_1 + g_{12}a_2)
\end{aligned}$$

Similarly, (**) can be simplified as:

$$\begin{aligned}
\sum_{j=1}^{2} h_{2j}(p)a_j &= \lambda\sum_{j=1}^{2} g_{2j}(p)a_j \\
\implies h_{21}a_1 + h_{22}a_2 &= \lambda(g_{21}a_1 + g_{22}a_2)
\end{aligned}$$

Note that one can rewrite the system

$$\begin{aligned}
h_{11}a_1 + h_{12}a_2 &= \lambda(g_{11}a_1 + g_{12}a_2) \\
h_{21}a_1 + h_{22}a_2 &= \lambda(g_{21}a_1 + g_{22}a_2)
\end{aligned}$$

in matrix form:

$$\underbrace{\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}}_{[h]}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \lambda\underbrace{\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}}_{[g]}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

As a regular surface, $[g]$ must be invertible and so we have:

$$[g]^{-1}[h]\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \lambda\begin{bmatrix} a_1 \\ a_2 \end{bmatrix},$$

which shows $[a_1 \;\; a_2]^T$ is an eigenvector of $[g]^{-1}[h]$ with certain eigenvalue λ.

We are left to show that the eigenvalue λ is exactly the normal curvature at p along T. Recall that $\sum_{i,j=1}^{2} g_{ij}a_i a_j = 1$, or in other words:

$$[a_1 \;\; a_2]\,[g]\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = 1.$$

Since we have
$$[h]\begin{bmatrix}a_1\\a_2\end{bmatrix}=\lambda[g]\begin{bmatrix}a_1\\a_2\end{bmatrix},$$
we get
$$[a_1\ \ a_2]\,[h]\begin{bmatrix}a_1\\a_2\end{bmatrix}=\lambda[a_1\ \ a_2]\,[g]\begin{bmatrix}a_1\\a_2\end{bmatrix}=\lambda$$
and it implies
$$k_n(p,\mathsf{T})=\sum_{i,j=1}^{2}h_{ij}(p)a_ia_j=[a_1\ \ a_2]\,[h]\begin{bmatrix}a_1\\a_2\end{bmatrix}=\lambda.$$
It completes the proof. □

B.2.5. Shape Operator. Proposition B.17 asserts that the principal curvatures and directions are the eigenvalues and eigenvectors of the matrix $[g]^{-1}[h]$ respectively. Since the matrix $[g]^{-1}[h]$ encodes the principal curvatures of the regular surface, we are going to study this matrix in more detail in this subsection.

Given a smooth local parametrization $\mathsf{F}(u_1,u_2)$ of an orientable regular surface M, g_{ij} is the (i,j)-th entry of the first fundamental form matrix $[g]$. From now on, we denote g^{ij} to be the (i,j)-th component of the inverse $[g]^{-1}$.

Given two square matrices A and B with their (i,j)-th entries denoted by A_{ij} and B_{ij} respectively, then the (i,j)-th entry of the matrix product AB is given by:
$$[AB]_{ij}=\sum_k A_{ik}B_{kj}.$$
Therefore, the (i,j)-th entry of the matrix $[g]^{-1}[h]$ is given by:
$$\sum_{k=1}^{2}g^{ik}h_{kj}.$$
Regarding $[g]^{-1}[h]$ as a linear map on T_pM, we define the following important operator in Differential Geometry:

Definition B.18 (Shape Operator)**.** Let M be an orientable regular surface with Gauss map N, and $\mathsf{F}(u_1,u_2)$ be a smooth local parametrization of M. The *shape operator* is a linear map $S_p : T_pM \to T_pM$ at each $p \in M$ defined by:
$$S_p\left(\frac{\partial\mathsf{F}}{\partial u_j}(p)\right):=\sum_{i=1}^{2}\left(\sum_{k=1}^{2}g^{ik}(p)h_{kj}(p)\right)\frac{\partial\mathsf{F}}{\partial u_i}(p).$$
In other words, the matrix representation of S with respect to the basis $\left\{\frac{\partial\mathsf{F}}{\partial u_1},\frac{\partial\mathsf{F}}{\partial u_2}\right\}$ is given by $[g]^{-1}[h]$.

If the point p is clear from the context, we may omit the subscript p when writing S_p. We may write the (i,j)-th component of the matrix representation $[S]$ as S^i_j. The lower index j refers to the column number, and the upper index i refers to the row number. We use this convention partly because $S^i_j=\sum_k g^{ik}h_{kj}$ and such a convention will preserve the positions of i and j on both sides. According to Proposition B.17, the eigenvalues of $[S^i_j]$ are the principal curvatures of M at p.

B.3. Curvatures

B.3.1. Gauss Curvature and Mean Curvature. Let's first summarize the definitions and results presented in the previous section. We first define the normal curvature $k_n(p, \mathsf{T})$ which measures the curvedness of the surface in a given direction T. Under a smooth local parametrization $\mathsf{F}(u_1, u_2)$, we showed in Proposition B.12 that $k_n(p, \mathsf{T}) = \sum_{i,j} h_{ij} a_i a_j$ where $\mathsf{T} = \sum_i a_i \frac{\partial \mathsf{F}}{\partial u_i}$. Then, we define principal curvatures $k_1(p)$ and $k_2(p)$ as the minimum and maximum possible normal curvatures, and using Lagrange's Multiplier we showed in Proposition B.17 that they are the eigenvalues of the matrix $[g]^{-1}[h]$. Due to this connection, we define the shape operator S, whose matrix representation is given by $[g]^{-1}[h]$.

In this section, we discuss two important types of curvatures in Differential Geometry, namely the Guass curvature and mean curvature. They are curvatures which are generated by the principal curvatures k_1 and k_2.

The Gauss curvature is defined to be the *product* $k_1 k_2$ of the two principal curvatures. Since k_i's are eigenvalues of the matrix representation $[g]^{-1}[h]$ of the shape operator, from Linear Algebra the product $k_1 k_2$ is given by the determinant of $[g]^{-1}[h]$. Since

$$\det[g]^{-1}[h] = \det[g]^{-1} \det[h] = \frac{\det[h]}{\det[g]},$$

the Gauss curvature of a regular surface can be defined as follows:

Definition B.19 (Gauss Curvature)**.** Given an orientable regular surface M with Gauss map N, and a point $p \in M$ with smooth local parametrization $\mathsf{F}(u_1, u_2)$, then the *Gauss curvature* $K(p)$ at the point p is equivalently defined as:

$$K(p) := k_1(p)k_2(p) = \det[S_p] = \frac{\det[h]}{\det[g]}(p) = \frac{h_{11}h_{22} - h_{12}^2}{g_{11}g_{22} - g_{12}^2}(p).$$

Although the Gauss curvature is defined using both first and second fundamental forms, we will later on show that it depends *only* on the first fundamental form. This is a remarkable and surprising result, proved by Gauss, and is commonly called *Theorema Egregium* (in Latin).

The mean curvature, as the name suggests, is the average $\frac{1}{2}(k_1 + k_2)$ of principal curvatures[1]. Since k_i's are eigenvalues of $[g]^{-1}[h]$, the sum $k_1 + k_2$ is in fact the *trace* of $[g]^{-1}[h]$. Let's give the precise definition:

Definition B.20 (Mean Curvature)**.** Given an orientable regular surface M with Gauss map N, and a point $p \in M$ with smooth local parametrization $\mathsf{F}(u_1, u_2)$, the *mean curvature* $H(p)$ at the point p is equivalently defined as:

$$H(p) := \frac{k_1(p) + k_2(p)}{2} = \frac{1}{2}\mathrm{Tr}[S_p] = \frac{1}{2}\mathrm{Tr}\left([g_p]^{-1}[h_p]\right) = \frac{1}{2}\sum_{i,j=1}^{2} g^{ij}(p)h_{ji}(p).$$

The mean curvature is important in Differential Geometry since it is related to surfaces that minimize area.

B.3.2. Invariance under Rigid-Body Motion. A rigid-body motion in $\mathbb{R}^3$ is a map that preserves distance between any two points in $\mathbb{R}^3$. Precisely,

[1]Some textbooks define the mean curvature as the sum $k_1 + k_2$ of principal curvatures.

Definition B.21 (Rigid-Body Motion)**.** A map $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ is said to be a *rigid body motion* if $\Phi(\mathsf{r}) = A\mathsf{r} + \mathsf{r}_0$ for any $\mathsf{r} \in \mathbb{R}^3$, where A is a 3×3-matrix satisfying $A^T A = I$ and r_0 is a fixed point in $\mathbb{R}^3$.

It is expected that principal curvatures, and hence Gauss and mean curvatures, are invariant under any rigid-body motion (up to signs). The following proposition shows that it is indeed the case:

Proposition B.22. *Let M be an orientable regular surface with Gauss map N, and $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ be a rigid-body motion. Denote $\widetilde{k}_n$, $\widetilde{k}_i$, $\widetilde{K}$ and $\widetilde{H}$ to be the normal, principal, Gauss and mean curvatures of the surface $\widetilde{M} := \Phi(M)$, then we have for any $p \in M$ and $i, j \in \{1, 2\}$:*

$$\widetilde{g}_{ij}(\Phi(p)) = g_{ij}(p) \quad \textit{and} \quad \widetilde{h}_{ij}(\Phi(p)) = \pm h_{ij}(p),$$

and hence for any $p \in M$ and unit tangent $\mathsf{T} \in T_pM$:

$$\widetilde{k}_n(\Phi(p), \Phi(\mathsf{T})) = \pm k_n(p, \mathsf{T}) \qquad \widetilde{k}_i(\Phi(p)) = \pm k_i(p)$$
$$\widetilde{K}(\Phi(p)) = K(p) \qquad \widetilde{H}(\Phi(p)) = \pm H(p)$$

Proof. Let $\mathsf{F}(u_1, u_2)$ be a smooth local parametrization of M, then $\Phi \circ \mathsf{F}$ is a smooth local parametrization of $\widetilde{M}$. As a rigid-body motion in $\mathbb{R}^3$, Φ can be expressed as $\Phi(\mathsf{r}) = A\mathsf{r} + \mathsf{r}_0$ for some $\mathsf{r}_0 \in \mathbb{R}^3$ and some 3×3 matrix A satisfying $A^T A = I$. We will use the fact that $A^T A = I$ to prove that F and $\Phi \circ \mathsf{F}$ have the same first and second fundamental forms at each pair of corresponding points p and $\Phi(p)$.

Under the smooth local parametrization $\Phi \circ \mathsf{F}$ for $\widetilde{M}$, the induced basis for the tangent plane at is given by $\left\{ \dfrac{\partial}{\partial u_i}(\Phi \circ \mathsf{F}) \right\}_{u=1,2}$. Denote A_{ij} to be the (i, j)-th entry of A which are constants, we can show for any $i, j = 1, 2$, we have

$$\begin{aligned}
\frac{\partial(\Phi \circ \mathsf{F})}{\partial u_i} &= \frac{\partial}{\partial u_i}(A\mathsf{F} + \mathsf{r}_0) = A\frac{\partial \mathsf{F}}{\partial u_i} \\
\frac{\partial(\Phi \circ \mathsf{F})}{\partial u_i} \cdot \frac{\partial(\Phi \circ \mathsf{F})}{\partial u_j} &= A\frac{\partial \mathsf{F}}{\partial u_i} \cdot A\frac{\partial \mathsf{F}}{\partial u_j} = \left(A\frac{\partial \mathsf{F}}{\partial u_i}\right)^T \left(A\frac{\partial \mathsf{F}}{\partial u_j}\right) \\
&= \left(\frac{\partial \mathsf{F}}{\partial u_i}\right)^T A^T A \frac{\partial \mathsf{F}}{\partial u_j} = \left(\frac{\partial \mathsf{F}}{\partial u_i}\right)^T \frac{\partial \mathsf{F}}{\partial u_j} \\
&= \frac{\partial \mathsf{F}}{\partial u_i} \cdot \frac{\partial \mathsf{F}}{\partial u_j}
\end{aligned}$$

Therefore, we get:

$$\widetilde{g}_{ij} = g_{ij} \quad \text{for any } i, j = 1, 2.$$

Next, we claim that the Gauss map $\widetilde{\mathsf{N}}$ for $\widetilde{M}$ is related to that of M at the corresponding point by the relation:

$$\widetilde{\mathsf{N}}(\Phi(p)) = \pm A\mathsf{N}(p).$$

To prove this, we first show that $A\mathsf{N}$ is orthogonal to the basis $\left\{ \dfrac{\partial}{\partial u_i}(\Phi \circ \mathsf{F}) \right\}_{u=1,2}$ of the tangent $T_{\Phi(p)}\widetilde{M}$:

$$\frac{\partial}{\partial u_i}(\Phi \circ \mathsf{F}) \cdot A\mathsf{N} = A\frac{\partial \mathsf{F}}{\partial u_i} \cdot A\mathsf{N} = \left(A\frac{\partial \mathsf{F}}{\partial u_i}\right)^T (A\mathsf{N}) = \left(\frac{\partial \mathsf{F}}{\partial u_i}\right)^T A^T A\mathsf{N} = \left(\frac{\partial \mathsf{F}}{\partial u_i}\right)^T \mathsf{N}.$$

Here we use again the fact that $A^T A = I$. Since

$$\left(\frac{\partial \mathsf{F}}{\partial u_i}\right)^T \mathsf{N} = \underbrace{\frac{\partial \mathsf{F}}{\partial u_i}}_{\text{tangent}} \cdot \underbrace{\mathsf{N}}_{\text{normal}} = 0,$$

we have proven that $\frac{\partial}{\partial u_i}(\Phi \circ \mathsf{F})$ is orthogonal to $A\mathsf{N}$ for each $i = 1, 2$. In other words, $A\mathsf{N}$ is orthogonal to the tangent plane $T_{\Phi(p)}\widetilde{M}$, and hence must be a normal vector at $\Phi(p)$. Furthermore, $A\mathsf{N}$ is unit since

$$|A\mathsf{N}|^2 = A\mathsf{N} \cdot A\mathsf{N} = (A\mathsf{N})^T(A\mathsf{N}) = \mathsf{N}^T A^T A\mathsf{N} = \mathsf{N}^T\mathsf{N} = |\mathsf{N}|^2 = 1.$$

Therefore, either $A\mathsf{N}$ or $-A\mathsf{N}$ can be taken to be the Gauss map $\widetilde{N}$ for $\widetilde{M}$.

Using the fact that $\frac{\partial}{\partial u_i}(\Phi \circ \mathsf{F}) = A\frac{\partial \mathsf{F}}{\partial u_i}$, $\widetilde{N} = \pm A\mathsf{N}$ (at corresponding points) and $A^T A = I$, one can show:

$$\widetilde{h}_{ij} = \frac{\partial^2(\Phi \circ \mathsf{F})}{\partial u_i \partial u_j} \cdot \widetilde{N} = \pm\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j} \cdot \mathsf{N} = \pm h_{ij}$$

for any $i, j = 1, 2$ at corresponding points. This is left as an exercise for readers.

Since all of the quantities k_n, k_i, K and H are uniquely determined by the first and second fundamental forms, it follows easily that they are the same (up to signs) for both $\widetilde{M}$ and M at each pair of corresponding points. □

B.3.3. Curvatures of Graphs. In Examples B.3 and B.15, we computed the first and second fundamental forms of the graph Γ_f of a function f. Using these, it is not difficult to compute various curvatures of the graph. In this subsection, we are going to discuss the geometric meaning of each curvature in this context, especially at the point where the tangent plane is horizontal.

Proposition B.23. *Let Γ_f be the graph of a function $f(u_1, u_2)$. Suppose p is a point on Γ_f such that the tangent plane $T_p\Gamma_f$ is horizontal, i.e. p is a critical point of f. Suppose the Gauss map N is taken to be upward-pointing, then*

- *$K(p) > 0$ and $H(p) > 0 \implies p$ is a local minimum of f*
- *$K(p) > 0$ and $H(p) < 0 \implies p$ is a local maximum of f*
- *$K(p) < 0 \implies p$ is a saddle of f*

Proof. At a critical p of f, we have $\nabla f(p) = 0$. From Examples B.3 and B.15, we have computed:

$$g_{ij}(p) = \delta_{ij} + \underbrace{\frac{\partial f}{\partial u_i}(p)\frac{\partial f}{\partial u_j}(p)}_{=0} = \delta_{ij}$$

$$h_{ij}(p) = \frac{\frac{\partial^2 f}{\partial u_i \partial u_j}(p)}{\sqrt{1 + |\nabla f(p)|^2}} = \frac{\partial^2 f}{\partial u_i \partial u_j}(p)$$

Note that the Gauss map N was taken to be upward-pointing in Example B.15, as required in this proposition.

Therefore, we have:

$$K(p) = \frac{\det[h]}{\det[g]}(p) = \det\left[\frac{\partial^2 f}{\partial u_i \partial u_j}(p)\right] = \left(f_{11}f_{22} - f_{12}^2\right)(p)$$

$$H(p) = \frac{1}{2}\sum_{i,j=1}^{2} g^{ij}(p)h_{ij}(p) = \frac{1}{2}\sum_{i,j=1}^{2} \delta^{ij}\frac{\partial^2 f}{\partial u_i \partial u_j}(p) = \frac{1}{2}\left(f_{11} + f_{22}\right)(p)$$

From the second derivative test in multivariable calculus, given a critical point p, if $f_{11}f_{22} - f_{12}^2 > 0$ and $f_{11} + f_{22} > 0$ at p, then p is a local minimum of f. The other cases can be proved similarly using the second derivative test. □

Given any regular surface M (not necessarily the graph of a function) and any point $p \in M$, one can apply a rigid-motion motion $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ so that T_pM is transformed into a horizontal plane. Then, the new surface $\Phi(M)$ will becomes locally a graph of a function f near the point $\Phi(p)$. Recall that the Gauss curvatures of p and $\Phi(p)$ are the same as given by Proposition B.22. If $K(p) > 0$ (and hence $K(\Phi(p)) > 0$), then Proposition B.23 asserts that $\Phi(p)$ is a local maximum or minimum of the function f and so the surface $\Phi(M)$ is locally above or below the tangent plane at $\Phi(p)$. In other words, near p the surface M is locally on one side of the the tangent plane T_pM. On the other hand, if $K(p) < 0$ then no matter how close to p the surface M would intersect T_pM at points other than p.

B.3.4. Surfaces of Revolution. Surfaces of revolution are surfaces obtained by revolving a plane curve about a central axis. They are important class of surfaces, examples of which include spheres, torus, and many others. In this subsection, we will study the fundamental forms and curvatures of these surfaces.

For simplicity, we assume that the z-axis is the central axis. A surface of revolution (about the z-axis) is defined as follows.

Definition B.24 (Surfaces of Revolution). Consider the curve $\gamma(t) = (x(t), 0, z(t))$, where $t \in (a,b)$, on the xz-plane such that $x(t) \geq 0$ for any $t \in (a,b)$. The surface of revolution generated by γ is obtained by revolving γ about the z-axis, and it can be parametrized by:

$$\mathsf{F}(t,\theta) = (x(t)\cos\theta,\, x(t)\sin\theta,\, z(t)), \quad (t,\theta) \in (a,b) \times [0, 2\pi].$$

It is a straight-forward computation to verify that:

$$\frac{\partial \mathsf{F}}{\partial t} = \left(x'(t)\cos\theta,\, x'(t)\sin\theta,\, z'(t)\right) \tag{B.5}$$

$$\frac{\partial \mathsf{F}}{\partial \theta} = \left(-x(t)\sin\theta,\, x(t)\cos\theta,\, 0\right) \tag{B.6}$$

$$\frac{\partial \mathsf{F}}{\partial t} \times \frac{\partial \mathsf{F}}{\partial \theta} = \left(-x(t)\,z'(t)\cos\theta,\; -x(t)\,z'(t)\sin\theta,\; x(t)\,x'(t)\right) \tag{B.7}$$

Exercise B.1. Verify (B.5)-(B.7) and show that:

$$\left|\frac{\partial \mathsf{F}}{\partial t} \times \frac{\partial \mathsf{F}}{\partial \theta}\right| = \left|x(t)\,\gamma'(t)\right|.$$

Under what condition(s) will F be a smooth local parametrization when (t,θ) is restricted to $(a,b) \times (0, 2\pi)$?

Under the condition on $\gamma(t) = (x(t), 0, z(t))$ that its surface of revolution is smooth, one can easily compute that the first fundamental form is given by:

(B.8)
$$[g] = \begin{bmatrix} (x')^2 + (z')^2 & 0 \\ 0 & x^2 \end{bmatrix} \qquad \text{(matrix notation)}$$
$$g = \left[(x')^2 + (z')^2\right] dt^2 + x^2 d\theta^2 \qquad \text{(tensor notation)}$$

and the second fundamental form with respect to the Gauss map $\mathsf{N} := \frac{\frac{\partial \mathsf{F}}{\partial t} \times \frac{\partial \mathsf{F}}{\partial \theta}}{\left|\frac{\partial \mathsf{F}}{\partial t} \times \frac{\partial \mathsf{F}}{\partial \theta}\right|}$ is given by:

(B.9)
$$[h] = \frac{1}{\sqrt{(x')^2 + (z')^2}} \begin{bmatrix} x'\,z'' - x''\,z' & 0 \\ 0 & x\,z' \end{bmatrix} \qquad \text{(matrix notation)}$$
$$h = \frac{1}{\sqrt{(x')^2 + (z')^2}} \left[\left(x'\,z'' - x''\,z'\right) dt^2 + x\,z' d\theta^2\right] \qquad \text{(tensor notation)}$$

Exercise B.2. Verify that the first and second fundamental forms of a surface of revolution with parametrization
$$\mathsf{F}(t,\theta) = (x(t)\cos\theta,\, x(t)\sin\theta,\, z(t)), \quad (t,\theta) \in (a,b) \times [0, 2\pi]$$
are given as in (B.8) and (B.9).

As both $[g]$ and $[h]$ are diagonal matrices, it is evident that the principal curvatures, i.e. the eigenvalues of $[g]^{-1}[h]$, are:
$$k_1 = \frac{x'z'' - x''z'}{\left[(x')^2 + (z')^2\right]^{3/2}} = \frac{x'z'' - x''z'}{|\gamma'|^3}$$
$$k_2 = \frac{z'}{x\sqrt{(x')^2 + (z')^2}} = \frac{z'}{x\,|\gamma'|}$$

Note that here we are not using the convention that $k_1 \leq k_2$ as in before, since there is no clear way to tell which eigenvalue is larger.

Therefore, the Gauss and mean curvatures are given by:

(B.10)
$$K = k_1 k_2 = \frac{(x'z'' - x''z')\,z'}{x\,|\gamma'|^4}$$

(B.11)
$$H = \frac{1}{2}(k_1 + k_2) = \frac{1}{2}\left(\frac{x'z'' - x''z'}{|\gamma'|^3} + \frac{z'}{x\,|\gamma'|}\right)$$

B.4. Covariant Derivatives

B.4.1. Vector Fields. A vector field X on a regular surface M is an assignment of a tangent vector $\mathsf{X}(p)$ to each point $p \in M$. Given a smooth local parametrization $\mathsf{F}(u_1, u_2)$, there is a natural basis $\{\frac{\partial \mathsf{F}}{\partial u_1}, \frac{\partial \mathsf{F}}{\partial u_2}\}$ for the tangent plane T_pM. A vector field X is often expressed as a linear combination of this basis instead of the standard basis in $\mathbb{R}^3$:

$$\mathsf{X}(p) = \sum_i X^i(p) \frac{\partial \mathsf{F}}{\partial u_i}(p).$$

As a convention set by geometers, the upper index i is used for the coefficient X^i. There is an important reason of doing so but we will not discuss it in this course. However, readers should not be confused X^i with "X to the power of i".

Under the local parametriation $\mathsf{F}(u_1, u_2)$, we can regard X, X^i and $\frac{\partial \mathsf{F}}{\partial u_i}$ to be functions of (u_1, u_2). While X is a function of $p \in M$, we can pre-compose it by F so that $\mathsf{X} \circ \mathsf{F}$ is a function of (u_1, u_2). With abuse of notations, we will write:

$$\frac{\partial \mathsf{X}}{\partial u_j}(p) := \frac{\partial(\mathsf{X} \circ \mathsf{F})}{\partial u_j}(u_1, u_2)$$

where (u_1, u_2) is the point corresponding to p, i.e. $\mathsf{F}(u_1, u_2) = p$. Likewise, we can also denote

$$\frac{\partial X^i}{\partial u_j}(p) := \frac{\partial(X^i \circ \mathsf{F})}{\partial u_j}(u_1, u_2).$$

As in Multivariable Calculus, other than partial derivatives, one can also talk about directional derivatives:

Definition B.25 (Directional Derivatives: along curves)**.** Let M be a regular surface, and $\gamma(t) : (a, b) \to M$ be a smooth curve on M. Given a vector field X on M, we define the directional derivative of X at $p \in M$ along γ to be

$$D_{\gamma'}\mathsf{X}(p) := \frac{d}{dt}\mathsf{X}(\gamma(t))\bigg|_{t=t_0}$$

where t_0 is a time such that $\gamma(t_0) = p$.

Example B.26. When $\gamma(t)$ is a u_1-coordinate curve $\gamma(t) = \mathsf{F}(t, 0)$ and X is any vector field, then $\mathsf{X}(\gamma(t)) = \mathsf{X}(\mathsf{F}(t, 0))$ and so

$$D_{\gamma'}\mathsf{X} = \frac{d}{dt}(\mathsf{X} \circ \mathsf{F})(t, 0) = \frac{\partial}{\partial u_1}(\mathsf{X} \circ \mathsf{F}) =: \frac{\partial \mathsf{X}}{\partial u_1}.$$

In particular, if $\mathsf{X} = \frac{\partial \mathsf{F}}{\partial u_2}$, then we have:

$$D_{\gamma'}\mathsf{X} = \frac{\partial}{\partial u_1}\frac{\partial \mathsf{F}}{\partial u_2} = \frac{\partial^2 \mathsf{F}}{\partial u_1 \partial u_2}.$$

□

Suppose $\mathsf{X} = \sum_i X^i \frac{\partial \mathsf{F}}{\partial u_i}$ and given any curve $\gamma(t) = \mathsf{F}(u_1(t), u_2(t))$ on M, then by the chain rule we have:

$$D_{\gamma'}\mathsf{X} = \frac{d}{dt}\mathsf{X}(\gamma(t)) = \frac{d}{dt}(\mathsf{X} \circ \mathsf{F})(u_1(t), u_2(t)) \tag{B.12}$$
$$= \sum_i \frac{\partial(\mathsf{X} \circ \mathsf{F})}{\partial u_i}\frac{du_i}{dt} =: \sum_i \frac{\partial \mathsf{X}}{\partial u_i}\frac{du_i}{dt}$$
$$= \sum_i \frac{\partial}{\partial u_i}\left(\sum_j X^j \frac{\partial \mathsf{F}}{\partial u_j}\right)\frac{du_i}{dt}$$
$$= \sum_{i,j}\left(\frac{\partial X^j}{\partial u_i}\frac{\partial \mathsf{F}}{\partial u_j} + X^j \frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}\right)\frac{du_i}{dt}$$

Note that when fixing the smooth local parametrization F, the quantities $\frac{\partial X^j}{\partial u_i}\frac{\partial \mathsf{F}}{\partial u_j} + X^j \frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}$ are uniquely determined by the vector field X whereas $\frac{du_i}{dt}$ are uniquely determined by the tangent vector γ' of the curve. Now given another vector field $\mathsf{Y} = \sum_i Y^i \frac{\partial \mathsf{F}}{\partial u_i}$. The Existence Theorem of ODEs guarantees there is a curve γ that flows along Y, or precisely $\gamma'(t) = \mathsf{Y}(\gamma(t))$ for all t. Since $D_{\gamma'}\mathsf{X}$ depends on the tangent vector γ' but not on the curve γ, we can also define directional derivatives along a vector field:

Definition B.27 (Directional Derivatives: along vector fields)**.** Let M be a regular surface, and X and Y be two vector fields on M. The directional derivative of X at $p \in M$ along Y is defined to be:

$$D_{\mathsf{Y}}\mathsf{X}(p) = D_{\gamma'}\mathsf{X}(p)$$

where γ is a curve on M which solves the ODE $\gamma'(t) = \mathsf{Y}(\gamma(t))$ and $\gamma(0) = p$.

Express γ in the above definition as $\mathsf{F}(u_1(t), u_2(t))$, then by the chain rule we get:

$$\gamma'(t) = \sum_i \frac{\partial \mathsf{F}}{\partial u_i}\frac{du_i}{dt}.$$

If Y is expressed $\sum_i Y^i \frac{\partial \mathsf{F}}{\partial u_i}$, then $\gamma'(t) = \mathsf{Y}(\gamma(t)) = (\mathsf{Y} \circ \mathsf{F})(u_1(t), u_2(t))$ is equivalent to saying:

$$Y^i = \frac{du_i}{dt}.$$

From (B.12), we get the local expression for $D_{\mathsf{Y}}\mathsf{X}$:

$$D_{\mathsf{Y}}\mathsf{X} = \sum_{i,j}\left(\frac{\partial X^j}{\partial u_i}\frac{\partial \mathsf{F}}{\partial u_j} + X^j \frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}\right)Y^i. \tag{B.13}$$

Definition B.28 (Covariant Derivatives)**.** Let M be a regular surface with Gauss map N, and $\gamma(t)$ be a smooth curve on M. Given two vector fields $\mathsf{X} : M \to TM$ and $\mathsf{Y} : M \to TM$, we define the covariant derivative of X at $p \in M$ along Y to be

$$\nabla_{\mathsf{Y}}\mathsf{X}(p) := (D_{\mathsf{Y}}\mathsf{X}(p))^{\text{tangent}} = D_{\mathsf{Y}}\mathsf{X}(p) - (D_{\mathsf{Y}}\mathsf{X}(p) \cdot \mathsf{N}(p))\,\mathsf{N}(p).$$

Here $(D_{\mathsf{Y}}\mathsf{X}(p))^{\text{tangent}}$ represents the projection of $D_{\mathsf{Y}}\mathsf{X}(p)$ onto the tangent plane T_pM.

By (B.13), we can derive the local expression for $\nabla_{\mathsf{Y}}\mathsf{X}$:

Proposition B.29. *Let M be a regular surface parametrized by $\mathsf{F}(u_1, u_2)$ with Gauss map N, and X and Y be two vector fields on M with local expressions: $\mathsf{X} = \sum_i X^i \frac{\partial \mathsf{F}}{\partial u_i}$ and $\mathsf{Y} = \sum_i Y^i \frac{\partial \mathsf{F}}{\partial u_i}$. Then, the covariant derivative $\nabla_{\mathsf{Y}}\mathsf{X}$ can be expressed in terms of local coordinates (u_1, u_2) as:*

$$\nabla_{\mathsf{Y}}\mathsf{X} = \underbrace{\sum_{i,j}\left(\frac{\partial X^j}{\partial u_i}\frac{\partial \mathsf{F}}{\partial u_j} + X^j \frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}\right) Y^i}_{D_{\mathsf{Y}}\mathsf{X}} - \underbrace{\left(\sum_{i,j} h_{ij} X^i Y^j\right)\mathsf{N}}_{(D_{\mathsf{Y}}\mathsf{X}\cdot\mathsf{N})\mathsf{N}} \tag{B.14}$$

where h_{ij} is the second fundamental form defined as $h_{ij} := h\left(\frac{\partial \mathsf{F}}{\partial u_i}, \frac{\partial \mathsf{F}}{\partial u_j}\right)$.

Proof. It suffices to show that $(D_{\mathsf{Y}}\mathsf{X}\cdot\mathsf{N})\,\mathsf{N} = \left(\sum_{i,j} h_{ij} X^i Y^j\right)\mathsf{N}$ □

Exercise B.3. Suppose X, $\widetilde{\mathsf{X}}$, Y and $\widetilde{\mathsf{Y}}$ are vector fields on a regular surface M, and φ is a smooth scalar functions defined on M. Verify that:

(a) $D_{\varphi\mathsf{Y}}\mathsf{X} = \varphi D_{\mathsf{Y}}\mathsf{X}$

(b) $D_{\mathsf{Y}}(\varphi\mathsf{X}) = (D_{\mathsf{Y}}\varphi)\,\mathsf{X} + \varphi D_{\mathsf{Y}}\mathsf{X}$

(c) $D_{\mathsf{Y}+\widetilde{\mathsf{Y}}}\mathsf{X} = D_{\mathsf{Y}}\mathsf{X} + D_{\widetilde{\mathsf{Y}}}\mathsf{X}$

(d) $D_{\mathsf{Y}}(\mathsf{X}+\widetilde{\mathsf{X}}) = D_{\mathsf{Y}}\mathsf{X} + D_{\mathsf{Y}}\widetilde{\mathsf{X}}$

Verify that (a)-(d) also hold if all D's are replaced by ∇'s.

B.4.2. Christoffel Symbols. The second derivatives $\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}$ play two very important roles in these derivatives. For one thing, they are directional derivatives $D_{\mathsf{Y}}\mathsf{X}$ with $\mathsf{X} = \frac{\partial \mathsf{F}}{\partial u_i}$ and $\mathsf{Y} = \frac{\partial \mathsf{F}}{\partial u_j}$. For another, they are important quantities in both (B.13) and (B.14) since both $D_{\mathsf{Y}}\mathsf{X}$ and $\nabla_{\mathsf{Y}}\mathsf{X}$ are in terms of them. At each $p \in M$, there is a natural basis $\left\{\frac{\partial \mathsf{F}}{\partial u_1}, \frac{\partial \mathsf{F}}{\partial u_2}, \mathsf{N}\right\}$ of $\mathbb{R}^3$. It is a much better basis than the standard one when dealing the surface M since the first two basis vectors are tangents and the last basis vector is the normal. When projecting a vector onto the tangent plane T_pM, we may simply drop the N-component. In this connection, we are going to express $\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}$ in terms of this tangent-normal basis.

Suppose we have:

$$\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j} = a\frac{\partial \mathsf{F}}{\partial u_1} + b\frac{\partial \mathsf{F}}{\partial u_2} + c\mathsf{N}.$$

It is not difficult to show that $c = h_{ij}$ (Exercise). We denote the coefficients a and b by the following symbols:

Definition B.30 (Christoffel Symbols)**.** Given a smooth local parametrization $\mathsf{F}(u_1, u_2)$ of a regular surface M, we define the Christoffel symbols Γ_{ij}^k to be the tangent coefficients of $\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j}$. Precisely, we have:

$$\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j} = \sum_k \Gamma_{ij}^k \frac{\partial \mathsf{F}}{\partial u_k} + h_{ij}\mathsf{N}. \tag{B.15}$$

Since $\dfrac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j} = D_{\frac{\partial \mathsf{F}}{\partial u_j}} \left(\dfrac{\partial \mathsf{F}}{\partial u_i} \right)$, we have:

$$\nabla_{\frac{\partial \mathsf{F}}{\partial u_j}} \left(\frac{\partial \mathsf{F}}{\partial u_i} \right) = \left(\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_j} \right)^{\text{tangent}} = \sum_k \Gamma^k_{ij} \frac{\partial \mathsf{F}}{\partial u_k}.$$

For simplicity, from now on we will denote $D_{\frac{\partial \mathsf{F}}{\partial u_j}} \mathsf{X}$ and $\nabla_{\frac{\partial \mathsf{F}}{\partial u_j}} \mathsf{X}$ simply by:

$$D_j \mathsf{X} := D_{\frac{\partial \mathsf{F}}{\partial u_j}} \mathsf{X}$$
$$\nabla_j \mathsf{X} := \nabla_{\frac{\partial \mathsf{F}}{\partial u_j}} \mathsf{X}$$

The Christoffel symbols can be shown to be depending only on the first fundamental form g. We will use it to prove that Gauss curvature depends also only on g but not on h.

Lemma B.31. *Let M be a regular surface with smooth local parametrization $\mathsf{F}(u_1, u_2)$. Then for any i, j and k, the Christoffel symbols Γ^k_{ij}'s can be locally expressed in terms of the first fundamental form as:*

$$\Gamma^k_{ij} = \frac{1}{2} \sum_l g^{kl} \left(\frac{\partial g_{jl}}{\partial u_i} + \frac{\partial g_{il}}{\partial u_j} - \frac{\partial g_{ij}}{\partial u_l} \right). \tag{B.16}$$

Proof. First recall that $g_{ij} = \frac{\partial \mathsf{F}}{\partial u_i} \cdot \frac{\partial \mathsf{F}}{\partial u_j}$. By differentiating both sides respect to u_l, we get:

$$\frac{\partial g_{ij}}{\partial u_l} = \frac{\partial^2 \mathsf{F}}{\partial u_l \partial u_i} \cdot \frac{\partial \mathsf{F}}{\partial u_j} + \frac{\partial^2 \mathsf{F}}{\partial u_l \partial u_j} \cdot \frac{\partial \mathsf{F}}{\partial u_i}.$$

Using (B.15), we get:

$$\begin{aligned} \frac{\partial g_{ij}}{\partial u_l} &= \underbrace{\left(\sum_k \Gamma^k_{li} \frac{\partial \mathsf{F}}{\partial u_k} + h_{li} \mathsf{N} \right)}_{\frac{\partial^2 \mathsf{F}}{\partial u_l \partial u_i}} \cdot \frac{\partial \mathsf{F}}{\partial u_j} + \underbrace{\left(\sum_k \Gamma^k_{lj} \frac{\partial \mathsf{F}}{\partial u_k} + h_{lj} \mathsf{N} \right)}_{\frac{\partial^2 \mathsf{F}}{\partial u_l \partial u_j}} \cdot \frac{\partial \mathsf{F}}{\partial u_i} \\ &= \sum_k \left(\Gamma^k_{il} g_{kj} + \Gamma^k_{lj} g_{ki} \right) \end{aligned} \tag{*}$$

By cyclic permutation of indices $\{i, j, l\}$, we also get:

$$\frac{\partial g_{il}}{\partial u_j} = \sum_k \left(\Gamma^k_{ij} g_{kl} + \Gamma^k_{jl} g_{ki} \right) \tag{**}$$

$$\frac{\partial g_{jl}}{\partial u_i} = \sum_k \left(\Gamma^k_{ji} g_{kl} + \Gamma^k_{il} g_{kj} \right) \tag{***}$$

Recall that $\Gamma^k_{ij} = \Gamma^k_{ji}$ and $h_{ij} = h_{ji}$ for any i, j and k. By considering (**)+(***)-(*), we get:

$$\frac{\partial g_{il}}{\partial u_j} + \frac{\partial g_{jl}}{\partial u_i} - \frac{\partial g_{ij}}{\partial u_l} = 2 \sum_k \Gamma^k_{ij} g_{kl} \tag{B.17}$$

Finally, consider the sum $\sum_l g_{kl} g^{lq}$ which is the (k, q)-entry of the matrix $[g][g]^{-1} = I$. Therefore, we have:

$$\sum_l g_{kl} g^{lq} = \delta^q_k = \begin{cases} 1 & \text{if } k = q \\ 0 & \text{if } k \neq q \end{cases}$$

Multiplying g^{lq} on both sides of (B.17) and summing up over all l, we get:

$$\begin{aligned}\sum_l g^{lq}\left(\frac{\partial g_{il}}{\partial u_j}+\frac{\partial g_{jl}}{\partial u_i}-\frac{\partial g_{ij}}{\partial u_l}\right) &= 2\sum_l\sum_k g^{lq}\Gamma^k_{ij}g_{kl}\\ &= 2\sum_k \Gamma^k_{ij}\sum_l g_{kl}g^{lq} = 2\sum_k \Gamma^k_{ij}\delta^q_k\\ &= 2\Gamma^q_{ij}.\end{aligned}$$

The last equality follows from the fact when summing up over all k, the only "survivor" is the only one that has $k = q$. By rearranging terms and relabelling the index q by k, we complete the proof of (B.16). $\square$

B.5. Theorema Egregium

In this section we will introduce and prove an important theorem in Differential Geometry – the Gauss's Theorema Egregium. This surprising theorem asserts that even though the definition of Gauss curvature involves both the first and second fundamental form, it in fact depends only on the first fundamental form! We will dissect this complicated proof into several components. First we will show that the derivative of the Gauss map is the negative of the shape operator. Using this, we derive the Gauss-Codazzi's equations, and from there we can show the Gauss curvature K depends only on g_{ij}'s.

B.5.1. Shape Operator Revisited. Let N be the Gauss map of an orientable regular surface M and $\mathsf{F}(u_1, u_2) : \mathcal{U} \to M$ be a smooth local parametrization. With a bit abuse of notations, we abbreviate the map $\mathsf{N} \circ \mathsf{F} : \mathcal{U} \to \mathbb{S}^2$ simply by N and so it makes sense to denote $\frac{\partial \mathsf{N}}{\partial u_j}$. In this subsection, we will find out what this partial derivative exactly is.

Since N is unit, we have $|\mathsf{N}|^2 = 1$. Therefore, $\frac{\partial \mathsf{N}}{\partial u_j}$ is orthogonal to N for any j, and so one can write $\frac{\partial \mathsf{N}}{\partial u_j}$ as a linear combination of tangent vectors $\frac{\partial \mathsf{F}}{\partial u_i}$'s. Let A^i_j be functions of (u_1, u_2) such that:

$$\frac{\partial \mathsf{N}}{\partial u_j} = \sum_i A^i_j \frac{\partial \mathsf{F}}{\partial u_i}.$$

Taking dot product with $\frac{\partial \mathsf{F}}{\partial u_k}$ on both sides, we get

$$\frac{\partial \mathsf{N}}{\partial u_j} \cdot \frac{\partial \mathsf{F}}{\partial u_k} = \sum_i A^i_j \frac{\partial \mathsf{F}}{\partial u_i} \cdot \frac{\partial \mathsf{F}}{\partial u_k} = \sum_i A^i_j g_{ik} \qquad \text{for any } j, k = 1, 2 \tag{B.18}$$

On the other hand, as $\mathsf{N} \cdot \frac{\partial \mathsf{F}}{\partial u_k} = 0$, by differentiation we get:

$$0 = \frac{\partial}{\partial u_j}\left(\mathsf{N} \cdot \frac{\partial \mathsf{F}}{\partial u_k}\right) = \frac{\partial \mathsf{N}}{\partial u_j} \cdot \frac{\partial \mathsf{F}}{\partial u_k} + \underbrace{\mathsf{N} \cdot \frac{\partial^2 \mathsf{F}}{\partial u_j \partial u_k}}_{h_{jk}}.$$

Hence $\frac{\partial \mathsf{N}}{\partial u_j} \cdot \frac{\partial \mathsf{F}}{\partial u_k} = -h_{jk}$. Substitute it back to (B.18), we get:

$$-h_{kj} = \sum_i g_{ki} A^i_j \quad \text{for any } j, k = 1, 2.$$

In matrix form, it is equivalent to saying that:

$$-[h] = [g][A] \implies [A] = -[g]^{-1}[h] = -[S]$$

where $[S]$ is the shape operator. To summarize, we have proved:

Lemma B.32. *Let M be a regular surface with Gauss map N and smooth local parametrization $\mathsf{F}(u_1, u_2)$. Then, we have:*

$$\frac{\partial \mathsf{N}}{\partial u_j} = -\sum_i S^i_j \frac{\partial \mathsf{F}}{\partial u_i} = -\sum_{i,k} g^{ik} h_{kj} \frac{\partial \mathsf{F}}{\partial u_i}. \tag{B.19}$$

B.5.2. Gauss-Codazzi's Equations. We derive the Gauss-Codazzi's equations in this subsection. The Gauss equation is important in Differential Geometry not only because it leads to the proof of the *Theoerma Egregium*, but also motivated the development of Riemannian Geometry.

Theorem B.33 (Gauss-Codazzi's Equations).

$$\frac{\partial \Gamma^q_{jk}}{\partial u_i} - \frac{\partial \Gamma^q_{ik}}{\partial u_j} + \sum_l \Gamma^l_{jk}\Gamma^q_{il} - \sum_l \Gamma^l_{ik}\Gamma^q_{jl} = \sum_l g^{ql}\left(h_{jk}h_{li} - h_{ik}h_{lj}\right) \qquad \text{(Gauss)}$$

$$\frac{\partial h_{jk}}{\partial u_i} - \frac{\partial h_{ik}}{\partial u_j} + \sum_l \left(\Gamma^l_{jk}h_{il} - \Gamma^l_{ik}h_{jl}\right) = 0 \qquad \text{(Codazzi)}$$

Proof. The key step of the proof is to start with the fact that:

$$\frac{\partial^3 \mathsf{F}}{\partial u_i \partial u_j \partial u_k} = \frac{\partial^3 \mathsf{F}}{\partial u_j \partial u_i \partial u_k}$$

for any i, j, k, and then rewrite both sides in terms of the tangent-normal basis $\left\{\frac{\partial \mathsf{F}}{\partial u_1}, \frac{\partial \mathsf{F}}{\partial u_2}, \mathsf{N}\right\}$ of $\mathbb{R}^3$. The Gauss's equation follows from equating the tangent coefficients, and the Codazzi's equation is obtained by equating the normal coefficient.

By (B.15), we have:

$$\frac{\partial^2 \mathsf{F}}{\partial u_j \partial u_k} = \sum_l \Gamma^l_{jk}\frac{\partial \mathsf{F}}{\partial u_l} + h_{jk}\mathsf{N}.$$

Differentiating both sides with respect to u_i, we get:

$$\begin{aligned}
\frac{\partial^3 \mathsf{F}}{\partial u_i \partial u_j \partial u_k} &= \frac{\partial}{\partial u_i}\left(\sum_l \Gamma^l_{jk}\frac{\partial \mathsf{F}}{\partial u_l} + h_{jk}\mathsf{N}\right) \\
&= \sum_l \left(\frac{\partial \Gamma^l_{jk}}{\partial u_i}\frac{\partial \mathsf{F}}{\partial u_l} + \Gamma^l_{jk}\frac{\partial^2 \mathsf{F}}{\partial u_i \partial u_l}\right) + \frac{\partial h_{jk}}{\partial u_i}\mathsf{N} + h_{jk}\frac{\partial \mathsf{N}}{\partial u_i} \\
&= \sum_l \left(\frac{\partial \Gamma^l_{jk}}{\partial u_i}\frac{\partial \mathsf{F}}{\partial u_l} + \Gamma^l_{jk}\left(\sum_q \Gamma^q_{il}\frac{\partial \mathsf{F}}{\partial u_q} + h_{il}\mathsf{N}\right)\right) + \frac{\partial h_{jk}}{\partial u_i}\mathsf{N} + h_{jk}\frac{\partial \mathsf{N}}{\partial u_i} \\
&= \underbrace{\sum_q \frac{\partial \Gamma^q_{jk}}{\partial u_i}\frac{\partial \mathsf{F}}{\partial u_q}}_{l \mapsto q} + \sum_{l,q}\Gamma^l_{jk}\Gamma^q_{il}\frac{\partial \mathsf{F}}{\partial u_q} + \left(\frac{\partial h_{jk}}{\partial u_i} + \sum_l \Gamma^l_{jk}h_{il}\right)\mathsf{N} + h_{jk}\frac{\partial \mathsf{N}}{\partial u_i} \\
&= \sum_q \left(\frac{\partial \Gamma^q_{jk}}{\partial u_i} + \sum_l \Gamma^l_{jk}\Gamma^q_{il}\right)\frac{\partial \mathsf{F}}{\partial u_q} + \left(\frac{\partial h_{jk}}{\partial u_i} + \sum_l \Gamma^l_{jk}h_{il}\right)\mathsf{N} + h_{jk}\frac{\partial \mathsf{N}}{\partial u_i}
\end{aligned}$$

Now every except the last term is in terms of the tangent-normal basis. To handle the last term, recall from (B.19) that:

$$\frac{\partial \mathsf{N}}{\partial u_i} = -\sum_q S^q_i \frac{\partial \mathsf{F}}{\partial u_q} = -\sum_{l,q} g^{ql}h_{li}\frac{\partial \mathsf{F}}{\partial u_q}.$$

Here we relabelled the indices to avoid repetitions. Finally, we showed:

$$\begin{aligned}\frac{\partial^3 \mathsf{F}}{\partial u_i \partial u_j \partial u_k} &= \sum_q \left(\frac{\partial \Gamma^q_{jk}}{\partial u_i} + \sum_l \Gamma^l_{jk}\Gamma^q_{il} \right) \frac{\partial \mathsf{F}}{\partial u_q} + \left(\frac{\partial h_{jk}}{\partial u_i} + \sum_l \Gamma^l_{jk} h_{il} \right) \mathsf{N} - \sum_{l,q} h_{jk} g^{ql} h_{li} \frac{\partial \mathsf{F}}{\partial u_q} \\ &= \sum_q \left(\frac{\partial \Gamma^q_{jk}}{\partial u_i} + \sum_l \Gamma^l_{jk}\Gamma^q_{il} - \sum_l h_{jk} g^{ql} h_{li} \right) \frac{\partial \mathsf{F}}{\partial u_q} + \left(\frac{\partial h_{jk}}{\partial u_i} + \sum_l \Gamma^l_{jk} h_{il} \right) \mathsf{N}\end{aligned}$$

By switching i and j, we get:

$$\frac{\partial^3 \mathsf{F}}{\partial u_j \partial u_i \partial u_k} = \sum_q \left(\frac{\partial \Gamma^q_{ik}}{\partial u_j} + \sum_l \Gamma^l_{ik}\Gamma^q_{jl} - \sum_l h_{ik} g^{ql} h_{lj} \right) \frac{\partial \mathsf{F}}{\partial u_q} + \left(\frac{\partial h_{ik}}{\partial u_j} + \sum_l \Gamma^l_{ik} h_{jl} \right) \mathsf{N}$$

The Gauss-Codazzi's equations can be obtained by equating the coefficients of each tangent and normal component. $\square$

B.5.3. Gauss Curvature is Intrinsic! In the previous subsection, we derived the Gauss-Codazzi's equations (Theorem B.33). It is worthwhile the note that the LHS of the Gauss's equation involves only Christoffel's symbols and their derivatives:

$$\underbrace{\frac{\partial \Gamma^q_{jk}}{\partial u_i} - \frac{\partial \Gamma^q_{ik}}{\partial u_j} + \sum_l \Gamma^l_{jk}\Gamma^q_{il} - \sum_l \Gamma^l_{ik}\Gamma^q_{jl}}_{\text{depends only on } \Gamma^k_{ij}\text{'s}} = \sum_l g^{ql} \left(h_{jk} h_{li} - h_{ik} h_{lj} \right)$$

From (B.16) we also know that the Christoffel symbols depend only on the first fundamental form g but not on h. In this connection, we denote:

$$R^q_{kij} := \frac{\partial \Gamma^q_{jk}}{\partial u_i} - \frac{\partial \Gamma^q_{ik}}{\partial u_j} + \sum_l \Gamma^l_{jk}\Gamma^q_{il} - \sum_l \Gamma^l_{ik}\Gamma^q_{jl}.$$

The lower and upper indices for R^q_{kij} are chosen so as to preserve their positions in the RHS expression (q being upper, and i, j, k being lower). However, different authors may use different conventions for the order of the lower indices when writing R_{kij}.

We are now in a position to give a proof of Gauss's Theorema Egregium, a very important result in Differential Geometry which leads to the development of Riemannian Geometry later on.

Theorem B.34 (Theorema Egregium)**.** *The Gauss curvature K of any a regular surface M depends only on its first fundamental form g. In other words, K is intrinsic.*

Proof. Consider the Gauss's equation, which asserts that for any i, j, k and q:

$$R^q_{kij} = \sum_l g^{ql} \left(h_{jk} h_{li} - h_{ik} h_{lj} \right).$$

Multiply both sides by g_{pq}, and sum up all q's, we get:

$$\sum_q g_{pq} R^q_{kij} = \sum_q g_{pq} \left(\sum_l g^{ql} \left(h_{jk} h_{li} - h_{ik} h_{lj} \right) \right) = \sum_l \sum_q g_{pq} g^{ql} \left(h_{jk} h_{li} - h_{ik} h_{lj} \right).$$

Note that $\sum_q g_{pq} g^{ql}$ is the (p, l)-entry of $[g][g]^{-1}$, which is simply the identity matrix. Therefore,

$$\sum_q g_{pq} g^{ql} = \delta^l_p = \begin{cases} 1 & \text{if } p = l \\ 0 & \text{if } p \neq l \end{cases}$$

Substitute this back in, we get:

$$\sum_q g_{pq} R^q_{kij} = \sum_l \delta^l_p \left(h_{jk} h_{li} - h_{ik} h_{lj} \right).$$

When summing over l, the fact δ^l_p is non-zero only when $l = p$. Therefore, we get:

$$\sum_q g_{pq} R^q_{kij} = \underbrace{h_{jk} h_{pi} - h_{ik} h_{pj}}_{\text{the only survivor}}.$$

The above result is true for any i, j, k and p. In particular, when $(i, j, k, p) = (1, 2, 2, 1)$, we get:

$$\sum_q g_{1q} R^q_{212} = h_{22} h_{11} - h_{12}^2 = \det[h].$$

This shows $\det[h]$ depends only on g since R^q_{212} is so.

Finally, recall that the Gauss curvature is given by:

$$K = \frac{\det[h]}{\det[g]}.$$

Therefore, we have completed the proof that K depends only g. □

B.5.4. Isometric Surfaces. Let M and $\widetilde{M}$ be two regular surfaces. They are considered to be *topologically* the same if there exists a bijective map $\Phi : M \to \widetilde{M}$ such that both Φ and Φ^{-1} are continuous. This map Φ is said to be a *homeomorphism*, and M and $\widetilde{M}$ are said to be *homeomorphic*. If furthermore such a map Φ and its inverse Φ^{-1} are smooth (see Definition 1.17), we say Φ is a *diffeomorphism*, and the two surfaces are said to be *diffeomorphic*. In Differential Geometry, we deal with diffeomorphisms much more often than homeomorphisms.

Consider two diffeomorphic regular surfaces M and $\widetilde{M}$ via a diffeomorphism $\Phi : M \to \widetilde{M}$. M and $\widetilde{M}$ are then *topologically* equal to each other in a *smooth* way. In this subsection, we will discuss a condition on Φ under which M and $\widetilde{M}$ are *geometrically* equal, i.e. same area, same curvature, etc.

Exercise B.4. Let M and $\widetilde{M}$ be two diffeomorphic regular surfaces via the diffeomorphism $\Phi : M \to \widetilde{M}$. Suppose $\mathsf{F}(u_1, u_2)$ is a smooth local parametrization of M, show that $(\Phi \circ \mathsf{F})(u_1, u_2)$ is a smooth local parametrization of $\widetilde{M}$.

If $\mathsf{F}(u_1, u_2)$ is a smooth local parametrization of M, then $(\Phi \circ \mathsf{F})(u_1, u_2)$ is a smooth local parametrization of $\widetilde{M}$ (see Exercise B.4). Under this pair of smooth local parametrizations, the corresponding points p and $\Phi(p)$ will have the same (u_1, u_2)-coordinates. Since many geometric quantities including surface area and Gauss curvature are uniquely determined by the first fundamental form, we want to seek a condition on Φ under which the first fundamental forms of M and $\widetilde{M}$ are equal at corresponding points.

Recall that their first fundamental forms are defined by:

$$g_{ij} = \frac{\partial \mathsf{F}}{\partial u_i} \cdot \frac{\partial \mathsf{F}}{\partial u_j} \qquad \text{(first fundamental form of } M\text{)}$$

$$\widetilde{g}_{ij} = \frac{\partial (\Phi \circ \mathsf{F})}{\partial u_i} \cdot \frac{\partial (\Phi \circ \mathsf{F})}{\partial u_j} \qquad \text{(first fundamental form of } \widetilde{M}\text{)}$$

Should they be equal, the two surfaces M and $\widetilde{M}$ will have the same surface area and Gauss's curvature. In this connection, we define:

Definition B.35 (Isometric Surfaces)**.** Let M and $\widetilde{M}$ be two orientable regular surfaces. Then M and $\widetilde{M}$ are said to be *isometric* if there exists a diffeomorphism $\Phi : M \to \widetilde{M}$ such that when $\mathsf{F}(u_1, u_2)$ is a smooth local parametrization of M, we have:

$$\frac{\partial(\Phi \circ \mathsf{F})}{\partial u_i} \cdot \frac{\partial(\Phi \circ \mathsf{F})}{\partial u_j} = \frac{\partial \mathsf{F}}{\partial u_i} \cdot \frac{\partial \mathsf{F}}{\partial u_j} \quad \text{for any } i, j.$$

In such case, Φ is called an *isometry* between M and $\widetilde{M}$.

Example B.36. Let $\Pi = \{(x, y, 0) : 0 < x < 2\pi \text{ and } y \in \mathbb{R}\}$ be an open subset of the xy-plane in $\mathbb{R}^3$. It can be parametrized by

$$\mathsf{F}(u_1, u_2) = (u_1, u_2, 0), \qquad (u_1, u_2) \in (0, 2\pi) \times \mathbb{R}.$$

Then, clearly we have:

$$\frac{\partial \mathsf{F}}{\partial u_i} \cdot \frac{\partial \mathsf{F}}{\partial u_j} = \delta_{ij}.$$

Let $\Sigma = \{(x, y, z) : x^2 + y^2 = 1\}$ be the infinite cylinder with radius 1 centered at the z-axis. Then, we remove the straight-line $L = \{(1, 0, z) : z \in \mathbb{R}\}$ from Σ, and define the map $\Phi : \Pi \to \Sigma \backslash L$ by:

$$\Phi(x, y, 0) = (\cos x, \sin x, y)$$

It is easy to show that Φ is one-to-one and onto (left as an exercise for readers).

Then, we have:

$$(\Phi \circ \mathsf{F})(u_1, u_2) = \Phi(u_1, u_2, 0) = (\cos u_1, \sin u_1, u_2).$$

By direct computations, we can verify that:

$$\frac{\partial(\Phi \circ \mathsf{F})}{\partial u_i} \cdot \frac{\partial(\Phi \circ \mathsf{F})}{\partial u_j} = \delta_{ij} \quad \text{for any } i, j.$$

Therefore, Π and $\Sigma \backslash L$ are isometric.

Hence, by Theorema Egregium, the Gauss's curvature of Π at p is the same as that of $\Sigma \backslash L$ at $\Phi(p)$, i.e. they are both zero. Of course, one can also find out the second fundamental form of $\Sigma \backslash L$ and compute the Gauss curvature directly. $\square$

B.6. Geodesics and Minimal Surfaces (work in progress)

B.6.1. Parallel Transport.

B.6.2. Geodesic Equation.

B.6.3. Geodesic Curvature.

B.6.4. Constant Mean Curvature Surfaces.

B.7. Gauss-Bonnet's Theorem (work in progress)

B.7.1. A Beautiful Theorem.

B.7.2. Applications and Significance.

Bibliography

[dC76] Manfredo P. do Carmo, *Differential geometry of curves and surfaces*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1976, Translated from the Portuguese. MR 0394451

[dC94] ______, *Differential forms and applications*, Universitext, Springer-Verlag, Berlin, 1994, Translated from the 1971 Portuguese original. MR 1301070

[Küh05] Wolfgang Kühnel, *Differential geometry, curves-surfaces-manifolds*, Student Mathematical Library, vol. 16, American Mathematical Society, Providence, RI, 2005. MR 3443721

[Lee09] Jeffrey M. Lee, *Manifolds and differential geometry*, Graduate Studies in Mathematics, vol. 107, American Mathematical Society, Providence, RI, 2009. MR 2572292

[Lee13] John M. Lee, *Introduction to smooth manifolds*, second ed., Graduate Texts in Mathematics, vol. 218, Springer, New York, 2013. MR 2954043

[Mun00] James Munkres, *Topology*, 2 ed., Prentice Hall, 2000.