Geometry *

Zhiyuan Bai

Compiled on June 20, 2021

This document serves as a set of revision materials for the Cambridge Mathematical Tripos Part IB course *Geometry* in Lent 2021. However, despite its primary focus, readers should note that it is NOT a verbatim recall of the lectures, since the author might have made further amendments in the content. Therefore, there should always be provisions for errors and typos while this material is being used.

Contents

1 Topological and Smooth Surfaces **2**
- 1.1 Topological Surfaces 2
- 1.2 Smooth Surfaces 5
- 1.3 Orientability 8
- 1.4 Euler Characteristic 10

2 Geometry of Surfaces in Three-Dimensional Space **11**
- 2.1 Local Isometries and the First Fundamental Form 11
- 2.2 The Second Fundamental Form 14
- 2.3 The Gaussian Curvature 15

3 Geodesics and Abstract Riemannian Metrics **18**
- 3.1 Geodesics 19
- 3.2 Riemannian Metric 24

4 Hyperbolic Surfaces **26**
- 4.1 Möbius Maps and Inversions 26
- 4.2 Hyperbolic Models 27
- 4.3 Hyperbolic Geometry 29
- 4.4 Compact Hyperbolic Surfaces 30

5 Closing Remarks **32**
- 5.1 The Gauss-Bonnet Theorem and its Consequences 32
- 5.2 Moduli of Surfaces 34

*Based on the lectures under the same name taught by Prof. I. Smith in Lent 2021.

1 Topological and Smooth Surfaces

1.1 Topological Surfaces

A surface is, well, what you think it should be.

Definition 1.1. A topological surface is a second-countable (i.e. has a countable basis) Hausdorff topological space Σ such that every $p \in \Sigma$ has a neighbourhood $U \ni p$ which is homeomorphic to $\mathbb{R}^2$ ("locally Euclidean").

Remark. 1. Since $\mathbb{R}^2$ is homeomorphic to any open disc, an altenative definition of being locally Euclidean is to say every points in the space has a neighbourhood that is homeomorphic to an open subset of $\mathbb{R}^2$.
2. There exists topological spaces that is locally Euclidean but not Hausdorff or second-countable. But they are somewhat ungeometrical.

Lemma 1.1. *Subspaces of a Hausdorff space are Hausdorff; Subspaces of second countable space are second-countable.*

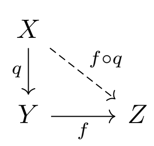
Since $\mathbb{R}^n$ is Hausdorff and second-countable, any subsets of it inherit these properties as well.

Proof. Exercise. $\square$

Example 1.1. 1. $\mathbb{R}^2$ is a topological surface.
2. Open subsets of $\mathbb{R}^2$ are topological surfaces.
3. If U is open in $\mathbb{R}^2$ and $f : U \to \mathbb{R}$ is a continuous function, then the graph of f, i.e. $\Gamma_f = \{(x, f(x)) : x \in U\} \subset \mathbb{R}^3$ is homeomorphic to U (via the correspondence $x \leftrightarrow (x, f(x))$), hence a topological surface as well.
4. $S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\} \subset \mathbb{R}^3$ is a surface with the homeomorphisms given by stereographic projections. What's nice is that it is also compact.

Definition 1.2. If Σ is a topological surface and $p \in \Sigma$, a chart on Σ is a pair (U, ϕ) where U is an open neighbourhood of p in Σ and $\phi : U \to V$ (for V open in $\mathbb{R}^2$) is a homeomorphism. The map ϕ^{-1} is called a local parameterisation (or frame of reference) of Σ near p.

A nice way to construction surfaces is to glue in pairs in the edges of a plane polygon. Naturally, this technique requires the theory of quotient spaces. Recall that if X is a space, Y a set and $q : X \to Y$, the quotient topology on Y induced by q is defined by saying $V \subset Y$ is open iff $q^{-1}(V) \subset X$ is. It can also be characterised by the universal property: If Z is any other space, then $f : Y \to Z$ is continuous iff $f \circ q : X \to Z$ is continuous.



In particular, if $\sim$ is an equivalence relation on X, then $X/\sim$ is defined to be the space of equivalence classes equipped with the quotient topology defined by the quotient map.

Also recall that we have the pasting lemma: If $X = A \cup B$ where A, B are closed in X and we have continuous $f_A : A \to Y$ and $f_B : B \to Y$ such that $f_A = f_B$ on $A \cap B$, then they define a continuous map $X \to Y$.
There is also the topological inverse function theorem: If X is a compact space and Y is Hausdorff, then any continuous bijection $f : X \to Y$ is a homeomorphism.
Consider $X = [0,1]^2$ with the minimal equivalence relation that contains $(x, 0) \sim (x, 1), (0, y) \sim (1, 1-y)$. This surface is called the Klein bottle.

Proposition 1.2. *The Klein bottle K is a topological surface.*

Proof. Quite obviously it is locally Euclidean around (the equivalence class of) any points in the interior of $[0,1]^2$. For (the equivalence class of) a point in the form $(0, y)$ where $y \neq 0, 1$, choose $\delta > 0$ such that $(0, y)$ has the distance $> \delta$ away from any vertex. Define $V_- = \overline{B((0,y),\delta)} \cap [0,1]^2, V_+ = \overline{B((1,1-y),\delta)} \cap [0,1]^2$, then $V = V_- \cap V_+$ (taken as its image under the quotient map) is compact and the map $\phi : V \to \overline{B(0,\delta)}$ given by

$$\phi(u, v) = \begin{cases} (u, v-y), \text{ if } (u,v) \in V_- \\ (u-1, 1-v-y), \text{ if } (u,v) \in V_+ \end{cases}$$

is well-defined and continuous by pasting lemma. It is also a bijection, hence also a homeomorphism as V is compact. Restricting ϕ to its interior gives the desired chart.
A similar strategy works around any $(0, x) \in K$ if $x \neq 0, 1$. We are left with the vertices, which are represented by just one point r in K. So just define a chart similarly by sensibly gluing together the 4 quarter-discs (instead of 2 semi-discs) of radius, say, $1/4$ around the vertices, and it works.
It is obviously second-countable and Hausdorff, so it is a topological surface. □

Using the same technique but slight variations, we can obtain some more topological surfaces.

Example 1.2. 1. $[0,1]^2/\sim$ with $(x,0) \sim (x,1), (0,y) \sim (1,y)$ gives the torus as an topological surface.
2. $[0,1]^2/\sim$ with $(x,0) \sim (0,x), (y,1) \sim (1,y)$, which is nto really a familiar object, is a topological surface.
3. For any polygon, if each edge has a unique edge with the same pairing (in the equivalence relation), then the resulting identification is a topological surface.
4. If we take a octagon and identify its edges by the labels $a, b, \bar{a}, \bar{b}, c, d, \bar{c}, \bar{d}$ (in that order, where $\bar{x}$ means label x but reversed orientation) then we obtain a torus but with two holes ("genus 2" surface). This surface, of course, can also be obtained by gluing together two regular tori along a cut-open hole.

Of course, we don't have to use up all the edges.

Example 1.3. 1. Take $[0,1] \times (0,1)/\sim$ with $(x,0) \sim (1-x,1)$ is the topological surface by the name of a Möbius band.
2. Glue two "closed Möbius bands" (Möbius bands but quotient from $[0,1] \times [0,1]$) together along their boundaries and, guess what, you got the Klein bottle.

Definition 1.3. The real projective plane is defined as $\mathbb{RP}^2 = [0,1]^2/\sim$ where $\sim$ is the minimal equivalence relation containing $(x,0) \sim (1-x,1), (0,y) \sim (1,1-y)$.

There are, of course, other ways to identify it. For example, $\mathbb{RP}^2$ can also be viewed as the unit disk with the two semicircles on the boundary identified in reverse direction. Another way to construct it is to take S^2 and identify antipodal points together. Observe that in this way, we can then identify each point as a line through the origin (by considering the line through it and its antipodal point) in $\mathbb{R}^3$.
Just another interesting thing before we move on: If we remove a disk from $\mathbb{RP}^2$ and join two copies together along the hole, we get the Klein bottle.

Remark. In these construction of surfaces by identifying sides, the precise parameterisation of $[0, 1]$ on any sidee does not affect the resulting topological surface.

We now turn to the space of lines X in $\mathbb{R}^2$. Let X_h be the set of non-vertical lines and X_v the set of non-horizontal lines in $\mathbb{R}^2$. We can uniquely specify a line in X_h with a parameter pair $(\theta, v) \in (-\pi/2, \pi/2) \times \mathbb{R}$ where θ is the slope and v the signed distance of the line from the origin. Then, in order to extend it to X, we can allow θ to reach its limit $\pm\pi/2$, we can collect the vertical lines as well. But there are some obvious redundancies, so what we want it to be is actually the surface $([-\pi/2, \pi/2] \times \mathbb{R})/\sim$ where $\sim$ is the minimal equivalence relation containing $(\pi/2, v) \sim (-\pi/2, -v)$. But what is it? The Möbius band, of course.
What would happen if we approach the problem from X_v? Exactly the same surface will pop out, but this time extending from $(0, \pi) \times \mathbb{R}$ (which X_v can be identified as).
Our aim is to give X a topology such that it becomes a topological surface that, per our expectation, homeomorphic to the Möbius band. This can be done by saying U is open iff $\phi_h(U \cap X_h) \subset (-\pi/2, \pi/2) \times \mathbb{R}$ and $\phi_v(U \cap X_v) \subset (0, \pi) \times \mathbb{R}$ are open where ϕ_u, ϕ_v are the respective identifications $X_u \to (-\pi/2, \pi/2) \times \mathbb{R}, X_v \to (0, \pi) \times \mathbb{R}$.
Globally, consider the set of lines that are neither horizontal nor vertical $X_{hv} = X_h \cap X_v$. Then there is an obvious map to consider:

$$\phi_{hv} = \phi_v \circ \phi_h^{-1} : \phi_h(X_{hv}) \to \phi_v(X_{hv})$$

which is a homeomorphism called the transition map.

Definition 1.4. The affine group $\text{Aff}(\mathbb{R}^2)$ is the group of transformations of $\mathbb{R}^2$ of the form $v \mapsto Av + b$ where $b \in \mathbb{R}^2, A \in \text{GL}_2(\mathbb{R})$.

This acts on X by homemomorphisms (under the topology we discussed earlier). Inside $\text{Aff}(\mathbb{R}^2)$, we have a subgroup $\text{Isom}(\mathbb{R}^2) = \{v \mapsto Av + b : A \in O(2)\}$ of isometries. These are precisely the affine transformations that preserves the usual Euclidean distance in $\mathbb{R}^2$.

Proposition 1.3. *There is no metric on X that both induces the topology on X and invariant under* $\text{Isom}(\mathbb{R}^2)$.

Proof. Suppose such a metric d exists. Given l, l' lines in $\mathbb{R}^2$ meeting at angle θ, we can translate them so they intersect at 0. Say $d(l, l') = f(\theta)$, then f has to be continuous and $f(\theta) \to 0$ as $\theta \to 0$. Take two parallel lines l, l' and suppose L traverse them at angle θ, then

$$0 < \delta = d(l, l') \leq d(l, L) + d(l', L) = 2f(\theta) \to 0$$

as $\theta \to 0$. Contradiction. □

Having seen enough examples, let's go back to some proper theory. If Σ is a topological surface, then by definition each $p \in \Sigma$ has an open neighbourhood homeomorphic to an open subset of $\mathbb{R}^2$. A chart (U, ϕ) for Σ, as we have seen, describes such a pair, where U is open in Σ and $\phi : U \to V$ is a homeomorphism where V is open in $\mathbb{R}^2$. If we have two charts $(U, \phi), (U', \phi')$ with nonempty intersection, then naturally we can consider the transition map $\phi' \circ \phi^{-1} : \phi(U \cap U') \to \phi'(U \cap U')$ which is by definition a homeomorphism.

Definition 1.5. An atlas for Σ is a collection of charts whose domains cover Σ.

Example 1.4. For S^2, we have an atlas with two charts given by stereographic projections from its north and south poles. One can compute the transition maps, but we'll not do it here. There are also many other possible charts we can think about. For example, we can project small enough neighbourhoods of points to a distant (non-degenrate) plane. Also, if we describe it in polar coordinates (ϕ, θ), then we can consider the natural chart $(\phi, \theta) = (\phi, \sin\theta)$ (which, notably, preserves area) or $(\phi, \theta) \mapsto (\phi, \log(\tan((\pi + 2\theta)/4)))$ which describes Mercatori projection.

Proposition 1.4. *Let Σ be a set given as a union of a countable collection of subsets $(U_\alpha)_{\alpha \in A}$. Suppose for each α we have a bijection $\phi_\alpha : U_\alpha \to V_\alpha$ where V_α open in $\mathbb{R}^2$ such that for any α, β, $\phi_\alpha(U_\alpha \cap U_\beta)$ is open in V_α.*
Then we can define a topology on Σ by saying $U \subset \Sigma$ is open iff for any α, $\phi_\alpha(U \cap U_\alpha) \subset V_\alpha$ is open in V_α. Furthermore, Σ is a topological space if the transition maps $\phi_\beta \circ \phi_\alpha^{-1}$ are all homeomorphisms and $\{(x, x) : x \in U_\alpha \cap U_\beta\} \subset U_\alpha \times U_\beta$ is closed. for all α, β.

This is analogous to how we built the space of lines.

Proof. Quite straightforward. □

Atlases don't seem much useful in the context of topological surfaces, then why do we care? Because they come in handy when we are dealing with smooth surfaces.

1.2 Smooth Surfaces

Definition 1.6. Let $U \subset \mathbb{R}^n, V \subset \mathbb{R}^m$ be open. We say $f : U \to V$ is smooth if it is infinitely differentiable. For arbitrary subsets $X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m$ is smooth if, locally, it is the restriction of a smooth map defined on an open subset of $\mathbb{R}^n$.

Definition 1.7. For $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$, a map $f : X \to Y$ is a diffeomorphism if it is a smooth homeomorphism with smooth inverse.

Definition 1.8. An abstract smooth surface Σ is a topological surface with an atlas of charts $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ such that the transition maps are diffeomorphisms.

Example 1.5. 1. S^2 with the atlas given by stereographic projections through the north and south poles is smooth surface.
2. The Klein bottle and the torus are smooth surfaces using the charts we defined before. Note that it is quite obvious that we can find a much smaller atlas, but the transition maps might turn out to be terrible.
3. The space of lines is a smooth surface using the atlas we discussed earlier as well.

Remark. 1. Usually, what we want to talk about is maximal atlases, i.e. including all charts for the surface that are compatible. This will not be discussed in technical detail in this course.
2. Being a topological surface is a structure of the topology, while being a smooth surface is data: It is possible to find surfaces admitting more than one compatible smooth structures.

Naturally, we want to talk about functions on smooth surfaces.

Definition 1.9. Suppose $f : \Sigma \to \mathbb{R}^2$ is continuous. We say f is smooth at $p \in \Sigma$ if $f \circ \phi^{-1}$ is smooth at $\phi(p)$ where (U, ϕ) is a chart around p.

Remark. The smoothness of f at p does not depend on the particular chart we choose since the transition maps are diffeomorphisms.

Similarly,

Definition 1.10. Suppose $f : \Sigma_1 \to \Sigma_2$ is a continuous map between smooth surfaces, then f is smooth at $p \in \Sigma_1$ if $\psi \circ f \circ \phi^{-1}$ is smooth at $\phi(p)$ for any charts (U, ϕ) around p and (V, ψ) around $f(p)$.

The same remark, of course, still applies.

Definition 1.11. Two abstract smooth surfaces are diffeomorphic if there exists a homeomorphism between them such that both the map and its inverse are smooth.

Sometimes, the surface we are interested in are embedded in some Euclidean space.

Definition 1.12. A smooth surface in $\mathbb{R}^3$ is a subset $\Sigma \subset \mathbb{R}^3$ which is locally diffeomorphic to $\mathbb{R}^2$.

Note that a smooth surface in $\mathbb{R}^3$ is, in particular, a topological surface and an abstract smooth surface. It is easy to tell whether a continuous map $V \to \Sigma$ is smooth where $V \subset \mathbb{R}^3$ is open – just take it as a map to $\mathbb{R}^3$ instead. The converse (i.e. determining whether a continuous map $g : U \to V$ is smooth where $U \subset \Sigma$ is open), however, is not very obvious.

Theorem 1.5. *The following are equivalent:*
(a) Σ *is a smooth surface in* $\mathbb{R}^3$.
(b) Σ *is locally the graph of a smooth function over one of the coordinate planes.*
(c) Σ *is locally cut out by the vanishing set of a smooth function with nonzero derivative. That is,* $\forall p \in \Sigma$, *there is an open* $\hat{U} \ni p$ *in* $\mathbb{R}^3$ *such that* $\Sigma \cap \hat{U} = f^{-1}(0)$ *where* $f : \hat{U} \to \mathbb{R}$ *is smooth and* $Df|_p \neq 0$.
(d) Σ *is locally the image of an "allowable" parameterisation, i.e. some* $\sigma : V \to \Sigma$ *where* $V \subset \mathbb{R}^2$ *is open such that* $\sigma(V) = U$ *is open in* Σ *and* $D\sigma$ *has full rank thoughout* V.

Proof. Quite obviously (b) implies everything else and (a) implies (d). It remains to show that (c) implies (b) and (d) implies (a), (b), both can be deduced from inverse function theorem (stated below). ☐

Theorem 1.6 (Inverse Function Theorem)**.** *Let* $U \subset \mathbb{R}^n$ *be open and* $f : U \to \mathbb{R}^n$ *be continuously differentiable. Suppose* $f(p) = q$ *and* $Df|_p$ *is invertible, then there exists an open neighbourhood* $V \ni q$ *and a differentiable map* $g : V \to \mathbb{R}^n$ *whose image is an open neighbourhood* $U' \subset U$ *of* p *such that* $f \circ g = \mathrm{id}_V$.

Remark. 1. We can replace "differentiable" by "smooth".
2. Once the premises of the theorem is satisfied, we have $Dg|_q = (Df|_p)^{-1}$.

If we have, however, a map $\mathbb{R}^n \to \mathbb{R}^m$ with $n > m$, then Df_p can't be invertible so we cannot possibly use the inverse function theorem. But we can do something almost as good.

Theorem 1.7 (Submersion Theorem). *Let $U \subset \mathbb{R}^n$ be open and suppose $f : U \to \mathbb{R}^m$ is continuously differentiable with $Df|_p$ surjective, then there exists open $U' \subset \mathbb{R}^n, V \subset \mathbb{R}^m$ containing the respective origins and diffeomorphisms $\phi : U' \to U, \psi : V \to \mathbb{R}^m$ with $\phi(0) = p, \psi(0) = q$ such that*

$$\begin{array}{ccc} U & \xrightarrow{f} & \mathbb{R}^m \\ \phi\uparrow & & \uparrow\psi \\ U' & \xrightarrow[F]{} & V \end{array}$$

commutes with $F(x_1, \ldots, x_n) = (x_1, \ldots, x_m)$.

Proof. WLOG $p = q = 0$ and that the first m columns of $Df|_p$ are linearly independent. Then $\Phi : U \to \mathbb{R}^n$ with

$$\Phi(x_1, \ldots, x_n) = (f(x_1, \ldots, x_n), x_{m+1}, \ldots, x_n)$$

has $D\Phi|_{p=0}$. By inverse function theorem, Φ is locally invertible near $\Phi(0) = 0$ and we obviously have $f \circ \Phi^{-1} = F$. □

So if we are interested in understanding the locus $f^{-1}(c)$ for $c \in \mathbb{R}^m$, then we can look at

$$\Phi^{-1}(c_1, \ldots, c_m, x_{m+1} \ldots, x_n) = (g_1(c), \ldots, g_m(c), x_{m+1} \ldots, x_n)$$

where g_i's acts like the inverses of f.

Theorem 1.8 (Implicit Function Theorem). *If $U \subset \mathbb{R}^k \times \mathbb{R}^m$ is open and $f : U \to \mathbb{R}^m$ is smooth with $Df|_p$ isomorphism when restricted to the coordinates $\{0_k\} \times \mathbb{R}^m$. Then there exists open $V \ni x_0$ in $\mathbb{R}^k$ and smooth $g : V \to \mathbb{R}^m$ such that $f(x, y) = 0 \iff y = g(x)$ for any $(x, y) \in (V \times \mathbb{R}^n) \cap U$.*

Proof. Rephrase the submersion theorem. □

Example 1.6. Take $f : \mathbb{R}^2 \to \mathbb{R}$ and $f(x_0, y_0) = 0$. Suppose $\partial f/\partial y|_{x_0,y_0} \neq 0$, then there exists smooth g defined locally near x_0 such that $g(x_0) = y_0$ such that $f(x, y) = 0 \iff y = g(x)$ for (x, y) near (x_0, y_0). We might not have an explicit formula for g right away, we can differentiate

$$f(x, g(x)) = 0 \implies \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} g'(x) = 0 \implies g'(x) = -\frac{f_x}{f_y}$$

Example 1.7. If we have $f : \mathbb{R}^3 \to \mathbb{R}$ with $f(x_0, y_0, z_0) = 0$. Assume $\partial f/\partial z|_{x_0,y_0,z_0} \neq 0$, then there exists an open neighbourhood $V \ni (x_0, y_0)$ in $\mathbb{R}^2$ and $g : V \to \mathbb{R}$ smooth such that locally $f'(0) = \{(x, y, g(x, y))\}$.

These conclude the proof of Theorem 1.5.

Remark. If Σ is a smooth surface in $\mathbb{R}^3$, then we know that it has a smooth atlas. Furthermore, any $p \in \Sigma$ has an open neighbourhood which is a graph over one of the coordinate planes. Suppose (U, ϕ) is a chart over Σ which is inverse to an allowable parameterisation of the form $(x, y, g(x, y))$, the transition maps between two such charts are of the form $(x, y) \mapsto (x, y), (x, y) \mapsto (x, g(x, y))$ or $(x, y) \mapsto (y, g(x, y))$ for some smooth g.

Example 1.8. We can take $S^2 \subset \mathbb{R}^3$ as $f^{-1}(0)$ where $f(x, y, z) = x^2+y^2+z^2-1$.

An interesting class of examples is the family of surfaces of revolution. Let $\gamma : (a, b) \to \mathbb{R}^2_{x,z} \subset \mathbb{R}^3$ via $\gamma(t) = (f(t), 0, g(t))$ be smooth with $\forall t \in (a, b), \gamma'(t) \neq 0, f(t) > 0$. Define $\sigma(u, v) = (f(u)\cos(v), f(u)\sin(v), g(u))$ where $(u, v) \in (a, b) \times (\theta, \theta + 2\pi)$ for some fixed θ. Now by calculation $\|\sigma_u \times \sigma_v\|^2 = f^2((f')^2 + (g')^2) \neq 0$, so $D\sigma$ is invertible. σ is also injective, so it gives an allowable parameterisation which implies that its image is a smooth surface in $\mathbb{R}^3$.

1.3 Orientability

Definition 1.13. Let U, V be connected open subsets of $\mathbb{R}^2$. A diffeomorphism $f : U \to V$ is orientation-preserving if $\det Df|_p > 0$ for all $p \in U$. We say f is orientation-reversing if $\det Df|_p > 0$ for all $p \in U$.

A diffeomorphism is either orientation-preserving or orientation-reversing since it cannot have non-invertible derivative anywhere.

Definition 1.14. An abstract smooth surface Σ is orientable if it admits a smooth atlas whose transition maps are all orientation-preserving. A choice of such a smooth atlas is called an orientation of Σ. If such a choice is made, we say Σ is oriented.

Example 1.9. 1. By computation, the transition maps for S^2 with stereographic projections along the north and south poles are orientation-preserving. So S^2 is orientable.
2. The torus is also orientatable with the atlas we discussed earlier.
3. The atlas we wrote down for the Klein bottle does not have all orientation-preserving transition maps. In fact, it is not orientable.

Let Σ be a smooth surface in $\mathbb{R}^3$ and $p \in \Sigma$. Let $\sigma : V \to \Sigma$ be an allowable parameterisation near p where $V \subset \mathbb{R}^2$ is open.

Definition 1.15. The tangent plane $T_p\Sigma$ to Σ at p is the 2-dimensional subspace of $\mathbb{R}^3$ given by the image $(D\sigma)|_p$.

Basically we want $p + T_p\Sigma$ to be the affine subspace (i.e. subspace but shifted) that is "tangent" to the surface.

Remark. $T_p\Sigma$ does not, in fact, depend on the choice of local parameterisation σ. If $\sigma, \hat{\sigma}$ are both local parameterisations near p. Then $\sigma = (\sigma \circ \hat{\sigma}^{-1}) \circ \hat{\sigma}$, so $D\sigma$ and $D\hat{\sigma}$ essentially have the same image.

Definition 1.16. At $p \in \Sigma$ where Σ is a smooth surface in $\mathbb{R}^3$, a normal vector to Σ is a vector in $\mathbb{R}^3$ that is orthogonal (via the usual inner product) to $T_p\Sigma$.

So there are exactly 2 unit normal vector to Σ at any $p \in \Sigma$.

Definition 1.17. A smooth surface Σ in $\mathbb{R}^3$ is two-sided if it admits a global continuous choice of unit normal vectors.

Lemma 1.9. *A smooth surface in $\mathbb{R}^3$ is orientable if and only if it is two-sided.*

Proof. Let $\sigma : V \to \Sigma$ be an allowable parameterisation. Define the positive normal $n_\sigma(p)$ be such that $(\sigma_u, \sigma_v, n_\sigma(p))$ form a right-handed basis of $\mathbb{R}^3$. Then $n_\sigma(p) = \sigma_u \times \sigma_v / \|\sigma_u \times \sigma_v\|$. If $\tilde{\sigma}$ is a different allowable parameterisation near p, then let $f = \tilde{\sigma}^{-1} \circ \sigma$ be the transition map.

$$Df|_p = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \implies \begin{cases} \sigma_u = \alpha \tilde{\sigma}_u + \gamma \tilde{\sigma}_v \\ \sigma_v = \beta \tilde{\sigma}_u + \delta \tilde{\sigma}_v \end{cases}$$

Therefore $\sigma_u \times \sigma_v = \det(Df|_p) \hat{\sigma}_u \times \hat{\sigma}_v$, which provides the equivalence. □

Example 1.10. A Möbius band in $\mathbb{R}^3$ admits the parameterisation

$$\sigma(t, \theta) = ((1 - t\sin(\theta/2))\cos\theta, (1 - t\sin(\theta/2))\sin\theta, t\cos(\theta/2))$$

for $(t, \theta) \in V_1 \cap V_2$ where $V_1 = \{t \in (-1/2, 1/2), \theta \in (0, 2\pi)\}, V_2 = \{t \in (-1/2, 1/2), \theta \in (-\pi, \pi)\}$. It is not quite a surface of revolution, but we can obtain it by revolving a line segment around an axis while at the same time rotate itself by half the speed.
One can check that $\sigma, D\sigma$ are injective on each of V_i for $i = 1, 2$. By calculation, $\sigma_t \times \sigma_\theta = (-\cos\theta\cos(\theta/2), -\sin\theta\cos(\theta/2), -\sin(\theta/2)) = n_{-\theta}$. But as $\theta \to 0^+, n_\theta \to (-1, 0, 0)$ whilst $\theta \to 2\pi^-, n_\theta \to (1, 0, 0)$. So the Möbius band (or rather, this particular Möbius band) is not orientable.

Lemma 1.10. *Orientability preserves under diffeomorphisms.*

Proof. Construct (or rather, refine) the atlas on one via the diffeomorphism from the other which already admits an orientation-preserving atlas. □

Proposition 1.11. *1. An abstract smooth surface Σ is orientable if and only if it does not contain a subsurface homeomorphic to a Möbius band.*
2. A compact smooth surface in $\mathbb{R}^3$ does not contain a Möbius band, hence is orientable.

Proof. Omitted (duh!). □

Remark. The "only if" part of (a) is trivial. The idea for the "if" part is following: WLOG the surface is connected. If we cover the surface by (finitely many – compactness!) charts and assign an orientation on one of them, then there is at most one choice on every other open set by looking at the intersections. Either it works, in which case we get an orientable surface, or it doesn't, where we will end up with a loop-like subsurface which one can refine to a Möbius band.
For (b), suppose we can find a compact surface in $\mathbb{R}^3$ that contains a Möbius band, then we can find a simple curve in $\mathbb{R}^3$ such that it hits the surface at odd number of points "transversely", which, as one can prove using more advanced tools, cannot happen.

Definition 1.18. A topological surface is orientable if and only if it does not contain a subsurface homeomorphic to the Möbius band.

This, of course, preserves under homeomorphisms.

1.4 Euler Characteristic

Definition 1.19. A subdivision of a compact topologcial surface Σ consists of
(i) A finite subset $V \subset \Sigma$ called the vertices;
(ii) A finite collection of continuous embeddings (edges) $\{e_i : [0,1] \to \Sigma\}$, each of which has endpoint in V, and any two of which meet at endpoints.
(iii) The connected components (faces) of $\Sigma \setminus (V \cup \bigcup_i e_i([0,1]))$ are homeomorphic to open discs.

Definition 1.20. A subdivision is a triangulation if the boundary of each face contains exactly 3 vertices and 3 edges. Also, (the closures of) any two faces are either disjoint, meet in a single vertex, or meet in a single edge.

Example 1.11. For S^2, we can take a single vertex on it, which is a subdivision with one vertex and one face. Or, we can have two vertices and link them with two edges. This gives two vertices, two edges and two faces. Alternatively, we can model any convex polyhedron on it with the respective number of vertices, edges, and faces. Amongst them, the tetrahedron is a triangulation. We can also find some subdivisions for torus (look it up!). Note that many of the trivial ones, although are subdivisions, are not triangulations.

Example 1.12. If Σ_1 and Σ_2 have triangulations with v_i, e_i, f_i vertices, edges and faces (for $i = 1, 2$), we can form a connected sum $\Sigma_1 \# \Sigma_2$ by removing one open face from each, and gluing the boundary edges and vertices together. So we end up with $v_1 + v_2 - 3, e_1 + e_2 - 3, f_1 + f_2 - 2$ vertices, edges and faces respectively.

Definition 1.21 (Euler Characteristic)**.** The Euler characteristic of a subdivision is $\chi = V - E + F$ where V, E, F are the number of vertices, edges and faces respectively.

Theorem 1.12. *(a) Every compact topological surface admits a triangulation.*
(b) The Euler characteristic of a compact topological surface is independent of triangulation, and is a homeomorphism invariant.
(c) Compact topological surfaces are completely determined (up to homeomorphism) by its Euler characteristic and orientability.

Proof. Nah. □

Example 1.13. We have $\chi(\Sigma_1 \# \Sigma_2) = \chi(\Sigma_1) + \chi(\Sigma_2) - 2$, so the Euler characteristic of a torus with g "holes" is $2 - 2g$. g here is called the genus of it.

Proposition 1.13. *If Σ is a compact orientable surface, then $\chi(\Sigma)$ is even.*

In fact, the orientable surface with genus g has Euler characteristic $2 - 2g$.

Sketch of proof. Fix a triangulation of Σ and let D be the (finite) set of directed edges of the triangulation. There are some natural permutations on D, namely $s : D \to D$ that reverses the direction of all edges, $\sigma : D \to D$ that coherently rotates the respective (directed) triangles anticlockwisely (which we can do this by orientability), and $\rho : D \to D$ that cycles the edges (directed away) on vertices clockwisely.
We have $\rho = \sigma \circ s$. Now s is a product of E transpositions, so $\operatorname{sgn} s = (-1)^E$.

Similarly $\operatorname{sgn}\sigma = (-1)^F$, but F is even since $2E = 3F$, so $\operatorname{sgn}\sigma = 1$. Also $\operatorname{sgn}\rho = (-1)^{V_{\text{even}}}$ where V_{even} is the number of vertices with even valence. So $\rho = \sigma \circ s$ gives $E \equiv V_{\text{even}} \pmod 2$. But by elementary combinatorics $V_{\text{odd}} \equiv 0 \pmod 2$, so $V \equiv E \pmod 2$, which implies that $\chi(\Sigma) = V - E + F$ is even. $\square$

2 Geometry of Surfaces in Three-Dimensional Space

2.1 Local Isometries and the First Fundamental Form

Definition 2.1. A smooth curve in $\mathbb{R}^3$ is a smooth $\gamma : (a, b) \to \mathbb{R}^3$ which is injective and has nowhere vanishing derivative. It's length is

$$L(\gamma) = \int_a^b \|\dot{\gamma}(t)\| \, \mathrm{d}t$$

The length is, of course, independent of parameterisation.

Definition 2.2. We say γ is parameterised by arc length if $\|\dot{\gamma}(t)\| = 1$ for all t.

It is easy to see that it is always possible to find a parameterisation by arc length.

Definition 2.3. Let Σ_1, Σ_2 be smooth surfaces in $\mathbb{R}^3$. We say Σ_1, Σ_2 are isometric if there is a diffeomorphism (called the isometry) $f : \Sigma_1 \to \Sigma_2$ such that any smooth curve $\gamma : (a, b) \to \Sigma_1 \subset \mathbb{R}^3$ has $L(f \circ \gamma) = L(\gamma)$.

Example 2.1. Any rigid motion, i.e. $v \mapsto Av + b, A \in O(3), b \in \mathbb{R}^3$, is an isometry.

Definition 2.4. Let Σ_1, Σ_2 be smooth surfaces in $\mathbb{R}^3$ and let $p \in \Sigma_1, q \in \Sigma_2$. We say Σ_1, Σ_2 are locally isometric near p, q is they admit open neighbourhoods $U_1 \subset \Sigma_1, U_2 \subset \Sigma_2$ such that U_1, U_2 are isometric.

Definition 2.5. Let Σ be a smooth surface in $\mathbb{R}^3$ and $p \in \Sigma$. The first fundamental form of Σ at p is the symmetric bilinear form on $T_p\Sigma$ defined by $\mathrm{I} : T_p\Sigma \times T_p\Sigma \to \mathbb{R}$ via $\mathrm{I}(v, w) = \langle v, w \rangle$ where $\langle v, w \rangle$ is the usual inner product.

We can define this more concretely in terms of allowable parameterisations. Suppose V is open in $\mathbb{R}^2$ and $\sigma : V \to \Sigma$ is an allowable parameterisation near p (say $0 \in V, \sigma(0) = p$), then $T_p\Sigma = \operatorname{Im}(D\sigma|_0) = \operatorname{span}(\sigma_u, \sigma_v)$ and

$$\mathrm{I} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}, E = \langle \sigma_u, \sigma_u \rangle, F = \langle \sigma_u, \sigma_v \rangle, G = \langle \sigma_v, \sigma_v \rangle$$

Sometimes we also write $\mathrm{I} = E \, \mathrm{d}u^2 + 2F \, \mathrm{d}u \, \mathrm{d}v + G \, \mathrm{d}v^2$.

Example 2.2. 1. For $\sigma(u, v) = (u, v, 0)$, $\mathrm{I} = \mathrm{d}u^2 + \mathrm{d}v^2$.
2. $\sigma(r, \theta) = (r \cos\theta, r \sin\theta, 0)$ has $\mathrm{I} = \mathrm{d}r^2 + r^2 \, \mathrm{d}\theta^2$.
3. The sphere of radius a can be parameterised in spherical polar coordinate by $\sigma(u, v) = (a \cos u \cos v, a \cos u \sin v, a \sin u)$. Then $\mathrm{I} = a^2 \, \mathrm{d}u^2 + a^2 \cos^2 u \, \mathrm{d}v^2$.

Suppose $\gamma : (a,b) \to \mathbb{R}^3$ lies on a smooth surface $\Sigma \subset \mathbb{R}^3$. Then locally we can write $\gamma(t) = \sigma(u(t), v(t))$, so $\|\dot\gamma(t)\| = \|\sigma_u \dot u + \sigma_v \dot v\| = \sqrt{E\dot u^2 + 2F\dot u\dot v + G\dot v^2} = \sqrt{\mathrm{I}/\mathrm{d}t^2}$. Consequently, smooth surfaces with the same fundamental forms in some local parameterisation are locally isometric there. The converse is, of course, also true.

Lemma 2.1. *If Σ_1, Σ_2 are locally isometric, say near $p \in \Sigma_1, q \in \Sigma_2$, then there exists local parameterisations near p, q with the same fundamental form.*

Proof. Suppose we have a local parameterisation $\sigma : V \to \Sigma_1$ a local parameterisation near p such that WLOG $0 \in V, \sigma(0) = p$. Let $\gamma_\epsilon : [0,\epsilon] \to \Sigma_1$ be $t \mapsto \sigma(t, 0)$. Then

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} L(\gamma_\epsilon) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} \int_0^\epsilon E(\sigma(t,0))^{1/2}\,\mathrm{d}t \to \sqrt{E(\sigma(0))} = \sqrt{E(p)}$$

as $\epsilon \to 0$. Similarly, $t \mapsto \sigma(0,t)$ determines $\sqrt{G}$ and $t \mapsto \sigma(t,t)$ determines $\sqrt{E + 2F + G}$, so the length behaviours of these local curves determines the first fundamental form. □

Example 2.3. Consider a cylinder of radius a with parameterisation $\sigma(u,v) = (a\cos v, a\sin v, u)$. Then $\mathrm{I} = \mathrm{d}u^2 + a^2\,\mathrm{d}v^2$. Consequently, there exists a parameterisation $\sigma(u', v')$ (by e.g. $u' = u, v' = av$) under which $\mathrm{I} = \mathrm{d}u'^2 + \mathrm{d}v'^2$. So the cylinder is locally isometric to the plane.

Example 2.4. Consider a cone $\sigma(u,v) = (\sqrt{2}u\cos(\sqrt{2}v), \sqrt{2}u\sin(\sqrt{2}v), \sqrt{2}u)$ which has $\mathrm{I} = 4\,\mathrm{d}u^2 + 4u^2\,\mathrm{d}v^2$. If we parameterise the plane by $\tilde\sigma(u,v) = (2u\cos v, 2u\sin v, 0)$, then it has the same first fundamental form as the cone exhibited. Consequently the cone is also locally isometric to the plane.

Definition 2.6. A ruled surface is a smooth surface in $\mathbb{R}^3$ swept by the segment of a straight line moving in space. Formally, a ruled surface is a surface that admits an allowable parameterisation of the form $\sigma(u,v) = \gamma(u) + v\delta(u)$ where $\gamma : (a,b) \to \mathbb{R}^3$ is a smooth curve with nonvanishing derivative and $\delta : (a,b) \to \mathbb{R}^3$ is smooth.

Remark. Note that for this parameterisation to be allowable, we need $(\gamma' + v\delta') \times \delta \neq 0$.

Example 2.5. 1. We can take all values of δ parallel. In this case, we call the surface a generalised cylinder. A special case of this is when δ is constant. In this case, this is an allowable parameterisation iff γ is never tangent to δ.
2. If all the ruling lines pass through a point $p \in \mathbb{R}^3$, we call the surface a generalised cone. Say $\delta(u) = \gamma(u) - p$, then $\sigma(u,v) = (1+v)\gamma(u) - vp$ which is injective iff p is not in the image of γ and no line through p meets γ twice. $D\sigma$ is injective (i.e. σ is allowable) provided that the ruling is not tangent to γ and $v \neq -1$.
3. A developable surface is a ruled surface which admits an allowable parameterisation of the form $\sigma(u,v) = \gamma(u) + v\gamma'(u)$. A necessary condition for this to work is that $v \neq 0$.
4. The hyperboloid of one sheet is $\Sigma = \{x^2 + y^2 = 1 + z^2\} \subset \mathbb{R}^3$ which can be parameterised by $\sigma(u,v) = (\cos u - v\sin u, \sin u + v\cos u, v)$ which can be written as $\sigma = \gamma(u) + v\delta(u)$ where $\gamma = (\cos u, \sin u, 0), \delta = (-\sin u, \cos u, 1) = \gamma'(u) + e_3$.

A natural question to ask is which smooth surfaces in $\mathbb{R}^3$ are locally isometric to the plane.

Lemma 2.2. *Developable surfaces are locally isometric to the plane.*

Proof. Let $\gamma : (a,b) \to \mathbb{R}^3$ be a smooth curve parameterised by arc length. Then the surface Σ parameterised by $\sigma(u,v) = \gamma(u) + v\gamma'(u)$ has (note that $\|\gamma'\| = 1$)

$$\mathrm{I} = (1 + v^2\kappa^2)\,\mathrm{d}u^2 + 2\,\mathrm{d}u\,\mathrm{d}v + \mathrm{d}v^2$$

where $\kappa = \|\gamma''\|$ is the curvature of γ. So the lemma will follow if we can find a plane curve (parameterised by arc length) $\rho(u) = (x(u), y(u), 0)$ whose curvature agrees locally with γ. This reduces to the differential equation

$$\begin{cases} \ddot{x} = \kappa\dot{y} \\ \ddot{y} = -\kappa\dot{x} \end{cases}, \dot{x}(t_0)^2 + \dot{y}(t_0)^2 = 1$$

which admits a local solution by the Picard-Lindelöf theorem. □

We can also use the first fundamental form to measure angles and areas. For $v, w \in \mathbb{R}^3$, then angle θ between v, w satisfies $\cos\theta = \langle v, w\rangle / \|v\|\|w\|$. So in a local parameterisation $\sigma : V \to \Sigma$ with $0 \in V, \sigma(0) = p, D\sigma|_0(\hat{v}) = v, D\sigma|_0(\hat{w}) = w$, we have

$$\begin{aligned} \cos\theta &= \frac{\mathrm{I}(v,w)}{\sqrt{\mathrm{I}(v,v)}\sqrt{\mathrm{I}(w,w)}} \\ &= \left(\hat{v}^\top \begin{pmatrix} E & F \\ F & G \end{pmatrix} \hat{w}\right) \Big/ \left(\left(\sqrt{\hat{v}^\top \begin{pmatrix} E & F \\ F & G \end{pmatrix} \hat{v}}\right)\left(\sqrt{\hat{w}^\top \begin{pmatrix} E & F \\ F & G \end{pmatrix} \hat{w}}\right)\right) \end{aligned}$$

Corollary 2.3. *Let $\Sigma \subset \mathbb{R}^3$ be a smooth surface in $\mathbb{R}^3$ and let $\sigma : V \to \Sigma$ be an allowable parameterisation for Σ near $p \in \Sigma$. Then σ preserves angles at p iff $E = G$ and $F = 0$ at p. Such a parameterisation is called conformal.*

Proof. Suppose $t \mapsto (u(t), v(t))$ and $t \mapsto (\tilde{u}(t), \tilde{v}(t))$ be two curves in V passing through WLOG $0 \in V$ at $t = 0$. Then, suppose their images meet at an angle θ on Σ, then

$$\cos\theta = \frac{E\dot{u}\dot{\tilde{u}} + F(\dot{u}\dot{\tilde{v}} + \dot{v}\dot{\tilde{u}}) + G\dot{v}\dot{\tilde{v}}}{\sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}\sqrt{E\dot{\tilde{u}}^2 + 2F\dot{\tilde{u}}\dot{\tilde{v}} + G\dot{\tilde{v}}^2}}$$

The "if" part follows immediately. For the "only if" part, we can see F have to vanish by taking the lines $t \mapsto (t, 0)$ and $t \mapsto (0, t)$. Similarly, $t \mapsto (t, t)$ and $t \mapsto (t, -t)$ gives $E = G$. □

Remark. If $E = G$ and $F = 0$, then the first fundamental form is just a scalar multiple of the identity $E(\mathrm{d}u^2 + \mathrm{d}v^2)$.

What other geometrical notions we might be interested in? Area, of course. Note that a parallelogram spanned by $v, w \in \mathbb{R}^2$ has area

$$\|v \times w\| = \sqrt{\langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle^2}$$

So for an allowable parameterisation σ of Σ, the "infinitesimal parallelogram" on the tangent space of Σ has area $\sqrt{EG - F^2} = \sqrt{\det \mathrm{I}}$.

Definition 2.7. If $U = \sigma(V) \subset \Sigma$ is a subset of a smooth surface in $\mathbb{R}^3$ with allowable parameterisation σ, then the area of U is

$$\int_V \sqrt{EG - F^2}\,\mathrm{d}A = \int_V \sqrt{\det \mathrm{I}}\,\mathrm{d}A$$

We, of course, would not be happy if this depends on σ.

Proposition 2.4. *The area does not depend on parameterisation.*

Proof. Suppose $\sigma : V \to \Sigma$ and $\tilde{\sigma} : \tilde{V} \to \Sigma$ are two different parameterisations, then there is a diffeomorphism (transition map) $f : \tilde{V} \to V$ such that

$$\begin{array}{ccc} \tilde{V} & \xrightarrow{f} & V \\ {\scriptstyle\tilde{\sigma}}\downarrow & \swarrow_{\sigma} & \\ \Sigma & & \end{array}$$

commutes. Let $\mathrm{I}, \tilde{\mathrm{I}}$ be the respective first fundamental forms, then

$$\tilde{\mathrm{I}} = (D\tilde{\sigma})^\top D\tilde{\sigma} = D(\sigma \circ f)^\top D(\sigma \circ f) = (Df)^\top \mathrm{I}(Df)$$

So

$$\int_{\tilde{V}} \sqrt{\det \tilde{\mathrm{I}}}\,\mathrm{d}\tilde{A} = \int_{\tilde{V}} \sqrt{\det \mathrm{I}} \det(Df)\,\mathrm{d}\tilde{A} = \int_V \sqrt{\det \mathrm{I}}\,\mathrm{d}A$$

by the change-of-coordinate formula. $\square$

It is also easy to see that the area is additive the same way as we expect it to be.

Example 2.6. Let Σ be the graph of a smooth function $(u, v, f(u, v))$ and $f : \mathbb{R}^2 \to \mathbb{R}$ smooth, then $\sqrt{EG - F^2} = \sqrt{1 + f_u^2 + f_v^2}$. So for example the area of the part of Σ that hoods over a disk is at least the area of that disk with equality iff that part of Σ is indeed planar.

2.2 The Second Fundamental Form

We also want to measure how much a surface in $\mathbb{R}^3$ is "bended" (i.e. deviate from its affine tangent space). This is not at all easy from the first fundamental form, so we shall introduce another notion for it.
As usual, let Σ be a smooth surface in $\mathbb{R}^3$ and $\sigma : V \to U \subset \Sigma$ be an allowable parameterisation. Over U, we can define the positive unit normal by

$$n = \frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|}$$

which does not depend on the specific parameterisation except maybe a sign since it is the normal to the tangent space. Near a point (u, v), we have

$$\sigma(u + h, v + l) = \sigma(u, v) + h\sigma_u + l\sigma_v + \frac{1}{2}(h^2\sigma_{uu} + 2hl\sigma_{uv} + l^2\sigma_{vv}) + O(h^3, l^3)$$

So the distance from $\sigma(u + h, v + l)$ to $T_p\Sigma$ is approximately

$$\langle n, \sigma(u + h, v + l) - \sigma(u, v)\rangle = \frac{1}{2}(\langle n, \sigma_{uu}\rangle h^2 + 2\langle n, \sigma_{uv}\rangle hl + \langle n, \sigma_{vv}\rangle l^2)$$

ignoring higher order terms. This prompts the theory on

Definition 2.8. The second fundamental form of Σ in the parameter σ is $\mathrm{II} = L\,\mathrm{d}u^2 + 2M\,\mathrm{d}u\,\mathrm{d}v + N\,\mathrm{d}v^2$ where $L = \langle n, \sigma_{uu}\rangle, M = \langle n, \sigma_{uv}\rangle, N = \langle n, \sigma_{vv}\rangle$. Alternatively, it is the symmetric bilinear form with matrix

$$\begin{pmatrix} L & M \\ M & N \end{pmatrix}$$

on $T_{\sigma(u,v)}\Sigma$.

Note that since $\langle n, n\rangle = 1, \langle n, n_u\rangle = \langle n, n_v\rangle = 0$. Also by definition $\langle n, \sigma_u\rangle = \langle n, \sigma_v\rangle = 0$. So we can write the second fundamental form alternatively as $L = -\langle n_u, \sigma_u\rangle, N = -\langle n_v, \sigma_v\rangle, M = -\langle n_u, \sigma_v\rangle = -\langle n_v, \sigma_u\rangle$. Recall that We can write the first fundamental form nicely in the form $\mathrm{I} = (D\sigma)^\top D\sigma$. Analogously, with this alternative form for the second fundamental form, we can write

$$\mathrm{II} = \begin{pmatrix} L & M \\ M & N \end{pmatrix} = -\begin{pmatrix} \langle n_u, \sigma_u\rangle & \langle n_u, \sigma_v\rangle \\ \langle n_v, \sigma_u\rangle & \langle n_v, \sigma_v\rangle \end{pmatrix} = -(Dn)^\top D\sigma$$

Remark. The first fundamental form is always non-degenerate, but the second fundamental form is in no way guaranteed to be non-degenerate. It can even vanish identically when (and only when) the surface is part of an affine plane.

Suppose Σ is an oriented surface in $\mathbb{R}^3$ and $p \in \Sigma$. We have $T_p\Sigma \subset \mathbb{R}^3$ and the unit normal vector $n(p) \in S^2$.

Definition 2.9. The Gauss map of an oriented smooth surface $\Sigma \subset \mathbb{R}^3$ is the map $\mathcal{N} : \Sigma \to S^2, p \mapsto n(p)$.

Remark. If σ is a parameterisation near p, then $\mathcal{N}(p) = \sigma_u \times \sigma_v / \|\sigma_u \times \sigma_v\|$. So when Σ is oriented, $\mathcal{N}(p)$ is globally well-defined over Σ and is smooth.

The second fundamental form can be written in terms of the derivative $D\mathcal{N}|_p : T_p\Sigma \to T_{\mathcal{N}(p)}S^2$ of the Gauss map. But what is it? Viewing $T_p\Sigma$ as the image of $D\sigma|_p$ (for $\sigma : V \to \Sigma, \sigma(0) = p$), then we have

$$\mathcal{N} \circ \sigma : (u, v) \mapsto \frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|}$$

We can differentiate this to get $D(\mathcal{N} \circ \sigma)|_0 : V \to \mathbb{R}^3$ which has image $T_{\mathcal{N}(p)}S^2$. This shows that $T_p\Sigma$ and $T_{\mathcal{N}(p)}S^2$ are actually parallel. So we can view $D\mathcal{N}|_p$ as a linear map on a fixed space $T_p\Sigma = T_{\mathcal{N}(p)}S^2$.

Definition 2.10. The shape operator $\mathbb{S} : T_p\Sigma \to T_p\Sigma$ of a smooth oriented surface Σ in $\mathbb{R}^3$ at $p \in \Sigma$ is $\mathbb{S} = -D\mathcal{N}|_p$ where $\mathcal{N} : \Sigma \to S^2$ is the Gauss map.

Then $\mathrm{II}(v, w) = \mathrm{I}(\mathbb{S}v, w)$ for $v, w \in T_p\Sigma$.

Lemma 2.5. $\mathbb{S}$ *is self-adjoint wrt* I.

Proof. I, II are both symmetric. □

2.3 The Gaussian Curvature

Definition 2.11. The Gauss curvature of a smooth surface Σ is the function $K(p) = \det(\mathbb{S}|_p) = \det(-D\mathcal{N}|_p) = \det(D\mathcal{N}|_p)$.

In local coordinates, as $v^\top \mathbb{S}^\top \mathrm{I}\, w = v^\top \mathrm{II}\, w$, we know that $\mathbb{S} = \mathrm{I}^{-1}\mathrm{II}$, so essentially $K = (LN - M^2)/(EG - F^2)$.

Example 2.7. The cylinder $\sigma(u, v) = (a\cos u, a\sin u, v)$ has $\sigma_{uv} = \sigma_{vv} = 0$, hence $K = (LN-M^2)/(EG-F^2) = 0$ locally. Another way to see it is to observe that the image of the Gauss map is contained in the equator, consequently $D\mathcal{N}|_p$ always has one dimensional image, hence have zero determinant.

Example 2.8. Suppose Σ can be locally parameterised as the graph of a function $\sigma(u, v) = (u, v, f(u, v))$ where $f : V \to \mathbb{R}$ is smooth. Then by tiring computation we have

$$K = \frac{f_{uu}f_{vv} - f_{uv}^2}{(1 + f_u^2 + f_v^2)^2} = \frac{\det H}{(1 + f_u^2 + f_v^2)^2}$$

where H is the Hessian of f. If $f(u, v) = \sqrt{r^2 - u^2 - v^2}$ (i.e. a piece of the sphere with radius r), then at $(0, 0, r)$, $f_{uu} = f_{vv} = -1/r$ and $f_{uv} = 0$, so $K = 1/r^2$. The group of rigid motions of $\mathbb{R}^3$ acts transitively on the sphere, so $K = 1/r^2$ identically. This is unsurprising since the Gauss map in this case is simply scaling by $1/r$.

Definition 2.12. If Σ is a smooth surface in $\mathbb{R}^3$, then $p \in \Sigma$ is
(a) Elliptic if $K(p) > 0$.
(b) Hyperbolic if $K(p) < 0$.
(c) Parabolic if $K(p) = 0$.

Lemma 2.6. *In a sufficiently small neighbourhood of an elliptic point, the surface Σ lies on one side of its tangent plane; In any neighbourhood of a hyperbolic point, the surface meets both sides of its tangent plane.*

Proof. Take a local paramterisation $\sigma : V \to \Sigma$ near p such that $(0, 0) \in V, \sigma(0, 0) = p$. Recall that if $W = u\sigma_u|_p + v\sigma_v|_p \in T_p\Sigma$, then $\mathrm{II}(w, w)/2$ is an infinitesimal measure of distance from $\sigma(u, v)$ to $T_p\Sigma$. If p is elliptic, then II is locally either positive definite or negative definite. So the distance from $\sigma(u, v)$ to $T_p\Sigma$ has constant sign locally, which means that Σ locally lies on one side of $T_p\Sigma$. If p is hypernbolic, then $\det \mathrm{II} < 0$, which means that $\mathrm{II}(w, w)$ can take both signs in any neighbourhood of p, i.e. Σ intersects both sides of $T_p\Sigma$. □

In the parabolic case, however, both can happen. The cylinder gives an example where the surface always lies on one side of its tangent plane; Whereas the "monkey saddle" $\sigma(u, v) = (u, v, u^3 - 3v^2u)$ intersects both sides of its tangent point at $(0, 0, 0)$.

Proposition 2.7. *If $\Sigma \subset \mathbb{R}^3$ is a compact smooth surface, then Σ has an elliptic point.*

Proof. By compactness, Σ can be contained inside $\bar{B}(0, r)$ for any sufficiently large r. Take $R = \inf_{\Sigma\subset\bar{B}(0,r)} r$. Then $\bar{B}(0, r)$ touches Σ. WLOG they intersect at $p = (0, 0, R)$. Then locally near $(0, 0) \in \mathbb{R}^2 \subset \mathbb{R}^3$, Σ is the graph of a smooth function f such that $f - \sqrt{R^2 - u^2 - v^2} \le 0$. Taylor expansion gives

$$\frac{1}{2}(f_{uu}u^2 + 2f_{uv}uv + f_{vv}v^2) + \frac{1}{2R}(u^2 + v^2) \le 0$$

which exactly means that II is locally negative definite. In particular, $K(p) > 0$. $\square$

Definition 2.13. We say a surface Σ is flat if $K = 0$ everywhere on Σ.

How to check if a surface is flat? We answer this question by working out what the shape operator is actually doing in terms of the parameterisation. Suppose $\sigma : V \to U \subset \Sigma$ is a parameterisation of an open set $U \subset \Sigma$. Let $p \in U$ and as usual we assume WLOG that $(0,0) \in V, \sigma(0,0) = p$ Let $\gamma(t) = \sigma(u(t), v(t))$ where $u(0) = v(0) = 0$. We observe

$$\begin{aligned} D\mathcal{N}|_p(\sigma_u \dot{u}(0) + \sigma_v \dot{v}(0)) = D\mathcal{N}|_p(\dot{\gamma}(0)) &= \frac{\mathrm{d}}{\mathrm{d}t} n(u(t), v(t))\bigg|_{t=0} \\ &= n_u \dot{u}(0) + n_v \dot{v}(0) \end{aligned}$$

So by varying the curve we see that $D\mathcal{N}|_p$ takes σ_u to n_u and σ_v to n_v. In particular, the curvature K of Σ at $p = \sigma(0,0)$ vanishes iff $n_u \times n_v = 0$ at p.

Theorem 2.8. *Let Σ be a smooth surface in $\mathbb{R}^3$ and $p \in \Sigma$ be such that $K(p) \neq 0$. Let U be an open neighbourhood of Σ and a sequence of open sets $(A_i), p \in A_i \subset U$ that shrinks to p (in the sense that $\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall i > N, A_i \subset B(p, \epsilon)$). Then*

$$\frac{\mathrm{Area}_{S^2}(\mathcal{N}(A_i))}{\mathrm{Area}_{\Sigma}(A_i)} \to |K(p)|$$

as $i \to \infty$.

To get the sign of the Gaussian curvature, observe that the Gauss map reverses local orientation near hyperbolic points and preserves it near elliptic points.

Proof. WLOG K does not change sign on U. Fix local parameterisation $\sigma : V \to U$ such that (WLOG) $(0,0) \in V, \sigma(0,0) = p$. Identify $A_i \subset U$ with $V_i \subset V$, then

$$\mathrm{Area}_{\Sigma}(A_i) = \int_{V_i} \sqrt{EG - F^2}\, \mathrm{d}u\, \mathrm{d}v = \int_{V_i} \|\sigma_u \times \sigma_v\|\, \mathrm{d}u\, \mathrm{d}v$$

The Gauss map $\mathcal{N}$ is a local diffeomorphism on U as K does not vanish on U. So $\mathcal{N} \circ \sigma : V \to S^2$ is an allowable parameterisation for an open subset of S^2 around $\mathcal{N}(p) \in S^2$, which gives

$$\begin{aligned} \mathrm{Area}_{S^2}(\mathcal{N}(A_i)) &= \int_{V_i} \|n_u \times n_v\|\, \mathrm{d}u\, \mathrm{d}v = \int_{V_i} \|D\mathcal{N}(\sigma_u) \times D\mathcal{N}(\sigma_v)\|\, \mathrm{d}u\, \mathrm{d}v \\ &= \int_{V_i} K(u,v) \|\sigma_u \times \sigma_v\|\, \mathrm{d}u\, \mathrm{d}v \end{aligned}$$

The continuity of K then yields the conclusion. $\square$

This gives a nice geometrical intuition of Gaussian curvature when it is nonzero. Naturally, we then want to see what is going on with flat surfaces. Recall that ruled surfaces are those surfaces swept by a moving line segment $\sigma(u,v) = \gamma(u) + v\delta(u)$. Examples of this includes hyperboloid of one sheet, generalised cyliners, generalised cones and developable surfaces. Note that just the surface itself does not allow us to uniquely recover γ, so there are certain choices to be made here.

Lemma 2.9. *Ruled surfaces have nonpositive Gaussian curvature.*

Proof. $\sigma_{vv} = 0 \implies N = 0 \implies K = -M^2/\det \mathrm{I} \leq 0$. □

But which ones are actually flat? We know generalised cyliners, generalised cones and developable surfaces are, but the hyperboloid of one sheet is not.

Proposition 2.10. *Let Σ be a smooth ruled surface in $\mathbb{R}^3$ which is not a generalised cylinder or generalised cone. Then Σ is flat iff it is a developable surface.*

Proof. Σ is flat iff $M = 0 \iff 0 = \langle n, \sigma_{uv} \rangle = \langle \sigma_u \times \sigma_v, \delta' \rangle$ which happens exactly when $\langle \delta', \gamma' \times \delta \rangle = 0$. WLOG δ is a unit vector (so $\langle \delta, \delta' \rangle = 0$) and Σ is not a generlised cylinder (so $\delta' \neq 0$ locally). Then $\gamma' = g\delta' + f\delta$. for some smooth f, g.
If $f = g'$, then $\gamma' = (g\delta)'$, which means that $\gamma = g\delta + a$ for some constant a. The change of coordinate $\tilde{u} = u, \tilde{v} = v + g(u)$ then reveals that Σ is a generalised cone.
Otherwise, define $\tilde{\gamma}(u) = \gamma(u) - g(u)\delta(u)$ and $\tilde{v} = (v + g(u))/(f(u) - g'(u))$ (which is valid at least locally), then $\sigma(u, v) = \tilde{\gamma}(u) + \tilde{v}\tilde{\gamma}'(u)$ is developable. □

Corollary 2.11. *Flat ruled surfaces are also locally isometric to the plane.*

3 Geodesics and Abstract Riemannian Metrics

The geometry of a smooth surface in $\mathbb{R}^3$ is constrained by three rigidity theorems:

Theorem 3.1 (The Fundamental Theorem of Surfaces in $\mathbb{R}^3$)**.** *A connected smooth surface in $\mathbb{R}^3$ is determined by its first and second fundamental forms up to rigid motion.*

Proof. Sketched in example sheet. □

Theorem 3.2 (Theorema Egregium)**.** *Isometric surfaces have the same Gaussian curvature.*

That is, Gaussian curvature is invariant under "bending" the surface.

Sketch of proof. We will show later in the section that each smooth surface in $\mathbb{R}^3$ admits parameterisations such that $\mathrm{I} = \mathrm{d}u^2 + G(u, v)\,\mathrm{d}v^2$. This can be achieved by using the coordinate given by some "nice curves" (in fact, geodesics). Under this situation, we can simply verify that $K = -(\sqrt{G})_{uu}/\sqrt{G}$ which gives the theorem. □

Theorem 3.3 (Gauss-Bonnet Theorem for Embedded Surfaces)**.** *If Σ is a compact smooth surface in $\mathbb{R}^3$ with area form $\mathrm{d}A$, then*

$$\int_\Sigma K \,\mathrm{d}A = 2\pi\chi(\Sigma)$$

Proof. Omitted. □

3.1 Geodesics

Let $\gamma : [a, b] \to \mathbb{R}^3$ be a smooth curve, recall that we defined

$$L(\gamma) = \int_a^b \|\dot{\gamma}(t)\| \,\mathrm{d}t$$

Definition 3.1. The energy of γ is

$$\mathrm{E}(\gamma) = \int_a^b \|\dot{\gamma}(t)\|^2 \,\mathrm{d}t$$

Definition 3.2. Let $\gamma : [a, b] \to \Sigma$ be a smooth map with Σ a smooth surface in $\mathbb{R}^3$. A one-parameter variation (with fixed endpoints) of γ is a smooth map $\Gamma : (-\epsilon, \epsilon) \times [a, b] \to \Sigma$ (write $\gamma_s(t) = \Gamma(s, t)$) such that $\gamma_0 = \gamma$ and $\gamma_s(a), \gamma_s(b)$ constant in s.

Definition 3.3. A smooth curve $\gamma : [a, b] \to \Sigma$ is a geodesic (from $\gamma(a)$ to $\gamma(b)$) if for every one-parameter variation (γ_s) of γ with fixed endpoints,

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathrm{E}(\gamma_s)\bigg|_{s=0} = 0$$

Suppose that $\gamma : [a, b] \to \Sigma$ has image or embedded curve lying inside the image of an allowable parameterisation $\sigma : V \to \Sigma$. Then we can write $\gamma_s(t) = \sigma(u(s, t), v(s, t))$ and

$$\mathrm{E}(\gamma_s) = \int_a^b E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2 \,\mathrm{d}t$$

So by differentiating under the integral sign and integration by part,

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathrm{E}(\gamma_s)\bigg|_{s=0} = \int_a^b A\frac{\partial u}{\partial s} + B\frac{\partial v}{\partial s} \,\mathrm{d}t\bigg|_{s=0}$$

where

$$A = E_u\dot{u}^2 + 2F_u\dot{u}\dot{v} + G_u\dot{v}^2 - 2\frac{\mathrm{d}}{\mathrm{d}t}(E\dot{u} + F\dot{v})$$
$$B = E_v\dot{u}^2 + 2F_v\dot{u}\dot{v} + G_v\dot{v}^2 - 2\frac{\mathrm{d}}{\mathrm{d}t}(F\dot{u} + G\dot{v})$$

Corollary 3.4. *A smooth curve $\gamma(t) = \sigma(u(t), v(t))$ is a geodesic if and only if it satisfies the geodesic equations*

$$\frac{\mathrm{d}}{\mathrm{d}t}(E\dot{u} + F\dot{v}) = \frac{E_u\dot{u}^2 + 2F_u\dot{u}\dot{v} + G_u\dot{v}^2}{2}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}(F\dot{u} + G\dot{v}) = \frac{E_v\dot{u}^2 + 2F_v\dot{u}\dot{v} + G_v\dot{v}^2}{2}$$

Remark. Unlike $L(\gamma)$, $\mathrm{E}(\gamma)$ does depend on the parameterisation of γ. Nonetheless, we always have $L(\gamma)^2 \leq \mathrm{E}(\gamma)(b-a)$ by Cauchy-Schwartz with equality hold iff γ has constant speed (e.g. parameterised by arc-length).

Corollary 3.5. *1. If γ locally minimises length and has constant speed, then γ is a geodesic.*
2. If γ globally minimises energy, then it globally minimises length with the same endpoints and has constant speed.

Example 3.1. 1. On the plane $\sigma(u,v) = (u,v,0)$, the geodesic equations boils down to $\ddot{u} = \ddot{v} = 0$, so the curve has to be straight lines.
2. On the sphere $\sigma(u,v) = (\cos u \cos v, \cos u \sin v, \sin u)$ the first fundamental form is $\mathrm{d}u^2 + \cos^2 u \,\mathrm{d}v^2$, so the geodesic equations become

$$\begin{cases} 0 = \ddot{u} + \dot{v}^2 \sin u \cos u \\ 0 = \ddot{v} - 2\dot{u}\dot{v} \tan u \end{cases}$$

Assuming that the geodesics is parameterised at unit speed (we will justify it later), then we have $\dot{u}^2 + \dot{v}^2 \cos^2 u = 1$. Solving the second equation gives

$$\dot{v} = \frac{c}{\cos^2 u} \implies \dot{u} = \sqrt{\frac{\cos^2 u - c^2}{\cos^2 u}}$$

where c is a constant. So

$$\begin{aligned} v &= \int \frac{\mathrm{d}v}{\mathrm{d}u} \,\mathrm{d}u = \int \frac{\dot{v}}{\dot{u}} \,\mathrm{d}u = \int \frac{c \sec^2 u}{\sqrt{1 - c^2 \sec^2 u}} \,\mathrm{d}u \\ &= \int \frac{\mathrm{d}w}{\sqrt{1 - w^2}}, w = \frac{c \tan u}{\sqrt{1 - c^2}} \\ &= \sin^{-1}(w) + \delta = \sin^{-1}(\lambda \tan u) + \delta \end{aligned}$$

where $\lambda = c/\sqrt{1 - c^2}$ and δ is a constant. Consequently,

$$\sin(v - \delta) = \lambda \tan u \implies (\sin v \cos u) \cos \delta - (\cos v \cos u) \sin \delta - \lambda \sin u = 0$$

Hence γ lies on the intersection of a plane through origin and S^2, i.e. a great circle. One can check that all arcs of great circles are indeed geodesics on S^2, either manually or using the more general argument that will be covered later.
3. Consider the torus $\sigma(u,v) = ((a + \cos u) \cos v, (a + \cos u) \sin v, \sin u)$. The first fundamental form is $\mathrm{d}u^2 + (a + \cos u)^2 \,\mathrm{d}v^2$, so the geodesic equations are

$$\begin{cases} 0 = \ddot{u} + \dot{v}^2 (a + \cos u) \sin u \\ 0 = \ddot{v} - 2\dot{u}\dot{v} \sin u/(a + \cos u) \end{cases}$$

subject to the unit speed condition $\dot{u}^2 + (a + \cos u)^2 \dot{v}^2 = 1$. Following exactly like what we did for the sphere gives

$$\frac{\mathrm{d}v}{\mathrm{d}u} = \frac{c\sqrt{a + \cos u}}{\sqrt{(a + \cos u)^2 - c^2}}$$

for some constant c. Sadly, this doesn't seem to integrate to a closed form in terms of elementary functions.

The straight lines that are geodesics on the plane are, by our argument, locally the shortest among a one-parameter family of variations. They are also locally "straightest", but what does that even mean?

Proposition 3.6. *Let Σ be a smooth surface in $\mathbb{R}^3$ and $\gamma : [a,b] \to \Sigma$ be a smooth curve on Σ. Then γ is a geodesics if and only if the derivative of the tangent vector to γ is everywhere normal to Σ.*

Proof. Suppose we have a local parameterisation $\sigma : V \to U$ where $U \subset \Sigma$ is open. We write $\gamma(t) = \sigma(u(t), v(t))$. Then $\dot\gamma(t) = \sigma_u \dot u(t) + \sigma_v \dot v(t)$. $\ddot\gamma$ is normal to Σ exactly when it is orthogonal to σ_u, σ_v, i.e.

$$\left\langle \frac{\mathrm{d}}{\mathrm{d}t}(\sigma_u \dot u + \sigma_v \dot v), \sigma_u \right\rangle = 0 = \left\langle \frac{\mathrm{d}}{\mathrm{d}t}(\sigma_u \dot u + \sigma_v \dot v), \sigma_v \right\rangle$$

The first equality gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \sigma_u \dot u + \sigma_v \dot v, \sigma_u \rangle - \left\langle \sigma_u \dot u + \sigma_v \dot v, \frac{\mathrm{d}\sigma_u}{\mathrm{d}t} \right\rangle = 0$$

which translates to

$$\frac{\mathrm{d}}{\mathrm{d}t}(E\dot u + F\dot v) - \langle \sigma_u \dot u + \sigma_v \dot v, \sigma_{uu}\dot u + \sigma_{vv}\dot v \rangle = 0$$

Observe that $E_u = 2\langle \sigma_u, \sigma_{uu} \rangle$ and $G_u = 2\langle \sigma_v, \sigma_{uv} \rangle$, so this equation is just

$$\frac{\mathrm{d}}{\mathrm{d}t}(E\dot u + F\dot v) = \frac{1}{2}(E_u \dot u^2 + 2F_u \dot u \dot v + G_u \dot v^2)$$

which is the first geodesic equation. We can similarly obtain the second geodesic equation from the second equality. □

Remark. 1. In particular, this proves that all geodesics have constant speed.
2. Another way to understand this is to view geodesics as trajectories of "free particles" on Σ.

Corollary 3.7. *Suppose we have a surface Σ in $\mathbb{R}^3$ and a plane $\Pi \subset \mathbb{R}^3$ such that $C = \Pi \cap \Sigma$ is a smooth curve and Σ is preserved under reflection in Π. Then C is a geodesic on Σ when parameterised at constant speed.*

Proof. Observe that we can write $\mathbb{R}^3$ as $\Pi \oplus \Pi^\perp$, but we can also write $\mathbb{R}^3$ as $T_p\Sigma \oplus \mathbb{R}n_p$ for $p \in \Sigma$. The linear transformation as the reflection across Π acts as $\mathrm{id}_\Pi \oplus (-\mathrm{id}_{\Pi^\perp})$ in the first description. Also, if $p \in C$, we know that this reflection preserves $T_p\Sigma$ and is identity on $\mathbb{R}n_p$. In particular, $n_p \in \Pi$. If we locally parameterise C via $\gamma : (-\epsilon, \epsilon) \to \Sigma$ with $\gamma(0) = p$, then $\dot\gamma, \ddot\gamma$ have images both contained in Π. As we parameterised γ at constant speed, $\langle \dot\gamma, \ddot\gamma \rangle = 0$, so necessarily $\ddot\gamma(0)$ is in the direction of n_p. Hence γ is a geodesic by the preceding proposition. □

Corollary 3.8. *Arcs of great circles on S^2 are geodesics.*

A nice class of example of surfaces we always use is, of course, surfaces of revolution. Recall that for a curve $\eta(u) = (f(u), 0, g(u))$ in the xz plane such that η is smooth and injective with $f(u) > 0$, its surface of revolution has the parameterisation $\sigma(u,v) = (f(u)\cos v, f(u)\sin v, g(u))$. There are some canonical curves on a surface of revolution.

Definition 3.4. A parallel is a circle obtained by rotating a point of η (i.e. fixing u). A meridian is a curve obtained from η by intersecting it with a vertical plane through the origin (i.e. fixing v).

What does geodesics on such a surface look like? The meridians on Σ have to be geodesics because of Corollary 3.7. In general, the surface has the first fundamental form $((f')^2+(g')^2)\,\mathrm{d}u^2+f^2\,\mathrm{d}v^2$. If we parameterise η by arc length, then this reduces to $\mathrm{d}u^2+f^2\,\mathrm{d}v^2$, so the geodesics equations become

$$\ddot{u}=\dot{v}^2 f\frac{\mathrm{d}f}{\mathrm{d}u},\ \frac{\mathrm{d}}{\mathrm{d}t}(f^2\dot{v})=0$$

The constant speed hypothesis for a geodesics adds the condition $\dot{u}^2+f^2\dot{v}^2=1$.

Lemma 3.9. *A parallel $u=u_0$ (when parameterised at constant speed) is a geodesic if and only if $f'(u_0)=0$.*

Proof. Just look at the geodesic equations. ☐

Another quantitative property of geodesics on surfaces of revolution is known as Clairaut relation. For a curve γ on the surface, at each point on it, we can obtain a unique parallel on which the point reside. Let ρ be the radius of the parallel and θ be the angle made by γ and the parallel at the point. Then

Proposition 3.10. *If γ is a geodesic, then $\rho\cos\theta$ is constant along γ.*

Proof. Write $\gamma(t)=\sigma(u(t),v(t))$, then σ_v is exactly the tangent vector to the parallel, so

$$\cos\theta=\frac{\langle\sigma_v,\sigma_u\dot{u}+\sigma_v\dot{v}\rangle}{\|\sigma_v\|\|\sigma_u\dot{u}+\sigma_v\dot{v}\|}$$

Assuming that γ is parameterised by arc length, then this reduces to $\cos\theta=|f(u)\dot{v}|=\rho\dot{v}$. We then obtain the result by the second geodesic equation $(\mathrm{d}/\mathrm{d}t)(f^2\dot{v})=0$. ☐

Remark. We have a partial converse to this: If γ satisfies the Clairaut relation and does not run along a parallel, then γ is a geodesic. The proof is exercise.

Example 3.2. Consider the ellipsoid of revolution given by rotating an arc of an ellipse which meets the z-axis orthogonally. The "waist curve" (the parallel in the middle) is then the only parallel that is a geodesic. For other geodesics, the Clairaut relation tells us that $\rho\cos\theta$ is constant along γ with value, say, $\rho_0\cos\theta_0$ where ρ_0,θ_0 is the respective values of ρ,θ at a chosen point on the curve. Suppose γ is not a meridian, then we can assume that $\theta_0\in[0,\pi/2)$. Then ρ has to be bounded below, so the geodesics are essentially trapped between two parallels!

Note that it is natural to allow geodesics to self-intersect, as we only need it to have a well-defined tangent vector everywhere. So $\gamma:(a,b)\to\Sigma$ is not necessarily an embedding, but being an immersion suffices.
How do we know the existence theory of geodesics? This, of course, follows from Picard-Lindelöf again. Why can we use this? Note that the geodesics equation can be written alternatively as

$$\begin{pmatrix}E & F\\ F & G\end{pmatrix}\begin{pmatrix}\ddot{u}\\ \ddot{v}\end{pmatrix}=\text{some smooth expression involving } u,v,\dot{u},\dot{v}$$

But the first fundamental form is positive definite hence invertible, therefore we can arrive at a system

$$\begin{cases} \ddot{u} = A(u, v, \dot{u}, \dot{v}) \\ \ddot{v} = B(u, v, \dot{u}, \dot{v}) \end{cases}$$

for some smooth A, B. The usual substitution $p = \dot{u}, q = \dot{v}$ then turns the ODE into the form where Picard-Lindelöf applies. The corresponding Lipschitz condition follows from the smoothness of A, B (which guarantees a lower bound on $\|DA\|$ and $\|DB\|$).

Corollary 3.11. *If Σ is a smooth surface in $\mathbb{R}^3$ and $p \in \Sigma, v \in T_p\Sigma \setminus \{0\}$, then there exists $\epsilon > 0$ and a unique geodesic $\gamma : [0, \epsilon) \to \Sigma$ parameterised at unit speed such that $\gamma(0) = p, \dot{\gamma}(0) = v$. Furthermore, γ depends smoothly on (p, v).*

Now fix $p \in \Sigma$ and consider a geodesic arc (parameterised by arc length) γ starting at p. For $t > 0$ small, we can find geodesic (by the existence theorem as above) γ_t also parameterised by arc length such that $\gamma_t(0) = \gamma(t)$ and $\gamma_t'(0) \perp \gamma'(t)$. Let $\sigma(u, v) = \gamma_v(u)$ defined on some small $[0, \epsilon) \times [0, \delta)$.

Lemma 3.12. *σ defines an allowable parameterisation.*

Proof. γ is smooth because of the smooth dependence of a geodesic on the initial conditions. Also $D\sigma \neq 0$ on a small enough open set since $\gamma_t'(0) \perp \gamma'(t)$. These suffice to imply the result. □

Corollary 3.13. *Any smooth surface Σ in $\mathbb{R}^3$ admits local parameterisations with respect to which the first fundamental form has the form $\mathrm{d}u^2 + G(u, v)\,\mathrm{d}v^2$.*

Proof. Just take the local parameterisation as the σ as above. Then $E = \langle \sigma_u, \sigma_u \rangle = 1$ because $u \mapsto \gamma_{v_0}(u)$ is parameterised by arc length for each v_0. As for F, consider

$$\frac{\mathrm{d}}{\mathrm{d}t}(F\dot{u} + G\dot{v}) = \frac{E_v \dot{u}^2 + 2F_v \dot{u}\dot{v} + G_v \dot{v}^2}{2}$$

On the curve $v = v_0$ constant and $u(t) = t$ (i.e. one of the constituent geodesics), we have $\dot{F} = F_u \dot{u} = F_u = 0$. So F is independent of t, hence $F \equiv 0$ since it is zero at $u = 0$. □

Remark. The local coordinates arising from the parameterisation as above are called geodesic normal coordinates.

Proposition 3.14. *In a local parameterisation with $E = 1, F = 0$, we have $K = -(\sqrt{G})_{uu}/\sqrt{G}$.*

Theorema Egregium follows.

Proof. Just computation. □

As you must have felt, surfaces with constant curvature are quite significant. To analyse them, it suffices to consider the cases where K is constantly 0 or ± 1 since the dilation $(x, y, z) \mapsto (ax, ay, az), a \neq 0$ scales the Gaussian curvature by a^{-1}.

Theorem 3.15. *Suppose Σ is a smooth surface in $\mathbb{R}^3$ with constant curvature K which is either 0 or ± 1. Then:*
1. If $K = 0$, then Σ is locally isometric to the plane.
2. If $K = 1$, then Σ is locally isometric to the unit sphere.
3. If $K = -1$, then Σ is locally isometric to the tractoid.

Proof. Near $p \in \Sigma$, take the geodesic normal parameterisation $\sigma : [0, \epsilon) \times [0, \delta)$ near p, with respect to which we know the first fundamental form has the form $\mathrm{d}u^2 + G(u, v)\,\mathrm{d}v^2$. We actually also know that $G(0, v) = 1, G_u(0, v) = 0$ quite directly from the fact that fixing v always give a geodesic at unit speed.
If $K = 0$, then $(\sqrt{G})_{uu} = 0 \implies \sqrt{G} = A(v)u + B(v)$ for some A, B. THe conditions $G(0, v) = 1, G_u(0, v) = 0$ then gives $B = 1, A = 0$ and thus the first fundamental form is $\mathrm{d}u^2 + \mathrm{d}v^2$, the same as the plane.
Similarly, for $K = 1$ we have $(\sqrt{G})_{uu} + \sqrt{G} = 0$, so $\sqrt{G} = A(v)\sin u + B(v)\cos u$ and the boundary conditions give $A = 0, B = 1$ and the first fundamental form $\mathrm{d}u^2 + \cos^2(u)\,\mathrm{d}v^2$, the first fundamental form of the sphere.
With the same method, $K = -1$ gives the first fundamental form $\mathrm{d}u^2 + \cosh^2(u)\,\mathrm{d}v^2$. Not quite obvious, but it is the first fundamental form on the tractoid as parameterised by

$$\sigma(u, v) = \left(e^{-u}\cos v, e^{-u}\sin v, -\int \sqrt{1 - e^{-2u}}\,\mathrm{d}u\right)$$

which one can transform to the usual parameterisation of the tractoid by the substitution $e^{-u} = \sin\theta, \theta \in (\pi/2, \pi)$. ☐

In fact, one can transform the first fundamental form on the tractoid as parameterised as above to $(\mathrm{d}\tilde{u}^2 + \mathrm{d}\tilde{v}^2)/\tilde{v}^2$ by the change of variable $u = \log\tilde{v}, v = \tilde{u}$. Or, from the first fundamental form $\mathrm{d}u^2 + \cosh^2(u)\,\mathrm{d}v^2$, use the substitution $\tilde{u} = e^v \tanh u, \tilde{v} = e^v \operatorname{sech} u$.

Remark. The tractoid is also called the pseudosphere due to the obvious reason. It is quite sad as a surface since it is not closed nor geometrically homogeneous. This prompts us to find better ways to deal with negative curvature surfaces.

3.2 Riemannian Metric

Definition 3.5. Let $V \subset \mathbb{R}^2$ be an open set. An abstract Riemannian metric on V is a map sending $z \in V$ to a positive-definite symmetric bilinear form

$$g_z = \begin{pmatrix} E(z) & F(z) \\ F(z) & G(z) \end{pmatrix}, E > 0, G > 0, EG - F^2 > 0$$

which we almost always insist to be smooth.
On such an open set equipped with a Riemannian metric, the length of a curve $\gamma = (u, v) : [0, 1] \to V$ is

$$L(\gamma) = \int_0^1 \sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}\,\mathrm{d}t$$

Definition 3.6. Suppose Σ is an abstract smooth surface with atlas (ϕ_i, U_i) where $\phi_i : U_i \to V_i$ are the respective homeomorphisms onto open sets of $\mathbb{R}^2$. A

Riemannian metric on Σ is a choice of Riemannian metric on each V_i such that they are compatible in the sense that the transition maps are isometries, i.e.

$$D(\tilde{\sigma}^{-1} \circ \sigma)^\top \begin{pmatrix} \tilde{E} & \tilde{F} \\ \tilde{F} & \tilde{G} \end{pmatrix} D(\tilde{\sigma}^{-1} \circ \sigma) = \begin{pmatrix} E & F \\ F & G \end{pmatrix}$$

Example 3.3. Let's start with the torus with the atlas we defined before. The transition maps are just the translations on the plane. We can simply equip each V_i with the usual Euclidean flat metric (i.e. $E = G = 1, F = 0$) which preserves under translation. So the torus inherits a flat metric.

However, we have seen that every compact surface Σ in $\mathbb{R}^3$ has elliptic points.

Corollary 3.16. *The flat metric on the torus is not the induced metric from $\mathbb{R}^3$ under any of its smooth embeddings into $\mathbb{R}^3$.*

Remark. Local geometrical notions like length, area, angle, geodesics and curvature can all be expressed and computed with just the first fundamental form. By replacing the first fundamental form with a Riemannian metric, these quantities all make sense on an abstract smooth surface equipped with a Riemannian metric.

Example 3.4. The Klein bottle also admits a flat metric (constructed with the same idea we used for torus).

Remark. If Σ is a smooth surface and a finite group Γ acts freely on it by diffeomorphisms, then Σ/Γ has the natural structure of a smooth surface inherited from the quotient map. Furthermore, observe that if g is a Riemannian metric on Σ and $f : \Sigma \to \Sigma$ is a diffeomorphism, then there is a pullback metric f^*g defined by $(f^*g)|_p = g|_{f(p)}$. Then $f : (\Sigma, g) \to (\Sigma, f^*g)$ is an isometry in the sense that it preserves the local lengths of curves. This allows the quotient to inherit a natural metric.

Example 3.5. The real projective plane admits a Riemannian metric with constantly positive curvature by viewing it as the quotient of S^2 by the antipodal action.

Definition 3.7. Given a Riemannian metric g on a nonempty connected abstract smooth surface, the distance between $p, q \in \Sigma$ is

$$d_g(p, q) = \inf_\gamma L(\gamma)$$

where γ varies over all piecewise smooth paths from p to q.

Proposition 3.17. *d_g is a metric on Σ and induces the topology on Σ.*

Proof. To see d_g is a metric, what we really need to verify is that $d_g(p, q) = 0 \implies p = q$. For any $p \neq q$ in Σ, we can find a chart (ϕ, U) around p (WLOG $\phi(p) = 0$) and $\epsilon > 0$ with $q \notin V = \phi^{-1}(\overline{B(0, \epsilon)}) \subset U$. Consequently, any path from p to q must escape V. Let g be the Riemannian metric on U, then there exists $\delta > 0$ such that $g - \delta I$ is still positive definite on $\overline{B(0, \epsilon)}$, so any path γ from 0 to the $\partial B(0, \epsilon)$ has length at least $\delta\epsilon > 0$. Therefore $d_g(p, q) > 0$.
It is easy to check that this does induce the topology on Σ. □

4 Hyperbolic Surfaces

We have seen that the embedded surface in $\mathbb{R}^3$ with constant curvature is either locally isometric to S^2, $\mathbb{R}^2$ or the tractoid. The first two are nice enough, e.g. they are extremely symmetric as the respective isometry groups act transitively on the surfaces. But the tractoid looks terrible. This prompts us to study other, possibly more symmetric, models of surfaces with constant negative curvature (hyperbolic surfaces).
From now on, by circle we mean either an ordinary circle or a line.

4.1 Möbius Maps and Inversions

Recall that the Möbius group $\mathcal{M} = \mathrm{PSL}_2(\mathbb{C})$ acts transitively on $\mathbb{C}_\infty$ (and also on the set of circles in $\mathbb{C}_\infty$) by Möbius transformations. Also, for any distinct $a, b, c \in \mathbb{C} \cup \{\infty\}$, there is a unique $T \in \mathcal{M}$ such that $T(a) = 0, T(b) = 1, T(c) = \infty$. In this case, we can define $T(d) = [a, b, c, d]$ to be the cross ratio of the four points. In addition, we know that $\mathcal{M}$ is generated by $z \mapsto \lambda z, \lambda \neq 0, z \mapsto z + a, a \mapsto 1/z$.

Definition 4.1. Let $\Gamma \subset \mathbb{C}_\infty$ be a circle. We say points z, z' are inverse points for Γ if each circle orthogonal to Γ and passes through z also passes through z'.

The existence and uniqueness of an inverse point are not completely obvious.

Example 4.1. If $\Gamma = \mathbb{R}_\infty = \mathbb{R} \cup \{\infty\}$, then z, z' are inverse points for Γ iff $z' = \bar{z}$.

Lemma 4.1. *Suppose $T \in \mathcal{M}$ is a Möbius map and $\Gamma \subset \mathbb{C}_\infty$ is a circle. If z, z' are inverse for Γ, then Tz and Tz' are inverses for $T(\Gamma)$.*

Proof. Möbius maps send circles to circles and preserve angles. □

Corollary 4.2. *Inverse points always exist and are unique.*

Definition 4.2. The map $J = J_\Gamma$ mapping a point z to the inverse point of z for Γ is called the inversion in Γ.

In particular, J_Γ can be given by the conjugation of $z \mapsto \bar{z}$ by a Möbius map T sending $\mathbb{R} \cup \{\infty\}$ to Γ. This, in turn, shows that the composition of two inversions is a Möbius map.

Example 4.2. $J_{S^1}(z) = 1/\bar{z}, J_{c+rS^1}(z) = c + r^2/(\bar{z} - \bar{c})$.

Proposition 4.3. *If $T \in \mathcal{M}$ maps the unit disk D to itself, then T has the form*

$$T(z) = \frac{az + b}{\bar{b}z + \bar{a}}, |a|^2 - |b|^2 = 1$$

The collection of Möbius that preserves a region U is denoted $\mathcal{M}(U)$.

Proof. If T preserves D, then it also preseres S^1, hence pairs of inverse points for S^1. Let $J = J_{S^1}$, then $J \circ J = \mathrm{id}$, then this just means that J and T commutes. Therefore $J \circ T \circ J = J \circ (T \circ J) = J \circ (J \circ T) = (J \circ J) \circ T = T$. Expanding this gives the result. □

The exact same idea (or, alternatively, via a suitable conjugation) also allows us to find the group of Möbius maps preserving other (possibly punctured) disks in $\mathbb{C}_\infty$.

Corollary 4.4. *Suppose* $\mathfrak{h} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ *is the upper half-plane. Then* $\mathcal{M}(\mathfrak{h}) = \mathrm{PSL}_2(\mathbb{R})$.

Proposition 4.5. *There is a unique smooth metric (in the topological sense)* $d : D \times D \to \mathbb{R}_{\geq 0}$*:*
1. It is invariant under $\mathcal{M}(D)$*, i.e.* $d(T(x), T(y)) = d(x, y)$ *for any* $T \in \mathcal{M}(D)$*.*
2. Intervals on $\mathbb{R} \cap D$ *are paths of shortest length.*

What we really want to get to is that this metric is natually induced by the abstract Riemannian metric $g_{\text{hyp}} = 4(\mathrm{d}u^2 + \mathrm{d}v^2)/(1 - u^2 - v^2)^2$ of a hyperbolic surface on which $\mathbb{R} \cap D$ is a geodesic.

Proof. T acts transitively on D and also contains all rotations, therefore its value on $\{0\} \times (0, 1)$ completely determines d. Let $p : (0, 1) \to \mathbb{R}_{>0}$ be $p(a) = d(0, a)$. Distance has to be additive along length-minimising paths, so for $0 < a < b < 1$,

$$p(a) + p\left(\frac{b - a}{1 - ab}\right) = p(b)$$

Differentiating both sides wrt b and set $b = a$ gives $p'(a) = p'(0)/(1 - a^2)$ which solves to $p(a) = 2\tanh^{-1}(a)$. Hence d is uniquely determined and has the form

$$d(a, b) = 2\tanh^{-1}\left(\frac{|b - a|}{|1 - \bar{a}b|}\right)$$

It remains to show that this satisfies the triangle inequality as other axioms follows immediately. It suffices to show that $d(a, b) \leq d(a, 0) + d(0, b)$ where WLOG $a \in \mathbb{R}_+, b \in e^{i\theta}\mathbb{R}_+$. Let $\alpha = d(a, 0), \beta = d(0, b)$ and $\gamma = d(a, b)$, then the inequality shall follow immediately from the hyperbolic cosine formula

$$\cosh\gamma = (\cosh\alpha)(\cosh\beta) - (\sinh\alpha)(\sinh\beta)(\cos\theta)$$

How do we prove this? We have $a = \tanh(\alpha/2), b = e^{i\theta}\tanh(\beta/2), |b - a|/|1 - \bar{a}b| = \tanh(\gamma/2)$. Playing around gives

$$\cosh\alpha = \frac{1 + |a|^2}{1 - |a|^2}, \cosh\beta = \frac{1 + |b|^2}{1 - |b|^2}, \cosh\gamma = \frac{|1 - \bar{a}b|^2 + |b - a|^2}{|1 - \bar{a}b|^2 - |b - a|^2}$$

Expanding the last expression gives what we wanted. □

Definition 4.3. The metric as in above is called the hyperbolic metric d_{hyp}.

Remark. For small α, β, γ, we have $\sinh\alpha \sim \alpha, \cosh\alpha \sim 1 + \alpha^2/2$, so the hyperbolic cosine formula reduces to the (flat) cosine formula $\gamma^2 = \alpha^2 + \beta^2 - 2\alpha\beta\cos\theta$.

4.2 Hyperbolic Models

Observe that if $z, z + \delta z \in D$, then $d_{\text{hyp}}(z, z + \delta z) \sim 2|\delta z|/(1 - |z|^2)$ as $\delta z \to 0$.

Definition 4.4. The hyperbolic Riemannian metric on D is the abstract Riemannian metric

$$g = g_{\text{hyp}} = 4\frac{\mathrm{d}u^2 + \mathrm{d}v^2}{(1-u^2-v^2)^2} = 4\frac{|\mathrm{d}z|^2}{(1-|z|^2)^2}$$

We call (D, g_{hyp}) the hyperbolic disc.

Lemma 4.6. *g_{hyp} is $\mathcal{M}(D)$-invariant.*

Proof. $\mathcal{M}(D)$ is generated by $z \mapsto e^{i\theta}z$ and $z \mapsto (z-a)/(1-\bar{a}z)$ for $\theta \in \mathbb{R}, |a| < 1$. The former clearly preserves g_{hyp}. As for the latter, let $w = (z-a)/(1-\bar{a}z)$, then

$$\mathrm{d}w = \frac{\mathrm{d}z}{1-\bar{a}z} + \frac{z-a}{(1-\bar{a}z)^2}\bar{a}\,\mathrm{d}z = \frac{1-|a|^2}{(1-\bar{a}z)^2}\,\mathrm{d}z$$

So

$$\frac{|\mathrm{d}w|}{1-|w|^2} = \frac{(1-|a|^2)|\mathrm{d}z|}{|1-\bar{a}z|^2 - |z-a|^2} = \frac{|\mathrm{d}z|}{1-|z|^2}$$

Yay. □

Remark. Suppose $\gamma = (u,v) : [0,1] \to D$ is a smooth path from $0 \in D$ to $a \in D \cap \mathbb{R}_+$. Then

$$L(\gamma) = \int_0^1 2\frac{|\dot{\gamma}(t)|}{1-|\gamma(t)|^2}\,\mathrm{d}t = \int_0^1 2\frac{\sqrt{\dot{u}^2+\dot{v}^2}}{1-u^2-v^2}\,\mathrm{d}t \geq \int_0^1 2\frac{|\dot{u}|}{1-u^2}\,\mathrm{d}t$$

with equality iff $\dot{v} \equiv 0$. Thus,

$$L(\gamma) \geq \int_0^1 2\frac{|\dot{u}|}{1-u^2}\,\mathrm{d}t \geq \int_0^1 \frac{2\dot{u}}{1-u^2}\,\mathrm{d}t = 2\tan^{-1}(a)$$

where the second inequality holds when $\dot{u}$ is nonnegative.

Corollary 4.7. $d_{g_{\text{hyp}}} = d_{\text{hyp}}$.

There are, of course, other models of hyperbolic geometry.

Definition 4.5. The hyperbolic upper half-plane $(\mathfrak{h}, g_{\text{hyp}})$ is the set $\mathfrak{h} = \{z : \operatorname{Im} z > 0\} \subset \mathbb{C}$ with the abstract metric $g_{\text{hyp}} = (\mathrm{d}x^2 + \mathrm{d}y^2)/y^2$.

In situations where possible ambiguities might occur, we write $g_{\text{hyp}}^{\mathfrak{h}}$ to denote the hyperbolic metric on $\mathfrak{h}$ and g_{hyp}^{D} to denote the one on D.

Proposition 4.8. *$(\mathfrak{h}, g_{\text{hyp}}^{\mathfrak{h}})$ can be obtained from (D, g_{hyp}^{D}) via the transformation*

$$D \to \mathfrak{h}, z \mapsto i\frac{1+z}{1-z}$$

So the two models of hyperbolic geometry are globally isometric, so any of our discussion about one applies automatically to the other.

Proof. Write $T : \mathfrak{h} \to D$ as the inverse of that transformation, i.e. $T(w) = (w-i)/(w+i)$, then $T'(w) = 2i/(w+i)^2$, so

$$\frac{|\mathrm{d}(T(w))|}{1-|T(w)|^2} = \frac{|T'(w)||\mathrm{d}w|}{1-|T(w)|^2} = \frac{2|\mathrm{d}w|}{|w+i|^2(1-|w-i|/|w+i|)^2} = \frac{|\mathrm{d}w|}{2\operatorname{Im} w}$$

Done. □

We can prove this, alternatively, by recalling that $\mathcal{M}(\mathfrak{h}) = \mathrm{PSL}_2(\mathbb{R})$ (which is generated by $w \mapsto w + c, w \mapsto aw, w \mapsto -1/w$ for $c \in \mathbb{R}, a \in \mathbb{R} \setminus \{0\}$) under which $g_{\text{hyp}}^{\mathfrak{h}}$ is invariant.

Corollary 4.9. *Let $a, b \in \mathfrak{h}$. Then there is a unique hyperbolic geodesic between a, b which is the arc of a circle perpendicular to $\partial\mathfrak{h}$ through a, b.*

Also, $\mathcal{M}(\mathfrak{h})$ acts transitively on the set of geodesics.

Proof. $\mathcal{M}(\mathfrak{h})$ sends circles to circles and preserves orthogonality. □

Lemma 4.10. *Any orientation-preserving isometry of $\mathfrak{h}$ is an element of $\mathcal{M}(\mathfrak{h})$. The full isometry group is generated by these and the set of inversions in hyperbolic geodesics.*

Proof. The metric $(\mathrm{d}x^2 + \mathrm{d}y^2)/y^2$ is obviously invariant under $y \mapsto y, x \mapsto -x$, i.e. the inversion in the geodesic $i\mathbb{R}_+$. As $\mathcal{M}(\mathfrak{h})$ acts transitively on the set of geodesics, all inversions are isometries. So the group generated by the inversions and Möbius maps is indeed contained in the group of isometries.
To see these are all, it is easier to work with the disk model. Suppose $\alpha : D \to D$ is an isometry. Composing it with a suitable element in $\mathcal{M}(D)$, we can assume WLOG that α fixes 0 and $D \cap \mathbb{R}_+$. Reflecting in $D \cap \mathbb{R}_+$ if necessary, we can also assume WLOG that it actually fixes both $D \cap \mathbb{R}$ and $D \cap i\mathbb{R}$. But since isometries fix distances, α is in fact (well, up to reflections) identity on $D \cap (\mathbb{R} \cup i\mathbb{R})$. It is not hard to see from there that α has to be the identity on D. □

Remark. The hyperbolic plane has constant curvature -1 since the change of variables $x = e^v \tanh(u), y = e^v \operatorname{sech}(u)$ gives $(\mathrm{d}x^2 + \mathrm{d}y^2)/y^2 = \mathrm{d}u^2 + \cosh^2(u)\,\mathrm{d}v^2$.

The region $\{-\pi < x < \pi, y > 1\}$ is actually isometric to the complement of a meridian in the tractoid. So this gives a correspondence to our old model for hyperbolic geometry.
The discussion reveals that talking about hyperbolic geometry in $\mathfrak{h}$ and D are entirely equivalent. From now on, we denote our model of hyperbolic geometry (in either $\mathfrak{h}$ or D) by $\mathbb{H}^2$ and denote its set of orientation-preserving isometries by $\mathcal{M}(\mathbb{H}^2)$.

Definition 4.6. Let $\phi \in \mathcal{M}(\mathbb{H}^2)$. We say ϕ is elliptic if if has a fixed point in $\mathbb{H}^2$. It is parabolic if it fixes exactly one point on $\partial\mathbb{H}^2$. It is hyperbolic if it fixed two points on $\partial\mathbb{H}^2$.

4.3 Hyperbolic Geometry

Definition 4.7. Two geodesics in $\mathbb{H}^2$ are:
1. Parallel, if they meet in $\partial\mathbb{H}^2$ but not in $\mathbb{H}^2$.
2. Ultraparallel, if they are disjoint in $\mathbb{H}^2 \cup \partial\mathbb{H}^2$.
3. Intersecting, otherwise.

Remark. The parallel postulate completely fails in the hyperbolic plane, but the other postulates still hold.

Definition 4.8. A hyperbolic triangle is the region in $\mathbb{H}^2$ cut out by three geodesics, any pair of which either intersect or are parallel. If a vertex of a hyperbolic triangle lies on $\partial\mathbb{H}^2$, it is called ideal vertices.

Remark. The metric we had on either model of the hyperbolic plane is conformal ($E = G, F = 0$), so angles in $\mathbb{H}^2$ are just the same as in $\mathbb{R}^2$ under the inclusions $D \hookrightarrow \mathbb{R}^2, \mathfrak{h} \hookrightarrow \mathbb{R}^2$.

Proposition 4.11. *Let T be a hyperbolic triangle with internal angles α, β, γ (in particular, the internal angle at an ideal vertex is 0), then the area of T is $\pi - (\alpha + \beta + \gamma)$.*

Proof. $\mathcal{M}(\mathbb{H}^2)$ acts triply transitively on points of $\partial\mathbb{H}^2$, so there is a unique triangle with all vertices ideal up to isometry. In the upper half-plane model, we can just take it as the one with sides $\{x = -1\}, \{x = 1\}$ and $\{|x + iy| = 1\}$ which has area

$$\int_{-1}^{1} \int_{\sqrt{1-x^2}}^{\infty} \frac{1}{y^2} \, \mathrm{d}y \, \mathrm{d}x = \pi = \pi - (0 + 0 + 0)$$

If we have a triangle with angles $0, 0, \alpha$, then its area $A(\alpha)$ has to be continuous and decreasing in α. Also $A(\alpha) + A(\beta) = A(\alpha + \beta) + \pi$ by drawing a diagram. This means that $\pi - A(\alpha)$ is additive and increasing, so $\pi - A(\alpha) = \lambda\alpha$ for some constant $\lambda > 0$. So $A(\alpha) = \pi - \lambda\alpha$. But $A(\alpha) + A(\pi - \alpha) = \pi$ by drawing yet another diagram, so $\lambda = 1$. It is not hard to see the general case from here (by e.g. extending the sides). $\square$

4.4 Compact Hyperbolic Surfaces

Recall that compact connected orientable abstract smooth surfaces are classified by their genus g. For $g = 0$, we get the sphere which admits a metric with constantly positive Gaussian curvature. For $g = 1$, we obtain the torus which can inherit a flat metric from the embedding $S^1 \times S^1 \subset \mathbb{R}^4$. For $g \geq 2$, we know that

$$\int_{\Sigma} K \, \mathrm{d}A < 0$$

by Gauss-Bonnet. Naturally, it is tempting to hypothesise that they admit a constantly negative Gaussian curvature.

Theorem 4.12. *For each $g \geq 2$, the compact connected orientable abstract smooth surface Σ_g of genus g admits a metric with constant Gaussian curvature -1.*

It is not hard to guess the idea: When we find a flat metric on the torus, what we did can be alternatively described as to make use of the natural chart that maps open subsets on the torus (as an identification space descended from the unit square) to open subsets on $\mathbb{R}^2$ via translation. This hints that we might be able to do something just as similar for surfaces with $g \geq 2$, since they can be described as the identification space of a $4g$-gon.

Recall that the area of a hyperbolic triangle has area $\pi - (\alpha + \beta + \gamma)$ where α, β, γ are its internal angles. Via an obvious trianglation, we know that a hyperbolic n-gon has area

$$(n-2)\pi - \sum_{i=1}^{n} \alpha_i$$

where α_i are the internal angles of the polygon.

Lemma 4.13. *For each $g \geq 2$, there is a regular $4g$-gon in $\mathbb{H}^2$ with internal angles $2\pi/(4g) = \pi/(2g)$.*

What we want from here is that the internal angles of this polygon sum to 2π so that the identification space on it works well.

Proof. Let's play with the disk model. First consider the $4g$-gon with vertices at the $\exp(2\pi ik/(4g)), k = 0, \ldots, 4g-1$. This polygon has area $(4g-2)\pi$. Then we shrink the polygon so that at each time $t \in [0, 1)$ we get a polygon with vertices $(1-t)\exp(2\pi ik/(4g)), k = 0, \ldots, 4g-1$. This shrinking is obviously continuous and the area of the polygon goes to 0 as $t \to 1^-$, so we can get a polygon with area $(4g-2)\pi - 4g(\pi/2g) = (4g-4)\pi \in (0, (4g-2)\pi]$ at a certain time. That is the polygon we want. $\square$

It is obvious what we want to do next: Take Σ_g as the identification space of a regular $4g$-gon on the hyperbolic plane and inherit its metric.

Lemma 4.14. *On $\mathbb{H}^2$, if two sets of data comprising of:*
1. An oriented geodesic.
2. A point on the said geodesic.
3. A region to one side of the geodesic.
are given, then there exists an isometry (which is in fact unique) taking one to the other.

Proof. Exercise. $\square$

Combining the two lemmas gives a locally hyperbolic atlas on Σ_g with transition maps given by isometries on $\mathbb{H}^2$. Theorem 4.12 follows.

Remark. Another way to view the torus is to take it as the quotient $\mathbb{R}^2/\mathbb{Z}^2$. The flat metric on it can then be identified as the metric descended from the flat metric on $\mathbb{R}^2$ under the quotient. Analogously, we can get a hyperbolic metric on Σ_g by viewing it as $\mathbb{H}^2/\Gamma$ where Γ is a group generated by suitably chosen isometries that pairs the corresponding sides.

Another way to construct compact orientable hyperbolic surfaces is by what's known as the pairs-of-pants decomposition.

Definition 4.9. A right-angled hyperbolic hexagon is a hyperbolic hexagon with all internal angles equal to $\pi/2$.

Lemma 4.15. *For each triple $l_\alpha, l_\beta, l_\gamma \in \mathbb{R}_{>0}$, there is a right-angled hyperbolic hexagon $H_{\alpha\beta\gamma} \subset \mathbb{H}^2$ such that a set of alternating sides of $H_{\alpha\beta\gamma}$ has lengths α, β, γ in that order.*

Proof. Start with $t > 0$ and a pair of ultraparallel geodesics such that they are distant t apart along their unique common perpendicular. Walking along the orthogonal intersections A, B to the same direction with distance l_α, l_β respectively to reach points C, D, and then consider geodesics starting from C, D that are orthogonal to the initial two geodesics. By a continuity argument, there exists t_0 such that the new geodesics are parallel. Consider $t \in (t_0, \infty)$, then the unique common perpendicular between the new geodesics has length varying continuously from 0 to ∞. This gives l_γ by a suitable choice of t. $\square$

Definition 4.10. A pair of pants is any topological surface (with boundary) homeomorphic to S^2 removing the interiors of three disjoint closed disks.

We can, in fact, glue two hyperbolic hexagons to obtain a hyperbolic pair of pants with geodesic boundary. For $l_\alpha, l_\beta, l_\gamma \in \mathbb{R}_+$, take a hexagon as in the preceding lemma. Suppose the hexagon has side lengths $l_\alpha, t_{\alpha\beta}, l_\beta, t_{\beta\gamma}, l_\gamma, t_{\gamma\alpha}$ in that order. Then Lemma 4.14 allows us to glue two copies of this hexagon together along the sides $t_{\alpha\beta}, t_{\beta\gamma}, t_{\gamma\alpha}$ to form a hyperbolic pair of pants whose holes have perimeters $2l_\alpha, 2l_\beta, 2l_\beta$ respectively. Furthermore, the gluing also ensures that there are geodesics between the respective holes having lengths $t_{\alpha\beta}, t_{\beta\gamma}, t_{\gamma\alpha}$ respectively.
A genus $g \geq 2$ (smooth) surface can be built from gluing together (smooth) pairs of pants (draw a picture to pursuade yourself of this!), and if we use those hyperbolic pairs of pants we just constructed, then the surface inherits a hyperbolic metric and any of our choice of girth!

5 Closing Remarks

5.1 The Gauss-Bonnet Theorem and its Consequences

We have seen that a spherical triange with internal angles α, β, γ has area $\alpha + \beta + \gamma - \pi$ and a hyperbolic triangle has area $\pi - (\alpha + \beta + \gamma)$. But we can make sense of the notion of geodesic triangles on any abstract smooth surface with a Riemannian metric, so how do they behave?
For simplicity, we only consider geodesic triangles those area is bounded by one chart (i.e. can be bounded by a disc).

Theorem 5.1 (Gauss-Bonnet Theorem for Triangles)**.** *Let R be a geodesic triangle with internal angles α, β, γ, then*

$$\int_R K \, \mathrm{d}A = (\alpha + \beta + \gamma) - \pi$$

More generally, we can consider a geodesic polygon that bounds a disc in Σ, in which case we have

Theorem 5.2 (Gauss-Bonnet Theorem for Polygons)**.** *Let R be a geodesic polygon with internal angles $\{\alpha_i\}$, then*

$$\int_R K \, \mathrm{d}A = \sum_{i=1}^{n} \alpha_i - (n-2)\pi$$

This theorem, along with several other special cases of Gauss-Bonnet we've seen in example sheets and previous lectures, of course follows from the global Gauss-Bonnet theorem

Theorem 5.3 (Gauss-Bonnet)**.** *If Σ is a compact smooth surface with abstract metric g and Gaussian curvature K, then*

$$\int_\Sigma K \, \mathrm{d}A = 2\pi\chi(\Sigma)$$

Remark. This is amazing in the sense that it linked the analytic properties of a surface to the topological property of it. Ideas like this has been the heart of geometry for several centuries, other theorems (sometimes in hierachy of each other) of the same flavour includes the Chern theorem, the Riemann-Roch Theorem, the Hirzebruch Signature Theorem, and the Atiyah-Singer Index Theorem.

There is a lot of consequences of this theorem.

Corollary 5.4. *1. $\chi(\Sigma)$ does not depend on a choice of triangulation or subdivision.*
2. If Σ admits a metric with $K > 0$ everywhere, then $\chi(\Sigma) > 0$. Consequently, if Σ is orientable, then it is homeomorphic to S^2.
3. No closed geodesic on a flat surface can bound a disc.

How do you prove Gauss-Bonnet? If the surface can be subdivided by geodesic polygons, then it is easy to deduce the general Gauss-Bonnet from the polygonal version. How do we show that we can always have a subdivision? This follows from (the compactness and) the existence of local geodesics we discussed way earlier.
So how about the polygonal Gauss-Bonnet? Suppose Σ is a smooth surface, say WLOG in $\mathbb{R}^3$ (we can always immerse and perturb it) and $\Gamma \subset \Sigma$ a geodesic polygon that lies in the a geodesic normal parameterisation $\sigma : V \to \Sigma$. The idea is to reduce it to – surprise surprise – Green's theorem on the plane. We take a moving frame given by

$$e = \sigma_u, f = \frac{\sigma_v}{\sqrt{G}}$$

Let γ be the (piecewise smooth geodesic) boundary of R (pulled back to V) parameterised by arc length, we are interested in

$$I = \int_\gamma \langle e, \dot{f} \rangle \,\mathrm{d}t = \int_\gamma \langle e, f_u \rangle \,\mathrm{d}u + \langle e, f_v \rangle \,\mathrm{d}v$$

Now,

$$\langle e, f_v \rangle_u - \langle e, f_u \rangle_v = \langle e_u, f_v \rangle - \langle e_v, f_u \rangle = -(\sqrt{G})_{uu} = K\sqrt{G}$$

So indeed

$$I = \int_R K \,\mathrm{d}A$$

On the other hand, let θ be the angle between $\dot{\gamma}$ and e, then

$$\dot{\gamma}(t) = e\cos\theta + f\sin\theta, \ddot{\gamma} = \dot{e}\cos\theta + \dot{f}\sin\theta + \eta\dot{\theta}$$

where $\eta = -e\sin\theta + f\cos\theta$. Since γ is geodesic, $\langle \ddot{\gamma}, \eta \rangle = 0$, so $\dot{\theta} = \langle e, \dot{f} \rangle$. Consequently,

$$\int_R K \,\mathrm{d}A = I = \int_\gamma \langle e, \dot{f} \rangle \,\mathrm{d}t = \int_\gamma \dot{\theta}(t) \,\mathrm{d}t = 2\pi - \sum(\text{external angles of } R)$$

which yields the polygonal Gauss-Bonnet.

5.2 Moduli of Surfaces

When we built hyperbolic metrics on surfaces of genus $g \geq 2$, we had a lot of choices. For example, we are able to choose the perimeters of the holes during pair of pants decompositions. Also, when gluing each pair of holes together, we had a choice of tilting the orientation of the circles. Does our choices affect the isometric class of these surfaces? This is a very interesting topic to discuss, but let's look at a simpler case instead.
Recall that we can construct a flat metric on T^2 by taking it as the identification space of the unit square. But there is nothing special about the unit square! We can also give it a flat metric by taking it as the identification space of a long rectangle, or even any parallelograms, in an analogous way. But are they the same metric?

Lemma 5.5. *If Q is a parallelogram in $\mathbb{R}^2$ that induces a flat metric g_Q on the torus as above, then the area of the torus from g_Q equals to the area of the parallelogram in the usual Euclidean metric.*

Proof. Obvious. □

Remark. Hyperbolic surfaces of genus g, however, have their total area determined by the topology (by e.g. Gauss-Bonnet).

The lemma immediately shows that parallelograms of different areas give rise to non-isometric flat tori. But are parallelograms of the same area necessarily induce the same metric?

Lemma 5.6. *Let Q_1 be the unit square and Q_2 be the rectangle with side lengths* 10 *and* 1/10*, then g_{Q_1} and g_{Q_2} gives non-isometric tori.*

Proof. Look at the shortest possible length of a closed geodesic (which can be pulled back to straight lines of rational slope on $\mathbb{R}^2$). □

So, indeed, there is a lot of freedom when we try to find a parallelogram to induce a flat metric on the torus. Of course, shifting and rotating a parallelogram does not change the metric in any way, an interesting class of parallelogram we shall pay our attention to would be the one parameter family $Q_\tau = \{1, \tau, 1+\tau, 0\}$ where τ is on the upper half-plane $\mathfrak{h} = \{\tau \in \mathbb{C} : \operatorname{Im} \tau > 0\}$. Denote the flat metric induced by it by g_τ.
Recall that if $f : \Sigma \to \Sigma$ is a diffeomorphism of an abstract smooth surface and g is an abstract Riemannian metric on Σ, then we can canonically obtain a pullback metric f^*g such that f is an isometry. Explicitly, this is given by $\langle v, w \rangle_{f^*g} = \langle Df(v), Df(w) \rangle_g$. If we take the torus to be the identification space of the unit square (i.e. $g = g_i$), then $\mathrm{SL}_2(\mathbb{Z})$ acts on the torus by smooth invertible diffeomorphisms. Pulling back the original metric by one of those diffeomorphisms gives another flat metric on the torus, which is not exactly the one we started with but nonetheless is isometric to it. Which quadrilateral induce that metric? Well, take $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, then the pullback (flat) metric is exactly g_{1+i}. Generalising this, we can conclude that $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathfrak{h}$ without changing the isometric class of the induced flat torus. So the map from $\mathfrak{h}$ to flat metrics on the torus via $\tau \mapsto g_\tau$ induces a map ϕ from $\mathfrak{h}/\mathrm{SL}_2(\mathbb{Z})$ to the quotient of the set of flat metrics on torus by pullback actions by diffeomorphisms. The former is, of course, naturally an object of hyperbolic geometry.

Theorem 5.7. *ϕ is, in fact, a correspondence.*

We say $\mathfrak{h}/\operatorname{SL}_2(\mathbb{Z})$ is the moduli space of flat metrics on the torus.

Remark. We say $A \in \operatorname{SL}_2(\mathbb{R})$ is elliptic if $|\operatorname{tr} A| < 2$, i.e. it fixes a point of $\mathfrak{h}$. If $A \in \operatorname{SL}_2(\mathbb{Z})$, then this means that A has finite order.
We say A is parabolic if $|\operatorname{tr} A| = 2$, i.e. it fixes a point on $\partial\mathfrak{h} = \mathbb{R} \cup \{\infty\}$.
A is hyperbolic if $|\operatorname{tr} A| > 2$, i.e. it fixes two points on $\partial\mathfrak{h}$.

The hyperbolic triangle with vertices ρ, ρ^2, ∞ (where $\rho = \exp(\pi i/3)$), as one can verify, is a fundamental domain (after removing certain parts of its boundary) for $\mathfrak{h}/\operatorname{SL}_2(\mathbb{Z})$.