CHINESE UNIVERSITY OF HONG KONG

MATH4030
Differential Geometry

香 港 中 文 大 學
The Chinese University of Hong Kong

Instructor: LUKE PEACHEY

Term 1, 2024/25

Contents

Introduction

This course will cover the basics of classical differential geometry: the study of local and global properties of curves and surfaces embedded within an extrinsic Euclidean space. This is distinguished from the modern approach to differential geometry, which deals mainly with intrinsically defined objects instead.

Despite the names classical and modern, the classical theory still has many uses in modern mathematics, as well as being essential for a deep intuitive understanding of the theory as a whole.

This course will introduce the basic concepts of differential geometry and prove many fundamental results regarding their structure. In turn, this will motivate the more modern approach to the subject. A fundamental result, the Theorem Egregium (1827) originally due to Gauss, states that

The Gaussian curvature of a surface depends only on intrinsic properties of the surface, and not on how such a surface is embedded in three space.

We will see a more precise formulation of the above theorem later in the course.

A remark on notation:
Often, I will use the notation associated with the more modern approach to the subject, rather than the classical notation in a lot of older textbooks. For example, in this course we will denote the first and second fundamental forms by g and A, as opposed to the notation I and $\mathrm{I\!I}$ found in the literature.

1 Curves

Recall, an interval is a non-empty connected subset of $\mathbb{R}$.

Definition 1.1. *A **curve** in $\mathbb{R}^3$ is a continuous map $\gamma : I \to \mathbb{R}^3$, where $I \subseteq \mathbb{R}$ is an open interval. We say that γ is a **smooth curve** if γ is a smooth function. The image $\gamma(I) \subseteq \mathbb{R}^3$ is called the **trace** of γ.*

Remark.

- *Replacing $\mathbb{R}^3$ with $\mathbb{R}^n$ is a valid generalisation of the above definition, although in this course we restrict our attention to the cases $n = 2$ or 3.*
- *We specify that our domain is open since we want our curve to be locally given by a section of the real line. Also, for smooth curves, its derivatives will then exist everywhere.*
- *We could replace the condition of being smooth with being C^k for any $k \geq 1$ instead, and all of the theory will carry through. However, to avoid unnecessary technicalities, we shall always work in the category of smooth objects in this course.*

Lets begin with some simple examples.

Example 1.2. *The function $\gamma : \mathbb{R} \to \mathbb{R}^3$ given by*

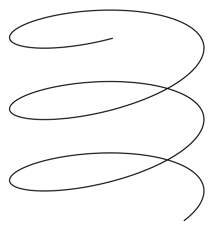
$$\gamma(t) = \begin{cases} (-1, t, 0) & : t \leq 0, \\ (1, t, 0) & : t > 0, \end{cases}$$

*is **not** a curve, since the function is not continuous.*

Example 1.3. *Fix $\alpha, \beta > 0$ and consider the curve $\gamma : \mathbb{R} \to \mathbb{R}^3$, given by*

$$\gamma(t) = (\alpha \cos t, \alpha \sin t, \beta t), \quad \forall t \in \mathbb{R}.$$

It is clear that γ is smooth. Its trace $\gamma(\mathbb{R})$ is a helix.



Example 1.4. *Consider the curve* $\gamma : \mathbb{R} \to \mathbb{R}^2$*, given by*

$$\gamma(t) = (t^3, t^2), \quad \forall t \in \mathbb{R}.$$

Note that γ *is smooth, however, its trace appears to have a singularity at the origin - there is no unique tangent line to the curve at this point. We will return to this example later.*

Example 1.5. *Consider the curve* $\gamma : \mathbb{R} \to \mathbb{R}^2$ *defined by*

$$\gamma(t) = (t^3 - 4t, t^2 - 4), \quad \forall t \in \mathbb{R}.$$

It is clear that γ *is smooth. However,* γ *is not injective; there are points of self-intersection*

$$\gamma(2) = (0, 0) = \gamma(-2).$$

Example 1.6. *The curve* $\gamma : \mathbb{R} \to \mathbb{R}^2$ *defined by*

$$\gamma(t) = (t, |t|), \quad \forall t \in \mathbb{R},$$

is ***not*** *smooth, since the map* $t \mapsto |t|$ *is not differentiable at* $t = 0$.

Exercise. *Find a smooth curve which has the same trace as the above non-smooth curve. What can be said about the derivative of this curve at the origin?*

In order to remove singularities such as the cusp like singularity from Example 1.4, we introduce an extra condition on our curves.

1.1 Regular curves

Definition 1.7. *Let* $\gamma : I \to \mathbb{R}^3$ *be a smooth curve. We say that* γ *is* ***regular*** *if*

$$\|\gamma'(t)\| \neq 0, \quad \forall t \in I.$$

If we interpret $\gamma(t)$ as the position of a particle at time t, then γ is regular if the speed of the particle is never zero. Equivalently, γ is regular if its velocity is never the zero vector, i.e

$$\gamma'(t) \neq (0, 0, 0), \quad \forall t \in I.$$

If γ is a regular curve, then for every $t \in I$, there is a unique straight line in $\mathbb{R}^3$ tangent to the curve $\gamma(I)$ at the point $\gamma(t)$, given parametrically as

$$\{\gamma(t) + s\gamma'(t) \in \mathbb{R}^3 : s \in \mathbb{R}\}.$$

Remark. *Examples 1.3 & 1.5 are regular curves, and Example 1.4 is not regular.*

As we shall see in this course, the existence of a canonical tangent line (or more generally tangent plane/tangent space) is essential for well-defined consistent formulations of calculus on our objects. We therefore restrict our attention to regular curves for the remainder of this section.

1.2 Arc length

Given a closed interval $[a, b] \subseteq I$, the arc length of γ restricted to the interval $[a, b]$ is given by

$$L(\gamma|_{[a,b]}) = \int_a^b \|\gamma'(t)\| dt.$$

Example 1.8. *Returning to one of the curves from Example 1.3, for $\gamma : \mathbb{R} \to \mathbb{R}^3$ given by*

$$\gamma(t) = (3\cos t, 3\sin t, 4t), \quad \forall t \in \mathbb{R},$$

we have

$$\|\gamma'(t)\| = \sqrt{9\sin^2 t + 9\cos^2 t + 16} = 5,$$

and hence

$$L(\gamma|_{[a,b]}) = \int_a^b 5\, dt = 5(b-a).$$

Definition 1.9. *Suppose I, J are open intervals, $\gamma : I \to \mathbb{R}^3$ is a curve, and $f : J \to I$ a continuous function. Then the composition $\tilde{\gamma} := \gamma \circ f : J \to \mathbb{R}^3$ is called a **reparameterisation** of γ. The two curves have the same trace*

$$\gamma(I) = \tilde{\gamma}(J).$$

The following example demonstrates how arc-length is invariant under reparameterisations of *regular* curves.

Example 1.10. *Suppose I, J are open intervals, $\gamma : I \to \mathbb{R}^3$ is a smooth regular curve, and $f : J \to I$ is smooth with $f' > 0$. Define the new regular smooth curve $\tilde{\gamma} : J \to \mathbb{R}^3$ via $\tilde{\gamma} = \gamma \circ f$. It follows from the change of variables formula that*

$$\begin{aligned} L(\tilde{\gamma}|_{[a,b]}) &= \int_a^b |(\gamma \circ f)'(s)|\, ds \\ &= \int_a^b |\gamma'(f(s))| f'(s) ds \\ &= \int_{f(a)}^{f(b)} |\gamma'(t)|\, dt \\ &= L(\gamma|_{[f(a), f(b)]}). \end{aligned}$$

Exercise. *Let $\gamma_1, \gamma_2 : \mathbb{R} \to \mathbb{R}^2$ be the smooth curves*

$$\gamma_1(t) = (t, 0), \quad \gamma_2(t) = (2t^3 - t, 0), \quad \forall t \in \mathbb{R}.$$

Show that $\gamma_1([-1,1]) = \gamma_2([-1,1])$, but $L(\gamma_1|_{[-1,1]}) < L(\gamma_2|_{[-1,1]})$. That is, for non-regular parameterisations of curves, the arc length is dependent on the parameterisation, since we may retrace parts of the curve.

Amongst all reparameterisation of a smooth regular curve, those traversed at unit speed are particularly important.

Definition 1.11. *A smooth regular curve $\gamma : I \to \mathbb{R}^3$ is* ***parameterised by arc-length*** *if*

$$\|\gamma'(s)\| = 1, \quad \forall s \in I.$$

Lemma 1.12. *Every smooth regular curve $\gamma : I \to \mathbb{R}^3$ admits an arc-length reparameterisation.*

Proof. Fix $t_0 \in I$ and define the map $s : I \to \mathbb{R}$ via

$$s(t) := \int_{t_0}^{t} \|\gamma'(x)\| dx. \tag{1.1}$$

Since γ is smooth, s is a smooth map (Exercise) with the image of s an open interval $J \subseteq \mathbb{R}$. As $s'(t) = \|\gamma'(t)\| > 0$, by the inverse function theorem, s admits a smooth inverse $t : J \to I$. Consider the reparameterisation $\tilde{\gamma} := \gamma \circ t : J \to \mathbb{R}^3$. For all $s \in J$, we have

$$\|\tilde{\gamma}'(s)\| = \|\gamma'(t(s))\| \left|\frac{dt}{ds}\right| = \|\gamma'(t)\| \cdot \|\gamma'(t)\|^{-1} = 1. \quad \square$$

Example 1.13. *For the curve $\gamma(t) = (3\cos t, 3\sin t, 4t)$ from Example 1.8, since its speed is constant, we find a simple arc-length parameterisation*

$$\gamma(t) = \left(3\cos\frac{t}{5}, 3\sin\frac{t}{5}, \frac{4t}{5}\right).$$

Example 1.14. *Returning to the curve $\gamma(t) = (t^3 - 4t, t^2 - 4)$ from Example 1.5, an arc-length parameterisation is given by the composition of γ with the inverse of the function*

$$s(t) = \int_0^t \sqrt{9x^4 - 20x^2 + 16}\, dx.$$

This example demonstrates that, although such an arc-length parameterisation always exists, even in simple cases it is difficult to express explicitly.

1.3 Curvature, Torsion, and the Frenet formulas

Given a regular smooth curve $\gamma : I \to \mathbb{R}^3$ parameterised by arc length, we note that differentiating the equation $\|\gamma'\| \equiv 1$ gives the identity

$$\langle \gamma', \gamma'' \rangle = 0.$$

Example 1.15. *Consider $\gamma : \mathbb{R} \to \mathbb{R}^2$ with*

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

This curve respresents a particle moving around the unit circle at unit speed. Differentiating we find

$$\gamma'(t) = (-\sin t, \cos t) \perp (-\cos t, -\sin t) = \gamma''(t), \quad \forall t \in \mathbb{R},$$

or the acceleration of the particle is perpendicular to the velocity. In particular, since the magnitude of the velocity is constant, the acceleration measures precisely the change in direction of the velocity.

The magnitude of the acceleration of a curve parameterised by arc-length is known as the curvature of the curve.

Definition 1.16. *Given a regular smooth curve $\gamma : I \to \mathbb{R}^3$ parameterised by arc length, the number $\kappa(s) := \|\gamma''(s)\| \geq 0$ is called the **curvature** of γ at $s \in I$.*

At a point where $\kappa(s) > 0$, we define the orthogonal unit vectors

$$T(s) := \gamma'(s), \quad N(s) := \frac{\gamma''(s)}{\|\gamma''(s)\|}, \quad B(s) := T(s) \times N(s).$$

We call T the tangent vector, N the normal vector, and B the binormal vector at $\gamma(s)$. Define $\mathrm{Span}\{T(s), N(s)\}$ to be the **osculating plane** at $\gamma(s)$, and the orthonormal frame $\{T(s), N(s), B(s)\}$ the **Frenet frame** at $\gamma(s)$.

Theorem 1.17. *Let $\gamma : I \to \mathbb{R}^3$ be a regular smooth curve parameterised by arc length with curvature $\kappa > 0$ (non-degenerate). Then the Frenet frame $\{T, N, B\}$ satisfies the differential equation*

$$\begin{pmatrix} T \\ N \\ B \end{pmatrix}' = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix} \begin{pmatrix} T \\ N \\ B \end{pmatrix},$$

*where τ is a smooth function called the **torsion** of γ.*

Proof. From the definition of curvature, $T' = \kappa N$. Since N and B have unit length, $N' \perp N$ and $B' \perp B$. In particular, we find that

$$\begin{aligned} N' &= \langle N', T\rangle\, T + \langle N', B\rangle\, B, \\ B' &= \langle B', T\rangle\, T + \langle B', N\rangle\, N. \end{aligned}$$

We first note that, by the product rule for the dot product

$$\langle N', T\rangle = \cancel{\langle N, T\rangle'} - \langle N, T'\rangle = -\kappa.$$

If we define the torsion $\tau := \langle N', B\rangle$, we can conclude that $N' = -\kappa T + \tau B$ as required. Finally, we finish the proof by calculating

$$\begin{aligned} \langle B', T\rangle &= \langle (T \times N)', T\rangle = \langle \cancel{T' \times N}, T\rangle + \langle T \times N', T\rangle = 0, \\ \langle B', N\rangle &= \langle T \times N', N\rangle = \langle T \times (-\kappa T + \tau B), N\rangle = \tau\, \langle T \times B, N\rangle = -\tau. \end{aligned}$$ □

As a consequence of Theorem 1.17, we see that

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

and hence the torsion is a quantative measure of how quickly the binormal vector moves in the direction of the normal vector. Alternatively, one may view the torsion as a measure of twisting of the osculating plane, with a fixed osculating plane along the curve corresponding to zero torsion.

As we shall in this chapter, the torsion and curvature completely characterise the geometry of a curves trace. We begin with the following simple lemma for particularly special cases of curvature and torsion.

Lemma 1.18. *Suppose $\gamma : I \to \mathbb{R}^3$ is a regular curve parameterised by arc length. Then*

(i) $\kappa \equiv 0 \iff \gamma(I)$ *is a straight line.*

(ii) $\kappa > 0$ *and* $\tau \equiv 0 \iff \gamma(I)$ *is a plane curve.*

(iii) $\kappa = r_0^{-1} > 0$ *is constant and* $\tau \equiv 0 \iff \gamma(I)$ *is a circular arc of radius* r_0*.*

Proof. (i) If the curvature vanishes, then $\gamma'' \equiv 0$. This second order differential equation has general solution $\gamma(t) = at + b$ for some $a, b \in \mathbb{R}^3$. The converse is obvious.

(ii) Recall, the torsion vanishing is equivalent to the binormal vector B being fixed. If γ is a plane curve, then the osculating plane and binormal vector B are fixed. Conversely, if B is a fixed vector, fix $t_0 \in I$, and consider the smooth function

$$f(t) := \langle \gamma(t) - \gamma(t_0), B \rangle, \quad \forall t \in I.$$

Differentiating, we have $f'(t) = \langle T, B \rangle = 0$, and as $f(t_0) = 0$, $f \equiv 0$ or $\gamma(t) - \gamma(t_0) \perp B$ for all $t \in I$. Therefore $\gamma(I)$ lies in the plane containing $\gamma(t_0)$ orthogonal to B.

(iii) By the previous part, we may assume $\gamma(I)$ lies in the plane $\{z = 0\}$ with $B \equiv (0, 0, 1)$. Consider the smooth function

$$f(t) = \gamma(t) + r_0 N(t), \quad \forall t \in I.$$

Differentiating

$$f'(t) = T(t) - r_0 \kappa(t) T(t) = 0,$$

and so $f \equiv a$, for some fixed vector $a \in \mathbb{R}^3$. In particular, $\|\gamma - a\| = r_0$. The converse is again obvious. □

Although the formulas for curvature and torsion are expressed simply with respect to an arc-length parameterisation, in practise, such a parameterisation is impractical. The following lemma gives the appropriate formulas for a general regular smooth curve.

Lemma 1.19. *Given a regular smooth curve $\gamma : I \to \mathbb{R}^3$, not necessarily parameterised by arc length, we have the following formulas for the curvature and torsion:*

$$\kappa(t) = \frac{\|\gamma'(t) \times \gamma''(t)\|}{\|\gamma'(t)\|^3}, \quad \tau(t) = \frac{\langle \gamma'(t) \times \gamma''(t), \gamma'''(t) \rangle}{\|\gamma'(t) \times \gamma''(t)\|^2}, \quad \forall t \in I.$$

Proof. Let $s : I \to J$ be as in (1.1) so that $\gamma \circ s^{-1}$ is an arc-length reparameterisation of γ. In particular, $\frac{ds}{dt} = \|\gamma'(t)\|$. Applying the chain rule, we have

$$\gamma'(t) = \frac{d\gamma}{ds}(s(t)) \cdot \frac{ds}{dt} = T(t) \cdot \|\gamma'(t)\|,$$
$$T'(t) = \frac{dT}{ds}(s(t)) \cdot \frac{ds}{dt} = \kappa(t)N(t) \cdot \|\gamma'(t)\|.$$

Combining these formulas, we find that

$$\gamma'' = (\ \|\gamma'\|T\ )' = \|\gamma'\|'T + \|\gamma'\|^2\kappa N,$$

and therefore

$$\gamma' \times \gamma'' = \|\gamma'\|T \times \left(\|\gamma'\|'T + \|\gamma'\|^2\kappa N\right) = \|\gamma'\|^3\kappa B,$$

which taking the length of and rearranging, gives the formula for κ. Next, we note that

$$\gamma''' = \left(\|\gamma'\|'T + \|\gamma'\|^2\kappa N\right)' = \|\gamma'\|''T + \|\gamma'\|'T' + (\kappa\|\gamma'\|^2)'N + \kappa\|\gamma'\|^2N'.$$

Using the Frenet formulas and the chain rule, we have that

$$\kappa\|\gamma'\|^2N' = \kappa\|\gamma'\|^3(-\kappa T + \tau B),$$

and so

$$\gamma''' = fT + gN + \tau\kappa\|\gamma'\|^3B,$$

for some smooth functions $f, g : I \to \mathbb{R}$. Therefore

$$\begin{aligned}\langle\gamma' \times \gamma'', \gamma'''\rangle &= \left\langle\kappa\|\gamma'\|^3B, fT + gN + \tau\kappa\|\gamma'\|^3B\right\rangle \\ &= \tau(\kappa\|\gamma'\|^3)^2 = \tau(\|\gamma' \times \gamma''\|^2).\end{aligned}$$

□

1.4 Isometries of Euclidean space

In order to classify curves inside of $\mathbb{R}^3$, we need to be able to say when two curves are 'equivalent'. For example, if the trace of two different curves are related to one another by a translation of $\mathbb{R}^3$, then despite them having different traces, their geometry is the same, and we would like to identify these two curves as being equivalent. In particular, the two curves are related by an *isometry* of $\mathbb{R}^3$; a bijective map into itself which preserves distances.

Definition 1.20. *An* ***isometry*** *of* $\mathbb{R}^n$ *is a bijective function* $\varphi : \mathbb{R}^n \to \mathbb{R}^n$*, such that*

$$\|\varphi(x) - \varphi(y)\| = \|x - y\|, \quad \forall x, y \in \mathbb{R}^n.$$

Suppose $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry that preserves the origin ($\varphi(0) = 0$). Then, by the polarisation identity, φ must also preserve angles:

$$\langle\varphi(x), \varphi(y)\rangle = \frac{\|\varphi(x)\|^2 + \|\varphi(y)\|^2 - \|\varphi(x) - \varphi(y)\|^2}{2} = \frac{\|x\|^2 + \|y\|^2 - \|x - y\|^2}{2} = \langle x, y\rangle .$$

Suppose $M : \mathbb{R}^n \to \mathbb{R}^n$ is a linear isometry (that is, M is both a linear map and an isometry). We can rewrite the condition that M preserves angles as

$$\left\langle (M^T M - I_n)x, y \right\rangle = \langle Mx, My \rangle - \langle x, y \rangle = 0, \quad \forall x, y \in \mathbb{R}^n,$$

from which it follows that $M^T M = I_n$.

Definition 1.21. *We define the orthogonal group*

$$O(n) := \{M \in \mathbb{R}^{n \times n} : M^T M = I_n\}.$$

In fact, it turns out that modulo translation, every isometry of Euclidean space is precisely an element of the orthogonal group.

Theorem 1.22. *Any isometry $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is of the form*

$$\varphi(x) := Mx + b, \quad \forall x \in \mathbb{R}^n,$$

where $M \in O(n)$ and $b \in \mathbb{R}^n$.

Sketch of Proof. Translations are isometries, and the composition of isometries is an isometry. Therefore, setting $b = \varphi(0)$, we consider the new isometry $\psi(x) := \varphi(x) - b$, which preserves the origin. To see that ψ is a linear map, for any $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, we have

$$\begin{aligned} \|\psi(\lambda x + y) - \lambda\psi(x) - \psi(y)\|^2 &= \|\psi(\lambda x + y)\|^2 + \lambda^2\|\psi(x)\|^2 + \|\psi(y)\|^2 \\ &\quad - 2\langle \psi(\lambda x + y), \lambda\psi(x) + \psi(y)\rangle + 2\lambda \langle \psi(x), \psi(y)\rangle \\ &= \|\lambda x + y\|^2 + \lambda^2\|x\|^2 + \|y\|^2 - 2\langle \lambda x + y, \lambda x + y\rangle + 2\lambda\langle x, y\rangle = 0. \end{aligned}$$

Therefore, ψ is a linear isometry of $\mathbb{R}^n$, and so must be an element of $O(n)$. □

The following lemma shows that isometries preserve the arc length, curvature and torsion of a smooth non-degenerate curve.

Lemma 1.23. *Let $\gamma : I \to \mathbb{R}^3$ be a regular smooth non-degenerate curve parameterised by arc length, and $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ an isometry. Then, $\tilde{\gamma} := \varphi \circ \gamma : I \to \mathbb{R}^3$ is also parameterised by arc length, and has the same curvature and torsion.*

Proof. By the previous theorem

$$\tilde{\gamma}(s) = M\gamma(s) + b, \quad \forall s \in I,$$

for some $M \in O(3)$ and $b \in \mathbb{R}^3$. It follows that for any $s \in I$, we have

$$\|\tilde{\gamma}'(s)\|^2 = \langle M\gamma'(s), M\gamma'(s)\rangle = \left\langle \underbrace{M^T M}_{I_3} \gamma'(s), \gamma'(s) \right\rangle = \|\gamma'(s)\|^2 = 1.$$

Next, since $\tilde{\gamma}$ is also parameterised by arc-length,

$$\tilde{\kappa}(s) = \|\tilde{\gamma}''(s)\| = \|M\gamma''(s)\| = \|\gamma''(s)\| = \kappa(s).$$

To show that the torsion is preserved, we use the formula from Lemma 1.19

$$\tau = \frac{\langle \gamma' \times \gamma'', \gamma''' \rangle}{\|\gamma' \times \gamma''\|^2}.$$

Using that $M \in O(3)$ we have

$$\begin{aligned}
\|\tilde{\gamma}' \times \tilde{\gamma}''\|^2 &= \|\tilde{\gamma}'\|^2 \|\tilde{\gamma}''\|^2 - |\langle \tilde{\gamma}, \tilde{\gamma}'' \rangle|^2 \\
&= \|M\gamma'\|^2 \|M\gamma''\|^2 - |\langle M\gamma, M\gamma'' \rangle|^2 \\
&= \|\gamma'\|^2 \|\gamma''\|^2 - |\langle \gamma, \gamma'' \rangle|^2 \\
&= \|\gamma' \times \gamma''\|^2.
\end{aligned}$$

Given $v_1, v_2, v_3 \in \mathbb{R}^3$, let $[v_1, v_2, v_3]$ denotes the 3×3 matrix whose columns are given by the three vectors. Then

$$\begin{aligned}
\langle \tilde{\gamma}' \times \tilde{\gamma}'', \tilde{\gamma}''' \rangle &= \det[\tilde{\gamma}', \tilde{\gamma}'', \tilde{\gamma}'''] \\
&= \det[M\gamma', M\gamma'', M\gamma'''] \\
&= \det M \cdot \det[\gamma', \gamma'', \gamma'''] \\
&= \langle \gamma' \times \gamma'', \gamma''' \rangle .
\end{aligned}$$

Thus, $\tilde{\tau}(s) = \tau(s)$. □

1.5 Existence and Uniqueness of Linear ODEs

Let $A : I \to \mathbb{R}^{n \times n}$ be a smooth family of $n \times n$ matrices. Fix $t_0 \in I$ and $\gamma_0 \in \mathbb{R}^n$, and consider the initial value problem (IVP)

$$\begin{cases} \gamma'(t) = A(t) \cdot \gamma(t), \quad \forall t \in I, \\ \gamma(t_0) = \gamma_0. \end{cases} \tag{1.2}$$

Theorem 1.24. *There exists a unique smooth solution $\gamma : I \to \mathbb{R}^n$ to the above IVP.*

Sketch of Proof. Without loss of generality, we may assume $t_0 = 0$. Define the following functions inductively: $\gamma_0(t) \equiv \gamma_0$, and

$$\gamma_n(t) := \gamma_0 + \int_0^t A(s)\gamma_{n-1}(s)ds, \quad \forall t \in I, \quad \forall n \in \mathbb{N}.$$

Fix a compact subset $0 \in K \Subset I$. Since A is continuous, $C := \sup_K \|A\| < \infty$, and hence for $t \in K$, we have

$$\begin{aligned}
\|\gamma_{n+1}(t) - \gamma_n(t)\| &\leq \|\int_0^t A(s)(\gamma_n(s) - \gamma_{n-1}(s))ds\| \\
&\leq \left| \int_0^t \|A(s)\| \|\gamma_n(s) - \gamma_{n-1}(s)\| ds \right| \\
&\leq MC \left| \int_0^t \|\gamma_n(s) - \gamma_{n-1}(s)\| ds \right| .
\end{aligned}$$

Setting $L := \sup_K \|\gamma_1 - \gamma_0\| < \infty$ and iterating the above procedure, we find that

$$\begin{aligned}\|\gamma_{n+1}(t) - \gamma_n(t)\| &\leq C^n \left| \int_0^t \int_0^{s_{n-1}} \cdots \int_0^{s_1} \|\gamma_1(s) - \gamma_0\| ds ds_1 \cdots ds_{n-1} \right| \\ &\leq C^n L \left| \int_0^t \int_0^{s_{n-1}} \cdots \int_0^{s_1} ds ds_1 \cdots ds_{n-1} \right| \\ &= \frac{C^n L\,|t|^n}{n!} \leq \frac{C^n L\,|K|^n}{n!}.\end{aligned}$$

Therefore the series $\gamma_{n+1}(t) - \gamma_n(t)$ is locally uniformly absolutely convergent, and hence there exists a function $\gamma_\infty : I \longrightarrow \mathbb{R}^n$, with $\gamma_n \longrightarrow \gamma_\infty$ locally uniformly. In particular, we conclude that

$$\gamma_\infty(t) = \gamma_0 + \int_0^t A(s)\gamma_\infty(s) ds, \quad \forall t \in I,$$

from which it follows easily that γ_∞ is a smooth solution to the IVP.

To show uniqueness, suppose we have two solutions $\gamma, \tilde{\gamma} : I \longrightarrow \mathbb{R}^n$ of the same IVP. Since the ODE is linear, their difference $\eta(t) := \gamma(t) - \tilde{\gamma}(t)$ is a smooth solution to the differential equation $\eta'(t) = A(t) \cdot \eta(t)$, with $\eta(0) = 0$. Note that for $t \in K$, by Cauchy Schwarz we have

$$\frac{d}{dt}\|\eta(t)\|^2 = 2\,\langle A(t)\eta(t), \eta(t)\rangle \leq 2M\|\eta(t)\|^2.$$

Therefore, for $t \in K$ we can conclude that

$$\frac{d}{dt}\left(e^{-2Mt}\|\eta(t)\|^2\right) \leq 0,$$

from which it is easy to see that $\eta \equiv 0$, or equivalently, $\gamma \equiv \tilde{\gamma}$. □

1.6 Fundamental Theorem of Curves

Definition 1.25. *Given two subsets $X, Y \subseteq \mathbb{R}^3$, we say that X and Y are isometric, written $X \cong Y$, if there exists an isometry $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ such that $\varphi(X) = Y$.*

The following theorem states that every non-degenerate regular curve in $\mathbb{R}^3$ is uniquely determined by its curvature and torsion.

Theorem 1.26. *Let $\kappa : I \longrightarrow (0, \infty)$ be a smooth positive function and $\tau : I \longrightarrow \mathbb{R}$ a smooth function. Then, there exists a regular non-degenerate curve $\gamma : I \longrightarrow \mathbb{R}^3$ parameterised by arc length such that the curvature and torsion of γ are precisely the functions κ and τ. Moreover, if $\eta : I \longrightarrow \mathbb{R}^3$ is any other curve parameterised by arc length with the same curvature and torsion, then $\gamma(I) \cong \eta(I)$.*

Proof. To show existence, we first fix $t_0 \in I$ and recall the Frenet formula from Theorem 1.17

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

Note, defining the smooth family of 9×9 matrices $A : I \longrightarrow \mathbb{R}^{9\times 9}$ by

$$A(t) := \begin{pmatrix} 0 & \kappa(t)I_3 & 0 \\ -\kappa(t)I_3 & 0 & \tau(t)I_3 \\ 0 & -\tau(t)I_3 & 0 \end{pmatrix},$$

where I_3 denotes the 3×3 identity matrix, the Frenet formula then corresponds to the ODE

$$\gamma(t)' = A(t)\gamma(t),$$

with

$$\gamma(t) = (T_1, T_2, T_3, N_1, N_2, N_3, B_1, B_2, B_3),$$

the Cartesian coordinates of the Frenet frame put into a single vector in $\mathbb{R}^9$. Therefore, given any initial data point $\gamma(t_0)$, by Theorem 1.24 there exists a solution to the corresponding IVP. We choose our initial data to be

$$\gamma(t_0) = (1, 0, 0, 0, 1, 0, 0, 0, 1),$$

so that the Frenet frame at t_0 is the standard basis of $\mathbb{R}^3$.

Claim. *$\gamma(t)$ defines an orthonormal frame at each time $t \in I$.*

Proof of Claim. Setting $T = (\gamma_1, \gamma_2, \gamma_3)$, $N = (\gamma_4, \gamma_5, \gamma_6)$ and $B = (\gamma_7, \gamma_8, \gamma_9)$, we see that

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

Note that this linear ODE has a solution

$$T \cdot T = N \cdot N = B \cdot B \equiv 1, \quad T \cdot N = T \cdot B = N \cdot B \equiv 0,$$

with initial condition $(1, 1, 1, 0, 0, 0)$. Therefore, by the uniqueness in Theorem 1.24, this is precisely the solution γ generates of the above system. □

We currently have a Frenet frame satisfying the Frenet formula for the corresponding functions κ and τ. To generate a curve from the Frenet frame, we simply integrate it up. To be more precise, define

$$\gamma(t) := \int_{t_0}^{t} T(s)ds, \quad \forall t \in I,$$

where the integral is evaluated componentwise. By the fundamental theorem of calculus, we find that γ has precisely the curvature κ and torsion τ required.

To show uniqueness, suppose $\gamma, \tilde{\gamma} : I \to \mathbb{R}^3$ are two such curves. Consider their Frenet frames $\{T, N, B\}$, $\{\tilde{T}, \tilde{N}, \tilde{B}\}$. Since $O(3)$ acts transitively on orthonormal frames, there exists $M \in O(3)$ such that at t_0,

$$T = M(\tilde{T}), \quad N = M(\tilde{N}), \quad B = M(\tilde{B}).$$

Since $\{T, N, B\}$ and $\{M(\tilde{T}), M(\tilde{N}), M(\tilde{B})\}$ both solve the same linear ODE with the same initial condition, by the uniqueness in Theorem 1.24, they agree on all of I. Therefore, $(\gamma - M\tilde{\gamma})' \equiv 0$ and so there exists $b \in \mathbb{R}^3$ such that

$$\gamma(t) = M\tilde{\gamma}(t) + b, \quad \forall t \in I. \quad \square$$

2 Regular Surfaces

Recall, for an open subset $U \subseteq \mathbb{R}^n$ and a differentiable function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, the derivative of f at $x \in U$ is the linear map $df(x) : \mathbb{R}^n \to \mathbb{R}^m$, given in Cartesian coordinates by the Jacobian matrix

$$df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdot & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdot & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

Definition 2.1. *For an open subset $U \subseteq \mathbb{R}^n$ and a differentiable function $f : U \to \mathbb{R}^m$, we say that f is an* ***immersion*** *if, at every point $x \in U$, the derivative $df(x) : \mathbb{R}^n \to \mathbb{R}^m$ is an injective linear map.*

Example 2.2. *In the case $n = 1$ and $m = 3$, a smooth curve $\gamma : I \to \mathbb{R}^3$ is an immersion if and only if $\gamma' \neq 0$, which is precisely the definition of γ being regular.*

We see that a (local) parameterisation being an immersion is the correct way to generalise being regular to higher dimensions.

Definition 2.3. *$S \subseteq \mathbb{R}^3$ is a regular surface if, for every $p \in S$, there exists open sets $U \subseteq \mathbb{R}^2$ and $p \in V \subseteq \mathbb{R}^3$, and a smooth map $X : U \to V \cap S \subseteq \mathbb{R}^3$, such that*

(i) X is an immersion:

$$dX(q) : \mathbb{R}^2 \to \mathbb{R}^3 \text{ is injective, for all } q \in U.$$

(ii) X is a homeomorphism:

$$X \text{ is bijective with both } X \text{ and } X^{-1} \text{ continuous.}$$

The map $X : U \to V \cap S$ is called a local parameterisation of S, or local coordinates on S. The set $V \cap S$ is called a local coordinate chart on S. That is, a regular surface is any subset $S \subseteq \mathbb{R}^3$ which can be covered by local coordinate charts.

Remark.

- *In constrast to curves, we have defined a regular surface as a subset of $\mathbb{R}^3$; this would be equivalent to defining curves via their trace. In particular, unlike for regular curves, regular surfaces do* <u>*not*</u> *have points of self-intersection.*
- *Our local coordinates being an immersion ensures the existence of a tangent plane at every point on our surface.*
- *Requiring our local coordinates to be a homeomorphism forces every point in our surface to have a neighbourhood which is topologically equivalent to a small neighbourhood in the plane. This will be essential later for a consistent definition of what it means for a function to be differentiable over a surface.*

2.1 Local Graphs

The simplest examples of regular surfaces are given by a single global parameterisation.

Example 2.4. *Consider the subset* $S = \{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2\}$*. Letting* $X : \mathbb{R}^2 \to S$ *be the smooth function defined by* $X(x, y) = (x, y, x^2 + y^2)$*, we see that*

(i)

$$dX(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2x & 2y \end{pmatrix},$$

which has full rank, so X *is an immersion.*

(ii) $X^{-1} : S \to \mathbb{R}^2$ *is given by* $X^{-1}(x, y, x^2 + y^2) = (x, y)$*, which is continuous. So* X *is a homeomorphism.*

Therefore, S *is a regular surface with a global parameterisation. Defining the function* $f : \mathbb{R}^2 \to \mathbb{R}$ *by* $f(x, y) = x^2 + y^2$*, we see that* $S = \mathrm{Graph}(f)$.

In general, we see that the graph of any smooth function over an open subset of the plane lying in $\mathbb{R}^3$ is a regular surface.

Lemma 2.5. *Let* $U \subseteq \mathbb{R}^2$ *to an open subset and* $f : U \to \mathbb{R}$ *be a smooth function. Then*

$$\mathrm{Graph}(f) = \{(x, y, f(x, y)) \in \mathbb{R}^3 : (x, y) \in U\},$$

is a regular surface.

Proof. Define the smooth function $X : U \to \mathrm{Graph}(f)$ by $X(x, y) = (x, y, f(x, y))$. Then X is a homeomorphism with continuous inverse $X^{-1}(x, y, f(x, y)) = (x, y)$, and X is an immersion with

$$dX(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ f_x(x, y) & f_y(x, y) \end{pmatrix}.$$

Therefore, $\mathrm{Graph}(f)$ is a regular surface with a single global coordinate chart. □

Most regular surfaces do not admit a single global parameterisation like a graph does.

Example 2.6. *Consider the unit sphere*

$$\mathbb{S}^2 := \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}.$$

Setting $D = \{(u, v) \in \mathbb{R}^2 : u^2 + v^2 < 1\}$ *to be the unit disk in the plane, we have local charts* $X_1, X_2 : D \to \mathbb{R}^3$ *defined by*

$$X_1(u, v) = (u, v, \sqrt{1 - u^2 - v^2}),$$
$$X_2(u, v) = (u, v, -\sqrt{1 - u^2 - v^2}).$$

These charts cover everything on the unit sphere except for the equator. In order to cover the entire sphere, consider the same charts under rotations of the sphere: $X_3, X_4, X_5, X_6 : D \to \mathbb{R}^3$ *defined by*

$$\begin{aligned}
X_3(u,v) &= (u, \sqrt{1-u^2-v^2}, v),\\
X_4(u,v) &= (u, -\sqrt{1-u^2-v^2}, v),\\
X_5(u,v) &= (\sqrt{1-u^2-v^2}, u, v),\\
X_6(u,v) &= (-\sqrt{1-u^2-v^2}, u, v).
\end{aligned}$$

Example 2.7. *There are multiple ways to take local coordinates on a surface. Lets consider again the sphere* $\mathbb{S}^2$*. Using spherical coordinates, we can find a chart on the sphere covering everything but half a great circle. That is, take*

$$\begin{aligned}
U &= \{(\theta,\varphi) \in \mathbb{R}^2 : \theta \in (0,2\pi), \varphi \in (0,\pi)\},\\
V &= \mathbb{R}^3 \setminus \{(x,0,z) \in \mathbb{R}^3 : x \geq 0\},
\end{aligned}$$

and $X : U \to V \cap \mathbb{S}^2$ *by*

$$X(\theta,\varphi) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi).$$

Checking, we find that X *is smooth with*

$$\begin{aligned}
\frac{\partial X}{\partial\theta} &= (-\sin\varphi\sin\theta, \sin\varphi\cos\theta, 0),\\
\frac{\partial X}{\partial\varphi} &= (\cos\varphi\cos\theta, \cos\varphi\sin\theta, -\sin\varphi),
\end{aligned}$$

which are linearly dependent iff $\sin\varphi = 0$ *iff* $\varphi \in \pi\mathbb{Z}$*. Therefore,* X *is an immersion on* U*.*

Exercise. *Show that* X *is a homeomorphism.*

We can then perform the same trick as before of rotating the sphere to find two more charts which completely cover the sphere.

Example 2.8. *There is another very important example of a pair of charts on* $\mathbb{S}^2$ *which each cover all of the sphere except a single point. These are known as stereographic projection.*

Define the charts $\varphi_1, \varphi_2 : \mathbb{R}^2 \to \mathbb{S}^2$ *via the formulas*

$$\begin{aligned}
X_1(u,v) &= \left(\frac{2u}{u^2+v^2+1}, \frac{2v}{u^2+v^2+1}, \frac{u^2+v^2-1}{u^2+v^2+1}\right),\\
X_2(u,v) &= \left(\frac{2u}{1+u^2+v^2}, \frac{2v}{1+u^2+v^2}, \frac{1-u^2-v^2}{1+u^2+v^2}\right).
\end{aligned}$$

Exercise. *Derive these formulas from the geometry and prove they define charts on* $\mathbb{S}^2$*.*

From the previous examples we see that the unit sphere is a regular surface which cannot be represented by a single graph. However, being able to be locally represented by a graph is indeed a characterisation of regular surfaces.

Lemma 2.9. *Let $S \subseteq \mathbb{R}^3$ be a regular surface and $p \in S$. Then there exists an open subset $V \subseteq \mathbb{R}^3$ with $p \in V$, such that $S \cap V$ is the graph of a function of two variables.*

Proof. Fix $p \in S$ and let $X : U \to V \cap S$ be local coordinates on a neighbourhood of p, with $q \in U$ such that $X(q) = p$. We decompose X using Cartesian coordinates on U and V in the form

$$X(u, v) = (x(u, v), y(u, v), z(u, v)), \quad \forall (u, v) \in U.$$

Since $dX(q)$ is injective, the vectors $\frac{\partial X}{\partial u}(q), \frac{\partial X}{\partial v}(q) \in \mathbb{R}^3$ are linearly independent. In particular, one of the three submatrices

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}, \quad \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{pmatrix}, \quad \begin{pmatrix} \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{pmatrix},$$

must have non-zero determinant, where all three are evaluated at the point $q \in U$. Without loss of generality, let us assume that the first submatrix has non-zero determinant. In particular, if we define $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ to be the projection map onto the first two coordinates $\pi(x, y, z) = (x, y)$, then $\pi \circ X : U \to \mathbb{R}^2$ has invertible derivative

$$d(\pi \circ X)(q) = d\pi(X(q)) \circ dX(q) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

By the inverse function theorem, after possibly shrinking U and V if necessary, there exists a smooth inverse $(\pi \circ X)^{-1} : \pi(V) \to U$. Note that $\pi : V \to \pi(V)$ is now bijective, with

$$\pi^{-1} = X \circ (\pi \circ X)^{-1}.$$

Therefore, defining the function $f : \pi(V) \to \mathbb{R}$ to be

$$f(x, y) := z((\pi \circ X)^{-1}(x, y)),$$

we have $S \cap V = \mathrm{Graph}(f)$. □

Combining Lemmas 2.5 and 2.9:

S is a regular surface iff S is locally a graph.

Lemma 2.9 can also be helpful in showing that a given subset is not a regular surface.

Example 2.10. *The cone $\{(x, y, z) \in \mathbb{R}^3 : z = \sqrt{x^2 + y^2}\}$ is not a regular surface. If it were, then in a neighbourhood of the origin, it would be a graph of a smooth function of two variables by Lemma 2.9. Since its projection onto the $\{x = 0\}$ and $\{y = 0\}$ planes is not injective, it must be a graph over the $\{z = 0\}$ plane. That is, locally near $(x, y) = (0, 0)$, the cone is given by the graph of $f(x, y)$. However, it must be that $f(x, y) = \sqrt{x^2 + y^2}$ which is not smooth at the origin. So the cone cannot be a regular surface.*

The following technical lemma follows from the same idea as in the proof of Lemma 2.9. It states that, if we know that S is a regular surface a priori, then for a smooth immersion X to be a coordinate chart, we only need to check that X is bijective - we do not need to also check that X^{-1} is continuous.

Lemma 2.11. *Let $S \subseteq \mathbb{R}^3$ be a regular surface. If $X : U \subseteq \mathbb{R}^2 \to S \subseteq \mathbb{R}^3$ is a smooth injective immersion, then it is a homeomorphism onto its image, and hence a coordinate chart on S.*

Proof. Fix $q \in U$. By the same reasoning as in the proof of Lemma 2.9, there exists open subsets $V_1, V_2 \subseteq \mathbb{R}^2$ with $q \in V_1 \subseteq U$ and $\pi \circ X(q) \in V_2$, such that

$$\pi \circ X : V_1 \to V_2,$$

is a smooth diffeomorphism. Since X is injective, it is a bijection onto its image, with

$$X^{-1} = (\pi \circ X)^{-1} \circ \pi : X(V_1) \to V_1.$$

Therefore, X^{-1} is continuous at $X(q)$, and since q was chosen arbitrarily, X^{-1} is a homeomorphism onto its image as required. □

2.2 Level sets

As we saw in the previous section, the unit sphere $\mathbb{S}^2$ is a regular surface. One natural definition of the unit sphere is as a level set of a smooth function. To be more precise, consider the smooth function $f : \mathbb{R}^3 \to \mathbb{R}$ given by $f(x) = \|x\|^2$. Note that the unit sphere (and indeed all scaled versions of it) are level sets of f. i.e $\mathbb{S}^2 = f^{-1}(1)$, and $f^{-1}(\lambda)$ is a regular surface, for all $\lambda > 0$. However, the level set $f^{-1}(0) = \{0\} \in \mathbb{R}^3$ is <u>not</u> a regular surface. This leads to the following question:

Given a smooth function $f : \mathbb{R}^3 \to \mathbb{R}$, for which values $\lambda \in \mathbb{R}$ are the level sets $f^{-1}(\lambda)$ regular surfaces?

Definition 2.12. *For an open subset $U \subseteq \mathbb{R}^n$ and a differentiable function $f : U \to \mathbb{R}^m$, we say that $x \in U$ is a **critical point** of f if the rank of the linear map $df(x) : \mathbb{R}^n \to \mathbb{R}^m$ is <u>not</u> maximal. The image of a critical point $f(x) \in \mathbb{R}^m$ is called a **critical value** of f. A value $\lambda \in \mathbb{R}^m$ which is not a critical value is called a **regular value** of f. That is, $\lambda \in \mathbb{R}^m$ is a regular value of f if $df(x)$ has maximal rank for every $x \in f^{-1}(\lambda)$.*

Remark. *In the case $n \geq m$, $df(x)$ having maximal rank is equivalent to $df(x)$ being surjective.*

Lemma 2.13. *Let $U \subseteq \mathbb{R}^3$ be an open subset and $f : U \to \mathbb{R}$ be a smooth function. If $\lambda \in \mathbb{R}$ is a regular value of f, then the level set*

$$f^{-1}(\lambda) := \{(x, y, z) \in U : f(x, y, z) = \lambda\},$$

is a regular surface.

Proof. Fix $(x_0, y_0, z_0) \in f^{-1}(\lambda)$. Without loss of generality, we may assume that $f_z(x_0, y_0, z_0) \neq 0$. Therefore, applying the implicit function theorem, there exists open subsets $U \subseteq \mathbb{R}^2$, $\tilde{U} \subseteq \mathbb{R}$, with $(x_0, y_0) \in U$ and $z_0 \in \tilde{U}$, and a smooth function $\varphi : U \to \tilde{U}$ with $\varphi(x_0, y_0) = z_0$, such that

$$f(x, y, \varphi(x, y)) = \lambda, \quad \forall (x, y) \in U.$$

In particular, $f^{-1}(\lambda)$ is given as the graph of the smooth function φ locally about the point (x_0, y_0, z_0), and hence by the results of the previous section, $f^{-1}(\lambda)$ is a regular surface. □

Lemma 2.13 allows us to very efficiently check if certain subsets are regular surfaces. Thanks to Lemma 2.11, this in turns makes finding coordinate charts easier also.

Example 2.14. *For the function $f = \|\cdot\|^2$, we see that*

$$f(x, y, z) = x^2 + y^2 + z^2, \quad df(x, y, z) = (2x, 2y, 2z),$$

which is non-zero away from the origin, which lies in the zero level set. Thus, $f^{-1}(\lambda)$ is a regular surface for all $\lambda > 0$.

Example 2.15. *For $a, b, c > 0$ consider a variation of the previous function*

$$f(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}, \quad df(x, y, z) = (\frac{2x}{a^2}, \frac{2y}{b^2}, \frac{2z}{c^2}),$$

which is again non-zero away from the origin. Thus, the ellipsoids $f^{-1}(\lambda)$ are regular surfaces for all $\lambda > 0$.

Example 2.16. *Consider the smooth function*

$$f(x, y, z) = -x^2 + -y^2 + z^2, \quad df(x, y, z) = (-2x, -2y, 2z).$$

Thus, the hyperboloid of two sheets $f^{-1}(1)$ is a regular surface.

Exercise. *Which quadric surfaces are regular surfaces?*

2.3 Surfaces of Revolution

Suppose $\gamma : \mathbb{R} \to \mathbb{R}^2$ is a smooth regular curve parameterised by arc-length. We say that γ is a **closed curve of length** $L > 0$ if

$$\gamma(s_1) = \gamma(s_2) \iff s_2 - s_1 \in L \cdot \mathbb{Z}.$$

With respect to Cartesian coordinates, we can write $\gamma(s) = (f(s), y(s))$. Since the curve is closed, after possibly translating, we may assume that $f(s) > 0$. We now rotate the curve about the y-axis to generate a surface S. That is, define the coordinate charts $X, Y, Z : (0, L) \times (0, 2\pi) \to \mathbb{R}^3$, by

$$\begin{aligned}
X(s, \theta) &= (f(s) \cos\theta, y(s), f(s) \sin\theta), \\
Y(s, \theta) &= \left(f(s + \frac{L}{3}) \cos(\theta + \frac{\pi}{2}), y(s + \frac{L}{3}), f(s + \frac{L}{3}) \sin(\theta + \frac{\pi}{2}) \right). \\
Z(s, \theta) &= \left(f(s + \frac{2L}{3}) \cos(\theta + \pi), y(s + \frac{2L}{3}), f(s + \frac{2L}{3}) \sin(\theta + \pi) \right).
\end{aligned}$$

We note that these are smooth and cover all of S. We now check that X is a coordinate chart (the arguments for Y and Z are identical). To check that X is an immersion we calculate

$$dX(s,\theta) = \begin{pmatrix} f'(s)\cos\theta & -f(s)\sin\theta \\ y'(s) & 0 \\ f'(s)\sin\theta & f(s)\cos\theta \end{pmatrix}.$$

Note that X fails to be an immersion if and only if at some point (s,θ), all three submatrix of $dX(s,\theta)$ have zero determinant, which is equivalent to the three equations

$$f(s)y'(s)\sin\theta = f(s)^2 f'(s)^2 = f(s)y'(s)\cos\theta = 0.$$

Diving through by $f(s) \neq 0$ and summing the equations together, we find that

$$\|\gamma'(s)\|^2 = f'(s)^2 + y'(s)^2 = 0,$$

which is a contradiction to the original curve being regular. We have shown X is an immersion. Next, it is clear from our choice of domain for our parameterisations that X is bijective onto its image. To show X^{-1} is continuous, we need to show that s and θ are continuous functions of (x,y,z). We first note that s is a continuous function of y and $\sqrt{x^2+z^2}$, and thus s is a continuous function of (x,y,z). To check the continuity of θ, we split our argument into two cases:

$(\theta \neq \pi)$: In this case, we note that

$$\tan\frac{\theta}{2} = \frac{2\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{2\cos^2\frac{\theta}{2}} = \frac{\sin\theta}{1+\cos\theta} = \frac{f(s)\sin\theta}{f(s)+f(s)\cos\theta} = \frac{z}{\sqrt{x^2+z^2}+x},$$

and hence

$$\theta = 2\arctan\frac{z}{\sqrt{x^2+z^2}+x}.$$

$(\theta = \pi)$: We repeat the argument but for the cotangent

$$\cot\frac{\theta}{2} = \frac{2\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{2\sin^2\frac{\theta}{2}} = \frac{\sin\theta}{1-\cos\theta} = \frac{f(s)\sin\theta}{f(s)-f(s)\cos\theta} = \frac{z}{\sqrt{x^2+z^2}-x},$$

and hence

$$\theta = 2\,\mathrm{arccot}\frac{z}{\sqrt{x^2+z^2}-x}.$$

We have shown that surfaces of revolution are regular surfaces.

Example 2.17. *Take $0 < a < b$, and let $\gamma : \mathbb{R} \to \mathbb{R}^2$ be the circle centred at $(b,0)$ of radius a*

$$\gamma(s) := (b + a\cos\frac{s}{a}, \sin\frac{s}{a}), \quad \forall s \in \mathbb{R}.$$

Note, this is a closed curve of length $2\pi a$. Rotating about the y-axis in $\mathbb{R}^3$ generates a torus.

2.4 Differentiable Functions

Consider a function $f : S \to \mathbb{R}$ defined over a regular surface. Since S admits local coordinates, we should be able to use these coordinates to define what it means for such a function f to be smooth. In particular, for each coordinate chart $X : U \to S$, we could consider the function $f \circ X : U \to \mathbb{R}$, which is now from an open subset of the plane to the reals.

At first glance there is a problem with this idea:

What if on a different coordinate chart, the composition is not smooth?

The following lemma rules out such behaviour.

Lemma 2.18. *Let $S \subseteq \mathbb{R}^3$ be a regular surface and $X : U \subseteq \mathbb{R}^2 \to S$, $Y : V \subseteq \mathbb{R}^2 \to S$ be two coordinate charts, with $p \in X(U) \cap Y(V) = W \subseteq S$. Then the change of coordinates function $h : X^{-1} \circ Y : Y^{-1}(W) \to X^{-1}(W)$ is a smooth diffeomorphism. That is, h is a smooth bijection with smooth inverse h^{-1}.*

Proof. Using property (ii) of coordinate charts, h is a composition of homeomorphisms and hence is a homeomorphism itself. We note that we cannot use the same argument to conclude that h is a diffeomorphism, as we do not yet know what it means for X^{-1} to be smooth as a function on a regular surface: X^{-1} is only defined on a codimension one subset, and not on an open set of $\mathbb{R}^3$.

Instead, fix $r \in Y^{-1}(W)$ and let $q = h(r) \in X^{-1}(W)$. Again, if we write our coordinate chart

$$X(u, v) = (x(u, v), y(u, v), z(u, v)), \quad \forall (u, v) \in U,$$

we may assume without loss of generality, that the matrix

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix},$$

is invertible at q.

We now find an invertible extension of X^{-1} defined on a cylindrical neighbourhood over S at $X(q)$. More precisely, we define $\hat{X} : U \times \mathbb{R} \to \mathbb{R}^3$ via

$$\hat{X}(u, v, t) := X(u, v) + (0, 0, t) \in \mathbb{R}^3, \quad \forall t \in \mathbb{R}.$$

Since

$$\det(d\hat{X}(q, 0)) = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \neq 0,$$

by the inverse function theorem, there exists an open set $\tilde{N} \subseteq \mathbb{R}^3$ containing $\hat{X}(q, 0) = X(q)$, an open set $\tilde{U} \subseteq U$ containing q, $\delta > 0$, and a smooth inverse

$$\hat{X}^{-1} : \tilde{N} \to \tilde{U} \times (-\delta, \delta).$$

As $Y(r) = X(q) \in \tilde{N}$, by the continuity of Y, there exists an open neighbourhood $r \in N \subseteq V$ such that $Y(N) \subseteq \tilde{N}$. Therefore we may write

$$h|_N = \hat{X}^{-1} \circ Y|_N,$$

and by the composition of smooth functions, we can conclude that h is smooth at $r \in N$. Since r was arbitrary, h is smooth on $Y^{-1}(W)$. Repeating the exact same argument but swapping X and Y, we conclude that h^{-1} is smooth on $X^{-1}(W)$. □

Remark. *The proof of Lemma 2.18 relies heavily on assumption (ii) in our definition of a coordinate chart. Without the homeomorphism assumption, the lemma would fail and the subsequent definition would not be consistent.*

Definition 2.19. *Let $S \subseteq \mathbb{R}^3$ be a regular surface and $f : S \to \mathbb{R}$ a function. f is said to be smooth at $p \in S$ if, for some coordinate chart $X : U \subseteq \mathbb{R}^2 \to V \cap S \subseteq \mathbb{R}^3$, with $p \in V$, the composition $f \circ X : U \to \mathbb{R}$ to smooth at $X^{-1}(p)$. We say that f is smooth if f is smooth at every $p \in S$.*

Remark. *Using Lemma 2.18, we see that the definition is independent of the coordinate chart chosen. Thus, in the definition above, we could equivalently require that $f \circ X$ is smooth at $X^{-1}(p)$ for every coordinate chart about p.*

Example 2.20. *Let $S \subseteq \mathbb{R}^3$ be a regular surface and $O \subseteq \mathbb{R}^3$ be an open subset with $S \subseteq O$. Suppose $f : O \subseteq \mathbb{R}^3 \to \mathbb{R}$ is a smooth function in the usual sense. It follows from the composition of smooth functions that $f \circ X$ is smooth for any coordinate chart X on S, and hence the restriction $f|_S$ is a smooth function on S.*

Example 2.21. *Fix $v \in \mathbb{R}^3$ and consider the smooth function $f : \mathbb{R}^3 \to R$, given by $f(x) = \langle x, v \rangle$. We call f a height function, as it measures the perpendicular distance from the plane with normal vector v. For any regular surface S, the restriction $f|_S$ is a smooth function on S by Example 2.20.*

Example 2.22. *Fix $x_0 \in \mathbb{R}^3$ and consider the smooth function $f : \mathbb{R}^3 \to R$, given by $f(x) = \|x - x_0\|^2$. f is the distance (squared) function from the point x_0. Again, for any regular surface S, the restriction $f|_S$ is a smooth function on S by Example 2.20.*

Our definition of smooth functions from a surface to the reals can easily be extended to functions between surfaces.

Definition 2.23. *Let $S_1, S_2 \subseteq \mathbb{R}^3$ be a pair of regular surfaces and $f : S_1 \to S_2$. f is said to be smooth at $p \in S_1$ if there exists a pair of coordinate charts $X_1 : U_1 \to V_1 \cap S_1$ and $X_2 : U_2 \to V_2 \cap S_2$ about p and $f(p)$ respectively, such that the composition*

$$X_2^{-1} \circ f \circ X_1 : U_1 \to U_2,$$

is a smooth function at $X_1^{-1}(p)$. f is smooth if f is smooth at every $p \in S_1$.

Definition 2.24. *Let $S_1, S_2 \subseteq \mathbb{R}^3$ be a pair of regular surfaces. A **diffeomorphism** between S_1 and S_2 is a smooth bijection $f : S_1 \to S_2$ with a smooth inverse $f^{-1} : S_2 \to S_1$. If a diffeomorphism between two surfaces exists, we say that S_1 is diffeomorphic to S_2, and denote this by $S_1 \cong S_2$.*

Example 2.25. *If $S \subseteq \mathbb{R}^3$ is a regular surface and $X : U \to S$ is a coordinate chart on S, then Lemma 2.18 implies that $X^{-1} : X(U) \to U$ is smooth, and hence $U \cong X(U)$ for any coordinate chart.*

Remark. *The previous example leads to the following characterisation of regular surfaces:*

$$S \subseteq \mathbb{R}^3 \text{ is a regular surface if and only if it is locally diffeomorphic to } \mathbb{R}^2.$$

2.5 Tangent Planes

Definition 2.26. *Let $S \subseteq \mathbb{R}^3$ be a regular surface and $p \in S$. A* ***tangent vector*** *to S at p is a vector $\gamma'(0) \in \mathbb{R}^3$, where $\gamma : (-\epsilon, \epsilon) \to S$ is a smooth curve in S, with $\gamma(0) = p$. The collection of all tangent vectors to S at p is called the* ***tangent plane*** *of S at p, denoted by $T_pS \subseteq \mathbb{R}^3$.*

Lemma 2.27. *Suppose $X : U \to S$ is a coordinate chart on a regular surface S. Then*

$$T_{X(q)}S = \mathrm{Im}(dX(q)), \quad \forall q \in U.$$

Remark. *Since coordinate charts are immersions, the above lemma implies that the tangent space T_pS is a well-defined two-dimensional subspace of $\mathbb{R}^3$. Moreover, the plane $\mathrm{Im}(dX(X^{-1}(p)))$ is independent of the coordinate chart X.*

Proof. If $w \in T_{X(q)}S$, then there exists $\epsilon > 0$ and a smooth curve $\gamma : (-\epsilon, \epsilon) \to X(U) \subseteq S$ with $\gamma(0) = X(q)$ and $\gamma'(0) = w$. By definition, the curve $\eta := X^{-1} \circ \gamma : (-\epsilon, \epsilon) \to U$ is smooth with $\eta(0) = q$, and therefore, by the chain rule

$$w = \gamma'(0) = (X \circ \eta)'(0) = dX(q) \cdot \eta'(0) \in \mathrm{Im}(dX(q)).$$

On the other hand, given $q \in U$ and $v \in \mathbb{R}^2$, consider the curve $\eta : (-\epsilon, \epsilon) \to U$ defined by

$$\eta(t) = q + tv, \quad \forall t \in (-\epsilon, \epsilon).$$

Then, $\gamma := X \circ \eta : (-\epsilon, \epsilon) \to S$ is a smooth curve with $\gamma(0) = X(q)$, and therefore

$$dX(q) \cdot v = dX(\eta(0)) \cdot \eta'(0) = (X \circ \eta)'(0) = \gamma'(0) \in T_{X(q)}S. \qquad \square$$

Given local coordinates $X : U \to S$ on a neighbourhood of $p \in S$, writing an element of the domain as $(u_1, u_2) \in U$ in Cartesian coordinates, if $q = X^{-1}(p)$, we generate a basis $\{\frac{\partial X}{\partial u_1}(q), \frac{\partial X}{\partial u_2}(q)\}$ of the tangent space T_pS called the basis associated with X. We use the shorthand $\{X_1(q), X_2(q)\}$ for this basis.

Given a vector $w \in T_pS$, there is a smooth curve $\gamma : (-\epsilon, \epsilon) \to S$ with $\gamma(0) = p$ and $\gamma'(0) = w$. The curve $\eta := X^{-1} \circ \gamma : (-\epsilon, \epsilon) \to U$ is then a representation of γ with respect to the coordinate chart X, given by $\eta(t) = (u_1(t), u_2(t))$, with $\eta(0) = q$. Therefore,

$$\begin{aligned} \gamma'(0) &= (X \circ \eta)'(0) \\ &= \frac{d}{dt} X(u_1(t), u_2(t))|_{t=0} \\ &= X_1(q)u_1'(0) + X_2(q)u_2'(0) = w. \end{aligned}$$

That is, in the basis $\{X_1(q), X_2(q)\}$, w has coordinates $(u_1'(0), u_2'(0))$, where $(u_1(t), u_2(t))$ is a parameterisation of the tangent curve γ with respect to the local coordinate chart X.

Consider now a smooth function $f : S_1 \to S_2$ between regular surfaces. Again, for $p \in S_1$ and $w \in T_pS_1$, let $\gamma : (-\epsilon, \epsilon) \to S_1$ be a smooth curve with $\gamma(0) = p$ and $\gamma'(0) = w$. Note that that composition $f \circ \gamma : (-\epsilon, \epsilon) \to S_2$ is a smooth curve in S_2 with $f \circ \gamma(0) = f(p)$, and so the derivative of this curve at zero should give a tangent vector in $T_{f(p)}S_2$. If we could apply the chain rule, then this tangent vector should be equal to the image of w under the derivative of f at p. We therefore make the following definition.

Definition 2.28. *Suppose $S_1, S_2 \subseteq \mathbb{R}^3$ are regular surfaces and $f : S_1 \to S_2$ a smooth function. For each $p \in S_1$, define the derivative of f at p to be the map $df(p) : T_pS_1 \to T_{f(p)}S_2$ given by the formula*

$$df(p) \cdot w := (f \circ \gamma_w)'(0), \quad \forall w \in T_pS_1,$$

where $\gamma_w : (-\epsilon, \epsilon) \to S_1$ denotes a smooth curve with $\gamma_w(0) = p$ and $\gamma_w'(0) = w$.

Before we can be confident in our definition, we need to check that the derivative $(f \circ \gamma_w)'(0)$ used in the definition is independent of our choice of curve γ_w, and instead only depends on the choice of tangent vector w. This is the content of the following lemma.

Lemma 2.29. *Let $S_1, S_2 \subseteq \mathbb{R}^3$ be a pair of regular surfaces, $f : S_1 \to S_2$ a smooth function and $p \in S_1$. Then the derivative of f at p given above is a well-defined linear map.*

Proof. Choose $X : U \to S_1$ and $Y : V \to S_2$ to be coordinate charts about the points p and $f(p)$ respectively. With respect to these coordinates, we can write

$$f(u_1, u_2) = (\underbrace{f_1(u_1, u_2)}_{v_1}, \underbrace{f_2(u_1, u_2)}_{v_2}), \quad \forall (u_1, u_2) \in U,$$

and $\gamma_w(t) = (u_1(t), u_2(t))$. If $q = X^{-1}(p) \in U$ and $r = Y^{-1}(f(p)) \in V$, with respect to the basis $\{X_1(q), X_2(q)\}$ we have $w = (u_1'(0), u_2'(0))$. Moreover, with respect to the basis $\{Y_1(r), Y_2(r)\}$ we have

$$\begin{aligned}
(f \circ \gamma_w)'(0) &= \frac{d}{dt}(f_1(u_1(t), u_2(t)), f_2(u_1(t), u_2(t)))|_{t=0} \\
&= (\frac{\partial f_1}{\partial u_1}(q) \cdot u_1'(0) + \frac{\partial f_1}{\partial u_2}(q) \cdot u_2'(0), \frac{\partial f_2}{\partial u_1}(q) \cdot u_1'(0) + \frac{\partial f_2}{\partial u_2}(q) \cdot u_2'(0)) \\
&= \begin{pmatrix} \frac{\partial f_1}{\partial u_1}(q) & \frac{\partial f_1}{\partial u_2}(q) \\ \frac{\partial f_2}{\partial u_1}(q) & \frac{\partial f_2}{\partial u_2}(q) \end{pmatrix} \begin{pmatrix} u_1'(0) \\ u_2'(0) \end{pmatrix}.
\end{aligned}$$

The formula above demonstrates that $(f \circ \gamma)'(0)$ depends only on $w = (u_1'(0), u_2'(0))$. Moreover, with respect to our bases $\{X_1(q), X_2(q)\}$ on T_pS_1 and $\{Y_1(r), Y_2(r)\}$ on $T_{f(p)}S_2$, we can rewrite our definition in the form

$$df(p) \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} \end{pmatrix}\Bigg|_q \begin{pmatrix} w_1 \\ w_2 \end{pmatrix},$$

for all $w = (w_1, w_2)$. In particular, $df(p)$ is linear. □

3 Geometry of Surfaces

Now that we have the basic definitions under our belt, we may begin to study the geometric properties of a regular surface.

In $\mathbb{R}^3$, the notion of distances and angles is given by the standard dot-product. We also used the dot product at each point in $\mathbb{R}^3$ to define lengths of curves: recall that, given a smooth regular curve $\gamma : I \longrightarrow \mathbb{R}^3$, its arc-length over some compact interval $[a, b] \subseteq I$ is given by the formula

$$L(\gamma|_{[a,b]}) = \int_a^b \|\gamma'(t)\| dt = \int_a^b \langle \gamma'(t), \gamma'(t) \rangle^{\frac{1}{2}} \, dt.$$

At every point of our curve $\gamma(t)$, we look at the size of the vector $\gamma'(t)$ using the dot-product on $\mathbb{R}^3$, and integrate this value over the domain.

3.1 First Fundamental Form

We want the intrinsic geometry of a regular surface $S \subseteq \mathbb{R}^3$ to be inherited from the geometry of its ambient space. More precisely, given any smooth regular curve $\gamma : I \longrightarrow S$, we require that its arc-length inside of S is the same as if we measured it as a curve inside of $\mathbb{R}^3$. As such, we see that for each $p \in S$, the tangent plane T_pS inherits a natural inner-product $\langle \cdot, \cdot \rangle_p$ as a subspace of $\mathbb{R}^3$.

Definition 3.1. *Let $S \subseteq \mathbb{R}^3$ be a regular surface. For each $p \in S$, we define the* ***first fundamental form*** *of S at p to be the non-degenerate quadratic form $g_p : T_pS \longrightarrow [0, \infty)$, given by*

$$g_p(v) := v \cdot v, \quad \forall v \in T_pS.$$

Remark.

- *The first fundamental form is non-denegerate in the sense that $g_p(v) = 0$ if and only if $v = 0$.*
- *The quadratic form uniquely determines a non-degenerate symmetric bilinear form $T_pS \times T_pS \longrightarrow \mathbb{R}$, also denoted by g_p, via the formula*

$$g_p(v, w) := \frac{1}{2} \left(g_p(v + w) - g_p(v) - g_p(w) \right), \quad \forall v, w \in T_pS. \tag{3.1}$$

- *Notice that* ***all*** *of the intrinsic information of the surface (i.e lengths of curves, angles of tangent vectors, areas of regions) is captured by the first fundamental form.*

Suppose $X : U \to S$ is a coordinate chart on a neighbourhood of $p \in S$, with $q = X^{-1}(p) \in U$. For any tangent vector $v \in T_pS$, there exists a smooth curve $\gamma : (-\epsilon, \epsilon) \to U$ such that

$$X \circ \gamma(0) = p, \quad (X \circ \gamma)'(0) = v.$$

Writing $\gamma(t) = (u_1(t), u_2(t))$, we find that

$$\begin{aligned} g_p(v) &= (X \circ \gamma)'(0) \cdot (X \circ \gamma)'(0) \\ &= \left\langle X_1(q)u_1'(0) + X_2(q)u_2'(0), X_1(q)u_1'(0) + X_2(q)u_2'(0) \right\rangle_p \\ &= \langle X_1, X_1 \rangle_p \, u_1'(0)^2 + 2 \langle X_1, X_2 \rangle_p \, u_1'(0)u_2'(0) + \langle X_2, X_2 \rangle_p \, u_2'(0)^2. \end{aligned}$$

In particular, if we let $\{X_1(q), X_2(q)\}$ be the basis of T_pS associated with X, then g_p can be expressed with respect to this basis as the symmetric matrix

$$g_p = \begin{pmatrix} \langle X_1, X_1 \rangle_p & \langle X_1, X_2 \rangle_p \\ \langle X_2, X_1 \rangle_p & \langle X_2, X_2 \rangle_p \end{pmatrix} =: \begin{pmatrix} g_{11}(p) & g_{12}(p) \\ g_{21}(p) & g_{22}(p) \end{pmatrix},$$

so that if $v = (v_1, v_2)$ with respect to this basis, then our equation becomes

$$g_p(v) = g_{11}v_1^2 + 2g_{12}v_1v_2 + g_{22}v_2^2 = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}^T \left.\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}\right|_p \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

This local expression for g can now be used to calculate the length of curves in S using only local coordinates. That is, suppose $\gamma : I \to S$ is a smooth regular curve in S. Let $[a, b] \subseteq I$ and $X : U \to S$ be local coordinates on S such that $\gamma([a, b]) \subseteq X(U)$. Then, if we write

$$\gamma(t) = X(u_1(t), u_2(t)), \quad \forall t \in [a, b],$$

we find that

$$\begin{aligned} L(\gamma|_{[a,b]}) &= \int_a^b \|\gamma'(t)\| dt \\ &= \int_a^b g_{\gamma(t)}(\gamma'(t))^{\frac{1}{2}} dt \\ &= \int_a^b \left(\sum_{i,j=1}^{2} g_{ij}(\gamma(t)) \cdot u_i'(t)u_j'(t) \right)^{\frac{1}{2}} dt. \end{aligned}$$

Example 3.2. *Consider an affine plane $P \subseteq \mathbb{R}^3$ with a global parameterisation $X : \mathbb{R}^2 \to P$ given explicitly by*

$$X(u_1, u_2) = x + u_1 w_1 + u_2 w_2, \quad \forall (u_1, u_2) \in \mathbb{R}^2,$$

where $x \in \mathbb{R}^3$ and $w_1, w_2 \in \mathbb{R}^3$ are orthonormal vectors.

At any point $p \in P$, we see that $X_1 = w_1$ and $X_2 = w_2$. Therefore, our first fundamental form is given in this parameterisation by the identity matrix everywhere

$$g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Example 3.3. *Consider the cylinder $C \subseteq \mathbb{R}^3$ with cross-section the unit circle $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. The cylinder admits a coordinate chart $X : (0, 2\pi) \times \mathbb{R} \to C$ given by*

$$X(u_1, u_2) = (\cos u_1, \sin u_1, u_2) \quad \forall (u_1, u_2) \in (0, 2\pi) \times \mathbb{R}.$$

Notice that

$$X_1 = (-\sin u_1, \cos u_1, 0), \quad X_2 = (0, 0, 1).$$

Therefore,

$$g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and the first fundamental form is also given by the identity matrix.

Example 3.4. *For $\beta > 0$, consider the helix (as in Example 1.3 with $\alpha = 1$)*

$$\gamma(t) = (\cos t, \sin t, \beta t), \quad \forall t \in \mathbb{R}.$$

For each $t \in \mathbb{R}$, consider the line parallel to the plane $\{z = 0\}$ connecting the point $\gamma(t)$ and the z-axis. The surface this generates H is called a helicoid, and admits a global coordinate chart $X : U = \mathbb{R}^2 \to H$ given by

$$X(u_1, u_2) = (u_1 \cos u_2, u_1 \sin u_2, \beta u_2), \quad \forall (u_1, u_2) \in U.$$

Exercise. *Show that H is a regular surface for any $\beta > 0$.*

Since $X_1 = (\cos u_2, \sin u_2, 0)$ and $X_2 = (-u_1 \sin u_2, u_1 \cos u_2, \beta)$, we see that the first fundamental form is given in these coordinates by

$$g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & u_1^2 + \beta^2 \end{pmatrix},$$

The curve $\gamma : \mathbb{R} \to U$ defined in our coordinate system by $\gamma(t) = (\alpha, t)$ will map under X to a curve which has trace the helix with parameters α and β. Since $u_1' = 0$ and $u_2' = 1$, using the first fundamental form, we find its length to be

$$L(\gamma|_{[a,b]}) = \int_0^{2\pi} \sqrt{g_{22}(\alpha, t)} = \int_0^{2\pi} \sqrt{\alpha^2 + \beta^2} = 2\pi\sqrt{\alpha^2 + \beta^2},$$

as expected.

Example 3.5. *Recall from Example 2.7, the unit sphere $\mathbb{S}^2$ admits a coordinate chart $X : U = (0, \pi) \times (0, 2\pi) \to \mathbb{S}^2$ given by spherical coordinates*

$$X(u_1, u_2) = (\sin u_1 \cos u_2, \sin u_1 \sin u_2, \cos u_1), \quad \forall (u_1, u_2) \in U.$$

Since

$$\begin{aligned} X_1(u_1, u_2) &= (\cos u_1 \cos u_2, \cos u_1 \sin u_2, -\sin u_1), \\ X_2(u_1, u_2) &= (-\sin u_1 \sin u_2, \sin u_1 \cos u_2, 0), \end{aligned}$$

it follows that the first fundamental form in this coordinate chart is given by

$$g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 u_1 \end{pmatrix}.$$

Consider the two curves $\gamma, \eta : (-\epsilon, \epsilon) \to U$ *given by*

$$\eta(t) = (c, \pi + t), \quad \gamma(t) = (c + t, \pi + t),$$

for some constant $c \in (0, \pi)$*. Note that, with respect to the basis* $\{X_1(c, \pi), X_2(c, \pi)\}$*, we have*

$$\eta'(0) = (0, 1), \quad \gamma'(0) = (1, 1), \quad g_{(c,\pi)} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 c \end{pmatrix}.$$

and so we have

$$g_{(c,\pi)}(\eta'(0)) = \sin^2 c, \quad g_{(c,\pi)}(\gamma'(0)) = 1 + \sin^2(c), \quad g_{(c,\pi)}(\eta'(0), \gamma'(0)) = \sin^2(c),$$

from which we see that the angle between the curves is given by

$$\theta = \arccos\left(\frac{g_{(c,\pi)}(\eta'(0), \gamma'(0))}{g_{(c,\pi)}(\eta'(0))^{\frac{1}{2}}\, g_{(c,\pi)}(\gamma'(0))^{\frac{1}{2}}} \right) = \arccos\left(\frac{\sin c}{\sqrt{1 + \sin^2 c}} \right).$$

Example 3.6. *Let* $U \subseteq \mathbb{R}^2$ *be an open subset and* $f : U \to \mathbb{R}$ *a smooth function. In Lemma 2.5 we showed that* Graph(f) *is a regular surface with global coordinates*

$$X(u_1, u_2) = (u_1, u_2, f(u_1, u_2)), \quad \forall (u_1, u_2) \in U.$$

As $X_1 = (1, 0, f_1)$ *and* $X_2 = (0, 1, f_2)$*, the first fundamental form is given in these coordinates by*

$$g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} 1 + f_1^2 & f_1 f_2 \\ f_1 f_2 & 1 + f_2^2 \end{pmatrix}.$$

3.2 Area

As well as length of curves in our surface, our first fundamental form can also be used to find the area of subsets within a regular surface.

Let $X : U \to S$ be a coordinate chart on S about a point $p \in S$, and consider the derivative $dX(q) : \mathbb{R}^2 \to T_pS$, where $q = X^{-1}(p)$. Note that $dX(q)$ maps the standard basis vectors $e_1 = (1, 0), e_2 = (0, 1)$ to the vectors $X_1(q), X_2(q) \in T_pS$ respectively.

As such, the image of the unit square spanned by e_1 and e_2 is mapped to the parallelogram spanned by $X_1(q), X_2(q)$, which has area $\|X_1(q) \times X_2(q)\|$, and thus, the infinitesimal area of our surface is given in the local coordinates X by the formula

$$dA = \|X_1(q) \times X_2(q)\| du_1 du_2.$$

Suppose $\Omega \subset X(U) \subseteq S$ is a compact subset within our coordinate chart, or equivalently, $X^{-1}(\Omega)$ is a compact subset of U. Consider the integral of our infinitesimal area over this region

$$\int_{(u_1,u_2)\in X^{-1}(\Omega)} \|X_1(u_1,u_2)\times X_2(u_1,u_2)\| du_1 du_2.$$

We now show that this integral is independent of the choice of coordinate chart X.

Suppose $X : U \longrightarrow S$ and $Y : V \longrightarrow S$ are a pair of coordinate charts on S, and that without loss of generality, $X(U) = Y(V)$ as subsets of S. Given coordinates $(u_1,u_2) \in U$ and $(v_1,v_2) \in V$, we can consider the change of coordinate map $h := Y^{-1} \circ X : U \longrightarrow V$ to be the expression

$$h(u_1,u_2) = (v_1(u_1,u_2), v_2(u_1,u_2)), \quad \forall (u_1,u_2) \in U.$$

Fix $q \in U$ and $r \in V$ so that $X(q) = Y(r)$. Since $X = Y \circ h$, we can apply the chain rule to find that

$$\begin{aligned}
X_1(q) &= dX(q)\cdot e_1 \\
&= dY(r)\cdot dh(q)\cdot e_1 \\
&= dY(r)\cdot \begin{pmatrix} \frac{\partial v_1}{\partial u_1} & \frac{\partial v_1}{\partial u_2} \\ \frac{\partial v_2}{\partial u_1} & \frac{\partial v_2}{\partial u_2} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} | & | \\ Y_1(r) & Y_2(r) \\ | & | \end{pmatrix}\begin{pmatrix} \frac{\partial v_1}{\partial u_1} \\ \frac{\partial v_2}{\partial u_1} \end{pmatrix} \\
&= \frac{\partial v_1}{\partial u_1}(q) Y_1(r) + \frac{\partial v_2}{\partial u_1}(q) Y_2(r).
\end{aligned}$$

Similarly, we find that

$$X_2(q) = \frac{\partial v_1}{\partial u_2}(q) Y_1(r) + \frac{\partial v_2}{\partial u_2}(q) Y_2(r).$$

Therefore, their cross product satisfies

$$\begin{aligned}
X_1 \times X_2 &= \left(\frac{\partial v_1}{\partial u_1} Y_1 + \frac{\partial v_2}{\partial u_1} Y_2\right) \times \left(\frac{\partial v_1}{\partial u_2} Y_1 + \frac{\partial v_2}{\partial u_2} Y_2\right) \\
&= \frac{\partial v_1}{\partial u_1} Y_1 \times \frac{\partial v_2}{\partial u_2} Y_2 + \frac{\partial v_2}{\partial u_1} Y_2 \times \frac{\partial v_1}{\partial u_2} Y_1 \\
&= \left(\frac{\partial v_1}{\partial u_1}\frac{\partial v_2}{\partial u_2} - \frac{\partial v_2}{\partial u_1}\frac{\partial v_1}{\partial u_2}\right) Y_1 \times Y_2 \\
&= \det(dh)\cdot Y_1 \times Y_2,
\end{aligned}$$

and by the change of variable formula for multi-variable integration, we have

$$\begin{aligned}
&\int_{(u_1,u_2)\in X^{-1}(\Omega)} \|X_1(u_1,u_2)\times X_2(u_1,u_2)\| du_1 du_2 \\
=&\int_{(u_1,u_2)\in X^{-1}(\Omega)} \|Y_1(v_1,v_2)\times Y_2(v_1,v_2)\| \, |\det dh(u_1,u_2)| \, du_1 du_2 \\
=&\int_{(v_1,v_2)\in Y^{-1}(\Omega)} \|Y_1(v_1,v_2)\times Y_2(v_1,v_2)\| dv_1 dv_2.
\end{aligned}$$

We have shown that the infinitesimal area form dA is independent of the choice of coordinate chart. In fact, it can be expressed purely in terms of the first fundamental form

$$\begin{aligned}
\|X_1(q)\times X_2(q)\| &= \sqrt{\|X_1(q)\|^2\|X_2(q)\|^2 - \langle X_1(q), X_2(q)\rangle^2} \\
&= \sqrt{g_{11}g_{22} - g_{12}g_{21}} \\
&= \sqrt{\det g_p},
\end{aligned}$$

and hence

$$dA = \sqrt{\det g_{(u_1,u_2)}} du_1 du_2.$$

Definition 3.7. *Let $S \subseteq \mathbb{R}^3$ be a regular surface, $X : U \to S$ be local coordinates on S, and $\Omega \subseteq X(U)$ be a compact subset of S lying in the image of the coordinate chart X. Then the integral*

$$\int_\Omega dA := \int_{(u_1,u_2)\in X^{-1}(\Omega)} \sqrt{\det g_{(u_1,u_2)}} du_1 du_2,$$

is a well-defined (independent of coordinate chart) non-negative real number known as the ***area*** *of Ω.*

Remark. *Although we have only defined the area of subsets contained within a single coordinate chart, for a general subset, we can simply decompose the subset into a disjoint union of subsets, with each component contained within a single coordinate chart. We can then define the area of the original subset to be the sum of the areas of the components.*

Example 3.8. *Let us return to Example 3.6. Since the first fundamental form is given by*

$$g = \begin{pmatrix} 1+f_1^2 & f_1 f_2 \\ f_1 f_2 & 1+f_2^2 \end{pmatrix},$$

we see that the infinitesimal area form is given by

$$\begin{aligned}
dA &= \sqrt{\det g} du_1 du_2 \\
&= \sqrt{(1+f_1^2)(1+f_2^2) - f_1^2 f_2^2} du_1 du_2 \\
&= \sqrt{1+f_1^2+f_2^2} du_1 du_2 \\
&= \sqrt{1+\|df\|^2} du_1 du_2.
\end{aligned}$$

Example 3.9. *Consider the torus T from Example 2.17, with $a = 1$, and $b = 2$. That is, we have the coordinate chart $X : U = (0, 2\pi) \times (0, 2\pi) \to T$ given by*

$$X(u_1, u_2) = ((2 + \cos u_1)\cos u_2, \sin u_1, (2 + \cos u_1)\sin u_2), \quad \forall (u_1, u_2) \in U,$$

which covers everything in the torus except for a meridian and a parallel. Since

$$\begin{aligned} X_1 &= (-\sin u_1 \cos u_2, \cos u_1, -\sin u_1 \sin u_2), \\ X_2 &= (-(2 + \cos u_1)\sin u_2, 0, (2 + \cos u_1)\cos u_2), \end{aligned}$$

it follows that the first fundamental form is given in these coordinates as

$$g = \begin{pmatrix} 1 & 0 \\ 0 & (2 + \cos u_1)^2 \end{pmatrix}.$$

For any $r \in (0, \pi)$, let $\Omega_r := X([r, 2\pi - r]^2) \subseteq T$. By taking the determinant of the first fundamental form, we find that the infinitesimal area is given in these coordinates as

$$dA = \sqrt{\det g_{(u_1,u_2)}} du_1 du_2 = (2 + \cos u_1) du_1 du_2,$$

and hence

$$\begin{aligned} \int_{\Omega_r} dA &= \int_r^{2\pi - r} \int_r^{2\pi - r} (2 + \cos u_1) du_1 du_2 \\ &= 2(\pi - r)\left(2u_1 + \sin u_1\right|_r^{2\pi - r} \\ &= 2(\pi - r)\left(4(\pi - r) + \sin(2\pi - r) - \sin r\right). \end{aligned}$$

By taking $r \downarrow 0$, we see that the area of the entire torus is given by the limit

$$\lim_{r \downarrow 0} \int_{\Omega_r} dA = \lim_{r \downarrow 0} 2(\pi - r)\left(4(\pi - r) + \sin(2\pi - r) - \sin r\right) = 8\pi^2.$$

3.3 Orientability

Let $S \subseteq \mathbb{R}^3$ be a regular surface and $X : U \to S$ local coordinates on S. For each $p \in S$, the tangent space T_pS is a 2-dimensional subspace of $\mathbb{R}^3$, and hence admits a 1-dimensional orthogonal subspace (with respect to the dot-product). In particular, if $q \in U$ is such that $X(q) = p$, then we can take the vector $X_1(q) \times X_2(q)$ as a non-zero vector lying within this orthogonal subspace. Normalising this vector, we get a smooth map $N : X(U) \to \mathbb{R}^3$, given by

$$N_p = \frac{X_1 \times X_2}{\|X_1 \times X_2\|}\Big|_{X^{-1}(p)}, \quad \forall p \in X(U). \tag{3.2}$$

That is, for each $p \in X(U)$, we have a unit normal vector $N_p \perp T_pS$, $\|N_p\| = 1$.

Remark. *If we swapped the order of X_1 and X_2, then we would reverse the sign of N. This detail will be important later.*

We now have a well-defined smooth normal vector locally. This raises the following question

Can we extend this map to the whole of S in a smooth way?

Given a real n-dimensional vector space V, consider the collection of ordered bases of V

$$\mathcal{B} := \{(e_1, \dots, e_n) \in V^n : e_1, \dots, e_n \text{ form a basis of } V\}.$$

Given $b_1 = (e_1, \dots, e_n), b_2 = (f_1, \dots, f_n) \in \mathcal{B}$, there exists a unique linear transformation, known as a change of basis matrix, $A : V \to V$, such that $Ae_i = f_i$, for $i = 1, \dots, n$. i.e, A maps b_1 to b_2. Note that A is invertible and so $\det A \neq 0$. We define an equivalence relation on $\mathcal{B}$ by declaring $b_1 \sim b_2$ iff $\det A > 0$.

Definition 3.10. *The space of orientations on V is defined as the quotient space* $Or(V) := \mathcal{B}/\sim$.

Lemma 3.11. *For any real vector space V, the space of orientations on V is a set of exactly two elements* $Or(V) \simeq \{\pm 1\}$.

Proof. Let $b_+ = (e_1, \dots, e_n)$ be any ordered basis of V. Define a new ordered basis by swapping the first two elements around. That is, let $b_- = (e_2, e_1, e_3, \dots, e_n)$. It follows that the change of basis matrix S from b_+ to b_- satisfies $\det S = -1$, and hence $[b_+], [b_-] \in Or(V)$ are distinct elements. Let $b = (f_1, \dots, f_n)$ be any other ordered basis of V. Let A denote the change of basis matrix from b to b_+. It follows that $S \circ A$ is the change of basis matrix from b to b_-. If $\det A > 0$, then $b \sim b_+$, otherwise $\det A < 0$ and hence $\det SA = \det S \det A = -\det A > 0$, so $b \sim b_-$. Therefore $Or(V) = \{[b_+], [b_-]\} \simeq \{\pm 1\}$. □

Given a regular surface S admitting local coordinates $X : U \to S$ about $p \in S$, the orientation of the tangent space T_pS with respect to X is the choice of orientation corresponding to the ordered basis $\{X_1(q), X_2(q)\}$ (where $q = X^{-1}(p)$). More precisely, we make the choice of orientation

$$[(X_1(q), X_2(q))] \in Or(T_pS).$$

Given different local coordinates $Y : V \to S$ about p, with $Y(r) = p$, we have a different choice of orientation on T_pS with respect to Y. Let $W := X(U) \cap Y(V)$. Note that the two bases are related in the following way

$$\begin{aligned} Y_1(r) &= \frac{\partial u_1}{\partial v_1} X_1(q) + \frac{\partial u_2}{\partial v_1} X_2(q), \\ Y_2(r) &= \frac{\partial u_1}{\partial v_2} X_1(q) + \frac{\partial u_2}{\partial v_2} X_2(q). \end{aligned}$$

In particular, the orientations on T_pS with respect to X and Y agree if and only if $(X_1(q), X_2(q)) \sim (Y_1(r), Y_2(r))$, if and only if

$$\det \begin{pmatrix} \frac{\partial u_1}{\partial v_1} & \frac{\partial u_2}{\partial v_1} \\ \frac{\partial u_1}{\partial v_2} & \frac{\partial u_2}{\partial v_2} \end{pmatrix} > 0.$$

i.e, the Jacobian matrix of the change of coordinates map $X^{-1} \circ Y : Y^{-1}(W) \to X^{-1}(W)$ has positive determinant at the point r.

Remark. *Since the change of coordinate function $X^{-1} \circ Y : Y^{-1}(W) \to X^{-1}(W)$ is a smooth diffeomorphism, the determinant of its Jacobian matrix is a smooth non-zero function on $Y^{-1}(W)$, and hence has locally constant sign. Therefore, if the orientations with respect to X and Y agree at the point $p \in W$, and if W is connected, then they must agree everywhere in $W \subseteq S$.*

Definition 3.12. *A regular surface $S \subseteq \mathbb{R}^3$ is called **orientable** if we can cover S by a collection of local coordinate charts, such that the orientation on each tangent space is independent of the coordinate chart chosen from the collection.*

*A choice of such a collection of charts on S is called an **orientation** on S.*

Example 3.13. *A regular surface given by the graph of a smooth function $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ is an orientable surface, since it has a single global coordinate chart.*

Example 3.14. *The two sphere $\mathbb{S}^2 \subseteq \mathbb{R}^3$ is an orientable surface. To see why, cover the sphere via a pair of stereographic projections*

$$X : \mathbb{R}^2 \to \mathbb{S}^2 \setminus \{N\}, \quad Y : \mathbb{R}^2 \to \mathbb{S}^2 \setminus \{S\},$$

where N, S denote the north pole and south pole respectively. Note that $W := \mathbb{S}^2 \setminus \{N, S\}$ is a connection open subset of the sphere. Consider the change of coordinate function $h := Y^{-1} \circ X : X^{-1}(W) \to Y^{-1}(W)$, given in local coordinates by

$$h = (h_1(u_1, u_2), h_2(u_1, u_2)), \quad \forall (u_1, u_2) \in X^{-1}(W).$$

Fix a point $p \in W$. If $\det(dh|_{X^{-1}(p)}) > 0$, then by our earlier remark, since W is connected, $\det(dh)$ is positive everywhere in $X^{-1}(W)$, and so $\mathbb{S}^2$ is orientable with this pair of charts.

Alternatively, if we find that $\det(dh|_{X^{-1}(p)}) < 0$, then we replace the chart $Y : \mathbb{R}^2 \to \mathbb{S}^2 \setminus \{S\}$ with the chart $\tilde{Y} : \mathbb{R}^2 \to \mathbb{S}^2 \setminus \{S\}$, defined by swapping the coordinates around

$$\tilde{Y}(v_1, v_2) := Y(v_2, v_1), \quad \forall (v_1, v_2) \in \mathbb{R}^2.$$

In particular, we find that the change of coordinates function $\tilde{h} = \tilde{Y}^{-1} \circ X$ is given in local coordinates by

$$\tilde{h} = (h_2(u_1, u_2), h_1(u_1, u_2)),$$

and so we find that $\det(d\tilde{h}|_{X^{-1}(p)}) = -\det(dh|_{X^{-1}(p)}) > 0$, and the same argument as before follows.

The following lemma shows that orientability is a topological property of the surface S; it does not depend on how we embed S into the ambient $\mathbb{R}^3$.

Lemma 3.15. *Let $S_1, S_2 \subseteq \mathbb{R}^3$ be regular surfaces which are smoothly diffeomorphic $S_1 \cong S_2$. Then S_1 is orientable if and only if S_2 is orientable. That is, orientability is a topological invariant.*

Proof. Let $f : S_1 \to S_2$ be a smooth diffeomorphism, and assume S_1 is orientable. Then, from the definition of orientability, we can cover S_1 with charts such that on their intersection, the Jacobian of the change of coordinate functions have positive determinant. For every such chart

$X : U \to S_1$, precompose with the diffeomorphism $f \circ X : U \to S_2$. Since f is a bijection, this new family of coordinate charts covers S_2. Moreover, given any two such charts $f \circ X : U \to S_2$ and $f \circ Y : V \to S_2$, the change of coordinate function is given by

$$(f \circ X)^{-1} \circ (f \circ Y) = X^{-1} \circ \underbrace{f^{-1} \circ f}_{=\mathrm{id}_{S_1}} \circ Y = X^{-1} \circ Y,$$

and its Jacobian has positive determinant. So S_2 is also orientable. □

Example 3.16. *Recall, for $a, b, c > 0$ we showed previously that the unit sphere $\mathbb{S}^2$ is diffeomorphic to the ellipsoid*

$$E := \{(x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\}.$$

Therefore, using Example 3.14, every such ellipsoid E is also orientable.

Before we provide an example of a non-orientable surface, we give a more geometric description of orientability using normal vectors. Note, this characterisation is specific to regular surfaces lying in $\mathbb{R}^3$.

Lemma 3.17. *A regular surface $S \subseteq \mathbb{R}^3$ is orientable if and only if there exists a smooth global choice of unit normal vectors $N : S \to \mathbb{R}^3$. That is $\|N_p\| = 1$, $N_p \perp T_pS$, for every $p \in S$.*

Proof. If S is orientable, for each coordinate chart in our orientable collection $X : U \to S$, define $N^X : X(U) \to \mathbb{R}^3$ as in (3.2)

$$N^X_p = \frac{X_1 \times X_2}{\|X_1 \times X_2\|}|_{X^{-1}(p)}.$$

If $Y : V \to S$ is another coordinate chart in our orientable collection, $p \in X(U) \cap Y(V)$, and if $h := Y^{-1} \circ X$ denotes the change of coordinate function, then

$$N^X_p = \frac{X_1 \times X_2}{\|X_1 \times X_2\|}|_{X^{-1}(p)} \tag{3.3}$$

$$= \frac{\det(dh) Y_1 \times Y_2}{\|\det(dh) Y_1 \times Y_2\|}|_{Y^{-1}(p)} \tag{3.4}$$

$$= \frac{\det(dh)}{|\det(dh)|} N^Y_p = N^Y_p, \tag{3.5}$$

and so the maps agree on their intersection, and hence piece together to give a well-defined global map $N : S \to \mathbb{R}^3$.

Conversely, assume we have a global map $N : S \to \mathbb{R}^3$, and consider local coordinate charts $X : U \to S$ covering S. Note that, after possibly splitting these charts up into more charts, we may always assume U (and hence $X(U)$) is connected. For each such chart, fix $p \in X(U)$. After possibly swapping the coordinates $(u_1, u_2) \in U$, we end up with

$$N_p = \frac{X_1 \times X_2}{\|X_1 \times X_2\|}|_{X^{-1}(p)}.$$

Since the inner product $\left\langle N_\bullet, \frac{X_1\times X_2}{\|X_1\times X_2\|}|_{X^{-1}(\bullet)}\right\rangle : X(U) \to \{\pm 1\}$ is continuous on a connected set, it must be constant, and hence

$$N_p = \frac{X_1 \times X_2}{\|X_1 \times X_2\|}|_{X^{-1}(p)}, \quad \forall p \in X(U).$$

In particular, for any two of these coordinate charts, by a similar calculation to (3.3), we see that the change of coordinate function $h = Y^{-1} \circ X$ must have positive determinant. That is, we have found a collection of coordinate charts covering S which give the same orientation on each tangent space, and hence S is orientable. □

Example 3.18. *For $k \in \mathbb{Z}$, we define the following family of regular surfaces*

$$C_k := \{((2 - v\sin(ku))\sin(2u), (2 - v\sin(ku))\cos(2u), v\cos(ku)) : u \in [0, \pi], v \in (-1, 1)\}.$$

Exercise. *Check that $C_k \subseteq \mathbb{R}^3$ is a regular surface for each $k \in \mathbb{Z}$.*

We first consider C_0 which is the cylinder $\{x^2 + y^2 = 4, |z| < 1\}$. By defining the outward normal vector $N(x, y, z) = \frac{(x,y,0)}{2}$, we see that C_0 is orientable.

Next, consider C_1, which is a Möbius band. We can cover C_1 with two charts $X, Y : (0, \pi) \times (-1, 1) \to C_1$ each omitting a single interval in C_1

$$X(u_1, u_2) = ((2 - u_2\sin(u_1))\sin(2u_1), (2 - u_2\sin(u_1))\cos(2u_1), u_2\cos(u_1)),$$
$$Y(v_1, v_2) = (-(2 - v_2\cos(v_1))\sin(2v_1), -(2 - v_2\cos(v_1))\cos(2v_1), -v_2\sin(v_1)).$$

The intersection of these coordinate charts is disconnected $W = W_1 \sqcup W_2$, where

$$W_1 = X\left((0, \pi/2) \times (-1, 1)\right) = Y((\pi/2, \pi) \times (-1, 1)),$$
$$W_2 = X((\pi/2, \pi) \times (-1, 1)) = Y((0, \pi/2) \times (-1, 1)).$$

The change of coordinate function is given by

$$(v_1, v_2) = (u_1 + \frac{\pi}{2}, u_2), \quad \text{on } W_1,$$
$$(v_1, v_2) = (u_1 - \frac{\pi}{2}, -u_2), \quad \text{on } W_2.$$

Therefore, the Jacobian is positive on W_1, but negative on W_2.

To show that C_1 is non-orientable, suppose for a contradiction that there is a smooth global unit normal field $N : C_1 \to \mathbb{R}^3$. Interchanging u_1 and u_2 if necessary, we may assume

$$N_p = \frac{X_1 \times X_2}{\|X_1 \times X_2\|}|_{X^{-1}(p)}, \quad \forall p \in X(U).$$

Similarly, after potentially swapping v_1 and v_2, we may also assume

$$N_p = \frac{Y_1 \times Y_2}{\|Y_1 \times Y_2\|}|_{Y^{-1}(p)}, \quad \forall p \in Y(U).$$

However, the Jacobian of the change of coordinates must be -1 in either W_1 or W_2. This implies that $N_p = -N_p$ at any point on this component of the intersection, which is a contradiction. Therefore, the Möbius strip C_1 is non-orientable.

Exercise. *For which values of $k \in \mathbb{Z}$ is the surface C_k orientable?*

To finish this section, we show that all of the regular surfaces we constructed in §2 are orientable.

Lemma 3.19. *Suppose $U \subseteq \mathbb{R}^3$ is an open subset, $f : U \to \mathbb{R}$ is a smooth function and $\lambda \in \mathbb{R}$ a regular value of f. Then the regular surface $S = f^{-1}(\lambda)$ is orientable*

Proof. As λ is a regular value of f,

$$\nabla f(p) = (f_x(p), f_y(p), f_z(p)) \in \mathbb{R}^3,$$

is non-zero at every $p \in S$. Moreover, by applying the chain rule, we find that $\nabla f(p) \perp T_pS$ at every $p \in S$. Therefore, we can define a smooth global unit normal field $N : S \to \mathbb{R}^3$ via the map

$$N_p := \frac{\nabla f(p)}{\|\nabla f(p)\|}, \quad \forall p \in S.$$

□

Exercise. *Show that any surface of revolution is orientable.*

3.4 Gauss Map and Shape Operator

In this section, we take $S \subseteq \mathbb{R}^3$ to be an oriented regular surface. That is, S is a orientable regular surface equipped with a specific orientation.

Due to Lemma 3.17, this is equivalent to S coming equipped with a smooth global map $N : S \to \mathbb{R}^3$ such that $\|N_p\| = 1$ and $N_p \perp T_pS$ for every $p \in S$. The condition $\|N_p\| = 1$ for every $p \in S$ is equivalent to saying that the image lies inside the unit sphere $N : S \to \mathbb{S}^2$.

Definition 3.20. *For an oriented regular surface $S \subseteq \mathbb{R}^3$, the map $N : S \to \mathbb{S}^2$ defined above is known as the* ***Gauss map****.*

Remark. *The Gauss map depends on the choice of orientation on the surface; changing the orientation will alter the Gauss map.*

Since locally the Gauss map is defined as in equation (3.2) using coordinate charts from our orientation, it is clear that the Gauss map is a smooth map between regular surfaces. Hence, we can consider its derivative at any point

$$dN_p : T_pS \to T_{N_p}\mathbb{S}^2, \quad \forall p \in S.$$

We first observe that T_pS and $T_{N_p}\mathbb{S}^2$ are both 2-dimensional subspaces of $\mathbb{R}^3$ perpendicular to N_p. It follows that they must be the same space, and so we can think of the derivative as a linear map

$$dN_p : T_pS \to T_pS, \quad \forall p \in S.$$

Definition 3.21. *For an oriented regular surface $S \subseteq \mathbb{R}^3$, the negative derivative of the Gauss map $-dN_p : T_pS \to T_pS$ at $p \in S$ is known as the* ***shape operator*** *at p.*

Remark. *The inclusion of a different sign on the derivative is just a convention.*

Let us unwind the definition slightly to see what the shape operator is measuring. Suppose $\gamma : (-\epsilon, \epsilon) \to S$ is a curve in S with $\gamma(0) = p \in S$. Then the curve $N \circ \gamma : (-\epsilon, \epsilon) \to \mathbb{S}^2$ describes the normal vector to S along γ. Then $-dN_p \cdot \gamma'(0) = -(N \circ \gamma)'(0)$ measures the rate of change of this normal vector along the curve γ at the point p.

For curves, the derivative of the normal vector (with respect to arc-length) is precisely the curvature of the curve $\kappa \in \mathbb{R}$. For surfaces we shall also think of as the curvature as the derivative of the normal vector (i.e the shape operator), however, this is now a 2×2 matrix.

Example 3.22. *Returning to Example 3.2, for an affine plane P with global parameterisation $X : \mathbb{R}^2 \to P$*

$$X(u_1, u_2) = x + u_1 w_1 + u_2 w_2, \quad \forall (u_1, u_2) \in \mathbb{R}^2,$$

where $x \in \mathbb{R}^3$ and $w_1, w_2 \in \mathbb{R}^3$ are orthonormal, the Gauss map $N : P \to \mathbb{S}^2$ is the constant map

$$N_p = w_1 \times w_2, \quad \forall p \in P,$$

so the shape operator $-dN_p$ vanishes everywhere and P has 'no curvature'.

Example 3.23. *Consider $\mathbb{S}^2$ equipped with the orientation corresponding to the Gauss map $N : \mathbb{S}^2 \to \mathbb{S}^2$ given by the identity on the sphere. In particular, $dN_p : T_pS \to T_pS$ is the identity map at any point $p \in \mathbb{S}^2$, and so $\mathbb{S}^2$ has 'constant curvature'*

$$dN_p = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Example 3.24. *Returning to Example 3.18, consider the cylinder*

$$C_0 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 4, |z| < 1\}.$$

With respect to an appropriate orientation, this cylinder has Gauss map $N : C_0 \to \mathbb{S}^2$,

$$N(x, y, z) = \frac{1}{2}(x, y, 0), \quad \forall (x, y, z) \in C_0.$$

Along any curve $\gamma(t) = (x(t), y(t), z(t)) \in C_0$ with $\gamma(0) = p$, we have

$$N \circ \gamma(t) = \frac{1}{2}(x(t), y(t), 0),$$

and hence

$$dN_p(x'(0), y'(0), z'(0)) = (N \circ \gamma)'(0) = \frac{1}{2}(x'(0), y'(0), 0).$$

Since T_pC_0 is spaned by a pair of vectors v_1, v_2, where v_1 is parallel to the $\{z = 0\}$ plane and $v_2 = (0, 0, 1)$, we find that $dN_p \cdot v_1 = \frac{1}{2}v_1$, and $dN_p v_2 = 0$. Thus, with respect to this basis of T_pC_0, the cylinder has shape operator

$$-dN_p = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}.$$

Example 3.25. *Consider the hyperbolic paraboloid $H = \{z = y^2 - x^2\}$ which is a graph given by the global parameterisation*

$$X(u_1, u_2) = (u_1, u_2, u_2^2 - u_1^2), \quad \forall (u_1, u_2) \in \mathbb{R}^2.$$

With respect to this parameterisation we have

$$X_1 = (1, 0, -2u_1), \quad X_2 = (0, 1, 2u_2),$$

and hence the Gauss map is

$$N_{(u_1,u_2)} = \left(\frac{2u_1}{\sqrt{1 + 4u_1^2 + 4u_2^2}}, \frac{-2u_2}{\sqrt{1 + 4u_1^2 + 4u_2^2}}, \frac{1}{\sqrt{1 + 4u_1^2 + 4u_2^2}} \right).$$

Consider a curve $\gamma(t) = X(u_1(t), u_2(t))$ in H with $\gamma(0) = (0, 0, 0) = p$. Then

$$\gamma'(0) = X_1 u_1'(0) + X_2 u_2'(0) = (u_1'(0), u_2'(0), 0),$$

in Cartesian coordinates. Therefore

$$(N \circ \gamma)'(0) = (2u_1, -2u_2, 0)\,,$$

and so with respect to the basis $X_1 = (1, 0, 0), X_2 = (0, 1, 0)$, we have

$$-dN_p = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Lemma 3.26. *The shape operator $-dN_p : T_pS \to T_pS$ at a point $p \in S$ is a self-adjoint linear map. That is*

$$\left\langle dN_p \cdot w_1, w_2 \right\rangle = \left\langle w_1, dN_p \cdot w_2 \right\rangle, \quad \forall w_1, w_2 \in T_pS. \tag{3.6}$$

Proof. Since the shape operator is linear, it suffices to check (3.6) for a single basis $\{w_1, w_2\}$ of T_pS. Lets choose $w_1 = X_1$, $w_2 = X_2$ for some local coordinates $X : U \to S$ about p.

If $\gamma : (-\epsilon, \epsilon) \to X(U) \subseteq S$ is a smooth curve with $\gamma(0) = p$, expressing it in local coordinates as

$$\gamma(t) = X(u_1(t), u_2(t)), \quad \forall t \in (-\epsilon, \epsilon),$$

we find that

$$\begin{aligned} dN_p(X_1 u_1'(0) + X_2 u_2'(0)) &= dN_p \cdot \gamma'(0) \\ &= \frac{d}{dt} \left(N(u_1(t), u_2(t))\right|_{t=0} \\ &= N_1 u_1'(0) + N_2 u_2'(0). \end{aligned}$$

In particular, we have $dN_p \cdot X_1 = N_1$ and $dN_p \cdot X_2 = N_2$. Thus, it suffices to show that

$$\langle N_1, X_2 \rangle = \langle X_1, N_2 \rangle\,.$$

As we have $\langle N, X_1\rangle = \langle N, X_2\rangle = 0$ locally in $X(U)$, differentiating these quantities yields

$$0 = \frac{\partial}{\partial u_2}\left(\langle N, X_1\rangle\right) = \langle N_2, X_1\rangle + \langle N, X_{12}\rangle,$$
$$0 = \frac{\partial}{\partial u_1}\left(\langle N, X_2\rangle\right) = \langle N_1, X_2\rangle + \langle N, X_{21}\rangle.$$

Since X is smooth, partial derivatives commute, and hence

$$\langle N_2, X_1\rangle = -\left\langle N, X_{12}\right\rangle = -\left\langle N, X_{21}\right\rangle = \langle N_1, X_2\rangle,$$

as required. □

Let $\{v_1, v_2\}$ be an orthonormal basis of T_pS. As the shape operator is self-adjoint, the matrix of dN_p with respect to this basis is symmetric:

$$(dN_p)_{ij} = \left\langle dN_p \cdot v_i, v_j\right\rangle = \left\langle v_i, dN_p \cdot v_j\right\rangle = (dN_p)_{ji}.$$

Therefore, by standard Linear Algebra, dN_p can be diagonalised by an orthonormal basis of eigenvectors. i.e. there exists constants $\kappa_1, \kappa_2 \in \mathbb{R}$ and an orthonormal basis $\{e_1, e_2\}$ of T_pS such that, with repsect to this basis

$$-dN_p = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix}. \tag{3.7}$$

Definition 3.27. *$\kappa_1, \kappa_2 \in \mathbb{R}$ as defined above are known as the* ***principal curvatures*** *of S at p.*

Remark. *Without loss of generality, we always assume $\kappa_1 \geq \kappa_2$.*

There are precisely two invariant polynomials on the space of 2×2 matrices (under conjugation by $GL(2, \mathbb{R})$) - the determinant and the trace. In each case, they can be written explicitly in terms of the principal curvatures.

Definition 3.28. *Let $S \subseteq \mathbb{R}^3$ be an oriented regular surface with Gauss map $N : S \to \mathbb{S}^2$. At any point $p \in S$, we define the* ***Gaussian curvature*** *of S at p to be the determinant of the shape operator*

$$K := \det(-dN_p) = \kappa_1\kappa_2,$$

and the ***mean curvature*** *of S at p to be the one half of the trace of the shape operator*

$$H := \frac{1}{2}\operatorname{Tr}(-dN_p) = \frac{\kappa_1 + \kappa_2}{2}.$$

Remark. *Note that, although switching the orientation on S will potentially reverse the sign of the mean curvature, the Gaussian curvature is independent of the orientation chosen.*

Example 3.29. *Returning to our earlier examples (equipped with the appropriate orientations) from before, we see that for the*

- *Affine plane P, $\kappa_1 = \kappa_2 = 0$, $H = 0$, $K = 0$;*
- *Sphere $\mathbb{S}^2$, $\kappa_1 = \kappa_2 = -1$, $K = 1$, $H = -1$;*
- *Cylinder C_0, $\kappa_1 = 0$, $\kappa_2 = -\frac{1}{2}$, $K = 0$, $H = -\frac{1}{2}$;*
- *Hyperbolic paraboloid H, $\kappa_1 = 2$, $\kappa_2 = -2$, $K = -4$, $H = 0$.*

3.5 Second Fundamental Form

Since the shape operator is self-adjoint, this allows us to consider the quadratic form associated to it at each point.

Definition 3.30. *Let $S \subseteq \mathbb{R}^3$ be an oriented regular surface with Gauss map $N : S \to \mathbb{S}^2$. For each $p \in S$, the **second fundamental form** of S at p is the quadratic form $h_p : T_pS \to \mathbb{R}$, given by*

$$h_p(v) := \left\langle -dN_p \cdot v, v \right\rangle, \quad \forall v \in T_pS.$$

Remark. *Unlike the first fundamental form, note that the second fundamental form can be degenerate.*

Recall, if $\{e_1, e_2\}$ is our orthonormal basis of eigenvectors for the shape operator $-dN_p$, so that (3.7) holds with respect to this basis, then for any $v \in T_pS$, writing $v = v_1e_1 + v_2e_2$ for some $v_1, v_2 \in \mathbb{R}$, we have

$$h_p(v) = \langle \kappa_1 v_1 e_1 + \kappa_2 v_2 e_2, v \rangle = \kappa_1 v_1^2 + \kappa_2 v_2^2.$$

Thus, we find that the principal curvatures are precisely the maximum and minimum values of the second fundamental form on the unit circle in T_pS:

$$\begin{aligned} \kappa_1 &= \max\{h_p(v) : v \in T_pS,\ \|v\| = 1\}, \\ \kappa_2 &= \min\{h_p(v) : v \in T_pS,\ \|v\| = 1\}. \end{aligned}$$

As with the first fundamental form, we find an expression for h using local coordinates. That is, suppose $X : U \to S$ are local coordinates about $p \in S$. Using the basis of T_pS associated with X, we find that

$$\begin{aligned} h_p(X_1) &= \left\langle -dN_p \cdot X_1, X_1 \right\rangle = -\langle N_1, X_1 \rangle = \langle N, X_{11} \rangle, \\ h_p(X_2) &= \left\langle -dN_p \cdot X_2, X_2 \right\rangle = \langle N, X_{22} \rangle, \\ h_p(X_1, X_2) &= \langle N, X_{12} \rangle, \\ h_p(X_2, X_1) &= \langle N, X_{21} \rangle. \end{aligned}$$

and so, since $N = \frac{X_1 \times X_2}{\|X_1 \times X_2\|} = \frac{X_1 \times X_2}{\sqrt{\det g}}$, we have that h_p can be expressed as a matrix with respect to the basis of T_pS associated with X as

$$h_p = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} = \begin{pmatrix} \langle N, X_{11} \rangle & \langle N, X_{12} \rangle \\ \langle N, X_{21} \rangle & \langle N, X_{22} \rangle \end{pmatrix} = \frac{1}{\sqrt{\det g}} \begin{pmatrix} (X_1, X_2, X_{11}) & (X_1, X_2, X_{12}) \\ (X_1, X_2, X_{21}) & (X_1, X_2, X_{22}) \end{pmatrix}, \tag{3.8}$$

where $(v_1, v_2, v_3) := \langle v_1 \times v_2, v_3 \rangle$ denotes the triple product.

Example 3.31. *Consider again the torus T with coordinate chart $X : U = (0, 2\pi) \times (0, 2\pi) \to T$ given by*

$$X(u_1, u_2) = ((2 + \cos u_1)\cos u_2, \sin u_1, (2 + \cos u_1)\sin u_2), \quad \forall (u_1, u_2) \in U,$$

We calculated previously that

$$X_1 = (-\sin u_1 \cos u_2, \cos u_1, -\sin u_1 \sin u_2),$$
$$X_2 = (-(2 + \cos u_1) \sin u_2, 0, (2 + \cos u_1) \cos u_2).$$

Differentiating the vector X again to find the second derivatives, we have

$$\begin{aligned} X_{11} &= (-\cos u_1 \cos u_2, -\sin u_1, -\cos u_1 \sin u_2), \\ X_{12} = X_{21} &= (\sin u_1 \sin u_2, 0, -\sin u_1 \cos u_2), \\ X_{22} &= (-(2 + \cos u_1) \cos u_2, 0, -(2 + \cos u_1) \sin u_2), \end{aligned}$$

and since

$$N = \frac{X_1 \times X_2}{\|X_1 \times X_2\|} = (\cos u_1 \cos u_2, \sin u_1, \cos u_1 \sin u_2)\,,$$

plugging everything into (3.8) *we have*

$$\begin{aligned} h_p &= \begin{pmatrix} \langle N, X_{11}\rangle & \langle N, X_{12}\rangle \\ \langle N, X_{21}\rangle & \langle N, X_{22}\rangle \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 \\ 0 & -(2 + \cos u_1) \cos u_1 \end{pmatrix}. \end{aligned}$$

We now find an expression for the shape operator (and hence the Gaussian and Mean curvatures) purely in terms of the 1st and 2nd fundamental forms:

Let $S \subseteq \mathbb{R}^3$ be an oriented regular surface with Gauss map $N : S \to \mathbb{S}^2$, and fix some local coordinates $X : U \to S$ about $p \in S$.

With respect to the basis $\{X_1, X_2\}$ of T_pS associated to X, we can express the shape operator at p as the 2×2 matrix

$$[-dN_p]_X = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

i.e. $-dN_p \cdot X_i = a_{i1}X_1 + a_{i2}X_2$, for $i = 1, 2$. With respect to the same basis of T_pS, it follows that

$$\begin{aligned} h_{ij} &= \left\langle -dN_p \cdot X_i, X_j \right\rangle \\ &= \left\langle a_{i1}X_1 + a_{i2}X_2, X_j \right\rangle \\ &= a_{i1}g_{1j} + a_{i2}g_{2j}, \end{aligned}$$

for any $i, j \in \{1, 2\}$. In particular, with respect to the basis of T_pS associated to X, we have the matrix relation

$$[h_p]_X = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = [-dN_p]_X \cdot [g_p]_X.$$

Since g_p is non-degenerate, its matrix with respect to the basis associated with X, $[g_p]_X$ is invertible, and hence

$$[-dN_p]_X = [h_p]_X \cdot [g_p]_X^{-1}. \tag{3.9}$$

Taking the determinant and trace of (3.9), we have the following lemma.

Lemma 3.32. *If $S \subseteq \mathbb{R}^3$ is an oriented regular surface with local coordinates $X : U \to S$ about $p \in S$, then the Gaussian curvature at p is given by*

$$K(p) = \frac{h_{11}h_{22} - h_{12}h_{21}}{g_{11}g_{22} - g_{12}g_{21}},$$

and the Mean curvature at p given by

$$H(p) = \frac{h_{11}g_{22} - h_{12}g_{21} - h_{21}g_{12} + h_{22}g_{11}}{2(g_{11}g_{22} - g_{12}g_{21})},$$

where g_p and h_p are expressed with respect to the local coordinates X.

Proof. Taking the determinant of (3.9) yields

$$\begin{aligned} K(p) &= \det[-dN_p]_X \\ &= \det([h_p]_X \cdot [g_p]_X^{-1}) \\ &= \det[h_p]_X \det([g_p]_X)^{-1} \\ &= \frac{h_{11}h_{22} - h_{12}h_{21}}{g_{11}g_{22} - g_{12}g_{21}}. \end{aligned}$$

Next, taking the trace of (3.9) yields

$$\begin{aligned} 2H(p) &= \operatorname{tr}[-dN_p]_X \\ &= \operatorname{tr}([h_p]_X \cdot [g_p]_X^{-1}) \\ &= \frac{1}{\det[g_p]_X} \operatorname{tr}\left(\begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{pmatrix}\right) \\ &= \frac{1}{g_{11}g_{22} - g_{12}g_{21}} \operatorname{tr}\begin{pmatrix} h_{11}g_{22} - h_{12}g_{21} & -h_{11}g_{12} + h_{12}g_{11} \\ h_{21}g_{22} - h_{22}g_{21} & -h_{21}g_{12} + h_{22}g_{11} \end{pmatrix} \\ &= \frac{h_{11}g_{22} - h_{12}g_{21} - h_{21}g_{12} + h_{22}g_{11}}{g_{11}g_{22} - g_{12}g_{21}}. \end{aligned}$$

□

Example 3.33. *For the torus T from Example 3.31, we have that in local coordinates*

$$\begin{aligned} g_p &= \begin{pmatrix} 1 & 0 \\ 0 & (2 + \cos u_1)^2 \end{pmatrix}, \\ h_p &= \begin{pmatrix} -1 & 0 \\ 0 & -(2 + \cos u_1)\cos u_1 \end{pmatrix}. \end{aligned}$$

Plugging these into our formulas, we conclude that

$$\begin{aligned} K(p) &= \frac{\cos u_1 (2 + \cos u_1)}{(2 + \cos u_1)^2} = \frac{\cos u_1}{2 + \cos u_1}, \\ H(p) &= \frac{-(2 + \cos u_1)^2 - \cos u_1(2 + \cos u_1)}{2(2 + \cos u_1)^2} = -\frac{1 + \cos u_1}{2 + \cos u_1}. \end{aligned}$$

In this example, we note that

$$\begin{aligned}\int_T K dA &= \int_0^{2\pi}\int_0^{2\pi} K\sqrt{\det g}\, du_1 du_2 \\ &= \int_0^{2\pi}\int_0^{2\pi} \cos u_1 du_1 du_2 = 0.\end{aligned}$$

That is, the torus has total curvature zero. In fact, as we shall see at the end of the course, this is true regardless of the way we embed a torus into $\mathbb{R}^3$*; if* S *is any regular surface diffeomorphic to* T*, then* $\int_S K dA = 0$ *also.*

4 Curvature

4.1 Gaussian Curvature

In §3 we defined all of the operators and quantities describing the geometry of a regular surface. We now investigate the geometric significance of the Gaussian curvature $K : S \to \mathbb{R}$. We begin with the following definition regarding the sign of K at each point.

Definition 4.1. *A point p on a regular surface S is called:*

1. *elliptic if $K(p) > 0$;*
2. *hyperbolic if $K(p) < 0$;*
3. *parabolic if $K(p) = 0$, but $dN_p \neq 0$;*
4. *planar if $dN_p = 0$.*

Warning: Contrary to the name, the shape operator vanishing at a single point does not imply that the surface is a plane.

Example 4.2. *Consider the surface S given by the graph of the smooth function $(x, y) \mapsto (x^2+y^2)^2$. This is a regular surface with global coordinate chart*

$$X(u, v) = (u, v, (u^2 + v^2)^2).$$

It follows that

$$X_u = (1, 0, 4u(u^2 + v^2)), \quad X_v = (0, 1, 4v(u^2 + v^2)),$$

$$X_{uu} = (0, 0, 12u^2 + 4v^2), \quad X_{uv} = (0, 0, 8uv), \quad X_{vv} = (0, 0, 12v^2 + 4u^2).$$

Therefore, the point $p = (0, 0, 0)$ is planar.

Since the Gaussian curvature is the product of the principal curvatures, its sign indicates whether the signs of the principal curvatures agree. In particular, this tells us geometrically how the surface bends locally about the tangent plane.

Lemma 4.3. *Let p be a point in a regular surface S. If p is an elliptic point, then there exists an open set V in S containing p such that all of the points inside of V lay on the same side of the affine plane $p + T_pS$. If p is a hyperbolic point, then for any open set V in S containing p, there exists points in V on either side of $p + T_pS$.*

Proof. Let $X : U \to S$ be local coordinates about p with $X(0,0) = p$, and normal vector N_p. Define the signed distance to $p + T_pS$ by $d : U \to \mathbb{R}$,

$$d(u_1, u_2) := \left\langle X(u_1, u_2) - p, N_p \right\rangle .$$

Since X is smooth, we can approximate it about the origin using Taylor's theorem

$$X(u_1, u_2) = \underbrace{X(0,0)}_{p} + X_1 u_1 + X_2 u_2 + \frac{1}{2}\left(X_{11}u_1^2 + 2X_{12}u_1u_2 + X_{22}u_2^2\right) + \underbrace{\varepsilon(u_1, u_2)}_{o(u_1^2+u_2^2)},$$

where the error function $\varepsilon(u_1, u_2)$ is a smooth function $U \to \mathbb{R}^3$ such that

$$\lim_{(u_1,u_2)\to(0,0)} \frac{\varepsilon(u_1, u_2)}{u_1^2 + u_2^2} = 0.$$

It follows that the signed distance is given by

$$\begin{aligned} d(u_1, u_2) &= \frac{1}{2}\left(\left\langle X_{11}, N_p\right\rangle u_1^2 + 2\left\langle X_{12}, N_p\right\rangle u_1u_2 + \left\langle X_{22}, N_p\right\rangle u_2^2\right) + \left\langle \varepsilon(u_1, u_2), N_p\right\rangle \\ &= \frac{1}{2}\left(h_{11}u_1^2 + 2h_{12}u_1u_2 + h_{22}u_2^2\right) + \left\langle \varepsilon(u_1, u_2), N_p\right\rangle \\ &= \frac{1}{2}h_p(w) + \left\langle \varepsilon(u_1, u_2), N_p\right\rangle \\ &= \frac{\|w\|^2}{2}\left(h_p\left(\frac{w}{\|w\|}\right) + \left\langle \frac{2\varepsilon(u_1, u_2)}{u_1^2 + u_2^2}, N_p\right\rangle\right), \end{aligned}$$

where $w = u_1X_1 + u_2X_2 \in T_pS$.

If p is elliptic, the signs of κ_1 and κ_2 agree, and hence $h_p(w/\|w\|)$ has a fixed sign. Without loss of generality, assume that the principal curvatures are positive, and so $h_p(w/\|w\|) \geq \kappa_2 > 0$ for any non-zero $w \in T_pS$. After possibly shrinking U, we can assume that

$$\left|\left\langle \frac{2\varepsilon(u_1, u_2)}{u_1^2 + u_2^2}, N_p\right\rangle\right| \leq \kappa_2,$$

and hence $d \geq 0$ on U. The conclusion for p elliptic follows by taking $V = X(U)$.

If p is hyperbolic, then $\kappa_1 > 0$ and $\kappa_2 < 0$, and we may again shrink U so that

$$\left|\left\langle \frac{2\epsilon(u_1, u_2)}{u_1^2 + u_2^2}, N_p\right\rangle\right| \leq \frac{1}{2}\min\{\kappa_1, -\kappa_2\}.$$

Let $e_1, e_2 \in T_pS$ denote the orthonormal basis of eigenvectors corresponding to the eigenvalues κ_1, κ_2. If $e_1 = a_1X_1 + a_2X_2$ and $e_2 = b_1X_1 + b_2X_2$, then for $\delta > 0$ sufficiently small, $(\delta a_1, \delta a_2), (\delta b_1, \delta b_2) \in U$, and we have

$$\begin{aligned} d(\delta a_1, \delta a_2) &\geq \frac{\delta^2}{4}\kappa_1 > 0, \\ d(\delta b_1, \delta b_2) &\leq \frac{\delta^2}{4}\kappa_2 < 0. \end{aligned}$$

Since δ can be made arbitrarily small, the case when p is hyperbolic follows. □

In fact, the sign of the Gaussian curvature at a point allows us to classify its local second order asymptotics.

Lemma 4.4. *Let $p \in S$ and $e_1, e_2 \in T_pS$ be the orthonormal basis of eigenvectors for the shape operator $-dN_p$. After applying a translation and rotation so that $p = (0,0,0)$, $e_1 = (1,0,0)$ and $e_2 = (0,1,0)$, the surface S is locally given by the graph*

$$f(x,y) = \frac{1}{2}(\kappa_1 x^2 + \kappa_2 y^2) + o(x^2 + y^2). \tag{4.1}$$

That is, the second order asymptotics of S at p are given by

- *An elliptic paraboloid if p is elliptic;*
- *A hyperbolic paraboloid if p is hyperbolic;*
- *A parabolic cylinder if p is parabolic.*

Proof. By Lemma 2.9, S is locally a graph of two variables at p. Moreover, as $N_p = (0,0,1)$ lies in both of the planes $\{x = 0\}$ and $\{y = 0\}$, S must locally be a graph over $\{z = 0\}$. That is, for some open subset $U \subseteq \mathbb{R}^2$ containing $(0,0)$, S is locally the graph of some smooth function $f : U \to \mathbb{R}$. By our assumptions, $T_pS = \{z = 0\}$, which implies that $f_1(0,0) = f_2(0,0) = 0$.

Locally, we have the coordinate chart $X : U \to \mathbb{R}^3$ given by $X(u_1, u_2) = (u_1, u_2, f(u_1, u_2))$. It follows that

$$X_1 = (1, 0, f_1), \quad X_2 = (0, 1, f_2), \quad N = \frac{(-f_1, -f_2, 1)}{\sqrt{1 + f_1^2 + f_2^2}}.$$

Differentiating the normal yields

$$-N_1 = \frac{1}{\sqrt{1 + f_1^2 + f_2^2}}(f_{11}, f_{21}, 0) - \frac{f_1 f_{11} + f_2 f_{21}}{(1 + f_1^2 + f_2^2)^{\frac{3}{2}}}(f_1, f_2, 0),$$
$$-N_2 = \frac{1}{\sqrt{1 + f_1^2 + f_2^2}}(f_{12}, f_{22}, 0) - \frac{f_1 f_{12} + f_2 f_{22}}{(1 + f_1^2 + f_2^2)^{\frac{3}{2}}}(f_1, f_2, 0),$$

which at p evaluate to

$$-N_1 = (f_{11}, f_{12}, 0), \quad -N_2 = (f_{12}, f_{22}, 0).$$

Using the formula $-N_i = -dN_p \cdot X_i = \kappa_i e_i$ at p, we also have

$$-N_1 = (\kappa_1, 0, 0), \quad -N_2 = (0, \kappa_2, 0).$$

Equating the two expressions for $-N_1$ and $-N_2$, we find that

$$f_{11} = \kappa_1, \quad f_{12} = f_{21} = 0, \quad f_{22} = \kappa_2.$$

Equation (4.1) follows from Taylor's theorem. □

Consider a regular surface $S \subseteq \mathbb{R}^3$ equipped with local coordinates $X : U \to S$ about p. Recall that the infinitesimal area form on S is given by

$$dA_S = \|X_u \times X_v\| dudv,$$

with respect to $X(u, v)$. Locally, we have the normal map $N : X(U) \to \mathbb{S}^2$, and so we can consider the smooth map $N \circ X : U \to \mathbb{S}^2$. It follows that infinitesimal area form on the sphere is given by

$$dA_{\mathbb{S}^2} = \|N_u \times N_v\| dudv,$$

with respect to $N(u, v)$. Recall, in §3.2 on area, we showed that

$$\|N_u \times N_v\| = \left|\det dN_{(u,v)}\right| \cdot \|X_u \times X_v\| = |K| \cdot \|X_u \times X_v\|,$$

and therefore, the size of the Gaussian curvature can be thought of as the distortion of the map $dA_S \mapsto dA_{\mathbb{S}^2}$ induced by the Gauss map. By integrating this quantity up and taking limits, we have the following lemma.

Lemma 4.5. *Let S be a regular surface and $X : U \to S$ local coordinates about $p \in S$. Suppose $p \in B_n \subseteq X(U)$ are sequence of compact subsets on the surface with*

$$\lim_{n\to\infty} \sup_{q\in B_n} \|p - q\| = 0.$$

Then the size of the Gaussian curvature at p is given by the limiting ratio of the areas

$$|K(p)| = \lim_{n\to\infty} \frac{\int_{N(B_n)} dA_{\mathbb{S}^2}}{\int_{B_n} dA_S}.$$

Proof. Let $X(U_n) = B_n$, so that by the discussion previous to the lemma, we have that the ratio is given in local coordinates as

$$\frac{\int_{N(B_n)} dA_{\mathbb{S}^2}}{\int_{B_n} dA_S} = \frac{\int_{U_n} \|N_u \times N_v\| dudv}{\int_{U_n} \|X_u \times X_v\| dudv} = \frac{\int_{U_n} |K(u,v)| \cdot \|X_u \times X_v\| dudv}{\int_{U_n} \|X_u \times X_v\| dudv}.$$

In particular, we have

$$\begin{aligned}\left|\frac{\int_{N(B_n)} dA_{\mathbb{S}^2}}{\int_{B_n} dA_S} - |K(p)|\right| &= \left|\frac{\int_{U_n} (|K(u,v)| - |K(p)|) \cdot \|X_u \times X_v\| dudv}{\int_{U_n} \|X_u \times X_v\| dudv}\right| \\ &\leq \frac{\int_{U_n} ||K(u,v)| - |K(p)|| \cdot \|X_u \times X_v\| dudv}{\int_{U_n} \|X_u \times X_v\| dudv} \\ &\leq \|K(u,v) - K(p)\|_{L^\infty(U_n)}.\end{aligned}$$

Since K is smooth, taking $n \uparrow \infty$, the right hand size is null and the conclusion follows.

Example 4.6. *The previous lemma gives us another interpretation of why the cylinder C has zero Gaussian curvature: indeed, the Gauss map of a cylinder traces out the equator $N(C) = \mathbb{S}^1 \subseteq \mathbb{S}^2$, and since the area of the equator is zero inside the sphere, the Gaussian curvature $K \equiv 0$ on C.*

□

4.2 Principal curvatures

We return to the principal curvatures $\kappa_1, \kappa_2 : S \to \mathbb{R}$ defined at each point $p \in S$ to be the eigenvalues (with corresponding orthonormal eigenvectors $e_1, e_2 \in T_pS$) of the shape operator $-dN_p$.

Consider a smooth regular curve $\gamma : I \to S$ parameterised by arc-length inside of our surface. For each $s \in I$, the derivative $\tau_{\gamma(s)} := \gamma'(s)$ is a unit vector laying within $T_{\gamma(s)}S$. If $N_{\gamma(s)}$ denotes the unit normal vector to S at $\gamma(s)$, then the triple $\{\tau_{\gamma(s)}, N_{\gamma(s)}, G_{\gamma(s)} := \tau_{\gamma(s)} \times N_{\gamma(s)}\}$ is a orthonormal basis of $\mathbb{R}^3$, with $G_{\gamma(s)} \in T_{\gamma(s)}S$ for each $s \in I$. In particular, the second derivative $\gamma''(s)$ is then a vector laying within the span of G and N at $\gamma(s)$.

Definition 4.7. *With the above set-up, define the smooth functions $\kappa_G, \kappa_N : I \to \mathbb{R}$ via the relationship*

$$\gamma''(s) = \kappa_G(s) \cdot G_{\gamma(s)} + \kappa_N(s) \cdot N_{\gamma(s)}, \quad \forall s \in I. \tag{4.2}$$

*We refer to the values $\kappa_G(s)$ and $\kappa_N(s)$ as the **geodesic curvature** and **normal curvature** of γ at s respectively.*

Remark.

- *Changing the orientation on S will change the sign of both κ_G and κ_N.*
- *For a fixed Gauss map N, changing the direction in which we traverse the curve does **not** change γ'', but will swap the sign of τ and hence G. So reversing the direction changes the sign of κ_G, but not κ_N. i.e. κ_N is independent of the direction the curve $\gamma(I)$ is traversed.*
- *Since the vector $G_{\gamma(s)} \perp N_{\gamma(s)}$ everywhere, taking the norm of (4.2) we find that*

  $$\kappa(s)^2 = \kappa_G(s)^2 + \kappa_N(s)^2,$$

  where $\kappa(s)$ denotes the usual curvature of the curve γ as a curve in $\mathbb{R}^3$. In particular, we if choose θ to be the angle formed between v, the normal vector of the curve γ, and N, the normal vector to S (that is, $\cos(\theta) = \langle v, N \rangle$), then it is true that

  $$\kappa_N = \kappa \cos\theta, \quad \kappa_G = \kappa \sin\theta.$$

Example 4.8. *Given any smooth regular curve $\gamma : I \to P$ in an affine plane P, we see that the normal to the curve v is always perpendicular to the normal to the plane P, and hence $\kappa_N \equiv 0$ and $\kappa_G \equiv \kappa$.*

All of the curvature of γ is due to its bending within P, and not due to the bending of P in the ambient space.

Example 4.9. *Consider the equator $\gamma : \mathbb{R} \to \mathbb{S}^2$, $\gamma(s) = (\cos s, \sin s, 0)$. In this example, the normal to the curve v is parallel to the normal to the sphere N, and hence $\kappa_G = 0$ and $\kappa_N \equiv \kappa$.*

All of the curvature of γ is due to the bending of $\mathbb{S}^2$ in the ambient space and not due to its bending within $\mathbb{S}^2$.

The following lemma shows that the normal curvature of a curve depends only on its first order information. That is, for any two curves $\gamma, \eta : I \to S$ with $\gamma(s_0) = \eta(s_0)$ and $\gamma'(s_0) = \eta'(s_0)$, then the normal curvatures of γ and η agree at s_0.

Lemma 4.10. *The normal curvature κ_N of a curve $\gamma : I \to S$ depends only on the tangent vector to the curve γ. Moreover, the normal curvature is bounded by the principal curvatures at each point*

$$\kappa_2(s) \leq \kappa_N(s) \leq \kappa_1(s), \quad \forall s \in I.$$

Proof. Recall, the normal curvature is given by the inner product

$$\begin{aligned}\kappa_N(s) &= \left\langle \gamma''(s), N_{\gamma(s)} \right\rangle \\ &= \langle \gamma'(s), (N \circ \gamma)'(s) \rangle \\ &= \left\langle -dN_{\gamma(s)} \cdot \gamma'(s), \gamma'(s) \right\rangle \\ &= h_{\gamma(s)}(\gamma'(s)),\end{aligned}$$

which depends only on $\gamma(s)$ and $\gamma'(s)$, showing the first part. The second part follows from the previously shown fact that the second fundamental form h_p is bounded by the principal curvatures on the unit circle in T_pS. □

Definition 4.11. *A point $p \in S$ is called **umbilic** if $\kappa_1(p) = \kappa_2(p)$.*

Lemma 4.12. *Let S be a regular surface and $X : U \to S$ be local coordinates on S with U connected. If for some function $\lambda : X(U) \to \mathbb{R}$ the following equation holds*

$$-dN_p = \lambda(p) \cdot \mathrm{id}_{T_pS}, \quad \forall p \in X(U),$$

then $X(U)$ is contained either within a plane or a sphere. Such a region $X(U)$ is called umbilical.

Remark. *A priori, the function λ given in the lemma can change from point to point, even discontinuously. The fact λ is constant is a consequence of the lemma.*

Proof. Since we may write $\lambda(p) = \frac{\langle -dN_p \cdot X_u, X_u \rangle}{\|X_u\|^2}$, it follows that λ is a smooth function. Next, differentiating the identities

$$-N_u = \lambda \cdot X_u, \quad -N_v = \lambda \cdot X_v,$$

we find that

$$-N_{uv} = \lambda_v \cdot X_u + \lambda \cdot X_{uv}, \quad -N_{vu} = \lambda_u \cdot X_v + \lambda \cdot X_{vu}.$$

As X, N are both smooth, partial derivatives commute, and we conclude

$$\lambda_v \cdot X_u - \lambda_u \cdot X_v = 0.$$

But since X_u, X_v are linearly independent, this implies $\lambda_u = \lambda_v = 0$ on $X(U)$ connected, and hence λ is constant. In particular, it follows that the quantity $N + \lambda X$ is constant. We now split the final analysis into two cases

- If $\lambda \equiv 0$, then the normal vector is constant, and hence $X(U)$ is contained within the hyperplane with normal N passing through a point $X(p)$.

- Otherwise, we label the fixed vector $\lambda X_0 := N + \lambda X$. Rearranging gives $X = X_0 + \lambda^{-1}N$, or X lies in the sphere $X_0 + \lambda^{-1}\mathbb{S}^2$. □

Corollary. *If S is a regular connected umbilical surface: there exists a function $\lambda : S \to \mathbb{R}$ such that*

$$-dN_p = \lambda(p) \cdot \mathrm{id}_{T_pS}, \quad \forall p \in S,$$

then S is contained with a plane or a sphere.

Proof. Fix $p_0 \in S$. By the previous lemma applied to the connected components of coordinate charts (which cover S), we see that λ is a smooth function. Note that the set

$$\Omega = \{p \in S : \lambda(p) = \lambda(p_0)\}$$

is a non-empty subset of S. It is closed since λ is smooth. By the previous lemma, it is open. Since S is connected, we conclude $\Omega = S$ and λ is constant. The result then follows via an identical argument. □

4.3 Mean curvature

We now find a geometric interpretation of the mean curvature via variations of regular surfaces. In particular, we look at the case when the mean curvature vanishes

Definition 4.13. *A regular surface S is called* minimal *if $H \equiv 0$ everywhere.*

Let $S \subseteq \mathbb{R}^3$ be an oriented regular surface with Gauss map $N : S \to \mathbb{S}^2$. We will consider compactly supported variations of S. More precisely, fix a compactly supported smooth function $f \in C_c^\infty(S)$, and cover the support of f by coordinate charts. For simplicity, we will assume that there is a single coordinate chart $X : U \to S$ such that $\mathrm{supp}(f) \Subset X(U)$. We then look at the variation $\hat{X} : U \times \mathbb{R} \to \mathbb{R}^3$, given by

$$\hat{X}(u, v, t) := X(u, v) + tf(u, v) \cdot N_{(u,v)}.$$

Claim. *There exists $\varepsilon > 0$ sufficiently small such that $\hat{X}(\cdot, t)$ is an immersion for every $|t| < \varepsilon$.*

Proof of Claim. Calculating the partial derivatives of $\hat{X}(\cdot, t) : U \to \mathbb{R}^3$ we have

$$\hat{X}_u(\cdot, t) = X_u + tf_uN + tfN_u,$$
$$\hat{X}_v(\cdot, t) = X_v + tf_vN + tfN_v,$$

and therefore, as f, N are smooth on $\mathrm{supp}(f)$ compact,

$$\hat{X}_u(\cdot, t) \times \hat{X}_v(\cdot, t) = X_u \times X_v + O(t),$$

which is non-zero for t sufficiently small. i.e. the vectors $\hat{X}_u(\cdot, t), \hat{X}_v(\cdot, t)$ are linearly independent. □

Exercise. *Show that for t sufficiently small, $\hat{X}(\cdot, t)$ is a homeomorphism onto its image.*

Combining this exercise with the previous claim, we see that for $|t| < \varepsilon$, $\hat{X}(\cdot, t)$ defines local coordinates for a new regular surface $S_t \subseteq \mathbb{R}^3$. Consider the family of first fundamental forms $g(t)$ on S_t. With respect to the local coordinates $\hat{X}(\cdot, t)$ we have

$$\begin{aligned} g_{ij}(t) &= \left\langle \hat{X}_i(\cdot, t), \hat{X}_j(\cdot, t) \right\rangle \\ &= \left\langle X_i + tf_iN + tfN_i, X_j + tf_jN + tfN_j \right\rangle \\ &= g_{ij} + tf\left(\left\langle N_i, X_j \right\rangle + \left\langle N_j, X_i \right\rangle \right) + O(t^2) \\ &= g_{ij} - 2tfh_{ij} + O(t^2), \end{aligned}$$

which we write as matrix formula

$$[g(t)]_{\hat{X}(\cdot,t)} = [g]_X - 2tf[h]_X + O(t^2). \tag{4.3}$$

We want to look at how the area is changing under this compact variation, and so we need the following result.

Claim. *For any $n \times n$ matrix M,*

$$\frac{d}{dt} \det(I_n + tM)|_{t=0} = \operatorname{tr}(M).$$

Proof of Claim. Since all of the off-diagonal entries of the matrix $I_n + tM$ are $O(t)$, we have

$$\begin{aligned} \det(I_n + tM) &= \prod_{i=1}^{n} (1 + tM_{ii}) + O(t^2) \\ &= 1 + t \underbrace{\sum_{i=1}^{n} M_{ii}}_{\operatorname{tr}(M)} + O(t^2). \end{aligned}$$ □

Applying the determinant to (4.3) we have

$$\begin{aligned} \det[g(t)]_{\hat{X}(\cdot,t)} &= \det([g]_X - 2tf[h]_X + O(t^2)) \\ &= \det([g]_X - 2tf[h]_X) + O(t^2) \\ &= \det[g]_X \cdot \det(I - 2tf[g]_X^{-1}[h]_X) + O(t^2), \end{aligned}$$

and so by the claim and equation (3.9)

$$\begin{aligned} \frac{d}{dt} \det[g(t)]_{\hat{X}(\cdot,t)}|_{t=0} &= 2f \det[g]_X \operatorname{tr}([g]_X^{-1}[h]_X) \\ &= 2f \det[g]_X \operatorname{tr}([-dN]_X) \\ &= 2fH \det[g]_X \end{aligned}$$

On the level of infinitesimal area forms, this corresponds to the relationship

$$\begin{aligned}\frac{d}{dt}dA_{S_t}|_{t=0} &= \frac{d}{dt}\sqrt{\det[g(t)]_{\hat{X}(\cdot,t)}}|_{t=0}\, dudv \\ &= fH\sqrt{\det[g]_X}\, dudv \\ &= fHdA_S.\end{aligned}$$

Integrating this quantity up, we have the first variation formula for the area of an oriented regular surface under compactly supported variations.

Theorem 4.14. *Let S be an oriented regular surface with Gauss map $N : S \to \mathbb{S}^2$. For any $f \in C_c^\infty(S)$, let S_t denote the regular surfaces generated by variations along fN as above. Then we have the formula*

$$\frac{d}{dt}\int_{\mathrm{supp}(f)} dA_{S_t}|_{t=0} = \int_{\mathrm{supp}(f)} -fHdA_S.$$

Corollary. *An oriented regular surface is a critical point of the area functional under compactly supported variations if and only if it is minimal.*

Proof. The if direction follows immediately from Theorem 4.14. For the reverse direction, suppose there is a point $q \in S$ with $H(q) \neq 0$. Let φ be any smooth compactly supported non-negative function on S with $\varphi(q) > 0$ and let $f = \varphi H \in C_c^\infty(S)$. Then, considering the variation with respect to f, Theorem 4.14 implies that

$$\frac{d}{dt}\int_{\mathrm{supp}(f)} dA_{S_t}|_{t=0} = \int_{\mathrm{supp}(\varphi)} -\varphi H^2 dA_S < 0. \quad \square$$

There are special coordinates on any regular surface under which checking minimality is much simpler.

Definition 4.15. *Local coordinates $X : U \to S$ are called* isothermal coordinates *if there exists a smooth function $\lambda : U \to \mathbb{R}$ such that the first fundamental form is given in these local coordinates by*

$$[g_{(u,v)}]_X = \begin{pmatrix} \lambda^2(u,v) & 0 \\ 0 & \lambda^2(u,v) \end{pmatrix}, \quad \forall (u,v) \in U.$$

Fact: By solving a local system of partial differential equations, one can show that every regular surface S can be covered by isothermal coordinate charts. Recall the following definition of a harmonic function

Definition 4.16. *A function $X : U \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ is* harmonic *if*

$$\Delta X(u,v) := X_{uu}(u,v) + X_{vv}(u,v) = 0, \quad \forall (u,v) \in U.$$

Exercise. *Let S be a regular surface. Then S is minimal if and only if, all isothermal coordinates $X : U \to S$ on S are harmonic.*

Appendix

Theorem (Inverse Function Theorem)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be open, $f : \Omega \to \mathbb{R}^n$ be a C^∞ function, and $f(a) = b$. Suppose $Df(a)$ is invertible (as an $n \times n$ matrix). Then there exists open sets $U, V \subseteq \mathbb{R}^n$ with $a \in U$ and $b \in V$, and a unique function $g : V \to U$ with $g(b) = a$ such that*

$$\begin{aligned} g \circ f(y) &= y, \quad \forall y \in U, \\ f \circ g(x) &= x, \quad \forall x \in V. \end{aligned}$$

That is g is a local inverse to f. Moreover, g is also a C^∞ function with

$$Dg(x) = Df(g(x))^{-1}, \quad \forall x \in V.$$

Theorem (Implicit Function Theorem)**.** *Let $\Omega \subseteq \mathbb{R}^{n+k}$ be open and $F : \Omega \to \mathbb{R}^k$ be a C^∞-function. Denote $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$, and*

$$F(x, y) = \begin{pmatrix} F_1(x, y) \\ \vdots \\ F_k(x, y) \end{pmatrix} = \begin{pmatrix} F_1(x_1, \ldots, x_n, y_1, \ldots, y_k) \\ \vdots \\ F_k(x_1, \ldots, x_n, y_1, \ldots, y_k) \end{pmatrix}.$$

Suppose $(a, b) \in \Omega$ is such that $F(a, b) = c \in \mathbb{R}^k$, and that the $k \times k$ matrix

$$\frac{\partial F}{\partial y}(a, b) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(a, b) & \cdots & \frac{\partial F_1}{\partial y_k}(a, b) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_k}{\partial y_1}(a, b) & \cdots & \frac{\partial F_k}{\partial y_k}(a, b) \end{pmatrix},$$

is invertible. Then, there exists open sets $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^k$ with $a \in U$ and $b \in V$, and a unique function $\varphi : U \to V$ such that $\varphi(a) = b$ and

$$F(x, \varphi(x)) = c, \quad \forall x \in U.$$

Moreover, φ is a C^∞ function with Jacobian matrix

$$\underbrace{\left(\frac{\partial \varphi}{\partial x}\right)}_{k \times n} = -\underbrace{\left(\frac{\partial F}{\partial y}\right)^{-1}}_{k \times k} \cdot \underbrace{\left(\frac{\partial F}{\partial x}\right)}_{k \times n}, \quad \forall x \in U.$$

Written in full, this is the equation

$$\begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1}(x) & \cdots & \frac{\partial \varphi_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_k}{\partial x_1}(x) & \cdots & \frac{\partial \varphi_k}{\partial x_n}(x) \end{pmatrix} = -\begin{pmatrix} \frac{\partial F_1}{\partial y_1}(x, \varphi(x)) & \cdots & \frac{\partial F_1}{\partial y_k}(x, \varphi(x)) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_k}{\partial y_1}(x, \varphi(x)) & \cdots & \frac{\partial F_k}{\partial y_k}(x, \varphi(x)) \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x, \varphi(x)) & \cdots & \frac{\partial F_1}{\partial x_n}(x, \varphi(x)) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_k}{\partial x_1}(x, \varphi(x)) & \cdots & \frac{\partial F_k}{\partial x_n}(x, \varphi(x)) \end{pmatrix}.$$