Characterizing Geodesics on Surfaces of Revolution

by

Olivia Grace Rigatti

A thesis submitted in partial fulfillment of the requirements
for graduation with Honors in Mathematics.

Whitman College
2023

Certificate of Approval

This is to certify that the accompanying thesis by Olivia Grace Rigatti has been accepted in partial fulfillment of the requirements for graduation with Honors in Mathematics.

Albert Schueller

Whitman College
May 9, 2023

Contents

List of Figures

Abstract

When we want to get between two places as fast as possible, we walk between them on the shortest path. This shortest path is called a geodesic. But what do we do if we are not on a flat surface? What if we are on a sphere or a donut, we can not walk in a straight line because these are curved surfaces. This paper will characterize the geodesics on different surfaces of revolution, which are surfaces in 3D that are constructed by rotating a continuous and differentiable curve around an axis that lies in the same plane. This includes the sphere, the torus, and the cylinder, among others.

1 Introduction

In everyday life, we are often trying to optimize our time. We want to stand in the quickest line in a store, get in and out when buying groceries, and take the shortest route to and from work. When we are walking places to get somewhere as fast as possible, we walk in a straight line. We do this because we know that the shortest distance between two points on a flat surface is a line. This shortest path that we choose to take is a geodesic. On any surface, a geodesic is the shortest path between two points. People seem to intuitively know to walk in a straight line to minimize distance. But what if a person is not on a flat surface? What if they were standing on a cylinder or a donut shaped surface? What path should this individual take if they want to walk the shortest distance possible? This paper will examine these questions and provide some answers.

The first section of this paper will introduce a formal definition for geodesics and relate them to surfaces of revolution. Surfaces of revolution are surfaces in 3D that are constructed by rotating a continuous and differentiable curve around an axis that lies in the same plane. Once we have this knowledge, we introduce tools that are helpful when examining surfaces and curves, including the Frenet Frame, the tangent plane $T_p(M)$, the unit normal $U(p)$, and the basis vectors for surfaces of revolution. Once these ideas have been established, we move into constructing the geodesic equations. This is done by looking at the behavior of a curve on surfaces of revolution and seeing what requirements exist for the curve to be a geodesic. We then derive the Euler-Lagrange equation.

The second section of this paper applies the information from the first section to specific examples. We begin by looking at geodesics on the cylinder. Here the geodesic equations will be solved directly. We then move into geodesics on the sphere. We characterize the geodesics in two different ways, one of which directly uses the geodesic equations. After the sphere, the geodesics of a torus are characterized, using the Clairaut relation. Lastly, we look at geodesics on the cone by introducing a different parameterization for the surface of revolution. Here we bring in knowledge relating to the Frenet Frame and the Euler-Lagrange equation. The final section of the paper will introduce the metric tensor. We show that the metric tensor can be found in several different ways and then use it to find the arc length of two different curves. Some applications of geodesics are then briefly touched on.

2 Background Information

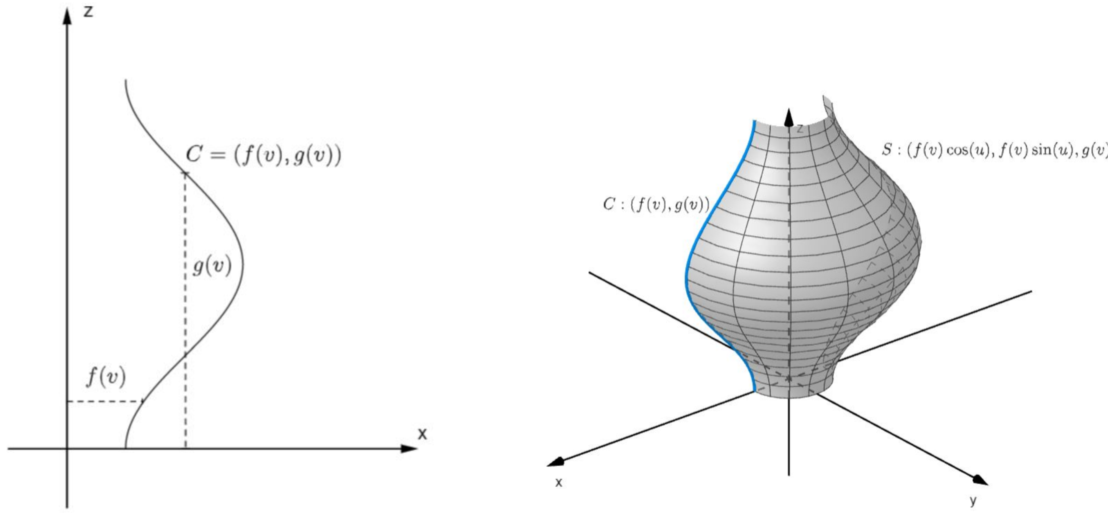
Before we begin finding geodesics on surfaces of revolution we first must know what geodesics and surfaces of revolution are, along with some other information about surfaces. This section aims to introduce geodesics and surfaces of revolution, along with the tools that are used to find geodesics, like the Frenet Frame, the tangent plane, the geodesic equations, and the Euler-Lagrange equation.

2.1 Geodesics and Surfaces of Revolution

When working with a surface, whether it be a plane, sphere, or hyperbolic paraboloid, we often want to know the shortest path between two points. This is where geodesics are useful.

Definition 2.1. (Geodesic) A geodesic is the shortest curve between two points on any given surface [2].

Intuitively, a geodesic can be thought of as a curve that only curves as much as necessary to stay on a surface. If we are working on a plane, the shortest path between two points will be a straight line. However, if we change surfaces a straight line might not be a geodesic. For instance, on the sphere a straight line is not the shortest path between two points, as the surface is curved. So, geodesics change with



(a) Profile Curve (b) Surface of Revolution

Figure 1: Profile Curve and Corresponding Surface of Revolution

the surface and can become more difficult to find for different surfaces. There are many different types of surfaces, but this paper will focus specifically on geodesics on surfaces of revolution.

Definition 2.2. (Surface of Revolution) A surface of revolution, S, is a surface obtained by rotating a (C^2 smooth) curve, C, around an axis that lies in the same plane [1]. The curve C is called the profile curve of S.

In this paper, the profile curve will be rotated around the z-axis. Therefore, we parameterize the curve C by $(x, z) = (f(v), g(v))$ (Figure 1a). Since we obtain the surface of revolution by rotating C around the z-axis, the parameterization for the surfaces we will examine is given by

$$\mathbf{x}(u, v) = (f(v)\cos(u), f(v)\sin(u), g(v)), \quad 0 \leq u < 2\pi. \tag{1}$$

We multiply $f(v)$ by $\cos(u)$ and $\sin(u)$ for the x and y coordinate, respectively, because that is the parameterization for the unit circle. Multiplying by these terms rotates the profile curve around the z-axis.

2.2 The Frenet Formulas

Now that we have a definition for surfaces of revolution, we can start to develop tools to describe and analyze curves on these surfaces. One such tool is the Frenet Formulas. In order to understand the Frenet Formulas, it helps to think of a robot 'hand' that has three 'fingers'. Imagine that the robot must move its' 'hand' through space to grab something. The movement of the 'hand' follows the path of a curve through space. So, as the 'hand' is moving along the curve the three 'fingers' are also moving. Now consider an actual curve in space. The robot's three 'fingers' become three vectors that describe the geometry of the curve at each point along the path. These vectors are orthogonal, unit, and will give a 'simple' way to describe the space around the curve. The first 'finger' of the robot's 'hand' can be thought of as the unit tangent vector to the curve.

Definition 2.3. (Unit Tangent Vector) Let $\alpha : I \to \mathbb{R}^3$ be a unit speed curve (that is, $|\alpha'(t)| = 1$ for all t). Then the unit tangent vector to α is given by $T = \alpha'$ [2].

With the definition of the unit tangent vector of a curve, we can build the other two vectors that will help describe the geometry of the curve. As α is a unit speed curve we know that $|\alpha'| = |T| = 1$ for every point on α. Therefore, T' only measures the rate of change of the direction of T. So, T' is a good candidate to measure any change in geometry of α, as the direction of T changes with α. Another important aspect of T' is that it is orthogonal to T [2].

Proposition 2.1. *T is orthogonal to T'.*

Proof. We know $|T| = \sqrt{T \cdot T} = 1$. Squaring both side of the equation we get that $T \cdot T = 1$. Since the derivative of a constant is zero we can see that

$$0 = (T \cdot T)' = T' \cdot T + T \cdot T' = 2(T' \cdot T).$$

Therefore, $T' \cdot T = 0$ and T' is perpendicular to T. $\square$

We know that $|T| = 1$ because α is a unit speed curve. However, there is no reason for T' to be unit speed. As a matter of convenience we instead consider a vector that is unit speed. Thus, let the curvature of α be

$$k(s) = |T'(s)|.$$

Now that we have the magnitude of T' we can construct our second unit vector, the principal normal vector, also called the normal vector.

Definition 2.4. (Normal Vector) Given $\alpha : I \to \mathbb{R}^3$ and $T = \alpha'$ the principal normal vector of α is given by $N = \frac{1}{k(s)}T'$ [2].

Since T and N are orthogonal to each other, and we are working in 3D, we want our third vector to be orthogonal to T and N. We know that the cross-product of two vectors is orthogonal to both the original vectors. Therefore, to construct a vector perpendicular to both T and N we take their cross product as the third vector, which we call the binormal vector.

Definition 2.5. (Binormal Vector) Given $\alpha : I \to \mathbb{R}^3$, $T = \alpha'$ and $N = \frac{1}{k(s)}T'$, we define the binormal vector along α to be $B = T \times N$ [2].

Note that $|B| = |T||N| \sin(\frac{\pi}{2}) = 1$ so B is already a unit vector.

These three vectors are called the Frenet Frame. Now that we know them, we can develop the Frenet Formulas. The measurement of how T, N, and B vary as we move along α will tell us how α itself curves through space. Thus, the Frenet Formulas involve the derivatives of T, N, and B. From Definition 2.4 we know that $T' = kN$. So, $k(s)$ describes the variation in the direction of T. Now we just need to find N' and B'. As T, N, and B are orthonormal vectors in $\mathbb{R}^3$ any other vector in $\mathbb{R}^3$ can be written as a linear combination of them. Consider $B' = aT + bN + cB$. If we can calculate a, b, and c then we can find B'. This involves using what we already know about T, N, and B. First,

$$\begin{aligned} T \cdot B' &= aT \cdot T + bT \cdot N + cT \cdot B \\ &= a \cdot 1 + b \cdot 0 + c \cdot 0 \\ &= a. \end{aligned}$$

The simplification in step 2 comes from orthonormal vectors having a dot product of zero. Using the same method above we find $N \cdot B' = b$ and $B \cdot B' = c$. Thus,

$$B' = (T \cdot B')T + (N \cdot B')N + (B \cdot B')B \tag{2}$$

Now we can identify $T \cdot B'$. We know that $T \cdot B = 0$, so $0 = (T \cdot B)' = T' \cdot B + T \cdot B'$ by the product rule. Thus, using $N \cdot B = 0$ we have

$$\begin{aligned} T \cdot B' &= -T' \cdot B \\ &= -kN \cdot B \\ &= 0. \end{aligned}$$

We can also find $B \cdot B'$. We know $B \cdot B = 1$, so

$$\begin{aligned} 0 &= (B \cdot B)' \\ 0 &= B' \cdot B + B \cdot B' \\ 0 &= 2B \cdot B' \\ 0 &= B \cdot B'. \end{aligned}$$

We cannot directly identify $N \cdot B'$, so instead, we give it a name. Define $\tau = -N \cdot B'$ to be the *torsion* of the curve α [2]. The torsion measures the turnaround of the binormal vector. So, if it is very large we can think of the binormal vector as rotating faster around the axis of the tangent vector. Substituting all these values into equation (2) we get

$$B' = -\tau N.$$

We can find N' using the same process as we did above but with the equation

$$N' = (T \cdot N')T + (N \cdot N')N + (B \cdot N')B.$$

The same types of calculations give us

$$\begin{aligned} T \cdot N' &= -k \\ N \cdot N' &= 0 \\ B \cdot N' &= \tau. \end{aligned}$$

Thus, we have

$$N' = -kT + \tau B.$$

Given these derivatives, we now have the Frenet Formulas for a unit speed curve.

Theorem 2.1 (The Frenet Formulas)**.** *For a unit speed curve with* $k > 0$*, the derivatives of the Frenet frame are as follows.*

$$\begin{aligned} T' &= kN \\ N' &= -kT + \tau B \\ B' &= -\tau N. \end{aligned}$$

Given the three vectors, T, N, and B, we can also define three different planes. The plane spanned by $\{T, N\}$ is the *osculating plane* [3]. The *normal plane* is the plane spanned by $\{N, B\}$. Lastly, the plane spanned by $\{T, B\}$ is called the *rectifying plane.*

As a curve α has a position vector associated with it, by looking at which plane the position vector lies in, we can learn about the geometry of the curve. When the position vector of a curve always lies in the normal plane, the curve lies on a sphere. If the position vector only ever lies in the osculating plane, the curve lies in a plane [3]. If a curve's position vector is always in the rectifying plane, it is called a rectifying curve. What this means for the curve will be discussed later in the paper.

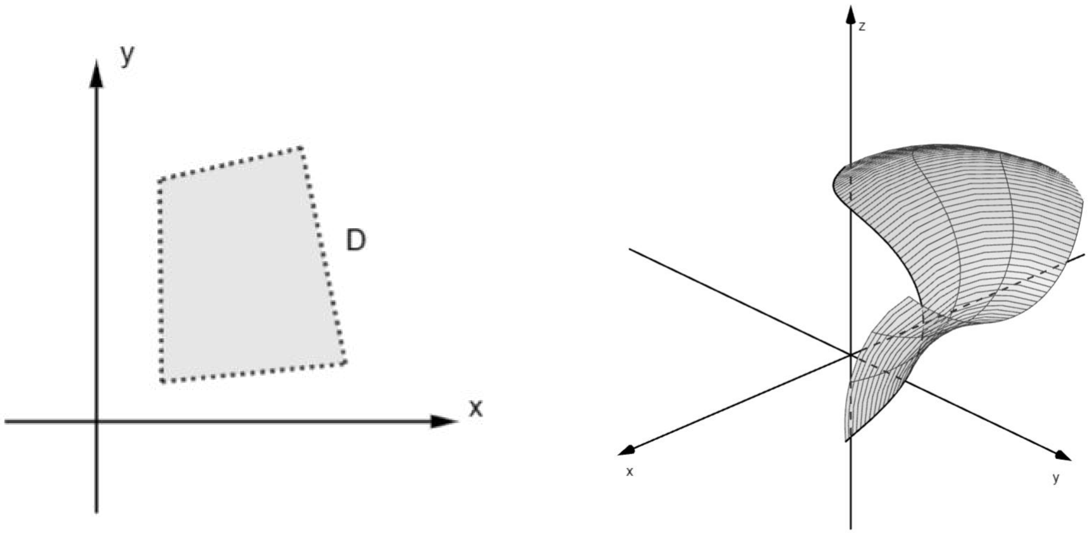
2.3 The Tangent Plane

To begin talking about the curvature of surfaces we need to have more formal ideas and definitions of what we are working with, including defining surfaces and curves. To fully define a surface we first must understand mappings and coordinate patches. Consider an open set $D \subseteq \mathbb{R}^2$. A *mapping* is a function that takes a point from D and sends it to a point in $\mathbb{R}^3$ [2]. We express this mapping as $\mathbf{x} : D \to \mathbb{R}^3$. A mapping is considered regular if $\mathbf{x}_u \times \mathbf{x}_v \neq 0$, where $\mathbf{x}_u$ and $\mathbf{x}_v$ are the derivatives of the parameterization of the mapping. A *coordinate patch* is very similar to a mapping. In fact, it is a one-to-one, regular mapping $\mathbf{x} : D \to \mathbb{R}^3$ of an open set $D \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ [2]. In other words, a coordinate patch is an open set in two dimensions, which has been mapped to an open set in three dimensions. In Figure 2 we can see an open set D (Figure 2a) that has been mapped to a coordinate patch (Figure 2b). This means that the surfaces that we are examining in $\mathbb{R}^3$ are locally planar and Euclidean.

Now that we have a formal idea of what a mapping and a coordinate patch are we can define a surface in $\mathbb{R}^3$.

Definition 2.6 (Surface)**.** A surface in $\mathbb{R}^3$ is some subset $M \subseteq \mathbb{R}^3$ such that each point of M has a neighborhood (in M) contained in the image of some coordinate patch $\mathbf{x} : D \to \mathbb{R}^3$ [2].

Informally, we may say that a surface, M, is some subset of $\mathbb{R}^3$ where every point of



(a) An Open Set in $\mathbb{R}^2$ (b) A Coordinate Patch in $\mathbb{R}^3$

Figure 2: A mapping from an open set to a coordinate patch

the surface is surrounded by other points that are part of the same coordinate patch. We include the requirement that a point has a neighborhood so we can take limits from all sides. If the neighborhood around the point did not exist, there might be instances when taking the limit from one direction is impossible.

Lastly, we need to know how to define curves on a surface. A curve, α, on a surface, M, is a mapping from some interval of real numbers $I = [a, b]$ to M, $\alpha : I \to M$ [2]. For our purposes, we need the curve on the surface to be smooth. If α is smooth, which implies it is differentiable, it can be parameterized such that $\alpha(t) = \mathbf{x}(u(t), v(t))$ where $\mathbf{x}$ is a mapping and $u(t), v(t)$ are unique functions in $\mathbb{R}^2$ such that $x = u(t)$ and $y = v(t)$ [2]. The uniqueness of these functions is important. Since they are the only functions for α we can analyze $u(t), v(t)$ in $\mathbb{R}^2$ rather than trying to analyze α in $\mathbb{R}^3$. If α were not smooth the functions would not be unique and we could not look at $u(t)$ and $v(t)$ in $\mathbb{R}^2$.

Now that we understand how a surface is created, we want to be able to talk about changes in its shape along different curves. In $\mathbb{R}^2$ when we want to see how a function of one variable changes we look at the tangent line of the function at some point. We do a similar thing with surfaces, but since we are looking at a 2D surface and not a 1D curve, rather than looking at the tangent line of a point, we look at vectors related to the tangent plane of a point. The tangent plane at a point p on a surface M is represented by $T_p(M)$. Although the tangent plane itself is important, the vectors associated with it tell us more about the change in the surface's geometry.

The two vectors that will be most important for our purpose are $\mathbf{x}_u$ and $\mathbf{x}_v$. These vectors are constructed by taking the derivative of equation (1) with respect to u and v. Thus, they have the form

$$\mathbf{x}_u = (-f(v)\sin(u), f(v)\cos(u), 0) \tag{3}$$

$$\mathbf{x}_v = (f'(v)\cos(u), f'(v)\sin(u), g'(v)). \tag{4}$$

By construction $\mathbf{x}_u$ and $\mathbf{x}_v$ span the tangent plane of the parameterized surface [2]. One should also note that $\mathbf{x}_u \cdot \mathbf{x}_v = 0$, so the vectors are orthogonal to each other and the tangent plane has an orthonormal basis [2]. These vectors become very important when trying to find general equations for geodesics on surfaces of revolution. To do this we must first define several more terms with regard to $\mathbf{x}_u$ and $\mathbf{x}_v$.

Let E, F, and G be defined as follows,

$$E = \mathbf{x}_u \cdot \mathbf{x}_u$$

$$F = \mathbf{x}_u \cdot \mathbf{x}_v$$

$$G = \mathbf{x}_v \cdot \mathbf{x}_v.$$

Note that as $\mathbf{x}_u$ and $\mathbf{x}_v$ are orthonormal F will always be zero when using equation (1) as the parameterization for a surface.

Another important vector that relates to the tangent plane is the unit normal vector at p, denoted $U(p)$. This vector is orthogonal to $T_p(M)$ and plays a role in detecting the curvature of a surface. The unit normal vector is purposefully constructed to be orthonormal to $T_p(M)$. We know that the tangent plane is spanned by $\mathbf{x}_u(p)$ and $\mathbf{x}_v(p)$. So to find a vector orthogonal to the tangent plane it must be orthogonal to $\mathbf{x}_u$ and $\mathbf{x}_v$. As the cross product of two vectors is orthogonal to both original vectors we consider $\mathbf{x}_u \times \mathbf{x}_v$. We also want this vector to be unit speed so the magnitude does not impact any calculations regarding curvature. Thus, we define the unit normal vector as

$$U = \frac{\mathbf{x}_u \times \mathbf{x}_v}{|\mathbf{x}_u \times \mathbf{x}_v|}. \tag{5}$$

2.4 Geodesic Equations

Now that we have language to describe surfaces and curves, we want to find the geodesics of a surface of revolution. To find these geodesics let $\alpha(t)$ be a geodesic in some coordinate patch $\mathbf{x}$. Then we know, by definition, that $\alpha(t) = \mathbf{x}(u(t), v(t))$ and $\alpha' = \mathbf{x}_u u' + \mathbf{x}_v v'$ by the chain rule. Taking the derivative of α' we get

$$\alpha'' = \mathbf{x}_{uu} u'^2 + \mathbf{x}_{uv} v' u' + \mathbf{x}_u u'' + \mathbf{x}_{vu} u' v' + \mathbf{x}_{vv} v'^2 + \mathbf{x}_v v''. \tag{6}$$

It is possible to break α'' up into two different components, a component that is normal to the surface of revolution, S, and one that is tangent to S [1]

$$\alpha'' = \alpha''^{\perp} + \alpha''^{\tan}.$$

The normal component, $\alpha''^{\perp}$, represents the curvature that is imposed on α by the surface S, this is called the normal curvature. On the other hand, the tangent component, $\alpha''^{\tan}$, represents any 'extra' curvature of α, this is called the geodesic curvature.

The tangent component of α'' can be expressed by

$$\alpha''^{\tan} = \alpha'' \cdot \mathbf{x}_u + \alpha'' \cdot \mathbf{x}_v. \tag{7}$$

The derivation of equation (7) is outside the scope of this paper, but is talked about in more detail in [1] [2]. As we are looking for geodesics we want the extra curvature of the curve to be zero, so the tangent component must be equal to zero. Since the tangent component is zero, we can think of the curve as only curving as much as necessary to stay on the surface, which is how we intuitively defined a geodesic. Therefore,

$$\alpha'' \cdot \mathbf{x}_u = 0, \;\; \alpha'' \cdot \mathbf{x}_v = 0. \tag{8}$$

The two dot products cannot be additive inverses as they are linearly independent, thus each term in equation (8) must be zero. We can compute the dot products in equation (8) using equation (6). To do this we need the following derivatives:

$$\begin{aligned}
\mathbf{x}_u &= (-f(v)\sin(u), f(v)\cos(u), 0) \\
\mathbf{x}_v &= (f'(v)\cos(u), f'(v)\sin(u), g'(v)) \\
\mathbf{x}_{uu} &= (-f(v)\cos(u), -f(v)\sin(u), 0) \\
\mathbf{x}_{vv} &= (f''(v)\cos(u), f''(v)\sin(u), g''(v)) \\
\mathbf{x}_{uv} &= \mathbf{x}_{vu} = (-f'(v)\sin(u), f'(v)\cos(u), 0).
\end{aligned}$$

We will also need the necessary inner products for the computation:

$$\mathbf{x}_u \cdot \mathbf{x}_u = f^2 \qquad \mathbf{x}_v \cdot \mathbf{x}_v = 1 \quad \mathbf{x}_u \cdot \mathbf{x}_v = 0 \tag{9}$$

$$\mathbf{x}_{uu} \cdot \mathbf{x}_u = 0 \qquad \mathbf{x}_{vv} \cdot \mathbf{x}_u = 0 \quad \mathbf{x}_{uv} \cdot \mathbf{x}_u = ff' \tag{10}$$

$$\mathbf{x}_{uu} \cdot \mathbf{x}_v = -ff' \quad \mathbf{x}_{vv} \cdot \mathbf{x}_v = 0 \quad \mathbf{x}_{uv} \cdot \mathbf{x}_v = 0. \tag{11}$$

Computing equation (8) using equations (6) and (9) we get two equations that all geodesics on surfaces of revolution must fulfill:

$$2ff'u'v' + f^2u'' = 0 \tag{12}$$

and

$$v'' - f'u'^2 = 0. \tag{13}$$

These are called the *geodesic equations*, and will play an important role in finding geodesics on the different surfaces we will examine. Along with these two equations comes the assumption that our curve α is unit speed. Therefore, it must be true that

$$1 = f^2u'^2 + v'^2. \tag{14}$$

With the geodesic equations we can begin to look for geodesics on different surfaces of revolution. We can also extrapolate some general geodesics on all surfaces of revolution. This will be done by looking at two different types of curves, meridians and parallels, A *meridian* is a curve $\alpha(t) = \mathbf{x}(u(t), v(t))$ where u is held constant, i.e. $u(t) = k$ [1]. On a surface of revolution they look like the profile curve. A *parallel* is a curve $\alpha(t) = \mathbf{x}(u(t), v(t))$ where v is held constant, i.e. $v(t) = l$ [1]. These curves are what would occur if one cut the surface parallel to the xy plane. In Figure 1b the meridians are the vertical lines on the surface, and the parallels are the horizontal lines on the surface. If we were looking at a globe the meridians would be the lines of longitude and the parallels the lines of latitude. Examining these curves we can get some general rules for geodesics that apply to all surfaces of revolution.

Since meridians occur when u is constant that means $u' = 0$ and $u'' = 0$. Thus, equation (12) is fulfilled. We see that equation (13) becomes

$$v'' = 0.$$

However, from the unit speed requirement we know that $v' = 1$ so the second geodesic requirement is also fulfilled. Thus, all meridians on surfaces of revolution are geodesics [1].

We can also examine the equations when looking at parallels. If v is constant then equation (14) implies that $fu' = 1$ so $f \neq 0$ and $u' \neq 0$. Therefore, equations (12) and (13) will be equivalent if and only if v is a constant such that $f'(v) = 0$. This means that parallel curves will be geodesics when v is a critical point of the function f. So, from the two geodesic equations we can find curves that are always geodesics, meridians, and develop a general rule for when parallels are geodesics. However, it is important to note that the other geodesics on surfaces of revolution can be more difficult to find and may require different approaches that vary with the surface.

2.5 The Euler-Lagrange Equation

Consider a curve given by $x = x(t)$, for $t_0 \leq t \leq t_1$, with fixed endpoints so $x(t_0) = x_0$ and $x(t_1) = x_1$. As a note, this curve might be in n-space. If this is the case we can think of the curve as $x(t) = (x^1(t), x^1(t), \ldots, x^n(t))$. If we are working in n-space this curve could be thought of as the path of a particle with respect to time t.

We can derive the Euler-Lagrange equation by trying to find the shortest distance

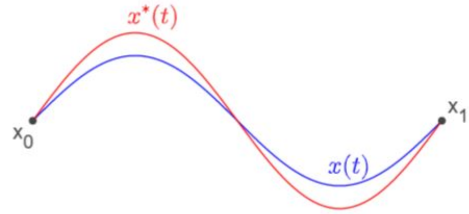


Figure 3: The curves $x(t)$ and $x^*(t)$

between the two fixed endpoints of $x(t)$. When we want to find the extreme of a function, say $f(x)$, we take the derivative $f'(x)$ and find where $f'(x) = 0$. This same approach will be used when trying to find the shortest distance between the two fixed endpoints. As a note, this method is best suited to finding local minima, not global thus it is necessary to have an interval for this approach.

Suppose that $x(t)$ is a curve between the two endpoints x_0, x_1, which minimizes the integral

$$J = \int_{t_0}^{t_1} f(t, x(t), \dot{x}(t))dt, \tag{15}$$

where the integrand is the arc length of the curve [2]. Now let $x^*(t)$ be a variation of $x(t)$ given by $x^*(t) = x(t) + \epsilon\eta(t)$ [2]. Think of ϵ as some small constant and η as some function of t such that $\eta(t_0) = \eta(t_1) = 0$ so the ends of $x(t)$ and $x^*(t)$ are the same. Note that $\dot{x}^*(t) = \dot{x}(t) + \epsilon\dot{\eta}(t)$ where the dot signifies the first derivative with respect to t. If this curve were in 2D space $x(t)$ and $x^*(t)$ could be visualized as in Figure 3. Note this figure illustrates how $x(t)$ and $x^*(t)$ differ, they are not curves of minimal length between the endpoints. As can be seen from Figure 3 the two curves are very similar, but $x^*(t)$ is slightly different from $x(t)$.

Now that we have introduced ϵ, we can think of the integral from equation (15) as a function of ϵ and rewrite it as

$$J(\epsilon) = \int_{t_0}^{t_1} f(t, x^*, \dot{x}^*)dt = \int_{t_0}^{t_1} f(t, x + \epsilon\eta, \dot{x} + \epsilon\dot{\eta})dt. \tag{16}$$

We do this so we can look at the derivative of $J(\epsilon)$. Without ϵ we have no variable to change that will allow us to get closer to the optimal arc length.

Because we want to minimize $J(\epsilon)$, we will take the derivative and find where it equals zero. From Lemma 7.1.4 in [2], we know that

$$\frac{d}{d\epsilon}\int_{t_0}^{t_1} f(t, x^*, \dot{x}^*)\, dt = \int_{t_0}^{t_1} \frac{\partial f(t, x^*, \dot{x}^*)}{\partial \epsilon}\, dt.$$

Using this property, we get

$$\frac{dJ}{d\epsilon} = \int_{t_0}^{t_1} \frac{\partial f}{\partial t}\frac{\partial t}{\partial \epsilon} + \frac{\partial f}{\partial x^*}\frac{\partial x^*}{\partial \epsilon} + \frac{\partial f}{\partial \dot{x}^*}\frac{\partial \dot{x}^*}{\partial \epsilon}\, dt.$$

As t does not rely on ϵ, we know that $\frac{\partial t}{\partial \epsilon} = 0$, so the first term disappears from the equation. Furthermore, from the definition of $x^*(t)$ and $\dot{x}^*(t)$ we can see that $\frac{\partial x^*}{\partial \epsilon} = \eta$ and $\frac{\partial \dot{x}^*}{\partial \epsilon} = \dot{\eta}$. Making these substitutions, the equation becomes

$$\frac{dJ}{d\epsilon} = \int_{t_0}^{t_1} \frac{\partial f}{\partial x^*}\eta + \frac{\partial f}{\partial \dot{x}^*}\dot{\eta}\, dt.$$

Since $x(t)$ is a minimum by hypothesis, the derivative of the arc length will be zero when $\epsilon = 0$ [2]. So evaluating at $\epsilon = 0$ we get

$$0 = \left.\frac{dJ}{d\epsilon}\right|_{\epsilon=0} = \int_{t_0}^{t_1} \frac{\partial f}{\partial x^*}\eta + \frac{\partial f}{\partial \dot{x}^*}\dot{\eta}\, dt. \tag{17}$$

The second term $\frac{\partial f}{\partial \dot{x}}\dot{\eta}$ can be integrated using parts. These calculations will not be shown as they are unnecessary for understanding the equation as a whole, but can be found in [2]. After integrating by parts and substituting the answer back into equation (17), the equation becomes

$$0 = \int_{t_0}^{t_1} \frac{\partial f}{\partial x^*} \eta dt - \int_{t_0}^{t_1} \eta \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) dt$$
$$0 = \int_{t_0}^{t_1} \eta \left[\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right)\right] dt.$$

This equality must be true for all functions $\eta(t)$. However, the only way for it to be true for all functions $\eta(t)$ is if

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) = 0. \tag{18}$$

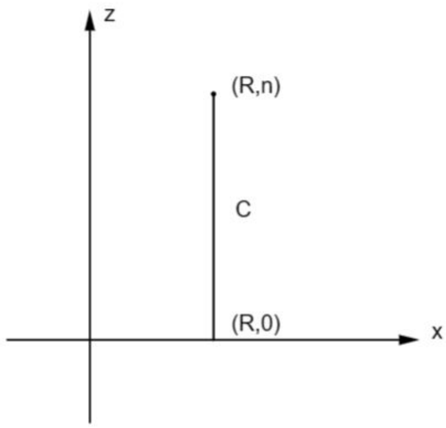
(Although this conclusion may not be obvious, it can be proved using proof by contradiction). This proof is also outside the scope of this paper and is thus omitted, but can be found in [2]. The condition found in equation (18) is the Euler-Lagrange equation. Since equation (18) was found trying to minimize the arc length of a curve, we can see that the Euler-Lagrange equation is a necessary requirement when working with minimizing curves in space, and thus has a natural connection to geodesics.

3 Characterizing Geodesics

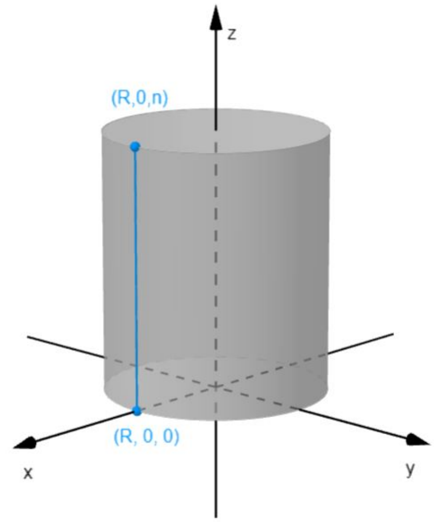
Now that we have knowledge about the geodesics of surfaces of revolution we aim to apply this knowledge and understand some specific surfaces. In this section we will look at geodesics on the cylinder, the sphere, the torus, and the cone.

3.1 The Cylinder

We start by looking at the cylinder as a surface of revolution, as it has a fairly simple profile curve and parameterization, and the geodesic equations can be solved directly.



(a) Profile Curve of the Cylinder



(b) The Cylinder

Figure 4: The Cylinder and its Profile Curve

The profile curve, C, for a right cylinder is given by

$$f(v) = R, \quad \text{and} \quad g(v) = v,$$

with $R \in \mathbb{R}, R \neq 0$, and $0 \leq v \leq n, n \in \mathbb{R}$ (Figure 4a).

Combining this parameterization of C with equation (1) we get that the parameterization for a right cylinder is

$$\mathbf{x}(u, v) = (R\cos(u),\ R\sin(u),\ v),\ \ 0 \leq u < 2\pi. \tag{19}$$

Thus, the profile curve has created a cylinder (Figure 4b). Now we can begin to characterize the geodesics on the cylinder. Before we use the geodesic equations let us look at the meridians and parallels. We know that all meridians on the cylinder are already geodesics, so we just need to find when $f'(v) = 0$ to see which parallels are geodesics. As $f(v) = R$ and $R \in \mathbb{R}$, we know that $f'(v) = 0$ for all v. Therefore, all the parallels of the cylinder are geodesics.

We will now use the geodesic equations to find the other geodesics of the cylinder. Since $f'(v) = 0$, equations (12) and (13) become

$$f^2 u'' = 0, \tag{20}$$

and

$$v'' = 0 \tag{21}$$

respectively. Since $f^2 = R^2$ is a constant, and $R \neq 0$, we know that equation (20) implies $u'' = 0$. Thus, we can integrate these equations with respect to t and get

$$u' = a,$$

and

$$v' = b,$$

where a and b are constants of integration. Integrating once more gives us

$$u = at + c,$$

and

$$v = bt + d,$$

where c and d are also constants of integration. With these results we have obtained the only possible functions for u and v if we want a curve α to be a geodesic. Substituting these functions into our parameterization of the cylinder we get

$$\mathbf{x}(u, v) = (R\cos(at + c),\ \ R\sin(at + c),\ \ bt + d).$$

If we try to visualize this we can see that as t increases, the curve α will go around

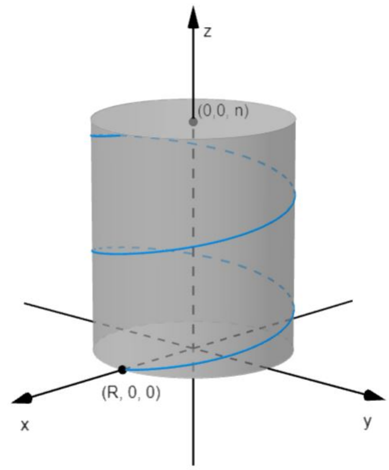
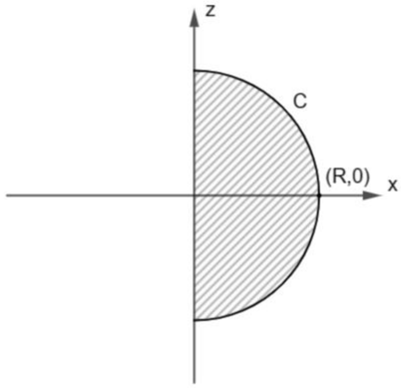


Figure 5: A Helix on the Cylinder

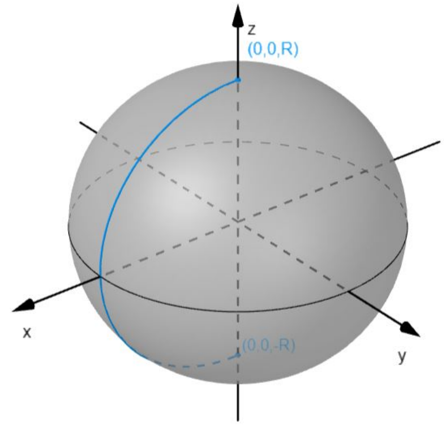
a circle with radius R while also moving upward on the z-axis. This shape is a helix. From the solutions to the geodesic equations we see that we can only express u and v as lines. Thus, the helix is the only non-parallel and non-meridian geodesic on the cylinder. In fact, it is worth noting that the parallels and meridians are just special cases of the helical geodesic. The meridians occur when u is constant and $a = 0$. They create vertical lines on the cylinder. The parallels occur when v is constant and $b = 0$. They create circles around the cylinder. So, all geodesics on a cylinder can be characterized by the helix which can be seen in Figure 5.

3.2 The Sphere

Now that we have solved some simpler geodesic equations, we can begin to look at a more complex surface of revolution, the sphere. There are multiple ways to characterize the geodesics of a sphere with radius R, denoted by S^2_R. In this paper we will focus on two different approaches: one which uses geometry, the unit normal, and a geodesic curve, and one which uses the geodesic equations. However, before we get into these two characterizations it is important to understand the basic parameterization of S^2_R.



(a) Profile Curve of the Sphere



(b) The Sphere (S^2_R)

Figure 6: The Sphere and its Profile Curve

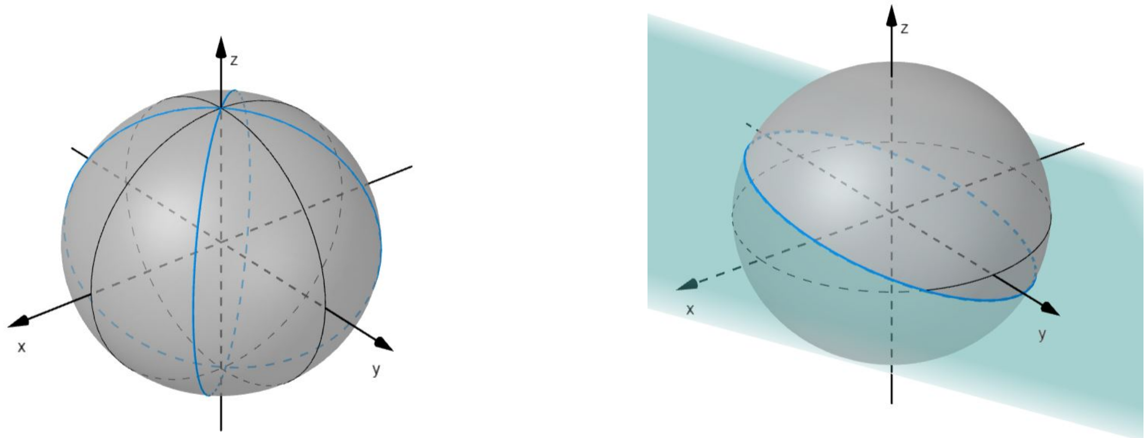
The profile curve, C, for S^2_R is given by $f(v) = R\cos(v)$ and $g(v) = R\sin(v)$ where $0 \leq v \leq \frac{\pi}{2}$ (Figure 6a). From here we use the parameterization for surfaces of revolution from equation (1) to get

$$\mathbf{x}(u, v) = (R\cos(v)\cos(u), R\cos(v)\sin(u), R\sin(v)).$$

With this parameterization we can find which parallels of the sphere are geodesics, as we already know all meridians are geodesics. As $f(v) = \cos(v)$ we get that $f'(v) = -\sin(v)$. Now we just need to find when $\sin(v) = 0$ for $-\frac{\pi}{2} \leq v \leq \frac{\pi}{2}$. The only instance when $\sin(v) = 0$ for the given interval is when $v = 0$. Substituting this value into our parameterization we see that it produces the equator of the sphere. Thus, the only parallel which is a geodesic on the sphere is the equator.

3.2.1 First Characterization of Geodesics on S^2_R

As previously stated, all meridians on the sphere are geodesics. For the sphere, the meridians can be visualized as the longitude lines from the North pole to the South pole. This is similar to our standard notation of longitude lines, e.g. prime meridian.



(a) Meridians on S^2_R (b) A Plane Intersecting S^2_R

Figure 7: Geodesics on the Sphere

As the sphere is symmetric about its center, by any appropriate rotation, we can think of it as a surface of revolution about any line which passes through the center [2]. This means that the meridians can be visualized as circles on the sphere which connect the points of intersection with the sphere. The points that the meridians are connecting are located at the North and South poles, under rotation. So the circles connecting them are great circles on S^2_R. Great circles are circles with radius R on a sphere with radius R, that have the same center as the sphere. This shows that all geodesics will be a section of some great circle. These meridians can be seen in Figure 7a.

The other part of this approach shows that all geodesics on S^2_R are great circles. Let α be a unit speed geodesic on the sphere. From previous calculations, we know that we can break up the second derivative of α into a tangent and a normal component. So we have $\alpha'' = \alpha''_{\text{tan}} + \alpha''_{\text{normal}}$. Using knowledge about curvature, U, and T, from the Frenet Formulas, we find that $\alpha''_{\text{normal}} = (\alpha'' \cdot U)U$ (for more detail see [2]). Substituting this into the formula for α'' we get

$$\alpha'' = \alpha''_{\text{tan}} + (\alpha'' \cdot U)U.$$

As the tangent component of α'' must be zero for α to be a geodesic we see that

$$\alpha'' = (\alpha'' \cdot U)U. \tag{22}$$

To simplify this equation we will need another way to express U. We can evaluate $U(\alpha(t))$ to get a substitution. As U is the unit normal, if we were at a point we would evaluate U at the point and make sure it was normalized. However, since we are looking at a curve we will get $U(\alpha(t)) = \frac{\alpha(t)}{R}$, where R comes from the normalization process [2]. Substituting this into equation (22) we get

$$\begin{aligned} \alpha'' &= \frac{\alpha}{R}(\alpha'' \cdot \frac{\alpha}{R}) \\ &= \frac{1}{R^2}(\alpha'' \cdot \alpha)\alpha. \end{aligned} \tag{23}$$

Equation (23) tells us that α'' is parallel to α. Now consider $(\alpha' \times U)'$. Using the same substitution we used above we get

$$\begin{aligned} (\alpha' \times U)' &= \frac{1}{R}(\alpha' \times \alpha)' \\ &= \frac{1}{R}(\alpha'' \times \alpha + \alpha' \times \alpha'). \end{aligned}$$

As α'' is parallel to α, and α' is parallel to itself, we get that both cross products are zero so

$$(\alpha' \times U)' = 0,$$

which means $\alpha' \times U$ is a constant vector. Let $\alpha' \times U = \mathcal{N}$. Since the cross product of two vectors is always orthogonal to the original vectors we know that $\mathcal{N}$ is perpendicular to U. We also know that $U = \frac{\alpha}{R}$. Thus, $\mathcal{N}$ is also perpendicular to α and $\alpha \cdot \mathcal{N} = 0$. Therefore, α lies in the plane that has $\mathcal{N}$ as a normal vector. Since $\mathcal{N}$ is perpendicular to U, we know that U also lies in this plane. As α and U lie in the same plane, we can get from any geodesic curve $\alpha(t)$ to the origin and stay in the plane [2]. Hence, the plane goes through (0,0,0). As α lies on S^2_R it is the intersection of the sphere with a plane that passes through the origin. Therefore, α is part of a great circle, as when planes intersect spheres through the center of the sphere they create great circles. This intersection can be seen in Figure 7b.

So, from this approach we can see that all great circles on S^2_R are geodesics, moreover they are the only geodesics on the sphere.

3.2.2 Second Characterization of Geodesics on S^2

It is worthwhile to directly solve the geodesic equations to characterize the geodesics on a sphere. For slightly more simplicity in the calculations we will use the unit sphere, S^2, but the results still hold for S^2_R.

The parameterization of S^2 is

$$\mathbf{x}(u, v) = (\cos(v)\cos(u), \cos(v)\sin(u), \sin(v)),$$

with $-\frac{\pi}{2} \leq v < \frac{\pi}{2}$ ad $0 \leq u \leq 2\pi$. Combining this with equations (12) and(13) we see that the geodesic equations for the unit sphere are

$$u'' - 2\tan(v)u'v' = 0,$$

and

$$v'' + \sin(v)\cos(v)u'^2 = 0$$

respectively. Although these equations look somewhat complex we can solve them directly using the unit speed requirement from equation (14).

Let $\alpha(t) = \mathbf{x}(u(t), v(t))$ be a unit speed geodesic on the sphere. With the unit speed requirement we have $\alpha' = u'\mathbf{x}_u + v'\mathbf{x}_v$, where $1 = f^2u'^2 + v'^2$. On S^2, this condition simplifies to $1 = \cos^2(v)u'^2 + v'^2$. Now that we have this equation we will be able to solve the geodesic equations. Starting with the first equation we get

$$
\begin{aligned}
u'' &= 2\tan(v)v'u' \\
\frac{u''}{u'} &= 2\tan(v)v' \\
\int \frac{u''}{u'}\,du &= \int 2\tan(v)v'\,dv \\
\ln u' &= -2\ln\cos(v) + C \\
u' &= \frac{c}{\cos^2(v)} \quad \text{where } c = e^C.
\end{aligned}
$$

This is where the unit speed requirement comes into play. We can use the value for u' that we just calculated and substitute it into the unit speed equation. This results in

$$
\begin{aligned}
1 &= \left(\frac{c}{\cos^2(v)}\right)^2 \cos^2(v) + v'^2 \\
1 &= \frac{c^2}{\cos^4(v)}\cos^2(v) + v'^2 \\
v'^2 &= 1 - \frac{c^2}{\cos^2(v)} \\
v' &= \pm\sqrt{\frac{\cos^2(v) - c^2}{\cos^2(v)}}.
\end{aligned}
$$

Now that we have equations for both u' and v' we can get a separable differential equation by dividing u' by v'. This gives us

$$
\frac{du}{dv} = \frac{\pm c}{\cos(v)\sqrt{\cos^2(v) - c^2}}.
$$

This equation can be integrated. These steps will be left out as they are not needed to understand the geodesic equations and are somewhat tedious. They can be found in full in [2]. Once the equation is integrated we get

$$\sin(u - d) = \lambda \tan(v),$$

where d is a constant of integration and $\lambda = \frac{c}{\sqrt{1-c^2}}$. We can expand the first term $\sin(u-d)$ into $\sin(u)\cos(d) - \sin(d)\cos(u)$ and substitute this into the equation. We can then find a common denominator, which will be $\cos(v)$. After some rearrangement and substitution where $x = \cos(u)\cos(v), y = \sin(u)\cos(v)$ and $z = \sin(v)$ we get

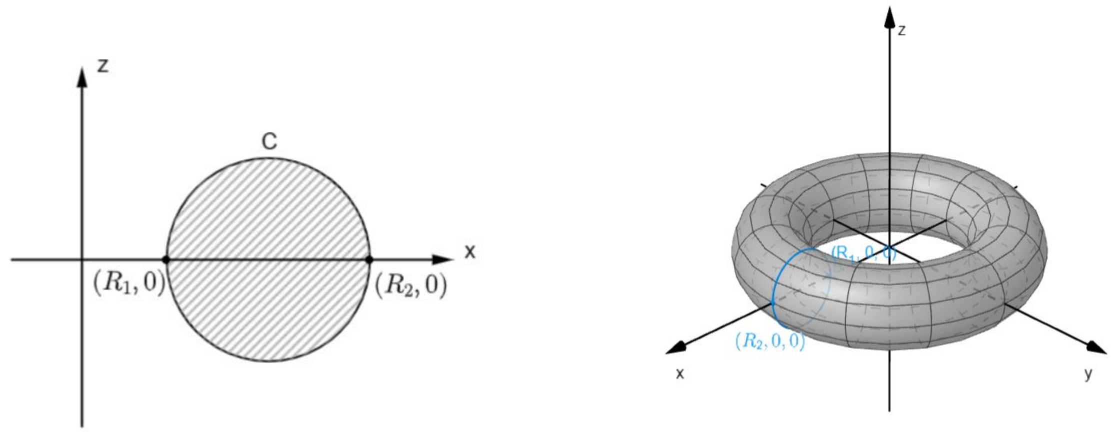
$$y\cos(d) - x\sin(d) - \lambda z = 0$$

as the final solution to the geodesic equations [2].

Notice that this is the equation for a plane $ax + by + cz = 0$ through the origin. Thus, this solution implies that α lies on such a plane. This means that α is contained in the intersection between a plane through the origin with the sphere. We have previously seen this is a great circle (Figure 7b). Thus, using a direct approach we have shown that all geodesics on a sphere follow the path of a great circle. This knowledge is readily applied to the paths of airplanes. If a plane is traveling between to airports, rather than following the lines of latitude, or some other path, the plane will want to follow the path of the great circle that connects the two airports. In this way, geodesics on the sphere are quite helpful in application.

3.3 The Torus

After looking at the sphere, a natural surface of revolution to move on to is the torus. The torus is denoted $T^2(R_1, R_2)$ where R_1 and R_2 are the inner and outer radii, respectively. We look at the torus right after the sphere as the profile curve, C, for the torus is a circle of radius $\frac{R_1+R_2}{2}$ which is centered at $(x, z) = (\frac{R_1+R_2}{2}, 0)$ (Figure



(a) Profile Curve of the Torus (b) The Torus ($T^2(R_1, R_2)$)

Figure 8: The Torus and its Profile Curve

8a).

Thus, we parameterize C by

$$f(v) = \frac{R_1 + R_2}{2} + \frac{R_2 - R_1}{2}\cos(v),$$
$$g(v) = \frac{R_2 - R_1}{2}\sin(v),$$

with $0 \leq v < 2\pi$. With this profile curve we parameterize $T^2(R_1, R_2)$ by

$$x(u, v) = (x, y, z)$$

where

$$x = \left(\frac{R_1 + R_2}{2} + \frac{R_2 - R_1}{2}\cos(v)\right)\cos(u),$$
$$y = \left(\frac{R_1 + R_2}{2} + \frac{R_2 - R_1}{2}\cos(v)\right)\sin(u),$$
$$z = \frac{R_2 - R_1}{2}\sin(v),$$

and $0 \leq u < 2\pi$. As we have established all the meridians on the torus will be geodesics. Now to classify the parallels of the torus. With $f(v)$ as previously defined we have $f'(v) = \frac{R_1 - R_2}{2}\sin(v)$ where $0 \leq v < 2\pi$. So $\sin(v) = 0$ when $v = 0$ or $v = \pi$. These values correspond to the inner and outer parallels. So, those are the only parallels which are geodesics for the torus. For the previous surfaces we examined we finished characterizing the geodesics by solving the geodesics equations. However, for the torus the geodesic equations are not directly solvable, so we have to come up with another approach to characterize the geodesics.

We will use Clairaut's relation with some other information to characterize the geodesics of the torus.

Theorem 3.1 (Clairaut's Relation)**.** *Consider a geodesics α on a surface of revolution S as parameterized in equation (1). If α intersects a parallel of S, let θ be the angle between α and that parallel and f be the radial distance the point of intersection is from the axis of revolution. Then along α, we have the constant relationship*

$$c = f\cos(\theta). \tag{24}$$

For a proof of Theorem 3.1 see [2]. Along with Clairaut's relation we will use the following theorem to help find other geodesics on the torus.

Theorem 3.2. *Except for the inner parallel, every geodesics on the torus must intersect the outer parallel [1]*

Proof. We know that all meridians intersect the outer parallel at right angles. Thus, we assume that a geodesic α intersects the parallels of $T^2(R_1, R_2)$ with a radial distance function f, and an angular function $\theta \neq \frac{\pi}{2}$. We know that f is bounded above by R_2, so we only need to show that α cannot asymptotically approach the outer parallel. If this were to happen, we would get $\lim_{f \to R_2}\theta = 0$. However, we see from equation (24) that f and θ increase together. Therefore, α must intersect the outer parallel. □

The last component we need to proceed with the classification of geodesics on the torus is the critical angle of the torus which we define as

$$\theta_c = \cos^{-1}\left(\frac{R_1}{R_2}\right).$$

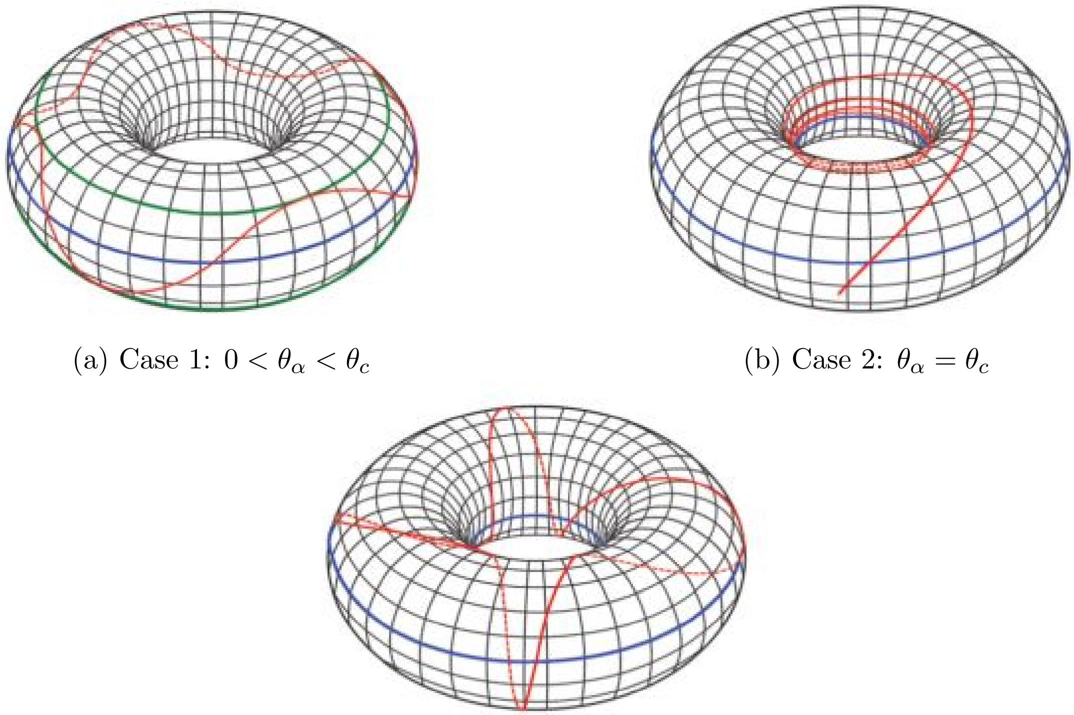
With this information we can now classify the geodesics on the torus according to the relationship of the critical angle and the geodesic's crossing angle. If we have a have a geodesic α which crosses the outer parallel of the torus at an angle $\theta_\alpha \in (0, \frac{\pi}{2})$ we have the following three cases.

3.3.1 Case 1. $0 < \theta_\alpha < \theta_c$

As cosine is a decreasing function, we know $\cos(\theta_\alpha) > \cos(\theta_c) = \frac{R_1}{R_2}$. This implies $R_2 \cos(\theta_\alpha) > R_1$. As R_2 is the distance at which α is crossing the outer parallel it corresponds to f from Theorem 3.1 and we get $c = R_2 \cos(\theta_\alpha)$. Therefore, $c > R_1$ and we see that the critical distance of α must always be greater than R_1. This implies that α will never cross the inner parallel. In fact, α will bounce between the top and bottom parallel of the torus (Figure 9a). This behavior arises from constraints related to the Clairaut relation and can be explored in more detail in [2].

3.3.2 Case 2. $\theta_\alpha = \theta_c$

This case means $\cos(\theta_\alpha) = \cos(\theta_c) = \frac{R_1}{R_2}$, so $R_2 \cos(\theta_\alpha) = R_2 \cos(\theta_c) = R_1$. As $c = R_2 \cos(\theta_\alpha)$ we can see that $c = R_1$. We know that R_1 is one of the two parallels that are geodesics. We also know that a geodesic is "locally uniquely determined by a point and direcion" [1]. So if α intersects the inner parallel both geodesics would share a tangent vector, which cannot happen. Therefore, α spirals asymptotically toward the inner parallel (Figure 9b).



(a) Case 1: $0 < \theta_\alpha < \theta_c$

(b) Case 2: $\theta_\alpha = \theta_c$

(c) Case 3: $\theta_c < \theta_\alpha$

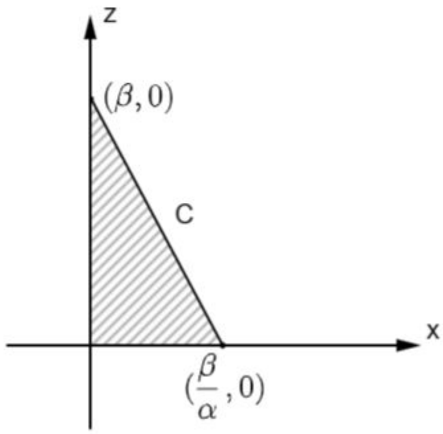
Figure 9: Geodesics on the Torus [1]

3.3.3 Case 3. $\theta_c < \theta_\alpha < \frac{\pi}{2}$

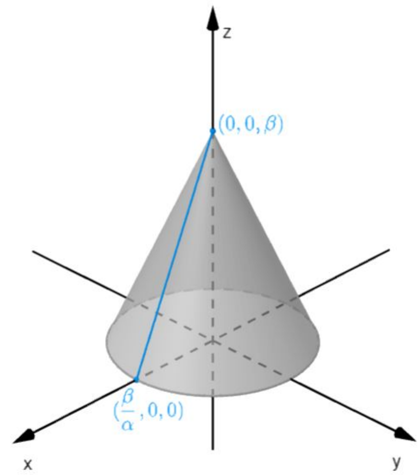
This case is opposite the first, it leads to $\cos(\theta_\alpha) < \cos(\theta_c)$. We have $c = R_2 \cos(\theta_\alpha)$ so we get

$$c = R_2 \cos(\theta_\alpha) < R_2 \cos(\theta_c) = R_1.$$

Therefore, $c < R_1$ and since no distance on the torus is less than R_1 the geodesic α can never reach its critical distance by intersection or asymptotic approach (Figure 9c) [1]. With those 3 cases, as well as the already classified meridians and parallels, we have characterized all the geodesics on the torus.



(a) Profile Curve of the Cone



(b) The Cone

Figure 10: Right Circular Cone and its Profile Curve

3.4 The Cone

We will examine geodesics on the right circular cone, although some of the conclusions we draw are valid for more general cones. We parameterize the profile curve, C, by

$$f(v) = v$$
$$g(v) = \alpha v + \beta,$$

with $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$. This parameterization produces the profile curve in Figure 10a. Thus, the parameterization for the cone is

$$\mathbf{x}(u, v) = (v\cos(u),\ v\sin(u),\ \alpha v + \beta)$$

and is visualized in Figure 10b. We know that all meridians on the cone are geodesics, so we only need to calculate $f'(v)$ to see which parallels are geodesics. As $f'(v) = 1$ and $1 \neq 0$ we see that there are no values for which $f'(v) = 0$ and therefore no parallels of the cone are geodesics.

To characterize the other geodesics, we will look at a different way to parameterize the cone. Consider a circular cone in Euclidean three-space (E^3). For a given curve, $y = y(t)$ is defined on an open interval, I, which lies on the unit sphere centered at the origin. Let C_y denote the cone with vertex at the origin over the curve y. We can parameterize C_y so that

$$C_y(t, u) = uy(t), \quad u \in \mathbb{R}^+[3]. \tag{25}$$

If we parameterize the cone as in equation (25), we define a ruling to be a curve $\beta(u) = C_y(t_0, u)\ \ u \in \mathbb{R}^+$ for some $t_0 \in I$ [3]. We will combine this new parameterization of the cone with the following two Lemmas to find the geodesic curves.

Lemma 3.3. *Every open portion of a ruling is a geodesic of the cone [3].*

Lemma 3.3 is saying that all meridians are geodesics, but using language concerning the new parameterization of a cone, where meridians are equivalent to rulings.

Lemma 3.4. *Let $y(t)$ be a unit speed curve on the unit sphere centered at the origin. Then, for each $c \in \mathbb{R}^+$ the curve $z(t) = cy(t)$ is not a geodesic of the cone defined by equation (25) [3].*

The proofs for Lemmas 3.3 and 3.4 have been omitted as they are not strictly necessary for understanding the characterization of the geodesics and are somewhat of a divergence of the focus of this paper. Full proofs can be found in [3].

Now that we have Lemmas 3.4 and 3.3, we can characterize the other geodesics on a cone. Let C_y be a cone parameterized as in equation (25). Without loss of generality, we can assume that $y(t)$ is a unit speed curve. Let z be a curve on C_y. If $z(t)$ is a meridian, we know that it is a geodesic, so let us assume that $z(t)$ is not a meridian. Since $z(t)$ does not lie in any meridian, we may say

$$z(t) = p(t)y(t), \ \ t \in I,$$

for some function $p(t)$, where I is an open interval [3].

If $p(t)$ is a nonzero constant, say c, then we have $z(t) = cy(t)$, which from Lemma 3.4 we know cannot be a geodesic. From this, we know that $p' \neq 0$, which will be important when solving the Euler-Lagrange equation. Now we must use the Euler-Lagrange equation. We want to evaluate the arc length of $z(t)$ and find when it is minimized. Thus, we can consider the function f for the arc length of $z(t)$. This function f will be given by $f(t,p,p') = \sqrt{p^2 + p'^2}$. For the curve $z(t)$ to be minimized f must satisfy the Euler-Lagrange equation. Thus, using equation (18) we get

$$pp'' - 2p'^2 - p^2 = 0.$$

This differential equation can be solved directly to obtain $p = a\sec(t+b)$ for some real numbers a, b where $a \neq 0$ [3]. Therefore,

$$z(t) = a\sec(t+b)y(t).$$

From [4] we know that this is the parameterization of a rectifying curve. Therefore, all curves which are not meridians are geodesics if and only if they are rectifying curves. An example of a rectifying curve on a cone can be seen in Figure 11. With that characterization, all geodesics on the cone have been identified.

4 The Metric Tensor and Arc Length

Although we now know the geodesics on different surfaces of revolution we have no way to tell the arc length of the geodesic. In this section we will look at a method for finding arc length no matter what surface we are on. Once we have established this method we will find the arc length of two different curves on the cylinder.

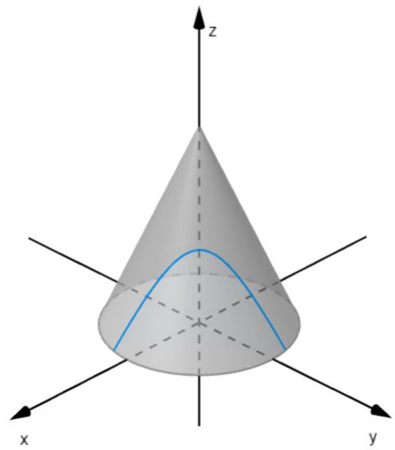


Figure 11: Geodesics on a Cone

4.1 The Metric Tensor

Now that we have classified the geodesics on different surfaces, it is natural to want to calculate the distance of a geodesic or any curve on the surface. One way to do this is to use the metric tensor for a surface. The metric tensor is a function that allows us to compute the distance between any two points on a surface [5]. When we are using a Cartesian coordinate system, in 2D, we know that $ds^2 = dx^2 + dy^2$. So finding the distance between two points is relatively simple. However, if we were to use polar coordinates this relationship no longer holds true, $ds^2 \neq dr^2 + d\theta^2$. We know that to convert from Cartesian to polar coordinates we have

$$x = r\cos(\theta), \;\; y = r\sin(\theta).$$

Taking the derivatives gives us

$$dx = \cos(\theta)dr - r\sin(\theta)d\theta$$
$$dy = \sin(\theta)dr + r\cos(\theta)d\theta.$$

If we substitute these values into the distance formula for Cartesian coordinates we get

$$ds^2 = dr^2 + r^2 d\theta^2.$$

So, if we wanted to calculate the distance between two points in polar coordinates we would have to use this equation instead of the Pythagorean Theorem.

Now consider a general coordinate system in 3D that is not the Cartesian system, and assume we want to find how far apart two points are in that coordinate system. Let

$$q_1 = q_1(x, y, z)$$
$$q_2 = q_2(x, y, z)$$
$$q_3 = q_3(x, y, z)$$

be our general coordinate system. Then, finding dx, dy, and dz like we did for polar coordinates, we get

$$\begin{aligned}
dx &= \frac{\partial x}{\partial q_1} dq_1 + \frac{\partial x}{\partial q_2} dq_2 + \frac{\partial x}{\partial q_3} dq_3 \\
dy &= \frac{\partial y}{\partial q_1} dq_1 + \frac{\partial y}{\partial q_2} dq_2 + \frac{\partial y}{\partial q_3} dq_3 \\
dz &= \frac{\partial z}{\partial q_1} dq_1 + \frac{\partial z}{\partial q_2} dq_2 + \frac{\partial z}{\partial q_3} dq_3.
\end{aligned}$$

Substituting these values into our distance formula gives us [5]

$$
\begin{aligned}
ds^2 &= g_{11}dq_1^2 + g_{12}dq_1dq_2 + g_{13}dq_1dq_3 \\
&\quad + g_{21}dq_2dq_1 + g_{22}dq_2^2 + g_{23}dq_2dq_3 \\
&\quad + g_{31}dq_3dq_1 + g_{32}dq_3dq_2 + g_{33}dq_3^2 \\
&= \Sigma_{ij}g_{ij}dq_idq_j.
\end{aligned}
\tag{26}
$$

The term g_{ij} in the sum is the entry in the ith row of the jth column of the metric tensor. Note that equation (26) is valid for all Reimannian spaces. Thus, with the metric tensor we can find the distance between any two points on any Reimannian Surface [5].

The entries for the metric tensor can be calculated several different ways. First we have [5]

$$
g_{ij} = \frac{\partial x}{\partial q_i}\frac{\partial x}{\partial q_j} + \frac{\partial y}{\partial q_i}\frac{\partial y}{\partial q_j} + \frac{\partial z}{\partial q_i}\frac{\partial z}{\partial q_j}.
$$

Although this equation works to get the metric tensor, we can also find g_{ij} by calculating the dot products of the basis vectors [6]. So, for a surface with n different basis vectors we get

$$
g = \begin{bmatrix}
\vec{e}_1 \cdot \vec{e}_1, & \vec{e}_1 \cdot \vec{e}_2, & \cdots & \vec{e}_1 \cdot \vec{e}_n \\
\vec{e}_2 \cdot \vec{e}_1, & \vec{e}_2 \cdot \vec{e}_2, & \cdots & \vec{e}_2 \cdot \vec{e}_n \\
\vdots & \vdots & \cdots & \vdots \\
\vec{e}_n \cdot \vec{e}_1, & \vec{e}_n \cdot \vec{e}_2, & \cdots & \vec{e}_n \cdot \vec{e}_n
\end{bmatrix}.
$$

For surfaces of revolution with parameterization matching equation (1) one possible set of basis vectors are $\mathbf{x}_u$ and $\mathbf{x}_v$. Thus, the metric tensor for a surface of revolution is

$$
\begin{bmatrix}
\mathbf{x}_u \cdot \mathbf{x}_u & \mathbf{x}_u \cdot \mathbf{x}_v \\
\mathbf{x}_v \cdot \mathbf{x}_u & \mathbf{x}_v \cdot \mathbf{x}_v
\end{bmatrix}.
$$

As we already have values for these dot products the metric tensor can be simplified to

$$\begin{bmatrix} E & F \\ F & G \end{bmatrix}. \tag{27}$$

Now that we have the metric tensor for a surface of revolution, we can find the distance between any two points. The metric tensor also provides another way to classify geodesics. As we have an equation for distance, we can take the general distance equation and try to find when a minimum occurs by taking the derivative and solving for zero. This method also involves taking integrals which can sometimes be difficult to calculate, and require lots of algebra, so it is not necessarily better than the approach we have been using.

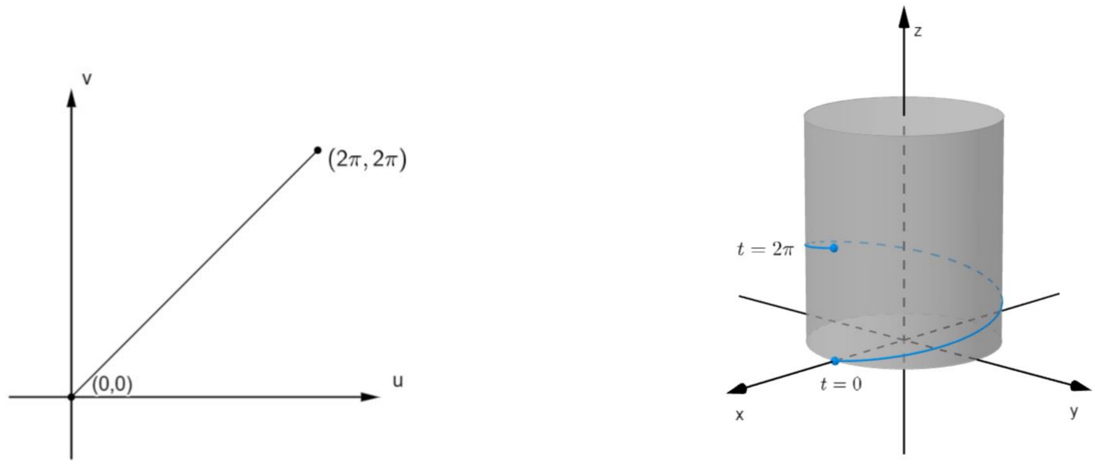
4.2 Finding Arc Length

Now that we have established the metric tensor we will examine how it can be used to find the arc length of a curve on a surface of revolution by going through two different examples. One example will use a simple parameterized curve, and the other will use a slightly more complex, but still nice, parametric curve. We will visualize the curve first in parameter space and then see how it appears on a given surface.

4.2.1 Example 1

Since the cylinder is one of the simplest surfaces of revolution, we will look at arc length on curves lying on the cylinder. Let $u(t) = v(t) = t$ where $0 \leq t \leq 2\pi$ so we get a line in parameter space (Figure 12a).

Now that we have our parameter curve, we can visualize it on the cylinder (Figure 12b). Using the metric tensor for surfaces of revolution found in equation (27), and



(a) Curve in Parameter Space (b) Curve on the Cylinder

Figure 12: Line Parameter Curve Graphs

the parameterization of the cylinder in equation (19) we get that the metric tensor is

$$\begin{bmatrix} R^2 & 0 \\ 0 & 1 \end{bmatrix}. \tag{28}$$

With the metric tensor, we can use equation (26) to get the squared distance of the curve. Using the values from the matrix in (28) we get

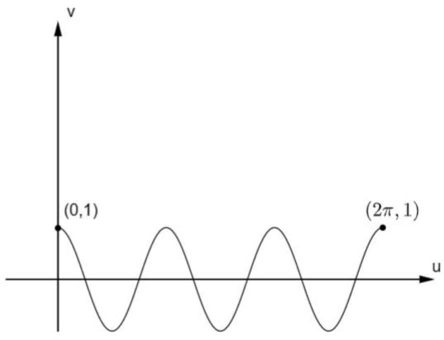
$$ds^2 = R^2 \frac{du}{dt} + 1 \cdot \frac{dv}{dt}. \tag{29}$$

Now, we just need to calculate $\frac{du}{dt}$ and $\frac{dv}{dt}$. As $u(t) = v(t) = t$, we have $\frac{du}{dt} = \frac{dv}{dt} = 1$. Therefore,

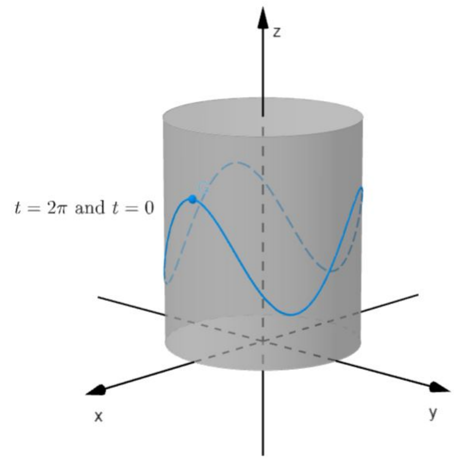
$$ds^2 = R^2 + 1.$$

The last step to find the arc length of the curve is to integrate over the distance. As we have calculated the squared distance, for the integral of the distance we get

$$\int_0^{2\pi} \sqrt{1 + R^2}\, dt.$$



(a) Curve in Parameter Space



(b) Curve on the Cylinder

Figure 13: Cosine Parameter Curve Graphs

Evaluating this integral, we find that the arc length of the curve is $2\pi\sqrt{1+R^2}$.

4.2.2 Example 2

To make the example slightly more complex we will look at a cosine curve in parameter space (Figure 13a). Using this curve, we have

$$u(t) = t, \;\; v(t) = \cos(3t), \quad 0 \leq t \leq 2\pi. \tag{30}$$

As we are still considering a curve on the cylinder we will use the same metric tensor, so the values R^2 and 1 will remain the same. However, as we have a different curve, we will need to recompute $\frac{du}{dt}$ and $\frac{dv}{dt}$. Using our parameterization from equation (30) we get

$$\frac{du}{dt} = 1, \;\; \frac{dv}{dt} = -3\sin(t).$$

Combining this with equation (29) yields

$$ds^2 = R^2 + (-3\sin(t))^2 = R^2 + 9\sin^2(t).$$

Substituting the square root of this expression into the integral we can see that the arc length of the curve is given by

$$\int_0^{2\pi} \sqrt{R^2 + 9\sin^2(t)}\, dt.$$

As the metric tensor relies on the radius of the cylinder R the arc length of the curve will change depending on the R value. For example, if we have a cylinder with a radius of 2 the arc length will be

$$\int_0^{2\pi} \sqrt{2^2 + 9\sin^2(t)}\, dt \approx 17.97.$$

However, if we have a cylinder with a radius of 5.5, the arc length will become

$$\int_0^{2\pi} \sqrt{5.5^2 + 9\sin^2(t)}\, dt \approx 36.99.$$

Thus, with these calculations we have found the general arc length of the curve from equation (30) on a cylinder. The approach that we used for this example can be used for any curve on a Reimannian surface to find the arc length of the curve.

5 Conclusion

Using the geodesic equations (equations (12) and (13)) and some extra knowledge about surfaces and curves, we were able to classify the geodesics on four different surfaces of revolution. We found that the geodesics for a cylinder are all helical, with the meridians and the parallels being special helical cases. For the sphere, all geodesics follow the path of a great circle. The torus was slightly more difficult to compute as we could not solve the geodesic equations directly. We found that the meridians and inner and outer parallels were geodesics. The other geodesics on the torus can be classified depending on the angle of intersection between the curve and the outer parallel. Lastly, we saw that except for the meridians, all geodesics on a cone are rectifying curves. We then saw that given any geodesic, or curve, on these surfaces we can find the arc length using the metric tensor for that surface. The metric tensor is not only useful for finding the arc length of specific curves, but provides another approach to finding geodesics.

Now that we know the shortest path between two points on several different surfaces, how can we apply this knowledge? Knowing the geodesics of a sphere is very useful for any sort of air or sea travel. When airplanes or water crafts want to take the shortest path possible between two places they just need to take a path that lies on a great circle. Geodesics are also applicable to general relativity. In fact, a free particle follows the path of a geodesic [6]. A free particle is a particle that has no external forces acting on it. For example, a particle that is approaching the event horizon of a black hole follows the path of a geodesic [6]. Using the fact that free particles follow geodesics, researchers are able to learn more about objects in space and how they manipulate the space-time around them. With just these applications as examples we can see that geodesics are not only relevant to math but other areas of life as well.

Acknowledgements

Special thanks to Sam Fix for reading all my progress journals and giving me valuable feedback, and to Benjamin Seashore Hobson for giving me cool math to read about every week. I would also like to thank Professor Balof for making class entertaining and informative. Lastly, thanks to Professor Schueller for answering all my questions and helping me learn about differential geometry.

References

[1] K. Celano, V. E. Coll, and J. Dodd, "Why curves curve: The geodesics on the torus," *Mathematics Magazine*, vol. 95, no. 3, p. 230–239, 2022.

[2] J. Oprea, *Differential geometry and its applications (classroom resource materials)*. Mathematical Association of America, 2007.

[3] B.-Y. Chen, "Rectifying curves and geodesics on a cone in the euclidean 3-space," *Tamkang Journal of Mathematics*, vol. 48, no. 2, p. 209–214, 2017.

[4] B.-Y. Chen, "When does the position vector of a space curve always lie in its rectifying plane?," *The American Mathematical Monthly*, vol. 110, no. 2, p. 147, 2003.

[5] G. Arfken, *Mathematical methods for physicists, 3rd edition (PBK)*. Academic Press, 1985.

[6] T. A. Moore, *A general relativity workbook*. University Science Books, 2013.