Geometry

July 5, 2008

Outline

The basic themes of this course are metrics (distances), lengths of curves, geodesics (the curve of minimum length between two fixed points), symmetries and groups of symmetries, and curvature (which then links into topology by the Gauss-Banet Theorem

Geometries

We shall consider the Euclidean, spherical, and torus geometries, then the hyperbolic plane, embedded surfaces in $\mathbb{R}^n$, and finally abstract surfaces.

1 Euclidean Geometry

We consider $\mathbb{R}^n$ equipped with the standard inner product $(x, y) = x \cdot y$ and distance function d. A map $f : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry (or rigid motion) of $\mathbb{R}^n$ if $d(f(P), f(Q)) = d(P, Q) \forall P, Q \in \mathbb{R}^n$. Recall than an $n \times n$' matrix is orthogonal if $AA^T = A^TA = I$; since $(Ax, Ay) = (Ax)^T(Ay) = x^TA^TAy = (x, A^TAy) = (A^TAx, y)$, A orthogonal $\Leftrightarrow (Ax, Ay) = (x, y) \forall x, y \in \mathbb{R}^n$. Since $(x, y) = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2)$, A is orthogonal $\Leftrightarrow \|Ax\| = \|x\| \forall x$. Thus if $f(x) = Ax + b$ then $d(f(x), f(y)) = \|A(x - y)\|$ so f is an isometry $\Leftrightarrow A$ is orthogonal.

1.1 Theorem

Any isometry $f : \mathbb{R}^n \to \mathbb{R}^n$ is of this form: let $e_1, \ldots, e_n$ be the standard basis, let $f(0) = b, f(e_i) - b = a_i$ for $i = 1, \ldots, n$, then $\|a_i\| = \|f(e_i) - f(0)\| = d(f(e_i), f(0)) = d(e_i, 0) = 1 \forall i$. For $i \neq j$ $(a_i, a_j) = -\frac{1}{2}(\|a_i - a_j\|^2 - \|a_i\|^2 - \|a_j\|^2) = -\frac{1}{2}(\|f(e_i) - f(e_j)\|^2 - 2) = -\frac{1}{2}(\|e_i - e_j\|^2 - 2) = 0$. Let A be the orthog mat w/ columns given by the orthonormal basis $a_1, \ldots, a_n$. Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be the isometry given by $g(x) = Ax + b$. $g(x) = f(x)$ for $x = 0, e_1, \ldots, e_n$. Now g has an inverse $g^{-1}(x) = A^{-1}(x - b) = A^T(x - b)$ so $h = g^{-1} \circ x$ is an isometry fixing $0, e_1, \ldots, e_n$.

Now, using a common technique, we claim $h = \iota$ and hence $f = g$; for general $x = \sum x_ie_i$ let $y = \sum_i y_ie_i = h(x)$; observe $d(x, e_i)^2 = \|x\|^2 + 1 - 2x_i, d(x, 0)^2 = \|x\|^2, d(y, e_i)^2 = \|y\|^2 + 1 - 2y_i, d(y, 0)^2 = \|y\|^2$. Since h is an isometry s.t. $h(0) = 0, h(e_i) = e_i \forall i, h(x) = y$ we have $\|x\|^2 = \|y\|^2$ and then $x_i = y_i \forall i$ and we are done.

) $Isom(\mathbb{R}^n)$ is a group.

Example

Reflections in affine hyperplanes: if $H \subset \mathbb{R}^n$ is an affine hyperplane defined by $\vec{u}\cdot\vec{x} = c$ for some unit vector $\vec{u}$ and constant c, define R_H by $\vec{x} \mapsto \vec{x} - 2(\vec{x}\cdot\vec{u} - c)\vec{u}$ is an isometry (see ExS1Q1), called the reflection in H; note $R(\vec{x}) = \vec{x}\forall\vec{x} \in H$. If $\vec{x} \in H$, $R(\vec{a} + t\vec{u}) = \vec{a} - t\vec{u}$.

Conversely, if S is an isometry fixing H (pointwise) and $\vec{a} \in H$ then the conjugate $R = T_{-\vec{a}} S T_{\vec{a}}$ (where $T_{\vec{a}}$ is translation by $\vec{a}$ i.e. $T_{\vec{a}}(\vec{x}) = \vec{x} + \vec{a}$) fixes the hyperplane $H' = T_{-\vec{a}}H$ through the origin. If H is given by $\vec{x}\cdot\vec{u} = c$ then this H' is given by $\vec{x}\cdot\vec{u} = 0$, since $c = \vec{a}\cdot\vec{u}$. Then $(R\vec{u}, \vec{x}) = (R\vec{u}, R\vec{x}) = (\vec{u}, \vec{x})\forall\vec{x} \in H$, so $R\vec{u} = \lambda\vec{u}$ for some λ. But $\|R\vec{u}\|^2 = \|\vec{u}\|^2 = 1$ so $\lambda^2 = 1$ and $\lambda = \pm 1$; $R = \iota$ or $R = R_{H'}$ by 1.1, so $S = \iota$ or $S = T_{\vec{a}} R_{H'} T_{-\vec{a}}$; $\vec{x} \mapsto \vec{x} - \vec{a} \mapsto \vec{x} - \vec{a} - 2(\vec{x}\cdot\vec{u} - \vec{a}\cdot\vec{u})\vec{u} \mapsto \vec{x} - 2(\vec{x}\cdot\vec{u} - c)\vec{u}$ i.e. $S = R_H$.

On the ExS we see that any isometry can be decomposed as a product of reflections.

$\exists$ a natural subgp of $\mathrm{Isom}(\mathbb{R}^n)$, the isometries fixing the origin, which by 1.1 is isomorphic to $O(n) = O(n, \mathbb{R})$, the orthog gp of $n \times n$ mats. If $A \in O(n)$, $\det(A)^2 = \det(AA^T) = 1 \Rightarrow \det A = \pm 1$. The subgp of $O(n)$ consisting of elts w/ det $= 1$ is called the special orthog gp $SO(n)$ e.g. $SO(2) \subset O(2)$. $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(2) \Leftrightarrow a^2 + c^2 = 1, d^2 + d^2 = 1, ab + cd = 0$. We set $a = \cos\theta, c = \sin\theta, b = -\sin\phi, d = \cos\phi$, then $\tan\theta = \tan\phi \Rightarrow \phi = \theta$ or $\theta \pm \pi$. In the first case $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, a rotation through argle θ, and in the second $A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$, a reflection in the line at angle $\frac{\theta}{2}$ above the x axis.

1.2 The group $O(3, \mathbb{R})$

If $\det A = 1$ then $\det(A - I) = \det(A^T - I) = \det A(A^T - I) = \det(I - A) \Rightarrow \det(A - I) = 0$ so 1 is an eigenvalue, i.e. $\exists v_1$ (with, scaling, $\|v_1\| = 1$) s.t. $Av_1 = v_1$. Let $W = \langle v_1 \rangle^\perp$; if $w \in W$ then $(Aw, v_1) = (Aw, Av_1) = (w, v_1) = 0$ so $A(W) \subset W$ and $A\mid_W$ is a rotation of the 2D space W; if $\{v_2, v_3\}$ is an orthonormal basis for W the action of A on $\mathbb{R}^3$ wrt the ON basis $\{v_1, v_2, v_3\}$ is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$.

Now suppose $\det A = -1$ there is an ON basis wrt which $-A$ is a rotation of the above form, so $A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}$ wrt some ON basis, where $\phi = \theta + \pi$, i.e. A is a rotated reflection; in the special case $\phi = 0$ A is a pure reflection.

1.3 Curves in $\mathbb{R}^3$

Defn

A <u>curve</u> Γ in $\mathbb{R}^n$ is a cnts function $\Gamma : [a, b] \to \mathbb{R}^n$. Consider dissections $\mathcal{D} : a = t_0 < t_1 < \cdots < t_N = b$ of $[a, b]$ w/ N arbitrary; set $P_i = \Gamma(t_i)$, $s_{\mathcal{D}} = \sum \|\overrightarrow{P_i P_{i+1}}\|$.

The length l of $\mathcal{D}$ is defined to be $\sup_{\mathcal{D}} s_{\mathcal{D}}$ if it exists.

If $\mathcal{D}'$ is a refinement of $\mathcal{D}$ the triangle inequality $\Rightarrow s_{\mathcal{D}} \leq s_{\mathcal{D}'}$, so (if it exists) $l = \lim_{\text{mesh}(\mathcal{D})\to 0} s_{\mathcal{D}}$, where $\text{mesh}(\mathcal{D}) = \max_i(t_i - t_{i-1})$.

Proposition 1.4

If Γ is cntsly diffable then length $\Gamma = \int_a^b |\Gamma'(t)|dt$: write $\Gamma(t) = (f_1(t), f_2(t), f_3(t))$, then $\forall s \neq t \in [a,b] \frac{\Gamma(t)-\Gamma(s)}{t-s} = (f_1'(\xi_1), f_2'(\xi_2), f_3'(\xi_3))$ for some $\xi_i \in (s,t)$; since the f_i are unifly cnts on $[a,b], \forall \epsilon > 0 \exists \delta > 0 : |t-s| < \delta \Rightarrow |f_i'(\xi_i) - f_i'(\xi)| < \frac{\epsilon}{3} \forall \xi \in (s,t)$, so if $|t-s| < \delta$ then $|\Gamma(t) - \Gamma(s) - (t-s)\Gamma'(\xi)| < \epsilon(t-s) \forall \xi \in (s,t)$; now take a dissection $\mathcal{D}$ w/ $\text{mesh}(\mathcal{D}) < \delta$ and by the triangle inequality $\sum(t_i - t_{i-1})|\Gamma'(t_i - 1)| - \epsilon(b-a) < s_{\mathcal{D}} < \sum(t_i - t_{i-1})|\Gamma'(t_i - 1)| + \epsilon(b-a)$; since $|\Gamma'(t)|$ is cnts it is integrable and $\sum(t_i - t_{i-1})|\Gamma'(t_{i-1})| \to \int_a^b : \Gamma'(T)|dt$ as $\text{mesh}(\mathcal{D}) \to 0$, so length $\Gamma := \lim_{\text{mesh}(\mathcal{D})\to 0} s_{\mathcal{D}} = \int_a^b |\Gamma'(t)|dt$; for (piecewise) cntsly diffable curves Γ we can obtain the length by integrating $|\Gamma'|$.

Now we look at $S = S^2 \subset \mathbb{R}^3$, the unit sphere centred on th origin. We are interested in great circles, defined as the intersection of S with a plane through $\vec{0}$. Through any two non-antipodal pts $P, Q \in S, \exists!$ great circle (also called spherical lines).

Defn

Distance $d(P,Q)$ is defined as the length of the shorter of the two segments PQ along the great circle (or π for antipodal points). Define $\vec{P} = \overrightarrow{OP}, \vec{Q} = \overrightarrow{OQ}$, then $d(P,Q) = \cos^{-1}(\vec{P} \cdot \vec{Q})$.

Say we have a "spherical triangle" ABC; A, B, C pts on the sphere, AB, BC, AC spherical line segments of length $< \pi$. Set $\vec{A} = \overrightarrow{OA}$ etc. Let $\vec{n}_1 = \frac{\vec{C}\times\vec{B}}{\sin a}, \vec{n}_2 = \frac{\vec{A}\times\vec{C}}{\sin b}, \vec{n}_3 = \frac{\vec{B}\times\vec{A}}{\sin c}$ [where a is the length of the side opposite A etc]. The angles of the spherical triangle are the angles between the defining planes of the sides, taking the angle with $0 < \alpha < \pi$ [α being the angle at A etc.]. Note that the angle between $\vec{n}_2$ and $\vec{n}_3$ is $\pi + \alpha$ (or $\pi - \alpha$), so $\vec{n}_2 \cdot \vec{n}_3 = \cos(\pi + \alpha) = -\cos\alpha$ (and similarly for the other $\vec{n}_i$).

Thm 2.1 (cosine formula)

$\sin a \sin b \cos\gamma = \cos c - \cos a \cos b$: recall $(\vec{C} \times \vec{B}) \cdot (\vec{A} \times \vec{C}) = (\vec{A} \cdot \vec{C})(\vec{B} \cdot \vec{C}) - (\vec{C} \cdot \vec{C})(\vec{B} \cdot \vec{A}) = (\vec{a} \cdot \vec{C})(\vec{B} \cdot \vec{C}) - (\vec{B} \cdot \vec{A})$ since $\|C\| = 1$, so $-\cos\gamma = \vec{n}_1 \cdot \vec{n}_2 = \frac{(\vec{C}\times\vec{B})\cdot(\vec{A}\times\vec{C})}{\sin a \sin b} = \frac{(\vec{A}\cdot\vec{C})(\vec{B}\cdot\vec{C}) - (\vec{B}\cdot\vec{A})}{\sin a \sin b} = \frac{(\cos b \cos a - \cos c)}{\sin a \sin b}$.

Corollary (Pythagoras)

When $\gamma = \frac{\pi}{2}, \cos c = \cos a \cos b$.

Thm 2.3 (sin formula)

$\frac{\sin a}{\sin\alpha} = \frac{\sin b}{\sin\beta} = \frac{\sin c}{\sin\gamma}$: use $(\vec{A} \times \vec{C}) \times (\vec{C} \times \vec{B}) = (\vec{C} \cdot (\vec{A} \times \vec{B}))\vec{C}$; the LHS$=-(\vec{n}_1 \times \vec{n}_2) \sin a \sin b$ where $\vec{n}_1 \times \vec{n}_2 = \vec{C} \sin\gamma$, so $\vec{C} \cdot (\vec{A} \times \vec{B}) = \sin a \sin b \sin\gamma$. The triple product is invariant under permutations so we have $\sin a \sin b \sin\gamma = \sin\alpha \sin b \sin c = \sin a \sin\beta \sin c$.

Assuming $a, b, c < \pi$, applying (2.1) we have $\cos c = \cos a \cos b + \sin a \sin b \cos \gamma$. Unless $\gamma = \pi$, i.e. C lies on the line segment AB and hence $c = a + b$, $\cos c > \cos a \cos b - \sin a \sin b = \cos(a + b) \Rightarrow c < a + b$.

Corollary (2.4) (Triangle inequality)

For $P, Q, R \in S^2$, $d(P, Q) + d(Q, R) \geq d(P, R)$, with equality iff Q is on the line segment PR (of shorter length). If $d(P, R) = \pi$ i.e. P, R antipodal, then the line PQ also pallel through R so $d(P, R) = D(P, Q) + d(Q, R)$.

Proposition 2.5

Given a curve $\Gamma : [0, 1] \to S^{`}2$ on S joining P to Q, $l = \text{length}(\Gamma) \geq d(P, Q)$; moreover, if $l = d(P, Q)$ the image of Γ is the spherical line segment PQ on S. An outline of the pf is: given a dissection $\mathcal{D}$ of $[0, 1]$ AS $0 = t_0 < t_1 < \cdots < t_N = 1$, let $P_i = \Gamma(t_i), S_{\mathcal{D}} = \sum_{i=1}^{N} |\overrightarrow{P_{i-1}P_i}| < S'_{\mathcal{D}} = \sum_{i=1}^{N} d(P_{i-1}, P_i)$. Since $\frac{\sin\theta}{\theta} \to 1$ as $\theta \to 0$, we have $2\sin\theta \leq 2\theta \leq (1 + \epsilon)2\sin\theta$ for θ sufficiently small, i.e. $S'_{\mathcal{D}} \leq (1 + \epsilon)S_{\mathcal{D}}$ for mesh $\mathcal{D}$ sufficiently small, so $S = S'$ in limit. By repeated triangle inequality $d(P, Q) \leq S'_{\mathcal{D}} \forall \mathcal{D}$, so $l \geq d(P, Q)$.

If Γ is of length $l = d(P, Q)$ then $\forall t \in [0, 1]$, $d(P, Q) = l = \text{length}\, \Gamma \mid_{[0,t]} + \text{length}\, \Gamma \mid_{[t,1]} \geq d(P, \Gamma(t)) + d(\Gamma(t), Q) \geq d(P, Q)$, so $d(P, Q) = d(P, \Gamma(t)) + d(\Gamma(t), Q) \forall t$, so $\Gamma(t)$ is on the shorter line segment $PQ \forall t$, so the image of Γ is the line segment as required.

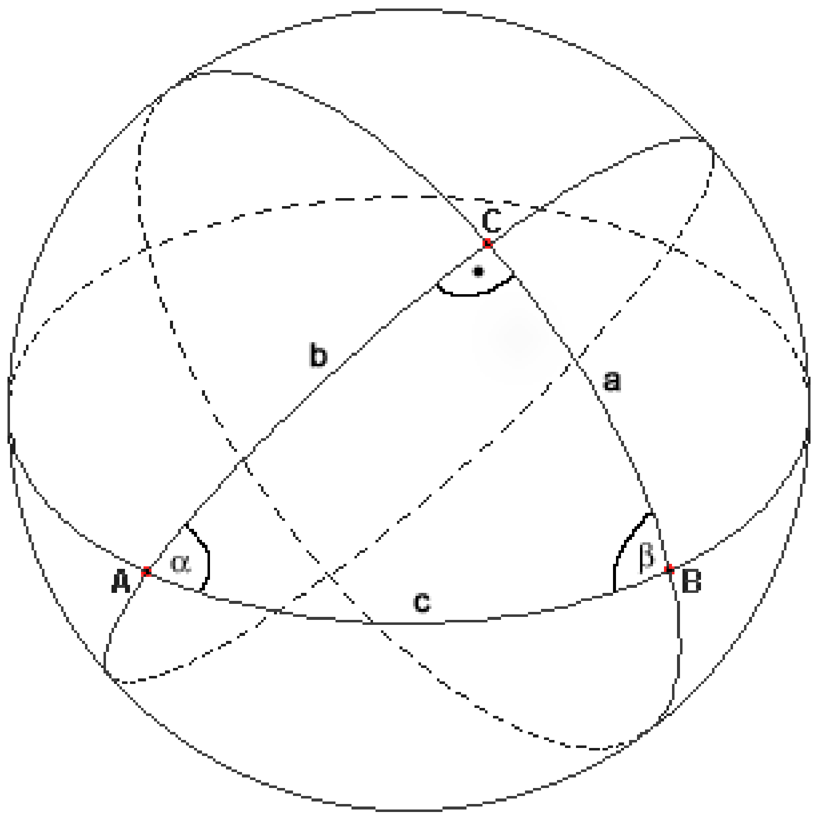
Rk 2.9

So any curve of minimum length joining P, Q is a spherical line segment. Moreover, from the above $\text{length}\, \Gamma \mid |[0, t] = d(P, \Gamma(t)) \forall t$ so the parameterization is monotonic. For now we call such Γ (minimizing) geodesics.

1.4 Area of spherical triangles (Gauss-Bonnet)

Proposition 2.10

If Δ is a spherical triangle w/ angles α, β, γ then its area is $(\alpha + \beta + \gamma) - \pi$: Def a double lune w/ angle α on S is the area cut out by two planes passing through antipodal pts of S, w/ the angle between the planes α; the area of this is 4α. A spherical triangle $\Delta = ABC$ is the intersection of three, or equivalently merely two, single lunes; Δ and its antipodal triangle Δ' are in all 3 of the double lunes (w/ areas $4\alpha, 4\beta, 4\gamma$), but any other point of the sphere is in only one; the diagram may make this clearer:



Thus $4(\alpha+\beta+\gamma) = 4\pi + 2 \times 2 \times A$ where $A =$ the area of Δ (= that of Δ'), since 4π is the total area of S, and we have the result.

Rk 2.11

For a spherical triangle, $\alpha+\beta+\gamma > \pi$; the limit as the area of $\Delta \to 0$ is $\alpha+\beta+\gamma = \pi$, the Euclidean case.

If M is a convex n-gon on S (for $n \geq 3$) (i.e. $\forall P, Q \in M$ the shorter line segment joining P and Q lies in M) and the interior angles are $\alpha_1, \ldots, \alpha_n$ then its area is $\sum \alpha_i - (n-2)\pi$, by subdividing into $n-2$ spherical triangles.

1.5 Möbius geometry

Let $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$ have coordinate ξ. Consider the stereographic projection $\pi : S^2 \to \mathbb{C}_\infty$ defined by $\pi(P) =$ the point of intersection of the line NP with $\mathbb{C}$ identified as the plane $z = 0$, where $N = (0,0,1)$, with $\pi(N) = \infty$; this is clearly a bijection; by similar triangles we can see $\pi(x,y,z) = \frac{x+iy}{1-z}$

Lemma 2.12

If π' is the stereographic projection from the south pole then $\pi'(P) = \frac{1}{\pi(P)}$, since if $P = (x, y, z)$ then $\pi(P) = \frac{x+iy}{1-z}, \pi'(P) = \frac{x+iy}{1+z} \therefore \overline{\pi(P)}\pi'(P) = \frac{x^2+y^2}{1-z^2} = 1$.

Rk 2.13

Thus $\pi' \circ \pi^{-1} : \mathbb{C}_\infty \to \mathbb{C}_\infty$ is just circular inversion $\xi \mapsto \frac{1}{\xi}$.

1.6 Antipodal Points

For $P = (x, y, z) \in S^2$, have $\pi(P) = \xi = \frac{x+iy}{1-z}, \pi(-P) = -\frac{x+iy}{1+z}$ ($-P = (-x, -y, -z)$, the antipodal point of P) so $\pi(P)\overline{\pi(-P)} = -\frac{x^2+y^2}{1-z^2} = -1$, so $\pi(-P) = -\frac{1}{\bar{\xi}}$.

1.7 Möbius transformations

Recall that $\mathbb{C}_\infty$ is acted on by the gp G of Möbius transformations acting on it; if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathbb{C})$ then it defs a Möbius transformation on $\mathbb{C}_\infty$ by $\xi \mapsto \frac{a\xi+b}{c\xi+d}$. For any $\lambda \in \mathbb{C}^\star = \mathbb{C} \setminus \{0\}$ note λA defs the same M transformation; conversely if A_1, A_2 def the same M trans we can easily find $\exists \lambda \in \mathbb{C}^\star : A_1 = \lambda A_2$, so $G = PGL(2, \mathbb{C}) := \frac{GL(2,\mathbb{C}}{\mathbb{C}^\star}$.

Alternatively we can always normalize A so $\det A = 1$; if $\det A_1 = 1 = \det A_2$ and $A_1 = \lambda A_2$ then we must have $\lambda = \pm 1$, so $G = PSL(2, \mathbb{C}) := \frac{SL(2,\mathbb{C})}{\{\pm 1\}}$. On S^2 we have the rotations $SO(3)$ (recall the full isometry gp is $O(3)$).

Given 4 distinct points $z_1, z_2, z_3, z_4 \in \mathbb{C}_\infty$, recall we have a unique M transformation h sending z_1, z_2, z_3 to $0, 1, \infty$ respectively. We define the cross-ratio [of z_1, z_2, z_3 and z_4] to be the image of z_4 under this h, which we can calculate to be $w = \frac{z_4-z_1}{z_4-z_3}\frac{z_2-z_3}{z_2-z_1}$; by the way we have defined this it is clearly invariant under M transformations.

Suppose $\Gamma \in \mathbb{C}$ is a circle or straight line containing z_1, z_2, z_3, then the image of it under the above h must be the real axis, so $z_4 \in \Gamma \Leftrightarrow h(z_4) = w \in \mathbb{R}$; four distinct points lie on a circle or straight line iff their cross-ratio is real.

On $\mathbb{C}_\infty$ we have the action of $PSU(2) = \frac{SU(2)}{\{\pm I\}}$, the group of M transformations defined by matricies of $SU(2) \subset SL(2, \mathbb{C})$; recall that $SU(2)$ consists of matricies of the form $\begin{pmatrix} a & -b \\ \bar{b} & \bar{a} \end{pmatrix}$ with $|a|^2 + |b|^2 = 1$ (geometrically these are on $S^3 \subset \mathbb{R}^4$); $PSU(2)$ consists of M transformations of the form $z \mapsto \frac{az-b}{\bar{b}z+\bar{a}}$.

Theorem 2.14

Via stereographic projection π every rotation of S^3 gives rise to a M transformation defined by an element of $SU(2) \subset SL(2, \mathbb{C})$:

Rotations $R(z, \theta)$ (i.e. rotation about the z axis $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ by clockwise angle θ) correspond to M transformations $z \mapsto e^{i\theta} z$ and hence to matricies $\begin{pmatrix} e^{\frac{i\theta}{2}} & 0 \\ 0 & e^{-\frac{i\theta}{2}} \end{pmatrix} \in$

$SU(2)$.

The rotation $r(y, \frac{\pi}{2})$ is given by the matrix $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$ i.e. $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} z \\ y \\ -x \end{pmatrix}$. In $\mathbb{C}_\infty$ the corresponding map is $\xi = \frac{x+iy}{1-z} \mapsto \xi' = \frac{z+iy}{1+x}$; this maps $-1 \to \infty, 1 \to 0, i \to i$ so if it is given by a M transformation it must be $\xi' = \frac{\xi-1}{\xi+1}$. Now we check this transformation: $\frac{\xi-1}{\xi+1} = \frac{x+iy-1+z}{x+iy+1-z} = \frac{x-1+z+iy}{x+1-(z-iy)} = \frac{(z+iy)(x-1+z+iz)}{(x+1)(z+iy)+x^2-1}$ (since $x^2+y^2+z^2=1$) $= \frac{(z+iy)(x-1+z+iy)}{(x+1)(z+iy+x-1)} = \xi'$ as required, so this is an M transformation corresponding to this rotation.

But $SO(3)$ is generated by $r(y, \frac{\pi}{2}$ and the set of rotations of the form $r(z,\theta), 0 \leq \theta < 2\pi$, since we observe $r(x,\phi) = r(y,\frac{\pi}{2})r(z,\phi)r(y,-\frac{\pi}{2}$ $[r(y,-\frac{\pi}{2} = r(y,\frac{\pi}{2})^3]$, and for any $\vec{v} \in S^2$, $\exists \phi, \psi$ s.t. $g = r(z,\psi)r(x,\phi)$ maps $\vec{v}$ to $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, by having the first rotation map $\vec{v}$ into the horizontal plane $z=0$. Then $r(\vec{v},\theta) = g^{-1}r(x,\theta)g$.

So any rotation can be written as a product of these two elements, so corresponds to a product of M transformations with matricies $\in SU(2)$, so we have the result.

Theorem 2.15

The group $SO(3)$ acting on S^2 corresponds precisely with the subgroup $PSU(2)$ of M transformations acting on $\mathbb{C}_\infty$, as given a M transformation $\in PSU(2)$ $g(z) = \frac{az-b}{\bar{b}z+\bar{a}}$ with $|a|^2+|b|^2 = 1$, if $g(0)=0$ then $b=0$ and $a\bar{a}=1$ $\therefore$ $a = e^{\frac{i\theta}{2}}$ for some θ and g corresponds to $r(z,\theta)$, in general (using a common technique) let $g(0) = w \in \mathbb{C}, Q \in S^2$ s.t. $\pi(Q) = w$. Choose a rotation A of S^2 with $A(Q) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$ and let α be the corresponding element of $PSU(2)$, so $\alpha(w) = 0$ $\therefore$ $\alpha \circ g$ fixes 0 and corresponds to a rotation $B = r(z,\theta)$, so g corresponds to the rotation $A^{-1}B$.

So there exist 2:1 maps $SU(2) \to PSU(2) \cong SO(3)$. This is important to physics; it is the central part of the concept of "spin". This is the reason there exists a non-closed path of translations in $SU(2)$ going from I to $-I$ which corresponds to a <u>closed</u> path in $SO(3)$, starting and ending at $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Defn

The torus $T = T^2$ may be defined as a set $\frac{\mathbb{R}^2}{\mathbb{Z}^2}$; points are represented by $(x,y) \in \mathbb{R}^2$ but identified under the equivalence relation $(x_1,y_1) \sim (x_2,y_2) \Leftrightarrow x_1 - x_2 \in \mathbb{Z}, y_1 - y_2 \in \mathbb{Z}$. Alternatively, if Q is a closed square in $\mathbb{R}^2$ with verticies $(a,b),(a+1,b),(a,b+1),(a+1,b+1)$ then T is given by identifying opposite sides of Q.

The distance function d is defined by $d(P_1, P_2) = \min_{x_1, x_2 \in \mathbb{R}^2 \text{ representing } P_1, P_2 \text{ respectively}} |\vec{x}_1 - \vec{x}_2|$; we can easily verify this is a metric.

Let $f : \mathbb{R}^2 \to T$ be the quotient map used in the definition; on the interior $\mathrm{Int}(Q)$ of a square of the above form $f : \mathrm{Int}(Q) \to T$ is a bijection onto an open subset U of T, the complement of two "circles". Clearly this map does not preserve distances; however, for any $P \in \mathrm{Int}(Q)$ the restricion of f to a small enough open ball about P is an isometry. Thus $f\,|_{\mathrm{Int}(Q)}$ is a homeomorphism; both f and its inverse are locally isometries and so continuous.

Note that this is the locally euclidean torus T; its distance is very different from that of the torus embedded in $\mathbb{R}^3$.

Defn

A topological triangle on $X = S, T$ (more generally on any metric space X) is the image $R \subset X$ of a closed Euclidean triangle $\Delta \subset \mathbb{R}^2$ under a homeomorphism (Exercise: a spherical triangle is a topological triangle). A (topological) triangulation of X consists of a collection of topological triangles which cover X with the following properties: (This definition is only for 2D, though it is of course possible to extend it to higher dimensions)

Two triangles are either disjoint, intersect at a common vertex, or intersect on a common edge.

Each edge is the edge of precisely two triangles

Definition

The Euler number is defined by $e = F - E + V$ where F is the number of triangles (faces), E the number of edges and V the number of verticies.

Fact

The Euler number is independent of the choice of triangulation, with $e(S^2) = 2, e(T^2) = 0$; a sketch proof may be given later in this course, a full proof appears in the part II course "Algebraic Topology".

Examples

Slice the sphere as by the three planes $x = 0, y = 0, z = 0$, dividing it into eight triangles (there was some ambiguity in our earlier definition of a spherical triangle; the spherical triangle is the smaller of the two areas meeting the definition, i.e. that with area $< 2\pi$. This has F=8,E=12,V=6 so $e = 2$.

Divide the torus into nine squares as for a "noughts and crosses" game, then draw all the diagonals running from top right to bottom left. This hase F=18,E=27,V=9. Note that if we only divide the torus into four squares initially, this is not strictly speaking a triangulization, as there are triangles whose intersection is two common verticies. However, this still obeys Euler's rule; F=8,E=12,V=4 so $e = 0$. Thus we can relax our definition somewhat from the strict one required in Algebraic Topology:

Definition

A (geodesic) polygonal decomposition of S^2 (respectively T^2) consists of a finite collection of geodesic polygons (each contained in an open hemisphere (respectively corresponding to a Euclidean polygon in $\mathrm{Int}(Q)$)) which cover X with the interiors (faces) of the polygons disjoint. The edges of the decomposition are defined to be those of the polygons, likewise verticies, and any edge is the edge of precisely two polygons of the collection, and no edge has a vertex as an "interior point". We can still define the Euler number:

Proposition 3.6

For any such geodesic triangle decomposition of S^2 (respectively T^2) the Euler number $e = F - E + V$ is 2 (respectively 0): if the triangles are $\Delta_1, \ldots, \Delta_f$ and the interior angles of each Δ_i sum to T_i respectively, then $\sum T_i = 2\pi V$ since at each vertex the angles sum to 2π. Also $3F = 2E$ since each triangle has three edges and each edge is common to two triangles. So $F = 2E - 2^F$. For S^2, by Gauss-Bonnet $\mathrm{area}(\Delta_i) = T_i - \pi$ so $4\pi = \sum_i = 1^F \mathrm{area}(\Delta_i) = \sum(T_i - \pi) = 2\pi V - \pi F = 2\pi(V - E + F$ so $e = 2$; for T we have $T_i = \pi \forall i$ so $\sum_i (T_i - \pi) = 0 = e$.

This proof is immediately also valid for any decomposition into convex polygons, by subdividing them into triangles (each subdivision increases both F and E by 1 each, so does not change e).

As a proof of an earlier exercise: a spherical triangle is the radial projection of a plane triangle onto S^2, and such projection is a homeomorphism, so a spherical triangle is a topological triangle.

Recall from analysis II the derivatives of maps: suppose $U \subset \mathbb{R}^n$ is open. A map $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ is smooth (or C^∞) if each f_i has partial derivatives of all orders. Clearly any such f is diffable since it has cnts partial derivs.

The derivation of f at $\vec{a} \in U$ is a linear map $df_{\vec{a}} : \mathbb{R}^n \to \mathbb{R}^m$ (in analysis II we called this $Df\mid_{\vec{a}}$, but there are good reasons for preferring this "little-d" notation; see part II) such that $\frac{\|f(\vec{a}+\vec{h}) - f(\vec{a}) - df_{\vec{a}}(h)\|}{\|h\|} \to 0$ as $h \to 0 \in \mathbb{R}^n$.

When $m = 1$, $df_{\vec{a}} : \mathbb{R}^n \to \mathbb{R}$ is determined by the partial derivatives of f at $\vec{a}$, $(\frac{\partial f}{\partial x_1}(\vec{a}), \ldots, \frac{\partial f}{\partial x_n}(\vec{a}))$ via matrix multiplication, i.e. it is $(h_1, \ldots, h_n) \mapsto \sum_i \frac{\partial f}{\partial x_i}(\vec{a}) h_i$. For general m $df_{\vec{a}} : \mathbb{R}^n \to \mathbb{R}^m$ is determined by the $m \times n$ matrix of partial derivatives $J(f) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdots & & \cdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$ at $\vec{a}$, the Jacobian matrix

Example

Analytic functions $f : U \subset \mathbb{C} \to \mathbb{C}$ in one complex variable x with U open have $\frac{|f(z+w) - f(z) - wf'(z)|}{|w|} \to 0$ as $w \to 0 \in \mathbb{C}$. So if we set $f'(z) = a + ib, w = h_1 + h_2$ then $f'(z)w = (ah_1 - bh_2) + i(ah_2 + bh_1)$; if we consider f as a smooth map $U \to \mathbb{R}^2$ then the linear map $df_x : \mathbb{R}^2 \to \mathbb{R}^2$ is given by $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Chain rule

Given smooth maps $f : U \subset \mathbb{R}^n \to \mathbb{R}^m, g : V \subset \mathbb{R}^p \to U$ with U, V open, $fg : V \to \mathbb{R}^m$ has derivative at $p \in V$ given by $d(fg)_p = df_{g(p)} \circ dg_p$. In terms of Jacobian matricies this is that $J(fg)_p = J(f)_{g(p)}J(g)_p$ where the subscripts are the points at which the matricies are evaluated and the product is matrix multiplication.

Riemannian metrics on open subsets of $\mathbb{R}^2$

Let $V \subset \mathbb{R}^2$ open, and take coordinates (u, v) on $\mathbb{R}^2$. A Riemannian metric on V, $Edu^2 + 2Fdudv + Gdv^2$, is defined by giving C^∞ functions E, F, G on V such that the matrix $\begin{pmatrix} E(P) & F(P) \\ F(P) & G(P) \end{pmatrix}$ is positive definite $\forall P \in V$ and hence (for $P \in V$) defines an inner product $\langle , \rangle_p$ on $\mathbb{R}^2$, varying smoothly with P.

Non-examinable: explanation of origin for notation

The coordinate functions $u : V \to \mathbb{R}, v : V \to \mathbb{R}$ have derivatives (at $P \in V$) which are denoted by du_P, dv_P where $du_P(h_1, h_2) = h_1, dv_P(h_1, h_2) = h_2$. Thus du_P, dv_P is the dual basis to the standard basis of $\mathbb{R}^2$. We drop the $_P$s as the maps are independent of P, then du, dv represents the dual basis at all points.

$du^2(\vec{h}, \vec{k}) := du(\vec{h})du(\vec{k}), dudv(\vec{h}, \vec{k}) := \frac{1}{2}(du(\vec{h})dv(\vec{k}) + dv(\vec{h})du(\vec{k})), dv^2(\vec{h}, \vec{k}) := dv(\vec{h})dv(\vec{k})$. These are symmetric bilinear forms, with corresponding matricies $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ respectively. So $Edu^2 + 2Fdudv + Gdv^2$ is just the symmetric bilinear form defined by $\begin{pmatrix} E & F \\ F & G \end{pmatrix}$.

Definition

Given a (piecewise) smooth curve $\gamma = (\gamma_1, \gamma_2) : [0, 1] \to V \subset \mathbb{R}^2$, define its <u>length</u> to be $\int_0^1 \sqrt{E\dot{\gamma}_1^2 + 2F\dot{\gamma}_1\dot{\gamma}_2 + G\dot{\gamma}_2^2}dt$. The <u>area</u> of a region $W \subset V$ is defined to be $\int_W \sqrt{EG - F^2}dudv$, [of course both of these are only defined] when this integral exists.

Example

$V = \mathbb{R}^2$, the Riemannian metric $\frac{4(du^2+dv^2)}{(1+u^2+v^2)^2}$. We have met this already; the stereographic projection $\pi : S^2 \to \mathbb{R}^2 = \mathbb{C}$ has for $P \in S^2 \setminus \{\text{north pole}\}, \pi(P) \in \mathbb{R}^2$ and an inner product $\langle , \rangle_{\pi(P)}$ defined as follows:

The tangent space to S^2 at P consists of vectors $\vec{x}$ such that $\vec{x} \cdot \overrightarrow{OP} = 0$ with the origin identified with P; this is a vector space. For $\vec{x}_1, \vec{x}_2$ in this space, $\vec{x}_1 \cdot \vec{x}_2 = \langle d\pi_P(\vec{x}_1), d\pi_P(\vec{x}_2)\rangle_{\pi(p)}$.

Suppose $\phi : V \to \tilde{V}$ is a <u>diffeomorphism</u>, that is, a smooth map with a smooth inverse, between open subsets of $\mathbb{R}^2$, and we have Riemannian metrics on $V, \tilde{V}$ giving ries to families of inner products $\langle , \rangle_P$ for $P \in V$, $\tilde{\langle , \rangle}_Q$ for $Q \in \tilde{V}$.

Definition

ϕ is called an isometry if $\forall P \in V, \langle \vec{x}, \vec{y} \rangle_P = \langle d\phi_P(\vec{x}), d\phi_P(\vec{y}) \rangle^{\sim}_{\phi(P)} \forall \vec{x}, \vec{y} \in \mathbb{R}^2$. If $\gamma : [0,1] \to V$ is a smooth curve then $\tilde{\gamma} = \phi \circ \gamma : [0,1] \to \tilde{V}$ is a smooth curve and $\langle \tilde{\gamma}'(t), \tilde{\gamma}'(t) \rangle^{\sim}_{\tilde{\gamma}(t)} = \langle d\phi_P(\gamma'(t), d\phi_P(\gamma'(t)) \rangle^{\sim}_{\phi(P)} = \langle \gamma'(t), \gamma'(t) \rangle_{\gamma(t)}$ where $P = \gamma(t)$. So if ϕ is an isometry, $\operatorname{length} \tilde{\gamma} = \operatorname{length} \gamma = \int_0^1 \sqrt{\langle \gamma^| prime(t), \gamma'(t) \rangle} dt$. Isometries also preserve areas, as shown in the printed notes for this course.

Disc model

The disc model for the hyperbolic plane is given by $V = D \subset \mathbb{C}$ the unit disc with Riemannian metric $\frac{4(du^2+dv^2)}{(1-u^2-v^2)^2}$. If $\zeta = u + iv$ we can write this as $\frac{4{:}d\zeta|^2}{(1-|\zeta|^2)^2}$ where $|d\zeta|^2 = du^2 + dv^2$. Lengths are scaled by $\frac{2}{1-r^2}$ where $r^2 = u^2 + v^2$, and areas are scaled by $\frac{4}{(1-r^2)^2}$.

Upper half-plane model

$H = \{z \in \mathbb{C} : \operatorname{Im}(z) > 0\}$ is conformally equivalent to D via the Möbius transformation $\zeta \mapsto \frac{i(1+\zeta)}{1-\zeta}$. Notice that this is equivalent to a rotation of the Riemannian sphere - we are mapping the "north" hemisphere to the "bottom" hemisphere. If we use z for complex coordinates on H, $z = x + iy$ we have $z = \frac{i(1+\zeta)}{1-\zeta}$ with inverse $\zeta = \frac{z-i}{z+i}$.

We can in general see hyperbolic geometry as "spherical geometry with the sign changed".

We want a metric on the upper half plane such that the above maps are isometries: if we had the Euclidean inner product on the unit disc as a subset of $\mathbb{C}$, we can write $\langle w_1, w_2 \rangle = \operatorname{Re}(w_1 \bar{w}_2) = \frac{1}{2}(w_1 \bar{w}_2 + \bar{w}_1 w_2)$ so at $z \in H$, the inner product on $\mathbb{R}^2 = \mathbb{C}$ induced from the Euclidean inner product at $\xi = \frac{z-i}{z+i}$, i.e. the inner product which makes the map an isometry, is given by forcing the isometry condition: $\langle w_1, w_2 \rangle_z = \langle \frac{d\xi}{dz} w_1, \frac{d\xi}{dz} w_2 \rangle_{\text{eucl}} = |\frac{d\xi}{dz}|^2 \operatorname{Re}(w_1 \bar{w}_2)$, i.e. we obtain the Riemannian metric $|\frac{d\xi}{dz}|^2 (dx^2 + dy^2) = \frac{4}{{:}z+i|^4}(dx^2 + dy^2)$ on H. From $\xi = \frac{z-i}{z+i}$ we can find $\frac{1}{1-|\xi|^2} = \frac{|z+i|^2}{4 \operatorname{Im} z}$, so the metric on H corresponding to the Riemannian metric $\frac{4|d\xi|^2}{(1-|\xi|^2)^2}$ is $4 \times \frac{4}{|z+i|^4} \left(\frac{|z+i|^2}{4+\operatorname{Im} z} \right)^2 |dz|^2 = \frac{|dz|^2}{(\operatorname{Im} z)^2} = \frac{dx^2+dy^2}{y^2}$ [check: should that be $4 + \operatorname{Im} z$ or $4 \operatorname{Im} z$?]

Consider $PSL(2, \mathbb{R})$, the group of M transformations $z \mapsto \frac{az+b}{cz+d}$ with $a, b, c, d \in \mathbb{R}, \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1$. The first condition gives the set of Möbius transformations mapping $R \cup \{\infty\}$ into itself. We also want to ensure the upper half-plane is mapped to itself (rather than to the lower half plane) wo we check: $i \mapsto \frac{ai+b}{ci+d} = \frac{(b+ai)(d-ci)}{c^2+d^2}$; the imaginary part of this is $> 0 \Leftrightarrow \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} > 0$, i.e. if and only if we can normalize $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ so its determinant is 1. Thus this group is a group of Möbius transformations mapping H into itself.

Proposition 4.4

The elements of $PSL(2,\mathbb{R})$ are isometries of H, i.e. they preserve length of curves.

Recall $PSL(2\mathbb{R})$ is generated by translations $z \mapsto z+a, a \in \mathbb{R}$, dilations $z \mapsto az, a \in \mathbb{R}+$ and the map $z \mapsto -\frac{1}{z}$. The first two of these are clearly isometries. For $z \mapsto -\frac{1}{z}$ the induced may on $\mathbb{C} = \mathbb{R}^2$ is given by multiplication by $\frac{d(-\frac{1}{z})}{dz} = \frac{1}{z^2}$, and note $\operatorname{Im}(-\frac{1}{z}) = -\frac{1}{|z|^2} \operatorname{Im} \bar{z} = \frac{Imz}{|z|^2}$, thus the metric is $\frac{\frac{1}{|z|^4}|dz|^2}{\frac{(\operatorname{Im} z)^2}{|z|^4}} = \frac{|dz|^2}{(\operatorname{Im} z)^2}$ as required.

$PSL(2,\mathbb{R})$ contains M transformations $z \mapsto az+b, a>0$ so acts transitively on H.

Let $l =$ the imaginary axis, $g \in PSL(2,\mathbb{R})$, then $g(l)$ is a circline orthogonal to the real axis $\mathbb{R} = g(\mathbb{R})$. Let $l+ = \{it : t > 0\}$; its image is either a vertical half-line or a semicircle (with ends on the real axis) (both in the upper half plane since g maps this into itself); these are called the hyperbolic lines.

Lemma (4.6)

Through any two points $z_1 \neq z_2 \in H, \exists!$ hyperbolic line l through z_1, z_2; this is clear if $\operatorname{Re} z_1 = \operatorname{Re} z_2$, otherwise the centre of the required semicircle is the intersection of the perpendicular bisector of z_1, z_2 with the real axis.

4.7

$PSL(2,\mathbb{R})$ acts transitively on the hyperbolic lines: given any hyperbolic line l, $\exists g \in PSL(2,\mathbb{R}) : g(l) = l+$; this is clear for l a vertical line, otherwise for l a semicircle with endpoints $s < t \in \mathbb{R}$ we take $g(z) = \frac{z-t}{z-s}$ (note $\det \begin{pmatrix} 1 & -t \\ 1 & -s \end{pmatrix} > 0$) and observe $g(t) = 0, g(s) = \infty$.

Remark

If we compose g with $z \mapsto -\frac{1}{z}$ we get $h \in PSL(2,\mathbb{R})$ such that $h(s) = 0, h(t) = \infty$. We can also choose that a given point $P \in l$ maps to i, by scaling.

We define $\rho(z_1, z_2) =$length along the hyperbolic line between z_1, z_2 (under our metric).

Given $z_1, z_2 \in H, \exists h \in PSL(2,\mathbb{R})$ sich that $h(z_1^\star) = 0, h(z_2^\star) = \infty$ where $z_i^\star$ is the endpoint of l as defined above nearer to z_i, so $h(z_1) = iu, h(z_2) = iv, u < v$. Since h is an isometry $\rho(z_1, z_2) = \rho(iu, iv)$.

Let $T : [0,1] \to H$ be such that $T(t) = if(t) \in l+$ where $T(0) = iu, T(1) = iv$ and $\frac{df}{dt}$ is continuous and $\geq 0 \forall t$. Then $\rho(z_1, z_2) = \rho(iu, iv) = \operatorname{length} T = \int_0^1 \frac{|\frac{df}{dt}|}{f} dt = \int_0^1 \frac{\frac{df}{dt}}{f} dt = \log \frac{v}{u}$.

The minimizing geodesics on H correspond to hyperbolic line segments:

Proposition 4.10

If $\gamma : [0,1] \to H$ is a piecewise continuously differentiable curve from z_1 to z_2, then $\operatorname{length} \gamma \geq \rho(z_1, z_2)$, with equality if and only if γ is a monotonic parameterization of the hyperbolic line segment $[z_1, z_2]$; without loss of generality take $z_1 = iu, z_2 = iv, u < v$. Suppose $\gamma = \gamma + i\gamma_2$ is as above, then length $\gamma = \int_0^1 \sqrt{(\frac{d\gamma_1}{dt})^2 + (\frac{d\gamma_2}{dt})^2} \frac{dt}{\gamma_2(t)} \geq \int_0^1 |\frac{d\gamma_2}{dt}| \frac{dt}{\gamma_2(t)} \geq \int_0^1 \frac{d\gamma_2}{dt} \frac{dt}{\gamma_2(t)} = [\log \gamma_2(t)]]_0^1 = \log(\frac{v}{u}) = \rho(z_1, z_2)$ with equality if and only if both these inequalities are equalities, i.e. $\frac{d\gamma_1}{dt} \equiv 0$ so $\gamma_1 \equiv 0$ (and γ runs along the imaginary axis) and $\frac{d\gamma_2}{dt} \geq 0$ so we have the result; since we considered general piecewise continuously diffentiable curves these statements may be false at the finitely many points where γ is not continuously differentiable, but this does not affect the conclusion.

Remark

Given z_1, z_2, z_3 we can consider a curve γ consisting of the hyperbolic line segments $[z_1, z_2]$ followed by $[z_2, z_3]$, then length $\gamma = \rho(z_1, z_2) + \rho(z_2, z_3) \geq \rho(z_1, z_3)$ with equality if and only if z_2 lies on $[z_1, z_3]$ (this is why we bothered considering piecewise continuously differentiable curves above). Thus ρ is a metric.

We can now use this to show that any continuous curve from z_1 to z_2 has length $\leq \rho(z_1, z_2)$, bu the same proof as in the Euclidean and spherical cases.

Geometry of disc model D

M transformations sending the unit circle to itself and D to itself correspond to elements of $PSL(2, \mathbb{R})$ acting on H and hence are isometries of D; they form a group G.

Hyperbolic lines in D are given by circle segments orthogonal to the unit circle, including diameters (since M transformations (such as the equivalence between D and H) preserve angles).

G acts transitively on the set of (hyperbolic line $l, P \in l$); cf (4.7).

D has Riemannian metric $\frac{4|dz|^2}{((1-|z|^2)^2}$; the elements of G are isometries of D.

On the disc:

(4.12)

Rotations $z \mapsto e^{i\theta} z$ are elements of G (this is trivial)

If $a \in D$ then $g(z) = \frac{z-a}{1-\bar{a}z}$ as a map is $\in G$: if $|z| = 1$ observe that $|1 - \bar{a}z| = |\bar{z}(1 - \bar{a}z)| = |\bar{z} - \bar{a}| = |z - a|$ so $|g(z)| = 1$, and $g(a) = 0 \in D$ so $g(D) = D$ and $g \in G$.

Remark

See example sheet 2; in fact any element of G is of the form $z \mapsto e^{i\theta} \left(\frac{z-a}{1-\bar{a}z}\right)$ for some θ and $a \in D$.

This group is "half" the isometries of D; it does not include reflections (and nor does $PSL(2)$ on H).

Proposition 4.14

If $0 \leq r < 1$ then $\rho(0, re^{i\theta}) = 2 \tanh^{-1} r$; in general for $z_1, z_2 \in D$, $\rho(z_1, z_2) = 2 \tanh^{-1} \left| \frac{z_1 - z_2}{1 - \bar{z}_1 z_2} \right|$; see example sheet 2 for a similar formula in H.

$\rho(0, re^{i\theta}) = \rho(0, r) = \int_0^r \frac{2dt}{1-t^2} = 2\tanh^{-1} r$; in general let l be the hyperbolic line through z_1, z_2 and apply $z \mapsto \frac{z-z_1}{1-\bar{z}_1 z} \in G$; this maps $z_1 \to 0, l \to$ a diameter of D, wlog the real axis, and under this $z_2 \to |\frac{z_1-z_2}{1-\bar{z}_1 z_2}|$ and we have the result.

Examples

Hyperbolic circles are Euclidean circles; transform the centre of the circle to 0 and this is clear, and since the transformation is a M transformation it transforms Euclidean circlines into Euclidean circlines; since the equivalence between D and H is also a M transformation this also means circles in H are Euclidean circles (in both cases we cannot have lines because the transformations map the interior of D into the interior of D (or H, and in this case recall the point at infinity is not in the interior of H)).

Given a point p and hyperbolic line l with $p \notin l, \exists!$ hyperbolic line l' with $l' \ni p, l' \perp l$ and the distance along this line is the minimum distance between p and l; this is clear by transforming p onto 0.

Lemma 4.17

Suppose g is an isometry of H fixing all points of $l+ = \{iy : y > 0\}$. Then $g = \iota$ or $g(z) = -\bar{z} \forall z \in H$, i.e. g is reflection in the y axis.

For any $p \notin l+, \exists!$ hyperbilic line l' through p and $\perp l+$, but the minimum distance between p and $l+$ is the distance from p to the point of intersection of $l', l+$, say q; q is mapped to itself and distances are preserved so p must be mapped to one of the two points p, p' (where p' is the reflection of p in $l+$). We now claim that if $\exists p \notin l+$ with $g(p) = p$ then $g = \iota$ and otherwise g is reflection in $l+$: take wlog $p \in$ the positive quadrant $H+$. If $g(P) = P$ then for any point A of $H+$ if $g(A) = A'$ then $\rho(A', P) = \rho(A, P$, but $\rho(A', P) = \rho(A', B) + \rho(B, P)$ for B the intersection of $[A, P'], [A', P]$, but by symmetry this is $\rho(A, B) + \rho(B, P)$ contradicting the triangle inequality [since B is not on the line $[A, P]$].

Let R be reflection in $l+$. For any hyperbolic line l choose $T \in PSL(2, \mathbb{R})$ such that $T(l) = l+$, then define $R_l = T^{-1}RT$; from this lemma this is the unique non-identity isometry fixing all points of l.

We can also define R_l geometrically; the image of P is the point on the perpendicular l' from P to l equidistant from l as P but $\neq P$.

4.19

Any isometry g of H is either an element of $PSL(2, \mathbb{R})$ or an element in its coset $R \cdot PSL(2, \mathbb{R})$ where R is reflection in $l+$; suppose $g(l+) = l$, choose T such that $Tl = l+$ and consider Tg, i.e. we can wlog take $g(l+) = l+$. Then composing if necessary with $z \mapsto -\frac{1}{z}$ and scaling by a real number, we can take $g(0) = 0, g(\infty) = \infty, g(i) = i$. Thus g fixes all the points of $l+$ and by (4.17) g is ι or R and we are done; by "unrolling" this process we have g of the required form. See the second example sheet for this course for a similar classification of the isometries of D.

We call the isometries of the form $z \mapsto \frac{az+b}{cz+d}, a, b, c, d \in \mathbb{R}, ad - bc > 0$ the direct isometries. By the same proof as in the Euclidean case, any isometry is the composition of at most three reflections, and direct isometries are the composition of at most two reflections.

For regions $R \subset H$, the area of R is $\iint_R \frac{dxdy}{y^2}$. We want to consider triangles; it is convenient to also include triangles with points on the boundary.

For a triangle $T = ABC$ with angles α, β, γ, possibly some of which are 0, the area of T is $\pi - (\alpha + \beta + \gamma)$; note this is - what it was in the spherical case; this is an example of curvature, which will be covered later in the course.

We first proove the result for $\gamma = 0$; we use H and can wlog take $C = \infty$ (notice that in this proof we implicitly assume isometries preserve areas; we shall proove this later). Then by translation and scaling by a real number we can take A, B to be on the unit semicircle (centre 0); by angle chasing we find B is $e^{i\beta}, A = e^{i(1-\alpha)}$. Then the side BC is the line $x = \cos\beta$ and similarly, so the area of T is $\int_{\cos(\pi-\alpha)}^{\cos\beta} \int_{\sqrt{1-x^2}}^{\infty} \frac{dydx}{y^2} = \int_{\cos(\pi-\alpha)}^{\cos\beta} \frac{dx}{\sqrt{1-x^2}} = [-\cos^{-1} x]_{\cos(\pi-\alpha)}^{\cos\beta} = \pi - \alpha - \beta$. Then for a general γ we transform the triangle so A, B lie on the unit circle and C lies on the line $x = \cos(\pi - \alpha)$; let δ be the angle between BC and $C\infty$. Let $\Delta_1 = AB\infty, \Delta_2 = CB\infty$, then the area of T is $\text{area}(\Delta_1) - \text{area}(\Delta_2) = \pi - \alpha - (\beta + \delta) - (\pi - \delta - (\pi - \gamma)) = \pi - \alpha - \beta - \gamma$ as required.

Parallel and Ultraparallel lines

In Euclidean geometry, for a line l and $P \notin l$ there is a unique line $l' \ni P$ not intersecting l; there is no such l' in the spherical case.

We define lines l_1, l_2 in D are <u>parallel</u> if they only meet on the boundary $|z| = 1$ and ultraparallel if they do not meet on the full set $|z| \leq 1$. In H we take wlog $l_1 = l+$, then the l_2 parallel to it are the vertical lines and the semicircles with one endpoint at 0, and the ultraparallel lines are the semicircles with both endpoints ¿0 or both endpoints ¡0.

On the second example sheet we proove that two hyperbolic lines are ultraparallel if and only if they have a common perpendicular. Note also that ultraparallel lines are some finite minimum distance apart, while parallel lines converge.

We saw that we can stereographicly project the sphere onto $\mathbb{C}[\cup\{\infty\}]$. We can also form D by projecting the hyperboloid S+ defined by $q(\vec{x}) = x^2 + y^2 - z^2 = -1$ for $z > 0$ onto $D \subset \mathbb{C}$ by projection with centre $\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$, $\pi(x, y, z) = \frac{x+iy}{1+z}$, with inverse $\sigma(u, v) = (\frac{2u}{1-r^2}, \frac{2v}{1-r^2}, \frac{1+r^2}{1-r^2})$ where $r^2 = u^2 + v^2$. $\sigma_u = d\sigma_{(u,v)}(e_1) = \frac{\partial\sigma}{\partial u} = \frac{2}{(1-r^2)^2}(1 + u^2 - v^2, 2uv, 2u), \sigma_v = \frac{2}{(1-r^2)^2}(2uv, 1 + v^2 - u^2, 2v)$; the reader may verify these are linearly independent for any (u, v) with $u^2 + v^2 < 1$. Then σ_u, σ_v generate the tangent space to S+ at $\sigma(u, v)$, and the Lorentzian inner product $\langle\vec{x}, \vec{x}\rangle = q(\vec{x})$ determines a symmetric bilinear form on this vector space, and hence via $d\sigma$, which identifies e_1 with σ_u and e_2 with σ_v, determines a bilinear form $Edu^2 + 2Fdudv + Gdv^2$ on D via $E = (e_1, e_1) := \langle d\sigma(e_1), d\sigma(e_1)\rangle = \langle\sigma_u, \sigma_u\rangle = \frac{4}{(1-r^2)^2}, F = (e_1, e_2) := \langle d\sigma(e_1), d\sigma(e_2)\rangle = \langle\sigma_u, \sigma_v\rangle = 0, G = (e_2, e_2) := \langle d\sigma(e_2), d\sigma(e_2)\rangle = \langle\sigma_v, \sigma_v\rangle = \frac{4}{(1-r^2)^2}$; this gives the hyperbolic metric $\frac{4(du^2+dv^2)}{(1-r^2)^2}$ on D.

Definition

$S \subset \mathbb{R}^3$ is a <u>smooth embedded surface</u> if each point of S has an open neighbourhood $U = W \cap S$ for W open in $\mathbb{R}^3$ and map $\sigma : V \to U$ from some open

$V \subset \mathbb{R}^2$ such that σ is a homeomorphism, i.e. $\sigma(u,v) = (x(u,v), y(u,v), z(u,v))$ is C^∞ (has continuous derivatives of all orders).

At each point $Q = \sigma(P) \in U$, the vectors $\sigma_u(P) = \frac{\partial \sigma}{\partial u}(P) = d\sigma_P(e_1)$ and $\sigma_v(P)$ are linearly independent.

σ is called a smooth parameterization of $U \subset S$; (u,v) are called smooth coordinates on U.

The tangent space $T_{S,U}$ is the subspace of $\mathbb{R}^3$ generated by σ_u, σ_v.

Proposition 5.2 [or possibly 5.3]

Suppose $\sigma : V \to U$ and $\tilde{\sigma} : \tilde{V} \to U$ are smooth parameterizations of U. Then the homeomorphism $\phi = \sigma^{-1} \circ \tilde{\sigma} : \tilde{V} \to V$ is a diffeomorphism, that is a smooth map with a smooth inverse, with $\tilde{\sigma} = \sigma \circ \phi$.

The Jacobian matrix $\begin{pmatrix} x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{pmatrix}$ has rank 2 everywhere since its columns are σ_u, σ_v. Since ϕ is a homeomorphism, it suffices to proove it is a diffeomorphism locally. Wlog take $\det \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} \neq 0$ (as the determinant of at least one such minor is $\neq 0$) at $(u_0, v_0) \in V$. Consider $F : V \to \mathbb{R}^2$ given by projection π of σ, i.e. $f(u,v) = (x(u,v), y(u,v))$. Now we apply a stronger form of the Inverse Function Theorem than was prooven in the Analysis II course and have that F is a local diffeomorphism at (u_0, v_0), i.e. $\exists$ open neighbourhoods $(u_0, v_0) \in N \subset V \subset \mathbb{R}^2, F(u_0, v_0) \in N' \subset \mathbb{R}^2$ such that $F\mid_N \mid N \to N'$ is a diffeomorphism. Now $\sigma\mid_N : N \to \sigma(N)$ is a homeomorphism onto an open subset of U; since $F\mid_N : N \to N'$ is a homeomorphism so too is the projection $\pi : \sigma(N) \to N' \subset \mathbb{R}^2$. Set $\tilde{N} = \tilde{\sigma}^{-1}(\sigma(N))$ open in $\tilde{V}$ and $\tilde{F} = \pi \circ \tilde{\sigma} : \tilde{N} \to N'$. Then $\phi\mid_{\tilde{N}} = \sigma^{-1} \circ \tilde{\sigma}$ is $\sigma^{-1} \circ \pi^{-1} \circ \pi \circ \sigma = F^{-1} \circ \tilde{F}$ on $\tilde{N}$; since F^{-1} and $\tilde{F}$ are C^∞ so is $\phi\mid_{\tilde{N}}$; similarly the same is true of $\phi^{-1}\mid_N$ so ϕ is a diffeomorphism.

Corollary 5.3

The tangent space $T_{S,Q}$ is independent of the choice of parameterization $\sigma : V \to U \ni Q$; (5.2) implies that given $\sigma : V \to U$ a smooth parameterization of S any other smooth parameterization is of the form $\tilde{\sigma} = \sigma \circ \phi$ with $\phi = (\phi_1, \phi_2) : \tilde{V} \to V \subset \mathbb{R}^2$ a diffeomorphism. Let $(\tilde{u}, \tilde{v})$ be coordinates on $\tilde{V}$, then $\tilde{\sigma}_{\tilde{u}} = \frac{\partial \phi_1}{\partial \tilde{u}} \sigma_u + \frac{\partial \phi_2}{\partial \tilde{u}} \sigma_v, \tilde{\sigma}_{\tilde{v}} = \frac{\partial \phi_1}{\partial \tilde{v}} \sigma_u + \frac{\partial \phi_2}{\partial \tilde{v}} \sigma_v$ where $J(\phi) = \begin{pmatrix} \frac{\partial \phi_1}{\partial \tilde{u}} & \frac{\partial \phi_2}{\partial \tilde{u}} \\ \frac{\partial \phi_1}{\partial \tilde{v}} & \frac{\partial \phi_2}{\partial \tilde{v}} \end{pmatrix}$ is invertible, and the result is clear.

Remark

$\tilde{\sigma}_{\tilde{u}} \times \tilde{\sigma}_{\tilde{v}} = \det(J) \sigma_u \times \sigma_v$. Define the unit normal to S at Q as $\vec{N} = \vec{N}_Q = \frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|}$; this is unique up to its sign.

The inverse $\theta : U \to V$ of $\sigma : V \to U \subset S$ is called a chart. A collection of charts covering S is called an atlas.

Given a smooth parameterization $\sigma : V \to U \subset S \subset \mathbb{R}^3$ we define a Riemannian metric on V by $\langle \vec{a}, \vec{b} \rangle_P = \langle d\sigma_P(\vec{a}), d\sigma_P(\vec{b}) \rangle_{\mathbb{R}^3}$, i.e. the Riemannian matrix at coordinates (u,v) has $E = |\sigma_u|^2 = \sigma_u \cdot \sigma_u, F = \sigma_u \cdot \sigma_v, G = |\sigma_v|^2$. We will refer to this as the first fundamental form; we may also see this as a family of inner products on the tangent spaces to points of S.

The inner product for the $\tilde{\sigma}$ chart $\langle \vec{a}, \vec{b} \rangle_P$ for $\vec{a}, \vec{b} \in \mathbb{R}^2$ is $\langle d\tilde{\sigma}_P(\vec{a}), d\tilde{\sigma}_P(\vec{b}) \rangle_{\mathbb{R}^3}$ (where $\langle , \rangle_{\mathbb{R}^3}$ denotes the standard inner product on $\mathbb{R}^3$) which is $\langle d\phi_P(\vec{a}), d\phi_P(\vec{b}) \rangle_{\phi(P)}$; similarly by the chain rule $d\tilde{\sigma}_P = d\sigma_{\phi(P)} \circ d\sigma(P)$ and by definition $\langle \vec{a}', \vec{b}' \rangle_{\phi(P)} = \langle d\sigma_{\phi(P)}(\vec{a}'), d\sigma_{\phi(P)}(\vec{b}') \rangle_{\mathbb{R}^3}$, so with regard to the induced Riemannian metrics on $V, \tilde{V}$, the diffeomorphism ϕ is an isometry.

If we have Riemannian metrics on $V, \tilde{V}$ by $Edu^2 + 2Fdudv + Gdv^2, \tilde{E}d\tilde{u}^2 + 2\tilde{F}d\tilde{u}\tilde{v} + \tilde{G}d\tilde{v}^2$ then this is the statement that $\vec{a}^t = \begin{pmatrix} \tilde{E} & \tilde{F} \\ \tilde{F} & \tilde{G} \end{pmatrix}_P \vec{b} = \vec{a}^t J^t \begin{pmatrix} E & F \\ F & G \end{pmatrix}_{\phi(P)} J\vec{b} \forall \vec{a}, \vec{b} \in \mathbb{R}^2$ where $J = J(\phi)$ the Jacobian matrix, i.e. $\begin{pmatrix} \tilde{E} & \tilde{F} \\ \tilde{F} & \tilde{G} \end{pmatrix} = J^t \begin{pmatrix} E \circ \phi & F \circ \phi \\ F \circ \phi & G \circ \phi \end{pmatrix} J$.

Length and energy

For $\Gamma : [a, b] \to S$ a smooth curve we define $\text{length}(\Gamma) = \int_a^b \|\Gamma'(t)\| dt, \text{energy}(\Gamma) = \int_a^b \|\Gamma'(t)\|^2 dt$ (many books insert a factor of $\frac{1}{2}$; cf kinetic energy. This is sometimes called action). If $\text{Im}(\Gamma) \subset U$ with $\theta : U \to V$ a chart given by $\theta = \sigma^{-1}$ set $\gamma = \theta \circ \Gamma : [a, b] \to V$, a smooth curve. Then by the chain rule and the fact that $\Gamma = \sigma \circ \gamma$, $\langle \Gamma'(t), \Gamma'(t) \rangle_{\mathbb{R}^3} = \langle \gamma'(t), \gamma'(t) \rangle_P$ so $\text{length}(\Gamma) = \text{length}(\gamma) = \int_a^b \sqrt{E(\gamma(t))\dot{\gamma}_1^2 + 2F(\gamma(t))\dot{\gamma}_1\dot{\gamma}_2 + G(\gamma(t))\dot{\gamma}_2^2} dt$ and similarly for energy. We could define the length of curves in this way, but doing so is senseless since we have a perfectly good definition of length in $\mathbb{R}^3$; however:

Area

With the same notation, for a "nice" region $T \subset U$ we define its area to be (when it exists) $\int_{\theta(T)} \sqrt{EG - F^2} dudv$; since $\|\sigma_u \times \sigma_v\|^2 + (\sigma_u \cdot \sigma_v)^2 = \|\sigma_u\|^2 \|\sigma_v\|^2$ this is $\int_{\theta(T)} \|\sigma_u \times \sigma_v\| dudv$; this definition is far more important than for length since we do not have a clear notion of area "upstairs" on the surface in $\mathbb{R}^3$. This means we need to proove:

Proposition

This definition is independent of the choice of parameterization $\sigma : V \to U$. A corollary of this is that we can define the area of regions not contained in the image of a single chart (since area is additive, and the area of a region "upstairs" is now well defined, we can just subdivide the region "upstairs"), though this is generally not very useful since we usually have a chart covering almost all of the surface. Given charts $\tilde{\theta}, \theta$ for U and ϕ the transition function $\sigma^{-1} \circ \tilde{\sigma} : \tilde{V} \to V$, for $P \in \tilde{V}$ $\begin{pmatrix} \tilde{E} & \tilde{F} \\ \tilde{F} & \tilde{G} \end{pmatrix}_P = J^t \begin{pmatrix} E \circ \phi & F \circ \phi \\ F \circ \phi & G \circ \phi \end{pmatrix}_{\phi(P)} J$ (*); by the change of variable formula for integrals on $\mathbb{R}^2$ for any continuous H on $\theta(T) = \phi(\tilde{\theta}(T))$, $\int_{\theta(T)} H dudv = \int_{\phi^{-1}(\theta(T)) = \tilde{\theta}(T)} H \circ \phi |J(\phi)| d\tilde{u}d\tilde{v}$; setting $H = \sqrt{EG - F^2}$ on V, $\tilde{H} = \sqrt{\tilde{E}\tilde{G} - \tilde{F}^2}$ on $\tilde{V}$ and taking determinants in (*) $\tilde{H} = (H \circ \phi)|\det J(\phi)|$ and $\int_{\theta(T)} H dudv = \int_{\tilde{\theta}(T)} \tilde{H} d\tilde{u}d\tilde{v}$ as required.

Suppose $V \subset \mathbb{R}^2$ is open with a Riemannian metric, and $\gamma = (\gamma_1, \gamma_2) : [a,b] \to V$ a smooth curve. Then:

Definition 6.1

γ is a <u>geodesic</u> curve if $\frac{d}{dt}(E\dot\gamma_1 + F\dot\gamma_2) = \frac{1}{2}(E_u\dot\gamma_1^2 + 2F_u\dot\gamma_1\dot\gamma_2 + G_u\dot\gamma_2^2)$, $\frac{d}{dt}(F\dot\gamma_1 + G\dot\gamma_2) = \frac{1}{2}(E_v\dot\gamma_1^2 + 2F_v\dot\gamma_1\dot\gamma_2 + G_v\dot\gamma_2^2)\forall t$; we have $E = E(\gamma_1(t), \gamma_2(t))$ etc. are functions of t; this may seem a rather abstract definition but we have the following:

Suppose $\gamma(a) = p, \gamma(b) = q$. Then a <u>proper</u> variation of γ is a smooth map $h : [a,b] \times)(-\epsilon, \epsilon) \subset \mathbb{R}^2 \to V$ such that $h(t,0) = \gamma(t)\forall t \in [a,b], h(a,\tau) = p, h(b,\tau) = q\forall\tau \in (-\epsilon,\epsilon)$. For each $\tau \in (-\epsilon,\epsilon)$ we have a smooth curve $\gamma_\tau : [a,b] \to V$ by $\gamma_\tau(t) = h(t,\tau)$.

Proposition 6.2

The curve γ satisfies the geodesic equations if and only if it represents a stationary point of the <u>energy</u> for all proper variations; let energy$(\gamma) = \int_a^b (E\dot u^2 + 2F\dot u\dot v + G\dot v^2)dt = \int_a^b I(t,u,v,\dot u,\dot v)dt$, then by the Euler-Lagrange equations γ is a stationary point for energy if and only if $\frac{d}{dt}(\frac{\partial I}{\partial \dot u} = \frac{\partial i}{\partial u}, \frac{d}{dt}(\frac{\partial I}{\partial \dot v} = \frac{\partial i}{\partial v}$. Since $\frac{\partial I}{\partial \dot u} = 2(E\dot u + F\dot v), \frac{\partial I}{\partial u} = E_u\dot u^2 + 2F_u\dot u\dot v + G_u\dot v^2$ and similarly for v, these are just the geodesic equations.

For $S \subset \mathbb{R}^3$ an embedded surface and $\sigma : V \to U \subset S, \theta = \sigma^{-1} : U \to V$, if $\gamma : [a,b] \to U$ then $\gamma = \theta \circ \Gamma$ is a smooth curve on V; call Γ a <u>geodesic</u> if and only if γ is a geodesic; this is the case if and only if Γ represents a stationary point for the energy $\int_a^b \|\gamma'(t)\|^2 dt$, so in particular this definition does not depend on the choice of chart; this also means we can define when an arbitrary curve $\Gamma : [a,b] \to S$ is a geodesic.

Corollary 6.3

If Γ on S minimizes the energy for curves joining $P = \Gamma(a)$ to $Q = \Gamma(b)$ then it is a geodesic in the above sense: for any $a < a_1 < b_1 < b$ the curve $\Gamma_1 = \Gamma\mid_{[a_1,b_1]}$ minimizes energy for curves joining $\Gamma(a_1)$ to $\Gamma(b_1)$. If a_1, b_1 are chosen so that $\mathrm{Im}(\Gamma_1) \subset$ some image of a chart U then Γ_1 is a geodesic since it represents a stationary point for energy; varying a_1, b_1 along the curve we have the result. In fact, if Γ <u>locally</u> minimises the energy (i.e. for any $t_0 \in (a,b)$, $\exists\epsilon > 0 : \Gamma\mid_{[t_0-\epsilon,t_0+\epsilon]}$ minimises energy for curves joining its endpoints, then Γ is a geodesic. Without proof, the conversely is also true: al geodesics minimise energy locally; see later for the relation between energy and length (and recall that the two geodesics joining two non-antipodal points on a sphere are the short and long great circle arcs between them).

Proposition 6.5

For a smooth curve Γ on an embedded surface S, the geodesic equations are equivalent to $\frac{d^2\Gamma}{dt^2}$ being always normal to S; wlog take $\Gamma : [a,b] \to U \subset S, \sigma : V \to U$ a parameterization, then $\Gamma = \sigma \circ \gamma$ with $\gamma(t) = \gamma_1(t)e_1 + \gamma_2(t)e_2$, so by the chain rule $\dot\Gamma(t) = (d\sigma)_{\gamma(t)}\dot\gamma(t) = (d\sigma)_{\gamma(t)}(\dot\gamma_1(t)e_1 + \dot\gamma_2(t)e_2 = \dot\gamma_1(t)\sigma_u + \dot\gamma_2\sigma_v$, so $\frac{d^2\Gamma}{dt^2} \perp \langle\sigma_u, \sigma_v\rangle \subset \mathbb{R}^2 \Leftrightarrow \frac{d}{dt}(\dot\gamma_1\sigma_u + \dot\gamma_2\sigma_v)\cdot\sigma_u = \frac{d}{dt}(\dot\gamma_1\sigma_u + \dot\gamma_2\sigma_v)\cdot\sigma_v = 0$, and we can show that this is equivalent to the geodesic equations. So if Γ is a geodesic on S then $\frac{d}{dt}(\dot{\Gamma}\cdot\dot{\Gamma}) = 2\dot{\Gamma}\cdot\ddot{\Gamma} = 0 \therefore \|\frac{d\Gamma}{dt}\|^2$ is constant. This is the difference between defining geodesics in terms of energy or length; if we were to define our geodesics just in terms of length the parameterization would be unrestricted, wheras this condition tells us something about the parameterization.

[lecture missed at this point]

For a surface of revolution S paramaterized by $\sigma(u,v) = (f(u)\cos v, f(u)\sin v, g(u))$ e.g. the embedded torus $f(u) = a + b\cos u, g(u) = b\sin u$ on V given by $\alpha < u < \alpha + 2\pi, \beta < v < \beta + 2\pi$ with $0 < b < a$, $\sigma_u = (f'\cos v, f'\sin v, g'), \sigma_v = (-f\sin v, f\cos v, 0)$; the first fundamental form with regard to this parameterization is $du^2 + f^2dv^2$ and the geodesic equations for $\gamma(t) = (u(t), v(t))$ are $\ddot{u} = f(u)\frac{df}{du}\dot{v}^2, \frac{d}{dt}(f(u)^2\dot{v}) = 0$. We assume without loss of generality that $\|\dot{\gamma}\| = 1$ i.e. $\dot{u}^2 + f(u)^2\dot{v}^2 = 1$ [lulz, the lecturing was at ludicrous speed by this point; apologies for any mistakes].

It is hard to find results about general geodesics, even in very simple cases like this, for example are the geodesics dense on the (embedded) torus (the answer is yes, but the proof of this is nonobvious).

Proposition 6.15

i) Every unit speed meridian (curve along an "unrevolved" shape) is a geodesic; we have v constant so the second equation is satisfied, and by the unit speed condition $\dot{u}$ is constant so the first equation is also satisfied.

ii) a (unit speed) parallel (line which is the revolution of a dot) $u = u_0$ is a geodesic if and only if $\frac{df}{du}\mid_{u_0} = 0$ i.e. u_0 is a stationary point of f: if $u = u_0$ for a unit speed parallel then $f(u)^2\dot{v}^2 = 1 \Rightarrow \dot{v} = \frac{\pm 1}{f(u_0)}$, a nonzero constant; the second equation is satisfied and the first is satisfied if and only if $\frac{df}{du}\mid_{u_0} = 0$.

Curvature

Suppose $\eta : [0,1] \to \mathbb{R}^2$ is a smooth curve with unit speed $\eta'\cdot\eta' = 1$ (note this implies $\eta'\cdot\eta'' = 0$; recall that the <u>curvature</u> κ at a point $\eta(s)$ is defined by $\eta''(s) = \kappa\vec{n}$ where $\vec{n}$ is a unit normal and $\kappa \geq 0$ (alternatively, we could define that $\vec{n}$ is such that $(\eta', \eta'', \vec{n})$ is always a right- or left-handed tuple and then κ may be negative.

If we reparamaterize by a smooth function $f : [c,d] \to [0,l]$ with $f'(t) > 0\forall t$ and set $\gamma(t) = \eta(f(t))$ then $\dot{\gamma}(t) = \frac{df}{dt}\eta'(f(t))$ (and hence $\|\dot{\gamma}\|^2 = \left(\frac{df}{dt}\right)^2$); moreover $\eta''(f(t)) = \kappa\vec{n}$ where κ is the curvature at $\gamma(t)$. $\gamma(t+\delta t) - \gamma(t) = \frac{df}{dt}\eta'(f(t))\delta t + \frac{1}{2}\left((\frac{d^2f}{dt^2})\eta'(f(t)) + (\frac{df}{dt})^2\eta''(f(t))\right)(\delta t)^2 + \ldots$. But $\eta'\cdot\eta' = 1 \Rightarrow \eta'\cdot\eta^{)}\prime\prime = 0 \Rightarrow \eta'\cdot\vec{n} = 0 \therefore (\gamma(t+\delta t) - \gamma(t))\cdot\vec{n} = \frac{1}{2}\kappa\|\dot{\gamma}\|^2(\delta t)^2$. Observe $\|\gamma(t+\delta t) - \gamma(t)\|^2 = \|\dot{\gamma}\|^2(\delta t)^2 + \ldots$ so $\frac{1}{2}\kappa$ is just the <u>ratio</u> of the quadratic terms of these two expansions, i.e. $\lim_{\delta t\to 0}\frac{(\gamma(t+\delta t)-\gamma(t))\cdot\vec{n}}{\|\gamma(t+\delta t)-\gamma(t)\|^2}$.

Given a parameterization $\sigma : V \to U \subset S$ for $V \subset \mathbb{R}^2$, Taylor's Theorem implies $\sigma(u+\delta u, v+\delta v) - \sigma(u,v) = \sigma_u\delta u + \sigma_v\delta v + \frac{1}{2}(\sigma_{uu}(\delta u)^2 + 2\sigma_{uv}\delta u\delta v + \sigma_{vv}(\delta v)^2) + \ldots$, so "deviation of σ from the tangent plane" is $(\sigma(u+\delta u, v+\delta v) - \sigma(u,v))\cdot\vec{N} = \frac{1}{2}(L(\delta u)^2 + 2M\delta u\delta v + N(\delta v)^2) + \ldots$ where $L = \sigma_{uu}\cdot\vec{N}, M = \sigma_{uv}\cdot\vec{N}, N = \sigma_{vv}\cdot\vec{N}$; the <u>second fundamental form</u> on V is $Ldu^2 + 2Mdudv + Ndv^2$. where L, M, N are C^∞ functions on V defined by this.

Now $\|\sigma(u+\delta u, v+\delta v) - \sigma(u,v)\|^2 = E(\delta u)^2 + 2F\delta u\delta v + g(\delta v)^2$ where $E = \|\sigma_u\|^2$ and similarly, i.e. $Edu^2 + Fdudv + Gdv^2$ is the first fundamental form.

Definition

The Gaussian Curvature of S at P is $K := \frac{LN-M^2}{EG-F^2}$ (note that we have not yet prooved this is independent of parameterization); $K > 0$ means the second fundamental form is positive or negative definite, $K < 0$ means it is indefinite and $K = 0$ means it is semidefinite but not definite; note this does not imply the surface is locally flat (however, if the second fundamental form is 0 this does imply the surface is locally flat).

Proposition 7.2

If $\vec{N}$ denotes the unit normal of the surface σ then at a given point, $\vec{N}_u = a\sigma_u + b\sigma_v, \vec{N}_v = c\sigma_u + d\sigma_v$ (†) where $\begin{pmatrix} L & M \\ M & N \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} E & F \\ F & G \end{pmatrix}$; in particular $K = ad - bc$: since $\vec{N}\cdot\vec{N} = 1$ we have $\vec{N}_u \cdot \vec{N} = 0 = \vec{N}_v \cdot \vec{N}$; similarly since $\vec{N}\cdot\sigma_u = 0$ we have $\vec{N}_u \cdot \sigma_u + \vec{N}\cdot\sigma_{uu} = 0 \Rightarrow \vec{N}u\cdot\sigma_u = -L$; similarly $\vec{N}_u \cdot \sigma_v = -M$ and so forth. Dotting (†) with σ_u, σ_v we have $-L = aE + bF, -M = cE + dF, -M = aF + bG, -N = cF + dG$ and we have the matrix equation, and taking determinants we have the particular result.

Corollary 7.3

K is independent of paramaterization; by 7.2, $N_u \times N_v = K\sigma_u \times \sigma_v$; if we reparamaterize U by $\tilde{\sigma} : \tilde{V} \to U$ with $\phi : \tilde{V} \to V$ a diffeomorphism, recall $\tilde{\sigma}_{\tilde{u}} \times \tilde{\sigma}_{\tilde{v}} = \det(J)\sigma_u \times \sigma_v$ where $J = J(\phi)$. $\tilde{N} = \pm N$ depending on the sign of $\det J$. By the chain rule, $N_{\tilde{u}} = \frac{\partial u}{\partial \tilde{u}} N_u + \frac{\partial v}{\partial \tilde{u}} N_v$ and similarly for $N_{\tilde{v}}$ so $N_{\tilde{u}} \times N_{\tilde{v}} = \det(J) N_u \times N_v$ so $\det(J)K\sigma_u \times \sigma_k = \det(J)N_u \times N_v = N_{\tilde{u}} \times N_{\tilde{v}} = \tilde{K}\tilde{\sigma}_{\tilde{u}} \times \tilde{\sigma}_{\tilde{v}} = \tilde{K}\det(J)\sigma_u \times \sigma_v$ and $K = \tilde{K}$.

Example

A surface of revolution obtained from a unit speed curve $\eta(u) = (f(u), 0, g(u))$ has $K = \frac{-f''}{f}$; see question 9 on the third example sheet for this course.

Theorem 7.4

Suppose S is an embedded surface with coordinate patch $\sigma : V \to U$ on which the first fundamental form takes the shape $du^2 + G(u,v)dv^2$, then the Gaussian curvature K is $\frac{-(\sqrt{G})_{uu}}{\sqrt{G}}$; this result should be proven on the third example sheet for this course. Also see the later coverage of geodesic polars.

For example, a surface of revolution created by revolving $(f(u), 0, g(u))$ about the z axis has $\sigma(u,v) = (f(u)\cos v, f(u)\sin v, g(u))$; the first fundamental form is $du^2 + f^2(u)dv^2$, so $K = \frac{-f''}{f}$.

Corollary

For $P \in S \subset \mathbb{R}^3$ with S an embedded surface, we have local geodesic polar coordinates (ρ, θ) at P and hence also $G(\rho, \theta)$ which depend only on the metric (which was before shown to be $d\rho^2 + G(\rho,\theta)d\theta^2$), so the curvature $K = \frac{-(\sqrt{G})_{\rho\rho}}{\sqrt{G}}$ (The point P which corresponds to $\rho = 0$ is not actually in the coordinate patch, but by continuity the result holds), so the curvature at P depends only on the metric and we have Gauss' Theorema Egregium: the curvature of an embedded surface depends only on the first fundamental form, and not on the embedding.

The proof is an example of Gauss' "Moving Frame" technique: we set $\vec{e} = \sigma_u, \vec{f} = \frac{\sigma_v}{\sqrt{G}}$ and together with $\vec{N}$ these form an orthonormal triad. Since $\vec{e}\cdot\vec{e} = 1$ we have $\vec{e}\cdot\vec{e}_u = 0$ and similar results, so we write [sometimes ommiting vector arrows from now on] $e_u = \alpha f + \lambda_1 \vec{N}, e_v = \beta f + \mu_1 \vec{N}, f_u = -\alpha' e + \lambda_2 \vec{N}, f_v = -\beta' e + \mu_2 \vec{N}$ $(\star)$. Since $e \cdot f = 0$ [and of course $e \cdot \vec{N} = 0$, and similarly] we have $e_u \cdot f + e \cdot f_u = 0, e_v \cdot f + e \cdot f_v = 0$ so in $(\star)$ we have $\alpha' = \alpha, \beta' = \beta$.

$\alpha = e_u \cdot f = \sigma_{uu} \cdot \frac{\sigma_v}{\sqrt{G}} = \frac{(\sigma_u \cdot \sigma_v)_u}{\sqrt{G}} - \frac{1}{2}\frac{(\sigma_u \cdot \sigma_u)_v}{\sqrt{G}} = 0 - 0 = 0, \beta = e_v \cdot f = \sigma_{uv} \cdot \frac{\sigma_v}{\sqrt{G}} = \frac{1}{2}\frac{G_u}{\sqrt{G}}$ (since $G = \sigma_v \cdot \sigma_v$) $= (\sqrt{G})_u$.

Apply $(\star)$ again: $\lambda_1\mu_2 - \lambda_2\mu_1 = e_u \cdot f_v - f_u \cdot e_v = \frac{\partial}{\partial u}(e \cdot f_v) - \frac{\partial}{\partial v}(e \cdot f_u) = -\beta_u - 0 = -(\sqrt{G})_{uu}$. Now $\vec{N}_u \times \vec{N}_v = (a\sigma_u + b\sigma_v) \times (c\sigma_u + d\sigma_v) = (ad - bc)\sigma_u \times \sigma_v = K(\sigma_u \times \sigma_v)$ where $\vec{N} = \frac{\sigma_u \times \sigma_v}{\sqrt{G}} = e \times f$. Now $\vec{N}_u \times \vec{N}_v \cdot \vec{N} = K\sqrt{G}$ but also $= (\vec{N}_u \times \vec{N}_v) \cdot (e \times f) = (N_u \cdot e)(N_v \cdot f) - (N_u \cdot f)(N_v \cdot e) = (N \cdot e_u)(N \cdot f_v) - (N \cdot f_u)(N \cdot e_v)$ since $N \cdot e = 0 \Rightarrow N_u \cdot e + N \cdot e_u = 0$ and similarly. So $K\sqrt{G} = \lambda_1\mu_2 - \lambda_2\mu_1 = (-\sqrt{G})_{uu} \therefore K = \frac{-(\sqrt{G})_{uu}}{\sqrt{G}}$ as required.

Definition

An abstract smooth surface S is a metric space (or equivalently a topological space, since we have a natural metric by the infinum of lengths of curves) with homeomorphisms $\theta_i : U_i \to V_i$ from open $U_i \subset S$ to open $V_i \subset \mathbb{R}^2$ such that $S = \bigcup_i U_i$ and $\forall i, j, \phi_{ij} := \theta_i \circ \theta_j^{-1} : \theta_j(U_i \cap U_j) \to \theta_i(U_i \cap U_j)$ is a diffeomorphism. We define charts and atlases as before. We say S is closed if it is compact.

For an abstract surface S equipped with an atlas, a Riemannian metric on S is given by Riemannian metrics on the images V_i of the charts $\theta_i : U_i \to V_i$ subject to the compatibility condition that $\forall i, j, \langle d\phi_P(\vec{a}), d\phi_P(\vec{b})\rangle_{\phi(P)}$ under the metric on $V_i = \langle \vec{a}, \vec{b}\rangle_P$ under the metric on $V_j \forall P \in \theta_j(U_i \cap U_j) \forall \vec{a}, \vec{b} \in \mathbb{R}^2$ where $\phi = \phi_{ij}$ i.e. the ϕ_{ij} are isometries. For $\Gamma : [0, 1] \to S$ a smooth curve we locally compose Γ with charts $\theta : U \to V$ to give smooth curves $\gamma = \theta \circ \Gamma$; we define $\|\Gamma'(t)\| = \langle \gamma'(t), \gamma'(t)\rangle_{\gamma(t)}^{\frac{1}{2}} = \sqrt{E\dot{\gamma}_1^2 + F\dot{\gamma}_1\dot{\gamma}_2 + G\dot{\gamma}_2^2}$ and this is independent of the choice of chart (for example for the locally Euclidean torus T our charts are simply translation onto squares in $\mathbb{R}^2$, clearly isometries of the locally Euclidean metric); then we define length$(\Gamma) = \int_0^1 \|\Gamma'(t)\| dt$; the reader should compare and contrast embedded and abstract smooth surfaces.

Examples

The standard geometries: $\mathbb{R}^2$ with metric $dx^2 + dy^2$, $S^2 \subset \mathbb{R}^3$ and indeed any embedded surfaces, and the disc $D \subset \mathbb{R}^2$ with metric $\frac{4(dx^2+dy^2)}{(1-x^2-y^2)^2}$ or upper half-plane $H \subset \mathbb{R}^2$ with metric $\frac{dx^2+dy^2}{y^2}$. Based on these examples we might ask why we make this definition when we could simply work with Riemannian metrics on open subsets of $\mathbb{R}^2$ and embedded surfaces, but we do need it because the locally Euclidean torus as introduced earlier is compact so not an open subset of $\mathbb{R}^2$, but cannot be realised as an embedded surface in $\mathbb{R}^3$ [as will be shown either later, or on the third example sheet for this course].

A map $f : X \to Y$ between abstract spaces is smooth if for any charts $\theta : U \to V$ on X and $\theta^\star : U^\star \to V^\star$ on V with $U \cap f^{-1}(U^\star)) \neq \emptyset$, the composite map $f : \theta(U \cap f^{-1}(U^\star)) \to V^\star = \theta^\star \circ f \circ \theta^{-1}$ is smooth; f is called a diffeomorphism if it is smooth with smooth inverse.

f is called a local isometry if for all pairs of charts as above the map f is an isometry, i.e. $\forall P \in \theta(U \cap f^{-1}(U^\star)) \forall \vec{a}, \vec{b} \in \mathbb{R}^2, \langle \vec{a}, \vec{b} \rangle_P = \langle d\vec{f}_P(\vec{a}), d\vec{f}_P(\vec{b}) \rangle^\star_{f(P)}$; the first inner product is the local inner product for the θ chart, the lecond that for the $\theta^\star$ chart.

The metric determines local geodesic polar coordinates (ρ, θ) at a given point P, with regard to which the metric takes the form $d\rho^2 + G(\rho,\theta)d\theta^2$ (with the metric at P itself found by the limit of this as the point in question tends to P); we can then define curvature as per (7.4), i.e. $K = \frac{-(\sqrt{G})_{\rho\rho}}{\sqrt{G}}$; this is equivalent to the usual definition of curvature in the embedded case. We take this to be the definition of curvature in the abstract case.

Examples

For these three examples we have a transitive isometry group, so we only need to find our results for one point.

$\mathbb{R}^2$ has $\rho = r$ and the metric is $d\rho^2 + \rho^2 d\theta^2$; $\sqrt{G} = \rho$ and $K = 0$.

S^2 has metric $d\rho^2 + \sin^2 \rho d\theta^2$; $\sqrt{G} = \sin \rho \Rightarrow (\sqrt{G})_{\rho\rho} = -\sqrt{G} \Rightarrow K = +1$.

The disc model of the hyperbolic plane has metric $(\frac{2}{1-r^2})^2(dr^2 + r^2 d\theta^2)$. Take geodesic polars (ρ, θ) where $\rho = 2\tanh^{-1} r$, then $d\rho^2 = (\frac{2}{1-r^2})^2 dr^2$ but $r = \tanh \frac{\rho}{2} \Rightarrow \frac{4r^2}{(1-r^2)^2} = \sinh^2 \rho$ so the metric is $d\rho^2 + \sinh^2 \rho d\theta^2$ $\therefore \sqrt{G} = \sinh \rho \Rightarrow (\sqrt{G})_{\rho\rho} = \sqrt{G} \Rightarrow K = -1$.

Lemma 7.8

Suppose S is a surface with Riemannian metric g with curvature K, then under the scaled metric $c^2 g$ on S with $c > 0$ the curvature becomes $\frac{K}{c^2}$, e.g. the sphere of radius $c > 0$ has constant curvature $\frac{1}{c^2}$.

Theorem 7.9

If S is equipped with a Riemannian metric g with constant curvature K, then after rescaling the metric S is locally isometric to an open subset of $S^2, \mathbb{R}^2$ or the hyperbolic plane, according as K is $> 0, = 0, < 0$ respectively.

Theorem 7.10 (Gauss-Bonnet)

If $T = ABC$ is a geodesic triangle on an abstract smooth surface S equipped with a Riemannian metric with angles α, β, γ then $\int_T K dA = (\alpha+\beta+\gamma) - \pi$. If S is a closed (i.e. compact) surface then $\int_S K dA = 2\pi e(s)$ where $e(s)$ is the Euler number [for triangulizations of this surface]. By the same argument as in 3.6, the first part of this theorem implies the second: $\int_S KdA = \sum_{\text{triangles } T_i} \int_{T_i} KdA = \sum_{T_i}(\alpha_i + \beta_i + \gamma_i - \pi) = \sum_{T_i}(\alpha_i - \beta_i - \gamma_i) - \pi F$ but $\sum_{T_i}(\alpha_i + \beta_i + \gamma_i) = 2\pi V$ and $3F = 2E$ so $\int_S KdA = 2\pi e(S)$. We shall not actually proove the result in this course, but as an "explanation of why it's true", consider decomposing a triangle into smaller triangles (cf the proof of Cauchy's Theorem in the Complex Aanlysis course)); for constant curvature we have already proven the result, so we take the limit as the triangles become small enough that the curvature is approximately constant on them (if we want an actual rigoorus proof, this is more easily obtained by applying calculus to the definition of curvature - cf the proof of Gauss-Bonnet in the hyperbolic case). Alternatively we can similarly consider small geodesic circles.

Assuming (7.10) we can <u>define</u> the Gaussian curvature K at a point P on a surface to be $\lim_{\text{area } \triangle \to 0} \frac{\sum \text{angles of } \triangle}{\text{area } \triangle}$; this is possibly the best way to show that K is well defined in the general case.

This is the end of this course; it leads particularly to the part II courses Geometry of Group Actions, Riemannian Surfaces, Differential Geometry, Algebraic Topology and General Relativity.