Geodesics on Surfaces of Constant Gaussian Curvature Using Mathematica

Jacob Lewis*

Columbia University

Abstract: We describe surfaces and geodesics without assuming prior knowledge of differential geometry. This involves selecting and presenting basic definitions and theorems. Included in this discussion are definitions of surface, coordinate patch, curvature, geodesic, etc. This summary closes with a proof of the length-minimizing properties of geodesics. Examples of surfaces of constant gaussian curvature are given and plotted in Mathematica. We also describe geodesics on these surfaces and plot select examples.

* The author wishes to thank Professor Charles Doran for his gracious help.

A well-known adage is, "The shortest distance between two points is a straight line." This is certainly true on the plane, but on other surfaces the adage proves to be false. For example, assume the earth to be a sphere. New York City and Madrid, Spain are both at latitudes of about 40°N. Yet an airplane taking the shortest distance between the two does not follow the 40^{th} parallel. Rather, it arcs north, following the great circle (i.e., circle centered at the sphere's center) between the two cities.

A *surface* in three-dimensional Euclidean space ($\mathbf{R}^3$) is a set of points in $\mathbf{R}^3$ that locally look like a plane—that is, given any point on the surface, there is a small neighborhood of that point which appears to be planar. Again, the earth's surface taken as a sphere is a good example. The earth's surface curves, yet by looking around, one cannot see this curvature. This is because the area of the earth one can see is a small enough neighborhood of the point where he/she is standing that this neighborhood appears flat. So the sphere is a surface in $\mathbf{R}^3$. More technically,

> **Definition:** $M \subset \mathbf{R}^3$ is a surface iff for any $x \in M$, there exists an open neighborhood $U \subset \mathbf{R}^3$ containing x, an open neighborhood $W \subset \mathbf{R}^2$, and a map $\mathbf{x}: W \to U \cap M$ that is differentiable with differentiable inverse. Such a function is called a parameterization or a coordinate patch since it allows us to assign coordinates to the surface corresponding to the coordinates of $\mathbf{R}^2$.

For a sphere of radius r centered at the origin, a coordinate patch for a hemisphere is $\mathbf{x}(u, v) = (r\cos(u)\cos(v), r\sin(u)\cos(v), r\sin(v))$, and patches for the northern, southern, eastern, and western hemispheres suffice to parameterize the entire sphere. A coordinate patch is said to be orthogonal if its first partial derivatives are orthogonal—that is, if $\mathbf{x}_u \bullet \mathbf{x}_v = 0$. If any two points on a surface can be connected by a curve contained in the surface, the surface is said to be connected.

Three useful and important constructs on a surface M are the tangent plane to the surface at a point p, a frame field on the surface, and the shape operator. Parameterize M in a neighborhood of p by $\mathbf{x}(u, v)$, with $\mathbf{x}(u0, v0) = p$. Then, the tangent plane to M at p—denoted by T_pM—is the two dimensional vector space spanned by $\{\mathbf{x}_u(u_0, v_0), \mathbf{x}_v(u_0, v_0)\}$. It is fairly easy to show that this space is equivalent to the space of all vectors $\mathbf{v}$ such that $\mathbf{v} = \alpha'(t_0)$, where α is a curve on M with $\alpha(t_0)=p$ (see [3]). Since T_pM is a vector space, an inner product can be defined on it. If an inner product is defined consistently on every tangent plane of M, then M is said to be a geometric surface.

A frame field on a surface M is a set of three orthogonal one-dimensional unit-vector fields E_1, E_2, E_3 on M defined at every point of M such that E_3 is everywhere normal to M. Given a vector $v \in T_pM$, the covariant derivative $\nabla_v E_3$ can be defined as follows: define a function F from the real numbers to the set of vectors tangent to M by $F = E_3(p + tv)$, then define $\nabla_v E_3 = F'(0)$. (This definition applies not only to normal fields on a surface, but to any vector field in $\mathbf{R}^3$.) Since the covariant derivative gives the rate of change of the unit normal of M in the direction of v, it tells something about the shape of the surface. For example, if $\nabla_v E_3$ is 0, then the surface is flat in the direction of v. Therefore, given a unit normal field E_3 to a surface M, for any point p of M, the shape operator $S_p: T_pM \rightarrow T_pM$ is defined as $S_p(v) = -\nabla_v E_3$. (The minus sign is conventional.) It can be shown that the shape operator is linear [2]. Since it is a linear map from a two-dimensional space to a two-dimensional space, it can be represented as a two-by-two matrix. The shape operator allows a way to measure how a surface curves. There are several different measures of curvature; the one used here is Gaussian curvature. Gaussian curvature is defined as the determinant of the shape operator, K = det S.

Given two surfaces M and N, a function F: M $\rightarrow$ N can be defined. If **x**: $\mathbf{R}^2 \rightarrow$ M is a patch for a neighborhood U of M, and **y**: $\mathbf{R}^2 \rightarrow$ N is a patch for a neighborhood V $\supset$ F(U) of N, note that $\mathbf{y}^{-1}F\mathbf{x}$: $\mathbf{R}^2 \rightarrow \mathbf{R}^2$. F is said to be differentiable if $\mathbf{y}^{-1}F\mathbf{x}$ is differentiable in the normal sense for $\mathbf{R}^2$. A differentiable function with a differentiable inverse is called a diffeomorphism. Given differentiable F: M $\rightarrow$ N and a point p$\in$M, define the tangent map F_*: $T_pM \rightarrow T_{F(p)}N$ as follows: let $\alpha \in$ M be curve on M such that α(t0) = p and $\alpha'(t_0) = v \in T_pM$. Let $\beta = F(\alpha)$, and define $F_*(v) = \beta'(t_0)$. If F_*: $T_pM \rightarrow T_{F(p)}N$ is one-to-one for every p$\in$M, then F is said to be regular. It is a useful fact that a one-to-one regular differentiable function M $\rightarrow$ N is a diffeomorphism.

Important curves on surfaces are curves called geodesics. Geodesics are essentially the extensions into M of straight lines in the plane—that is, relative to the surface, there appears to be no acceleration. Formally,

> **Definition:** For a surface M in Euclidean three-space, a geodesic is a curve α:[0, 1]$\rightarrow$M where α'' is always normal to M.

Given an orthogonal coordinate patch **x** in a geometric surface M and frame field E_1, E_2, E_3, geodesics can be defined by differential equations called, appropriately, the geodesic equations. Consider a curve α in M. Express α(t) = $\mathbf{x}$(u(t), v(t)). Then, $\alpha' = \frac{\partial \mathbf{x}}{\partial u} u'(t) + \frac{\partial \mathbf{x}}{\partial v} v'(t) = \mathbf{x}_u u' + \mathbf{x}_v v'$, and so

(1) $\alpha'' = \mathbf{x}_u u'' + u'(\mathbf{x}_{uu} u' + \mathbf{x}_{uv} v') + \mathbf{x}_v v'' + v'(\mathbf{x}_{uv} u' + \mathbf{x}_{vv} v')$

Since $E_1 = \mathbf{x}_u/||\mathbf{x}_u||$, $E_2 = \mathbf{x}_v/||\mathbf{x}_v||$, and $E_3 = E_1 \times E_2$ are a frame for T_pM, a curve α is a geodesic if and only if

(2) $\alpha''\bullet\mathbf{x}_u = 0$ and $\alpha''\bullet\mathbf{x}_v = 0$

So, using (1) and (2) and the fact that $\mathbf{x}_u\bullet\mathbf{x}_v = 0$, yields the differential system

(3) $\mathbf{x}_u u'' + u'^2\, \mathbf{x}_{uu}\bullet\mathbf{x}_u + u'v'\mathbf{x}_{uv}\bullet\mathbf{x}_u + v'^2\mathbf{x}_{vv}\bullet\mathbf{x}_u = 0$

(4) $\mathbf{x}_v v'' + v'^2\, \mathbf{x}_{vv}\bullet\mathbf{x}_v + u'v'\mathbf{x}_{uv}\bullet\mathbf{x}_v + u'^2\mathbf{x}_{uu}\bullet\mathbf{x}_v = 0$

which a curve must satisfy to be a geodesic. An immediate result of this system of differential equations is the following theorem:

Theorem: Given a regular surface M, a point $p\in M$, and vector $v\in T_pM$, there exists a unique geodesic γ such that γ(0)=p and γ′(0)=v.

Proof: Let γ(t)=$\mathbf{x}$(u(t), v(t)). Then γ(0)=p gives initial conditions u(0) and v(0). γ′(0)=v gives initial conditions u′(0) and v′(0). Then, by the fundamental existence and uniqueness theorems of ordinary differential equations [1], γ exists and is unique. ⌧

If every geodesic can be extended infinitely without leaving the surface, then the surface is called a complete surface.

Section II: Length-Minimizing Properties of Geodesics

An interesting question is, given two points p and q on a regular surface M, what is the shortest distance a curve between the two points could have?

Definition: A curve segment α is <u>a shortest segment</u> from p to q if for any other curve segment β from p to q, $L(\beta) \geq L(\alpha)$. α is <u>the</u> shortest segment from p to q if for any other curve segment β from p to q, $L(\beta)=L(\alpha)$ implies that β is a reparameterization of α.

Where p and q are relatively close, these shortest curves are always geodesics. To show this requires an examination of the relationship between M and its tangent plane at a point.

Definition: Given $p \in M$, define the <u>exponential map</u> $\exp_p: T_pM \to M$ by $\exp_p(v)=\gamma_v(1)$ where γ_v, defined on [0, 1] is the geodesic starting at p with initial velocity v.

Note that $\exp_p(t) = g_v(t)$. Since γ_v is a solution of the geodesics equations (3) and (4), it can be shown that exp is a differentiable function [3]. To see this, notice that for a fixed t, the curve $\beta(s) = \gamma_v(ts)$ is a geodesic with initial velocity $b'(0) = t\,\gamma_v'(0) = tv$, and the geodesic γ_{tv} also has, by definition, initial velocity tv. Thus, by the uniqueness of geodesics, $\beta = \gamma_{tv}$, and so $\exp_p(tv)=\gamma_{tv}(1)=\beta(1)=\gamma_v(t)$. Thus, exp sends radial lines $t \to tv$ of T_pM to <u>radial geodesics</u> $t \to \gamma_v(t)$ of M. Without loss of generality it can be assumed that radial geodesics are unit-speed, since otherwise the curve could be reparametrized to be unit-speed. This shows that

Theorem: exp maps a neighborhood U of T_pM diffeomorphically to a neighborhood N, centered at p, of M.

Proof: It is known that exp is differentiable, so it remains to show that it is one-to-one and regular. Assume for $v, w \in T_pM$ that $\exp(v) = \exp(w)$. Then

The neighborhood N above is called a normal neighborhood; if the radius of U is ε, N is called a normal ε-neighborhood.

Let $\mathbf{e}_1$, $\mathbf{e}_2$ be a frame for T_pM. Then $\mathbf{y}(u, v) = u \cos v\, \mathbf{e}_1 + u \sin v\, \mathbf{e}_2$ is a regular patch for T_pM.

Definition: For a normal ε-neighborhood N of M, define the <u>geodesic polar parameterization</u> by $\mathbf{x}(u, v) = \exp_p(\mathbf{y}(u, v))$.

Now, with these definitions, we may proceed to the central issue of this section:

Theorem: For each point q in a normal ε-neighborhood N of p, the radial geodesic from p to q is the shortest curve from p to q.

Proof: Let **x** be the geodesic polar parameterization of N. First, note that for constant v, $\mathbf{x}(u,v_0)$ is a radial geodesic. So $\mathbf{x}_u=1$, which implies that E=1 and $\mathbf{x}_{uu}=0$. Also, $F_u=\langle \mathbf{x}_u, \mathbf{x}_v\rangle_u=\langle \mathbf{x}_u, \mathbf{x}_{vu}\rangle=\langle \mathbf{x}_u, \mathbf{x}_{uv}\rangle=(1/2)E_v=0$, so F is constant. Since F(0, v)=0, then, F=0. Moreover, **x** is regular, so $EG - F^2=G>0$.

Let $q=(u_0, v_0)$. Then, $\gamma(u)=\mathbf{x}(u, v_0)$. Let α be an arbitrary curve from p to q. Assume $L(\alpha)\leq L(\gamma)$. Without loss of generality, α contains no loops, for if it did we could discard the loops, thereby shortening α. Therefore, we can write $\alpha(u) = \mathbf{x}(a_1(u), a_2(u))$, with $0 \leq u \leq u_0$. Since $\mathbf{x}(a_1(0), a_2(0)) = \mathbf{x}(0, 0) = p$, and $\mathbf{x}(a_1(u_0), a_2(u_0)) = \mathbf{x}(u_0, v_0) = q$, we see that $a_1(0)=a_2(0)=0$, $a_1(u_0)=u_0$, and $a_1(v_0)= v_0$.

Since E=1 and F=0,

(5) $$\|\alpha\| = \sqrt{Ea_1'^2 + 2Fa_1'a_2' + Ga_2'^2} = \sqrt{a_1'^2 + Ga_2'^2} \geq \sqrt{a_1'^2} = a_1'$$, so

(6) $$\mathrm{L}(\alpha) = \int_0^{u_0} \sqrt{(a_1')^2 + G(a_2')^2}\, du \geq \int_0^{u_0} a_1' du = u_0$$

Now γ has unit speed, so $\mathrm{L}(\gamma) = \int_0^{u_0} du = u_0$, so $L(\alpha)\geq L(\gamma)$. If $L(\alpha)=L(\gamma)$, then

(7) $$\int_0^{u_0} \sqrt{a_1'^2 + Ga_2'^2}\, du = \int_0^{u_0} a_1' du = u_0$$, so

$\sqrt{a_1'^2 + Ga_2'^2} = a_1'$. Then, since G>0, $a_2'=0$. So $a_2=v_0$, which implies that $\alpha(u)=\mathbf{x}(a_1(u), v_0)$, which is a reparametrization of γ. So γ is by definition the curve of smallest length. ∎

Geodesics, then, give a way of finding the distance between two points; geodesics give rise to a metric. If γ is a constant-speed geodesic from p to q, define the distance ρ between p and q on M as $\rho(p, q)=L(\gamma)$. An immediate result of this definition is that the image of a geodesic under an isometry (a map that preserves distances) is a geodesic.

Section III: Geodesics in Mathematica

A useful way to visualize surfaces and geodesics is to graph them in Mathematica. Mathematica can solve systems of first-order differential equations numerically, so it can provide approximations of solutions to the geodesic equations, and Mathematica's three-dimensional graphing capabilities are excellent. The definitive work on using Mathematica to study surfaces is [2], which includes a handy appendix giving the Mathematica code necessary to make many surfaces and curves, and to do various other things. The discussion here is largely an elaboration of the methods described in [2].

To graph a surface with patch $\mathbf{x}(u, v)$ in Mathematica is a relatively simple matter. First, input the equation for the patch. Second, select an appropriate rectangle P in $\mathbf{R}^2$, where $(u,v) \in P$, on which to graph. Lastly, graph using the built-in function ParametricPlot3D. The syntax is as follows: `ParametricPlot3D[x[u, v] // Evaluate, {u, u0, u1}, {v, v0, v1}, Lighting->True].` Note that the case of letters does matter in

Mathematica; `parametricplot3D` will not work. Several examples are given below.

To solve the geodesic equations in Mathematica, it is first necessary to convert the two second-order differential equations into a system of four first order differential equations as follows:

$$(8) \qquad \begin{cases} u' = p \\ v' = q \\ p' = \dfrac{-1}{\|\mathbf{x}_u\|^2}\left(\mathbf{x}_{uu} \bullet \mathbf{x}_u p^2 + 2\mathbf{x}_{uv} \bullet \mathbf{x}_u pq + \mathbf{x}_{vv} \bullet \mathbf{x}_u q^2\right) \\ q' = \dfrac{-1}{\|\mathbf{x}_v\|^2}\left(\mathbf{x}_{uu} \bullet \mathbf{x}_v p^2 + 2\mathbf{x}_{uv} \bullet \mathbf{x}_v pq + \mathbf{x}_{vv} \bullet \mathbf{x}_v q^2\right) \end{cases}$$

Now, these equations are inputted into Mathematica, as in Figure 1. To make this easier, the partial derivatives of **x** are designated as functions of **x** before inputting the equations.

```
xu[x_][u_, v_] := D[x[uu, vv], uu] /. {uu → u, vv → v}
xv[x_][u_, v_] := D[x[uu, vv], vv] /. {uu → u, vv → v}
xuu[x_][u_, v_] := D[xu[x][uu, vv], uu] /. {uu → u, vv → v}
xuv[x_][u_, v_] := D[xu[x][uu, vv], vv] /. {uu → u, vv → v}
xvv[x_][u_, v_] := D[xv[x][uu, vv], vv] /. {uu → u, vv → v}
geo[x_][p_, q_, u_, v_] :=
 {p, q, (-1/Simplify[xu[x][u, v].xu[x][u, v]])
   (Simplify[xuu[x][u, v].xu[x][u, v]] p^2 +
     2 Simplify[xuv[x][u, v].xu[x][u, v]] p q +
     Simplify[xvv[x][u, v].xu[x][u, v]] q^2),
  (-1/Simplify[xv[x][u, v].xv[x][u, v]])
   (Simplify[xuu[x][u, v].xv[x][u, v]] p^2 +
     2 Simplify[xuv[x][u, v].xv[x][u, v]] p q +
     Simplify[xvv[x][u, v].xv[x][u, v]] q^2)}
```

Figure 1

[2] and [3] both give more streamlined presentations of the geodesic equations by first introducing more auxiliary functions. Using these in Mathematica would be equally effective. As an example of the output, `geo` is calculated for a sphere in Figure 2. Above the black line is the input; below is the output.

```
sphere[r_][u_, v_] :=
 {r Cos[u] Cos[v], r Cos[v] Sin[u], r Sin[v]}
geo[sphere[1]][p, q, u, v]
```

```
{p, q, 2 p q Tan[v], -p^2 Cos[v] Sin[v]}
```

Figure 2

Mathematica is able to solve the equations numerically by evaluating at a sequence of points and interpolating to estimate the behavior between the points. The code in Figure 3 is a slight modification of that in [2].

```
solvegeoeqs[x_, a_: 0, {u0_: 0, v0_: 0}, ang_, tfin_,
  div_, optsnd___] :=
 Flatten[Map[NDSolve[#, {u, v}, {ss, 0, tfin}] &,
   Table[
    Join[MapThread[Equal,
      {{u''[ss], v''[ss]},
       Take[geo[x][u'[ss], v'[ss], u[ss], v[ss]], {3, 4}]}],
     {u[a] == u0, v[a] == v0, u'[a] == Cos[theta],
      v'[a] == Sin[theta]}],
    {theta, 2 Pi / div + ang, 2 Pi + ang, 2 Pi / div}]], 1]
```

Figure 3

The arguments of `solvegeoeqs` are as follows:

- x: The surface in question.
- a: The initial starting point of the sequence (almost always zero).
- {u0, v0}: The values of u and v such that **x**(u,v)=p, where p is the starting point of the geodesic.

- ang: The angle formed between the initial velocity vector of the geodesic and a horizontal line. Essential, {u0, v0} and ang are the initial conditions.
- tfin: The program solves from t=0 to t=tfin. Changing tfin changes the length of the solved geodesic.
- div: This program can give solutions for a variety of angles at once. div specifies the number of geodesics the program should solve for. To see one geodesic extending in both directions, set div=2.
- optsnd: Various options, such as lighting, can be given if desired.

In [2], it is recommended that `geo[x]` be solved in one notebook, and defined to be equal to the output in the first notebook in a second notebook. Then, `solvegeoeqs` should be run in the second notebook. This saves the time of `solvegeoeqs` recalculating `geo[x]`.

Section IV: Geodesics on Surfaces of Constant Curvature

An interesting class of surfaces on which to examine the geodesics is the class of connected surfaces with constant gaussian curvature. Particularly simple will be the complete surfaces with constant gaussian curvature.

There are only three simply connected complete surfaces with constant curvature [3]. The first is the sphere of radius r, which has curvature $1/r^2$. A patch for the sphere of radius r (minus the north and south poles) is **x**(u, v)=(r cosv cosu, r cosv sinu, r sinv). Figure 4 below shows a sphere of radius one plotted in Mathematica.

```
sphere[a_][u_, v_] :=
 a {Cos[u] Cos[v], Cos[v] Sin[u], Sin[v]}
SpherePic = ParametricPlot3D[
   sphere[1][u, v] // Evaluate, {u, 0, 2 Pi},
   {v, -Pi / 2, Pi / 2}, PlotPoints → {24, 20},
   Lighting → True];
```

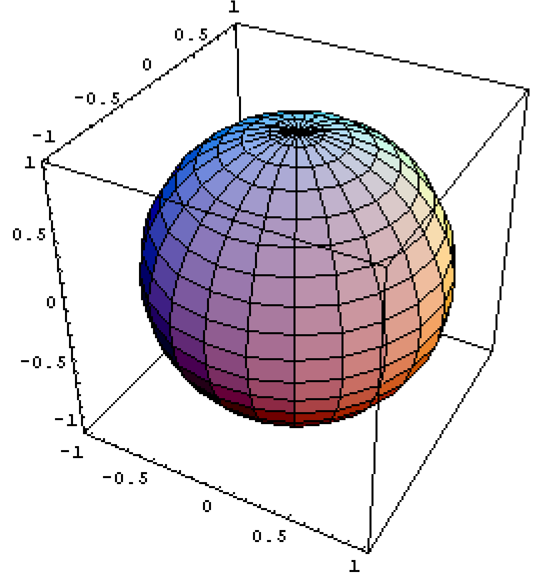

Figure 4

The plane also fits into this category. A plane has zero curvature. A patch for the plane consisting of points (a_1, b_1, c_1), (a_2, b_2, c_2), and $(a_1 + a_2, b_1 + b_2, c_1 + c_2)$ is $\mathbf{x}(u, v) = (a_1 u + a_2 v, b_1 u + b_2 v, c_1 u + c_2 v)$. Figure 5 below shows a plane passing through the points (2,1,0) and (1,2,3).

```
plane[x1_, y1_, z1_, x2_, y2_, z2_][u_, v_] :=
 {x1 u + x2 v, y1 u + y2 v, z1 * u + z2 * v}
PlanePic = ParametricPlot3D[
   plane[2, 1, 0, 1, 2, 3][u, v] // Evaluate,
   {u, -1, 1}, {v, -1, 1}, PlotPoints → {10, 10},
   Lighting → True];
```

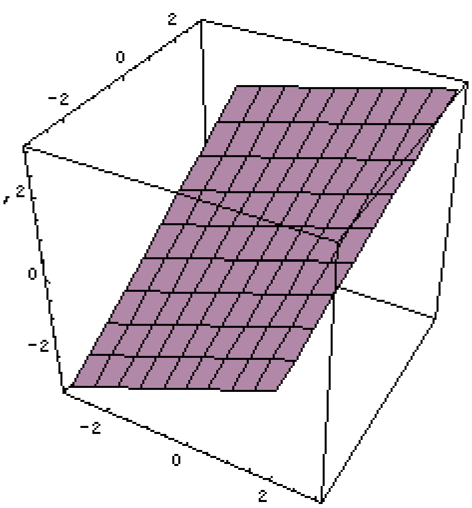

Figure 5

The third simply connected complete surface with constant curvature is the hyperbolic plane, which has curvature -1. The hyperbolic plane is the foundation of hyperbolic geometry. For more information, see [4]. In the hyperbolic plane, Euclid's fifth axiom fails; given a line (i.e. geodesic) in the hyperbolic plane and a point in the hyperbolic plane not on the line, there exist infinitely many lines through the point parallel to the line. The hyperbolic plane cannot be neatly pictured. This is because there is no way to embed the hyperbolic plane in $\mathbf{R}^3$ without distorting distances. There are, however, several useful models of the hyperbolic plane. One of these is the "upper half-plane" model. In this model, the hyperbolic plane is modeled based on the complex plane as $\mathbf{U}^2 = \{x + yi \in \mathbf{C} \mid y > 0\}$. Distances in this model, however, are not Euclidean distances. The distance between two points $z_1 = x_1 + y_1 i$, $z_2 = x_2 + y_2 i$ in $\mathbf{U}^2$ is given by $\mathrm{d}_{\mathrm{U}}(z_1, z_2) = \cosh^{-1}\left(1 + \frac{|z_1 - z_2|}{2y_1y_2}\right)$. The element of arc-length ds_U in this model is given in terms of the Euclidean arc-length element $\mathrm{d}s_E = \sqrt{\mathrm{dx}^2 + \mathrm{dy}^2}$ as $ds_U = ds_E/y$.

On the sphere, the geodesics are arcs of great circles. Any two non-antipodal points p, q on a sphere can be connected by two arcs of a great circle. To see that these arcs are geodesics, first consider the special case that the sphere is centered at the origin and p, q lie in the x-y plane. In this case, the equation for the great circle is $\alpha(t) = (r\cos(t), r\sin(t), 0)$. Setting $\alpha(t) = \mathbf{x}(u(t), v(t))$, then, implies that $u(t) = t$ and $v(t) = 0$. So u'', v', and v'' all equal zero. Now, $u'^2\mathbf{x}_u \bullet \mathbf{x}_{uu} = 1(-r\sin t, r\cos t, 0) \bullet (-r\cos t, -r\sin t, 0) = r^2\cos t\sin t - r^2\cos t\sin t = 0$, so this great circle satisfies the geodesic equations.

Note that for any p, q on a sphere of radius r centered at the origin, a rotation ϕ of the sphere can be found taking p, q to p', q' in the x-y plane. Note too that for a sphere of radius r centered at any point (x, y, z) in $\mathbf{R}^3$, a translation τ can be found taking this sphere to the sphere of radius r centered at the origin.

Next, note that rotations and translations are isometries. This is intuitively obvious; see [3] for details. Lastly, note that given p, q on a sphere of radius r at a point in $\mathbf{R}^3$, the great circle α connecting $p' = \tau^{-1}\phi^{-1}(p)$, $q' = \tau^{-1}\phi^{-1}(q)$ in the x-y plane is sent to the great circle γ connecting p and q by the isometry $\phi\tau$. So, since isometries preserve geodesics and α is a geodesic, γ is a geodesic. Geodesics are unique, so the two arcs of γ are the only two geodesics (ignoring reparametrization) connecting p and q. Of these two, one is shorter. This arc is the shortest curve between p and q. If the Euclidean angle between p and q is θ, $\rho(p, q)=L(\gamma)=r\theta$. Figure 6 below shows four geodesic great circles passing through a point on the sphere.

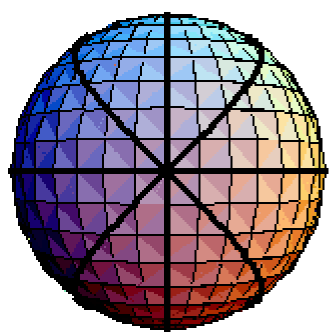

Figure 6

The code for producing this picture is in Figure 7 below. Note that the length of the geodesic (tfin) is less than $\pi/2$; at $\pi/2$, a problem arises because the patch used for the sphere does not include antipodal points.

```
geo[sphere[1]][p_, q_, u_, v_] := {p, q, 2 * p * q * Tan[v], Cos[v] * p^2 * Sin[v]}
sd1 = solvegeoeqs[sphere[1], 0, {0, 0}, 5 Pi / 12, 8];
g1 = ParametricPlot3D[Evaluate[Append[sphere[1] @@ Sequence[{u[t], v[t]}], AbsoluteThickness[3]] /. sd1], {t, 0, 5 Pi / 12},
   Axes → None, Boxed → False];
g2 = ParametricPlot3D[sphere[1][u, v], {u, 0, 2 Pi}, {v, 0, 2 Pi}, PlotPoints → {32, 32}, Axes → None, Boxed → False, Lighting → True];
Show[g1, g2, Lighting → True, ViewPoint → {4, 0, 0}];
```

Figure 7

On the plane, the mapping exp is the identity map. So geodesics are straight lines. This confirms the adage "The shortest distance between two points is a straight line." $\gamma(t)=p + t(q-p)$, so $\rho(p, q) = L(\gamma) = ||\gamma'||dt = ||q-p||$. So ρ is the Euclidean distance formula, as expected.

To examine what the geodesics in the upper half-plane model for the hyperbolic plane are, first consider the special case of a geodesic between two points y_1i, $y_2i \in \mathbf{U}^2$. Clearly, the shortest curve connecting these two is the curve $\alpha(t) = ti$. The isomorphisms of **U**2 are functions of the type $f(z) = (az + b)/(cz + d)$, where a,b,c,d are real numbers such that $ad-bc \neq 0$. Direct verification of this using the given metric is messy; see [4] for a neater method. Now consider two points $p,q \in \mathbf{U}^2$. If their real parts $Re(p) = Re(q) = x_0$, then the isomorphism $f(z) = z + x_0$ sends α to the geodesic $x_0 + ti$ passing through p and q. If $Re(p) \neq Re(q)$, then there exists a Euclidean circle β centered at a point on the real axis passing through p and q. Designate by x_1, x_2 the points at which the graph of this circle intersect the real axis. Now, the isometry $f(z) = (x_2z - x_1)/(z - 1)$ sends α to the upper half of β. So the geodesic of the upper half-plane are vertical lines and the upper halves of Euclidean circles centered on points of the real axis.

Aside from these three surfaces, there are many (in fact, infinitely many) other surfaces of constant curvature. The cylinder, for example, has zero curvature. A patch for the cylinder of radius r is **x**(u, v) = (r cos(u), r sin(u), v). In general, geodesics on the cylinder are segments of helices, as can be checked using the geodesic equations. The general form for a helix on a cylinder of radius r (centered around the z-axis) with slope m is α(t)=(r cos(t), r sin(t), mt). The initial velocity in this case is

(8) $$\|\alpha'(0)\| = \sqrt{\alpha'(0)\bullet\alpha'(0)} = \sqrt{r^2\sin^2 t + r^2\cos^2 t + m^2} = \sqrt{r^2+m^2}\ .$$

A constant speed reparameterization is

$$\beta(t) = \left(r\cos\left(\frac{t}{\sqrt{r^2+m^2}}\right), r\sin\left(\frac{t}{\sqrt{r^2+m^2}}\right), \frac{mt}{\sqrt{r^2+m^2}} \right).$$ Expressing β(t) as **x**(u(t), v(t)) in terms of the standard patch for the cylinder, **x**(u, v) = (r cos(u), r sin(u), v), it is clear that $u(t) = \frac{t}{\sqrt{r^2+m^2}}$ and $v(t) = \frac{mt}{\sqrt{r^2+m^2}}$. u''(t), v''(t), $\mathbf{x}_{uv}$, and $\mathbf{x}_{vv}$ are all zero, so the only nontrivial terms in the two geodesic equations are

$u'^2\mathbf{x}_u \bullet \mathbf{x}_{uv} = \frac{r^2\cos t\sin t - r^2\cos t\sin t}{r^2+m^2} = 0$ and $u'^2\mathbf{x}_{uu}\bullet\mathbf{x}_v$, which equals zero since $\mathbf{x}_{uu}$ and $\mathbf{x}_v$ are clearly orthogonal. So β(t) satisfies the geodesic equations. One example is given in Figure 8, where r=1, m=1. Two noteworthy degenerate cases are when the initial velocity is horizontal (m=0), and when it is vertical (m=∞). In the first case, the geodesic is a circle, as in Figure 9; in the second, the geodesic is a vertical line, as in Figure 10.

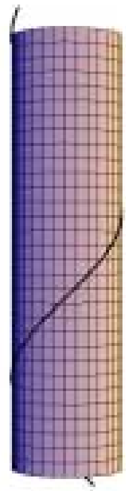
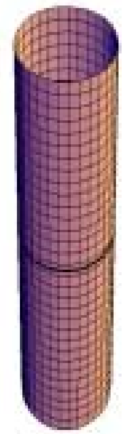
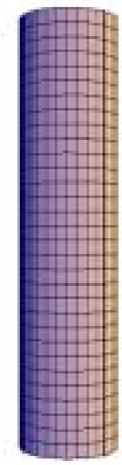

Figure 8 Figure 9 Figure 10

The distance between points p and q on a cylinder can be evaluated as the arc-length of the associated geodesic segment of a helix. For example, on the cylinder below the distance between (1,0,0) and (0,1,1) is $(1 + (\pi/2)^2)^{1/2}$.

Another example of a surface with zero curvature is the flat torus. The "donut-shaped" flat torus can be visualized as the product of two circles of radii a and b. A patch for the flat torus is **x**(u, v) = (a cos u, a sin u, b cos v, b sin v). The flat torus can be constructed from a rectangle (for simplicity's sake, a 1x1 square will be used) by identifying the edges and vertices as shown below:

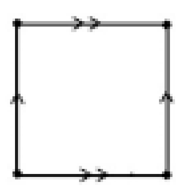

Figure 11

It is called "flat" because the Euclidean distances defined on the rectangle are preserved on the flat torus. (Note: the flat torus should not be confused with a torus in $\mathbf{R}^3$, formed from rotating a circle in the y-z plane about the z-axis. The flat torus cannot be put in $\mathbf{R}^3$.) Since distances are preserved, the identifications are isometries. As we have seen, geodesics are isometric invariants, so a geodesic on the rectangle will map to a geodesic on the flat torus under the identifications. An interesting question is, when are these geodesics closed curves? To answer this, assume without loss of generality that the geodesic (i.e. line) on the square begins at the vertex, and denote by θ the angle between this geodesic and the bottom of the square. In order for the geodesic on the flat torus to close, the line on the square must eventually reach the vertex. If $\tan\theta$ is rational, say a/b in lowest terms, then represent the flat torus by an lattice of (a x b) unit squares, where all sides and vertices of the squares are identified. For example, for $\tan\theta = 3/2$:

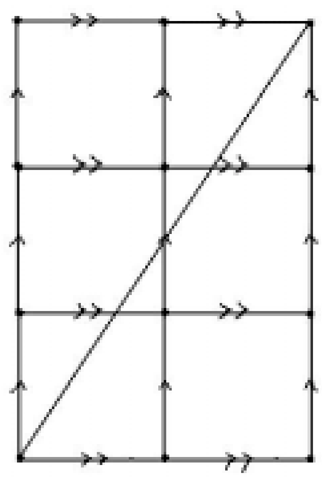

Figure 12

Then the line is the hypotenuse of a right triangle with legs a and b. This line goes from one vertex of the lattice to another. Therefore, under the identifications, this line maps to a closed geodesic. On the other hand, if $\tan\theta$ is irrational, then the line will never be the hypotenuse of a right triangle with legs of integral lengths. Therefore, the line will never return to a vertex, and the geodesic will never close. Two basic cases of closed geodesics on the torus are illustrated in Figure 9 below. One case is if tanθ = 0, in which case the geodesic is the longitude, the circle going around the outside of the torus. The other case shown is where tanθ is undefined. In this case, the geodesic is the meridian, the circle going vertically through the hole in the torus.

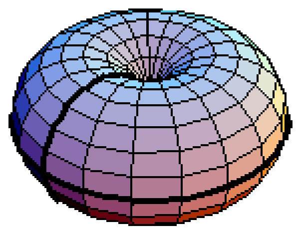

Figure 13

The concept of “line” is a very intuitive and fundamental concept in our everyday lives. Generalizing this concept to other surfaces is a very interesting mathematical exercise, and using the differential geodesic equations and solving them in Mathematica provides an easy way to see what “lines” on a surface will be.

References

[1] Braun, M. *Differential Equations and Their Applications*. New York: Springer-Verlag, 1983. p. 289.

[2] Gray, H. *Modern Differential Geometry of Curves and Surfaces with Mathematica®*. Second edition. Boca Raton: CRC Press, 1998. p. 613, 873-874, 939, 947, 960, 966-967.

[3] O’Neill, B. *Elementary Differential Geometry*. Second edition. San Diego: Academic, 1997. p. 98, 148, 320, 375, 418-419.

[4] Ratcliffe, J. *Foundations of Hyperbolic Manifolds*. New York: Springer-Verlag, 1994. § 4.6.