Geometry

Jarek Kędra

March 27, 2018

Contents

1 Introduction

Geometry is the study of metric spaces. There are many metric spaces around us: the surface of the table, the surfaces of walls and ceilings in the room we sit in, the surface of our bodies, the surface of the bonnet of a car, the surface of the Earth, the space and the space-time around us etc. These are examples of metric spaces we can see and/or we can touch. This course aims at presenting the part of geometry which is related and applicable to such spaces.

The surfaces around us can be approximated by surfaces built from rigid convex polygons. In the first part of the course we will deal with such ***polygonal surfaces*** while in the second part we will study smooth surfaces in the three-dimensional space. In the process we will also investigate more general and more abstract examples. We will generalise many geometric concepts we know from Euclidean geometry to the world of reasonably general metric spaces. These include the concept of a straight line, an angle, a triangle, the number π and others.

One of the guiding concepts for this course is *curvature*. We intuitively know whether a surface is curved either by looking at it or by touching it. This, however, can be misleading. For example, we will learn that the surface of a cylinder is *intrinsically* flat, so what we really see and feel is that it is only curved in the three-dimensional space. We will make the concept of the curvature mathematically precise and amenable to computations. We will also learn about the intimate connection between the curvature and topology.

Our spatial intuition and imagination play an important role in studying geometry. They are mostly related to two of our senses: the sight and the touch. However, a few millions years ago we reduced our quite involved three-dimensional intuition to nearly two-dimensional Euclidean one. Although it was highly beneficial from the general evolutionary point of view, we lost a lot. Just see and admire apes moving gracefully in the trees or tigers who *never get lost* and compare it with us getting lost in the mountains or cities not to mention even slightly complicated corridors.

Can we do something to regain our geometric intuition? One of the greatest geometers of all times, Bill Thurston, used to practice "visualising things" every day. Let me quote Benson Farb, a student of Thurston: *Bill was probably the best geometric thinker in the history of mathematics. Thus it came as a surprise when I found out that he had no stereoscopic vision, that* *is, no depth perception. Perhaps the latter was responsible somehow for the former? I once mentioned this theory to Bill. He disagreed with it, claiming that all of his skill arose from his decision, apparently as a first grader, to "practice visualizing things" every day.* – from ***On being Thurstonized***.

2 Recollection on metric spaces

Definition 2.1. A ***metric*** or a ***distance*** on a set X is a function

$$\mathrm{d}\colon X \times X \to \mathbf{R}$$

such that it satisfies each of the following conditions for all points $x, y, z \in X$:
1. $\mathrm{d}(x, y) \geq 0$ and $\mathrm{d}(x, y) = 0$ if and only if $x = y$;
2. $\mathrm{d}(x, y) + \mathrm{d}(y, z) \geq \mathrm{d}(x, z)$;
3. $\mathrm{d}(x, y) = \mathrm{d}(y, x)$.

The pair (X, d) is called a ***metric space***.

Remark 2.2. This definition reflects our intuitive understanding of a distance: it is always non-negative, and positive between distinct places (1), it satisfies the triangle inequality (2) and it is symmetric (3). Observe, however, that in practical situations distances are not always symmetric. For example, a distance in a town with one way roads, when we are driving a car.

2.3 L^p-metrics

Example 2.4. Let $\mathbf{R}^2$ denote the set of all pairs of real numbers. The function defined by

$$\mathrm{d}((x_1, x_2), (y_1, y_2)) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

is a metric which is easy to check. It is called the ***Euclidean*** metric on the plane. We imagine this metric space as a flat plane. ♣

Example 2.5. The function $\mathrm{d}_p\colon \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ defined by

$$\mathrm{d}_p((x_1, x_2), (y_1, y_2)) = \sqrt[p]{|x_1 - y_1|^p + |x_2 - y_2|^p}$$

is a metric which is (not so) easy to check. It is called the L^p-metric on the plane. Observe that d_2 is the Euclidean metric. These metrics have an obvious generalisation to spaces $\mathbf{R}^n$ for all natural n. Notice that on the set $\mathbf{R}$ of real numbers they are all equal. ♣

Example 2.6. The function $\mathrm{d}_\infty\colon \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ defined by

$$\mathrm{d}_\infty((x_1, x_2), (y_1, y_2)) = \max\{x_1 - y_1, x_2 - y_2\}$$

is a metric which is also easy to check. It is called the L^∞-metric on the plane. ♣

Definition 2.7. Let (X, d) be a metric space, let $x \in X$, and let $r \in \mathbf{R}$. The subset

$$B(r, x) := \{y \in X \mid \mathrm{d}(x, y) \leq r\}$$

is called the ***closed ball of radius*** r ***centered at*** x. The subset

$$S(r, x) := \{y \in X \mid \mathrm{d}(x, y) = r\}$$

is called the ***sphere of radius*** r ***centered at*** x.

2.8 Induced and intrinsic metrics

Definition 2.9. Let (X, d) be a metric space and let $Y \subset X$ be a subset. The ***induced metric*** d_Y on Y is defined by $\mathrm{d}_Y(x, y) := \mathrm{d}(x, y)$ for all $x, y \in Y$.

Example 2.10. Let $X = \mathbf{R}^2 - B(1, (0, 0))$ be the complement of a closed disc on the plane, a plane with a ***hole***. Suppose it is equipped with a metric induced from the Euclidean metric on the plane. Then the distance between $A = (-2, 0)$ and $B = (2, 0)$ is equal to four. However, our intuition is that in order to get from A to B we have to avoid the hole so the distance should be bigger. ♣

Example 2.11. Let $S^1 \subset \mathbf{R}^2$ be the unit circle on the plane. The distance between points x, y is defined to be the length of the shortest arc from x to y. Concretely, if $x = (\cos\alpha, \sin\alpha)$ and $y = (\cos\beta, \sin\beta)$, where $\alpha, \beta \in [0, 2\pi)$ then

$$\mathrm{d}(x, y) = \min\{|\alpha - \beta|, 2\pi - |\alpha - \beta|\}.$$

This is called the ***intrinsic*** metric on the circle.

We also have the metric induced from the plane for which the distance between x and y is the length of the line segment between them. That is

$$\mathrm{d}_{\mathsf{ind}}(x, y) = \sqrt{(\cos\alpha - \cos\beta)^2 + (\sin\alpha - \sin\beta)^2}.$$

♣

Example 2.12. Let $\mathbf{S}^2 = \{(x, y, z) \in \mathbf{R}^2 \,|\, x^2 + y^2 + z^2 = 1\}$ be the unit sphere in three dimensional space. The natural intrinsic metric is defined as follows. The distance between two points is the length of the shortest arc joining them and contained in the (unique) great circle through these points. This and the previous example generalise to higher dimensions. ♣

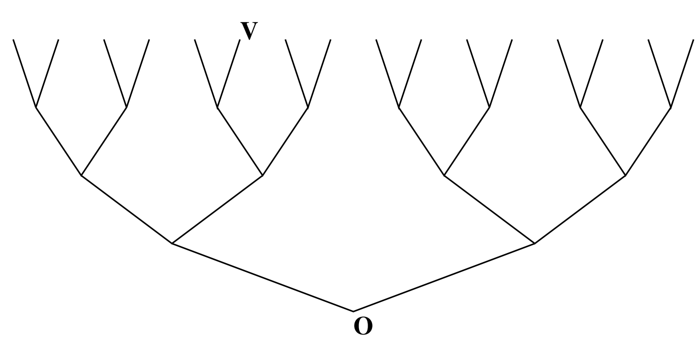
Convention: When there is no metric mentioned, the sphere $\mathbf{S}^n$ is always meant to be equipped with its intrinsic metric. Similarly, $\mathbf{R}^n$ is meant to be endowed with the Euclidean metric.

Example 2.13. Let X be a graph and we define a metric d on X by saying that every edge of X is isometric to the interval $[0, 1]$ with its natural metric induced from $\mathbf{R}$. Let's consider a few concrete examples:

1. Let X be the union of vertices and edges of a regular tetrahedron. The distance between any two distinct vertices is equal to one.
2. Let X be the union of vertices and edges of a cube. Then the distance between any two distinct vertices is equal eiter to one, two, or three.
3. Let X be the union of vertical and horizontal lines on the plane passing through point with integer coefficients (infinite squared paper). The distance from (k, l) to (m, n) for $k, l, m, n \in \mathbf{Z}$ is equal to $|k-m|+|l-n|$.
4. Let X be an infinite rooted tree pictured below.

   Observe that every vertex can be uniquely described as a string of letters L and R, or O if it is the origin. Namely, to get from the origin to a vertex we follow the unique path and list the left and right turns. For example, the vertex $\mathbf{v}$ is denoted by $LRLR$. The distance between $LRLRLLLLLRRR$ and $LRLRRRRLL$ is equal to thirteen. In general, in order to compute the distance we cross out the same beginings from both strings and sum up the amout of remaining letters.

♣



Rooted tree

2.14 Maps of metric spaces

Definition 2.15. A map $f\colon (X, \mathrm{d}_X) \to (Y, \mathrm{d}_Y)$ is called ***Lipschitz*** if there exists a number $C > 0$ such that

$$\mathrm{d}_Y(f(x_1), f(x_2)) \leq C\,\mathrm{d}_X(x_1, x_2)$$

holds for every $x_1, x_2 \in X$. The number C is called a ***Lipschitz constant*** of f.

Example 2.16. The inclusion $\iota\colon \mathbf{S}^1 \to \mathbf{R}^2$ of the circle in the Euclidean plane is Lipschitz with respect to the intrinsic metric on the circle. The Lipschitz constant of ι is equal to $2/\pi$. ♣

Definition 2.17. A map $f\colon (X, \mathrm{d}_X) \to (Y, \mathrm{d}_Y)$ is called an ***isometric embedding*** if

$$\mathrm{d}_X(x_1, x_2) = \mathrm{d}_Y(f(x_1), f(x_2))$$

holds for every $x_1, x_2 \in X$. If an isometric embedding f has an inverse $g\colon Y \to X$ which is also an isometric embedding then f is called an ***isometry***.

Example 2.18. The inclusion of a subset $A \subset X$ is an isometric embedding with respect to the metric induced from X.

EXERCISES

1. Read carefully all examples and provide more details and make relevant drawings if possible. For each example draw a few balls of various radii centered at various points.

   Make a list of things you don't understand and a list of mistakes and errors on this section.

2. Draw a circle of radius one centered at the origin of $(\mathbf{R}^2, \mathrm{d}_p)$ for $p \in \{1, 2, 3, 4, \infty\}$.
3. Draw two balls of radius one centered at the north pole of the unit sphere $\mathbf{S}^2$. The first with respect to the metric induced from the Euclidean metric on $\mathbf{R}^3$ and the second with respect to the intrinsic metric.
4. For which r the unit sphere $\mathbf{S}^1 \in \mathbf{R}^2$ endowed with the intrinsic metric is isometric to the circle of radius r with the induced metric?
5. Prove that the set of all isometries of a metric space is a group with respect to the composition.
6. For each example in this section give an example of a nontrivial isometry of finite (infinite) order in the group of isometries, provided such an isometry exists.
7. Prove that $f : \mathbf{R} \to \mathbf{R}^2$ given by $f(t) = x + t\mathbf{v}$ is an isometric embedding, provided $\mathbf{v}$ is a vector with (the Euclidean) norm equal to one and $x \in \mathbf{R}^2$ is a point.
8. Give an explicit formula for the intrinsic metric on the unit sphere $\mathbf{S}^2 \subset \mathbf{R}^3$. Generalise your solution to $\mathbf{S}^n \in \mathbf{R}^{n+1}$.
9. Prove that $f : \mathbf{S}^1 \to \mathbf{S}^2$ given by $f(\cos t, \sin t) = (\cos t, \sin t, 0)$ is an isometric embedding. Prove that the inclusion of any great circle is an isometric embedding.
10. Show that the inclusion $\iota : \mathbf{S}^1 \to \mathbf{R}^2$ is not an isometric embedding. Show that it is Lipschitz and compute the Lipschitz constant.

3 Geodesics

3.1 The length of a path

Definition 3.2. Let (X, d) be a metric space and let $[a, b] \in \mathbf{R}$ be an interval endowed with metric induced from the standard metric on the line $\mathbf{R}$. A continuous function

$$c \colon [a, b] \to X$$

is called a ***path*** or a ***curve***. If $c_1 \colon [a_1, b_1] \to X$ and $c_2 \colon [a_2, b_2] \to X$ are two paths then their ***concatenation*** is a path $c \colon [a_1, b_1 + b_2 - a_2] \to X$ defined by

$$c(t) := \begin{cases} c_1(t) & \text{if } t \in [a_1, b_1] \\ c_2(t + a_2 - b_1) & \text{if } t \in [b_1, b_1 + b_2 - a_2] \end{cases}$$

The concatenation of more than two paths is defined inductively.

Definition 3.3. Let $c \colon [a, b] \to (X, \mathrm{d})$ be a path. Its ***length*** is defined to be

$$\mathcal{L}(c) := \sup_{a=t_0<t_1<\cdots<t_n=b} \quad \sum_{i=1}^{n} \mathrm{d}(c(t_{i-1}), c(t_i)),$$

where the supremum is taken over all partitions (possibly infinite) of the interval $[a, b]$. A path c is called ***rectifiable*** if its length is finite.

Example 3.4. Let $c \colon [0, b] \to \mathbf{R}^n$ be a straight interval. That is, $c(t) = x + tv$, where $x \in \mathbf{R}^n$ is a point and v is a vector attached at x. Then $\mathcal{L}(c) = \mathrm{d}_2(x, y)$, where $y = x + bv$. Indeed, any partition of $[0, b]$ gives the same sum in the above definition. ♣

Definition 3.5. A metric space (X, d) is called a ***length space*** if the distance between any two points $x, y \in X$ is equal to the infimum of the lenghts of rectifiable paths joining them. The metric d of a length space is called the ***length metric***.

Example 3.6. The plane with a hole (equipped with the induced metric) is not a length space. The plane $\mathbf{R}^2$ with the metric d_p is a length space. Observe that a segment of a Euclidean straight line is a path of length equal to the distance between its endpoints. ♣

3.7 Geodesics

Definition 3.8. Let (X, d) be a metric space. A path $\gamma\colon [a,b] \to X$ is called a ***geodesic*** if for every $c \in [a,b]$ there exists an $\epsilon > 0$ such that for every $s,t \in (c-\epsilon, c+\epsilon)$ we have that $\mathrm{d}(\gamma(s),\gamma(t)) = |t-s|$.

Remark 3.9. If γ is a geodesic then $\mathcal{L}(\gamma|_{[s,t]}) = \mathrm{d}(\gamma(s),\gamma(t))$ for s,t as in the above definition. It means that a geodesic locally minimises the distance or that locally a geodesic is the shortest path between its endpoints. It is sometimes useful to imagine that the light in a metric space X travels along geodesics.

Example 3.10. A segment of a straight line in the Euclidean $\mathbf{R}^n$ is a geodesic. More precisely, a path $\gamma\colon [0,a] \to \mathbf{R}^n$ given by the formula $\gamma(t) = x + tv$, where $x, v \in \mathbf{R}^n$, is a geodesic. Notice that $\gamma'\colon [0,a/2] \to \mathbf{R}^n$ defined by $\gamma'(t) = x + 2tv$ has the same image as γ (which is a line segment) but it is not a geodesic. ♣

Remark 3.11. We should not confuse a path with its image. Intuitively speaking, we should imagine a path $\gamma\colon [a,b] \to X$ as a moving point in X where $t \in [a,b]$ can be thought of as time. Then the condition $\mathrm{d}(\gamma(s),\gamma(t)) = |t-s|$ tells us that a point is moving with the speed one.

Example 3.12. Let $\mathbf{S}^1$ be the unit circle with its intrinsic metric. A path $\gamma\colon [0,3] \to \mathbf{S}^1$ defined by $\gamma(t) = (\cos t, \sin t)$ is a geodesic. Observe, however, that the distance between $\gamma(0)$ and $\gamma(3)$ is not equal to three but to $2\pi - 3$. This example shows that, in general, a geodesic does not minimise the distance globally. ♣

Definition 3.13. Let X be a metric space. If for every two points $x, y \in X$ there exists a geodesic from x to y then X is called a ***geodesic*** metric space.

Example 3.14.

1. The plane with a hole is not a geodesic metric space.
2. Let $X \subset \mathbf{R}^n$ be a subset of the Euclidean space equipped with the induced metric. It is a geodesic metric space if and only if it is convex.
3. The circle with its intrinsic metric is a geodesic metric space.
4. A connected graph is a geodesic metric space.

♣

3.15 The angle between geodesics

Recall that if $\Delta \subset \mathbf{R}^2$ is a triangle on the Euclidean plane then the ***law of cosines*** states that

$$c^2 = a^2 + b^2 - 2ab\cos(\gamma),$$

where a, b, c are lengths of the sides of Δ and γ is the angle oposite side c. Thus if we know the lengths of the sides of an Eulidean triangle then we can compute the angle by

$$\gamma = \arccos \frac{a^2 + b^2 - c^2}{2ab}.$$

Definition 3.16. Let (X, d) be a metric space and let $\gamma, \gamma' \colon [0, a] \to X$ be two geodesics from a point p. That is, $\gamma(0) = \gamma'(0) = p$. If the limit

$$\lim_{t,s \to 0} \arccos \left(\frac{s^2 + t^2 - \mathrm{d}(\gamma(s), \gamma'(t))^2}{2st} \right)$$

exists then it is called the ***angle*** between γ and γ' at p and it is denoted by $\angle_p(\gamma, \gamma')$ and it is a number in $[0, \pi]$.

Example 3.17. The angle between two geodesics in the Euclidean space issuing from the same point is equal to their Euclidean angle. ♣

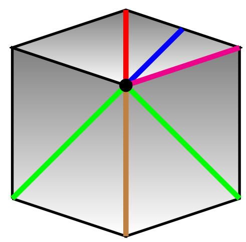
Lemma 3.18. *If $\gamma \colon [-1, 1] \to X$ is a geodesic then the angle at the midpoint $\gamma(0)$ between two parts of γ is equal to π.*

Proof. Let $\gamma_1, \gamma_2 \colon [0, 1] \to X$ be defined by $\gamma_1(t) = \gamma(t)$ and $\gamma_2(t) = \gamma(-t)$. It follows from definition of a geodesic, that there exists an $\epsilon > 0$ such that for every $s, t \in (-\epsilon, \epsilon)$ we have that $\mathrm{d}(\gamma(s), \gamma(t) = |s - t|$. Hence we have that $\mathrm{d}(\gamma_1(s), \gamma_2(t)) = \mathrm{d}(\gamma(s), \gamma(-t)) = |s - (-t)| = |s + t|$. Consequently,

$$\begin{aligned}
\angle(\gamma_1, \gamma_2) &= \arccos \left(\frac{s^2 + t^2 - \mathrm{d}(\gamma_1(s), \gamma_2(t))^2}{2st} \right) \\
&= \arccos \left(\frac{s^2 + t^2 - |s + t|^2}{2st} \right) \\
&= \arccos \left(\frac{s^2 + t^2 - s^2 - 2st - t^2}{2st} \right) \\
&= \arccos(-1) = \pi,
\end{aligned}$$

which proves the statement. □

Example 3.19. Let X be the surface of a cube and let $\gamma\colon [0,1] \to X$ be a (red) geodesic contained in a diagonal of the top face and such that $\gamma(0) = v$ is a vertex of X. Let $\gamma'\colon [0,1] \to X$ be a geodesic such that $\gamma'(0) = \gamma(0)$. Let us compute the angle between γ and γ' at v. The following figure shows several possibilities. Let us look at the angles:



Alexandrov angles

1. $\angle$(red,blue)= $\arccos(\frac{3}{\sqrt{10}}) \cong 0.32175$.
2. $\angle$(red,magenta)= $\pi/4$.
3. $\angle$(red,green)= $\pi/2$.
4. $\angle$(red,brown)= $3\pi/4$.

We figure out these angles without making any computations. We just unfold the neighbouring faces to make them flat and we look at the Euclidean angle. We see that the maximal possible angle is $3\pi/4 < \pi$. Consequently, no path passing through a vertex of a cube can be a geodesic, according to Lemma 3.18. Hikers know it very well: there is always a shortcut avoiding the summit. ♣

Example 3.20 (Slit plane)**.** Let $X = \mathbf{R}^2 \setminus D$, where $D = \{(x,y) \in \mathbf{R}^2 \mid x = y,\ x > 0\}$, be equipped with the length metric d. Let $\gamma_1, \gamma_2\colon [0,\infty) \to X$ be geodesics rays defined by $\gamma_1(t) = (t,0)$ and $\gamma_2(t) = (0,t)$. These are just geodesic parametrising the rays of coordinate axes. Observe that the distance

$\mathrm{d}\left((s,0),(0,t)\right) = s+t$. This implies that

$$\begin{aligned}\angle(\gamma_1,\gamma_2) &= \arccos\left(\frac{s^2+t^2-\mathrm{d}(\gamma_1(s),\gamma_2(t))^2}{2st}\right)\\ &= \arccos\left(\frac{s^2+t^2-|s+t|^2}{2st}\right)\\ &= \arccos\left(\frac{s^2+t^2-s^2-2st-t^2}{2st}\right)\\ &= \arccos(-1) = \pi.\end{aligned}$$

In simple words this example says that the angle between the positive rays of coordinate axes on the slit plane is equal to π. ♢

Proposition 3.21. *Let (X,d) be a metric space and let $\gamma_1,\gamma_2,\gamma_3\colon [0,1]\to X$ be geodesics issuing from the same point $p\in X$. Then*

$$\angle(\gamma_1,\gamma_3) \le \angle(\gamma_1,\gamma_2)+\angle(\gamma_2,\gamma_3),$$

provided that the angles are defined.

Proof. Suppose that the statement is not true. Then there exists a $\delta>0$ such that

$$\angle(\gamma_1,\gamma_3) > \angle(\gamma_1,\gamma_2)+\angle(\gamma_2,\gamma_3)+3\delta.$$

Let $f_{ij}(s,t) = \arccos\left(\frac{s^2+t^2-d(\gamma_i(s),\gamma_j(t))^2}{2st}\right)$ for $i,j\in\{1,2,3\}$. Since the angle $\angle(\gamma_i,\gamma_j)$ is the limit of f_{ij} as $s,t\to 0$, there exists $\epsilon>0$ such that

$$\begin{aligned}f_{12}(s,t) &< \angle(\gamma_1,\gamma_2)+\delta\\ f_{23}(s,t) &< \angle(\gamma_2,\gamma_3)+\delta\\ f_{13}(s,t) &> \angle(\gamma_1,\gamma_3)-\delta>0\end{aligned}$$

for all $0<s,t<\epsilon$. Let $\alpha>0$ be such that $f_{13}(t_1,t_3)>\alpha>\angle(\gamma_1,\gamma_3)-\delta$ for some $0<t_1,t_3<\epsilon$. Let $x_1,x_3\in\mathbf{R}^2$ be such that $d_{\mathbf{R}^2}(0,x_1)=t_1$, $d_{\mathbf{R}^2}(0,x_3)=t_3$ and the Euclidean angle between the segments $[0,x_1]$ and $[0,x_3]$ is equal to α. Observe that the inequality $f_{13}(t_1,t_3)>\alpha$ implies that

$$d_{\mathbf{R}^2}(x_1,x_3) < d(\gamma_1(t_1),\gamma_3(t_3)).$$

On the other hand, the inequality $\alpha>\angle(\gamma_1,\gamma_3)-\delta$ yields

$$\alpha > \angle(\gamma_1,\gamma_2)+\angle(\gamma_2,\gamma_3)+2\delta$$

which implies that there exists $x_2 \in [x_1, x_3]$ such that the angle α_1 between $[0, x_1]$ and $[0, x_2]$ is bigger than $\angle(\gamma_1, \gamma_2) + \delta$ (respectively the angle α_3 between $[0, x_2]$ and $[0, x_3]$ is bigger than $\angle(\gamma_2, \gamma_3) + \delta$). It follows that $d(\gamma_1(t_1), \gamma_2(t_2)) < d_{\mathbf{R}^2}(x_1, x_2)$ and $d(\gamma_2(t_2), \gamma_3(t_3)) < d_{\mathbf{R}^2}(x_2, x_3)$, where $t_2 = d_{\mathbf{R}^2}(0, x_2) < \epsilon$.

Finally we get that

$$\begin{aligned} d(\gamma_1(t_1), \gamma_3(t_3)) &> d_{\mathbf{R}^2}(x_1, x_3) = d_{\mathbf{R}^2}(x_1, x_2) + d_{\mathbf{R}^2}(x_2, x_3) \\ &> d(\gamma_1(t_1), \gamma_2(t_2)) + d(\gamma_2(t_2), \gamma_3(t_3)) \end{aligned}$$

which violates the triangle inequality in (X, d). This contradiction finishes the proof. $\square$

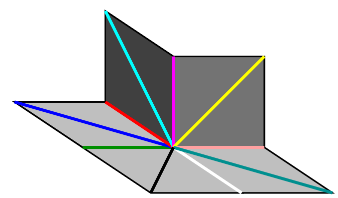
EXERCISES

1. Read carefully all examples and provide more details and make relevant drawings if possible.

   Make a list of things you don't understand and a list of mistakes and errors in this section.

2. Let $\mathbf{R}^2$ be the Euclidean plane and let $\gamma\colon \mathbf{R} \to \mathbf{R}^2$ be defined by $\gamma(t) = xt + y$, where $x = (x_1, x_2), y = (y_1, y_2) \in \mathbf{R}^2$. Prove that γ is a geodesic if and only if $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2} = 1$. State and prove the analogous criterion for the d_p metric on $\mathbf{R}^2$.

3. Consider the space pictured right and compute the angles of between the black geodesic and every other geodesic drawn. Which concatenations are geodesics?

   Compute the distance between any two vertices.

4. What is the circumference of a circle of radius r centered at the interior vertex of the space from the previous part.

5. Let γ_i for $i = 1, 2, 3$ be geodesics starting from the same point of a metric space. Suppose that the angle $\angle(\gamma_1, \gamma_2) = \angle(\gamma_2, \gamma_3) = 3\pi/4$. What can you say about the angle $\angle(\gamma_3, \gamma_1)$?

6. Draw all geodesics between diagonal vertices on the surface of the unit cube.

7. Draw balls of radii ranging between zero and three and centered at a vertex or at the centre of a face of the surfac eof the unit cube. Compute the circumference in several cases.

8. Let X be the lateral surface of a right cone of directrix d and unit generatrix. Let p and q be points in the base circle. Draw a geodesic between p and q and compute the distance between them.

9. Let $X = \mathbf{S}^1 \times \mathbf{R} \subset \mathbf{R}^3$ be the infinite cylinder based on the unit circle. Give a formula for the intrinsic metric and investigate geodesics.

10. State and prove the converse of Lemma 3.18.

11. Let $0 = t_0 < t_1 < \ldots < t_n < \ldots$ be an infinite sequence of numbers converging to 1. Let $X = [-1, 1]$ and let $c\colon [0, 1] \to X$ be any path

stisfying the following conditions: $c(0) = 0$ and $c(t_n) = \sum_{k=1}^{n}(-1)^{k+1}/k$. Show that c is not rectifiable.

Hint: Its length is bounded below by the sum of harmonic series.

12. The plane $\mathbf{R}^2$ equipped with the d_1 metric is called the ***Manhattan plane***. Recall that, $\mathrm{d}_1(x, y) = \mathrm{d}_1\left((x_1, x_2), (y_1, y_2)\right) = |x_1 - y_1| + |x_2 - y_2|$, where $x = (x_1, x_2)$ and $y = (y_1, y_2)$ are points on the plane $\mathbf{R}^2$. Moreover, let $\|x\|_1 = \mathrm{d}_2(x, 0) = |x_1| + |x_2|$.
    (a) Let $\gamma\colon \mathbf{R} \to \mathbf{R}^2$ be given by $\gamma(t) = xt + y$. Show that γ is a geodesic if and only if $\|x\|_1 = 1$.
    (b) Deduce from the previous item that the Manhattan plane is a geodesic metric space.
    (c) Let $\gamma, \gamma'\colon [0, \infty) \to \mathbf{R}^2$ be defined by $\gamma(t) = (t, 0)$ and $\gamma'(t) = (at, bt)$, where $a, b \geq 0$ and $a + b = 1$. Show that the angle between γ and γ' does not exist in general (for example, when $a = b = \frac{1}{2}$).
    (d) The ***upper angle*** between geodesics γ and γ' is defined similarly to the angle with the limit replaced by *limit superior*. Compute the upper angle between geodesics from the previous item.
    (e) How many geodesics are there between $(0, 0)$ and $(1, 1)$? How many of them are distance minimising?
    (f) How many geodesics are there between $(0, 0)$ and $(1, 0)$? How many of them are distance minimising?

4 Polygonal complexes

Definition 4.1. A ***polygonal complex*** X is a union of convex polygons (called faces of X)

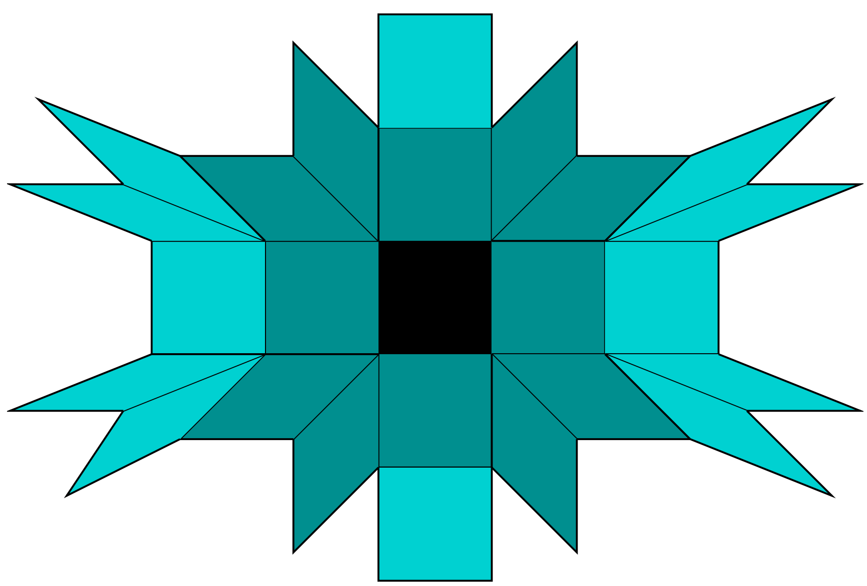
$$X = \bigcup_{i \in I} P_i$$

such that if $i \neq j$ then the intersection $P_i \cap P_j$ is either empty or it is a common edge or it is a common vertex of P_i and P_j.

Example 4.2. The surface of the cube is a polygonal complex. It has six faces, each being a square. Similarly the surface of any regular solid in the three dimensional space is a polygonal complex. ♣

Example 4.3. The plane can be subdivided (or tiled) into infinitely many squares such that exactly four meet at any vertex. This is also a polygonal complex. Similarly the plane can be regularly tiled by equilateral triangles or by regular hexagons. ♣

Example 4.4. Let X be a square complex (i.e. a complex in which all faces are squares) such that each vertex belongs to exactly five squares.



Square complex with five squares at each vertex

Let us try to understand this complex briefly by constructing it step by step. Start with a square (black) and attach the next generation of twelve squares (dark cyan); imagine a street intersection with four tower blocks around it. The next generation (light cyan, only few of them drawn) will consist of sixty squares. The physical construction would be very difficult to continue due to exponentially growing number of squares. This complex provides a good intuition of what does the hyperbolic plane look like. ♣

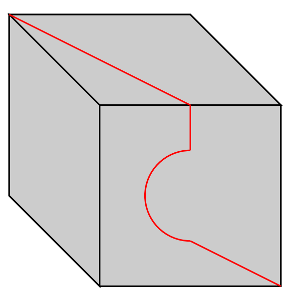
4.5 Metrics on polygonal complexes

The purpose of this section is to define a natural metric on a polygonal complex X. Let $c\colon [a,b] \to X$ be a path for which there exist numbers $a = a_0 < a_1 < a_2 < \cdots < a_n = b$ such that each restriction $c|_{[a_i,a_{i+1}]}$ is a rectifiable path in a face of X, where each face is equipped with the Euclidean metric. We define the length of c by

$$\mathcal{L}(c) = \sum_{i=1}^{n} \mathcal{L}(c|_{[a_{i-1},a_i]}),$$

where the lengths on the right hand side are computed with respect to the Euclidean metric on the faces on X. In simple words, we compute the Euclidean lengths of pieces of the path c contained entirely in the faces of X and take the sum.

Example 4.6. Suppose that the cube below has edges of length four. Then the length of the red curve is equal to $\sqrt{20} + 1 + \pi + \sqrt{5} = 1 + 3\sqrt{5} + \pi$. ♣



Length of a curve on a cube.

Definition 4.7. Let X be a polygonal complex. We define a metric d on X by

$$\mathrm{d}(x, y) = \inf_c \mathcal{L}(c),$$

where the infimum is taken over all paths $c\colon [a, b] \to X$ from x to y.

Remark 4.8. The polygonal metric d defined above is the maximal metric on X such that the inclusion of each face is an isometric embedding.

4.9 The Euler characteristic

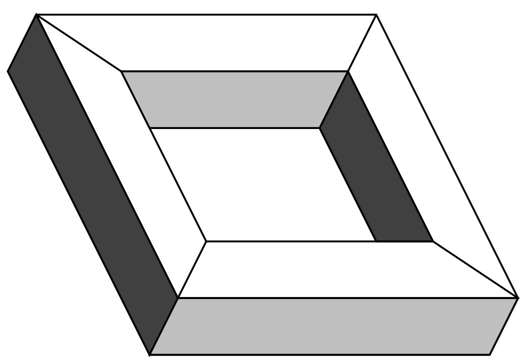
Definition 4.10. Let X be a finite polygonal complex or a finite graph. The number

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

where V is the number of vertices, E the number of edges, and F the number of faces, is called the ***Euler characteristic*** of X.

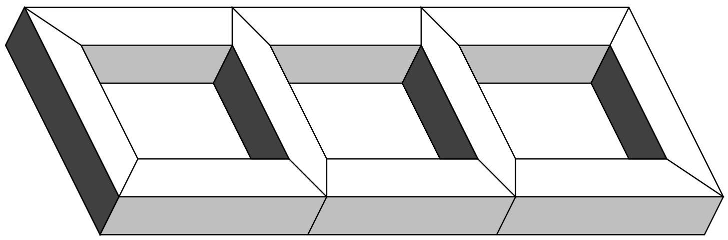
Example 4.11. The Euler characteristic of a regular tetrahedron, a cube, a regular octahedron, a regular dodecahedron and a regular icosahedron is equal to two. ♣

Example 4.12. Let X be a polygonal complex in pictured below. Its faces are trapezia and there are sixteen of them. Its Euler characteristic is equal to zero. ♣



Euler characteristic zero.

Example 4.13. The following figure shows a polygonal complex with Euler characteristic equal to -4. ♣

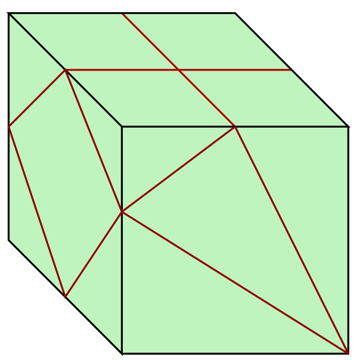


Euler characteristic -4.

Definition 4.14. Let X be a polygonal complex. A subdivision of X is a polygonal complex obtained from X by applying finitely many times the following two operations:

1. adding a vertex to an edge;
2. adding an edge contained in a face and joining two existing vertices.

Example 4.15. The following figure shows a cube subdivided in such a way that the top face consists of four squares, the left hand side face consists of a square and three triangles and the front face consists of four triangles. ♣

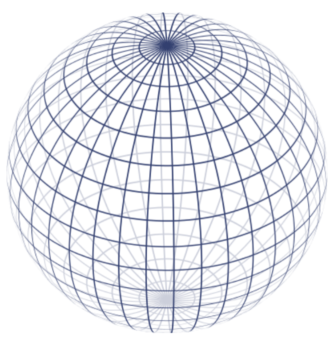


Subdivision of the cube.

Theorem 4.16. *Let X be a finite polygonal complex and let X' be a subdivision of X. Then $\chi(X) = \chi(X')$.*

Proof. The subdivision X' arises by adding vertices and edges to X. Observe that placing an additional vertex on an existing edge does not change the Euler characteristic. Indeed, this operation increase the number of vertices by one and the number of edges also by one. Next observe that adding an edge connectig two vertices and contained in a face also does not change the Euler characteristic. It increases the number of edges by one and the number of faces also by one. □

Example 4.17. The following figure shows the power of Theorem 4.16. The Euler characteristic of this polygonal complex is equal to two.



A polygonal sphere

All you need to observe is that it is a subdivision of, for example, a cube. Of course, this sphere has no shape of a cube, but what matters is the combinatorial structure (the number of vertices, edges and faces). ♣

Theorem 4.18. *If P is a finite polygonal complex isometric to a convex polygon then its Euler characteristic is equal to one.*

Proof. Let P be a finite polygonal complex isometric to a convex polygon Q. We can thus assume that $P \subset \mathbf{R}^2$ is a subset of the plane. Since P is finite, it has finitely many edges $E_1, E_2, \ldots, E_p$. Let $E \subset P$ be an edge and let $L \supset E$ be a straight line on the plane containing E. Let $E' = L \cap P$. The intersection $E \cap E_i$ is either a single point or the whole E_i. Let P_E be the subdivision of P obtained by adding vertices for each of the above sigle point intesection and adding the edges of the form $E \cap F_j$, where $F_j \subset P$ are faces of P. Repeat this procedure for every edge of P. Observe that the resulting subdivision P' is a subdivision of the polygon Q and hence $\chi(P) = \chi(P') = \chi(Q) = 1$. $\square$

Proposition 4.19. *The Euler characteristic of a tree is equal to one.*

Proof. Let T be a tree, i.e. a connected graph without circuits. The proof is by induction on the number of vertices. If T has one vertex (and hence no edges) then its Euler characteristic is obviously one. Suppose that the statement is true for every tree with n vertices. Let T be a tree with $n+1$ vertices. Then T has a vertex that belong to exactly one edge (a leaf). Discarding this vertex and its edge does not change the Euler characteristic and T becomes a tree with n vertices. It follows from the induction hypothesis that $\chi(T) = 1$, which proves the proposition. $\square$

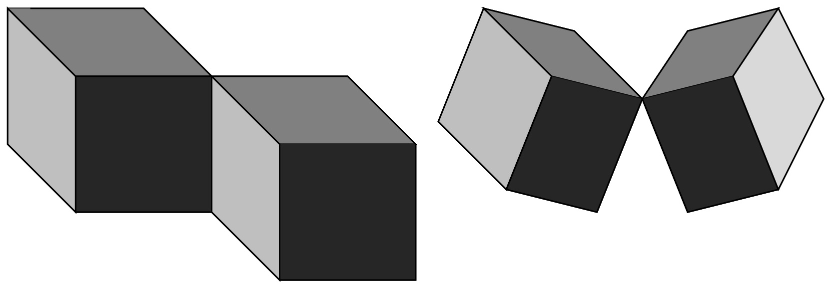
Corollary 4.20. *If G is a connected graph then its Euler characteristic is at most one.*

Proof. Let $T \subset G$ be a minimal subgraph of G containing all the vertices of G. Any such T is a tree because if it had a circuit then it would not be minimal. Thus T and G have the same number of vertices and G has at least as many edges as T which shows that $\chi(G) \leq \chi(T) = 1$, due to Proposition 4.19. $\square$

4.21 Polygonal surfaces

Definition 4.22. A polygonal complex X is called a ***polygonal surface*** if every point in X has a neighbourhood homeomorphic to a neighbourhood of a point in a convex polygon. The union of edges belonging to exactly one face is called the ***boundary*** of X. If X has finite number of faces and has no boundary then it is called ***closed***.

Example 4.23. The following polygonal complexes are not polygonal surfaces. The first has an edge belonging to four faces and the second has a vertex that belongs to faces which do not share an edge.



These are not surfaces.

Theorem 4.24. *If X is a closed connected polygonal surface then its Euler characteristic is at most* 2*.*

Proof. At a price of a suitable subdivision we may assume that X is a triangular closed surface. Remove one face from X; the Euler characteristic has decreased by one. We have now a triangular surface X' with boundary. Apply the following procedures as long as the Euler characteristic is not changed:

1. Remove a triangle with only one edge adjacent to the exterior. This removes one face and one edge and hence does not change the Euler characteristic.
2. Remove a triangle with two edges shared by the exterior. This removes one face, two edges and one vertex and so does not change the Euler characteristic.

Remark 4.25. The Euler characteristic could be changed if at some stage there was a vertex shared by *exactly two* triangles. Then removing one of these triangles would increase the Euler characteristic.

Let T be a triangle in the resulting complex. Apply the following procedures provided they don't change the Euler characteristic:

1. Remove a vertex, two edges and interior of T; it leaves just an edge of T.
2. Remove an edge and the interior of T; it leaves the union of two remaining edges.

In the proces our complex is not a polygonal complex anymore, it is a union of edges and faces. After the procedure terminates we are left with a connected graph G. Since $\chi(G) \leq 1$, due to Corollary 4.20, we obtain that

$$\chi(X) = \chi(X') + 1 = \chi(G) + 1 \leq 2$$

as claimed. □

4.26 Constructing surfaces with given Euler characteristic

Let X and Y be polygonal surfaces. Suppose that X and Y contain an *isometric* n-gonal face (for example, a triangular face which we can always achievie by a suitable subdivision). Remove such a face from both X and Y. We obtain surfaces X' and Y' with n-gonal holes. Define a new polygonal surface $X\#Y$ to be the union of X' and Y' with the boundaries of the n-gonal holes identified.

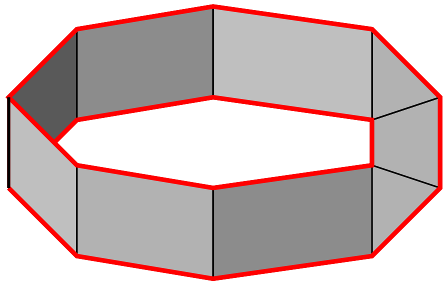
Definition 4.27. If X and Y are polygonal surfaces then the surface $X\#Y$ is called a ***connected sum*** of X and Y.

Proposition 4.28. *The Euler characteristic of a connected sum of surfaces X and Y is equal to*

$$\chi(X\#Y) = \chi(X) + \chi(Y) - 2.$$

Proof. The Euler characteristic of X (respectively Y) with a hole is equal to $\chi(X) - 1$ (respectively $\chi(Y) - 1$). Thus the Euler characterictic of their disjoint union is equal to $\chi(X) - 1 + \chi(Y) - 1$. The connected sum is obtained from the disjoint union by identifying suitable edges and vertices and the number of edges and vertices is decreased by n which do not change the Euler characteristic. □

Example 4.29. The g-fold connected sum of a polygonal torus has Euler characteristic equal to $2 - 2g$. Such a complex is called a polygonal surface of genus g. ♣



Polygonal Möbius band.

Example 4.30. The figure above shows a polygonal Möbius band. Its boundary (drawn in red) is a sixteen-gon. Let X be a closed polygonal surface obtained from the above Möbius band and a polygonal 16-gon by identifying the boundaries. In other words we take a polygonal 16-gon and attach it to the Möbius band along the boundary. This cannot be physically done in three dimensional space. We can quickly calculate that the Euler characteristic of the surface X is equal to one. Such a surface is called a polygonal ***projective plane***. ♣

Example 4.31. Taking a connected sum of a polygonal surface of genus g with a polygonal projective space we obtain a closed surface with Euler characteristic equal to $2 - 2g - 1$. ♣

The above examples prove the following result.

Theorem 4.32. *For any integer $k \leq 2$ there exists a closed polygonal surface X with $\chi(X) = k$.* □

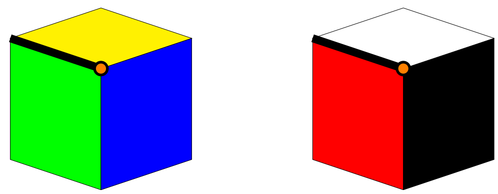
Remark 4.33. There is a theorem stating that the above examples are all possible closed surfaces up to a homeomorphism. The proof uses a very basic algebraic topology.

EXERCISES

1. Read carefully all examples and provide more details and make relevant drawings if possible.
2. Let X be a square surface obtained as follows. Take a solid cube subdivided into 27 equal cubes. Remove the central cube and each cube containing the centre of a face of the big cube (seven cubes removed). Let X be the surface of the remaining solid. Compute the Euler characteristic $\chi(X)$.
3. Construct a triangular closed surface with the Euler characteristic equal to zero.
4. Let X be a polygonal surface such that all faces of X are k-gons for some $k > 2$ and m faces meet at each vertex. Proof the following statements:

   (a) $2E = kF$; (b) $2E = mV$; (c) $kF = mV$.

5. Take a polygonal surface of genus g (for example $g = 2$), remove one face, subdivide, and glue in the Möbius band. What is the Euler characteristic of the resulting surface?
6. Take two Möbius bands and glue them together by identifying their boundaries. What is the Euler characteristic of the resulting surface?
7. The surface from the previous part has zero Euler characteristic. Prove that it is not homeomorphic to a torus.
8. Suppose that X is a triangular closed surface such that six triangles meet at each vertex. What is the Euler charateristic of X?
9. Suppose that X is a triangular closed surface such that k triangles meet at each vertex. What is the Euler charateristic of X?
10. Construct a closed square surface made from 12 squares in which 6 squares meet at each vertex. You can do this as follows. Take two cubes as in the picture below. Cut them open along the thick black lines. Glue the white face of the right hand side cube to the green face of the left hand one and the yellow face to the red one. Repeat the procedure at all edges parallel to the thick black one (four in total in each cube). Thus the two orange vertices of both cubes become one in which six squares meet: white-green-blue-yellow-red-black-white. This procedure is impossible in our three dimensional space. Compute the Euler characteristic of the constructed polygonal surface.



Construction of a square complex in which 6 squares meet at each vertex

11. Try to give a rigorous definition of the above surface. For example, the expressions *cut open* or *glue the face to another face* are mathematically meaningless. We could say that our surface is obtained from a disjoint union of 12 squares by introducing an appropriate equivalence relation on the set of the boundary points of the squares.
12. Let $\mathbf{C}$ be the surface of the unit cube and let $A \in \mathbf{C}$ be a vertex. Is there a geodesic $\gamma\colon [a, b] \to \mathbf{C}$ such that $\gamma(a) = \gamma(b) = A$?
13. The same question as in the previous exercise for the case of surfaces of other Platonic solids (tetrahedron, octahedron, dodecahedron and icosahedron).

5 The Gauss-Bonnet theorem for polygonal complexes

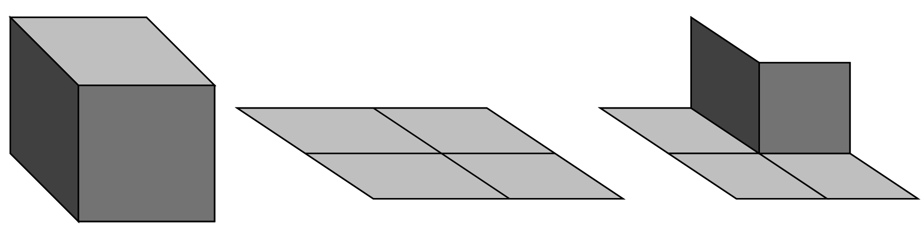
Definition 5.1. Let $v \in X$ be an interior vertex of a polygonal surface. The ***curvature*** of X at v is defined to be

$$K(v) := 2\pi - (\alpha_1 + \alpha_2 + \dots \alpha_n),$$

where α_i are all angles at the vertex v. We also say that X is

- ***positively curved*** at v if $K(v) > 0$;
- ***flat*** at v if $K(v) = 0$;
- ***negatively curved*** at v if $K(v) < 0$.

Example 5.2. The picture shows examples of positively curved (the curvature of a vertex in a cube is equal $\pi/2$), flat and negatively curved (here the curvature is $-\pi/2$) polygonal surfaces. ♣



Positive curvature **Flat** **Negative curvature**

The following theorem is a simple example of a profound connection between geometry and topology. A version of it was known to Gauss and a special case was published by Bonnet. The interplay between geometry and topology has been extensively studied since then.

Theorem 5.3. *Let X be a closed polygonal surface. The sum of curvatures at all vertices of X is equal to 2π times the Euler characteristic of X.*

Proof. Let α_i denote an angle and let $\alpha_{i(v)}$ denote an angle at a vertex v and let $\alpha_{i(f)}$ denote an angle of a face f. Let F_k denote the number of k-gonal faces of X.

$$
\begin{aligned}
\sum_v K(v) &= \sum_v (2\pi - \sum \alpha_{i(v)}) && (5.1)\\
&= 2\pi V - \sum \alpha_i && (5.2)\\
&= 2\pi V - \sum_{k=3}^{\infty} F_k(k-2)\pi && (5.3)\\
&= 2\pi V - \sum_{k=3}^{\infty} (\pi k F_k - 2\pi F_k) && (5.4)\\
&= 2\pi V + 2\pi F - \pi \sum_{k=3}^{\infty} kF_k && (5.5)\\
&= 2\pi V + 2\pi F - 2\pi E && (5.6)\\
&= 2\pi(V - E + F) = 2\pi\chi(X) && (5.7)
\end{aligned}
$$

Let us explain each equality in the above calculation:

1. This is just the definition of the curvature.
2. We are summing up all the angles at all vertices and hence we are summing up all the angles of X.
3. We rearange the sum in such a way we are summing up the angles of all faces. Since the complex is finite the sum is actually finite.
4. Direct computation.
5. Again.
6. We use the fact that the sum $\sum kF_k$ is equal to twice the number of edges (exercise).
7. The rest is obvious.

□

The amazing thing of the Gauss-Bonnet theorem is that it tells us what is the topology of a closed surface from from the geometric information about the curvature of vertices. For example, if we know that all vertices are nonegatively curved and some of them are positively curved then the surface has to be a polygonal sphere!

5.4 The Gauss-Bonnet theorem for non-closed surfaces

Definition 5.5. If v is a vertex on the boundary of a polygonal surface then its ***curvature*** is defined to be

$$K(v) = \pi - (\alpha_1 + \ldots + \alpha_k),$$

where α_i are angles at v.

The follwoing result is a general form of the Gauss-Bonnet theorem for polygonal surfaces. Its proof is left as an exercise.

Theorem 5.6. *If X is a finite polygonal surface then*

$$\sum_v K(v) = 2\pi\chi(X)$$

□

EXERCISES

1. Read carefully all examples and provide more details and make relevant drawings if possible.
2. Prove the general version of the Gauss-Bonnet theorem.
3. Let X be a closed triangular surface in which six triangles meet at each vertex. What is the Euler characteristic of X?
4. What can you say about the Euler characteristic of a closed polygonal surface made of k-gons where $k \geq 6$?
5. Let X be a closed heptagonal surface in which three faces meet at each vertex. Show that the number of vertices is divisible by seven.
6. Let X be a k-gonal surface in which three faces meet at each vertex. Show that $(6-k)V = 2k\chi(X)$. Derive interesting consequences for some concrete values of k. For example, if $k = 11$ then the number of vertices is divisible by 22 and $|\chi(X)|$ is divisible by five.

6 Geometric meaning of curvature

In this section we are going to examine some geometric features of polygonal surfaces and how the curvature affetcs them.

Definition 6.1. A polygonal surface is called:

- ***non-positively curved*** if all internal vertices have non-positive defects;
- ***non-negatively curved*** if all internal vertices have non-negative defects.

Remark 6.2. We don't define positively or negatively curved polygonal surfaces because they are flat almost everywhere (every face is a flat polygon). Later we will define a different notion of a curvature and we will see examples of nowhere flat surfaces (e.g. a sphere).

6.3 Perimetrer

We know that the perimeter of a disc of radius r on the Euclidean plane is equal to $2\pi r$, where $\pi \simeq 3.14....$ Let X be a polygonal surface and let $D(p, r)$ denote the ball of radius r centered at p. In the sequel, we shall call it a disc of radius r centered at p. Let $\pi_X \colon X \times [0, \infty) \to \mathbf{R}$ be a function defined by

$$\pi_X(p, r) = \frac{\text{Perimeter}(D(p, r))}{2r}.$$

In the case of the Euclidean plane the function is constant and equal to the number π.

Example 6.4 (The cube)**.** If X is the unit cube then we have the following situation. If $D(p, r)$ does not contain a vertex then its perimeter is equal to $2\pi r$ and the disc itself is isometric to the Euclidean disc of the same radius.

Suppose that p is a vertex and $r \leq 1$. We see that the perimeter of the disc is now equal to $\frac{3}{4}(2\pi r)$. For a radius bigger than one the formula for the perimeter is more complicated, although the values of π_{cube} are not bigger than π and that for $r \geq \sqrt{5}$ the value $\pi_{\text{cube}}(p, r) = 0$. ♣

Example 6.5 (Negative curvature)**.** Let X be a square surface such that five squares meet at each vertex as in Example 4.4. If p is a vertex then we see that the perimeter of a disc of radius at most one is equal to $\frac{5}{4}(2\pi r)$ and in general the function π_X grows exponentially fast with the radius. ♣

We see in the above examples that in the case of positive curvature disc tend to have smaller perimeter than flat discs of the same radius and in the case of negative curvature they tend to have bigger perimeter.

6.6 Geodesics and geodesic polygons

Let X be a geodesic metric space and let $x, y \in X$ be two points. Since X is a geodesic space there exists a path of length $d(x, y)$ from x to y. The image of this path is called a ***geodesic segment*** and it is denoted by $[x, y]$. Notice that the notation $[x, y]$ does not specify a unique geodesic segment as there may be many of them. A ***closed geodesic*** in an isometric embedding $\mathbf{S}^1(R) \to X$ of the circle of radius R.

Definition 6.7. Let X be a geodesic metric space. A ***geodesic*** n***-gon*** consists of n points $x_1, \ldots, x_n \in X$ called vertices and n geodesic segments $[x_1, x_2], \ldots, [x_{n-1}, x_n], [x_n, x_1]$ called edges. A geodesic n-gon is called ***degenerate*** if it has a pair of edges intersecting at infinitely many points.

A geodesic $(n-1)$-gon can be made into an n-gon by adding a vertex to the interior of an existing edge. If an n-gon is not obtained from an $(n-1)$-gon by this procedure it is called ***proper***. In other words, in a proper n-gon the angle between geodesic segments meeting at a vartex is different from π.

Example 6.8. A geodesic 3-gon is called a geodesic ***triangle***. A geodesic triangle in the Euclidean plane is the usual triangle. A geodesic triangle on a sphere is called ***spherical triangle***. A geodesic 2-gon is called ***digon*** or a ***diangle***. Only degenerate digons exist on the Euclidean plane. However, on the plane with the d_1-metric there are infinitely many pairwise distinct digons between any two distinct pooints. The same is true for the sphere. For example there are infinitely many digons between any pair of antipodal points. A 1-gon is called a ***unigon***. They do not exist on the euclidean plane. A 0-gon is the image of a closed geodesic. $\diamondsuit$

Let $\gamma, \gamma' \colon [0, \infty) \to X$ be two geodesic rays starting at $p = \gamma(0) = \gamma'(0)$. In the following examples we discuss a typical behaviour.

Example 6.9 (The cube)**.** Let γ and γ' be two distinct geodesic on the cube starting at a vertex v and initially contained in the same face. Assume that the angle of both γ and γ' with the diagonal of this face at v is $\alpha < \frac{\pi}{10}$. Observe that γ and γ' intersect again forming a digon.

Example 6.10. If we have two distinct geodesic rays on the Euclidean plane starting from the same point then they go away from each other and they never meet again. So there are no diangles on the Euclidean plane (as we always knew). $\clubsuit$

Let's investigate consequences of existing of a diangle in a polygonal surface. Let $D \subset X$ be a polygonal subcomplex such that its boundary is a diangle with vertices $p, q \in X$. Let α and β be the internal angles of D at p and q respectively. It follows from the Gauss-Bonnet theorem that

$$\begin{aligned} 2\pi\chi(D) &= \sum_v K(v) \\ &= \pi - \alpha + \pi - \beta + \sum_{v \in \mathrm{int}(D)} K(v) \\ &= 2\pi - (\alpha + \beta) + \sum_{v \in \mathrm{int} D} K(v). \end{aligned}$$

Since the Euler characteristic of a polygonal diangle is at most one we obtain the following inequality.

$$\sum_{v \in \mathrm{int} D} K(v) \leq \alpha + \beta.$$

That is, the sum of the curvatures at the internal vertices of D is at most $\alpha + \beta$.

Corollary 6.11. *If X be a non-positively curved polygonal surface then it does not contain polygonal diangles of Euler characteristic one.* □

Corollary 6.12. *If X is a simply connected non-positively curved polygonal surface then it does not contain diangles. In particular, two distinct geodesics starting from the same point never meet again.* □

Corollary 6.13. *Let X be a simply connected non-positively curved polygonal surface. If γ is a geodesic between p and q then it is unique.* □

6.14 Triangles

A ***polygonal triangle*** in a polygonal surface is a subcomplex whose boundary is a geodesic triangle whose edges intersect only at the vertices of the triangle.

We have a reasonably good intuition for triangles on the Euclidean plane. For example, we know that the sum of internal angles is equal to π and many other nice things. Let us try to gain some intuition for triangles in positive and negative curvature. We will do it by looking at our usual examples.

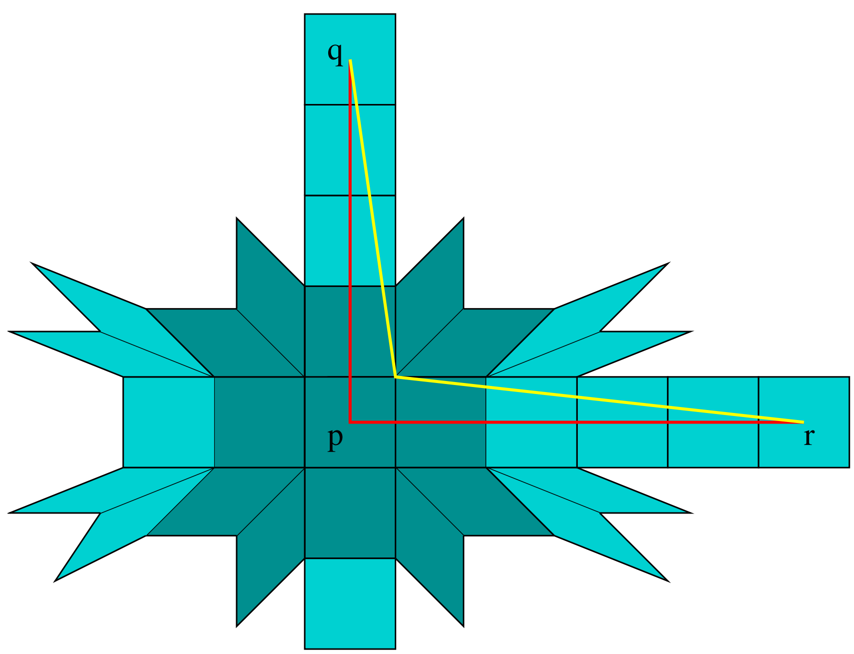
Example 6.15 (The cube)**.** Let T be a triangle on the unit cube with internal angles α, β, γ. It's Euler characteristic is equal to one and by applying the Gauss-Bonnet theorem we get that

$$2\pi = 3\pi - (\alpha + \beta + \gamma) + k\frac{\pi}{2},$$

where k is the number of vertices of the cube contained in the interior of T. This implies that

$$\alpha + \beta + \gamma = \pi + k\frac{\pi}{2}.$$

Hence if a triangle contains a vertex of the cube then the sum of its internal angles is bigger that π. ♣



A triangle in a negatively curved complex

Example 6.16 (Negative curvature)**.** Let X be the square complex from Example 4.4. That is, simply connected square complex in which five squares meet at each vertex. Let γ and γ' be perpendicular geodesic rays (drawn in red) starting at the middle p of a face and parallel to its sides. We are interested in the triangle with vertices p, $q := \gamma(s)$ and $r := \gamma'(t)$. If s, t are small then the triangle is isometric to an Euclidean triangle (its third side is drawn in bright green). However, if s, t are big then the triangle looks much different from an Euclidean triangle. Observe that the yellow line is a geodesic from q to r. The only thing we need to check is that the angle when it passes through a vertex near p is equal to π and this is an easy exercise.

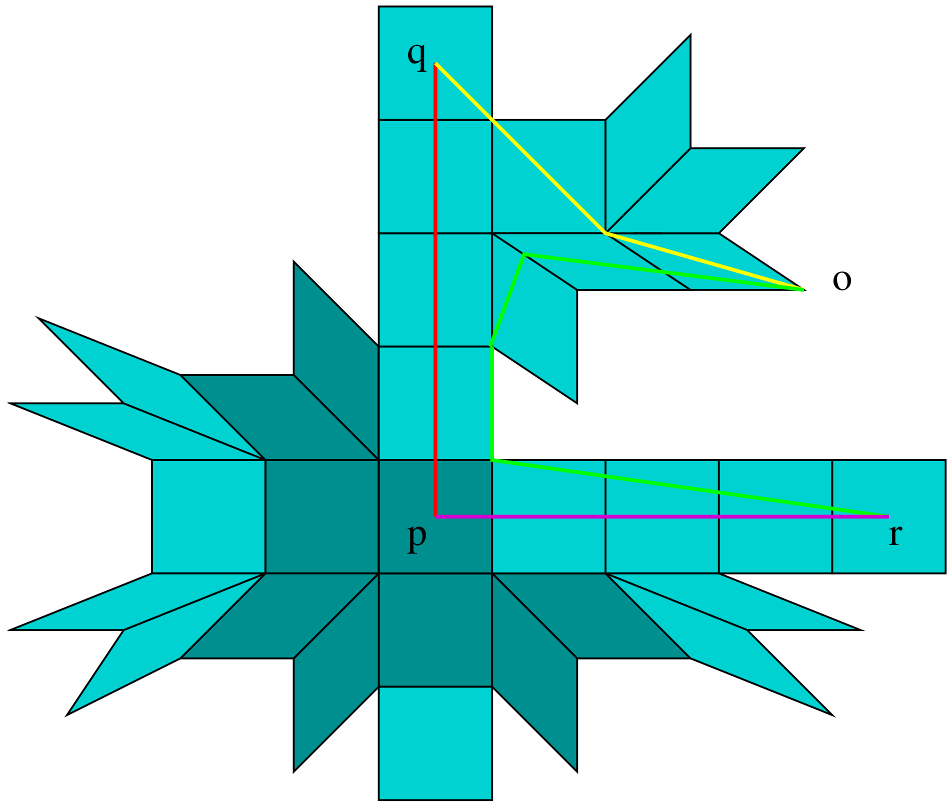
It means that the yellow 'shortcut' from q to r is almost of the same length as the red route. ♣

Example 6.17. Suppose we have a triangle with angles α, β and γ in a non-positively curved and simply connected polygonal surface. The simple connectivity implies that the subcomplex enclosed by the triangle has Euler characterictic equal to one. The Gauss-Bonnet formula implies that the sum of angles $\alpha + \beta + \gamma$ is at most π. ♣

6.18 A 'practical' application

Imagine a region in the Highlands and approximate it with a polygonal complex. Observe that positively curved vertices are the summits and kettles and most of the region are mountain passes which are negatively curved. This implies that there are hardy any shortcuts in the mountains and hence it is very easy to get lost if we rely on our Euclidean intuition. Let's look at an example presented in the following figure. Suppose we want to go for a hike from point r through p (where we take a right turn) to q and return. We take a decision based on our Euclidean intuition to take a 45° turn at q and go along the yellow path. At the point o we loose our confidence and we ask for a shortest path back to r. It is the green path which looks like we almost need to go back.

Of course, in the reality mountains are not as negatively curved as our example.



Getting lost in polygonal mountains

EXERCISES

1. Read carefully all examples and provide more details and make relevant drawings if possible.
2. Give a formula for $\pi_X(r, v)$ where X is the surface of

   (a) the cube,
   (b) the regular tetrahedron,
   (c) the regular octahedron,

   and v is a vertex and $r \geq 0$ is arbitrary. In the case of the cube work out the case when v is the centre a face.
3. Give a formula for $\pi_{\mathbf{S}^2}(r, p)$.
4. Give a formula for $\pi_{\mathbf{R}^2}(r, 0)$, where $\mathbf{R}^2$ is equipped with either L^1.
5. Construct a closed surface with all its vertices negatively curved.
6. Construct a closed surface in $\mathbf{R}^3$ with all its vertices negatively curved.

   Remark 6.19. When a professional geometer is asked a question:

   Does there exist a closed polygonal surface in $\mathbf{R}^3$ *with all vertices negatively curved?*

   most of them answer in the negative (including myself a few years ago). This is because the analogous statement for *smooth* surfaces in $\mathbf{R}^3$ is not true: a smooth closed surface in $\mathbf{R}^3$ always has regions where it is positively curved. An earlier version of this course had an exercise which asked for a proof that such a polygonal surface cannot exist in $\mathbf{R}^3$. Only when I was challenged by the students I came up with examples. Since then it is one of my favourite geometric problems. The first geometer who gave the right answer straight away was Anton Petrunin and he included this problem in his book.

7. Prove Corollaries 6.11, 6.12 and 6.13.
8. Let X be the lateral surface of a right circular cone with unit generatrix and directrix equal to $3\pi/2$. Is a neighbourhood of the apex isometric to a neighbourhood of a vertex of a cube?
9. Let X and Y be the lateral surfaces of right circular cones with unit generatrices. When do their apices have isometric neighbourhoods?

7 Spheres and hyperbolic spaces

7.1 The intrinsic metric on a sphere

Recall that the Euclidean scalar product on the space $\mathbf{R}^{n+1}$ is defined by

$$\langle x, y\rangle = \sum x_i y_i,$$

where $x = (x_0, x_1, \ldots, x_n)$ and $y = (y_0, y_1, \ldots, y_n)$. It can be used to define the Euclidean angle between nonzero vectors x and y by the following identity:

$$\angle(x, y) = \arccos \frac{\langle x, y\rangle}{\|x\|\|y\|},$$

where $\|x\| = \sqrt{\langle x, x\rangle}$. Let $\mathbf{S}^n(R) = \{x \in \mathbf{R}^{n+1} \mid \langle x, x\rangle = R^2\}$ be the sphere of radius R centered at the origin in the Euclidean space. The sphere of radius one will be simply denoted by $\mathbf{S}^n$. The following formula defines a metric on $\mathbf{S}^n(R)$:

$$d(x, y) = R \arccos \frac{\langle x, y\rangle}{R^2}.$$

In particular, if $R = 1$ then the distance between x and y is simply the angle between them.

Lemma 7.2. *The above formula defines a metric on* $\mathbf{S}^n(R)$.

Proof. Observe that it is enough to prove it for the unit sphere. We have that $d(x, y)$ if and only the scalar product $\langle x, y\rangle = 1$ which implies that the angle between x and y is zero which means that $x = y$. The symmetry follows from the symmetry of the scalar product and the triangle inequality follows from a general fact that if $\gamma_1, \gamma_2, \gamma_3$ are geodesics in a metric space with angles between all pairs of them defined then

$$\angle(\gamma_1, \gamma_3) \leq \angle(\gamma_1, \gamma_2) + \angle(\gamma_2, \gamma_3),$$

see Proposition 3.21 or the spherical Laws of Cosines we shall prove later. □

Lemma 7.3. *The following statements hold true for* $\mathbf{S}^n(R)$*:*

1. *The diameter of* $\mathbf{S}^n(R)$ *is equal to* πR.
2. *The geodesics are arclength parametrisations of the great circles.*

3. *The restriction of an othogonal linear transformation of $\mathbf{R}^{n+1}$ is an isometry of $\mathbf{S}^n(R)$.*
4. *There are diangles: for example, two meridians between the north and the south poles form a diangle.*
5. *Between any two antipodal point there are infinitely many geodesics (like meridians between the poles).*

Proof. Exercise. □

Example 7.4 (Visual thought experiment)**.** Imagine our $3D$-space is not the Euclidean $\mathbf{R}^3$ but the sphere $\mathbf{S}^3(R)$ of a large radius, so that we don't notice that we leave on the sphere. In the same way early people did not notice they were living on the surface of a sphere. Also assume that the radius R is not so big so that the light can travel between antipodal point within seconds or minutes. Remember that the light travels along geodesics. Imagine two of us staring at each other and assume the retina of your eye is exactly at the north pole of the spere. Imagine I am moving away from you towards the south pole. What do you see? Initially you roughly see whatever would happen in the Euclidean space: I get smaller and smaller linearly. This continue, although not linearly, until I reach the equator. When I move further, you see me bigger and bigger; initially only slightly bigger but then I grow much faster. From your point of view it is the same as if I moved towards you from the equator. When I approach the south pole I am covering almost all your visual horizon and when I reach the south pole you see me all over the place (as if I was touching your eye). This is because the geodesics (light rays) from the south pole to the north pole cover the whole sphere. ♢

7.5 The two dimensional unit sphere

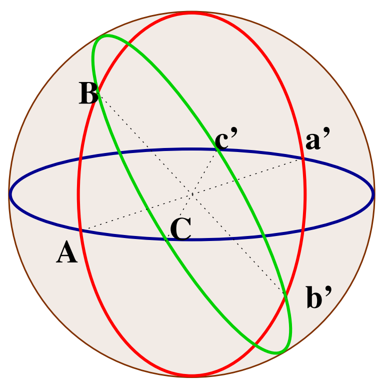
We restrict our attention to the two-dimensional unit sphere and we develop some basic trigoonometry in this section.

Proposition 7.6. *The area of a geodesic triangle with internal angles on the unit sphere is equal to $\alpha+\beta+\gamma-\pi$, where α, β and γ are internal angles of the triangle. In particular, the sum of angles is bigger than π.*

Remark 7.7. Notice that the notion of the area is not defined on a metric space in general. We shall precisely define it later for certain class of metric spaces. If you took the course on Measure Theory in the first semester then you may defined the area on the sphere as a certain Lebesgue measure. Alternatively, it can be defined as an appropriate surface integral.

Proof. Let's take for granted the fact (known to Archimedes) that the total area of the unit sphere is equal to 4π. (We can prove it by calculating the appropriate integral). It then follows that the are of a diangle D_α with internal angles equal to α is equal to 2α.

Let T denote a triangle with internal angles α, β and γ and vertices A, B and C respectively. In the following figure we see that the sphere is a union of six diangles of the respective areas equal to 2α, 2β and 2γ overlaping on the triangle with vertices A, B and C.



A spherical triangle

The triangle T and its antipodal twin (with vertices a', b', c') are covered by three layers from diangles. Thus when expressing the total area of the sphere as the sum of the areas of diangles we count the area of the triangle T four times too much,

$$4\pi = 4(\alpha + \beta + \gamma) - 4\,\mathrm{Area}(T),$$

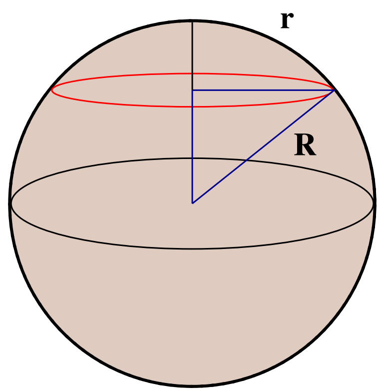
which gives that $\mathrm{Area}(T) = \alpha + \beta + \gamma - \pi$, as claimed. □

Example 7.8 (The function $\pi_{\mathbf{S}^2}$)**.** First notice that the function $\pi_{\mathbf{S}^2}$ does not depend on the choice of the point on $\mathbf{S}^2$. This is due to the fact that for every two points $p, q \in \mathbf{S}^2$ there exists an isometry f such that $f(p) = q$.

Observe that the circumference of a disc of radius r (the red circle in the figure below) on the sphere $\mathbf{S}^2$ is equal to $2\pi\sin(r)$ and hence we have

$$\pi_{\mathbf{S}^2}(p, r) = \frac{\pi\sin(r)}{r}.$$

This means that the perimeter of a circle on the sphere is always smaller than the perimeter of the Euclidean circle. We also see that if the radius r is very small then the function π is very close to the Euclidean π. ♣



The function π on the sphere.

Corollary 7.9. *The sphere $\mathbf{S}^2$ is not locally isometric to the Euclidean plane. This precisely means that there is no nonempty open subset of $\mathbf{S}^2$ isometric to a subset of the Euclidean plane.* □

In 'practice' this result tells us that we can't glue a post-stamp on a sphere. There will always be folds and wrikles (try it!). However, we can easily glue a stamp on a cylinder. A more serious practical observation is that it is impossible to make an accurate map of the Earth. By accurate we mean that it is an isometry up to a scaling factor. Since most of our geographical knowledge comes from studying maps our intuition is distorted. For example, we find it surprising that a plane from Heathrow to New York flies over Greenland.

Corollary 7.10. *An nonempty open subset $U \subset \mathbf{S}^2(r)$ is isometric to an open subset $V \in \mathbf{S}^2(R)$ if and only if $r = R$.* □

7.11 The spherical law of cosines

Theorem 7.12. *Let $T \subset \mathbf{S}^2$ be a spherical triangle with vertices A, B, C and the corresponding angles α, β, γ and side lengths a, b, c. Then the following two laws of cosines hold:*

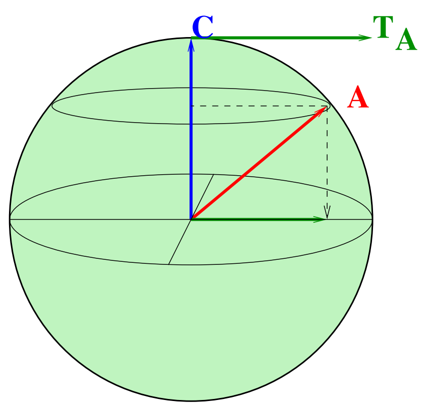
$$\cos(c) = \cos(a)\cos(b) + \sin(a)\sin(b)\cos(\gamma)$$

$$\cos(\alpha) = -\cos(\beta)\cos(\gamma) + \sin(\beta)\sin(\gamma)\cos(a).$$

Proof. We think of A, B, C as of unit vectors in $\mathbf{R}^3$. The angle γ is the angle between (unit vectors) T_A, T_B located at C and tangent to the geodesics joining C with A and B respectively. They can be computed in a straighforward way and we have that

$$T_A = \frac{A - C\langle A, C\rangle}{\|A - C\langle A, C\rangle\|} = \frac{A - C\langle A, C\rangle}{\sin a},$$

$$T_B = \frac{B - C\langle B, C\rangle}{\|B - C\langle B, C\rangle\|} = \frac{B - C\langle B, C\rangle}{\sin b}.$$



The vector T_A.

It then follows that

$$\begin{aligned}
\cos(\gamma) &= \langle T_A, T_B \rangle \\
&= \frac{\langle A - C\langle A, C\rangle, B - C\langle B, C\rangle\rangle}{\sin(a)\sin(b)} \\
&= \frac{\langle A, B\rangle - 2\langle A, C\rangle\langle B, C\rangle + \langle A, C\rangle\langle B, C\rangle}{\sin(a)\sin(b)} \\
&= \frac{\cos(c) - \cos(b)\cos(a)}{\sin(a)\sin(b)}
\end{aligned}$$

and we get that

$$\cos(c) = \cos(a)\cos(b) + \sin(a)\sin(b)\cos(\gamma).$$

□

Corollary 7.13 (Pythagoras's theorem on the sphere)**.** *If* $\gamma = \pi/2$ *then*

$$\cos(c) = \cos(a)\cos(b).$$

7.14 The hyperbolic space

The Lorentz scalar product on $\mathbf{R}^{n+1}$ is defined by

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i - x_{n+1}y_{n+1}.$$

The real hyperbolic space is defined to be

$$\mathbf{H}^n = \{x \in \mathbf{R}^{n+1} \mid \langle x, x\rangle = -1 \text{ and } x_{n+1} > 0\}.$$

The hyperbolic metric on $\mathbf{H}^n$ is defined by

$$\mathrm{d}(x, y) := \operatorname{arccosh}\langle x, y\rangle.$$

The metric space $(\mathbf{H}^2, \mathrm{d})$ is called the ***hyperbolic plane***. By a ***model*** of the hyperbolic plane we mean a metric space (X, d_X) which is isometric to $(\mathbf{H}^2, \mathrm{d})$.

7.15 The hyperbolic plane

We start by presenting two useful models of the hyperbolic plane: the half-plane model and the disc model. We will make a choice of the model according to the problem we investigate. For example, for understanding the group of isometries of the hyperbolic plane we will invastigate the half-plane and for computing the area of a triangle we will work with the disc model. Both models can serve as a definition of the hyperbolic plane.

The hyperplane model. Let $\mathbf{H} = \{z \in \mathbf{C} \mid \operatorname{Im}(z) > 0\}$ be the upper half of the complex plane. The hyperbolic metric is defined by

$$\mathrm{d}(z, w) = \operatorname{arccosh}\left(1 - \frac{2|z-w|^2}{(z-\overline{z})(w-\overline{w})}\right).$$

In the Euclidean coordinates it has the follwoing form

$$\mathrm{d}(z, w) = \operatorname{arccosh}\left(1 + \frac{(x_1-x_2)^2 + (y_1-y_2)^2}{2y_1y_2}\right),$$

where $z = x_1 + iy_1$ and $w = x_2 + y_2$.

The disc model. Let $\mathbf{D} \subset \mathbf{C}$ be the unit Euclidean disc on the complex plane. We equip it with the metric given by

$$\mathrm{d}(z, w) = \arccos\left(1 + \frac{2|z-w|^2}{(1-|z|^2)(1-|w|^2)}\right).$$

Isometries of the hyperbolic plane (in the half-plane model)

In this section we work with the half-plane model. The following result is straighforward to prove.

Lemma 7.16. *For every matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_+(2, \mathbf{R})$ with positive determinant the map $f_A \colon \mathbf{H} \to \mathbf{H}$ defined by*

$$f_A(z) = \frac{az+b}{cz+d} \tag{7.1}$$

is an isometry of the hyperbolic metric. □

The map given by (7.1) is called either a ***fractional transformation*** or a ***Möbius transformation***.

Example 7.17. Observe that the following maps are isometries of the hyperbolic plane (for each of them write a matrix of the fractional transormation):

- scaling: $z \mapsto az$, $a > 0$;
- horizontal translation: $z \mapsto z + b$, $b \in \mathbf{R}$;
- negative inverse: $z \mapsto -\frac{1}{z}$.

$\diamondsuit$

Example 7.18. Let $L \in \mathbf{H}$ be a vertical line, $x = a$. The reflection of $\mathbf{H}$ with respect to L is the map given by

$$z \mapsto z + 2(a - \mathrm{Re}(z)) = -\overline{z} + 2a.$$

This reflection is an isometry. Since horizontal translations are isometries it is enough to check that the map $z \mapsto -\overline{z}$ is an isometry, which is done by a direct computation. Observe that the above reflection is not a fractional transformation (why?). $\clubsuit$

Example 7.19. The isometry $f(z) = -1/z$ is the composition of the inversion in the unit circle with the reflection in the imaginary axis, $f = R \circ I$. Since the reflection is an isometry, so is the inversion $I = R^{-1} \circ f$. It then follows that the inversion in any circle centered in the real axis is an isometry because it is conjugate to the above inversion I by a composition of a translation and a scaling. $\clubsuit$

Corollary 7.20. *The inversion in a half-circle perpendicular to the real axis is an isometry of the hyperbolic plane.* $\square$

Remark 7.21. It is perfectly fine to think about a vertical line in $\mathbf{H}$ as of a half-circle (of infinite radius). The inversion is then the ordinary reflection which is an isometry.

Geodesics of the hyperbolic plane (in the half-plane model)

Proposition 7.22. *If $\gamma\colon [a, b] \to \mathbf{H}$ is a geodesic then its image is contained either in a vertical line or in a circle perpendicular to the real axis. More precisely, every geodesic $\gamma\colon \mathbf{R} \to \mathbf{H}$ is given either by*

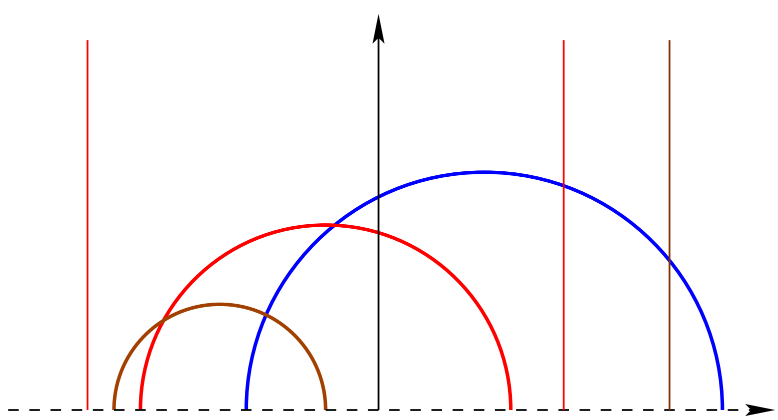
$$\gamma(t) = a + \mathbf{i}e^t \quad \textit{or by} \quad \gamma(t) = \frac{ae^{2t} + b}{e^{2t} + 1} + \frac{(a - b)e^t}{e^{2t} + 1}\mathbf{i},$$

where $a > b$ are real numbers.

Proof. Let us proof that $\gamma(t) = a + \mathbf{i}e^t$ defines a geodesic. We shall show that $\mathrm{d}(\gamma(s), \gamma(t)) = |s - t|$ for all $s, t \in \mathbf{R}$. Assume that $s \geq t$ and use the formula for the distance in the Euclidean coordinates in the following computation.

$$\begin{aligned} \mathrm{d}(\gamma(s), \gamma(t)) &= \mathrm{d}(a + \mathbf{i}e^s, a + \mathbf{i}e^t) \\ &= \operatorname{arccosh}\left(1 + \frac{(e^s - e^t)^2}{2e^s e^t}\right) \\ &= \operatorname{arccosh}\left(\frac{e^{2s} + e^{2t}}{2e^{s+t}}\right) \\ &= s - t \end{aligned}$$

This computation shows that $\gamma(t) = a + \mathbf{i}e^t$ defines a globally distance minimising geodesic for any $a \in \mathbf{R}$.



Geodesics on the hyperbolic plane

Let $a > b$ be two points on the real line and let $f_A \colon \mathbf{H} \to \mathbf{H}$ be given by $A = \left(\begin{smallmatrix} a & b \\ 1 & 1 \end{smallmatrix}\right) \in \mathrm{GL}_+(2, \mathbf{R})$. Take the geodesic defined by $\gamma(t) = \mathbf{i}e^t$ and compose it with f_A. Since f_A is an isometry the composition $f_A \circ \gamma$ is a geodesic and it is straightforward to check that its image is a half-circle perpendicular to the real axis and intersecting the real axis at a and b. It is given by the formula

$$(f_A \circ \gamma)(t) = \frac{ae^{2t} + b}{e^{2t} + 1} + \frac{(a - b)e^t}{e^{2t} + 1}\mathbf{i}.$$

We have to postpone the proof of the fact that all geodesics have the above form until we develop more tools. □

Remark 7.23. In the above proof we used the usual trick: we found one geodesic and then composed it with isometries in order to find many more.

Corollary 7.24. *The hyperbolic plane is a geodesic metric space.*

Proof. Let $w, z \in \mathbf{H}$. If their real parts are equal then there is a vertical geodesic between them. If the real parts are distinct then there is a half-circle containing w and z, and perpendicular to the real line. □

Corollary 7.25. *The geodesics in the disc model* $\mathbf{D}$ *of the hyperbolic plane are either diameters or arcs of circles perpendicular to the boundary of* $\mathbf{D}$.

Proof. Recall that an inversion carries circles to circles (here we think that a line is a circle of infinite radius) and preserves angles. Since the isometry between $\mathbf{H}$ and $\mathbf{D}$ is given by an inversion, the images of geodesics in $\mathbf{H}$ are lines and circles perpendicular to the boundary of $\mathbf{D}$. □

The hyperbolic area. In the case of a sphere or the Euclidean plane we intuitively know what is an area (roughly, the amount of paint you need to cover a subset of the physical sphere or the plane). In the case of the hyperbolic plane we don't know. We will define the notion of the area later. As in the case of the sphere where we used that the total area of the unit sphere was 4π, here we take for granted than the area of an ideal triangle is equal to π.

The area of a hyperbolic triangle

Definition 7.26. The ***ideal boundary*** of the hyperbolic plane is the boundary of the disc $\mathbf{D}$. An ***ideal triangle*** on the hyperbolic plane is the union three geodesics which meets on the ideal boundary.

In the hyperplane model the ideal boundary is the union of the real axis and $\{\infty\}$ (an ideal end of all vertical geodesics). As above an ideal triangle is the union of three geodesics meeting on the ideal boundary.

Lemma 7.27. *If* $\Delta_1, \Delta_2 \subset \mathbf{H}$ *are ideal triangles then there exists an isometry* $\psi\colon \mathbf{H} \to \mathbf{H}$ *such that* $\psi(\Delta_1) = \Delta_2$.

Proof. Let $x, y, z \in \mathbf{R} \cup \infty$ be ideal vertices of the triangle Δ_1. Suppose without losing generality that $x < y < z$. We will prove that there is a fractional transformation f_A such that $f_A(x) = 0,\ f_A(y) = 1,\ f_A(z) = \infty$.

First, there exists a horizontal translation f_1 such that $f_1(x) = 0$. Next there exists a scaling f_2 such that $f_2(0) = 0$ and $f_2(y) = 1$. It remains to show that there exists a fractional transformation f_A which satifies the following conditions: $f_A(0) = 0$ and $f_A(1) = 1$ and $f_A(z') = \infty$, where $z' = f_2(f_1(z)) > 1$. Let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. The first condition implies that $b = 0$, the second that $d = a - c$ and the third that $c = \frac{a}{1-z'}$. Thus f_A is associated with the matrix

$$A = \begin{pmatrix} z' - 1 & 0 \\ -1 & z' \end{pmatrix}$$

and the required isometry is the composition $f_A \circ f_2 \circ f_1$.

The same argument yields an isometry taking Δ_2 to the ideal triangle with vertices $0, 1, \infty$. The composition of the first with the inverse of the second gives the required ψ. □

Corollary 7.28. *The area of any ideal triangle is equal to the are of the triangle with ideal vertices* $0, 1, \infty$. □

Theorem 7.29. *Let* $T_{\alpha,\beta,\gamma} \subset \mathbf{D}$ *be a hyperbolic triangle with internal angles* α, β, γ. *Then its area is equal to*

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

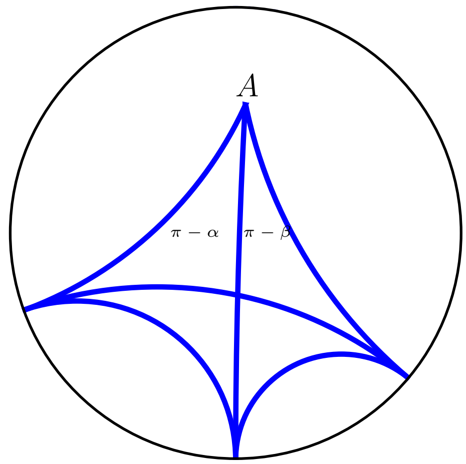
In particular, the sum of the internal angles in a hyperbolic triangle is smaller than π.

Proof. In this proof we consider the disc model of the hyperbolic plane. We take for granted the fact that the area of every ideal triangle is equal to π (we will prove this later).

The first step is to prove that $\mathrm{Area}(T_{\alpha,0,0}) = \pi - \alpha$. Let $f(\alpha) = \mathrm{Area}(T_{\pi-\alpha,0,0})$. The function f satisfies the following identity:

$$f(\alpha) + f(\beta) = f(\alpha + \beta - \pi) + \pi,$$

which can be proven by looking at the picture below.



Area of a triange with two ideal vertices

Moreover, we know that $f(0) = 0$ (degenerate triangle) and $f(\pi) = \pi$ (ideal triangle). Consider the following observations:

$$\begin{aligned} f\left(\frac{\pi}{2}\right) + f\left(\frac{\pi}{2}\right) &= f(0) + \pi = \pi & &\implies f\left(\frac{\pi}{2}\right) = \frac{\pi}{2} \\ 2f\left(\frac{3\pi}{4}\right) &= f\left(\frac{\pi}{2}\right) + \pi & &\implies f\left(\frac{3\pi}{4}\right) = \frac{3\pi}{4} \\ f\left(\frac{3\pi}{4}\right) + f\left(\frac{\pi}{2}\right) &= f\left(\frac{\pi}{4}\right) + \pi & &\implies f\left(\frac{\pi}{4}\right) = \frac{\pi}{4} \end{aligned}$$

In general, if $f\left(\frac{k\pi}{2^n}\right) = \frac{k\pi}{2^n}$ then $f\left(\frac{(2^{n+1}-1)\pi}{2^{n+1}}\right) = \frac{(2^{n+1}-1)\pi}{2^{n+1}}$ which is proven by induction. It follows that $f(\alpha) = \alpha$ for all $\alpha \in [0, \pi]$ which are dyadic multiples of π. The general case is a consequence of the continuity of the area.

Consider a general triangle $T_{\alpha,\beta,\gamma}$ and observe (by extending its edges to infinity) that is it contained in an ideal triangle and this ideal triangle is a union of $T_{\alpha,\beta,\gamma}$ and three triangles of the form $T_{\pi-\alpha,0,0}$,$T_{\pi-\beta,0,0}$ and $T_{\pi-\gamma,0,0}$. Consequently,

$$\pi = \text{Area}(T_{\alpha,\beta,\gamma}) + \alpha + \beta + \gamma,$$

which implies that $T_{\alpha,\beta,\gamma} = \pi - (\alpha + \beta + \gamma)$. $\square$

EXERCISES

1. Compute the radius of the Earth.

   Hint: Go to the beach and using the law of cosines measure the distance from your eyes to the horizon (you need two people and two protractors and a ship on the horizon). Then apply the Pythagoras theorem.
2. Compute the distance from Aberdeen to New York. Use the following data: Aberdeen coordinates - $57N, 2W$, New York coordinates - $40N, 74W$, and the radius of the Earth is approximately 6384km. *Hint:* Look at the spherical triangle with vertices in Aberdeen, New York and the North Pole.
3. Let $A = [a_1, a_2, a_3],\ B = [b_1, b_2, b_3],\ C = [c_1, c_2, c_3] \in \mathbf{S}^2$ be three points on the sphere. Give a formula for the area of a geodesic triangle with vertices A, B, C in terms of the coordinates of the vertices. Notice that you need to make choices for this problem. For example, a sensible choice to to assume that the edges of the triangle are the *shorter* geodesics between the vertices. Try special cases first.
4. Let $\mathrm{I}_{p,r} \colon \mathbf{R}^2 - \{p\} \to \mathbf{R}^2 - \{p\}$ be an inversion in the circle of radius r centered at $p \in \mathbf{R}^2$. Write the formula for $\mathrm{I}(x, y)$.
5. Prove that an inversion preserves circles and straight lines (the image of a circle is either a circle or a line and the image of a line is either a line or a circle).
6. Prove that the half-plane model and the disc model are isometric.
7. Let $A \in \mathrm{GL}_+(2, \mathbf{R})$ and let $f_A \colon \mathbf{H} \to \mathbf{H}$ be the corresponding Möbius transformation. Express f_A as the composition of translations, inversion, and scaling. Deduce that a Möbius map shares all good properties with an inversion.
8. Draw ten pairwise non-isometric geodesic triangles on the hypebolic plane.
9. Take for granted the fact that the angle between geodesics on the hyperbolic place is equal to the *Euclidean* angle between their tangent lines at the intersection point (we will prove it later). Compute the angle between the following geodesics:

   (a) $\gamma(t) = a + ie^t$ for $-1 \leq a \leq 1$ and the geodesic parametrising the unit circle.

   (b) The unit circle centered at the origin and the unit circle centered at $0 < a \leq 2$.

(c) The unit circle centered at the origin and the circle of radis $r > 0$ centered centered so that they intersect.

Convince yourself that you can compute the angle between any two geodesics.

10. Compute the area of the hyperbolic triangle with vertices:

    (a) $i, e^{i\frac{\pi}{4}}, \infty$;
    (b) $i, e^{i\frac{\pi}{3}}, \infty$;
    (c) $i, e^{i\frac{\pi}{6}}, \infty$;
    (d) i, e^{it}, ∞, where $0 \leq t \leq \frac{\pi}{2}$;
    (e) e^{is}, e^{it}, ∞, where $0 \leq s < t \leq \pi$.

11. Provide a detailed and efficient algorithm for computing the area of the hyperbolic triangle with vertices A, B, C. That is, your input consists of three complex numbers with positive imaginary part the output is the area of the triangle. If you have experience, write a suitable computer program.
12. Read wikipedia article on models and the history of the hyperbolic plane.
13. Google the artwork by M.C.Escher and admire those related to the hyperbolic plane.

8 Surfaces in the three dimensional space

The aim of this and the next sections is to develop tools for understanding smooth surfaces in the Euclidean 3-dimensional space. The first step is to define such a surface as a subset that can be locally parametrised by maps defined on open subsets of the plane. Next we restrict the standard Eulidean scalar product in $\mathbf{R}^3$ to tangent spaces of a surface $S \subset \mathbf{R}^3$. This allows to define lengths of paths on S and the metric. The geometry of the intrinsic metric of S is then investigated using linear algebra of tangent vectors. This is an example of a Faustian bargain: we sell geometry for algebra. After the deal is made we can calculate a lot of fantastic things but we can't see and touch anymore.

Throughout this section $\langle X, Y \rangle$ denotes the standard Euclidean scalar product on $\mathbf{R}^3$.

8.1 Basic definitions

Let $T_u\mathbf{R}^m$ denote the space of vectors located at $u \in \mathbf{R}^n$. It is canonicaly isomorphic to the vector space $\mathbf{R}^m$. If R^n is a differentiable map then its differential (or the tangent map) at $u \in \mathbf{R}^m$ is a linear map $df_u \colon T_u\mathbf{R}^m \to T_{f(u)}\mathbf{R}^n$. In the standard basis the differential is represented by the Jacobi matrix (df_i/du_j).

Example 8.2. Let $f \colon \mathbf{R}^2 \to \mathbf{R}^3$ be a differentiable map defined by $f(u,v) = (u^2, u+2v, e^u + v^3)$. Then its tangent map at (u,v) is given by the matrix $\begin{pmatrix} 2u & 0 \\ 1 & 2 \\ e^u & 3v^2 \end{pmatrix}$. ♣

Definition 8.3.

1. Let $U \subset \mathbf{R}^2$ be an open subset. A differentiable map $f \colon U \to \mathbf{R}^3$ such that the differential $df_u \colon T_u\mathbf{R}^2 \to T_{f(u)}\mathbf{R}^3$ is injective for all $u \in U$ is called a ***surface patch***. We say that a patch f ***parametrises*** $f(U)$.
2. A ***surface*** in $\mathbf{R}^3$ is a subset $S \subset \mathbf{R}^3$ satisfying the following conditions:
   (a) for every point $p \in S$ there exist a surface patch $f \colon U \to S$ such that $p \in f(U)$
   (b) if $f_i \colon U_i \to S$ are surface patches then the map
   $$f_{ij} \colon U_j \cap f_j^{-1}(f_i(U_i)) \to U_i \cap f_i^{-1}(f_j(U_j))$$
   given by $f_{ij} = f_i^{-1} \circ f_j$ is a diffeomorphism (see the figure below).

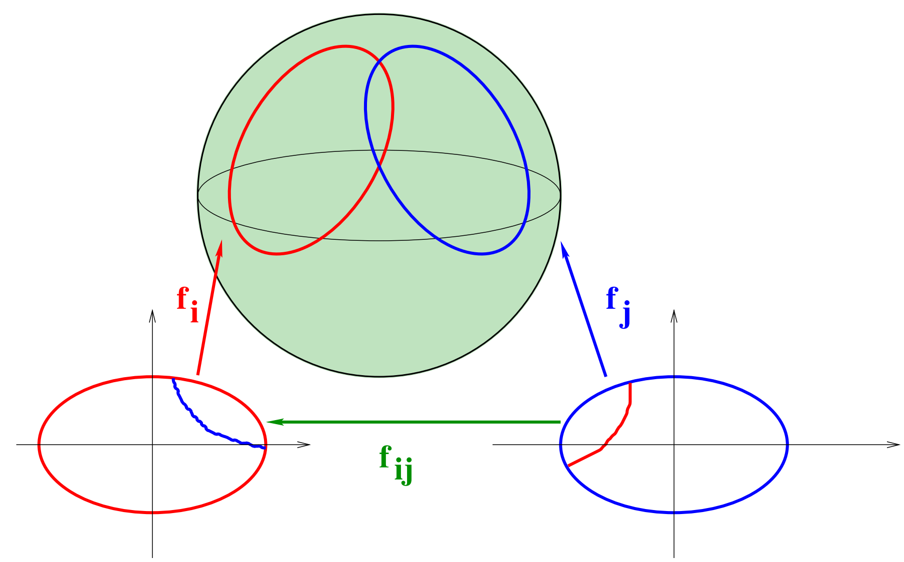
Notice that $f(U)$ is a surface.

3. The two-dimensional subspace $df_u(T_u\mathbf{R}^2) \subset T_{f(u)}\mathbf{R}^3$ is called the ***tangent space of*** S ***at*** $f(u)$ and it is denoted by $T_{f(u)}S$. Elements of $T_{f(u)}S$ are called ***tangent vectors*** to S.
4. The symmetric bilinear form $\mathbf{I}\colon T_pS \times T_pS \to \mathbf{R}$ defined by $\mathbf{I}(X, Y) = \langle X, Y \rangle$ is called the ***first fundamental form*** of S. Consequently, if $c\colon [a, b] \to S$ is a differentiable path then its length is computed by

$$\mathcal{L}(c) = \int_a^b \mathbf{I}(\dot{c}(t), \dot{c}(t))\; dt$$

(this is simply the Euclidean length of c considered as a path in $\mathbf{R}^3$). It follows that the intrinsic metric on S is defined as

$$d(x, y) = \inf\{\mathcal{L}(c) \in \mathbf{R} \mid c(a) = x,\; c(b) = y\}.$$

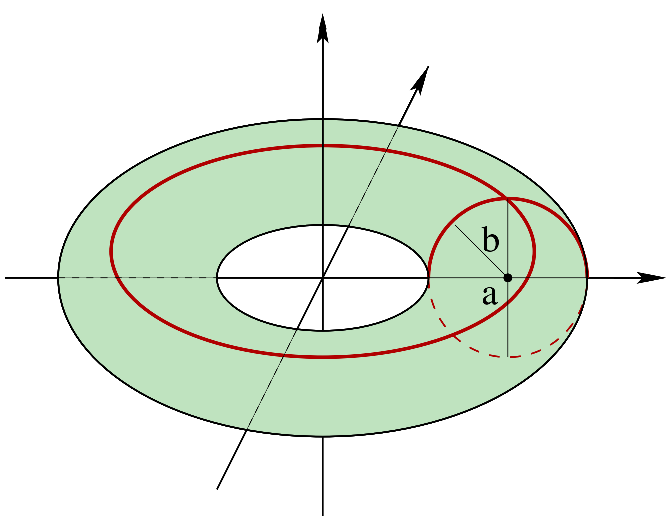


Patches of a sphere.

Example 8.4.

1. If $x, y \in \mathbf{R}^3$ are linearly independent vectors then $f\colon \mathbf{R}^2 \to \mathbf{R}^3$ defined by $f(u,v) = x_0 + ux + vy$ parametrises a plane in the three dimensional space.
2. Let $U = \{(u,v) \in \mathbf{R}^2 \,|\, u^2 + v^2 = 1\}$. The function defined by $f(u,v) = (u, v, \sqrt{1-u^2-v^2})$ parametrises the upper hemisphere of radius one centered at the origin.
3. Let $U = (-\pi/2, \pi/2) \times \mathbf{R}$. The function given by the formula $f(u,v) = (\cos u \cos v, \cos u \sin v, \sin u)$ parametrises the unit sphere minus the poles $\mathbf{S}^2 - \{0, 0, \pm 1\}$.
4. Let $f\colon \mathbf{R}^2 \to \mathbf{R}^3$ be defined by the following formula $f(u,v) = ((a + b\cos u)\cos v, (a + b\cos u)\sin v, b\sin u)$, where $a > b > 0$. This parametrises the torus. Also, this example shows that a patch does not have to be injective.
5. The function $f(u,v) = (v\cos u, -v\sin u, au)$ parametrises a helicoid.

♣



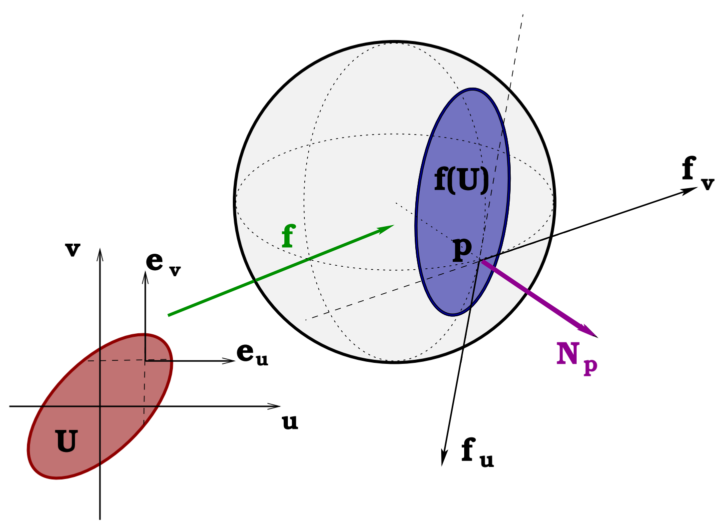
A torus

8.5 Vectors associated with a surface patch

Let $u, v\colon \mathbf{R}^2 \to \mathbf{R}$ be the standard coordinates on the plane and denote by $\mathbf{e}_u = (1,0), \mathbf{e}_v = (0,1) \in T_{(u,v)}\mathbf{R}^2$ the the basis of the tangent space at a point $(u,v) \in \mathbf{R}^2$. Let $f\colon U \to \mathbf{R}^2$ be a surface patch and let $\mathbf{f}_u = df(\mathbf{e}_u)$ and $\mathbf{f}_v = df(\mathbf{e_v})$, where the derivatives are computed at a point (u,v). The vectors $\mathbf{f}_u, \mathbf{f}_v$ are tangent to the surface $f(U)$ at $p = f(u,v)$. The ***normal unit vector*** to $f(U)$ at p is defined by

$$\mathbf{n}_p = \frac{\mathbf{f}_u \times \mathbf{f}_v}{\|\mathbf{f}_u \times \mathbf{f}_v\|},$$

where $p = f(u,v)$ and $\times$ denotes the vector product in $\mathbf{R}^3$.



The normal unit vector

The second derivatives $\mathbf{f}_{uu}, \mathbf{f}_{uv}, \mathbf{f}_{vv}$ are not tangent to the surface $f(U)$ in general. To measure their deviation from being tangent we project them onto the normal vector. That we compute the scalar products

$$\begin{aligned} L &= \langle \mathbf{f}_{uu}, \mathbf{n}_p \rangle, \\ M &= \langle \mathbf{f}_{uv}, \mathbf{n}_p \rangle = \langle \mathbf{f}_{vu}, \mathbf{n}_p \rangle, \\ N &= \langle \mathbf{f}_{vv}, \mathbf{n}_p \rangle \end{aligned}$$

Definition 8.6. The bilinear form on $T_p f(U)$ defined by the matrix $\left(\begin{smallmatrix} L & M \\ M & N \end{smallmatrix}\right)$ with respect to the basis $\mathbf{f}_u$, $\mathbf{f}_v$ is called the ***second fundamental form*** of f at p. It is denoted by $\mathbf{II}$. It can be also considered as a bilinear form on $T_{(u,v)}\mathbf{R}^2$ represented by the same matrix with respect to the basis $\mathbf{e}_u$, $\mathbf{e}_v$.

Example 8.7. Let $f\colon \mathbf{R}^2 \to \mathbf{R}^3$ be given by $f(u, v) = (u, v, 0)$. That is, f parametrises the horizontal plane through the origin. We have:

$$\begin{aligned} \mathbf{f}_u &= (1, 0, 0) & \mathbf{f}_{uu} &= (0, 0, 0) \\ \mathbf{f}_v &= (0, 1, 0) & \mathbf{f}_{uv} &= (0, 0, 0) \\ \mathbf{n} &= (0, 0, 1) & \mathbf{f}_{vv} &= (0, 0, 0) \end{aligned}$$

It follows that the first fundamental form is represented by the identity matrix and the second fundamental form is identically zero. ♢

Example 8.8. Consider the parametrisation $f\colon U = (-\pi/2, \pi/2) \times \mathbf{R} \to \mathbf{R}^3$ of the sphere (without poles) given by $f(u, v) = (\cos u \cos v, \cos u \sin v, \sin u)$. In this case we have

$$\begin{aligned} \mathbf{f}_u &= (-\sin u \cos v, -\sin u \sin v, \cos u) \\ \mathbf{f}_v &= (-\cos u \sin v, \cos u \cos v, 0) \\ \mathbf{I} &= \begin{pmatrix} 1 & 0 \\ 0 & \cos^2 u \end{pmatrix} \\ \mathbf{n} &= -(\cos u \cos v, \cos u \sin v, \sin u) \\ \mathbf{f}_{uu} &= (-\cos u \cos v, -\cos u \sin v, -\sin u) \\ \mathbf{f}_{uv} &= (\sin u \sin v, -\sin u \cos v, 0) \\ \mathbf{f}_{vv} &= (-\cos u \cos v, -\cos u \sin v, 0) \\ \mathbf{II} &= \begin{pmatrix} 1 & 0 \\ 0 & \cos^2 u \end{pmatrix} \end{aligned}$$

♢

The procedure in the above examples is as follows:

1. Write the formula for the surface patch.
2. Compute the first derivatives.
3. Get vectors $\mathbf{f}_u$ and $\mathbf{f}_v$ and compute $\mathbf{n}$ and $\mathbf{I}$.

```
4. Compute the second derivatives.
5. Get vectors f_uu, f_uv and f_vv.
6. Compute ⟨f_ij, n⟩ and get II.
```

8.9 Abstract surfaces

Let $f\colon U \to S \subset \mathbf{R}^3$ be a surface patch. The first fundamental form can be defined directly on U by

$$\mathbf{I}_{(u,v)}(X,Y) = \langle df_{(u,v)}(X), df_{(u,v)}(Y)\rangle$$

for any $X, Y \in T_{(u,v)}U$. This defines the lengths of paths $c\colon [a,b] \to U$ by

$$\mathcal{L}(c) = \int_a^b \sqrt{\mathbf{I}_{c(t)}(\dot{c}(t), \dot{c}(t))}\, dt$$

and hence a metric $d(x,y) = \inf\{\mathcal{L}(c) \mid c(a) = x,\ c(b) = y\}$. In other words, the set U is given a metric such that (U,d) is locally isometric to the surface $f(U)$ with its intrinsic metric.

HOWEVER

We can define a symmetric bilinear positive definiete form $\mathbf{g}_{(u,v)}\colon T_{(u,v)}U \times T_{(u,v)}U \to \mathbf{R}$ directly without pulling it back from the Euclidean space and define lengths of paths and a metric as above using $\mathbf{g}$ instead of $\mathbf{I}$. In this approach we get a metric on U which does not necessarily correspond to a surface in $\mathbf{R}^3$. Thus $\mathbf{g}$ is a function whose value at (u,v) is an inner product on $T_{(u,v)}U$. It is called a ***metric tensor***.

Example 8.10. Let $U = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times (a,b)$ and let $\mathbf{g}_{(u,v)}$ be defined by the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & \cos^2 u \end{pmatrix}.$$

The map $f\colon U \to \mathbf{R}^3$ parametrising the sphere without poles defines a local isometry between U and $\mathbf{S}^2 \setminus \{(0,0,\pm 1)\}$. If $|a-b| \le \pi$ then f is a genuine isometry between U and $f(U) \subset \mathbf{S}^2$. If $|a-b| > \pi$ then f is local isometry. For example, the distance in U between $(0,0)$ and $(0,2\pi)$ is 2π and the distance between $f(0,0) = (1,0,0)$ and $f(0,2\pi) = (1,0,0)$ is zero. If we take $(a,b) = \mathbf{R}$ then we obtain a metric on an infintie strip which is locally isometric to the sphere. $\diamondsuit$

Example 8.11. Let $\mathbf{H} = \{(x, y) \in \mathbf{R}^2 \mid y > 0\}$ be the upper half-plane. Let

$$\mathbf{g}_{(x,y)} = \frac{1}{y^2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

be the metric tensor. It is the standard Euclidean scalar product scaled by $\frac{1}{y^2}$. Thus the length of a curve $c\colon [a, b] \to \mathbf{H}$ is computed by

$$\begin{aligned} \mathcal{L}(c) &= \int_a^b \mathbf{g}_{c(t)}(\dot{c}(t), \dot{c}(t))\ dt \\ &= \int_a^b \frac{\sqrt{c_1'(t)^2 + c_2'(t)^2}}{c_2(t)}\ dt. \end{aligned}$$

For example, if $c(t) = (0, e^t)$ then

$$\mathcal{L}(c) = \int_a^b \frac{\sqrt{0^2 + e^{2t}}}{e^t}\ dt = \int_a^b dt = b - a.$$

In fact, the metric induced by $\mathbf{g}$ is isometric to the hyperbolic metric. $\diamondsuit$

8.12 The area form of a surface

Let $(U, \mathbf{g})$ be an abstract surface with a metric tensor $\mathbf{g}$. The area of a subset $A \subset U$ is defined by

$$\text{area}(A) = \int_A \sqrt{\det \mathbf{g}_{(u,v)}}\ dudv.$$

Observe that this formula gives the standard area of a subset of the Euclidean plane. Also, if $f\colon U \to \mathbf{R}^3$ is a surface patch then $(U, \mathbf{g})$ is isometric to $f(U)$ equipped with the intrinsic metric and the above formula corresponds to the standard surface intergral.

Proposition 8.13. *The area of an ideal hyperbolic triangle is equal to π.*

Proof. Consider the ideal triangle $T \in \mathbf{H}$ with vertices $(-1, 0), (1, 0), \infty$ in the half-plane model. Since for the hyperbolic plane we have $\det \mathbf{g} = \frac{1}{y^4}$ (see Example 8.11), we need to compute the integral $\int_T \frac{dxdy}{y^2}$ which, according to Green's Theorem, is equal to the path intgral

$$\int_{\partial T} \frac{dx}{y},$$

where ∂T is the path (in fact three paths) parametrising the edges of the triangle T. Since two of these edges are vertical the above integral will vanish on them and we are left with the integral over the edge which is a half-circle from $(-1,0)$ to $(1,0)$. It has the following parametrisation

$$c(t) = \left(\frac{e^{2t}-1}{e^{2t}+1}, \frac{2e^t}{e^{2t}+1} \right),$$

(see the section about geodesics on the hyperbolic plane). We now finish the computation of the area of the triangle T.

$$\begin{aligned} \text{area}(T) &= \int_{c(\mathbf{R})} \frac{dx}{y} = \int_{-\infty}^{\infty} \frac{c_1'(t)}{c_2(t)} dt \\ &= \int_{-\infty}^{\infty} \frac{\frac{4e^{2t}}{(e^{2t}+1)^2}}{\frac{2e^t}{e^{2t}+1}} \, dt \\ &= \int_{-\infty}^{\infty} \frac{2e^t}{e^{2t}+1} \, dt = \pi. \end{aligned}$$

Since all ideal hyperbolic trangles are isometric, the above computation finishes the proof. $\square$

EXERCISES

1. Let $A, B, C \in \mathbf{R}^3$ be point not on the same line. Write a parametrisation of the plane through A, B, C.
2. Let $A, B, C \in \mathbf{R}^3$ be as above. Prove that there exist infinitely many distinct spheres containing A, B, C.
3. Let $A, B, C, D \in \mathbf{R}^3$ be points not included in a plane. Prove that there exists a unique sphere containing these points. Determine the centre of this sphere and its radius.
4. Let $h\colon (a, b) \to \mathbf{R}$ be a differentiable function. Consider a parametrisation $f\colon (a, b) \times \mathbf{R} \to \mathbf{R}^3$ defined by $f(u, v) = (h(u)\cos v, h(u)\sin v, u)$. Draw the surface parametrised by f for h defined by:

   (a) $h(u) = u + 1$;
   (b) $h(u) = u^2 + 1$;
   (c) $h(u) = u^3 + 1$;
   (d) $h(u) = \cos u$;
   (e) $h(u) = \frac{1}{x+2}$.

   Observe that this surface is obtained by first drawing the graph of $x = h(z)$ in the xz-plane and then rotating it around the z-axis.
5. Let $h, k\colon (a, b) \to \mathbf{R}$ be differentiable functions and let $u \mapsto (h(u), 0, k(u))$ be a parametrisation of a curve in the xz-plane. Let $f\colon (a, b) \times \mathbf{R} \to \mathbf{R}^3$ be a parametrisation defined by $f(u, v) = (h(u)\cos v, h(u)\sin v, k(u))$. Draw the surface parametrised by f for h, k defined by:

   (a) $h(u) = \cos u + 2$, and $k(u) = \sin u$;
   (b) $h(u) = \alpha\cos u + 2$, and $k(u) = \beta\sin u$, where $0 < \alpha < 2$ and $\beta > 0$.

   Similarly, this surface is *swept out* by the curve $(h(u), k(u))$ in the xz-plane by rotating it around the z-axis.
6. Compute the second fundamental form of the following surfaces:

   (a) a plane;
   (b) a sphere of radius $r > 0$;
   (c) a general ***surface of revolution*** as in the previous exercise provided that $(h')^2 + (k')^2 \neq 0$ and $h \neq 0$. Evaluate your formula on all examples in the previous two exercises.

9 Curvature

9.1 The Gauss map

Let $f\colon U \to S \subset \mathbf{R}^3$ be a surface. The map $\nu\colon U \to \mathbf{S}^2$ defined by

$$\nu(u,v) = \mathbf{n}_{f(u,v)}$$

is called the ***Gauss map***. Its value is the unit normal vector at $p = f(u,v)$ relocated to the origin and hence defining a point on the unit sphere.

Proposition 9.2. *Let $f\colon U \to \mathbf{R}^3$ be a surface and let $p = f(u,v)$. The image of $d\nu_{(u,v)}\colon T_{(u,v)}\mathbf{R}^2 \to T_p\mathbf{R}^3$ is included in $T_pS \subset T_p\mathbf{R}^3$.*

Proof. The image is spanned by the vectors defined by the partial derivatives $\mathbf{n}_u := d\nu(\mathbf{e}_u)$ and $\mathbf{n}_v := d\nu(\mathbf{e}_v)$. Since the normal vector is of unit length we have that $\langle \nu(u,v), \nu(u,v)\rangle = 1$. Differentiating this equality yields

$$\begin{aligned}\langle \mathbf{n}_u(u,v), \nu(u,v)\rangle &= 0\\ \langle \mathbf{n}_v(u,v), \nu(u,v)\rangle &= 0.\end{aligned}$$

This means that both $\mathbf{n}_u$ and $\mathbf{n}_v$ are tangent to S as claimed. $\square$

9.3 The Gauss map and the second fundamental form

Proposition 9.4. *Let $f\colon U \to S \subset \mathbf{R}^3$ be a surface. The second fundamental form satisfies the following equality*

$$\mathbf{II}(X,Y) = -\langle d\nu(X), df(Y)\rangle.$$

Proof. Since $\mathbf{n}$ and $\mathbf{f}_u$ are perpendicular, we have the following computation

$$0 = \partial_u\langle \mathbf{n}, \mathbf{f}_u\rangle = \langle \mathbf{n}_u, \mathbf{f}_u\rangle + \langle \mathbf{n}, \mathbf{f}_{uu}\rangle = \langle \mathbf{n}_u, \mathbf{f}_u\rangle + L.$$

And similarly for M and N (see the definition of the second fundamental form). We obtain that

$$\begin{aligned}L &= -\langle \mathbf{n}_u, \mathbf{f}_u\rangle\\ M &= -\langle \mathbf{n}_v, \mathbf{f}_u\rangle\\ N &= -\langle \mathbf{n}_v, \mathbf{f}_v\rangle\end{aligned}$$

which means that the bilinear form

$$(X,Y) \mapsto -\langle d\nu(X), df(Y)\rangle$$

is represented by the same matrix as the second fundamental form with respect to the same basis which means the two froms are equal. $\square$

9.5 Principal directions and principal curvatures

Let $T_p^1 S = \{X \in T_p S \mid \langle X, X\rangle = 1\}$ be the circle of unit vectors tangent to S at p. A vector $X_0 \in T_p^1 S$ is called a ***principal direction*** if X_0 is a critical point of the function $\kappa\colon T_p^1 S \to \mathbf{R}$ defined by

$$\kappa(X) := \mathbf{II}(X, X).$$

The value $\kappa(X_0)$ is called a ***principal curvature*** at $p \in S$.

Proposition 9.6. *Let $X \in T_p^1 S$. It is a principal direction if and only if it is an eigenvector of the map $W_p\colon T_p S \to T_p S$ defined by*

$$W_p(X) = -(d\nu \circ df^{-1})(X).$$

*The map W_p is called the **Weingarten map**.*

Proof. Let X_0 be a principal direction. The differential of the function $X \mapsto \mathbf{II}(X, X) - \kappa(X_0)\langle X, X\rangle$ at X_0 is trivial (the first summand by definition of the principal direction and the second summand because $\langle X, X\rangle = 1$ is constant). On the other hand, both summands are quadratic forms and hence we get that

$$0 = d(\mathbf{II} - \kappa(X_0)\langle\ ,\ \rangle)_{X_0}(Y) = \mathbf{II}(X_0, Y) - \kappa(X_0)\langle X_0, Y\rangle.$$

Since $\mathbf{II}(X, Y) = \langle W_p(X), Y\rangle$ the above observation implies that

$$-\langle dn(df^{-1}(X_0)), Y\rangle = \mathbf{II}(X_0, Y) = \kappa(X_0)\langle X_0, Y\rangle,$$

for all $Y \in T_p^1 S$. Consequently $W_p(X_0) = \kappa(X_0)X_0$ which means that X_0 is an eigenvector of W_p with eigenvalue $\kappa(X_0)$.

Conversely, assume that $W(X_0) = \kappa(X_0)X_0$. Let $X_0 + \epsilon Y$ be a unit vector (notice that Y here varies with ϵ). We consider the Taylor expansion of κ at X_0:

$$\kappa(X_0 + \epsilon Y) = \kappa(X_0) + \epsilon d\kappa_{X_0}(Y) + \epsilon^2(\dots) + \dots$$

and our aim is to prove that $d\kappa_{X_0}$ vanishes, which means that X_0 is a principal direction.

In the following computation we use the fact that $X_0 + \epsilon Y$ is a unit vector, that is, $\langle X_0 + \epsilon Y, X_0 + \epsilon Y\rangle = 1$. Since X_0 is also a unit vector, we get that $2\epsilon\langle X_0, Y\rangle + \epsilon^2\langle Y, Y\rangle = 0$.

$$
\begin{aligned}
\kappa(X_0+\epsilon Y) &= \mathbf{II}(X_0+\epsilon Y, X_0+\epsilon Y)\\
&= \mathbf{II}(X_0,X_0) + 2\epsilon\mathbf{II}(X_0,Y) + \epsilon^2\mathbf{II}(Y,Y)\\
&= \kappa(X_0) - 2\epsilon\langle d\nu(df^{-1}X_0), Y\rangle + \epsilon^2\mathbf{II}(Y,Y)\\
&= \kappa(X_0)\langle X_0+\epsilon Y, X_0+\epsilon Y\rangle + 2\epsilon\langle W_p(X_0), Y\rangle + \epsilon^2\mathbf{II}(Y,Y)\\
&= \kappa(X_0)\Big(\langle X_0,X_0\rangle + 2\epsilon\langle X_0,Y\rangle + \epsilon^2\langle Y,Y\rangle\Big) + 2\epsilon\langle W_p(X_0),Y\rangle + \epsilon^2\mathbf{II}(Y,Y)\\
&= \kappa(X_0)\Big(\langle X_0,X_0\rangle - 2\epsilon\langle X_0,Y\rangle - \epsilon^2\langle Y,Y\rangle\Big) + 2\epsilon\langle \kappa(X_0)X_0,Y\rangle + \epsilon^2\mathbf{II}(Y,Y)\\
&= \kappa(X_0) - 2\epsilon\langle\kappa(X_0)X_0,Y\rangle + 2\epsilon\langle\kappa(X_0)X_0,Y\rangle + \epsilon^2\Big(\mathbf{II}(Y,Y) - \kappa(X_0)\langle Y,Y\rangle\Big)\\
&= \kappa(X_0) + \epsilon^2\Big(\mathbf{II}(Y,Y) - \kappa(X_0)\langle Y,Y\rangle\Big).
\end{aligned}
$$

This shows that the linear term in the Taylor expansion of κ at X_0 vanished which proves that $d\kappa_{X_0} = 0$ as claimed. □

Corollary 9.7. *The second fundamental form is either proportional to the induced Riemannian structure ($\mathbf{II} = \kappa g$), in which case every direction is principal, or there exists exactly two (up to sign) principal directions orthogonal to each other.*

Proof. Let κ_1 and κ_2 be the largest and the smallest principal curvatures associated with principal directions X_1 and X_2 respectively. We have

$$\kappa_1\langle X_1, X_2\rangle = \mathbf{II}(X_1, X_2) = \kappa_2\langle X_1, X_2\rangle.$$

Therefore either $\kappa_1 = \kappa_2$ which means that κ is constant, or $\kappa_1 > \kappa_2$ and $\langle X_1, X_2\rangle = 0$. If X_0 is a principal direction then either $\langle X, X_1\rangle = 0$ or $\langle X, X_2\rangle = 0$ which implies that either $X_0 = \pm X_1$ or $X_0 = \pm X_2$. □

9.8 The Gauss and the mean curvature

Definition 9.9. Let $f\colon U \to \mathbf{R}^3$ be a surface. The ***Gauss curvature*** $K\colon U \to \mathbf{R}$ and the ***mean curvature*** $H\colon U \to \mathbf{R}$ are functions defined by

$$K(u) := \kappa_1(u)\kappa_2(u) \qquad H(u) := \frac{1}{2}(\kappa_1(u) + \kappa_2(u)),$$

where κ_i are the principal curvatures.

Proposition 9.10. *Let $[g_{ij}]$ be the matrix of the first fundamental form of a surface $f\colon U \to \mathbf{R}^3$ and let $[g^{ij}]$ denote its inverse. Let $[h_{ij}]$ be the matrix of the second fundamental form. Then the matrix $[w_{ij}]$ representing the Weingarten map is given by*

$$w_{ij} = \sum_k h_{ik} g^{kj}.$$

Consequently,

$$\begin{aligned} K(u) &= \frac{\det \mathbf{II}}{\det g} = \frac{\det[h_{ij}]}{\det[g_{ij}]} = \frac{LN - M^2}{EG - F^2}, \\ H(u) &= \frac{1}{2} = \sum_{i,j} h_{ij} g^{ij}. \end{aligned}$$

Proof. Let $\mathbf{e}_i$ and $\mathbf{f}_i$ denote the standard basis vectors of $T_{(u,v)}U$ and $T_{f(u,v)}S$ respectively. We have

$$\sum_k w_{ij}\mathbf{f}_j = W(\mathbf{f}_i) = -dn(df^{-1}(\mathbf{f}_i) = -dn(\mathbf{e}_i) = -\mathbf{n}_i.$$

Taking the inner product with $\mathbf{f}_k$ we obtain

$$\sum w_{ij} g_{jk} = h_{ik}$$

This is equivalent to the matrix equation $[w_{ij}][g_{jk}] = [h_{ik}]$. Multiplying both sides by the inverse $[g^{kj}]$ we obtain the claimed formula for the matrix representing the Weingarten map.

The formulae for K and H follows because K is equal to the determinant of the Weingarten map and H is equal to the half of the trace. $\square$

EXERCISES

1. Let V be a real vector space and let $\beta\colon V \to \mathbf{R}$ be a quadratic form corresponding the a symmetric bilinear form $\beta'\colon V \times V \to \mathbf{R}$. Show that the differential $d\beta_X\colon T_XV = V \to T_{\beta(X)}\mathbf{R} = \mathbf{R}$ satisfies $d\beta_X(Y) = 2\beta'(X,Y)$.
2. Show that if $A = [a_{ij}]$ is a symmetric, positive definite, nondegenerate matrix (hence it defines a scalar product) then $\det A > 0$.

3. Write the formula for the induced Riemannian structure for the sphere, the torus, and a general surface of revolution.
4. Compute the Gauss and the mean curvature for the sphere, the torus and a general surface of revolution.
5. Give an example of a surface for which the Gauss curvature is: **(a)** negative, **(b)** zero, **(c)** positive, **(d)** constant negative, **(e)** nonconstant positive.
6. Consider the standard torus as in the previous examples. Determine the subsets of the torus for which the Gauss curvature is negative, zero, positive. ***Hint***: Since $\det[g_{ij}] > 0$ according to the exercise (2) all you need to check is the sign of the determinant of the second fundamental form.

10 Theorema Egregium

When computing the Gaussian curvature we consider the surface S as a geometric object in the Euclidean space. There is an important question if the curvature can be computed in terms of intrinsic data only. In other words, can the curvature be computed by a creature living entirely on the surface and being oblivious to the existence of the ambient Euclidean space? Gauss's ***Theorema Egregium*** says that the answer is yes.

There are many versions of the Theorema Egregium. We will provide a version that is related to functions π we saw in earlier sections. More precisely, we will prove that the Gauss curvature can be computed by measuring the circumference of circles on a surface. On the one hand the circumference of a circle is purely intrinsic and geometric feature of a surface considered as a metric space. On the other hand the Gauss curvature is an algebraic device which depends on the choice of the surface patch. The fact that it depends only on the metric of the surface is not obvious at all. Gauss was so fascinated by this fact that he gave this special name to the result. Nowadays we are more used to the fact that an abstract nonsense can have a nontrivial physical meaning.

10.1 The induced metric on a surface in $\mathbf{R}^3$

The length of a curve $c\colon [a,b] \to \mathbf{R}^3$ is equal to

$$\mathcal{L}(c) = \int_a^b \|\dot{c}(t)\| dt = \int_a^b \sqrt{\langle \dot{c}(t), \dot{c}(t) \rangle} dt.$$

Let $S \subset \mathbf{R}^3$ be a surface. The induced metric is defined by

$$\mathrm{d}(x,y) = \inf_c \mathcal{L}(c),$$

where the infimum is taken over all smooth curves $c\colon [a,b] \to S$ from x to y. Observe that $\|\dot{c}(t)\| = \mathbf{g}(\dot{c}(t), \dot{c}(t))$, where $\mathbf{g}$ is the induced Riemannian structure on S.

10.2 Covariant derivative and geodesics

Definition 10.3. Let $c\colon (a,b) \to S \subset \mathbf{R}^3$ be a curve and let $X\colon (a,b) \to TS$ be a smooth family of vectors tangent to S along c (that is, $X(t) \in T_{c(t)}S$).

Such an X is also called a vector field along c. The vector field

$$\frac{\nabla X}{dt} := \mathrm{pr}\left(\frac{dX}{dt}\right),$$

where $\mathrm{pr}\colon T_p\mathbf{R}^3 \to T_pS$ is the orthogonal projection is called the ***covariant derivative*** of X along c.

Proposition 10.4. *A curve $c\colon [a,b] \to S$ is a geodesic of the induced metric* d *if and only if*

$$\frac{\nabla \dot{c}(t)}{dt} = 0 \textit{ for all } t \in [a,b].$$

Proposition 10.5. *There exists a positive number $a > 0$ such that for every tangent vector $X \in T_pS$ there exists a geodesic $\gamma_X\colon (-a,a) \to S$ with $\dot{\gamma_X}(0) = X$.*

10.6 Exponential map and normal coordinates

Suppose that $X \in T_pS$ is a vector such that the corresponding geodesic γ_X is defined for $t = 1$. Let $U \subset T_pS$ be the set of all vectors with this property. Observe that this set U contains an open neighbourhood of the zero vector. The map $\exp\colon U \supset T_pS \to S$ defined by $\exp(X) = \gamma_X(1)$ is called the ***exponential map***.

Lemma 10.7. *The differential $d\exp\colon T_0(T_pS) = T_pS \to T_pS$ is equal to the identity.*

It follows from Lemma 10.7 (and the inverse function theorem) that there exist a negibourhood U of the zero vector in T_pS such that the restriction of the exponential map to U defines a surface patch. This patch is also known as ***normal coordinates***. More precisely, the surface patch is given by $f\colon (0,a)\times(0,2\pi) \to S$ by $f(r,\theta) = \exp(P(r,\theta))$, where $P\colon (0,a)\times(0,2\pi) \to T_pS$ are the polar coordinates in T_pS.

Observe that if $C(0,r) \subset T_pS = \mathbf{R}^2$ is a Euclidean circle of small radius r centered at the origin then its image $\exp(C(0,r))$ is a circle in S centered at p (of radius possibly different from r). This implies that the induced Riemannian structure has a very simple form in the polar coordinates (r,θ) on T_pS. Indeed, we have $\|\partial_r\| = 1$, $\langle \partial_r, \partial_\theta\rangle = 0$ and $\|\partial_\theta(r,\theta)\| = G(\exp(r,\theta))$.

Lemma 10.8. *In the polar coordinates the Gauss curvature has the following form*

$$K = -\frac{1}{\sqrt{G}} \frac{\partial^2 \sqrt{G}}{\partial r^2}.$$

Notice that G is a function on the surface patch.

Remark 10.9. The proof of this lemma is a lengthy computation which can be found in Spivak's "A Comprehensive Introduction to Differential Geometry II" on page 138.

Our task is to compute the circumference of a circle centered at $p \in S$ and express it using the Gauss curvature.

10.10 The circumference of a circle

Let (r, θ) be the polar coordinates in T_pS and let $c\colon [0, 2\pi] \to T_pS$ be the parametrisation of the circle of radius r. That is, $c(\theta) = (r, \theta)$. We are coing to compute the length of the curve $\exp \circ c$ which is a circle in S centered at p (the radius is in general different from r). Since

$$\begin{aligned}
\mathcal{L}(\exp \circ c) &= \int_0^{2\pi} \|d \exp_{(r,\theta)}(\partial_\theta)\| d\theta \\
&= \int_0^{2\pi} \sqrt{G}(\exp(r, \theta)) d\theta \\
&= \int_0^{2\pi} \left(r - \frac{K(\exp(r,\theta))r^3}{6} \right) d\theta + \int_0^{2\pi} o(r^3) d\theta \\
&= 2\pi \left(r - \frac{K(\exp(r,\theta))r^3}{6}) \right) + o(r^3) \\
&= 2\pi r + \frac{\pi K(\exp(r,\theta))r^3}{3} + o(r^3)
\end{aligned}$$

We thus have proved the following result.

Theorem 10.11 (Bertrand-Puiseux, 1848)**.** *Let $p \in S$ be a point on a surface in the Euclidean space $\mathbf{R}^3$. Let $C(p, r)$ be the circumference of the circle of radius r centered at p. Then*

$$C(p, r) = 2\pi r - \frac{\pi}{3} K(p) r^3 + o(r^3).$$

Corollary 10.12 (Theorema Egregium)**.** *The Gaussian curvature can be computed in terms if intrinsic geometry of the surface. More precisely,*

$$K(p) = \lim_{r \to 0} \frac{6\pi r - 3C(p,r)}{\pi r^3}.$$

Recall that the function $\pi_S \colon S \times [0, \infty) \to \mathbf{R}$ is defined by $\pi_S(p,r) = C(p,r)/2r$. It follows from the above version of the Theorema Egregium that if the Gauss curvature is positive (negative) at a point $p \in S$ then $\pi_S(p,r) < \pi$ ($\pi_S(p,r) > \pi$) for sufficiently small r. The case when the Gauss curvature is zero at a point is more complicated.

Corollary 10.13. *Let S and Σ be surfaces. If their Gauss curvatures are distinct then the surfaces cannot be isometric. More precisely, if $F \colon S \to \Sigma$ is a smooth map and $K_S(p) \neq K_\Sigma(F(p))$ then F is not a local isometry at p.*

Example 10.14. Let us discuss some more practical consequences of the Gauss theorem.

1. It is impossible to draw an accurately scaled map of the Earth on a flat piece of paper.
2. Bending a surface does not change the Gauss curvature. This has a nice experimental feature that a corrugated surface is rigid. If we want to lift a sheet of paper (almost) horizontally we create a little bent which makes the paper rigid. This is because the Gauss curvature is zero. On the other hand it is a product of principal curvatures. By creating a bent we make one of them nonzero and hence the other must remain zero. We use this fact when eatnig a wedge of pizza with hands. It is also the reason that a cardboard (or more precisely a corrugated fibreboard) is quite rigid.

♣

11 What does a surface look like?

Let $F\colon \mathbf{R}^2 \supset U \to \mathbf{R}$ be a smooth function. It defines a surface patch $f\colon U \to \mathbf{R}^3$ by $f(u,v) = (u, v, F(u,v))$ that parametrises the graph of F. Observe that in such a case the computation of the second fundamental form is relatively easy. Let us do a concrete example.

Example 11.1. Let $F(u,v) = \frac{1}{2}(au^2 + bv^2)$. We have

$$\mathbf{f}_u = (1, 0, au)$$
$$\mathbf{f}_v = (0, 1, bv)$$
$$\mathbf{g} = \begin{pmatrix} 1 + a^2u^2 & abuv \\ abuv & 1 + b^2v^2 \end{pmatrix}$$

$$\mathbf{f}_{uu} = (0, 0, a)$$
$$\mathbf{f}_{uv} = (0, 0, 0)$$
$$\mathbf{f}_{vv} = (0, 0, b)$$
$$\mathbf{II} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

Thus the Weingarten map at the point $(0,0)$ is equal to the second fundamental form and we get that the principal curvarures are a and b and hence the Gauss curvature $K(0,0) = ab$. If a and b are both nonzero and of the same sign the surface is an elliptic paraboloid, if they are both nonzero and of oposite signs then the surface is a hypebolic paraboloid. If either a or b is zero but not both then we have an parabolic cylinder and if they are both zero then we have a plane. ♣

The aim of this section is to prove that surfaces locally look ***more or less*** like those in the above example.

Let $f\colon U \to \mathbf{R}^3$ be a surface patch. We can apply an appropriate Euclidean translation an get that $f(0,0) = (0,0,0)$. We want to find a diffeomorphism $\psi\colon V \to U_0$ for some open neighbourhoods $V, U_0 \subset U$ of the origin such that the composed surface patch $f \circ \psi\colon V \to \mathbf{R}^3$ will have a nice form.

Let $\mathbf{n}_0 \in T_0\mathbf{R}^3$ be the unit vector normal to $f(U)$. The vectors $\mathbf{f}_1, \mathbf{f}_2$ and $\mathbf{n}_0$ form a basis of $\mathbf{R}^3$ and we want to express the surface patch with respect to this coordinate system. We thus have

$$f(u) = x(u)\mathbf{f}_1 + y(u)\mathbf{f}_2 + z(u)\mathbf{n}_0.$$

Since $\partial_i f(0) = \mathbf{f}_i$ we get that $\partial_1 x(0) = 1 = \partial_2 y(0)$ and $0 = \partial_2 x(0) = \partial_1 y(0) = \partial_i z(0)$. Moreover we have $x(0) = y(0) = z(0) = 0$. By computing the second partial derivatives we get that $\partial_{ij} z(0) = h_{ij}(0)$ are the coeffients of the second fundamental form. Since $z(0) = \partial_i z(0) = 0$ we obtain the Taylor series for z

$$z(u) = \frac{1}{2} \sum_{i,j} h_{ij}(u) u_i u_j + o(\|u\|^2).$$

Consequently, our surface patch has the following Taylor expansion

$$\begin{aligned} f(u) &= \sum_i u_i \partial_i f(0) + \frac{1}{2} \sum_{i,j} h_{ij}(0) u_i u_j \mathbf{n}_0 + o(\|u\|^2) \\ &= \sum_i u_i \mathbf{f}_i(0) + \frac{1}{2} \sum_{i,j} h_{ij}(0) u_i u_j \mathbf{n}_0 + o(\|u\|^2) \end{aligned}$$

which means that up to higher order terms the surface is equal to one of those in Example 11.1.

APPENDICES

12 Further definitions

12.1 Graphs

A ***graph*** is a pair (V, E), where V is a set and $E \subset V \times V$ is a symmetric relation. An element $v \in V$ is called a ***vertex*** and an element $(u, v) \in E$ is called an ***edge*** between vertices u and v.

The ***topological realisation*** of a graph (V, E) is the topological space $E \times [0,1]/\simeq$, where the relation $\simeq$ is the minimal equivalence relation satisfying $(u, v, 1) \simeq (u', v', 0)$ whenever $(v, u') \in E$. The topological realisation of a graph is also called a graph. And this is what usually we imagine as a graph (a collection of vertices joined by edges).

A ***path*** from v_0 to v_m in a graph (V, E) is a sequence of vertices $v_0, v_1, \ldots, v_n$ such that $(v_i, v_{i+1}) \in E$. A graph is called ***connected*** if for every pair of vertices there is a path between them.

A ***circuit*** in a graph is a path $v_0, v_1, \ldots, v_n, v_0$ such that $n \geq 2$ and the vertices v_i are pairwise distinct.

A ***tree*** is a connected graph without circuits.

12.2 Simply connected spaces

A path connected space X is called ***simply connected*** if every continuous map $f \colon \mathbf{S}^1 \to X$ extends to a map $\hat{f} \colon \mathbf{D}^2 \to X$. This means that the map $\hat{f}$ restricted to the boundary circle is equal to f.

A polygonal complex is simply connected if and only if every circuit made from edges bounds a subcomplex homeomorphic to a disc.

The following spaces are simply connected:

- the surface of a convex solid in $\mathbf{R}^3$;
- a sphere $\mathbf{S}^n$ for $n \geq 2$;
- $\mathbf{R}^n$;
- the hyperbolic plane;
- a tree.

The following spaces are not simply connected:

- a polygonal surface of positive genus;

- projective plane;
- the plane with a hole;
- the circle;
- a surface with nonempty boundary and not homeomorphic to a disc;

12.3 Curvature and torsion of a curve

Let $c\colon (a,b) \to \mathbf{R}^3$ be a differentiable curve parametrised by the arc length. The latter means that $\|\dot{c}(t)\| = 1$. The ***curvature*** of c at $t \in (a,b)$ is defined to be

$$\kappa(t) = \|\ddot{c}(t)\|.$$

It is the norm of the accelleration vector of c. For example, if c is a segment of a straight line then $\dot{c}$ is constant and hence $\ddot{c} = 0$. So the curvature measures the deviation of a curve of being straight.

The vector $\mathbf{n}(t) = \frac{\ddot{c}(t)}{\kappa(t)}$ is called the ***principal vector*** normal to c and $\mathbf{b}(t) = \dot{c}(t) \times \mathbf{n}(t)$ is called

12.4 Inversion

An ***inversion*** in a circle of radius r and centered at $p \in \mathbf{R}^2$ is a map $I\colon \mathbf{R}^2 - \{p\} \to \mathbf{R}^{-}\{p\}$ defined on the following figure.

An inversion has the following properties:

1. it is an involution, that is, $I \circ I = \mathrm{Id}$;
2. the image of a circle avoiding the centre is a circle avoiding the centre;
3. the image of a circle containing the centre is a straight line;
4. the image of a straight line is a circle containing the centre;
5. it preserves angles;

13 Trigonometric functions

$$
\begin{aligned}
\sinh(t) &= \frac{e^t - e^{-t}}{2} = \frac{e^{2t}}{2e^t - 1} \\
\cosh(t) &= \frac{e^t + e^{-t}}{2} = \frac{e^{2t}}{2e^t + 1} \\
e^t &= \sinh(t) + \cosh(t) \\
1 &= \sinh^2(t) - \cosh^2(t)
\end{aligned}
$$

14 Isometries

The proof of the following results is left as an exercise.

Theorem 14.1. *Let $f\colon (X, \mathrm{d}_X) \to (Y, \mathrm{d}_Y)$ be an isometry. Then the following statements hold:*

1. $\mathrm{Diameter}(X, \mathrm{d}_X) = \mathrm{Diameter}(Y, \mathrm{d}_Y)$;
2. *if $\gamma\colon [a,b] \to X$ is a rectifiable path then $f \circ \gamma$ is also rectifiable and $\mathcal{L}(\gamma) = \mathcal{L}(f \circ \gamma)$;*
3. *if $\gamma\colon [a,b] \to X$ is a geodesic then $f \circ \gamma$ is also a geodesic and $\mathcal{L}(\gamma) = \mathcal{L}(f \circ \gamma)$;*
4. *X is geodesic if and only if Y is geodesic;*
5. *if f is a polygonal isometry then for every vertex $v \in X$ the image $f(v)$ is a vertex in Y and $D(v) = D(f(v))$;*
6. *if $P \subset X$ is a geodesic polygon then $f(P)$ is a geodesic polygon in Y;*
7. *$\pi_X(p, r) = \pi_Y(f(p), r)$ for all $p \in X$*

15 Hints, answers and solutions

Warning 15.1. Reading a solution without trying solving the problem is pointless for at least two reasons. First, you learn less and second you are depriving yourself from finding your individual solutions which is often better then the one presented here. Also, I don't think that my slutions are the best.

Section 2

1. Read carefully all examples and provide more details and make relevant drawings if possible. For each example draw a few balls of various radii centered at various points.

   Make a list of things you don't understand and a list of mistakes and errors on this section.

2. Draw a circle of radius one centered at the origin of $(\mathbf{R}^2, \mathrm{d}_p)$ for $p \in \{1, 2, 3, 4, \infty\}$.

   Answer. https://en.wikipedia.org/wiki/Lp_space#The_p-norm_in_finite_dimensions

3. Draw two balls of radius one centered at the north pole of the unit sphere $\mathbf{S}^2$. The first with respect to the metric induced from the Euclidean metric on $\mathbf{R}^3$ and the second with respect to the intrinsic metric.

   Hint. The circle with respect to the intrinsic metric is included in the ball enclosed by the circle with respect to the induced metric.

4. For which r the unit circle $\mathbf{S}^1 \in \mathbf{R}^2$ endowed with the intrinsic metric is isometric to the circle of radius r with the induced metric?

   Hint. Look at points *between* two antipodal points in both spaces.

   Answer. For none.

   Solution. For none. Let $C(r) \subset \mathbf{R}^2$ be the circle of radius r centered at the origin. The diameter of $\mathbf{S}^1$ is equal to π and the diameter of $C(r)$ is equal to $2r$. Thus if the spaces were isometric we would have that $r = \frac{\pi}{2}$.

   Suppose that $\psi \colon \mathbf{S}^1 \to C\left(\frac{\pi}{2}\right)$ is an isometry. Since for both spaces the distance between x and $-x$ is equal to the diameter we obtain that ψ carries antipodal points to antipodal points. More precisely, we have that $\psi(-x) = -\psi(x)$.

Let $x = (0,1)$. If $y \in \mathbf{S}^1$ has the property that $d(x,y) = d(-x,y)$ then this distance is equal to $\frac{\pi}{2}$. Informally, a point between two antipodal points on $\mathbf{S}^1$ is within distance $\frac{\pi}{2}$ from both of them. In the induced metric, if $p \in C\left(\frac{\pi}{2}\right)$ is such that $d(\psi(x),p) = d(-\psi(x),p)$ this distance is equal to $\frac{\pi}{2}\sqrt{2}$ which implies that ψ cannot be an isometry. $\square$

5. Prove that the set of all isometries of a metric space is a group with respect to the composition.

   Hint. This is a standard exercise on verifying the defining axioms.

6. For each example in this section give an example of a nontrivial isometry of finite (infinite) order in the group of isometries, provided such an isometry exists.

7. Prove that $f\colon \mathbf{R} \to \mathbf{R}^2$ given by $f(t) = x + tv$ is an isometric embedding, provided v is a vector with (the Euclidean) norm equal to one and $x \in \mathbf{R}^2$ is a point.

   Solution: Let $x = (x_1, x_2)$ and $v = [v_1, v_2]$. Let $s, t \in \mathbf{R}$. The following computation shows that f is an isometric embedding.

$$\begin{aligned} \mathrm{d}_{\mathbf{R}^2}(f(s), f(t)) &= \mathrm{d}_{\mathbf{R}^2}((x_1 + sv_1, x_2 + sv_2), (x_1 + tv_1, x_2 + tv_2)) \\ &= \sqrt{(s-t)^2 v_1^2 + (s-t)^2 v_2^2} \\ &= |s-t|\|v\| = |s-t| = \mathrm{d}_{\mathbf{R}}(s,t). \end{aligned}$$

   $\square$

8. Give an explicit formula for the intrinsic metric on the unit sphere $\mathbf{S}^2 \subset \mathbf{R}^3$. Generalise your solution to $\mathbf{S}^n \in \mathbf{R}^{n+1}$.

   Hint. Do it for $\mathbf{S}^1$ first.

   Answer. $d(x,y) = \arccos\langle x, y\rangle$, where $\langle\ ,\ \rangle$ denotes the standard scalar product in $\mathbf{R}^{n+1}$.

9. Prove that $f\colon \mathbf{S}^1 \to \mathbf{S}^2$ given by $f(\cos t, \sin t) = (\cos t, \sin t, 0)$ is an isometric embedding. Prove that the inclusion of any great circle is an isometric embedding.

   Hint. The first part is a direct computation (using the formula from the previous exercise). The second uses the fact that any rotation of $\mathbf{R}^3$ with respect to an axis through the origin restricts to an isometry of $\mathbf{S}^2$.

10. Show that the inclusion $\iota\colon \mathbf{S}^1 \to \mathbf{R}^2$ is not an isometric embedding. Show that it is Lipschitz and compute the Lipschitz constant.

Solution. We have, for example, that $\mathrm{d}_{\mathbf{R}^2}((1,0),(0,1)) = \sqrt{2}$ and that $\mathrm{d}_{\mathbf{S}^1}((1,0),(0,1)) = \pi/2$, which shows tha the inclusiuon ι is not an isometric embedding.

We shall show that for every $x, y \in \mathbf{S}^1$ we have

$$\mathrm{d}_{\mathbf{R}^2}(x,y) \leq \mathrm{d}_{\mathbf{S}^1}(x,y).$$

Assume without loss of generality that $x = (1,0)$ and $y = (\cos(t), \sin(t))$. Then we have that $\mathrm{d}_{\mathbf{R}^2}(x,y) = \sqrt{2 - 2\cos(t)}$ and $\mathrm{d}_{\mathbf{S}^1}(x,y) = t$, so we need to show that $\sqrt{2 - 2\cos(t)} \leq t$ for $t \in [0,\pi]$ which is enough. Geometrically we are proving an obvious fact that a cord is shorter than the corresponding arc.

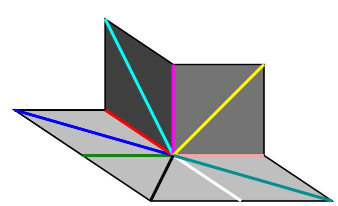
Observe that the derivative of the function $f\colon [0,\infty) \to \mathbf{R}$ given by $f(t) = t^2 + 2\cos(t) - 2$ is equal to $f'(t) = 2t - 2\sin(t)$. In particular $f'(0) = 0$ and $f'(t) > 0$ for $t > 0$. Thus the function f is increasing and $f(0) = 0$ which implies that it is nonegative $f(t) \geq 0$. We thus get that $t^2 \geq 2 - 2\cos(t)$ as required. □

Section 3

1. Read carefully all examples and provide more details and make relevant drawings if possible.
2. Let $\mathbf{R}^2$ be the Euclidean plane and let $\gamma\colon \mathbf{R} \to \mathbf{R}^2$ be defined by $\gamma(t) = xt + y$, where $x = (x_1, x_2), y = (y_1, y_2) \in \mathbf{R}^2$. Prove that γ is a geodesic if and only if $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2} = 1$. State and prove the analogous criterion for the d_p metric on $\mathbf{R}^2$.

   Hint. This is a standard verification. The condition for the second part is $\|x\|_p = 1$.

3. Consider the space pictured right and compute the angles of between the black geodesic and every other geodesic drawn. Which concatenations are geodesics?



Compute the distance between any two vertices.

Answer. The angles are

$$
\begin{aligned}
\angle(\text{black}, \text{white}) &= \frac{\pi}{4} & \angle(\text{black}, \text{greenish}) &= \frac{\pi}{2} \\
\angle(\text{black}, \text{pink}) &= \frac{3\pi}{4} & \angle(\text{black}, \text{yellow}) &= \pi \\
\angle(\text{black}, \text{purple}) &= \pi & \angle(\text{black}, \text{turquoise}) &= \pi \\
\angle(\text{black}, \text{red}) &= \frac{3\pi}{4} & \angle(\text{black}, \text{blue}) &= \frac{\pi}{2} \\
\angle(\text{black}, \text{green}) &= \frac{\pi}{4}
\end{aligned}
$$

4. What is the circumference of a circle of radius r centered in the interior vertex of the space from the previous part.

   Answer. $\frac{5\pi}{2}r$.

5. Let γ_i for $i = 1, 2, 3$ be geodesics starting from the same point of a metric space. Suppose that the angle $\angle(\gamma_1, \gamma_2) = \angle(\gamma_2, \gamma_3) = 3\pi/4$. What can you say about the angle $\angle(\gamma_3, \gamma_1)$?

   Answer. Nothing. For each $\alpha \in [0, \pi]$ you can construct a space (similar to the one in Exercise 3) for which $\angle(\gamma_3, \gamma_1) = \alpha$.

6. Draw all geodesics between diagonal vertices A, B on the surface of the unit cube.

   Hint. If you find one geodesic γ then you can obtain more geodesic as the compositions of γ with isometries preserving the vertices A, B. This way you should find 6 geodesic which are the shortest. Drawing all geodesics is impossible because there are infinitely many of them. However, determining all of them (for example, in terms of the angle such a geodesic makes with an edge containig the starting vertex) is possible and not too difficult.

7. Draw balls of radius $r \in [0, 3]$ centered at a vertex or at the centre of a face of the surfac eof the unit cube. Compute the circumference in several cases.

   Hint. A ball of suitably big radius centered at the center of a face is a diamond-like shape contained in the oposite face.

8. Let X be the lateral surface of a right cone of directrix d and unit generatrix. Let p and q be points in the base circle. Draw a geodesic between p and q and compute the distance between them.

   Solution. Let A be the apex of the cone and let 0 be a point fixed on the base of the cone. Let $p_1 = |Ap|$ and p_2 be the (positive) angle from the ray $A0$ and the ray Ap (and define q_1 and q_2 analogously).

   Cut the cone open along a ray Ap and flatten in on the plane. The distance between p and q is then equal to the length of the shortest seqment joining the points in the flattened cone (there may be two such segments). We obtain a triangle with vertices p, q, A with the angle at A equal to $|q_2 - p_2|$ or $2\pi - |q_2 - p_2|$ and sides from containing A of lenghts p_1 and q_1. The third side is the distance we are looking for and it can be computed by applying the cosine formula.

9. Let $X = \mathbf{S}^1 \times \mathbf{R} \subset \mathbf{R}^3$ be the infinite cylinder based on the unit circle. Give a formula for the intrinsic metric and investigate geodesics.

   Solution. Let (w, s) and (z, t) be two points on the cylinder $\mathbf{S}^1 \times \mathbf{R}^2$. Cut the cylinder open along $\{w\} \times \mathbf{R}$ and flatten it on the plane. The distance we are looking for is then the Euclidean distance and it is equal to $\sqrt{\mathrm{d}_{\mathbf{S}^1}(w,z)^2 + (s-t)^2}$, where $d_{\mathbf{S}^1}$ denotes the intrinsic metric on the unit circle.

   By flattening the cylinder we see that the geodesic are straigth lines. Formally, they are compositions of straight lines on the Eulclidean plane $\mathbf{R}^2$ with the projection $\mathbf{R}^2 \to \mathbf{S}^1 \times \mathbf{R}$ given by $(x, y) \mapsto (x_{\mathrm{mod} 2\pi}, y)$.

10. State and prove the converse of Lemma 3.18.

    Solution. If the angle between two geodesics $\gamma_1, \gamma_2\colon [0,1] \to X$ is equal to π then $\gamma\colon [-1,1] \to X$ defined by

    $$\gamma(t) = \begin{cases} \gamma_1(-t) & t \leq 0 \\ \gamma_2(t) & t \geq 0 \end{cases}$$

    is a geodesic.

11. Let $0 = t_0 < t_1 < \ldots < t_n < \ldots$ be an infinite sequence of numbers converging to 1. Let $X = [-1, 1]$ and let $c\colon [0,1] \to X$ be any path stisfying the following conditions: $c(0) = 0$ and $c(t_n) = \sum_{k=1}^{n} (-1)^{k+1}/k$. Show that c is not rectifiable.

    Hint. Its length is bounded below by the sum of harmonic series.

12. The plane $\mathbf{R}^2$ equipped with the d_1 metric is called the ***Manhattan plane***. Recall that, $\mathrm{d}_1(x,y) = \mathrm{d}_1\left((x_1,x_2),(y_1,y_2)\right) = |x_1 - y_1| + |x_2 - y_2|$, where $x = (x_1, x_2)$ and $y = (y_1, y_2)$ are points on the plane $\mathbf{R}^2$. Moreover, let $\|x\|_1 = \mathrm{d}_2(x, 0) = |x_1| + |x_2|$.

    (a) Let $\gamma\colon \mathbf{R} \to \mathbf{R}^2$ be given by $\gamma(t) = xt + y$. Show that γ is a geodesic if and only if $\|x\|_1 = 1$.

    (b) Deduce from the previous item that the Manhattan plane is a geodesic metric space.

    Solution. For $x, y \in \mathbf{R}^3$ a geodesic between them is, for example, $\gamma\colon [0, d_1(x,y)] \to \mathbf{R}^2$ given by

    $$\gamma(t) = \frac{y - x}{\|y - x\|_1} t + x.$$

    (c) Let $\gamma, \gamma'\colon [0,\infty) \to \mathbf{R}^2$ be defined by $\gamma(t) = (t, 0)$ and $\gamma'(t) = (at, bt)$, where $a, b \geq 0$ and $a + b = 1$. Show that the angle between γ and γ' does not exist in general (for example, when $a = b = \frac{1}{2}$).

    Hint. Show that the limit defining the angle does not exist by computing it for $s = t$ and for $s = 2t$.

    (d) The ***upper angle*** between geodesics γ and γ' is defined similarly to the angle with the limit replaced by *limit superior*. Compute the upper angle between geodesics from the previous item.

    Answer. $\angle(\gamma, \gamma') = \arccos(a - b)$.

    (e) How many geodesics are there between $(0,0)$ and $(1,1)$? How many of them are distance minimising?

    Answer. Infinitely many. Infinitely many.

    (f) How many geodesics are there between $(0,0)$ and $(1,0)$? How many of them are distance minimising?

    Answer. Infinitely many. One.

Section 4

1. Read carefully all examples and provide more details and make relevant drawings if possible.
2. Let X be a square surface obtained as follows. Take a solid cube subdivided into 27 equal cubes. Remove the central cube and each cube containing the centre of a face of the big cube (seven cubes removed). Let

X be the surface of the remaining solid. Compute the Euler characteristic $\chi(X)$.

Answer. $\chi(X) = -8$.

3. Construct a triangular closed surface with the Euler characteristic equal to zero.

   Hint. This can be done in many ways. Take, for example, two nets of a regular tetrahedron and try to construct a net of a surface in which 6 triangles meet at each vertex.

4. Let X be a closed polygonal surface such that all faces of X are k-gons for some $k > 2$ and m faces meet at each vertex. Proof the following statements:

   (a) $2E = kF$; (b) $2E = mV$; (c) $kF = mV$.

   Solution of (a). Each k-gonal face contributes k-edges to the total number of edges. Since an edge belongs to exactly two faces, each each has been counted twice in the previous count. This shows that $2E = kF$. This argument proves more. If X is any closed surface then $2E = \sum_{k=3}^{\infty} kF_k$, where F_k denotes the number of k-gonal faces.

   Hint. The remaining solutions are similar.

5. Take a polygonal surface of genus g (for example $g = 2$), remove one face, subdivide, and glue in the Möbius band. What is the Euler characteristic of the resulting surface?

   Answer. $1 - 2g$.

6. Take two Möbius bands and glue them together by identifying their boundaries. What is the Euler characteristic of the resulting surface?

   Answer. 0.

7. The surface from the previous part has zero Euler characteristic. Prove that it is not homeomorphic to a torus.

   Hint. This is not so easy without sophisticated tools. Hoever, the intuition is the following. Our surface contains the Möbius band as a subsurface. If it was homeomorphic to the torus then the torus would contain a Möbius band. If we consider the torus $\mathbf{T}^2$ as a surface in $\mathbf{R}^3$ then the complement $\mathbf{R}^3 \backslash \mathbf{T}^2$ has two connected components. This implies that the torus has two sides corresponding to this connected components (intuition: the torus can be painted white from the inside and black from the outside). If the torus contained the Möbius band it would have only one side (intuition: if you start painting the Möbius band with white then you paint is all white).

8. Suppose that X is a triangular closed surface such that six triangles meet at each vertex. What is the Euler charateristic of X?

   Answer. 0.

9. Suppose that X is a triangular closed surface such that k triangles meet at each vertex. What is the Euler charateristic of X?

   Solution. We have $3F = 2E = kV$. Thus

   $$F - E + V = F - 3F/2 + 3F/k = F(3/k - 1/2).$$

   This is all we can say. This identity has many interesting consequences. For example, if $k > 6$ the Euler characteristic is negative; if k is odd then F is even etc.

10. Construct a closed square surface made from 12 squares in which 6 squares meet at each vertex. Compute the Euler characteristic of the constructed polygonal surface.

    Answer. $\chi = -6$.

11. Try to give a rigorous definition of the above surface.

12. Let $\mathbf{C}$ be the surface of the unit cube and let $A \in \mathbf{C}$ be a vertex. Is there a geodesic $\gamma\colon [a, b] \to \mathbf{C}$ such that $\gamma(a) = \gamma(b) = A$?

13. The same question as in the previous exercise for the case of surfaces of other Platonic solids (tetrahedron, octahedron, dodecahedron and icosahedron).

Section 5

(5) We have $7F = 2E = 3V$ and we compute that $F - E + V = 3V/7 - 3V/2 + V = -V/14$.

(6) By the same reasoning as above we get that the Euler characteristic is equal to $(6 - k)V/2k$.

Section 6

(6) The answer is yes. Cut the cone open along a generatrix and flatten on the plane. Consider a disc of radius one centered at a vertex on the unit cube. Cut it open along the edge and flatten on the plane. It is isometric to the flatten cone.

(7) Again, cut open along generatrices, flatten on the plane. It is now clear that some neighbourhoods are isometric if and only if their directrices are equal.

Section 7

(3) Let $P = (p, q)$ and let $X = (x, y)$. Then

$$\mathrm{I}_{P,r}(x, y) = \frac{r^2}{(x-p)^2 + (y-q)^2}(x - p, y - q).$$

(5) See the section ***Decomposition and elementary properties*** in the wipedia article on Möbius transformation.

(7) The hyperbolic angle is equal to the Euclidean angle. Hence to compute the hyperbolic angle we draw tangent (Euclidean straght) lines to the sides of the triangle at a vertex and compute the Euclidean angle between them.

Section 8

(2) The points $A, B, C \in \mathbf{R}^3$ define uniquely a plane $P \subset \mathbf{R}^3$. Moreover, there is a unique circle on the plane P though the points A, B, C. Let $O \in P$ be the centre of this circle and let $L \subset \mathbf{R}^3$ be a line perpendicular to P and intersecting P at O. Any point on L is equidistant from A, B and C and defines uniquely a sphere in $\mathbf{R}^3$ containing the points A, B, C.

(3) Use the previous problem. Let $D \in \mathbf{R}^3$ be the fourth point. Find a point Q on the line L such that it is equidistant from A and D. This is the centre of the sphere we are looking for.

Section ??

Section 9

Section 10