B3.2 GEOMETRY OF SURFACES
MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, MT 2025

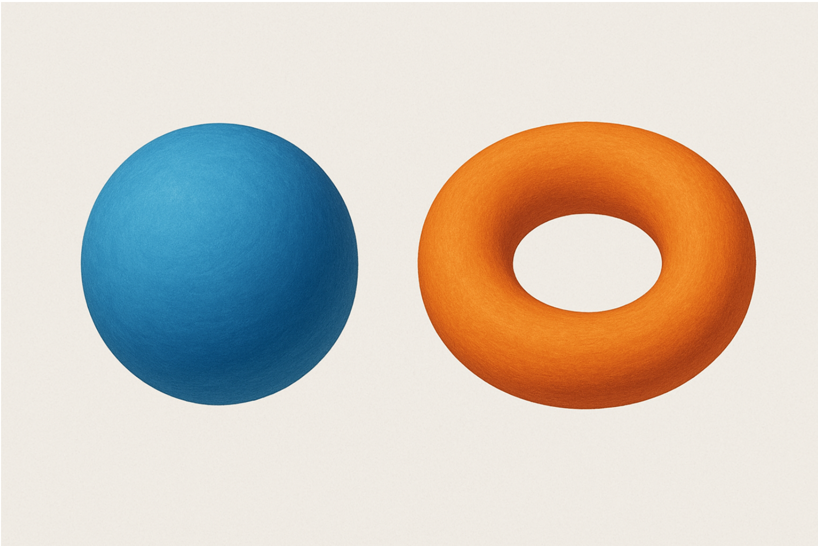
HÜLYA ARGÜZ

Contents

Acknowledgement These notes were prepared by TeX-typing Dominic Joyce's handwritten lecture slides from his recorded lectures in the 2020 COVID year. Nigel Hitchin's 2013 lecture notes serve as a second key source for this lecture.

1. Introduction

This is a course on "surfaces". Roughly speaking, surface is a shape that looks locally like a 2-dimensional disc — even if it may curve or twist when viewed as a whole (We will see a precise definition of a surface later on). Here are two examples of surfaces, a sphere and a torus:



Our goal is to study and relate three classes of surfaces:

(1) Topological surfaces (topological 2-manifolds),
(2) Riemann surfaces (complex 1-manifolds).
(3) Smooth surfaces (smooth real 2-manifolds),
   (a) embedded in $\mathbb{R}^3$,
   (b) abstractly (i.e.possibly without a choice of embedding into any $\mathbb{R}^n$),

Whether a surface locally looks like the disc continuously, holomorphically, or smoothly distinguishes the cases (1), (2), and (3) respectively.

2. Topological Surfaces

2.1. **Background from topology.** A5: Topology is essential for this course. We will recall some basic notions (several references for these, see for instance Wilson A. Sutherland. Introduction to metric and topological spaces. Oxford University Press, Oxford, second edition, 2009. Companion web site: `www.oup.com/uk/companion/metric.`).

Definition 2.1. A *topological space* is a set X together with a collection $\mathcal{T}$ of subsets of X, called *open sets* in X, such that:

(i) $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$,
(ii) If $U, V \in \mathcal{T}$, then $U \cap V \in \mathcal{T}$,
(iii) If $U_i \in \mathcal{T}$ for $i \in I$, for an index set I (could be finite or infinite), then $\bigcup_i U_i \in \mathcal{T}$.

We say that $\mathcal{T}$ is a *topology* for X. We sometimes use the notation $(X, \mathcal{T})$ for a topological space X with a topology $\mathcal{T}$.

- X is *compact* if every open cover of X has a finite subcover. A topological space is called *locally compact* if every point has a compact neighbourhood.
- A map $f: X \to Y$ of topological spaces is *continuous* if whenever $V \subseteq Y$ is open, then $f^{-1}(V) \subseteq X$ is open.

- X is *Hausdorff* if whenever $x, y \in X$, and $x \neq y$, there are open subsets U, V of X with $U \cap V = \emptyset$ such that $x \in U$ and $y \in V$.
- $f : X \to Y$ is a *homeomorphism* if it is continuous and invertible with continuous inverse. In this case, X and Y are called *homeomorphic*.
- If X is a topological space with topology $\mathcal{T}_X$ and $Y \subseteq X$, the *subspace topology* on Y is $\mathcal{T}_Y = \{Y \cap U \mid U \in \mathcal{T}_X\}$.

Definition 2.2. Given a topological space $(X, \mathcal{T})$, a *basis* for $\mathcal{T}$ is a subcollection $\mathcal{B} \subset \mathcal{T}$ such that every set in $\mathcal{T}$ is a union of sets from $\mathcal{B}$.

Many topological properties can be verified by just checking basis elements: for instance, to check a map $f \colon (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ of topological spaces is continuous, it suffices to check that $f^{-1}(B)$ is open in $\mathcal{T}_X$ for each set B in some basis $\mathcal{B}_Y$ for $\mathcal{T}_Y$. Although a topology may admit many different bases, it is often preferable in practice to work with one that uses as few sets as possible.

- A topological space X is called *second countable* if it has a countable basis for its topology. Every compact metric space is second countable.

You need to be familiar with these topological notions and with the basic constructions, such as the construction of a quotient topological space.

2.2. **Topological Surfaces.**

Definition 2.3. A *topological surface* (or just *surface*) X is a Hausdorff topological space such that each $x \in X$ has an open neighbourhood $U \subseteq X$ with a homeomorphism $\varphi \colon U \to V$ to an open subset $V \subseteq \mathbb{R}^2$. The triple (U, V, φ) is a *chart* on X.

Remarks 2.4. (a) In older books, a surface is called *closed* if it is compact.
(b) It is better to also require X to be second countable or paracompact (global topological conditions). As they are automatic for compact surfaces, which is what we mostly care about, for simplicity we won't worry about this.
(c) More generally, we can define a *topological manifold of dimension n* to be a (second countable) Hausdorff topological space X, locally homeomorphic to $\mathbb{R}^n$. Then, surfaces are topological manifolds of dimension 2. We will see later that compact surfaces can be completely classified. But for $n > 2$, classifying compact topological n-manifolds is very difficult.

Example 2.5. Any open subset X of $\mathbb{R}^2$ is a topological surface which can be covered by one chart (U, V, φ) with $U = V = X$, and $\varphi = \mathrm{id}$.

Example 2.6. The 2-sphere is

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

We can cover this by two charts (U_1, V_1, φ_1), (U_2, V_2, φ_2) with $U_1 = \{S^2 \setminus (0, 0, -1)\}$, $U_2 = \{S^2 \setminus (0, 0, 1)\}$, $V_1 = V_2 = \mathbb{R}^2$, and $\varphi_i : U_i \to V_i$ given by

$$\varphi_1(x, y, z) = \frac{1}{1+z}(x, y)$$
$$\varphi_2(x, y, z) = \frac{1}{1-z}(x, y).$$

Note that you can not cover S^2 with one chart, you need at least two. So, S^2 is locally homeomorphic to $\mathbb{R}^2$. It is Hausdorff (and second countable), as it is a subspace of $\mathbb{R}^2$ which is both. So, S^2 is a topological surface. Moreover, it is compact by Heine-Borel, as closed and bounded in $\mathbb{R}^3$.

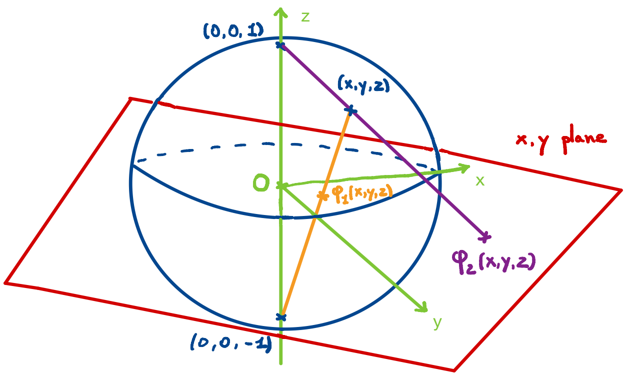


FIGURE 2.1. The 2-sphere S^2

2.3. Building surfaces as quotient topological spaces.

Definition 2.7. Let X be a topological space, and $\sim$ an equivalence relation on X. Write $X/\sim$ for the set of $\sim$ equivalence classes $[x]$ of points x in X. Write $\pi\colon X \to X/\sim$ for the surjective projection $\pi\colon x \mapsto [x]$.

The *quotient topology* on $X/\sim$ is defined by $U \subseteq X/\sim$ is open iff $\pi^{-1}(U) \subseteq X$ is open in X.

Warning: If $\sim$ is not well chosen, then $X/\sim$ may not be a nice topological space. For example, X Hausdorff does not imply $X/\sim$ Hausdorff (in fact, $X/\sim$ is Hausdorff iff $\{(x,y) \in X \times X \, : x \sim y\}$ is closed in $X \times X$).

X compact does imply $X/\sim$ is compact.

Example 2.8. Take $X = [0,1]$. Define an equivalence relation $\sim$ on X by $x \sim x$, $0 \sim 1$, $1 \sim 0$. Think of it as gluing 0 to 1. Then, $X/\sim$ is the circle.

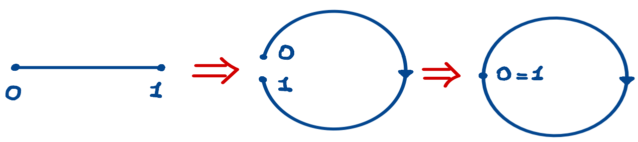


FIGURE 2.2. The circle as a quotient of the interval

Example 2.9. Take $X = [0,1]^2$. Define an equivalence relation $\sim$ on X by $(x,0) \sim (x,1)$ and $(0,y) \sim (1,y)$, for all $x, y \in [0,1]$. (We leave out $(x,y) \sim (x,y)$ and $(x,1) \sim (x,0)$, etc. If we just write some relations in an equivalence relation, we mean $\sim$ is the weakest equivalence relation inducing these relations). Draw this as in Figure 2.3.

The same kind of arrows means identify these sides in this direction. Then $X/\sim$ is a surface homeomorphic to the torus $T^2 = (S^1)^2$. Indeed, $X/\sim = ([0,1]/\sim) \times ([0,1]/\sim) \cong S^1 \times S^1$, where $0 \sim 1$ as in Example 2.8.

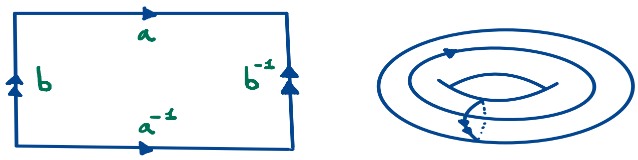


FIGURE 2.3. The torus T^2 as a quotient of $[0, 1]^2$

To convince yourself that $X/\sim$ is a surface, that is, it is locally homeomorphic to $\mathbb{R}^2$:

- Any point $(0, y)/\sim$ has an open neighbourhood homeomorphic to an open ball in $\mathbb{R}^2$, as illustrated below.

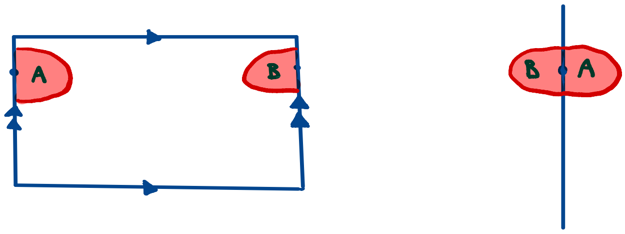


FIGURE 2.4. Open neighbourhood of a point $(0, y)/\sim$

- Any point $(0, 0)/\sim$ has an open neighbourhood homeomorphic to an open ball in $\mathbb{R}^2$. Note that all 4 vertices of X are identified in $X/\sim$, as illustrated below.

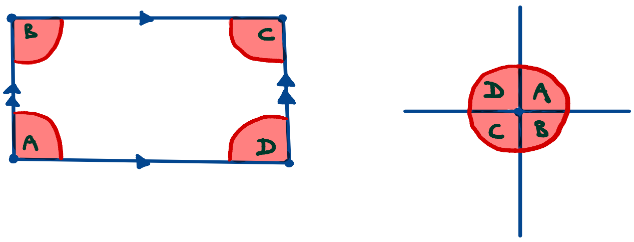


FIGURE 2.5. Open neighbourhood of a point $(0, 0)/\sim$

Later, we will define notation which describes the diagram in Figure 2.3 as $ab^{-1}a^{-1}b$. In fact, we can make other topological surfaces by identifying sides of $X = [0, 1]^2$.

Example 2.10. In Figure 2.6 we illustrate $X/\sim\ \cong\ S^2$.
Note that only 2 vertices get identified.
Later notation: $aa^{-1}bb^{-1}$.

Example 2.11. In Figure 2.7 we illustrate $X/\sim\ \cong\ \mathbb{RP}^2$, the real projective space.
$\mathbb{RP}^2 = S^2/\pm 1$, $(x, y, z) \sim (-x, -y, -z)$.
This is a *non-orientable* surface (explained later)
Later notation: $abab$.

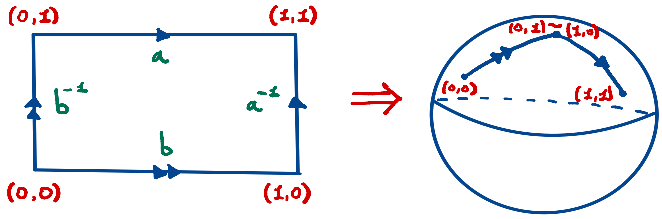


FIGURE 2.6. S^2 as a quotient of $[0,1]^2$

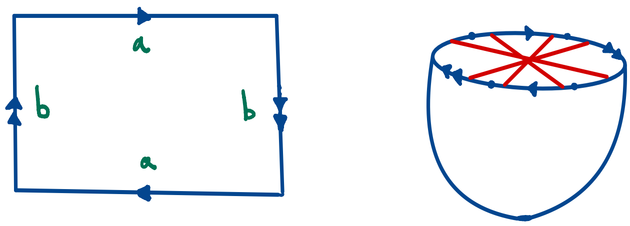


FIGURE 2.7. $\mathbb{RP}^2$ as a quotient of $[0,1]^2$ and as a quotient of S^2

One can think of $\mathbb{RP}^2$ as a half-sphere with opposite points on the equator identified. Note that $\mathbb{RP}^2$ cannot be embedded in $\mathbb{R}^3$ (since it is not orientable), so it is difficult to visualize.

Example 2.12. In Figure 2.8 we illustrate $X/\sim\, \cong K$, the Klein bottle.

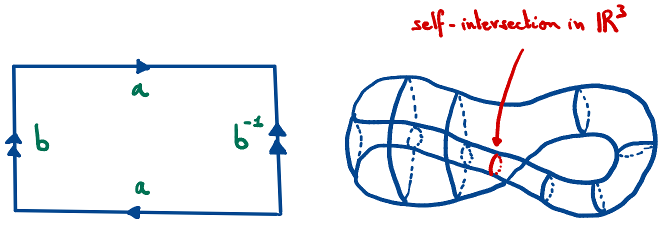


FIGURE 2.8. The Klein bottle

Note that all 4 vertices get identified by $\sim$.

Later notation: $ab^{-1}ab$.

The Klein bottle is also non-orientable, it cannot be embedded in $\mathbb{R}^3$.

Examples 2.8-2.12 are all different (non-homeomorphic) surfaces.

Generally, we do not need to work with squares: we can take any polygon X in the plane with an even number of sides, and identify sides in pairs, then $X/\sim$ will be a compact surface. One needs to do a little work to decide which subsets of vertices are identified. In the example below in Figure 2.9, all 6 vertices are identified. Later notation: $abc^{-1}a^{-1}cb$.

Every point in $X/\sim$ locally looks like in Figure 2.10. Note that $X/\sim$ is a surface which is compact, as X is.

Another example is provided in Figure 2.11 – later notation: $ab^{-1}ca^{-1}cb$. In this example, vertices are identified in 2 groups of 3.

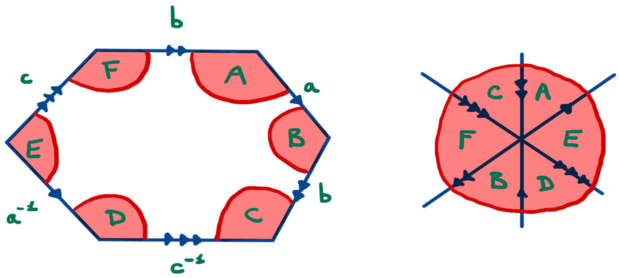


FIGURE 2.9. The compact surface $abc^{-1}a^{-1}cb$

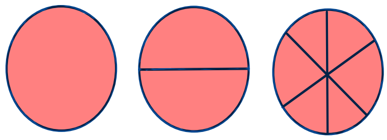


FIGURE 2.10. Local neighbourhoods of points in the surface $abc^{-1}a^{-1}cb$

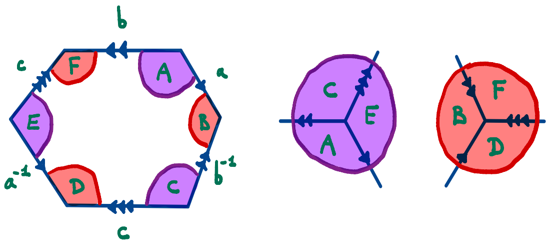


FIGURE 2.11. The compact surface $ab^{-1}ca^{-1}cb$

Learn how to calculate which groups of vertices are identified:

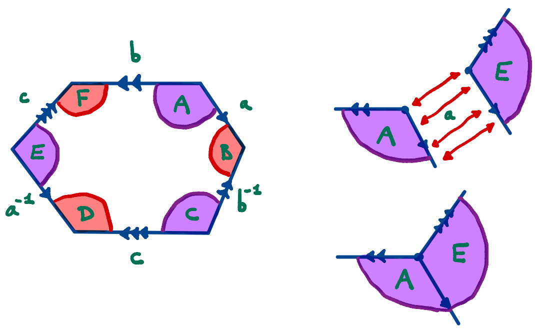


FIGURE 2.12. Calculating groups of vertices identified

Proceed with this process until all closes up, and you have got one group of vertices identified by $\sim$. Do this for all vertex groups.

We may not always start with a polygon. Generally, we may also allow curved sides. For instance, we can start with a 2-gon:

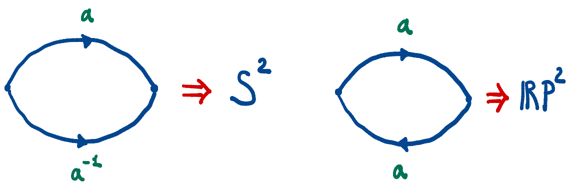


FIGURE 2.13. Surfaces from curved 2-gons

Here is some notation for describing polygons with side identifications. Label sides by $a, b, c, \ldots$ or $a^{-1}, b^{-1}, c^{-1}, \ldots$, where identified sides get the same label $a, b, c, \ldots$ and have $a, b, c, \ldots$ for sides with clockwise arrows and $a^{-1}, b^{-1}, c^{-1}, \ldots$ for sides with anti-clockwise arrows.

Then, starting from some vertex (it does not matter which), list the labels on the sides in clockwise order. For instance, for Figure 2.14 below, we get $abc^{-1}a^{-1}cb$, or equivalently $c^{-1}a^{-1}cbab$ – cyclic permutations do not make any difference.

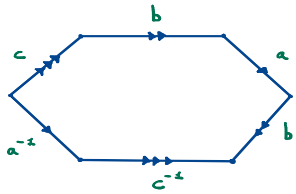


FIGURE 2.14. The surface $abc^{-1}a^{-1}cb$ is also the surface $c^{-1}a^{-1}cbab$

This is called a *planar model* for a surface. It provides a very succinct way to describe surfaces.

Note that different words can describe the same (i.e. homeomorphic) surfaces, and it is not trivial to decide when they do.

2.4. **Cellular decompositions and triangulations.** Let $n \in \{0, 1, 2, \ldots\}$. Denote by

$$D^n = \{\underline{x} \in \mathbb{R}^n \mid \|x\| \leq 1\}, \text{the closed unit disc,}$$

$$\mathring{D}^n = \{\underline{x} \in \mathbb{R}^n \mid \|x\| < 1\}, \text{the interior of } D^n, \text{and}$$

$$\partial D^n = \{\underline{x} \in \mathbb{R}^n \mid \|x\| = 1\}, \text{the boundary of } D^n.$$

Definition 2.13. Let X be a compact surface. A *cellular decomposition* of X is a finite collection of continuous maps, called *cells*, given by:

(i) Maps $v_i \colon D^0 \to X$ called *0-cells* or *vertices*,
(ii) Maps $e_j \colon D^1 \to X$ called *1-cells* or *edges*, and
(i) Maps $f_k \colon D^2 \to X$ called *2-cells* or *faces*,

satisfying the following conditions:

(a) Each map restricted to the interior $\mathring{D}^n$ is a homeomorphism onto its image.
(b) The image of ∂D^n is contained in the images of the cells of dimension strictly less than n.

(c) X is the disjoint union of the images of the interiors of the cells.

Remark 2.14. Definition 2.13 essentially coincides with what Hitchin's lecture notes refer to *subdivision*. In that setting, a subdivision of X is given by the collection of subsets $v_i(D^0), e_j(\mathring{D}^1)$, and $f_k(\mathring{D}^2)$, referred to as *vertices, edges* and *faces*, respectively, but the specific data of the mapss v_i, e_j, f_k is not recorded.

What this means: A *cellular decomposition* or *subdivision* of X is a division of X into polygons (the faces), with edges and vertices. Each edge ends at two vertices. However, the maps e_i, f_j are not required to be injective when restricted to the boundaries of their domains. So, an edge can end at the same vertex at both ends. A face can have repeated edges and vertices in its boundary.

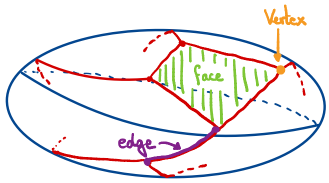


FIGURE 2.15. A cellular decomposition of a surface

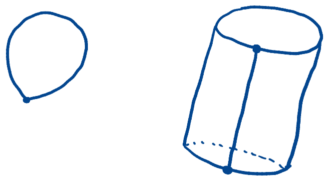


FIGURE 2.16. Repeated vertices and edges in cellular decompositions

A *planar model* of X, as introduced in the previous section, is simply a cellular decomposition (or subdivision) of X with a single face. A *triangulation* of X is a cellular decomposition (or subdivision) in which every face is adjacent to three edges and three vertices (allowing repetitions). Equivalently, a triangulation may be viewed as a decomposition of X into closed triangles whose edges are identified in pairs.

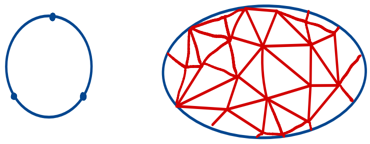


FIGURE 2.17. Triangle and triangulation

Here is a fact we will not prove in this course:

Theorem 2.15. *Every compact surface admits a triangulation.*

We will later discuss the construction of triangulations in the case where X is a smooth surface or a Riemann surface.

2.5. **The Euler characteristic.**

Definition 2.16. Let X be a compact surface. Choose a cellular decomposition (or subdivision) of X with V vertices, E edges and F faces. The *Euler characteristic* of X is:

$$\chi(X) = V - E + F \in \mathbb{Z}\,.$$

Theorem 2.17. *The Euler characteristic $\chi(X)$ depends only on X as a topological space, and not on the choice of subdivision.*

Sketch of the proof: (not examinable) The Euler charcteristic is invariant under the following two processes:

(A) Subdividing an edge into two edges, by adding a vertex at an interior point.
(B) Subdividing a face into two faces, by connecting two vertices by a new edge.

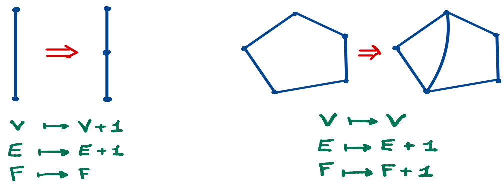


FIGURE 2.18. Moves between triangulations

Claim: Any two subdivisions of X can be linked by a finite sequence of moves A, B and their inverses, together with continuous deformations. So, $\chi(X)$ is independent of the subdivision.

Alternatively: One can define the homology groups $H_i(X)$, depending only on X as a topological space (see section 3.1 in Allen Hatcher. Algebraic topology. Cambridge University Press, Cambridge, 2002.) and show that $\chi(X) = \dim H^0(X) - \dim H^1(X) + \dim H^2(X)$.

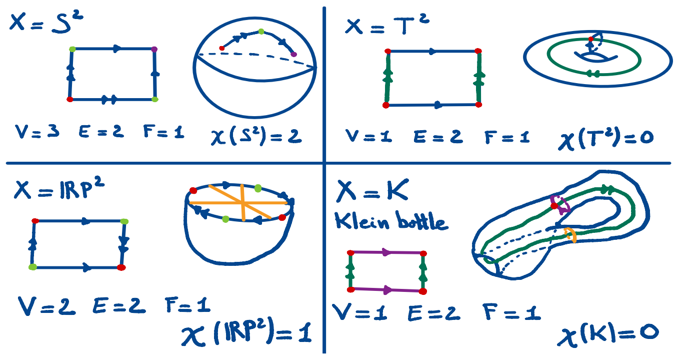


FIGURE 2.19. Euler characteristics of S^2, T^2, $\mathbb{RP}^2$, and K

Example 2.18. Consider the sphere $X = S^2$. We have $V = 3, E = 2, F = 1$. So, $\chi(X) = 2$ – see Figure 2.19.

Example 2.19. Consider the torus $X = T^2$. We have $V = 1, E = 2, F = 1$. So, $\chi(X) = 0$– see Figure 2.19.

Example 2.20. Consider the real projective plane $X = \mathbb{RP}^2$. We have $V = 2, E = 2, F = 1$. So, $\chi(X) = 1$ – see Figure 2.19.

Example 2.21. Consider the Klein bottle $X = K$. We have $V = 1, E = 2, F = 1$. So, $\chi(X) = 0$ – see Figure 2.19. Note that $\chi(T^2) = \chi(K) = 0$, but $T^2 \ncong K$.

2.6. **Connected sums.** Let X, Y be compact, connected surfaces. The *connected sum* $X\#Y$ is a compact, connected surface obtained by cutting out small open discs $\mathring{D} \subset X$, $\mathring{D}' \subset Y$ and gluing the S^1 boundaries of $X \setminus \mathring{D}$ and $Y \setminus \mathring{D}'$ (Technically, this depends on the orientation you glue the boundary S^1, but we ignore this).

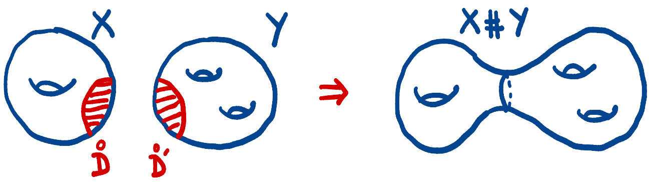


FIGURE 2.20. Connected sums of surfaces

Note that $X\#S^2 \cong X$, as it is obtained by cutting out a disc and gluing a disc.
Choose triangulations of X, Y such that the discs cut out are faces. Then,

$$\begin{aligned} V_{X\#Y} &= V_X + V_Y - 3 \\ E_{X\#Y} &= E_X + E_Y - 3 \\ F_{X\#Y} &= F_X + F_Y - 2\,. \end{aligned}$$

So, we obtain

$$\chi(X\#Y) = \chi(X) + \chi(Y) - 2\,. \tag{2.1}$$

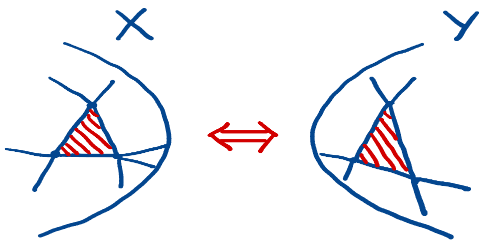


FIGURE 2.21. Gluing triangulations in a connected sum

Example 2.22. The surface Σ_g of genus $g \geq 0$, or *sphere with g holes*, is the multiple connected sum of g copies of T^2 if $g > 0$ (it is just S^2 if $g = 0$).

From (2.1), since $\chi(S^2) = 2$, $\chi(T^2) = 0$, we see by induction that

$$\chi(\Sigma_g) = 2 - 2g\,. \tag{2.2}$$

Note that this distinguishes Σ_g, for different g.

We can take connected sums of planar models:

If X is represented by a word $a_1 \dots a_{2k}$ and Y by a word $b_1 \dots b_{2\ell}$, then $X\#Y$ is represented by $a_1 \dots a_{2k} b_1 \dots b_{2\ell}$.

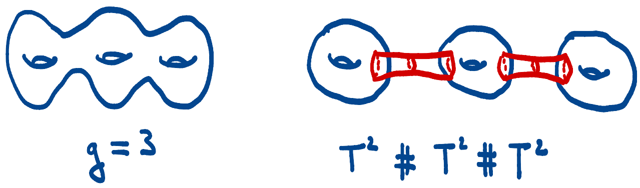


FIGURE 2.22. The genus $g = 3$ surface Σ_3

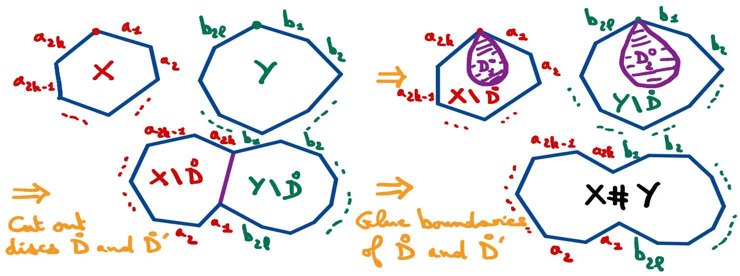


FIGURE 2.23. Connected sum of planar models

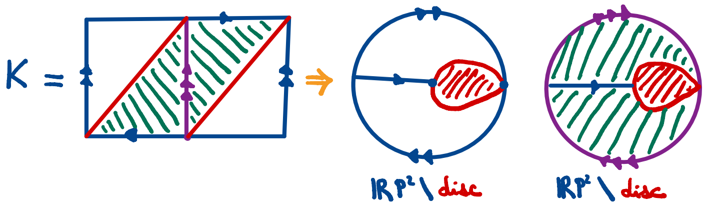


FIGURE 2.24. Proof of $\mathbb{RP}^2 \# \mathbb{RP}^2 \cong K$

Example 2.23. There is a homeomorphism $\mathbb{RP}^2 \# \mathbb{RP}^2 \cong K$, Klein bottle.

As illustrated in Figure 2.24, cut K along the red S^1 into two Möbius strips (white and green regions), which are each $\mathbb{RP}^2 \setminus$ disc.

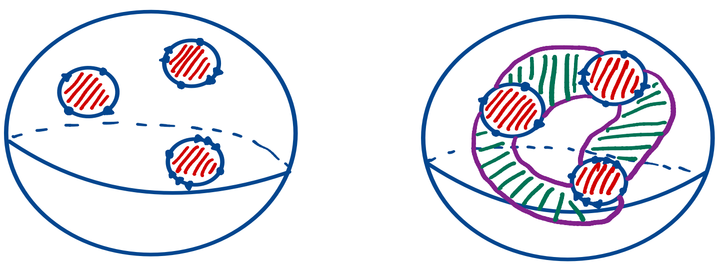


FIGURE 2.25. Proof of $\mathbb{RP}^2 \# \mathbb{RP}^2 \# \mathbb{RP}^2 \cong T^2 \# \mathbb{RP}^2$

Example 2.24. There are homeomorphisms $K\#\mathbb{RP}^2 \cong \mathbb{RP}^2\#\mathbb{RP}^2\#\mathbb{RP}^2 \cong T^2\#\mathbb{RP}^2$: The first homeomorphism here follows from the previous example. For the second one, note that $\mathbb{RP}^2\#\mathbb{RP}^2\#\mathbb{RP}^2 \cong S^2\#\mathbb{RP}^2\#\mathbb{RP}^2\#\mathbb{RP}^2$. Regard this as S^2 with 3 discs removed, with boundaries glued as shown. Note that taking a connected sum of a surface with $\mathbb{RP}^2$ amounts to removing a disc from it and then gluing in a Möbius strip.

Now to see $\mathbb{RP}^2\#\mathbb{RP}^2\#\mathbb{RP}^2 \cong T^2\#\mathbb{RP}^2$, as illustrated in Figure 2.25 cut along purple lines into two regions as shown. Green region is $\mathbb{RP}^2 \setminus$ disc and white region is $T^2 \setminus$ disc. So, $\mathbb{RP}^2\#\mathbb{RP}^2\#\mathbb{RP}^2 \cong T^2\#\mathbb{RP}^2$.

Note that $T^2\#\mathbb{RP}^2 \cong K\#\mathbb{RP}^2$, but $T^2 \ncong K$, so we can not do cancellation in connected sums.

2.7. **Orientations and orientability.**

Definition 2.25. The *Möbius strip* M is $[0,1]\times[0,1]/\sim$ where $(0,y)\sim(1,1-y)$.

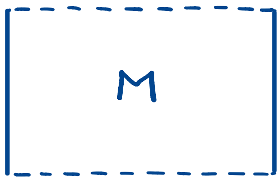


FIGURE 2.26. The Möbius strip

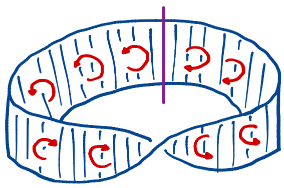


FIGURE 2.27. The Möbius strip

A surface X is called *orientable* if it does not contain any open subset homeomorphic to M.

Equivalent pointof view: An *orientation* on X is a consistent notion of "*clockwise*" everywhere on X.The Möbius strip has no orientation, as if we take a notion of clockwise and deform ariund the loop, it turns into anticlockwise. figure If X cannot be oriented, there is some loop in X, deforming round which turns clockwise into anticlockwise, and a neighbourhood of this loop is a Möbius strip.

For a planar model, X is orientable iff every pair of glued edges are oriented one clockwise and one anticlockwise (i.e. in the word like $ab^{-1}c^{-1}a^{-1}bc$, each symbol $a,b,c,\dots$ appears once as a and once as a^{-1} etc). Since otherwise a strip drawn from $a \implies a$ or $a^{-1} \implies a^{-1}$ is a Möbius strip.

Examples 2.26.
- $S^2 = aa^{-1}bb^{-1}$ is orientable.
- $T^2 = ab^{-1}a^{-1}b$ is orientable.
- $\mathbb{RP}^2 = abab$ is not orientable.

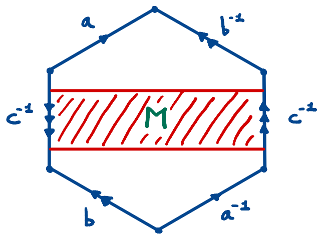


FIGURE 2.28. A Möbius strip in a non-orientable surface

- $K = ab^{-1}ab$ is not orientable.

Note that $\chi(T^2) = \chi(K) = 0$, but one can distinguish T^2 and K as one is orientable and one is not.

A connected sum $X\#Y$ is orientable iff X and Y are both orientable.

2.8. **The classification of surfaces.**

Theorem 2.27. *Let X be a compact, connected surface. Then, either:*

(a) *X is orientable, and then X is homeomorphic to a surface Σ_g for $g \geq 0$ (recall $\Sigma_0 \cong S^2$ and Σ_g is the connected sum of g copies of T^2's for $g \geq 1$), and $\chi(X) = 2 - 2g$, or*

(b) *X is not orientable, and then X is homeomorphic to a connected sum of h copies of $\mathbb{RP}^2$ for $h \geq 1$, $X = \mathbb{RP}^2\# \dots \#\mathbb{RP}^2$ and $\chi(X) = 2 - h$.*

We sketch a proof below (this is not examinable). See Hitchin's notes for details.

Sketch of the proof

Step 1: X admits a triangulation, by Theorem 2.15 in §2.13.

Step 2: X admits a planar model: Take a subdivision of X with minimal number of faces – such a subdivision exists, since any triangulation is one. If a subdivision contains more than one face, two adjacent faces can be glued together along a common edge, thereby reducing the total number of faces. Hence, the minimal number of faces is in fact 1.

Step 3: X admits a planar model with a single vertex, unless $X = S^2$; otherwise, one can cut and paste to reduce the number of vertices. For example, an edge can be contracted, as illustrated below.

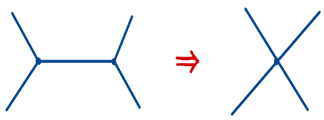


FIGURE 2.29. Contracting an edge

Hence: X has a subdivision with 1 vertex, 1 face and n edges, and can be obtained by sides of a $2n$-gon. Then, $\chi(X) = 2 - n$, so $\chi(X) \leq 2$. [Special case: for $X = S^2$ we have 2 vertices, 1 face and 1 edge. So, $\chi(X) = 2$ – we can not shrink the edge to reduce the number of vertices in this case.]

Step 4: If X is not orientable, then the planar model has two glued edges with same orientation. We can draw a Möbius strip $M = \mathbb{RP}^2 \setminus (\text{disc})$ joining these. Hence, we can write $X = Y\#\mathbb{RP}^2$ with $\chi(Y) = \chi(X) + 1$, $n_Y = n_X - 1$.

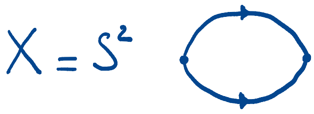


FIGURE 2.30. Step 3: $X = S^2$

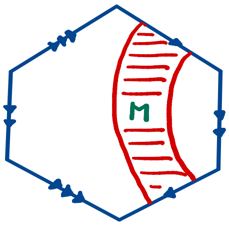


FIGURE 2.31. Step 4: Möbius strip in non-orientable surface

Step 5: If X is orientable, one can show that $X = Y\#T^2$, so that $\chi(Y) = \chi(X) + 2$, $n_Y = n_X - 2$, or $X = S^2$, by cutting and pastimg planar models.

Step 6: Induction on n now implies that, either

$$X = \#_g T^2, \text{for some } g \geq 0, \text{ if } X \text{ is orientable, and}$$
$$X = (\#_g T^2)\#(\#_k \mathbb{RP}^2), \text{for some } g, k \geq 0 \text{ if } X \text{ is orientable.}$$

In the latter case, since $T\#\mathbb{RP}^2 \cong \mathbb{RP}^2\#\mathbb{RP}^2\#\mathbb{RP}^2$, by Example 2.24, we obtain $X = \#_{2g+k}\mathbb{RP}^2$. This completes the proof.

Remark 2.28. For surfaces, knowing the Euler characteristic $\chi(X)$ and whether the surface is orientable or not completely determines it, up to homeomorphism. Things get far more complicated in higher dimensions — no such neat classification exists.

3. Riemann surfaces

A *Riemann surface* is a topological surface X with an extra geometric structure, a *complex structure* or *holomorphic atlas* $\mathcal{A}$. This gives a notion of *holomorphic function* $U \to \mathbb{C}$, for $U \subseteq X$ open, and more generally a notion of *holomorphic map* $f : X \to Y$ for Riemann surfaces X, Y.

A2: Complex analysis generalizes to Riemann surfaces. Riemann surfaces are 1-dimensional complex manifolds and generalize to *n-dimensional complex manifolds* locally modelled on $\mathbb{C}^n$.

3.1. The definition of Riemann surface.

Definition 3.1. Let X be a topological surface. A *complex chart* on X is a triple (U, V, φ) such that $U \subseteq X$, $V \subseteq \mathbb{C}$ are open and $\varphi\colon U \to V$ is a homeomorphism.

In Definition 3.1 since $\varphi\colon U \to V$ is a homeomorphism, $\varphi^{-1}\colon V \to U$ is also a homeomorphism. Notation for complex charts varies in different books: one can also write as pairs (U, φ), or as (V, φ^{-1}), since $V = \varphi(U)$ and $U = \varphi^{-1}(V)$.

As $\mathbb{C} \cong \mathbb{R}^2$, by definition of surfaces, every point $x \in X$ on a surface X admits a complex chart with $x \in U$ for some open $U \subset X$. Think of φ as a *holomorphic coordinate* on $U \subseteq X$.

We call two charts (U_1, V_1, φ_1) and (U_2, V_2, φ_2) *compatible* if $\varphi_2 \circ \varphi_1^{-1} \colon \varphi_1(U_1 \cap U_2) \to \varphi_2(U_1 \cap U_2)$ is a holomorphic map between open subsets of $\mathbb{C}$, with holomorphic inverse (note that $\varphi_1(U_1 \cap U_2) \subseteq V_1 \subseteq \mathbb{C}$ and $\varphi_2(U_1 \cap U_2) \subseteq V_2 \subseteq \mathbb{C}$). The map $\varphi_2 \circ \varphi_1^{-1}$ is called a *transition function*.

Note: $\varphi_2 \circ \varphi_1^{-1}$ is automatically a homeomorphism, so being holomorphic implies it has a holomorphic inverse by a theorem in complex analysis.

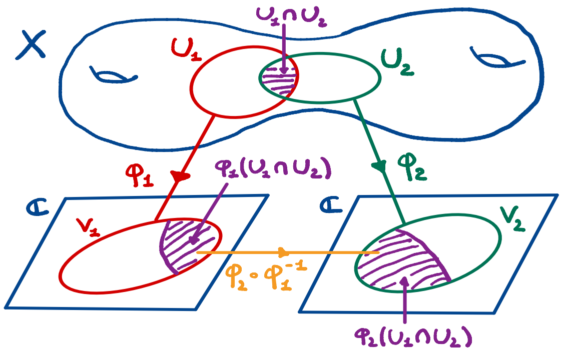


Figure 3.1. Riemann surface

A *holomorphic atlas* $\mathcal{A} = \{(U_i, V_i, \varphi_i) \colon i \in I, \text{ for some index set } I\}$ is a family of complex charts on X, such that:

(i) (U_i, V_i, φ_i) and (U_j, V_j, φ_j) are compatible for all $i, j \in I$,
(ii) $X = \bigcup_{i \in I} U_i$.

A *Riemann surface* $(X, \mathcal{A})$ is a topological surface X with a holomorphic atlas $\mathcal{A}$. Usually, we omit $\mathcal{A}$ from notation, and just say X is a Riemann surface. (Sometimes one requires $\mathcal{A}$ to be a *maximal* atlas, that is, not a proper subset of any other atlas. Any atlas is contained in a unique maximal atlas. This makes the definition a bit more canonical.)

Let X, Y be Riemann surfaces with holomorphic atlases $\mathcal{A} = \{(U_i, V_i, \varphi_i) \colon i \in I\}$, and $\mathcal{B} = \{(U'_j, V'_j, \varphi'_j) \colon j \in J\}$. We call a continuous map *holomorphic* if for all $i \in I$ and $j \in J$,

$$\varphi_i^{-1}(U_i \cap f^{-1}(U'_j)) \xrightarrow{\varphi'_j \circ f \circ \varphi_i^{-1}} V'_j$$

is a holomorüphic map between open subsets of $\mathbb{C}$ (note that $\varphi_i^{-1}(U_i \cap f^{-1}(U'_j)) \subseteq V_i \subseteq \mathbb{C}$, and $V_j \subseteq \mathbb{C}$). That is, f is holomorphic when written in local holomorphic coordinates φ_i on X and φ'_j on Y.

Example 3.2. $X = \mathbb{C}$ is a Riemann surface with atlas $\mathcal{A} = \{(\mathbb{C}, \mathbb{C}, \mathrm{id}_{\mathbb{C}})\}$ with one chart.

Aside: we could also take $\mathcal{A} = \{(\mathbb{C}, \mathbb{C}, f)\}$, where $f(x+iy) = x - iy$, or $f(x+iy) = x + 2iy$ etc. These give *non-equivalent* Riemann surface structures on $X = \mathbb{C}$. The atlas $\mathcal{A}$ in $(X, \mathcal{A})$ is essential data. It is not enough that there exists an atlas, we need a particular choice.

Example 3.3. We make $X = \mathbb{C} \cup \{\infty\}$ into a topological surface homeomorphic to S^2, with open sets $U \subseteq \mathbb{C}$ open, together with sets of the form $U \cup \{\infty\}$ for $U \subseteq \mathbb{C}$ such that $\mathbb{C} \setminus U$ is

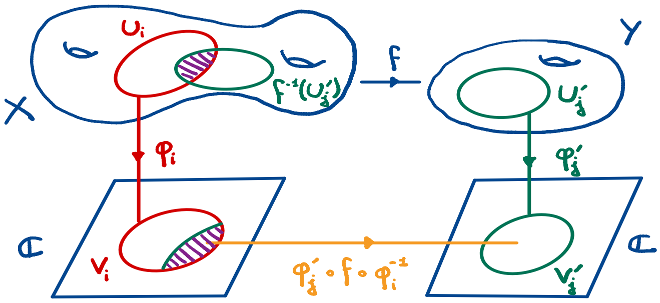


FIGURE 3.2. Holomorphic map between Riemann surfaces

compact. Then X is a Riemann surface with atlas $\mathcal{A} = \{(U_1, V_1, \varphi_1), (U_2, V_2, \varphi_2)\}$, where

$$U_1 = \mathbb{C} = X \setminus \{\infty\},\ V_1 = \mathbb{C},\ \varphi_1 = \mathrm{id}_{\mathbb{C}}$$

$$U_2 = (\mathbb{C} \setminus 0) \cup \{\infty\},\ V_2 = \mathbb{C},\ \varphi_2(z) = \begin{cases} \frac{1}{z} & \text{if } z \in \mathbb{C} \setminus \{0\} \\ 0 & \text{if } z = \infty \end{cases}$$

The transition function is,

$$\begin{aligned} \varphi_2 \circ \varphi_1^{-1} \colon \mathbb{C} \setminus \{0\} &\longrightarrow \mathbb{C} \setminus \{0\} \\ z &\longmapsto z^{-1}, \end{aligned}$$

which is holomorphic (no problem of a pole with z^{-1} as 0 is not in the domain). So, $(U_1, V_1, \varphi_1), (U_2, V_2, \varphi_2)$ are compatible. Moreover, as $X = U_1 \cup U_2$, they cover X. Thus, $(X, \mathcal{A})$ is a Riemann surface. One may write $\mathbb{C} \cup \{\infty\}$ as $\mathbb{CP}^1$ (complex projective line). We call $\mathbb{C} \cup \{\infty\}$ the *Riemann sphere*.

Example 3.4. Let $\omega_1, \omega_2 \in \mathbb{C}$ be linearly independent over $\mathbb{R}$. Write

$$\Lambda = \{m\omega_1 + n\omega_2 \colon m, n \in \mathbb{Z}\},$$

which is an additive subgroup of $\mathbb{C}$ (a *lattice*). Take $X = \mathbb{C}/\Lambda$ the quotient group, with the quotient topology. A fundamental domain for the action of Λ on $\mathbb{C}$ is the parallelogram as shown in Figure 3.3. Write vertices $o, \omega_1, \omega_1 + \omega_2, \omega_2$. Then X is obtained by identifying opposite sides of this parallelogram as shown, so topologically $X \cong T^2$, a torus. If $V \subseteq \mathbb{C}$

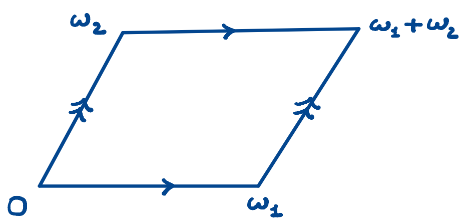


FIGURE 3.3. A fundamental domain for the action of Λ on $\mathbb{C}$

is open such that $(V + \lambda) \cap V = \emptyset$ for all $0 \neq \lambda \in \Lambda$ (this holds if V is a small ball), define

$U = \{v + \lambda \colon v \in V\} \subseteq \mathbb{C}/\Lambda$, and $\varphi \colon U \to V$ to be the inverse of the homeomorphism

$$\begin{aligned} \varphi^{-1} \colon V &\longrightarrow U \\ v &\longmapsto v + \Lambda \end{aligned}$$

Then, (U, V, φ) is a complex chart on X. Write $\mathcal{A}$ for the set of all charts of this form. If (U_1, V_1, φ_1), (U_2, V_2, φ_2) are charts of this form, one can show that the transition function $\varphi_2 \circ \varphi_1^{-1}$ is locally of the form $z \longmapsto z + \lambda$ for $\lambda \in \Lambda$, and is holomorphic. Thus the charts in $\mathcal{A}$ are pairwise compatible, and they cover X. So, $\mathcal{A}$ is a holomorphic atlas and $X = \mathbb{C}/\Lambda$ is a Riemann surface.

Aside: as a topological space $X \cong T^2$, but as a Riemann surface X depends on the lattice Λ: if $\Lambda \neq \alpha\Lambda'$ for $\alpha \in \mathbb{C} \setminus \{0\}$ then $\mathbb{C}/\Lambda \ncong \mathbb{C}/\Lambda'$.

Proposition 3.5. *Any Riemann surface X is orientable.*

Sketch of the proof: X has a holomorphic atlas $\{(U_i, V_i, \varphi_i) \colon i \in I\}$. Recall that X is orientable if it has a consistent notion of "*clockwise*". On $U_i \subset X$ we define clockwise by

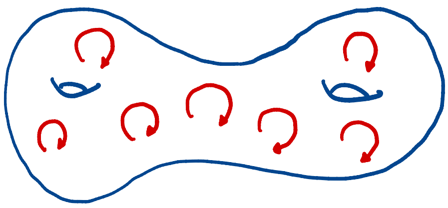


FIGURE 3.4. Natural orientation of a Riemann surface

identifying $\varphi_i \colon U_i \to V_i \subseteq \mathbb{C}$ and using standard notion of "*clockwise*" in $\mathbb{C}$. For two charts (U_i, V_i, φ_i), (U_j, V_j, φ_j) as

$$\varphi_j \circ \varphi_i^{-1} \colon \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$$

is holomorphic, it preserves angles, so it preserves the notion of "*clockwise*". Thus, the notions of "*clockwise*" on U_i, U_j agree on $U_i \cap U_j$. As $\{U_i \colon i \in I\}$ covers X, we have an orientation on X.

From the classification of surfaces (Theorem 2.27), we have:

Corollary 3.6. *Any compact, connected Riemann surface X is homeomorphic to a surface Σ_g of genus $g \geq 0$.*

3.2. **Meromorphic functions.**

Definition 3.7. Let X be a Riemann surface. A *meromorphic function* on X is a holomorphic function

$$f \colon X \to \mathbb{C} \cup \{\infty\}$$

to the Riemann sphere $\mathbb{C} \cup \{\infty\}$ as a Riemann surface.

Remark 3.8. We can also consider holomorphic functions $f \colon X \to \mathbb{C}$, of course. However, if X is a compact Riemann surface, then

$$|f| \colon X \to [0, \infty)$$

attains its maximum, and using the *maximum modulus theorem* in complex analysis, one can show f is constant. So, holomorphic functions $f \colon X \to \mathbb{C}$ are boring for compact X. It is a deep theorem, proved in B3.3 using the Riemann-Roch Theorem, that any compact Riemann surface has non-constant meromorphic functions (in fact, an infinite-dimensional family of them). So, they are a good thing to study.

Example 3.9. Let $p(z), q(z)$ be non-zero complex polynomials with no common factors. Define $f\colon \mathbb{C}\cup\{\infty\} \to \mathbb{C}\cup\{\infty\}$ by

$$f(z) = \begin{cases} \frac{p(z)}{q(z)} & \text{if } z \in \mathbb{C} \text{ and } q(z) \neq 0, \\ \infty & \text{if } z \in \mathbb{C} \text{ and } q(z) = 0. \end{cases}$$

$$f(\infty) = \begin{cases} \infty & \text{if} \deg(p) > \deg(q), \\ 0 & \text{if} \deg(p) < \deg(q), \\ \frac{\text{leading coefficient } p}{\text{leading coefficient } q} & \text{if} \deg(p) = \deg(q). \end{cases}$$

Then, $f\colon \mathbb{C}\cup\{\infty\} \to \mathbb{C}\cup\{\infty\}$ is a non-constant meromorphic function. One can show (in B3.3) that every meromorphic function on $\mathbb{C}\cup\{\infty\}$ is of this form.

Aside: Transcendental holomorphic functions such as $e^z\colon \mathbb{C}\to\mathbb{C}$ do not extend to holomorphic functions $\mathbb{C}\cup\{\infty\} \to \mathbb{C}\cup\{\infty\}$. E.g. $e^{\frac{1}{z}}$ has an "*essential singuarity*" at $z = 0$. We can't define $f(z) = e^{\frac{1}{z}}$ if $z \neq 0$ and $f(0) = \infty$, as this would not be continuous $f\colon \mathbb{C} \to \mathbb{C}\cup\{\infty\}$.

Example 3.10. Let $\omega_1, \omega_2 \in \mathbb{C}$ be linearly independent over $\mathbb{R}$. Write

$$\Lambda = \{m\omega_1 + n\omega_2 \colon m, n \in \mathbb{Z}\},$$

and $X = \mathbb{C}/\Lambda$, as in Example 3.4. Define the *Weierstrass $\wp$-function*,

$$\wp(z) = \begin{cases} \frac{1}{z^2} + \sum_{0\neq\omega\in\Lambda} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right) & \text{if } z \in \mathbb{C}\setminus\Lambda, \\ \infty & \text{if } z \in \Lambda. \end{cases}$$

One can prove:

- The sum converges uniformly on compact subsets of $\mathbb{C}\setminus\Lambda$ to a holomorphic function.
- $\wp(z)$ is meromorphic with a double pole at each $\omega \in \Lambda$, $\wp(z) = \frac{1}{(z-\omega)^2} + O(1)$, where $O(1)$ stands for all the terms that remain *bounded* as $z \to \omega$ (i.e. there is no first order pole of $\wp(z)$).
- $\wp(z) = \wp(z+\omega)$, for all $z \in \mathbb{C}$ and $\omega \in \Lambda$ (it's doubly periodic).

Hence, it descends to a meromorphic function

$$\wp\colon \mathbb{C}/\Lambda \longrightarrow \mathbb{C}\cup\{\infty\},$$

with one pole at $0+\Lambda$. One can use this to build other meromorphic functions on $\mathbb{C}/\Lambda$, e.g. $\wp', \wp'', \frac{1}{\wp+c}, \ldots$

3.3. **Branch points and ramification points.** *Recall some facts from complex analysis:*

Let $U \subseteq \mathbb{C}$ be open and $f : U \to \mathbb{C}$ be holomorphic. If f is not locally constant, the zeroes of f are isolated in U. Thus, the zeroes of $\frac{df}{dz}$ are also isolated.

For any $a \in U$, as f has a Taylor series at a, if f is not locally constant there is a unique $m \geq 1$ such that $\frac{d^m f}{dz^m}(a) \neq 0$ and $\frac{d^k f}{dz^k}(a) = 0$, for $k = 1, \ldots, m-1$.

Definition 3.11. Let X, Y be Riemann surfaces and $f\colon X \to Y$ a holomorphic map, which is not locally constant on X. Let $x \in X$ with $f(x) = y \in Y$, and choose charts (U, V, φ) on X and (U', V', φ') on Y with $x \in U$, $y \in U'$. Then,

$$\varphi' \circ f \circ \varphi^{-1} \colon \varphi(U \cap f^{-1}(U')) \longrightarrow V'$$

is holomorphic and not locally constant. Note that $\varphi(x) \in \varphi(U \cap f^{-1}(U'))$. We call x a _ramification point_ and y a _branch point_ of f, if $\frac{d}{dz}(\varphi' \circ f \circ \varphi^{-1})|_{\varphi(x)} = 0$.

One can check that Definition 3.11 is independent of the choice of charts. The _ramification index_ $\nu_f(x)$ of $x \in X$ is the unique $m \geq 1$ with $\left(\frac{d}{dz}\right)^m(\varphi' \circ f \circ \varphi^{-1})|_{\varphi(x)} \neq 0$ and $\left(\frac{d}{dz}\right)^k(\varphi' \circ f \circ \varphi^{-1})|_{\varphi(x)} \neq 0$, for $k = 1, \ldots, m-1$. Then, $\nu_f(x) > 1$ iff x is a ramification point.

Ramification points are isolated in X, since zeroes of $\frac{df}{dz}$ are isolated. Thus, if X is compact, then f has finitely many ramification points. If $f\colon X \to Y$ has ramification index m at x with $f(x) = y$, one can choose holomorphic coordinates w on X near x and z on Y near y with x at $w = 0$ and y at $z = 0$, such that $f\colon w \mapsto z = w^m$. That is, a ramification point $x \in X$ is a point where f looks locally like the function $w \mapsto w^m$ in holomorphic coordinates for $m > 1$. Notice that if y' is close to y, then $f^{-1}(y')$ contains m points close to x, with $m = \nu_f(x)$, as a small $\varepsilon \neq 0$ has m-many m^{th} roots.

Definition 3.12. Let X, Y be compact Riemann surfaces with Y connected, and $f\colon X \to Y$ be holomorphic and not locally constant. Then, f has finitely many ramification points $x_1, \ldots, x_k$ and so finitely many branch points $y_1, \ldots, y_k$ with $f(x_i) = y_i$ (The y_i need not be distinct). There is a number $d = \deg(f)$, called the _degree_ of f, such that if $y \in Y \backslash \{y_1, \ldots, y_k\}$ then $|f^{-1}(y)| = d$. This holds as on $X \backslash \{x_1, \ldots x_k\}$, f looks locally like a holomorphic function $f(z)$ with $\frac{df}{dz} \neq 0$, so it is locall invertible and locally maps d sheets to 1 sheet as illustrated in Figure 3.6. The number d is locally constant on $Y \setminus \{y_1, \ldots, y_k\}$ as X is compact, and globally constant as $Y \setminus \{y_1, \ldots, y_k\}$ is connected. So, $\deg f$ is well-defined.

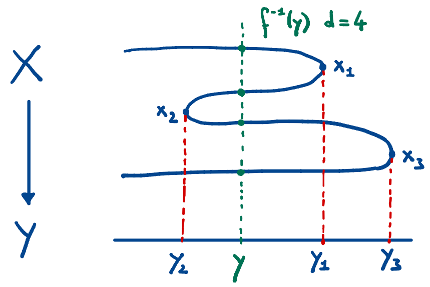


FIGURE 3.5. A degree $d = 4$ holomorphic map between Riemann surfaces

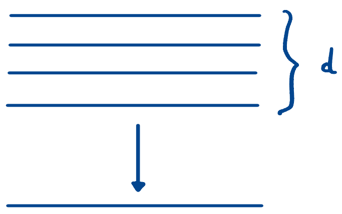


FIGURE 3.6. Local degree d map

In fact, for all $y \in Y$, we have

$$d = \sum_{\substack{x \in X \\ f(x)=y}} \nu_f(x)\,. \tag{3.1}$$

This holds for $y \in Y \setminus \{y_1, \dots y_k\}$ as $\nu_f(x) = 1$ except at $x_1, \dots, x_k$. At a ramification point x_i, as $y \to y_i$, $\nu_f(x_i)$ points in $f^{-1}(y)$ come together at x_i, so replace $1 + \dots + 1$ by $\nu_f(x_i)$ in the sum in (3.1).

Theorem 3.13. *Let X, Y be compact, connected Riemann surfaces, and $f\colon X \to Y$ a non-constant holomorphic map of degree d, with ramification points $x_1, \dots, x_k$. Then,*

$$\chi(X) = d \cdot \chi(Y) - \sum_{i=1}^{k} (\nu_f(x_i) - 1)\,.$$

This is known as the Riemann-Hurwitz formula.

Proof. Let $y_i = f(x_i)$, so $y_1, \dots, y_k$ are the (not necessarily distinct) branch points of f. Choose a triangulation of Y whose vertices include the branch points $y_1, \dots, y_k$ of f.

Over the interiors of edges and faces, f is locally invertible, and maps d points to one point. So, we can lift the triangulation of Y to a triangulation of X, in which each edge and face lifts to d edges and faces. Each vertex y lifts to $f^{-1}(y)$ points, where $\sum_{x \in f^{-1}(y)} \nu_f(x) = d$. Hence,

$$|f^{-1}(y)| = \sum_{x \in f^{-1}(y)} 1 = d - \sum_{x \in f^{-1}(y)} (\nu_f(x) - 1)\,.$$

Let the triangulation of Y have V, E, F vertices, edges, faces. Thus,

$$\begin{aligned}\chi(X) &= (dV - \sum_{i=i}^{k} (\nu_f(x_i) - 1)) - dE + dF \\ &= d(V - E + F) - \sum_{i=1}^{k} (\nu_f(x_i) - 1) \\ &= d \cdot \chi(Y) - \sum_{i=1}^{k} (\nu_f(x_i) - 1)\,.\end{aligned}$$

□

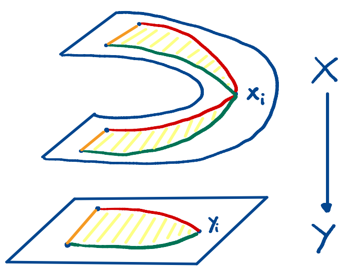


FIGURE 3.7. Lifting triangulations

Remark 3.14. Given a meromorphic function $f\colon X \to \mathbb{C} \cup \{\infty\} = S^2$, one can use this to construct a triangulation of X – compare Theorem 2.15.

3.4. **An example.** Let $w_1, w_2 \in \mathbb{C}$ be linearly independent over $\mathbb{R}$,

$$\Lambda = \{mw_1 + nw_2 \, : m, n \in \mathbb{Z}\} ,$$

and $X = \mathbb{C}/\Lambda$ as in Example 3.4.

Define $\wp\colon X \to \mathbb{C} \cup \{\infty\}$ as in Example 3.10, by

$$\wp(z) = \begin{cases} \frac{1}{z^2} + \sum_{0 \neq \omega \in \Lambda} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right) & \text{if } z \in \mathbb{C} \setminus \Lambda, \\ \infty & \text{if } z \in \Lambda. \end{cases}$$

Observe: there is a unique point $0 + \Lambda$ in X with $\wp(0 + \Lambda) = \infty$. As this is a double pole, $\wp(0 + \Lambda) = \frac{1}{z^2} + O(1)$. Using $\frac{1}{z}$ as the coordinate on $\mathbb{C} \cup \{\infty\}$ near ∞, we see that $\wp$ has a ramification index 2 at $0 + \Lambda$. Hence, the degree of $\wp$ is

$$d = \sum_{\substack{x \in X \\ \wp(x) = \infty}} \nu_\wp(x) = 2 \, .$$

What about other ramification points ? These occur when $\wp'(z) = 0$. Since $\wp(z) = \wp(-z)$ and $\wp(z) = \wp(w_1 + z)$, we have $\wp(\frac{w_1}{2} + z) = \wp(\frac{w_1}{2} - z)$, i.e. $\wp$ is even around $\frac{w_1}{2}$, so $\wp'(\frac{w_1}{2}) = 0$.

Hence, $\frac{w_1}{2} + \Lambda$ is a ramification point of $\wp$, of ramification index $\nu_\wp(\frac{w_1}{2} + \Lambda) \geq 2$. Similarly, $\frac{w_2}{2} + \Lambda$ and $\frac{w_1 + w_2}{2} + \Lambda$ are ramification points.

We have $\chi(X) = 0$ as $X \cong T^2$, $\chi(\mathbb{C} \cup \{\infty\}) = 2$ as $\mathbb{C} \cup \{\infty\} \cong S^2$. So, Riemann-Hurwitz gives

$$0 = 2 \cdot 2 - \sum_{i=1}^{k} (\nu_f(x_i) - 1) \, .$$

Hence, $0 + \Lambda$, $\frac{w_1}{2} + \Lambda$, $\frac{w_2}{2} + \Lambda$ and $\frac{w_1 + w_2}{2} + \Lambda$ are the only ramification points, and all have ramification index 2. Note that one can also see that the ramification index equals 2, as the ramification index is less than or equal to degree, by using $d = \sum_{x\colon f(x) = y} \nu_f(x)$.

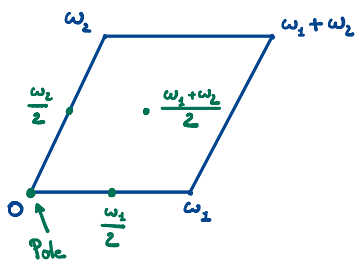


FIGURE 3.8. Zeroes and poles of the $\mathfrak{p}$ function

3.5. **Building Riemann surfaces as branched double covers.** In Example 3.4, $X = \mathbb{C}/\Lambda$ is a compact Riemann surface with a meromorphic map $f\colon X \to \mathbb{C} \cup \{\infty\}$ of degree 2 with ramification points $x_1, \ldots, x_4 \in X$ and branch points $y_1, \ldots, y_4 \in \mathbb{C} \cup \{\infty\}$. In fact, we can build X as a Riemann surface just out of $\mathbb{C} \cup \{\infty\}$ and $y_1, \ldots, y_4$. Choose "*cuts*" in $\mathbb{C} \cup \{\infty\}$ from y_1 to y_2 and y_3 to y_4. Take 2 copies $(\mathbb{C} \cup \{\infty\})_a$, $(\mathbb{C} \cup \{\infty\})_b$, cut both along the line segments $y_1 \to y_2$, $y_3 \to y_4$, and glue the cut edges together swapping over the $'a'$ and $'b'$ copies.

We can also do this using any even number of ramification points $y_1, \ldots, y_{2n}$ in $\mathbb{C} \cup \{\infty\}$. That is, we can define a compact, connected Riemann surface X with a holomorphic map

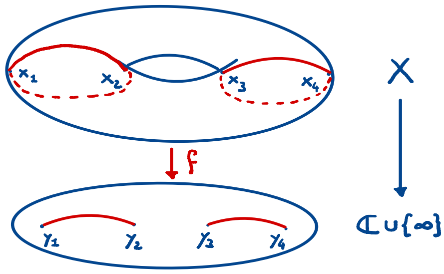


FIGURE 3.9. $X = \mathbb{C}/\Lambda$ as a double cover of $\mathbb{C} \cup \{\infty\}$ branched at 4 points

$f\colon X \to \mathbb{C} \cup \{\infty\}$ of degree 2 with branch points at $y_1, \ldots y_{2n}$. This X is unique up to isomorphism. Such Riemann surfaces are called *hyperelliptic*.

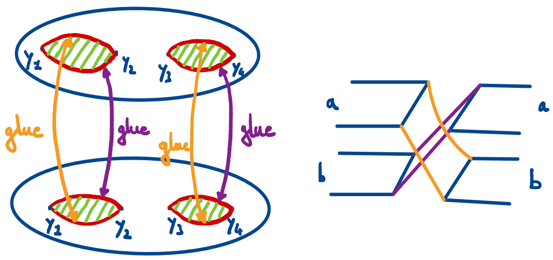


FIGURE 3.10. Recovering $X = \mathbb{C}/\Lambda$ from 4 points in $\mathbb{C} \cup \{\infty\}$

Here is a more algebraic way to define X. For simplicity, take $y_1, \ldots, y_{2n-1} \in \mathbb{C}$ and $y_{2n} = \infty$. Define

$$X = \{(w, x) \in \mathbb{C}^2 \, : w^2 = (x - y_1) \cdots (x - y_{2n-1})\} \cup \{(\infty, \infty)\} \, .$$

One can show that X has the natural structure of a Riemann surface, such that $f\colon X \to \mathbb{C} \cup \{\infty\}$, $f(w, x) = x$, is degree 2 meromorphic with branch points $y_1, \ldots y_{2n}$.

Note that $w = \sqrt{(x - y_1) \cdots (x - y_{2n-1})}$ is also a rational function on X. Hyperelliptic surfaces occur in problems involving *hyperelliptic integrals* $\int \frac{dx}{\sqrt{(x-y_1)\cdots(x-y_{2n-1})}}$.

4. Smooth surfaces

4.1. **Abstract smooth surfaces.** We define smooth surfaces as for Riemann surfaces, but replace "$\mathbb{C}$, holomorphic" by "$\mathbb{R}^2$, smooth".

Definition 4.1. Let X be a topological surface. A *(smooth) chart* on X is a triple (U, V, φ) with $U \subseteq X$, $V \subseteq \mathbb{R}^2$ open and $\varphi : U \to V$ a homeomorphism.

Two charts (U_1, V_1, φ_1) and (U_2, V_2, φ_2) are *compatible* if

$$\varphi_2 \circ \varphi_1^{-1} \colon \varphi_1(U_1 \cap U_2) \longrightarrow \varphi_2(U_1 \cap U_2)$$

is a smooth map between open subsets of $\mathbb{R}^2$, with smooth inverse. Here, *smooth*, or C^∞, means all partial derivatives $\frac{\partial^{k+\ell}(\varphi_2 \circ \varphi_1^{-1})}{\partial x^k \partial y^\ell}$ exist for all $k, \ell \geq 0$. A *(smooth) atlas* $\mathcal{A} = \{(U_i, V_i, \varphi_i) \, : i \in I\}$ on X is a family of smooth charts on X, such that

(i) (U_i, V_i, φ_i) and (U_j, V_j, φ_j) are compatible, for all $i, j \in I$,
(i) $X = \bigcup_{i\in I} U_i$.

A *(smooth) surface* $(X, \mathcal{A})$ is a topological surface X with a smooth atlas $\mathcal{A}$.

One can also define *smooth maps* $f\colon X \to Y$ between smooth surfaces X, Y, by f continuous and

$$\varphi_j^{-1} \circ f \circ \varphi_i^{-1} \colon \varphi_i^{-1}(\varphi_i \cap f^{-1}(U_j')) \longrightarrow V_j'$$

is smooth between open subsets of $\mathbb{R}^2$, for all charts (U_i, V_i, φ_i) on X and (U_j', V_j', φ_j') on Y.

Remarks 4.2. (a) One can generalize the above to *smooth manifolds* of dimension n by taking $V \subseteq \mathbb{R}^n$ open, not $V \subseteq \mathbb{R}^2$ – see C3.3 Differentiable Manifolds.
(b) Every Riemann surface is a smooth surface, by identifying $\mathbb{C} \cong \mathbb{R}^2$, and then transition functions $\varphi_j \circ \varphi_i^{-1}$ are holomorphic $\implies$ $\varphi_j \circ \varphi_i^{-1}$ smooth, so a holomorphic atlas is also a smooth atlas.

4.2. Smooth surfaces in $\mathbb{R}^3$.

Definition 4.3. A *smooth surface in* $\mathbb{R}^3$ is a subset $X \subset \mathbb{R}^3$ such that each point $x \in X$ has an open neighbourhood $x \in U \subseteq X$ and a map $\underline{r}\colon V \to X \subset \mathbb{R}^3$ from $V \subseteq \mathbb{R}^2$ open, such that:
(i) $\underline{r}\colon V \to U$ is a homeomorphism.
(ii) $\underline{r}(u, v) = (x(u, v), y(u, v), z(u, v))$ has derivatives of all orders.
(iii) At each $(u, v) \in V$, $\underline{r}_u = \frac{\partial}{\partial u}$ and $\underline{r}_v = \frac{\partial}{\partial v}$ are linearly independent in $\mathbb{R}^3$.

We call $\underline{r}$ satisfying $(i) - (iii)$ a *local parametrization* of the surface.

If $X \subset \mathbb{R}^3$ is a smooth surface, define

$$\mathcal{A} = \{(U, V, \varphi) \ : \text{for } U, V \text{ and } \underline{r} \text{ as above } \varphi = \underline{r}^{-1}\colon U \to V\}\,.$$

Then, each (U, V, φ) is a chart, and one can show that any two such charts are compatible (needs the *Inverse Function Theorem*). Then, $\mathcal{A}$ is a smooth atlas on X and makes X into an abstract smooth surface.

Example 4.4. (a) $S^2 = \{(x, y, z) \in \mathbb{R}^3 \ : x^2 + y^2 + z^2 = 1\}$, the unit sphere is a smooth surface in $\mathbb{R}^3$.

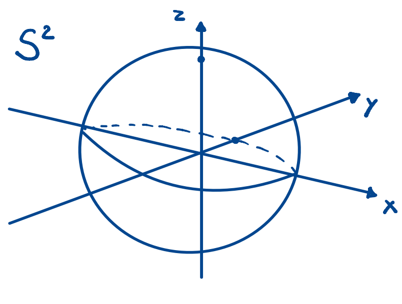


FIGURE 4.1. The unit sphere $S^2 \subset \mathbb{R}^3$

(b) The hyperboloids

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

are smooth surfaces in $\mathbb{R}^3$. But the cone $C = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2-z^2 = 0\}$ is not : no chart exists around $(0,0,0)$.

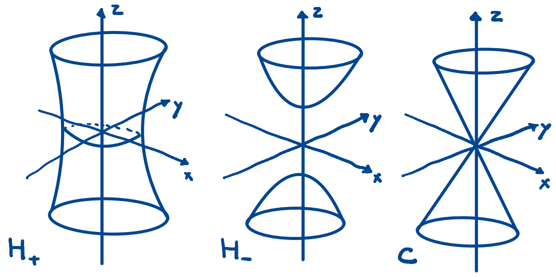


FIGURE 4.2. The hyperboloids H_+, H_-, and the cone C in $\mathbb{R}^3$

Definition 4.5. Let X be a surface in $\mathbb{R}^3$ and $x \in X$. Let $\underline{r}\colon V \to X \subset \mathbb{R}^3$ be a local parametrization of $\mathbb{R}^3$, with $\underline{r}(u,v) = x$. The <u>*tangent space*</u> T_xX is the vector subspace

$$T_xX = \langle \underline{r}_u|_{(u,v)}, \underline{r}_v|_{(u,v)} \rangle .$$

Note that $\underline{r}_u|_{(u,v)}, \underline{r}_v|_{(u,v)}$ are linearly independent by definition.

The two <u>*unit normals*</u> to X at x are

$$\pm \underline{n} = \pm \frac{\underline{r}_u \wedge \underline{r}_v}{|\underline{r}_u \wedge \underline{r}_v|} = \pm 1 .$$

Then $T_xX \subset \mathbb{R}^3$ and $\{\pm \underline{n}\}$ are independent of the choice of the parametrization $\underline{r}$, since any other parametrization $\underline{\hat{r}}$ is locally of the form

$$\underline{\hat{r}}(u,v) = \underline{r}(\hat{u}(u,v), \hat{v}(u,v)) ,$$

for $(\hat{u},\hat{v})\colon (\text{open in } \mathbb{R}^2) \to (\text{open in } \mathbb{R}^2)$ smooth and invertible. Then, $\begin{pmatrix} \underline{\hat{r}}_u \\ \underline{\hat{r}}_v \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \underline{r}_u \\ \underline{r}_v \end{pmatrix}$, where $\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \frac{\partial \hat{u}}{\partial u} & \frac{\partial \hat{u}}{\partial v} \\ \frac{\partial \hat{v}}{\partial u} & \frac{\partial \hat{v}}{\partial v} \end{pmatrix}$ is invertible, as $(\hat{u},\hat{v})$ is smooth and invertible, so T_xX and $\{\pm\underline{n}\}$ are independent of choices.

$$T_xX = \{w \in \mathbb{R}^3 : w \cdot \underline{n} = 0\} ,$$

orthogonal subspace to unit normal $\underline{n}$.

4.3. **The first fundamental form.** Let X be a surface in $\mathbb{R}^3$. Using the usual notion of distance in $\mathbb{R}^3$, we can define <u>*lengths of curves*</u> γ in X, and <u>*areas of regions*</u> $U \subseteq X$. We will do this using the <u>*first fundamental form*</u> of X.

<u>*Note:*</u> If X is an abstract smooth surface not embedded in $\mathbb{R}^3$, we can not define lengths of curves or areas of regions without choosing an extra structure, a <u>*Riemann metric*</u>, which is what the first fundamental form really is.

Definition 4.6. Let $X \subset \mathbb{R}^3$ be a smooth surface. Let $\underline{r}\colon V \to X$ be a smooth parametrization, for $V \subseteq \mathbb{R}^2$ open. Define smooth functions $E, F, G\colon V \to \mathbb{R}$ by

$$E = \underline{r}_u \cdot \underline{r}_u, \;\; F = \underline{r}_u \cdot \underline{r}_v = \underline{r}_v \cdot \underline{r}_u, \;\; G = \underline{r}_v \cdot \underline{r}_v ,$$

The <u>*first fundamental form*</u> of X is the expression

$$g = Edu^2 + 2Fdudv + Gdv^2 .$$

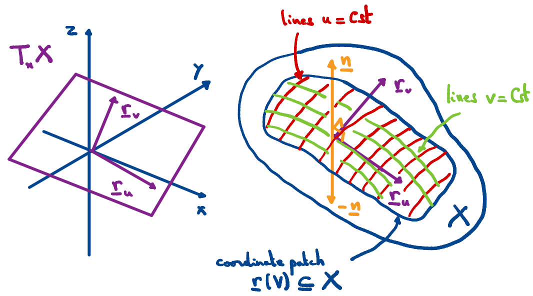


FIGURE 4.3. Tangent space and unit normals to a surface in $\mathbb{R}^3$

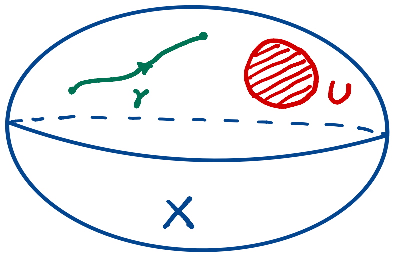


FIGURE 4.4. Curves and regions on a surface in $\mathbb{R}^3$

This is the quadratic form $Q(\underline{v}, \underline{v}) = \underline{v} \cdot \underline{v}$ on $T_xX \subset \mathbb{R}^3$ in the basis $\underline{r}_u, \underline{r}_v$.

We can write $g = \begin{pmatrix} du & dv \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}$, for $\begin{pmatrix} E & F \\ F & G \end{pmatrix}$ a symmetric, positive definite matrix of functions.

Remark 4.7. You probably wonder "What are $du, dv, du^2, dudv, dv^2$? " To do this properly: du, dv are smooth sections of the *cotangent bundle* $T^*X \to X$, and $g, du^2, dudv, dv^2$ are sections of the *tensor product* $\otimes^2 T^*X \to X$. Beyond the scope of this course, explained in C3.3.

For now: $du, du^2, \dots$ are formal symbols which make sense if you have a choice of local coordinates on X. They behave as you expect under change of coordinates. E.g. if you have two coordinates $(u, v), (x, y)$ on X with $u = u(x, y)$, $v = v(x, y)$, then $du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy$, $dv = \frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy$, $du^2 = \left(\frac{\partial u}{\partial x}\right)^2 dx^2 + 2 \frac{\partial u}{\partial x} \frac{\partial u}{\partial y} dxdy + \left(\frac{\partial u}{\partial y}\right)^2 dy^2$, etc. You may think of du as the "*derivative of u*", but without taking partial derivatives w.r.t. particular coordinates. At $x \in X$, du lies in the vector space $T_x^*X = \langle du, dv \rangle$ dual to T_xX.

We can use the first fundamental form to compute lengths of curves in X.

Definition 4.8. Let $X \subset \mathbb{R}^3$ be a smooth surface, and let $\gamma \colon [a, b] \to X \subset \mathbb{R}^3$ be a smooth curve. Suppose $\underline{r} \colon V \to X$ is a local parametrization of X, and γ factors through $\underline{r}$, i.e.

$\gamma(t) = \underline{r}(u(t), v(t))$. Then, the length of γ is

$$\begin{aligned}\ell(\gamma) = \int_a^b \left|\frac{d\gamma}{dt}\right| dt &= \int_a^b \left|\frac{d}{dt}(\underline{r}(u(t), v(t)))\right| dt \qquad (4.1)\\ &= \int_a^b \left(\left|\underline{r}_u \frac{du}{dt} + \underline{r}_v \frac{dv}{dt}\right|^2 \right)^{\frac{1}{2}} dt \\ &= \int_a^b \left(E(u,v)\left(\frac{du}{dt}\right)^2 + 2F\left(\frac{du}{dt}\right)\left(\frac{dv}{dt}\right) + G(u,v)\left(\frac{dv}{dt}\right)^2\right)^{\frac{1}{2}} dt\,.\end{aligned}$$

So, we can write lengths of curves using the first fundamental form $Edu^2 + 2Fdudv + Gdv^2$. Note that formally

$$\left(E(u,v)\left(\frac{du}{dt}\right)^2 + 2F\left(\frac{du}{dt}\right)\left(\frac{dv}{dt}\right) + G(u,v)\left(\frac{dv}{dt}\right)^2\right)^{\frac{1}{2}} dt = (Edu^2 + 2Fdudv + Gdv^2)^{\frac{1}{2}}\,,$$

if we cancel the dt's.

We can also write areas in terms of first fundamental forms:

Definition 4.9. Let $X \subset \mathbb{R}^3$ be a smooth surface, of finite area, and suppose X is covered by a single parametrization $\underline{r}\colon V \to X \subset \mathbb{R}^3$. Then,

$$\begin{aligned}\operatorname{area}(X) &= \int_V |\underline{r}_u \wedge \underline{r}_v| dudv \qquad (4.2)\\ &= \int_V (|\underline{r_u}|^2 |\underline{r}_v|^2 - |\underline{r}_u \cdot \underline{r}_v|^2)^{\frac{1}{2}} dudv \\ &= \int_V (EG - F^2)^{\frac{1}{2}} dudv \\ &= \int_V \det \left|\begin{pmatrix} E & F \\ F & G\end{pmatrix}\right|^{\frac{1}{2}} dudv\end{aligned}$$

4.4. **Riemann metrics on abstract surfaces.**

Definition 4.10. Let X be a smooth surface with atlas $\mathcal{A} = \{(U_i, V_i, \varphi_i)\colon i \in I\}$. A <u>*Riemann metric*</u> g on X is the data

$$E_i du^2 + 2F_i dudv + G_i dv^2\,,$$

on V_i where $E_i, F_i, G_i\colon V_i \to \mathbb{R}$ are smooth functions, for each $i \in I$ satisfying:

(a) $\begin{pmatrix} E_i & F_i \\ F_i & G_i\end{pmatrix}$ is a positive definite matrix at each point of V_i.

(b) Let $i, j \in I$, and write the map

$$\varphi_j \circ \varphi_i^{-1}\colon \varphi_i(U_i \cap U_j) \longrightarrow \varphi_j(U_i \cap U_j)\,,$$

as $(u, v) \mapsto (\tilde{u}, \tilde{v})$. Then,

$$E_j d\tilde{u}^2 + 2F_j d\tilde{u}d\tilde{v} + G_j d\tilde{v}^2 = E_j du^2 + 2F_j dudv + G_j dv^2\,,$$

with $d\tilde{u} = \frac{\partial \tilde{u}}{\partial u} du + \frac{\partial \tilde{u}}{\partial v} dv$ and $d\tilde{v} = \frac{\partial \tilde{v}}{\partial u} du + \frac{\partial \tilde{v}}{.} \partial v dv$. That is,

$$E_i|_{\text{open set}} = E_j \left(\frac{\partial \tilde{u}}{\partial u}\right)^2 + 2F_j \frac{\partial \tilde{u}}{\partial u}\frac{\partial \tilde{v}}{\partial u} + G_j \left(\frac{\partial \tilde{v}}{\partial u}\right)^2, \text{etc.}$$

Then, we can define <u>*lengths of curves*</u> $\gamma : [a,b] \to X$ and <u>*areas of regions*</u> $X' \subset X$ by formulae (4.1) and (4.2) in §4.3.

Any surface $X \subset \mathbb{R}^3$ has the natural structure of an abstract smooth surface, as in §4.2, and has a natural Riemannian metric as in §4.3. But Riemannian metrics on surfaces do not have to come from embeddings in $\mathbb{R}^3$ – indeed, some surfaces (e.g. $\mathbb{RP}^2, K$) can not be embedded in $\mathbb{R}^3$, but every smooth surface admits Riemannian metrics.

Remark 4.11. For an abstract surface X, one can define a <u>*tangent space*</u> T_xX at each $x \in X$, with basis $\frac{\partial}{\partial u}, \frac{\partial}{\partial v}$ if u, v are local coordinates on X near x, but can make T_xX coordinate independent. The <u>*cotangent space*</u> is $T_x^*X = (T_xX)^*$ with basis du, dv. One can interpret a Riemannian metric g as giving a positive definite quadratic form $\underline{v} \mapsto g_x(\underline{u}, \underline{v})$ on each tangent space T_xX varying smoothly with x, which determines the (squared) lengths of vectors. In the dual basis du, dv one can write $g = Edu^2 + 2Fdudv + Gdv^2$.

4.5. Examples of first fundamental forms.

Example 4.12. $X = \{(x,y,z) \in \mathbb{R}^3 \colon z = 0\}$, the (x,y)-plane, has parametrization $\underline{r}(u,v) = (u,v,0)$, $\underline{r}_u = (1,0,0)$, $\underline{r}_v = (0,1,0)$, and first fundamental form (1FF) $g = du^2 + dv^2$.

Example 4.13. Let $X = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = R^2\}$ be the sphere of radius $R > 0$ in $\mathbb{R}^3$. Define spherical polar coordinates

$$\underline{r} : (o,\pi) \times (0,2\pi) \longrightarrow X \subset \mathbb{R}^3 \text{ by}$$

$$\underline{r}(\theta,\varphi) = (R\sin\theta\cos\varphi, R\sin\theta\sin\varphi, R\cos\theta)\,.$$

Then,

$$\underline{r}_\theta = (R\cos\theta\cos\varphi, R\cos\theta\sin\varphi, -R\sin\theta),$$
$$\underline{r}_\varphi = (-R\sin\theta\sin\varphi, R\sin\theta\cos\varphi, 0),$$

(1FF) is $g = R^2(d\theta^2 + \sin^2\theta d\varphi^2)$.

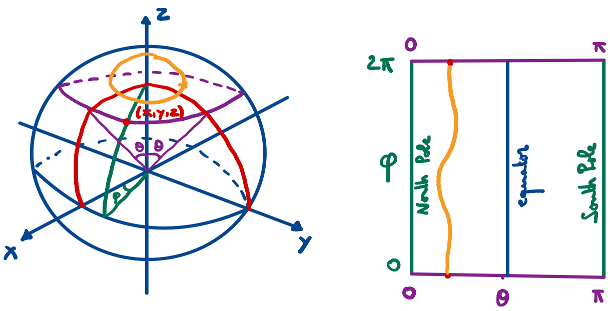


FIGURE 4.5. Spherical coordinates on the unit sphere. Think of this as a map of the Earth's surface, where (θ,φ) = (lattitude, longitude). The map distorts distances, angles and areas. Knowing the (1FF) $g = R^2(d\theta^2 + \sin^2\theta d\varphi^2)$ allows you to compute distances, angles, and areas on the Earth's surface, just using the map. You don't need to know how the map is embedded in $\mathbb{R}^3$.

4.6. **Isometric surfaces.**

Definition 4.14. Let $X, Y \subset \mathbb{R}^3$ be surfaces in $\mathbb{R}^3$. We call X, Y *isometric* if there is a smooth homeomorphism $f: X \to Y$ which maps curves $\gamma: [a,b] \to X$ to curves $f \circ \gamma: [a,b] \to Y$ of the same length. That is, isometries preserve lengths of curves. If (X, g) and (Y, h) are abstract smooth surfaces with Riemannian metrics, we also call $f: X \to Y$ and isometry if it is a smooth homeomorphism which preserves lengths of curves.

Proposition 4.15. *A smooth homeomorphism $f : X \to Y$ is an isometry iff whenever $\underline{r}: V \to X$ is a smooth local parametrization, $V \subseteq \mathbb{R}^2$ open, then $f \circ \underline{r}: V \to Y$ is a smooth local parametrization and $\underline{r}, f \circ \underline{r}$ have the same first fundamental form $Edu^2 + 2Fdudv + Gdv^2$.*

Proof. "if": obvious as lengths of curves are computed using (1FF).

"only if": Fix $(u, v) \in V$ and $(u', v') \in \mathbb{R}^2$. Define $\gamma_\epsilon: (0,1) \to V$ by

$$\gamma_\epsilon(t) = (u + \epsilon t u', v + \epsilon t v'),$$

for $\epsilon > 0$ small. Then,

$$\lim_{\epsilon \to 0} \frac{\text{length}(\underline{r} \circ \gamma_\epsilon)^2}{\epsilon^2} = E(u,v)(u')^2 + 2F(u,v)u'v' + G(u,v)(v')^2$$

One can recover $E(u,v), F(u,v), G(u,v)$ by $(u', v') = (1,0), (0,1), (1,1)$. If $f \circ \underline{r}$ is a smooth parametrization, then f is length preserving $\implies$ f identifies (1FF)s. (One can also show f must be a smooth local parametrization using Inverse Function Theorem, as otherwise one would have $0 \neq (u', v')$ in $\text{Ker}(df)$.) ☐

Two surfaces in $\mathbb{R}^3$ can be isometric even if they do not differ by an ambient isometry of $\mathbb{R}^3$.

Example 4.16. Consider the plane $P = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}$ with parametrization $\underline{r}(u, v) = (u, v, 0)$ and the cylinder

$$C = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1\},$$

with parametrization $\underline{r}'(u, v) = (\cos u, \sin u, v)$. Both have (1FF) $du^2 + dv^2$. So, mapping $\underline{r}(u, v) \mapsto \underline{r}'(u, v)$ gives a (local) isometry from P to C.

Explanation: One can make the cylinder C by rolling up a piece of paper P; "rolling up" does not change distances in P.

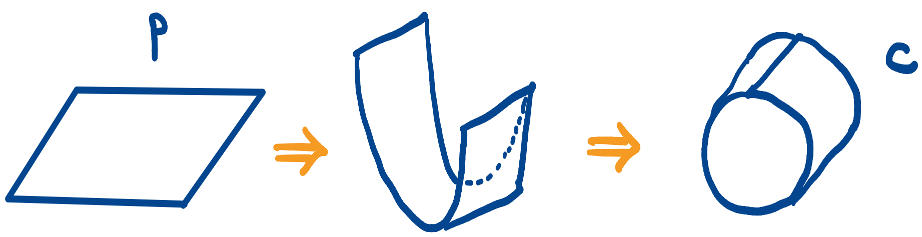


FIGURE 4.6. Rolling up a piece of paper P into a cylinder C

Remark 4.17. It is an important question when a surface $X \subset \mathbb{R}^3$ or (X, g) is locally isometric to the plane P. We will answer this using Gaussian curvature $\kappa: X \to \mathbb{R}$: X is locally isometric to $(\mathbb{R}^2, du^2 + dv^2)$ iff $K = 0$.

4.7. **The second fundamental form.**

Definition 4.18. Let $X \subset \mathbb{R}^3$ be a smooth surface, and $\underline{r}\colon V \to X$ be a local parametrization, for $V \subseteq \mathbb{R}^2$ open. As in §??, the unit normal is

$$\underline{n} = \frac{\underline{r}_u \wedge \underline{r}_v}{|\underline{r}_u \wedge \underline{r}_v|}\,.$$

(or minus of this: we now choose a sign, which is equivalent to choosing an orientation on X)

The second fundamental form of X is the expression

$$II = Ldu^2 + 2Mdudv + Ndv^2\,,$$

where $L, M, N\colon V \to \mathbb{R}$ are the smooth functions $L = \underline{r}_{uu} \cdot \underline{n}$, $M = \underline{r}_{uv} \cdot \underline{n}$, $N = \underline{r}_{vv} \cdot \underline{n}$.

Since $\underline{r}_u \cdot \underline{n} = \underline{r}_v \cdot \underline{n} = 0$, by differentiating we get alternative expressions

$$L = -\underline{r}_u \cdot \underline{n}_u,\ \ M = -\underline{r}_u \cdot \underline{n}_v = -\underline{r}_v \cdot \underline{n}_u,\ \ N = -\underline{r}_v \cdot \underline{n}_v\,.$$

We can write

$$II = \begin{pmatrix} du & dv \end{pmatrix} \begin{pmatrix} L & M \\ M & N \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}\,,$$

$\begin{pmatrix} L & M \\ M & N \end{pmatrix}$ a symmetric matrix of functions. The second fundamental form has the same behaviour under change of coordinates as the first fundamental form does (although $\begin{pmatrix} L & M \\ M & N \end{pmatrix}$ does not need to be positive definite). It is a geometric structure of the same kind (section of $\otimes^2 T^*X$).

Remark 4.19. The second fundamental form (2FF) depends on the embedding $X \hookrightarrow \mathbb{R}^3$. It does not make sense for abstract surfaces (X, g) with Riemannian metrics. Isometric surfaces in $\mathbb{R}^3$ need not have the same (2FF). Changing orientation changes the sign of the (2FF).

Example 4.20. The plane $P = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}$, with parametrization $\underline{r}(u, v) = (u, v, 0)$ and normal vector $\underline{n} = (0, 0, 1)$ has (1FF) $g = du^2 + dv^2$ and (2FF) $II = 0$. The cylinder $C = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1\}$ with parametrization $\underline{r}(u, v) = (\cos u, \sin v, v)$ and normal $\underline{n} = (\cos u, \sin u, 0)$ has $g = du^2 + dv^2$ and (2FF) $II = -du^2$. P and C are locally isometric but have different (2FF)'s.

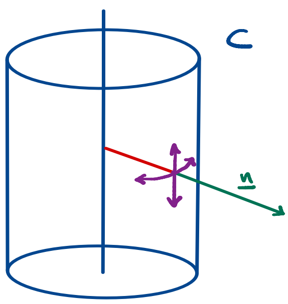


FIGURE 4.7. Cylinder

Definition 4.21. Let $X \subset \mathbb{R}^3$ be a smooth surface and $\underline{r}\colon V \to X$ be a local parametrization for $V \subseteq \mathbb{R}^2$ open. Then, we have the (1FF)

$$g = Edu^2 + 2Fdudv + Gdv^2\,,$$

and second fundamental form (2FF)

$$II = Ldu^2 + 2Mdudv + Ndv^2\,.$$

Fix $x = \underline{r}(u,v)$ in X. The principal curvatures κ_1, κ_2 of X at x are the solutions of

$$\det\left(\lambda\begin{pmatrix} E & F \\ F & G\end{pmatrix} - \begin{pmatrix} L & M \\ M & N\end{pmatrix}\right) = 0\,.$$

(unique if we require $\kappa_1 \leq \kappa_2$).

The Gaussian curvature is

$$\kappa = \kappa_1\kappa_2 = \frac{\det\begin{pmatrix} L & M \\ M & N\end{pmatrix}}{\det\begin{pmatrix} E & F \\ F & G\end{pmatrix}} = \frac{LN - M^2}{EG - F^2}\,.$$

The mean curvature is

$$H = \kappa_1 + \kappa_2 = \mathrm{Trace}\left(\begin{pmatrix} L & M \\ M & N\end{pmatrix}\begin{pmatrix} E & F \\ F & G\end{pmatrix}^{-1}\right)$$

Warning: Conventions differ. Some authors write $H = \frac{1}{2}(\kappa_1 + \kappa_2)$.

Note that both κ and H are smooth functions $X \to \mathbb{R}$, independent of the choice of local parametrization $\underline{r}$.

We call X a minimal surface if $H = 0$ (this is a partial differential equation (PDE) on X). A surface is minimal if it has stationary area with fixed boundary. That is, if you deform the surface in its interior a little bit, you do not change the area to first order (in general, the first order change is $\int_X H\underline{n}\cdot$ (normal deformation)).

A bubble spanning a loop of wire is a minimal surface, as surface tension minimizes the area.

A coordinate independent point of view is to regard g and II as quadratic forms on T_xX for $x \in X$, i.e. $g(u_1, u_2)$, $II(u_1, u_2)$ for $u_1, u_2 \in T_xX$.

The shape operator or the Weingarten map is the unique linear map $\mathcal{S}\colon T_xX \to T_xX$ such that

$$II(\underline{v}_1, \underline{v}_2) = g(\mathcal{S}(\underline{v}_1), \underline{v}_2)\,,$$

for all $\underline{v}_1, \underline{v}_2$ in T_xX.

In coordinates, it has matrix

$$\begin{pmatrix} E & F \\ F & G\end{pmatrix}^{\frac{-1}{2}}\begin{pmatrix} L & M \\ M & N\end{pmatrix}\begin{pmatrix} E & F \\ F & G\end{pmatrix}^{\frac{-1}{2}}\,.$$

This is a symmetric matrix. Then, κ_1, κ_2 are eigenvectors of $\mathcal{S}$, and $\kappa = \det(S)$, and $H = \mathrm{Trace}(\mathcal{S})$.

The principal directions $\underline{v}_1, \underline{v}_2$ in T_xX are the unit eigenvectors associated to the eigenvalues κ_1, κ_2 of $\mathcal{S}$. Then, $\underline{v}_1 \perp \underline{v}_2$.

Example 4.22. Let $X \subset \mathbb{R}^3$ be a smooth surface, and $x \in X$. Then, $\underline{v}_1, \underline{v}_2, \underline{n}$ at x are an orthonormal basis of $\mathbb{R}^3$. Apply a rotation/reflection and translation such that

$$x = (0,0,0),\ \ \underline{v}_1 = (1,0,0),\ \ \underline{v}_2 = (0,1,0),\ \ \underline{n} = (0,0,1)\,.$$

Then locally near x,mone can write X as the graph

$$\{(x, y, f(x, y)) \mid f(x, y) \in \mathbb{R}^2\}$$

of a smooth function $f(x, y)$ with $f(0,0) = f_x(0,0) = f_y(0,0)$.

We have $\underline{r}_x = (1, 0, f_x)$, $\underline{r}_y = (1, 0, f_y)$, and

$$\underline{n} = \frac{(-f_x, -f_y, 1)}{(1 + f_x^2 + f_y^2)^{\frac{1}{2}}}\,.$$

At x we have $g = dx^2 + dy^2$, and

$$II = f_{xx}(0,0)dx^2 + 2f_{xy}(0,0)dxdy + f_{yy}(0,0)dy^2\,.$$

So,

$$\mathcal{S} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix}$$

since $\mathcal{S}$ has eigenvalues κ_1, κ_2 with eigenvectors $(1,0,0), (0,1,0)$. Thus,

$$f(x, y) = \frac{1}{2}\kappa_1 x^2 + \frac{1}{2}\kappa_2 y^2 + O((x, y)^3)\,.$$

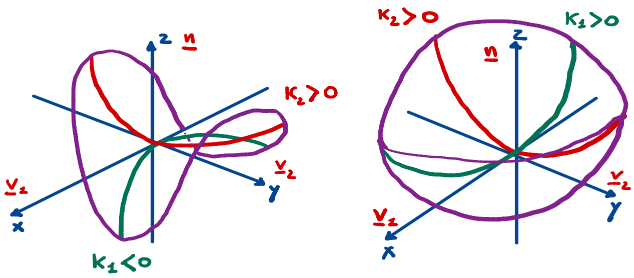


FIGURE 4.8. Left hand figure: $\kappa = \kappa_1\kappa_2 < 0$, negative Gaussian curvature, like a saddle point or Pringle. Necessary for $H = 0$, minimal. Right hand figure: $\kappa = \kappa_1\kappa_2 > 0$, positive Gaussian curvature, like a paraboloid, sphere, or ellipsoid (rugby ball) locally – locally convex, not minimal.

Example 4.23. The catenoid is the surface of revolution of the graph

$$y = \cosh v \qquad \text{(This is called the catenary)}\,.$$

Parametrize it as

$$\begin{aligned} \underline{r}(u, v) &= (\cos u \cosh v, \sin u \cosh v, v) \\ \underline{r}_u &= (-\sin u \cosh v, \cos u \cosh v, 0) \\ \underline{r}_v &= (\cos u \sinh v, \sin u \sinh v, 1)\,. \end{aligned}$$

So, we have

$$\begin{aligned}
\underline{n} &= (\cos u \ \mathrm{sech} v, \sin u \ \mathrm{sech} v, -\mathrm{tanh} v) \\
g &= \cosh^2 v(du^2 + dv^2) \\
II &= -du^2 + dv^2 \\
\kappa_1 &= \frac{-1}{\cosh^2 v}, \ v_1 = (-\sin u, \cos u, 0) \\
\kappa_2 &= \frac{1}{\cosh^2 v}, \ v_2 = (\cos u \mathrm{tanh} v, \sin u \mathrm{tanh} v, \mathrm{sech} v) \\
\kappa &= \frac{-1}{\cosh^4 v}, \ H = 0
\end{aligned}$$

So, the catenoid is a minimal surface.

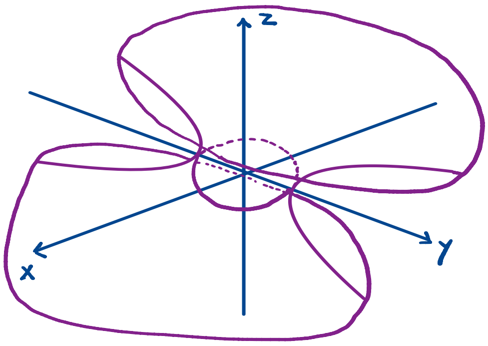


FIGURE 4.9. The catenoid

4.8. **Tangential derivatives and the Theorema Egregium.**

Definition 4.24. Let $X \subset \mathbb{R}^3$ be a smooth surface, and $\underline{r}\colon V \to X$ a local parametrization, $(u, v) \in V \subseteq \mathbb{R}^2$ open. Define smooth functions $\Gamma^c_{ab}\colon V \to \mathbb{R}$ for $a, b, c \in \{u, v\}$, called Christoffel symbols, by

$$\begin{aligned}
\underline{r}_{uu} &= L\underline{n} + \Gamma^u_{uu}\underline{r}_u + \Gamma^v_{uu}\underline{r}_v \\
\underline{r}_{uv} &= M\underline{n} + \Gamma^u_{uv}\underline{r}_u + \Gamma^v_{uv}\underline{r}_v \\
\underline{r}_{vu} &= M\underline{n} + \Gamma^u_{vu}\underline{r}_u + \Gamma^v_{vu}\underline{r}_v \\
\underline{r}_{vv} &= N\underline{n} + \Gamma^u_{vv}\underline{r}_u + \Gamma^v_{vv}\underline{r}_v
\end{aligned} \tag{4.3}$$

Note that as $\underline{r}_{uv} = \underline{r}_{vu}$, we have $\Gamma^c_{ab} = \Gamma^c_{ba}$.

Proposition 4.25. *The Christoffel symbols Γ^c_{ab} depend only on the 1FF $g = Edu^2 + 2Fdudv + Gdv^2$. Hence, the Γ^c_{ab} are also defined for an abstract surface with Riemannian metric (X, g).*

Proof. As $E = \underline{r}_u\underline{r}_u$, $F = \underline{r}_u\underline{r}_v$, we have

$$\begin{aligned}
E_u &= 2\underline{r}_u\underline{r}_{uu} \\
E_v &= 2\underline{r}_u\underline{r}_{uv} \\
F_u &= \underline{r}_{uu}\underline{r}_v + \underline{r}_u\underline{r}_{uv}
\end{aligned}$$

So, we have

$$E\ \Gamma^u_{uu} + F\ \Gamma^v_{uu} = \underline{r}_u \underline{r}_{uu} = \frac{1}{2}E_u$$

$$F\ \Gamma^u_{uu} + G\ \Gamma^v_{uu} = \underline{r}_v \underline{r}_{uu} = F_u - \frac{1}{2}E_v$$

Hence,

$$\begin{pmatrix} \Gamma^u_{uu} \\ \Gamma^v_{uu} \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{2}E_u \\ F_u - \frac{1}{2}E_v \end{pmatrix}$$

The proofs for the other Γ^a_{bc} are similar. □

Definition 4.26. Let $X \subset \mathbb{R}^3$ be a smooth surface, and $\underline{r}\colon V \to X$ a local parametrization. A vector field on X is a smooth map

$$\underline{a}\colon X \longrightarrow \mathbb{R}^3$$
$$x \longmapsto T_x X \subset \mathbb{R}^3\,,$$

for each $x \in X$. On V, by abuse of notation, we often write $\underline{a}$ for $\underline{a} \circ \underline{r}\colon V \to \mathbb{R}^3$. We have

$$\underline{a} = \underline{a} \circ \underline{r} = e\underline{r}_u + f\underline{r}_v\,,$$

where $e, f\colon V \to \mathbb{R}$ are smooth. Also, write this as

$$\underline{a} = e\frac{\partial}{\partial u} + f\frac{\partial}{\partial v}\,.$$

Then vector fields also make sense on abstract smooth surfaces X in local coordinates (u, v). The tnagential derivatives of $\underline{a}$ are

$$\nabla_u \underline{a} = \underline{a}_u - (\underline{n} \cdot \underline{a}_u)\underline{n} = \underline{a}_u + (\underline{n}_u \cdot \underline{a})\underline{n}$$
$$\nabla_v \underline{a} = \underline{a}_v - (\underline{n} \cdot \underline{a}_v)\underline{n} = \underline{a}_v + (\underline{n}_v \cdot \underline{a})\underline{n}$$

That is, they are the orthogonal projections of $\underline{a}_u, \underline{a}_v$ to the tangent spaces $T_x X$. One can also write them as $\nabla_{\frac{\partial}{\partial u}} \underline{a}, \nabla_{\frac{\partial}{\partial v}} \underline{a}$, i.e. we differentiate the vector field $\underline{a} = e\frac{\partial}{\partial u} + f\frac{\partial}{\partial v}$ in the directions of vector fields $\frac{\partial}{\partial u}, \frac{\partial}{\partial v}$. From (4.3), we see that

$$\nabla_u \underline{a} = e_u \underline{r}_u + f_u \underline{r}_v + e\Gamma^u_{uu}\underline{r}_u + e\Gamma^v_{uu}\underline{r}_v + f\Gamma^u_{vu}\underline{r}_u + f\Gamma^v_{vu}\underline{r}_v$$
$$\nabla_v \underline{a} = e_v \underline{r}_u + f_v \underline{r}_v + e\Gamma^u_{uv}\underline{r}_u + e\Gamma^v_{uu}\underline{r}_v + f\Gamma^u_{vv}\underline{r}_u + f\Gamma^v_{uv}\underline{r}_v\,.$$

That is, for $\underline{a} = e\frac{\partial}{\partial u} + f\frac{\partial}{\partial v}$, we have

$$\nabla_{\frac{\partial}{\partial u}} \underline{a} = e_u\frac{\partial}{\partial u} + f_u\frac{\partial}{\partial v} + e\Gamma^u_{uu}\frac{\partial}{\partial u} + e\Gamma^v_{uu}\frac{\partial}{\partial v} + f\Gamma^u_{vu}\frac{\partial}{\partial u} + f\Gamma^v_{vu}\frac{\partial}{\partial v}$$
$$\nabla_{\frac{\partial}{\partial v}} \underline{a} = e_v\frac{\partial}{\partial u} + f_v\frac{\partial}{\partial v} + e\Gamma^u_{uv}\frac{\partial}{\partial u} + e\Gamma^v_{uv}\frac{\partial}{\partial v} + f\Gamma^u_{vv}\frac{\partial}{\partial u} + f\Gamma^v_{uv}\frac{\partial}{\partial v}\,.$$

By Proposition 4.25 the Γ^a_{bc} depend only on the 1FF, and are also defined on general (X, g). Hence, ∇_u, ∇_v also make sense on a surface X with Riemannian metric g.

This defines a structure ∇ ("nabla") called the Levi-Civita connection on (X, g). It differentiates a vector field on X in the direction of another vector field, where vector fields are of the form $e\frac{\partial}{\partial u} + f\frac{\partial}{\partial v}$ in coordinates (u, v) on X.

Aside: Let X be a surface and $\underline{a}, \underline{b}$ be vector fields on X. We would like to define $\nabla_{\underline{b}}\underline{a}$, the derivative of $\underline{a}$ in the direction of $\underline{b}$. Heuristically, at $x \in X$, we would like

$$\nabla_{\underline{b}}\underline{a} = \lim_{\epsilon \to 0} \frac{\underline{a}|_{x+\epsilon\underline{b}} - \underline{a}|_x}{\epsilon}\,.$$

However, $\underline{a}|_{x+\epsilon \underline{b}} \in T_{x+\epsilon \underline{b}}X$ and $\underline{a}|_x \in T_x X$ lie in <u>different vector spaces</u>, so we cannot subtract them. Heuristically, the job of a connection is to identify nearby tangent spaces $T_x X = T_y X$ for x, y close in X, so we can make sense of this.

Proposition 4.27. *For any vector field $\underline{a}$ on $X \subset \mathbb{R}^3$ we have*

$$\nabla_v \nabla_u \underline{a} - \nabla_u \nabla_v \underline{a} = \frac{LN - M^2}{\sqrt{EG - F^2}}\ \underline{n} \wedge \underline{a} = K\sqrt{EG - F^2}\ \underline{n} \wedge \underline{a} \tag{4.4}$$

Proof. We have

$$\nabla_v \nabla_u \underline{a} = \underline{a}_{vu} - (\underline{n} \cdot \underline{a}_{vu})\underline{n} + (\underline{n}_u \cdot \underline{a})\underline{n}_v\,.$$

So,

$$\nabla_v \nabla_u \underline{a} - \nabla_u \nabla_v \underline{a} = (\underline{n}_u \cdot \underline{a})\underline{n}_v - (\underline{n}_v \cdot \underline{a})\underline{n}_u = (\underline{n}_u \wedge \underline{n}_v) \wedge \underline{a}\,.$$

Write $\underline{n}_u \wedge \underline{n}_v = \lambda \underline{n}$. Then,

$$\begin{aligned} \lambda \underline{n} \cdot (\underline{r}_u \wedge \underline{r}_v) &= (\underline{n}_u \wedge \underline{n}_v) \cdot (\underline{r}_u \wedge \underline{r}_v) \\ &= (\underline{n}_u \cdot \underline{r}_u)(\underline{n}_v \cdot \underline{r}_v) - (\underline{n}_u \cdot \underline{r}_v)(\underline{n}_v \cdot \underline{r}_u) \\ &= LN - M^2\,. \end{aligned}$$

We also have

$$\underline{n} \cdot (\underline{r}_u \wedge \underline{r}_v) = \sqrt{EG - F^2}\,.$$

Therefore,

$$\lambda = \frac{LN - M^2}{\sqrt{EG - F^2}}$$

<u>Note</u>: The right hand side of (4.4) involves no partial derivatives of $\underline{a}$. □

Corollary 4.28. *<u>Gauss' Theorema Egregium:</u> The Gaussian curvature κ of a surface X can be written solely in terms of teh 1FF $Edu^2 + 2Fdudv + Gdv^2$ and its first and second derivatives (one could give an explicit formula, but it is complicated).*

Proof. Proposition 4.27 implies that $\underline{a} \mapsto \nabla_v \nabla_u \underline{a} - \nabla_u \nabla_v \underline{a}$ is given by $\kappa\sqrt{EG - F^2}$ (rotation by 90°). But ∇_u, ∇_v depend only on 1FF by Proposition 4.25. Also, rotation by 90° also depends on the 1FF (actually, it also depends on orientation, but orientation is used to determine the order of u, v in $\nabla_v \nabla_u - \nabla_u \nabla_v$, so overall formula is orientation independent). So, $\kappa\sqrt{EG - F^2}$ only depends on the 1FF, and so does κ. Each ∇_u, ∇_v depends on 1FF and first derivatives, but taking second derivatives $\nabla_v \nabla_u, \nabla_u \nabla_v$ induces an extra derivative of the 1FF. □

This implies that the Gaussian curvature κ is also defined for a surface X with a Riemannian metric g.

4.9. **Geodesic curvature and geodesics.**

Definition 4.29. Let $X \subset \mathbb{R}^3$ be an (oriented) smooth surface, and $\gamma\colon [a, b] \to X$ be a smooth curve parametrized by arc-length s, i.e. $|\frac{d}{ds}(\gamma(s))| = 1$. Set $\underline{t} = \gamma' = \frac{d\gamma}{ds}$ and $\underline{t}' = \frac{d^2\gamma}{ds^2}$. Then, $\underline{t}$ is the unit tangent vector to γ at $\gamma(s)$. The <u>geodesic curvature</u> $K_g\colon [a, b] \to \mathbb{R}$ of γ is

$$K_g = \underline{t}' \cdot (\underline{n} \wedge \underline{t})\,,$$

where $\underline{n}$ is the unit normal to X at $\gamma(s)$ (we need an orientation to choose the sign of $\underline{n}$). We call γ a geodesic if $K_g = 0$. This is an O.D.E. on γ.

The geodesic equation is equivalent to

$$\underline{t}' = \frac{d^2\gamma}{ds^2} = \lambda \underline{n}\,,$$

for some $\lambda\colon [a,b] \to \mathbb{R}$. That is, the acceleration of $\gamma(s)$, as a moving point, is normal to X.

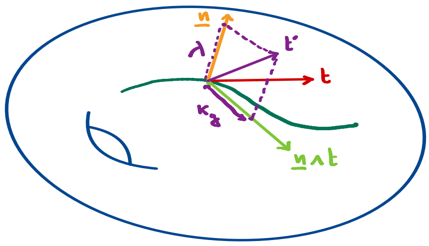


FIGURE 4.10. Geodesic curvature κ_g

Note that $\underline{t}, \underline{n}, \underline{n} \wedge \underline{t}$ form an orthonormal basis. Also, $\underline{t} \cdot \underline{t}' = 0$ as $|\underline{t}|^2 = 1$. So,

$$\underline{t}' = \lambda \underline{n} + K_g \; \underline{n} \wedge \underline{t}\,,$$

for some $\lambda, K_g \in \mathbb{R}$. So, $K_g = 0$ iff $\underline{t}' = \lambda \underline{n}$.

Note that the geodesic equation $K_g = 0$ appears in mechanics: the motion of a small ball rolling on a frictionless surface without gravity, so the only force is normal to the surface.

Geodesics are locally length-minimizing. That is, if $a \le t_1 \le t_2 \le b$ with $|t_1 - t_2|$ small, then $\gamma([t_1, t_2])$ is the shortest path from $\gamma(t_1)$ to $\gamma(t_2)$ in X.

Let $\underline{r}\colon V \to X$ be a local parametrization, and write $\gamma(s) = \underline{r}(u(s), v(s))$. Then, $\underline{t} = u'\underline{r}_u + v'\underline{r}_v$, and $\underline{t}' = u''\underline{r}_u + v''\underline{r}_v + ((u')^2\underline{r}_{uu} + 2u'v'\underline{r}_{uv} + (v')^2\underline{r}_{vv}) = u''\underline{r}_u + v''\underline{r}_v + ((u')^2 L + 2u'v'M + (v')^2 N)\underline{n} + (u')^2(\Gamma^u_{uu}\underline{r}_u + \Gamma^v_{uu}\underline{r}_v) + 2(u'v')(\Gamma^u_{uv}\underline{r}_u + \Gamma^v_{uv}\underline{r}_v) + (v')^2(\Gamma^u_{vv}\underline{r}_u + \Gamma^v_{vv}\underline{r}_v)$, where the last equality follows by (4.3). Write

$$\underline{n} \wedge \underline{t} = f\underline{r}_u + g\underline{r}_v\,.$$

Then, f, g depend only on 1FF,u',v', and orientation, as $\underline{n} \wedge \underline{t}$ is rotation of $\underline{t}$ by $90°$, and $K_g = (Ef + Fg)(u'' + (u')^2\Gamma^u_{uu} + 2u'v'\Gamma^u_{uv} + (v')^2\Gamma^u_{vv} + (Ff + Gg)(v'' + (u')^2\Gamma^v_{uu} + 2u'v'\Gamma^v_{uv} + (v')^2\Gamma^v_{vv}))$

Hence, by Proposition 4.25, the geodesic curvature κ_g of γ depends only on γ, 1FF and the orientation of X (orientation determines the sign). Thus, κ_g is also well defined for curves in surfaces (X, g) with Riemannian metrics.

Proposition 4.30. *Let $X \subset \mathbb{R}^3$ be a smooth surface, and $\underline{r}\colon V \to X$ a local parametrization. A curve $\gamma(s) = \underline{r}(u(s), v(s))$ parametrized by arc-length s is a geodesic iff*

$$\frac{d}{ds}(Eu' + Fv') = \frac{1}{2}(E_u(u')^2 + 2F_u u'v' + G_u((v')^2)$$
$$\frac{d}{ds}(Fu' + Gv') = \frac{1}{2}(E_v(u')^2 + 2F_v u'v' + G_v((v')^2)\,.$$

Proof. We have $\underline{t} = \underline{r}_u u' + \underline{r}_v v'$, and γ is a geodesic iff $\underline{t}'$ is normal, i.e.

$$\underline{t}' \cdot \underline{r}_u = \underline{t}' \cdot \underline{r}_v = 0\,.$$

But,

$$\underline{t}' \cdot \underline{r}_u = (\underline{t} \cdot \underline{r}_u)' - \underline{t} \cdot \underline{r}_u'\,,$$

So, the first equation is $(\underline{t} \cdot \underline{r}_u)' = \underline{t} \cdot \underline{r}_u'$, that is,

$$\frac{d}{ds}((u'\underline{r}_u + v'\underline{r}_v) \cdot \underline{r}_u) = (u'\underline{r}_u + v'\underline{r}_v) \cdot (u'\underline{r}_{uu} + v'\underline{r}_{uv}) \,.$$

That is,

$$\frac{d}{ds}(Eu' + Fv') = \frac{1}{2}(E_u(u')^2 + 2F_u u'v' + G_u(v')^2 \,,$$

as $E_u = (\underline{r}_u \cdot \underline{r}_u)_u = 2\underline{r}_u\underline{r}_{uu}$, etc. The second equation is similar. □

These geodesic equations also make sense on a surface with Riemannian metric (X, g). One can use them and results on O.D.E.'s to show that given a point $x \in X$ and a direction at x, there is a unique geodesic through x in this direction.

4.10. **The Gauss–Bonnet Theorem.**

Theorem 4.31. *Local Gauss–Bonnet: Let $X \subset \mathbb{R}^3$ be an oriented smooth surface, and $\underline{r}\colon V \to X$ a local parametrization. Let $\gamma\colon S^1 \to \underline{r}(V) \subseteq X$ be a smooth curve parametrized by arc-length s, which is the boundary of a compact disc shaped region R in $\underline{r}(V) \subseteq X$, and γ goes anti-clockwise around R. Then,*

$$\int_\gamma K_g ds = 2\pi - \int_R K dA \,,$$

where κ_g is the geodesic curvature of γ, κ is the Gaussian curvature of X, and $\int \ldots ds$ and $\int \ldots dA$ are integration w.r.t. arc length and area respectively.

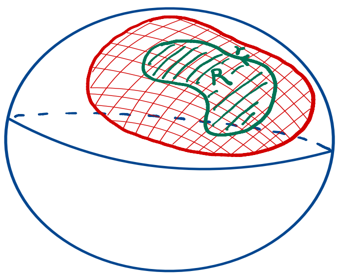


FIGURE 4.11. Illustration for the local Gauss-Bonnet theorem (Theorem 4.5)

Proof. (not examinable) Write $R = \underline{r}(D)$ for $D \subset V$ a closed disc-shaped region with smooth boundary ∂D, so $\gamma(S^1) = \underline{r}(\partial D)$. Parametrize ∂D by arc-length s in X, writing $(u(s), v(s)) \in \partial D$. Let $P, Q\colon V \to \mathbb{R}$ be smooth. Then, Green's formula gives

$$\int_{\partial D} (P(u(s), v(s))u' + Q(u(s), v(s))v')ds = \int_D (Q_u - P_v) du dv \,. \tag{4.5}$$

Choose a unit length tangent vector $\underline{e}$ on $\underline{r}(V)$, for instance $\underline{e} = \frac{\underline{r}_u}{\sqrt{E}}$. Then, $\nabla_u \underline{e}, \nabla_v \underline{e}$ are tangent vector fields orthogonal to $\underline{e}$, so there are smooth functions $P, Q\colon V \to \mathbb{R}$ with

$$\nabla_u \underline{e} = P \; \underline{n} \wedge \underline{e}$$
$$\nabla_v \underline{e} = Q \; \underline{n} \wedge \underline{e} \,.$$

Let $\underline{t}$ be the unit tangent to $\underline{\gamma}$, and write

$$\underline{t} = \cos\theta \underline{e} + \sin\theta \; \underline{n} \wedge \underline{e} \,,$$

for smooth

$$\theta\colon \partial D \to \frac{\mathbb{R}}{2\pi\mathbb{Z}}\,.$$

So, θ is the angle between $\underline{t}$ and $\underline{e}$. Then, the geodesic curvature is $K_g = \underline{t}' \cdot (\underline{n} \wedge \underline{t}) = \Big(\theta'(-\sin\theta\ \underline{e} + \cos\theta\ \underline{n}\wedge\underline{e})) + \cos\theta(u'\nabla_u\underline{e} + \underline{v}\nabla_v\underline{e}) + \sin\theta\ \underline{n}\wedge(u'\nabla_u\underline{e} + v'\nabla_v\underline{e}\Big) \cdot \Big(\underline{n}\wedge(\cos\theta\ \underline{e} + \sin\theta\ \underline{n}\wedge\underline{e})\Big) = \theta' + Pu' + Qv'$. Hence, the left hand side of (4.5) is

$$\int_{\partial D} (K_g - \theta')ds = \int_\gamma K_g ds - 2\pi\,,$$

since θ increases from 0 to 2π around $S^1 = \partial D$. For the right hand side of (4.5), we have

$$\nabla_v\nabla_u\underline{e} = \nabla_v(P\ \underline{n}\wedge\underline{e}) = P_v\ \underline{n}\wedge\underline{e} + P\ \underline{n}\wedge\nabla_v\underline{e} = P_v\ \underline{n}\wedge\underline{e} + PQ\ \underline{n}\wedge(\underline{n}\wedge\underline{e})\,,$$

So,

$$\nabla_v\nabla_u\underline{e} - \nabla_u\nabla_v\underline{e} = (P_v - Q_u)\ \underline{n}\wedge\underline{e} = K\sqrt{EG-F^2}\ \underline{n}\wedge\underline{e}\,,$$

by Proposition 4.27. Thus the right hand side of (4.5) is

$$\int_D K\sqrt{EG-F^2}dudv = \int_D KdA\,.$$

Hence, the Theorem follows. □

Remarks 4.32. a) Theorem 4.31 also holds for a smooth disc in a surface X with Riemannian metric g, not embedded in $\mathbb{R}^3$. We have shown κ_g and κ are defined then, as they depend on the 1FF.

b) We can extend Theorem 4.31 to allow γ <u>piecewise-smooth</u>. For instance, as illustrated in Figure 4.12 we can have a curvilinear polygon with n vertices, with internal angles $\alpha_1, \ldots, \alpha_n$. Then, we get

$$\begin{aligned}\int_\gamma K_g ds &= 2\pi - (\pi-\alpha_1) - \ldots - (\pi - \alpha_n) - \int_R KdA \\ &= (2-n)\pi + \alpha_1 + \ldots + \alpha_n - \int_R KdA\,,\end{aligned} \tag{4.6}$$

since smoothing off a corner adds $\pi - \alpha$ to $\int_\gamma K_g dA$ in the limit.

<u>Note</u>: in nice cases, if the sides are geodesics, we have $K_g = 0$, and

$$(n-2)\pi + \int_R KdA = \alpha_1 + \ldots + \alpha_n\,.$$

If $K = 0$, we get the usual formula for angles of a polygon in $\mathbb{R}^2$.

Theorem 4.33. <u>*Gauss–Bonnet Theorem*</u> *Let (X, g) be a compact, smooth surface with a Riemannian metric g and Gaussian curvature κ., Then,*

$$\int_X KdA = 2\pi\chi(X)\,.$$

This says that the total curvature $\int_X KdA$ is a topological invariant – it is independent of the metric g.

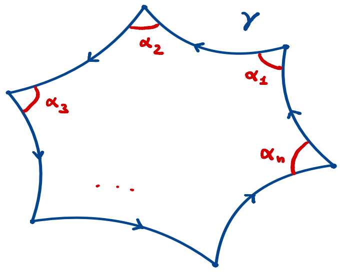


FIGURE 4.12. Curvilinear polygon

Proof. Choose a triangulation of X into smooth triangles $\Delta_1, \ldots, \Delta_{2n}$, each lying in a coordinate neighbourhood of X. There are in total V vertices $E = 3n$ edges, and $F = 2n$ faces, so

$$\chi(X) = V - E + F = V - n\,.$$

Write α_{ij} for the internal angles of Δ_i, $j \in \{1, 2, 3\}$ as in Figure **??**. Then, (4.6) gives

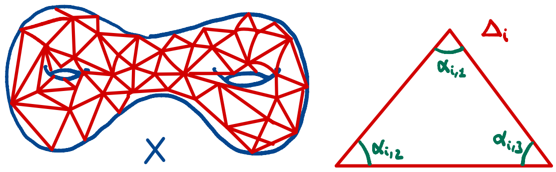


FIGURE 4.13. Illustration for the proof of Theorem 4.6

$$\int_{\partial\Delta_i} K_g ds = -\pi + \alpha_{i1} + \alpha_{i2} + \alpha_{i3} - \int_{\Delta_i} K dA\,.$$

Adding over $i = 1\ldots, 2n$ gives

$$\sum_{i=1}^{2n} \int_{\partial\Delta_i} K_g ds = -2\pi n + \sum_{i=1}^{2n}\sum_{j=1}^{3} \alpha_{ij} - \int_X K dA\,.$$

Each edge is the boundary of two triangles Δ_{i1}, Δ_{i2} and κ_g has opposite signs as the edges are oriented in opposite directions, so the integral of κ_g along pairs of edges cancel, giving

$$\sum_{i=1}^{2n} \int_{\partial\Delta_i} K_g ds = 0\,.$$

Also, $\sum_{i,j} \alpha_{ij} = 2\pi V$, as at each vertex the internal angles sum to 2π. Hence,

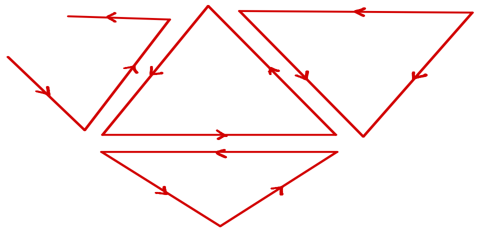


FIGURE 4.14. Illustration for the proof of Theorem 4.6

$$\begin{aligned}
0 &= \sum_{i=1}^{2n} \int_{\partial \Delta_i} K_g ds \\
&= -2\pi n + 2\pi V - \int_X K dA \\
&= 2\pi (V-n) - \int_X K dA \\
&= 2\pi \chi(X) - \int_X K dA \,.
\end{aligned}$$

□

4.11. **Critical points and the Euler characteristic.**

Definition 4.34. Let X be a smooth surface and $f \colon X \to \mathbb{R}$ a smooth function. A point $x \in X$ is a critical point if $f_u = f_v = 0$ in local coordinates (u, v).

The Hessian of f at X is

$$\operatorname{Hess}(f) = \begin{pmatrix} f_{uu} & f_{uv} \\ f_{vu} & f_{vv} \end{pmatrix} .$$

The critical point is non-degenerate if $\det \operatorname{Hess}(f) \neq 0$. Non-degenerate critical points are isolated. So, there are only finitely many if X is compact. They are divided into local minima, saddle points, and local maxima, if $\operatorname{Hess}(f)$ has signature $(+,+), (+,-), (-,-)$, respectively.

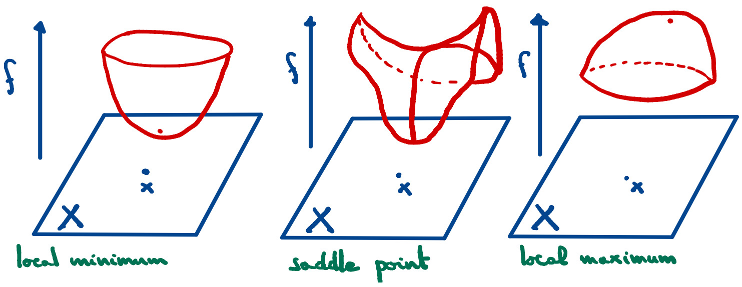


FIGURE 4.15. Non-degenerate critical points

Theorem 4.35. *Let X be a compact smooth surface, and $f \colon X \to \mathbb{R}$ a smooth function woth only non-degenerate critical points. Then,*

$$\chi(X) = (\#\textit{local min}) - (\#\textit{saddle points}) + (\#\textit{local max}) \,.$$

Proof. We adapt the proof of the Gauss–Bonnet Theorem in **??**. Choose a Riemannian metric g on X. Define a vector $\underline{a} = \nabla f$ on X with zeroes only at the critical points of f. Write $x_1, \ldots, x_k$ for the critical points. Choose a small disc D_i around x_i, with $i = 1, \ldots, k$. On $X \setminus (D_1 \cup \ldots \sup D_k)$, define $\underline{e} = \frac{\underline{a}}{|\underline{a}|}$, a unit vector field. On each D_i, choose a unit vector field $\underline{e}_i$.

The proof of Theorem 4.31 also shows: if $R \subset X$ is a region with smooth boundary $\gamma = \partial R$, and $\underline{e}$ is a unit vector field on R, then

$$\int_\gamma (K_g - \theta')ds = -\int_R K dA\,,$$

where θ is the angle between $\underline{t} = \frac{d\gamma}{ds}$ and $\underline{e}$. This holds for all closed regions R, not just discs.

Set $\gamma_i = \partial D_i$ and write $\overline{\gamma}_i$ for γ_i with the opposite orientation. We have, for $R = X \setminus (D_1 \cup \ldots \cup D_k)$,

$$\sum_{i=1}^{k} \int_{\overline{\gamma}_i} (K_g - \theta')ds = -\int_{X\setminus(D_1\cup\ldots\cup D_k)} K dA\,,$$

and for $R = D_i$,

$$\int_{\overline{\gamma}_i} (K_g - \theta_i')ds = -\int_{D_i} K dA\,,$$

where θ, θ_i are the angles between $\underline{t}$ and $\underline{e}, \underline{e}_i$, respectively.

FIGURE 4.16. Illustration for the proof of Theorem 4.7

As $\overline{\gamma}_i$ has opposite orientation to γ_i, we have

$$\int_{\overline{\gamma}_i} (K_g - \theta')ds = -\int_{\gamma_i} (K_g - \theta')ds\,.$$

So, adding gives

$$\sum_{i=1}^{k} \int_{\gamma_i} (\theta' - \theta_i')ds = -\int_X K dA = -2\pi\chi(X)\,,$$

by Gauss–Bonnet.

Now, $\theta - \theta_i$ increases by $-2\pi, 2\pi, -2\pi$ for a minimum, saddle, and maximum respectively. Hence,

$$2\pi((\#\text{local min}) - (\#\text{saddle points}) + (\#\text{local max})) = -2\pi\chi(X)\,.$$

□

Examples 4.36. The "Hairy Ball Theorem" says that if you have a 2-sphere S^2 with "hair" all over it (e.g. a hamster) you cannot comb the hair so it lies flat everywhere. That is, there are no vector fields $\underline{a}$ on S^2 with no zeroes, as then a variant of Theorem 4.35 would give

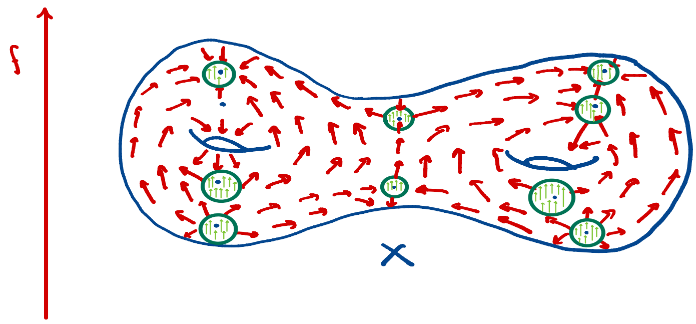


FIGURE 4.17. Blue points: $x_1, \ldots, x_k$, green discs: D_i, green arrows: $\underline{e}_i$, red arrows: $\underline{e}$

$\chi(S^2) = 0$, contradicting $\chi(S^2) = 2$. Thinking of $\underline{a}$ as the wind velocity on the surface of the Earth, there must be some point with no wind.

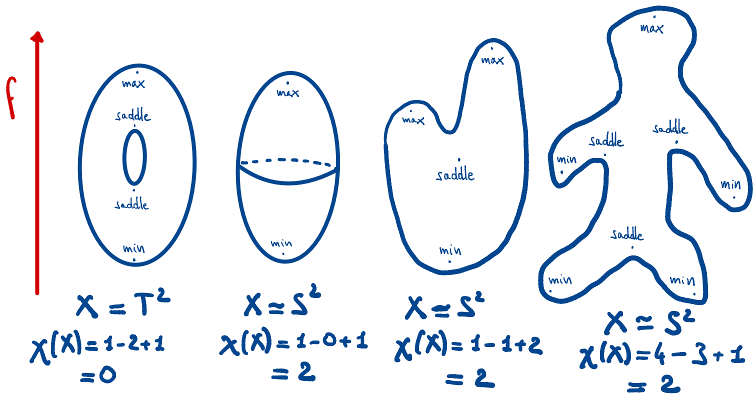


FIGURE 4.18. Examples illustrating Theorem 4.7

5. THE HYPERBOLIC PLANE

Example 5.1. Let

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

Define coordinates (u, v) on $S^2 \setminus (R, 0, 0)$ such that $(0, 0, R)$, (x, y, z) and $(u, v, 0)$ are collinear. This gives

$$\underline{r}(u, v) = \left(\frac{2R^2u}{R^2 + r^2}, \frac{2R^2v}{R^2 + r^2}, \frac{Rr^2 - R^3}{R^2 + r^2}\right),$$

where $r^2 = u^2 + v^2$. The first fundamental form is

$$g = \frac{4R^2}{(R^2 + r^2)^2}(du^2 + dv^2)\,.$$

The principal curvatures are $\kappa_1 = \kappa_2 = \frac{1}{R}$, so $\kappa = \frac{1}{R^2}$. Observe that g still makes sense if we take R to be imaginary, so $R^2 < 0$. If $R = i = \sqrt{-1}$, this gives

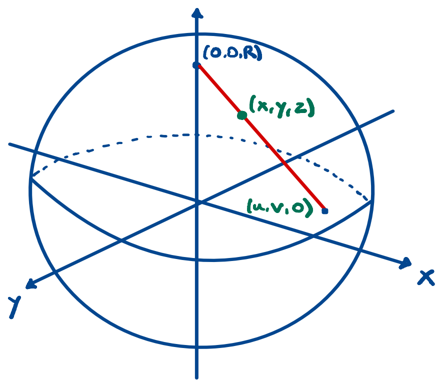


FIGURE 5.1. (u, v) coordinates on the sphere S^2_R

$$g = \frac{4}{(1-r^2)^2}(du^2 + dv^2)\,,$$

with $\kappa = -1$. As this blows up when $r = 1$, we set

$$X = \{(u, v) \in \mathbb{R}^2 \, : \, u^2 + v^2 < 1\}\,,$$

with metric

$$g = \frac{4}{(1-u^2-v^2)^2}(du^2 + dv^2)\,.$$

This is the Poincaré disc model of the hyperbolic plane. It has Gaussian curvature $\kappa = -1$. It is an abstract surface X with Riemannian metric g, which canbe thought of as the sphere with radius i. An alternative model is the upper half plane model

$$\mathbb{H} = \{(x, y) \in \mathbb{R}^2 \, : \, y > 0\}\,,$$

with metric $g = \frac{dx^2+dy^2}{y^2}$. The isometric transformation between them is

$$u + iv = \frac{x + iy - i}{x + iy + i} \text{ or } x + iy = \frac{i(1 + u + iv)}{1 - (u + iv)}\,.$$

We will usually work with the upper half plane model, as it is simpler. It is helpful to write $z = x + iy$ and consider $\mathbb{H} \subset \mathbb{C} \subset \mathbb{C} \cup \{\infty\}$.

Theorem 5.2. *Suppose* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *is real* 2×2 *matrix with* $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} > 0$. *Then,*

$$z \longmapsto \frac{az + b}{cz + d} \text{ and } z \longmapsto \frac{b - a\bar{z}}{d - c\bar{z}}$$

are isometries of $\mathbb{H}$. *All isometries are of this form.*

Proof. See Hitchin's notes §5. □

Note that $\mathbb{R}^* = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \, : \, 0 \neq a \in \mathbb{R} \right\}$ acts trivially, so the isometry group is $\mathbb{Z}_2 \ltimes PGL(2, \mathbb{R})$, where $PGL(2, \mathbb{R}) = \frac{GL(2,\mathbb{R})}{\mathbb{R}^*}$ is 3-dimensional. The unit sphere S^2 also has a 3-dimensional isometry group $O(3) = \mathbb{Z}_2 \ltimes SO(3)$.

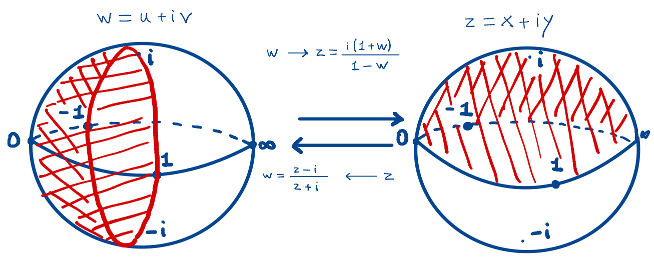


FIGURE 5.2. The unit disc model (on the left) and the upper half plane model (on the right) can both be viewed as hemispheres of $\mathbb{C} \cup \{\infty\}$

5.1. **Geodesics in the hyperbolic plane.** Let $\gamma(s) = (x(s), y(s))$ be a geodesic in $(\mathbb{H}, \frac{dx^2+dy^2}{y^2})$. Then, $\underline{t} = (x', y')$ is a unit vector, so $(x')^2+(y')^2 = y^2$. Also, γ satisfies the geodesic equations in Proposition 4.30. These reduce to

$$\frac{d}{ds}\left(\frac{x'}{y^2}\right) = 0, \ \frac{d}{ds}\left(\frac{y'}{y^2}\right) = \frac{(x')^2y' + (y')^3}{y^3}\,.$$

So, $x' = cy^2$, and $(x')^2 + (y')^2 = y^2$ gives $y' = \sqrt{y^2 - c^2y^4}$.

$$\frac{dy}{dx} = \frac{y'}{x'} = \sqrt{\frac{y^2 - c^2y^4}{c^2y^4}} \quad \Longrightarrow \quad \frac{cydy}{\sqrt{1-c^2y^2}} = dx\,.$$

This integrates to $-c^{-1}\sqrt{1-c^2y^2} = x - a$, i.e. $(x-a)^2 + y^2 = \frac{1}{c^2}$.

Geodesics are semicircles centered on the x-axis, plus vertical half lines (the case $c = 0$, which reduces to $x = a$.)

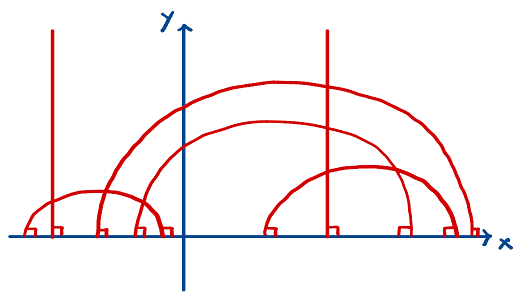


FIGURE 5.3. Hyperbolic geodesics in the half-plane model

In the unit disc model, geodesics are arcs of circles meeting the unit circle at right angles, plus diameters of the unit disc (straight lines through $(0, 0)$).

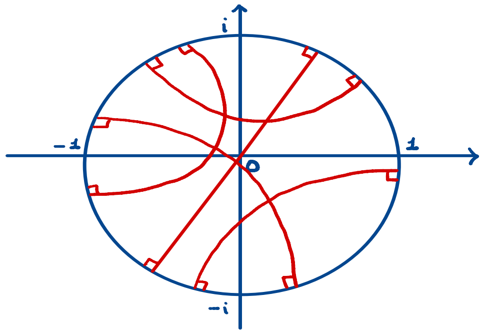


FIGURE 5.4. Hyperbolic geodesics in the disc model

5.2. **Hyperbolic triangles.** Consider a hyperbolic triangle Δ in $\mathbb{H}$, with sides segments of 3 geodesics. Let α, β, γ be the internal angles at the vertices. Note that angles in $(\mathbb{H}, g)$ are the same as angles measured in $\mathbb{R}^2$ in the coordinates (u, v), since $g = (\text{function}) \cdot (du^2 + dv^2)$ and multiplying by a function does not change angles.

Then, piecewise-smooth local Gauss-Bonnet gives

$$0 = (2-3)\pi + \alpha + \beta + \gamma - \int_\Delta \kappa dA\,.$$

Hence,

$$\text{Area}(\Delta) = \pi - (\alpha + \beta + \gamma)\,.$$

Note that $\text{Area}(\Delta) \leq \pi$ even for arbitrarily large Δ. Also, $\alpha + \beta + \gamma < \pi$ (In the Euclidean plane $\alpha + \beta + \gamma = \pi$).

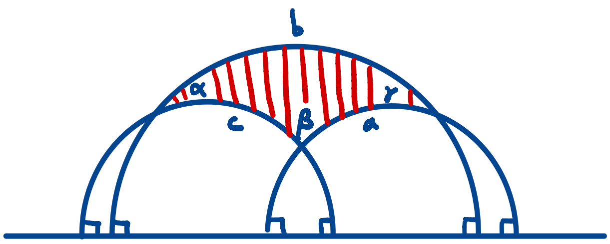


FIGURE 5.5. Hyperbolic triangle.

Let a, b, c be the lengths of the sides opposite α, β, γ. The hyperbolic cosine rule says that

$$\text{cosh}c = \text{cosh}a\ \text{cosh}b - \text{sinh}a\ \text{sinh}b \cos\gamma$$

If a, b, c were small $\text{cosh}a \simeq 1 + a^2$, $\text{sinh}a \simeq a$, $1 + c^2 \simeq (1 + a^2)(1 + b^2) - ab\cos\gamma \iff c^2 \simeq a^2 + b^2 - ab\cos\gamma$. So, one can recover the usual cosine rule in limit $a, b, c \to 0$.

The hyperbolic sine rule says that

$$\frac{\sin\alpha}{\sinh a} = \frac{\sin\beta}{\sinh b} = \frac{\sin\gamma}{\sinh c}\,.$$

If a, b, c are very small then we get $\frac{\sin\alpha}{\alpha} = \frac{\sin\beta}{\beta} = \frac{\sin\gamma}{\gamma}$.

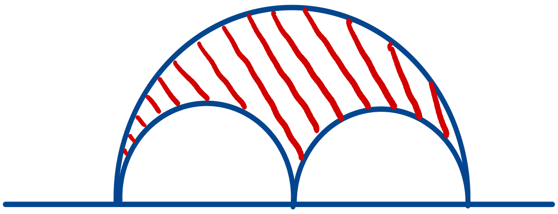


FIGURE 5.6. If $\alpha = \beta = \gamma = 0$ (side lengths infinite), then Area$(\Delta) = \pi$

Lots of geometry in $\mathbb{R}^2$ has nice extensions to $\mathbb{H}$. Tends to involve hyperbolic trigonometric functions $\sinh, \cosh, tanh$.

5.3. **The uniformization theorem.** Let X be a Riemann surface with complex structure J. Then, X is also a smooth surface, so we can consider Riemannian metrics g on X. Now X has tangent spaces $T_xX \cong \mathbb{R}^2$ for $x \in X$. The complex structure on X makes T_xX into a 1-dimensional complex vector space $T_xX \cong \mathbb{C}$. So, multiplication by i gives rotation by $90°$ in T_xX. The metric g gives an inner product on T_xX. There is a natural compatibility condition : the notions of rotation by $90°$ for J, g should be the same. Then, we say that (X, J) and (X, g) have the same conformal structure (conformal = notion of angle). If $z = x + iy$ is a complex coordinate, this happens if $g = E(dx^2 + dy^2)$, i.e. $E = G$ and $F = 0$.

Theorem 5.3. *Uniformization Theorem: Every compact, connected Riemann surface (X, J) has a Riemann metric g compatible with its conformal structure, with constant Gaussian curvature $\kappa = 1, 0, or - 1$.*

From Gauss–Bonnet we see that $\kappa = 1$ if $\chi(X) > 0$, i.e. $g = 0$; $\kappa = 0$ if $\chi(X) = 0$ i.e. $g = 1$, and $\kappa = -1$ if $\chi(X) < 0$ i.e. $g > 1$.

$$\begin{aligned}
\kappa = 1 &\implies X = (S^2, g) \text{ with round metric,}\\
\kappa = 0 &\implies X = \mathbb{R}^2/\Lambda, \text{ with Euclidean metric on } \mathbb{R}^2,\\
\kappa = -1 &\implies X = \mathbb{H}/\Gamma \text{ for } \Gamma \text{ an infinite group of isometries of } \mathbb{H}, \text{ acting freely on } \mathbb{H}.
\end{aligned}$$

So, we can understand Riemann surfaces of genus $g > 1$ using hyperbolic geometry. One can get them by gluing sides of a polygon in $\mathbb{H}$ with geodesic sides: the hyperbolic version of plane models.

Example 5.4. Take a regular octagon "O" in the hyperbolic plane, with geodesic sides, and indernal angels $45°$. Local Gauss–Bonnet gives

$$\int_O \kappa dA = (2-8)\pi + \frac{\pi}{4} + \ldots + \frac{\pi}{4} = -4\pi = 2\pi(-2)\,.$$

Gluing sides as shown in Figure 5.7 gives a genus 2 surface with a hyperbolic metric. ($V = 2, E = 4, F = 1 \implies \chi(X) = -2, g = 2$.)

The hyperbolic plane was historically important in the development of mathematics. Euclid studied geometry in $\mathbb{R}^2$ starting from axioms. His final axiom was called the "prallel postulate":
- Two lines L_1, L_2 are parallel if $L_1 \cap L_2 = \emptyset$. - Given a line L_1, and a point $p \in \mathbb{R}^2$ not on L_1, there is a unique line L_2 through p parallel to L_1.

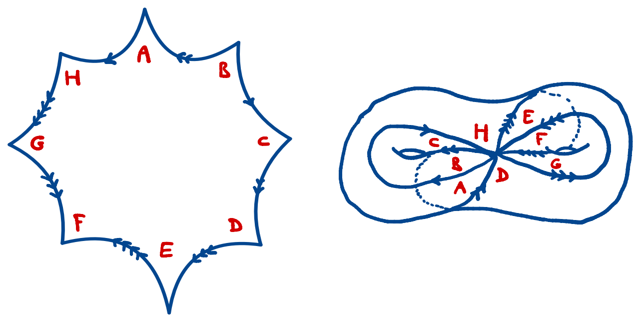


FIGURE 5.7. Hyperbolic genus 2 surface from a regular hyperbolic octogon

Euclid seemed to be embarrassed about the parallel postulate and avoided using it as much as possible. It was a long-standing problem to prove the parallel postulate from the other axioms. The hyperbolic plane satisfies all Euclid's axioms except the parallel postulate: there are many lines L_2 through p not intersecting L_1. This eventually led to a reassessment of what geometry was.

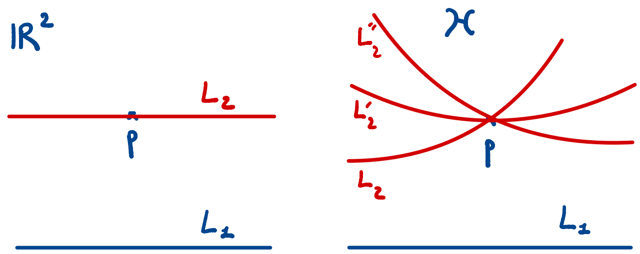


FIGURE 5.8. Parallel lines through a point in the Euclidean and hyperbolic planes