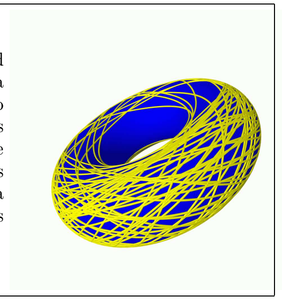
GEODESICS

Math118, O. Knill

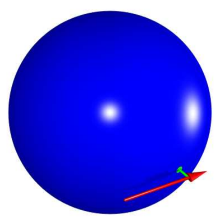
ABSTRACT. Light moves on shortest paths. The corresponding dynamical system is called the **geodesic flow**. We will see examples of geodesic flows which are integrable like the flow on a surface of revolution. This is an introduction to geodesic flows without Riemannian geometry which allows to go straight to the essential math without too much formalism.

ARCHIMEDES THEOREM. We have seen that the shortest distance between two points in Euclidean space is the line. We have proven this in the case of the plane without use of derivatives. This "Archimedes proof" can be generalized to higher dimensional Euclidean spaces too.

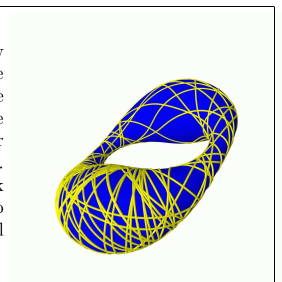
DEFINITION. Given a smooth surface in space, a point P on the surface and initial tangent velocity vector v. Define a path on the surface by letting a particle move freely in space under the influence of a force perpendicular to the surface in such a way that the particle stays on the surface. This defines a path on the surface called **geodesic flow**. This dynamical system can be described using differential equations too. However, for many of the examples considered here, we can work with the intuitive notion. If the surface has a boundary, then we have a **surface billiard**. In that case, we assume the mass point bounces off the boundary according to the usual billiard law.



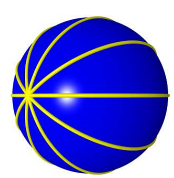
The force $F(x,v)$ perpendicular to the surface at the point x to the direction v can be computed by intersecting the plane spanned by the unit normal vector $\vec{n}$ and the vector v with the surface, leading to a curve with a **radius of curvature** r. Applying the centrifugal force $F(x,v) = |v|^2 n/r$ assures that the particle stays on the surface. The number $\kappa(x,v) = 1/r(x,v)$ is called the **sectional curvature** at the point in the direction v.



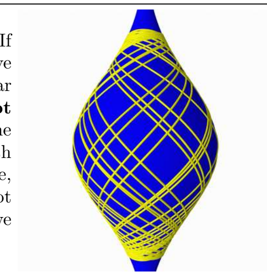
MOTIVATION. The numerical method, we used to compute the geodesic flow on some of the pictures on this page is a mechanical one. We constrain the free motion onto the surface. Given a surface X in space we look at the free evolution of the particle subject to a strong force which pulls the particle to the surface. That force is always perpendicular to the surface and so perpendicular to the velocity of the particle. Especially, it does not accelerate the particle. Do a free evolution in space for some time dt, then projection the vector back onto the surface. $X(u,v) \to X(u,v) + V \to X(u_1,v_1)$ This method is so efficient and simple, that we have let the ray-tracing program (Povray) do all the computation for the pictures.



EXAMPLE: GEODESICS ON THE SPHERE.
On a sphere, the mass-point is at any time subject to a force which goes through the center of the sphere. Angular momentum conservation $\frac{d}{dt}L = \frac{d}{dt} x \times v = 0$ implies that the particle stays on a plane spanned by the normal vector and the initial vector v. The geodesic curve is the intersection of the plane with the sphere: it is a grand circle. The plane can be seen as a limiting case of the sphere, when the radius goes to infinity. A particle which initially is on a plane and has a velocity tangent to the plane stays on the plane without any need of constraint. The geodesic curves consist of lines.

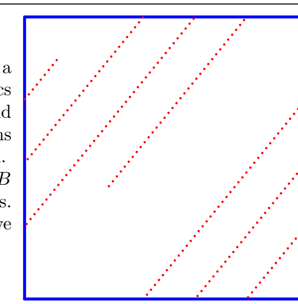


EXAMPLE: GEODESICS ON SURFACE OF REVOLUTION. If ϕ is the angle between a longitudinal line and the geodesic curve and r is the distance from the axes of rotation, then the angular momentum $L = r\sin(\phi)$ is conserved. It is called the **Clairot** integral. Examples of surfaces of revolution are the cylinder, the cone or the torus. If we write the torus as part of the plane with a space dependent metric which depends only on one coordinate, we have a geodesic flow on a surface of revolution. The Clairot integral $r\sin(\phi)$ is the analogue of Snells integral $g(x)\sin(\alpha)$ we have seen before.



METRIC AND DISTANCE. Consider a two-dimensional parametrized surface $(u,v) \mapsto r(u,v)$. At a point $(u,v,r(u,v)$, we have the tangent vectors $dx = r_u du, dy = r_v dv$ The distance element $ds = \sqrt{dx \cdot dx + dy \cdot dy}$ satisfies $ds^2 = (r_u du + r_v dv)^2 = r_u \cdot r_u du du + r_u \cdot r_v du dv + r_v \cdot r_u dv du + r_v \cdot r_v dv dv$. With $g = \begin{bmatrix} r_u \cdot r_u & r_u \cdot r_v \\ r_v \cdot r_u & r_v \cdot r_v \end{bmatrix}$, this can be written as $ds^2 = (du,dv) \cdot g(du,dv)$. A new dot product $< a,b >= a \cdot gb$ and length $||a|| = \sqrt{< a,a >}$ allows to write the length of a curve as $\int_a^b ||r'(t)||\, dt$. Riemanns view is to start with a two dimensional surface M and a symmetric matrix at each point $g_{ij}(x,y)$ defined so that both eigenvalues of g are positive everywhere. The pair (M,g) defines a **Riemannian manifold**. One can measure distances on it without referring to the ambient space in which the surface is embedded.

EXAMPLE: GEODESICS ON THE FLAT TORUS. Because a region in a flat torus can be seen as a region in the plane, geodesics on the flat torus are made of lines. With $g_{ij} = 1$ if $i = j$ and $g_{ij} = 0$ if $i \neq j$ as in the case of the plane, the differential equations for the geodesics are $\ddot{x}^k = \Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0$. There is no acceleration. The fact that the shortest connections between two points A, B on the flat plane are straight lines can be seen in different ways. The straight line gives a distance between the two points as we have seen before in the plane.



EXAMPLE: HILLY REGION. Let $r(u,v) = (u,v,f(u,v))$ be a parameterization of the graph of f. The metric is $g(u,v) = \begin{bmatrix} r_u \cdot r_u & r_u \cdot r_v \\ r_v \cdot r_u & r_v \cdot r_v \end{bmatrix} = \begin{bmatrix} 1+f_u^2 & f_u f_v \\ f_u f_v & 1+f_v^2 \end{bmatrix}$. So, if $r(u(t),v(t))$ is a curve on the surface, we can calculate its length. We should get the same result as if we would compute the length of the curve $r(t) = (u(t),v(t),f(u(t),v(t)))$ in three dimensional flat space. But with the internal formalism, it is possible to compute the length without using the third dimension.

CONNECTION. When minimizing the length of a curve, we have to find the Euler Lagrange equations. This involves differentiating the metric g further. The **Christoffel symbols** are defined as

$$\Gamma_{ijk} = \frac{1}{2}[\frac{\partial}{\partial x^i} g_{jk}(x) + \frac{\partial}{\partial x^j} g_{ik}(x) - \frac{\partial}{\partial x^k} g_{ij}(x)] \, .$$

For a parametrized surface, this is

$$\Gamma_{111} = r_{uu} \cdot r_u, \; \Gamma_{112} = r_{uu} \cdot r_v \qquad \Gamma_{211} = r_{vu} \cdot r_u, \; \Gamma_{212} = r_{vu} \cdot r_v$$
$$\Gamma_{121} = r_{uv} \cdot r_u, \; \Gamma_{122} = r_{uv} \cdot r_v \qquad \Gamma_{221} = r_{vv} \cdot r_u, \; \Gamma_{222} = r_{vv} \cdot r_v$$

FREE MOTION ON A SURFACE. A particle of momentum p has the Lagrangian $F(t,x,p) = \frac{1}{2} g_{ij}(x) p^i p^j$. We use **Einstein summation convention** to automatically sum over pairs of lower and upper indices. We want to minimize $I(x) = \int_a^b F(t,x,\dot{x}) dt = \int_{t_1}^{t_2} g_{ij}(x)\dot{x}^i \dot{x}^j \; dt$ With $F_{p_k} = g_{ki} p^i$ and $F_{x_k} = \frac{1}{2}\frac{\partial}{\partial x^k} g_{ij}(x) p^i p^j$ and the identities $\frac{1}{2}\frac{\partial}{\partial x^j} g_{ik}(x)\dot{x}^i \dot{x}^j = \frac{1}{2}\frac{\partial}{\partial x^i} g_{jk}(x)\dot{x}^i\dot{x}^j$ $g_{ki}\ddot{x}^i = -\Gamma_{ijk}\dot{x}^i\dot{x}^j$ and the definitions $g^{ij} = g_{ij}^{-1}$, $\Gamma^k_{ij} := g^{lk}\Gamma_{ijl}$ this can be written as

$$\ddot{x}^k = -\Gamma^k_{ij}\dot{x}^i\dot{x}^j$$

Because F is time independent, $H(p) = p^k F_{p^k} - F = p^k g_{ki} p^i - F = 2F - F = F(p)$ is constant along the orbit.

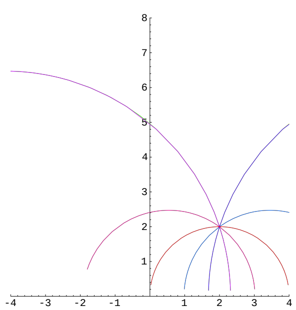
GEODESICS ON A SURFACE With $G(t,x,p) = \sqrt{g_{ij}(x)p^i p^j} = \sqrt{2F}$, the functional $I(\gamma) = \int_{t_1}^{t_2} \sqrt{g_{ij}(x)\dot{x}^i\dot{x}^j} \; dt$ is the **arc length** of γ. The Euler-Lagrange equations $\frac{d}{dt} G_{p^i} = G_{x^i}$ can using the previous function F be written as $\frac{d}{dt}\frac{F_{p^i}}{\sqrt{2F}} = \frac{F_{x^i}}{\sqrt{2F}}$ Which means $\frac{d}{dt} F_{p^i} = F_{x^i}$ because $\frac{d}{dt} F = 0$. Even so we got the same equations as for the free motion, they are not equivalent: a reparametrization of time $t \mapsto \tau(t)$ leaves only the first equation invariant and not the second. The distinguished parameterization for the extremal solution is proportional to the arc length. The relation between the two variational problems for energy and arc length is a special case of the **Maupertius principle**.

EXAMPLE: GEODESICS ON THE HYPERBOLIC PLANE. This is an example, where the surface is not given as an embedded surface in R^3. Instead, we assume that the distance on the upper half plane H is given by the formula

$$L(\gamma) = \int_a^b \frac{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}}{y(t)} \; dt \; .$$

THEOREM. On the hyperbolic plane, geodesics between two points P, Q is the circle through P, Q which hits the x axes in right angles.

PROOF. For points $P = (x,a), Q = (x,b)$ with the same x coordinate, the distance is $d(P,Q) = \int_a^b y'(t)/y(t) \; dt = |\log(b/a)|$. The geodesic connection is a line. Now see H as part of the complex plane and note that **Moebius transformation**



$$T(z) = \frac{(az+b)}{(cz+d)}$$

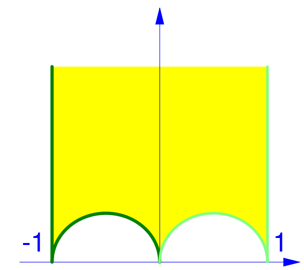
with $ad - bc = 1$ maps circles to circles or lines is an isometry: $d(P,Q) = d(T(P), T(Q))$. Indeed, the two formulas $\text{Im}(T(z)) = \text{Im}(z)/|cz+d|^2$ and $d/dt T(z(t)) = z'(t)/|cz+d|^2$ imply

$$\int_a^b \frac{d/dt T(z(t))}{\text{Im}(T(z(t))} \; dt = \int_a^b \frac{z'(t)}{\text{Im}(z(t))} \; dt \; .$$

To see that a Moebius transformation preserves circles, note that one can write T as a composition $T = T_2 I T_1$, where $T_1(z) = cz + d, T_2(z) = a/c + (ad - bc)z/c$ and where $I(z) = 1/z$ is the inversion at the unit circle. Because all three transformations preserve circles also A circle through the origin is maped into a line. If a, b, c, d are real, then T maps the upper half plane onto itself.

CHAOTIC GEODESIC FLOW. We have seen that the cat map $T(x,y) = (2x+y, x+y)$ is integrable and harmless on the plane. You have computed in a homework an integral, a function $F(x,y)$ which is invariant under T. When projecting the map onto the torus R^2/Z^2, then chaos happens. We have seen that the map allows a description by a symbolic dynamical system. Especially, it is chaotic in the sense of Devaney. A similar thing happens when we look at the geodesic flow on the upper half plane H. The orbits are circles. Even so you have sensitive dependence on initial conditions (as you can see in the picture above that if you start with different direction from the same point, the trajectories separate fast). We can do the analogue of the torus construction on the hyperbolic plane: take a discrete subgroup Γ of the group of all Möbius transformations.

For example Γ could be the subgroup of Möbius transformations with integer entries. It is called the **modular group**. An other subgroup is the **modular group lambda** Λ of all transformation $T(z) = (az+b)/(cz+d)$, where a, d are odd integers and b, d are even integers. The equivalent region to the square in the case of the torus is the **fundamental region** H/Λ which is displayed to the right. Billiard trajectories move on circles, when hitting the the boundary z of the region they enter at an other place $\gamma(z)$ similar than Pacman does for the torus. The corresponding flow is chaotic for any known notion of chaos.



THE DOUGHNUT. The rotationally symmetric torus in space is parameterized by

$r(u,v) = ((a + b\cos(2\pi v))\cos(2\pi u), (a + b\cos(2\pi v))\sin(2\pi u), b\sin$

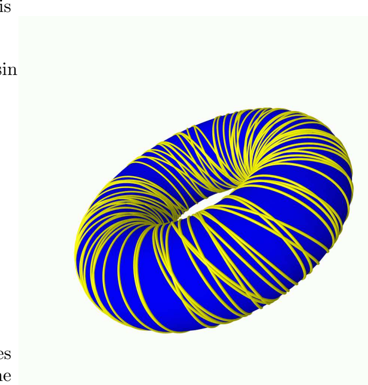
where $0 < b < a$. The metric is

$$\begin{aligned} g_{11} &= 4\pi^2(a + b\cos(2\pi v))^2 = 4\pi^2 r^2 \\ g_{22} &= 4\pi^2 b^2 \\ g_{12} &= g_{21} = 0 \end{aligned}$$

so that length of a curve is measured with the formula

$$\int_a^b 4\pi^2 (r(u(t), v(t)))^2 \dot{u}^2 + b^2\dot{v}^2) \; dt \; .$$

The circles $v = 0, v = 1/2$ are geodesics as are all the circles $u = u_0$. The surface is rotationally symmetric and one has the Clairot integral.



HOPF-RYNOV THEOREM ETC. The geodesic flow is defined for all times for closed complete surfaces without boundary. On every point on the surface and in any direction, there exists exactly one geodesic curve. Every geodesic subsegment of a geodesic curve is a geodesic curve. The shortest path between two points on the surface is a geodesic. But as the sphere shows, not every geodesic is the shortest path (you might go into the wrong direction on the grand circle). If two points are close enough, then the shortest geodesic connecting the two points is the shortest curve.

REMARKS. It is not custom to **define** the geodesic flow by constraining the free flow to the surface. But it is a useful fact and used for proving the integrability of the geodesic flow on the ellipsoid. The construction works in general: the **Nash embedding theorem** assures that any Riemannian surface can be embedded isometrically in an Euclidean space.