

Rose-Hulman Undergraduate Mathematics Journal

Volume 3
Issue 1

Article 3

Geodesics Using Mathematica

Jacob Lewis

Columbia University, jml104@columbia.edu

Follow this and additional works at: <https://scholar.rose-hulman.edu/rhumj>

Recommended Citation

Lewis, Jacob (2002) "Geodesics Using Mathematica," *Rose-Hulman Undergraduate Mathematics Journal*: Vol. 3 : Iss. 1 , Article 3.

Available at: <https://scholar.rose-hulman.edu/rhumj/vol3/iss1/3>

Geodesics 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, geodesic, etc. This summary closes with a proof of the length-minimizing properties of geodesics. Examples of surfaces are given and plotted in Mathematica. We also describe geodesics on these surfaces and plot select examples. The surfaces chosen include some with Clairaut patches, some without, and some surfaces in \mathbf{R}^3 and some not in \mathbf{R}^3 .

^{*} 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 40th parallel. Rather, it arcs north, following the great circle (i.e., circle centered at the sphere's center) between the two cities.

To generalize the adage--and along the way to explain why planes travel this way--we will introduce a special class of curves on surfaces, called geodesics. Geodesics have the useful property that the shortest curve segment connecting two points on a surface is a segment of a geodesic. As we shall see, great circles are geodesics on the sphere, and they therefore have the property that they are the "shortest" curves on the sphere. To examine geodesics, we will develop connections between differential geometry, differential equations, and vector calculus. In order to see geodesics, even when they cannot be found explicitly, the computer algebra system Mathematica will be used.

Differential Geometry

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 if 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 $\mathbf{x}: W \rightarrow U \cap M$ 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 .

Surfaces do not need to be in \mathbf{R}^3 ; replacing \mathbf{R}^3 with a general space X yields a valid definition. In fact, some of the surfaces we will examine cannot be placed in \mathbf{R}^3 .

For a sphere of radius r centered at the origin, a coordinate patch is $\mathbf{x}(u, v) = (r \cos(u) \cos(v), r \sin(u) \cos(v), r \sin(v))$, for $-\pi < u < \pi$, and $-\pi < v < \pi$. A coordinate patch is said to be orthogonal if its first partial derivatives are orthogonal—that is, if $\mathbf{x}_u \bullet \mathbf{x}_v = 0$. Clearly, for an orthogonal patch \mathbf{x} , $\mathbf{x}_u \times \mathbf{x}_v$ is never zero. This means it is possible to construct a unit normal $\mathbf{U} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{|\mathbf{x}_u \times \mathbf{x}_v|}$ at any point on the surface. Also, because \mathbf{x}_u and \mathbf{x}_v vary smoothly on M , so will \mathbf{U} . If any two points on a surface can be connected by a curve contained in the surface, the surface is said to be connected.

Using the sphere as an example, the parametrization of the sphere given above is orthogonal. $\mathbf{x}_u = (-r \sin(u) \cos(v), r \cos(u) \cos(v), 0)$, and $\mathbf{x}_v = (-r \cos(u) \sin(v), -r \sin(u) \sin(v), r \cos(v))$. So $\mathbf{x}_u \bullet \mathbf{x}_v = r^2 \sin(u) \sin(v) \cos(u) \cos(v) - r^2 \sin(u) \sin(v) \cos(u) \cos(v) + 0 = 0$. The unit normal at $\mathbf{x}(u, v)$ is

$\frac{1}{\sqrt{\cos^2(u) + \sin^2(v)}}(\cos^2(u), \cos(u)\sin(u), \sin(v))$, as the reader can easily verify either by hand or using Mathematica.

A useful and important construct on a surface M is the tangent plane to the surface at a point p . Parameterize M in a neighborhood of p by $\mathbf{x}(u, v)$, with $\mathbf{x}(u_0, v_0) = p$. Then, the tangent plane to M at p —denoted by $T_p M$ —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 $T_p M$ 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.

The Geodesic Equations

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 $\alpha: [0, 1] \rightarrow M$ where α'' is always normal to M .

Since α'' is always normal to M , that means that the dot product of α'' and a vector in $T_p M$ is always zero. In particular, $\alpha'' \cdot \alpha' = 0$. Let $s(t)$ be the speed of α at t ; $s(t) = \sqrt{\alpha'(t) \cdot \alpha'(t)}$. Differentiating s with respect to t ,

$$(1) \quad s'(t) = \frac{\alpha''(t) \cdot \alpha'(t)}{\sqrt{\alpha'(t) \cdot \alpha'(t)}} = 0$$

This shows that $s(t)$ is a constant function, and so a geodesic must be a constant-speed curve. In fact, we can parameterize any geodesic so that it is

unit-speed.

Given an orthogonal coordinate patch \mathbf{x} in a geometric surface M , geodesics can be defined by differential equations called, appropriately, the geodesic equations. Consider a curve α in M . Express $\alpha(t) = \mathbf{x}(u(t), v(t))$. Then,

$$\alpha' = \frac{\partial \mathbf{x}}{\partial u} u'(t) + \frac{\partial \mathbf{x}}{\partial v} v'(t) = \mathbf{x}_u u' + \mathbf{x}_v v', \text{ and so}$$

$$(2) \quad \alpha'' = \mathbf{x}_{uu} u'' + u'(\mathbf{x}_{uu} u' + \mathbf{x}_{uv} v') + \mathbf{x}_{vu} v'' + v'(\mathbf{x}_{uv} u' + \mathbf{x}_{vv} v')$$

If α is a geodesic, then it is normal to the surface, and hence

$$(3) \quad \alpha'' \bullet \mathbf{x}_u = 0 \text{ and } \alpha'' \bullet \mathbf{x}_v = 0$$

So, using (2) and (3) and the fact that $\mathbf{x}_u \bullet \mathbf{x}_v = 0$ yields the differential system

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

$$(5) \quad |\mathbf{x}_v|^2 v'' + v'^2 \mathbf{x}_{vv} \bullet \mathbf{x}_v + 2u'v' \mathbf{x}_{uv} \bullet \mathbf{x}_v + u'^2 \mathbf{x}_{uu} \bullet \mathbf{x}_v = 0$$

which a curve must satisfy to be a geodesic. These two equations are called the

geodesic equations, because solving them gives geodesics. In fact, these

equations allow us to generalize the definition of a geodesic to a surface that is

not in \mathbf{R}^3 . For such a surface, we merely define a geodesic to be a solution of

the geodesic equations. An immediate result of this system of differential

equations is the following theorem:

Theorem: Given a surface M in \mathbf{R}^3 with a unit normal, a point $p \in M$, and vector $v \in T_p M$, there exists a unique geodesic γ such that $\gamma(0) = p$ and $\gamma'(0) = v$.

Proof: Let $\gamma(t) = \mathbf{x}(u(t), v(t))$. Then $\gamma(0) = p$ gives initial conditions $u(0)$ and $v(0)$. $\gamma'(0) = v$ 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.

Usually, the geodesic equations cannot be solved by hand. For this reason, it is useful to be able to solve the geodesic equations numerically; we give a Mathematica procedure for this below. On a few surfaces, such as the sphere, explicit solutions can be found. One class of surfaces for which the geodesic equations can often be explicitly solved are surfaces given a Clairaut patch. A Clairaut patch is an orthogonal patch in which $|\mathbf{x}_u|$ and $|\mathbf{x}_v|$ are both

independent of v . This implies $\frac{\partial}{\partial v}(\mathbf{x}_u \bullet \mathbf{x}_u) = 2\mathbf{x}_{uv} \bullet \mathbf{x}_u = 0$ and

$\frac{\partial}{\partial v}(\mathbf{x}_v \bullet \mathbf{x}_v) = 2\mathbf{x}_{vv} \bullet \mathbf{x}_v = 0$, or simply $\mathbf{x}_{uv} \bullet \mathbf{x}_u = \mathbf{x}_{vv} \bullet \mathbf{x}_v = 0$. Also, since $\mathbf{x}_u \bullet \mathbf{x}_v = 0$,

$\frac{\partial}{\partial u}(\mathbf{x}_u \bullet \mathbf{x}_v) = \mathbf{x}_{uu} \bullet \mathbf{x}_v + \mathbf{x}_u \bullet \mathbf{x}_{uv} = \mathbf{x}_{uu} \bullet \mathbf{x}_v = 0$. Therefore, the geodesic equations

reduce in this case to

$$(6) \quad |\mathbf{x}_u|^2 u'' + u'^2 \mathbf{x}_{uu} \bullet \mathbf{x}_u + v'^2 \mathbf{x}_{vv} \bullet \mathbf{x}_u = 0$$

$$(7) \quad |\mathbf{x}_v|^2 v'' + 2u'v' \mathbf{x}_{uv} \bullet \mathbf{x}_v = 0$$

Isolating the v -terms in (7) gives

$$(8) \quad \frac{\frac{d^2 v}{dt^2}}{\frac{dv}{dt}} = \frac{v''}{v'} = -2u' \frac{\mathbf{x}_{uv} \bullet \mathbf{x}_v}{\mathbf{x}_v \bullet \mathbf{x}_v} = -\frac{du}{dt} \left(\frac{\frac{\partial G}{\partial u}}{G} \right), \text{ where } G = \mathbf{x}_v \bullet \mathbf{x}_v.$$

Integrating both sides of (8) with respect to t , and making the substitution $w =$

$$dv/dt, \int \frac{dw}{w} = -\int \frac{G_u}{G} du = -\int \frac{dG}{G} \Rightarrow \ln v' = -\ln G + c. \text{ Taking the exponential of both}$$

sides gives the equation

$$(9) \quad \frac{dv}{dt} = v' = \frac{k}{G}$$

for some constant k . Now we use the fact that we may assume the geodesic $\alpha(t) = \mathbf{x}(u(t), v(t))$ to be unit speed. This means that

$$1 = \alpha' \bullet \alpha' = (\mathbf{x}_u u' + \mathbf{x}_v v') \bullet (\mathbf{x}_u u' + \mathbf{x}_v v') = |\mathbf{x}_u|^2 u'^2 + G v'^2 = |\mathbf{x}_u|^2 u'^2 + \frac{k^2}{G}, \text{ or}$$

$$(10) \quad \frac{du}{dt} = u' = \pm \sqrt{\frac{G - k^2}{|\mathbf{x}_u|^2 G}}$$

Dividing (9) by (10) gives $\frac{dv}{du} = \pm \sqrt{\frac{|\mathbf{x}_u|^2 k^2}{G^2 - k^2 G}}$, and integrating this yields

$$(11) \quad v = \pm \int \sqrt{\frac{|\mathbf{x}_u|^2 k^2}{G^2 - k^2 G}} du$$

If for a particular surface with a Clairaut patch, the integral on the right-hand side of (11) can be solved explicitly, this gives an explicit solution to the geodesic equations for this surface. If $|\mathbf{x}_u|$ and $|\mathbf{x}_v|$ are independent of u , then the patch is said to be v -Clairaut, and equation (11) holds by interchanging u and v . In this case, we use the variable E instead of G .

Length Minimizing Properties of Geodesics

An important property of geodesics, alluded to earlier, is that if a shortest curve between two points on a surface exists, it is a segment of a geodesic. (A

shortest curve between two points might not exist if the surface is not geodesically complete.) To show this, we make the following definition:

Definition: A curve segment α is a shortest segment from p to q if for any other curve segment β from p to q , the length $L(\beta) \geq L(\alpha)$. α is the 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 .

Now, for any point p on a surface M , there exists a small neighborhood $N(p)$ of M around p for which we can give a patch $\mathbf{x}(u, v)$, called the geodesic polar patch, with the following properties: \mathbf{x} is orthogonal, $E = \mathbf{x}_u \bullet \mathbf{x}_u = 1$, and $G = \mathbf{x}_v \bullet \mathbf{x}_v > 0$. Details of how to define this patch are given in [4]. For q in $N(p)$, there is a geodesic segment connecting q and $N(p)$ which is contained completely in $N(p)$. If $q = \mathbf{x}(u_0, v_0)$, then this geodesic, called the “radial geodesic,” is given by $g(u) = \mathbf{x}(u, v_0)$. Now,

Theorem: For each point q in $N(p)$, the radial geodesic from p to q is the shortest curve from p to q .

Proof: Let \mathbf{x} be the above-mentioned parameterization of $N(p)$. 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 the patch is orthogonal,

$$(5) \quad \|\alpha'\| = \sqrt{E a_1'^2 + G a_2'^2} = \sqrt{a_1'^2 + G a_2'^2} \geq \sqrt{a_1'^2} = a_1', \text{ so}$$

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

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

$$(7) \quad \int_0^{u_0} \sqrt{a_1'^2 + G a_2'^2} du = \int_0^{u_0} a_1' du = u_0, \text{ so}$$

$\sqrt{a_1'^2 + G a_2'^2} = a_1'$. 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. ☒

This remarkable theorem gives an idea of why visualizing geodesics is so important to understanding the geometry of a surface. Just as lines on the plane give the shortest path between two points on the plane, geodesics perform this function on arbitrary surfaces. In a very real sense then, geodesics are the abstraction of the notion of a line to other surfaces. (Indeed, the theorem implies that straight lines are the geodesics on the plane.)

Mathematica and Geodesics

To be able to see geodesics on a surface—particularly useful in the absence of an explicit equation—we turn to the computer algebra system Mathematica. Mathematica can solve differential equations numerically using a built-in function named `NDSolve`, and it can graph a parametrized surface using a function named `ParametricPlot3D`. Note the capitalization; Mathematica commands are case-sensitive. 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 intended to streamline and generalize the methods described in [2] for dealing with geodesics.

To graph a surface with patch $\mathbf{x}(u, v)$ in Mathematica is a relatively simple matter. First, input the equation for the patch. To input a function, the argument goes in square brackets, an underscore follows the variables' names, and a colon precedes the equals sign. For example, to define the sphere, the command is: `sphere[u_,v_]:= {r Cos[u] Cos[v], r Sin[u] Cos[v], r Sin[v]}`. Second, select an appropriate rectangle $P = [u_0, u_1] \times [v_0, v_1]$ in \mathbb{R}^2 on which to graph. Lastly, graph using the built-in function `ParametricPlot3D`. The syntax is as follows: `ParametricPlot3D[x[u, v], {u, u0, u1}, {v, v0, v1}, Lighting->True]`. (The lighting is important; otherwise the graph is all black.) It may be convenient to view the surface from a particular viewpoint (x, y, z) ; in this case, after `Lighting->True`, a comma then `ViewPoint->{x,y,z}` is added.

The amount of time Mathematica takes to solve a system of differential equations increases as the order of the system increases. For this reason, it is convenient to reduce the order of the system by introducing the auxiliary variables $p = u'$ and $q = v'$. This yields the system

$$(12) \quad \begin{cases} u' = p \\ v' = q \\ p' = \frac{-1}{\|\mathbf{x}_u\|^2} (\mathbf{x}_{uu} \bullet \mathbf{x}_u p^2 + 2\mathbf{x}_{uv} \bullet \mathbf{x}_u pq + \mathbf{x}_{vv} \bullet \mathbf{x}_u q^2) \\ q' = \frac{-1}{\|\mathbf{x}_v\|^2} (\mathbf{x}_{uv} \bullet \mathbf{x}_v p^2 + 2\mathbf{x}_{vv} \bullet \mathbf{x}_v pq + \mathbf{x}_{vv} \bullet \mathbf{x}_v q^2) \end{cases}$$

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

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 geodesic equations numerically by evaluating at a sequence of points and interpolating to estimate the behavior between the points. In the following code in Figure 3, the variables $\{u_0, v_0\}$ represent the starting point of the geodesic; `ang` gives the initial direction of the geodesic, which combined with the fact that the geodesic is unit-speed, gives the initial

velocity. Therefore, $\{u_0, v_0\}$ and ang give the necessary initial conditions to give a unique geodesic.

```

solvetest1[x_, a_ : 0, {u0_ : 0, v0_ : 0}, ang_, tfin_,
  optsnd___] :=
NDSolve[
  {u'[ss] ==
    First[Take[geo[x][u'[ss], v'[ss], u[ss], v[ss]],
      {3, 3}]],
  v'[ss] ==
    First[Take[geo[x][u'[ss], v'[ss], u[ss], v[ss]],
      {4, 4}]], u[a] == u0, v[a] == v0, u'[a] == Cos[ang],
  v'[a] == Sin[ang]}, {u, v}, {ss, 0, tfin}]

```

Figure 3

`Take[... , {a,b}]` selects the a^{th} through b^{th} elements of a list and outputs them in another list; `First[...]` gives the first element of a list as a number. So `First[Take[geo..., {3,3}]` gives the numeric value of u'' .

To display the solution on a surface, two different plots are made and saved. First, the equations are solved: `solution1=solvetest1[x,0,{u0,v0},ang,tfin]`. Next, the geodesic is plotted:

```

g1=ParametricPlot3D[Evaluate[Append[x[u[t],
  v[t]], AbsoluteThickness[3]] /. Solution1],
  {t,0,tfin}],

```

and the surface is plotted:

```

g2=ParametricPlot3D[x[u,v], {u, u1, u2}, {v, v1,
  v2}, Lighting->True].

```

Lastly, the two plots are shown together: `Show[g1,g2,Lighting->True,ViewPoint->{x,y,z}]`. Finding the proper viewpoint depends on the

geodesic, and often requires trial and error. Now, we give specific examples of this procedure by finding and plotting geodesics on several surfaces.

The Sphere

For the patch for the sphere, we see that $G = \mathbf{x}_u \bullet \mathbf{x}_u = r^2$, and $E = \mathbf{x}_v \bullet \mathbf{x}_v = r^2 \cos^2 u$, so the patch for the sphere is a v-Clairaut patch. So, using a unit

sphere for simplicity, $v' = \sqrt{\frac{E - k^2}{EG}} = \sqrt{\frac{\cos^2 v - k^2}{\cos^2 v}}$ and $u' = \frac{k}{\cos^2 v}$. So

$$u = \int \frac{k}{\cos v \sqrt{\cos^2 v - k^2}} = \sin^{-1} \left(\frac{k \tan v}{\sqrt{1 - k^2}} \right) + c. \text{ This implies}$$

$$u = \int \frac{k}{\cos v \sqrt{\cos^2 v - k^2}} = \sin^{-1} \left(\frac{k \tan v}{\sqrt{1 - k^2}} \right) + c, \text{ and so } \frac{\sqrt{1 - k^2}}{k} \sin(u - c) = \tan v \Rightarrow$$

$$\frac{\sqrt{1 - k^2}}{k} \sin(u - c) \cos v = \sin v \Rightarrow \frac{\sqrt{1 - k^2}}{k} (\sin u \cos c - \cos u \sin c) \cos v = \sin v \Rightarrow$$

$$(13) \quad \frac{\cos c \sqrt{1 - k^2}}{k} \sin u \cos v - \frac{\sin c \sqrt{1 - k^2}}{k} \cos u \cos v = \sin v$$

Note that the components of the patch for the sphere are here in this last equation. Setting $x = \cos u \cos v$, $y = \sin u \cos v$, $z = \sin v$, we see that equation (13) implies that the points all lie on the same plane. So a geodesic on a sphere is the intersection of the sphere and a plane.

Moreover, note that if a point (x, y, z) on the sphere satisfies (13) so too does $(-x, -y, -z)$. So if the plane in question includes (x, y, z) , it includes $(-x, -y, -z)$. Hence it includes the line between them, and therefore the plane passes through the origin. So a geodesic on the sphere is the intersection of the sphere and a plane passing through the origin—a great circle! Choosing two points that

are not antipodal—that is, not on opposite sides of the sphere—allows one to solve for a unique k and c .

Using the Mathematica procedure above verifies the fact that the geodesics on a sphere are great circles. Figure 4 shows a few geodesics on a sphere.

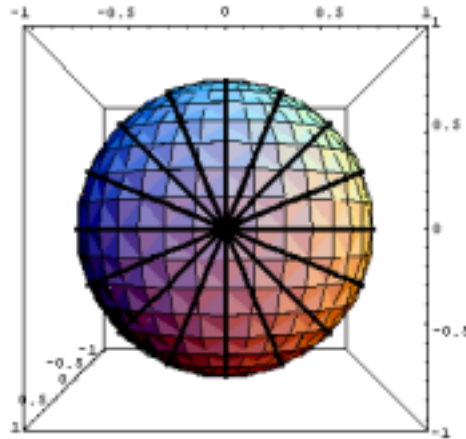


Figure 4

Returning to the example of a plane flying from New York City to Madrid, we can now plot its path and find the distance that it will travel. Madrid is located at $40^\circ 26'' = 40.43^\circ$ North latitude and $3^\circ 42'' = 3.70^\circ$ West longitude; New York City is at $40^\circ 47'' = 40.78^\circ$ North latitude and $73^\circ 58'' = 73.97^\circ$ West longitude, according to [5]. The patch given for the sphere is called the “geographical patch” because the point $\mathbf{x}(u,v)$, u and v give the longitude and latitude when converted into degrees. u gives longitude; a positive value is East longitude, and a negative one is West longitude. v gives latitude; a negative value is South latitude. This means that New York = $\mathbf{x}(-73.97 \pi/180, 40.78 \pi/180)$, and Madrid = $\mathbf{x}(-3.70 \pi/180, 40.43 \pi/180)$. Since the geodesic segment between them is an

arc of a great circle, the length of this arc is $r\theta$, where r is the radius of the earth and θ is the angle between the vector from the center of the earth to New York and the vector from the center of the earth to Madrid. Using Mathematica we can find θ by

```
theta = ArcCos[(sphere[-73.97 Pi/180, 40.78
Pi/180].sphere[-3.70 Pi/180, 40.43 Pi/180])/(r^2)]
```

This gives $\theta = 0.904385$. The radius of the earth can be estimated as 6378 km (3963 miles). So the distance from New York to Madrid along the shortest path is 5768.17 km (3589.50 miles). The radius of the circle around the fortieth

parallel is $\sqrt{x_{NYC}^2 + y_{NYC}^2 - z_{NYC}^2} = 2443.48$ km; the center of this circle is (0, 0,

4165.83). So the angle between the vector from the center of this circle to New York and the vector from the center of this circle to Madrid (assuming for simplicity both are on the 40th parallel) can be found in Mathematica as

```
angle=ArcCos[(earth[-73.97 Pi/180, 40 Pi/180]-
{0,0,4165.83}).(earth[-3.70 Pi/180, 40 Pi/180]-
{0,0,4165.83}))/ (rho^2)], where rho is the radius of the 40th parallel.
```

angle = 1.22644, so the distance from New York to Madrid along the 40th parallel is 5992.2 km. So the great circle method is significantly shorter.

The Cylinder

The cylinder can be parameterized by $\mathbf{x}(u, v) = (r \cos(u), r \sin(u), v)$.

$\mathbf{x}_u \bullet \mathbf{x}_u = r^2$, $\mathbf{x}_u \bullet \mathbf{x}_v = 0$, and $\mathbf{x}_v \bullet \mathbf{x}_v = 1$, so this is clearly a Clairaut patch. Solving the Clairaut integral for v in terms of u gives $v = ku$. Now, unit-speed implies $1 =$

$r^2 u'^2 + v'^2 = (r^2 + k^2) u'^2$. So $u(t) = \frac{t}{\sqrt{r^2 + k^2}} + c \Rightarrow v(t) = \frac{kt}{\sqrt{r^2 + k^2}} + c$, for some

constants k and c . So a geodesic α on a cylinder has the form

$$\alpha(t) = \left(r \cos\left(\frac{t}{\sqrt{r^2 + k^2}} + c\right), r \sin\left(\frac{t}{\sqrt{r^2 + k^2}} + c\right), \frac{kt}{\sqrt{r^2 + k^2}} + c \right).$$
 This is the equation

of a helix. Two interesting degenerate cases are where the initial velocity is horizontal ($k=0$)—in which case the geodesic is a circle going around the cylinder—and where the initial velocity is vertical ($k=\infty$)—in which case the geodesic is a vertical line up the side of the cylinder. These cases and one non-degenerate helical case are shown, using the Mathematica procedure, in Figure

5.

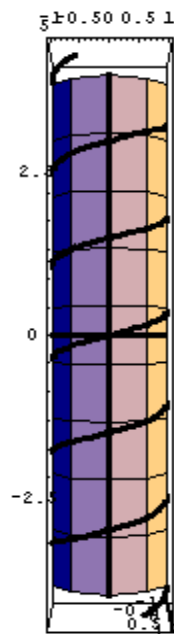


Figure 5

The Torus of Revolution

The torus of revolution is the surface formed by revolving a circle in the y - z plane, which does not touch the z -axis, about the z -axis. The surface looks like a doughnut. A patch for the torus of revolution is $\mathbf{x}(u, v) = ((a + r \cos v) \cos u, (a + r \cos v) \sin u, r \sin v)$. r here represents the radius of the circle being revolved; a represents the distance from the circle to the z -axis. The case $r=1$, $a=2$ is shown in Figure 6.

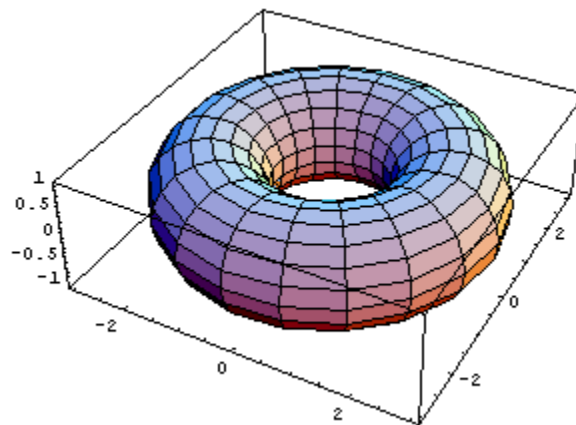


Figure 6

The patch here is a v -Clairaut patch. In fact, all surfaces of revolution admit a Clairaut patch. If the curve being revolved is $(g(v), h(v))$, the patch is $\mathbf{x}(u, v) = (h(v) \cos u, h(v) \sin u, g(v))$. In the case of the torus, $E = \mathbf{x}_u \cdot \mathbf{x}_u = (a + r \cos v)^2$, and $G = \mathbf{x}_v \cdot \mathbf{x}_v = r^2$. Yet the Clairaut integral cannot be solved explicitly. Here is where the implicit approach of the Mathematica procedure is useful. A few examples may be instructive. If the initial direction is $\pi/2$, the geodesic goes around the “doughnut hole” as in Figure 7

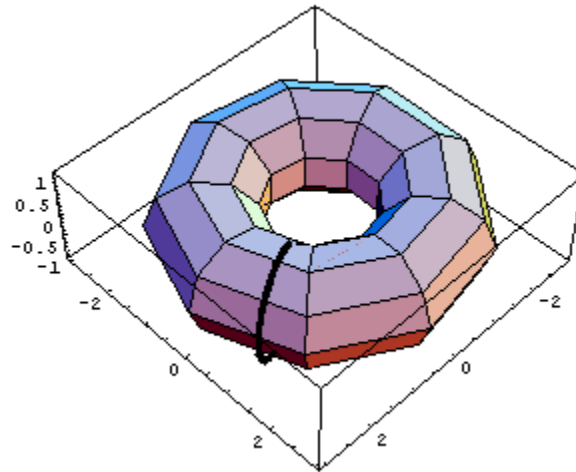


Figure 7

If the initial point is above the largest circle around the outside of the torus, and the initial direction is 0 (i.e. horizontal), then the geodesic turns downward to the bottom of the torus, then turns upward to the top, then downward, and so on.

Figure 8 is an example.

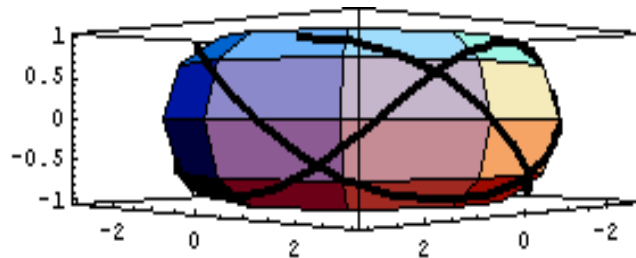


Figure 8

The “Two-Banana” Surface

As we saw with the torus of revolution, there are cases where with a Clairaut patch no explicit equations for the geodesics can be found. If the patch we are working with for a surface is not Clairaut, it becomes even less likely that a search for explicit geodesics will be successful. For example, consider the surface parameterized by $\mathbf{x}[u,v]=((2+\cos(u))\cos(v), \cos(v)\sin(u), \sin(v))$. We call

this surface the “Two-Banana” surface because it looks like two bananas meeting at their ends. This surface is shown in Figure 9.

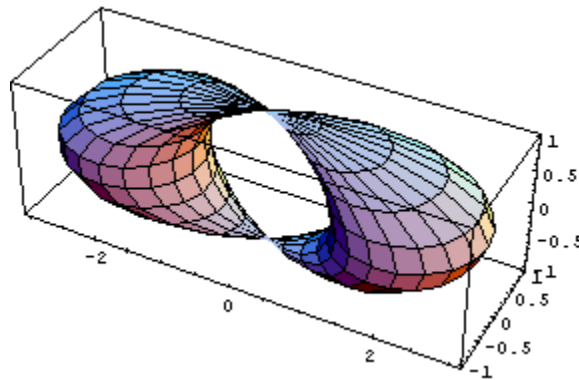


Figure 9

For this patch, $G = \mathbf{x}_v \cdot \mathbf{x}_v = \cos^2 v + (5 + 4\cos u)\sin^2 v$, so the patch is not Clairaut or v-Clairaut. We can use the Mathematica procedure `geo` to get an idea of what the geodesic equations look like in this case. The results, in Figure 10, are not pretty.

$$\begin{aligned} & \text{geo}[\text{twobanana}][p, q, u, v] \\ & \{p, q, -\sec[v]^2 (2 q^2 \cos[v]^2 \sin[u] - 2 p q \cos[v] \sin[v]), \\ & -(4 q^2 (1 + \cos[u]) \cos[v] \sin[v] + \\ & p^2 (1 + 2 \cos[u]) \cos[v] \sin[v] - 4 p q \sin[u] \sin[v]^2) / \\ & (\cos[v]^2 + (5 + 4 \cos[u]) \sin[v]^2)\} \end{aligned}$$

Figure 10

Mathematica, however, can still graph the geodesics implicitly. If the geodesic begins horizontally along the widest circle of a banana, it will go around that circle as in Figure 11. If it begins vertically along the outermost part of a banana, it will loop around both bananas as in Figure 12. In general, however, the geodesics stay on one banana and have complicated shapes. For example, Figure 13 shows the geodesic starting at $u=0, v=0$ and with initial direction $\pi/4$.

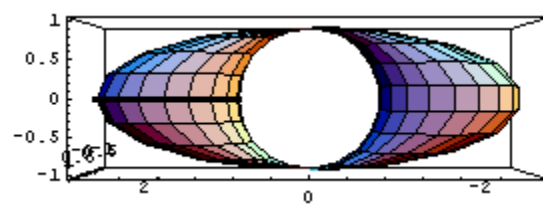


Figure 11

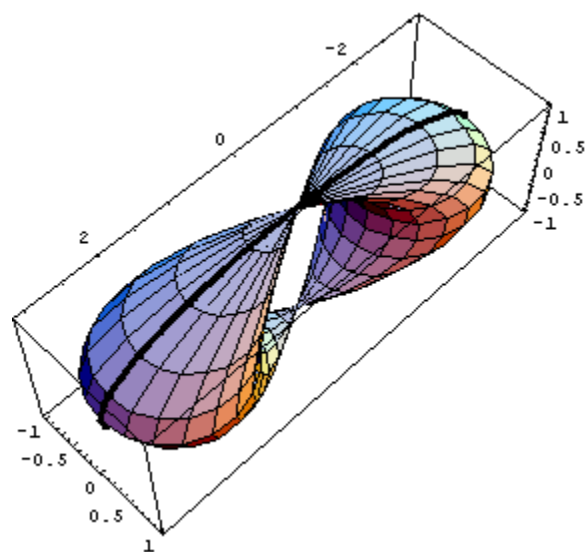


Figure 12

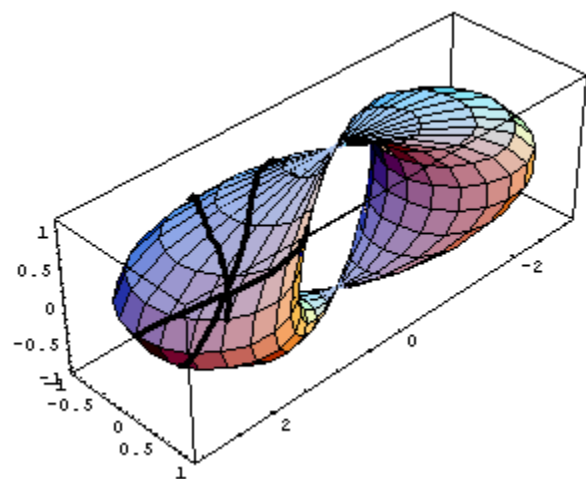


Figure 13

The Hyperbolic Plane

Thus far all the surfaces examined have been surfaces in \mathbf{R}^3 . For surfaces not in \mathbf{R}^3 the procedure requires modification. For example, we consider the hyperbolic plane. This is modeled as the upper half-plane \mathbf{U}^2 , the set of all (u, v) such that $v > 0$. The properties of the plane, however, are different from those of \mathbf{R}^2 . Here, the inner product depends on where it is applied. If two vectors w_1 and w_2 are applied at (u, v) , $w_1 \circ w_2 = (w_1 \bullet w_2)/v^2$, where \circ is the inner product in \mathbf{U}^2 , and \bullet is the standard dot-product in \mathbf{R}^2 . The upper half-plane with this inner product is called the Poincaré plane. So $\mathbf{x}(u, v) = (u, v)$, and $\mathbf{x}_u(u, v) = (1, 0)$, but $\mathbf{x}_u(u, v) \bullet \mathbf{x}_u(u, v) = 1/v^2$, and $\mathbf{x}_v(u, v) = (0, 1/v)$, but $\mathbf{x}_v(u, v) \bullet \mathbf{x}_v(u, v) = 1/v^2$. So this is a v -Clairaut patch, and so

$$u = \int \frac{k}{\sqrt{\frac{1}{v^2} - k^2}} dv \Rightarrow ku = \int \frac{k^2}{\sqrt{\frac{1}{v^2} - k^2}} dv \Rightarrow ku = -v \sqrt{\frac{1}{v^2} - k^2} + c \Rightarrow k^2 \left(u - \frac{c}{k} \right)^2 = 1 - k^2 v^2 \Rightarrow \frac{1}{k^2} = v^2 + \left(u - \frac{c}{k} \right)^2.$$

This gives the equation of a circle with radius $1/k$ centered at $(c/k, 0)$. Since we are dealing with the Poincaré plane, the geodesics are the upper halves of these circles. The initial conditions determine k and c . In the case $k=0$, the half-circle degenerates to a vertical half-line.

Given two points (u_1, v_1) and (u_2, v_2) in the Poincaré plane, if $v_1 = v_2$, then they are connected by a vertical half-line but by no circle. If $v_1 \neq v_2$, then the only geodesic passing through (u_1, v_1) and (u_2, v_2) is the half-circle centered at

$$\left(\frac{u_2^2 - u_1^2 + v_2^2 - v_1^2}{2(u_2 - u_1)}, 0 \right) \text{ with radius } \sqrt{v_1^2 - \frac{(u_1^2 - 2u_1u_2 + u_2^2 - v_1^2 + v_2^2)^2}{4(u_2 - u_1)^2}}.$$

This geodesic can be graphed in Mathematica using the procedure shown in Figure 14.

```

radius[uo_, vo_, ut_, vt_] :=
  Sqrt[{uo - (ut ^ 2 - uo ^ 2 + vt ^ 2 - vo ^ 2) / (2 (ut - uo))} ^ 2 +
    vo ^ 2]

gcenter[uo_, vo_, ut_, vt_] :=
  {(uo - (ut ^ 2 - uo ^ 2 + vt ^ 2 - vo ^ 2) / (2 (ut - uo))), 0}

geocirc[uo_, vo_, ut_, vt_, t_] :=
  {radius[uo, vo, ut, vt] Cos[t] -
    First[gcenter[uo, vo, ut, vt]],
    radius[uo, vo, ut, vt] Sin[t]}

ParametricPlot[geocirc[u1, v1, u2, v2, t], {t, 0, Pi}]

```

Figure 14

For example, Figure 15 shows the geodesic in the Poincaré half-plane passing through (1,2) and (3,4).

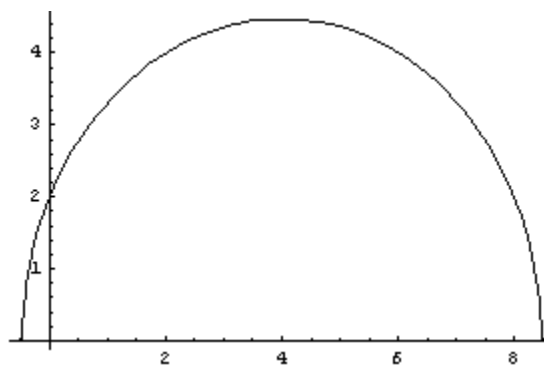


Figure 15

Conclusion

The concept of “line” is a very intuitive and fundamental concept in our everyday lives. While the concept of “geodesic” is less intuitive on many surfaces, it is no less fundamental. For this reason, it is critical to know what the geodesics look like on a surface in order to gain an intuitive understanding of the geometry of a surface. Mathematica is an invaluable aid in this process, as Mathematica bypasses the need to perform laborious calculations in solving, implicitly or explicitly, the geodesic equations.

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] Oprea, J. *Differential Geometry and Its Applications*. Upper Saddle River, NJ: Prentis Hall, 1997. 155, 160-162.