Classical Differential Geometry

Peter Petersen

Preface

This is an evolving set of lecture notes on the classical theory of curves and surfaces. Pictures will be added eventually. I recommend people download 3DXplorMath to check out the constructions of curves and surfaces with this app. It can also be used to create new curves and surfaces in parametric form. Other useful and free apps are Geogebra, Grapher (on Mac), and WolframAlpha.

At a minimum a one quarter course should cover sections 1.1, 1.2, 1.3, 2.1, 3.1, 3.2 and chapters 4, 5. In a semester course it'd be possible to cover more from chapter 2 and also delve into chapter 6. Chapters 6 and 7 can be covered in a second quarter class. Note that section 2.2 is a necessary prerequisite for proving the general Gauss-Bonnet in section 6.5.

An excellent reference for the classical treatment of differential geometry is the book by Struik [2]. The more descriptive guide by Hilbert and Cohn-Vossen [1] is also highly recommended. This book covers both geometry and differential geometry essentially without the use of calculus. It contains many interesting results and gives excellent descriptions of many of the constructions and results in differential geometry.

This text is fairly classical and is not intended as an introduction to abstract 2-dimensional Riemannian geometry. In fact we do not discuss covariant differentiation or parallel translation. Most proofs are local in nature and try to use only basic linear algebra and multivariable calculus. The only sense in which the text is more modern is in not using the language of differentials and infinitesimals as most of the classical texts do.

Some standard topics are not covered in the text. However, I hope most of them can be found among the exercises. As such, they can easily be incorporated into lectures as the instructor sees fit.

I'd like to thank Chadwick Sprouse and Michael Williams for trying out these notes and providing valuable feedback.

> "Reading your notes is like reading poetry, and I don't understand that either." Reed Douglas, UCLA student.

Contents

CHAPTER 1

General Curve Theory

One of the key aspects in geometry is *invariance*. This can be somewhat difficult to define, but the idea is that the properties or measurements under discussion should be described in such a way that they they make sense without reference to a special coordinate system. This idea has been a guiding principle since the ancient Greeks started formulating geometry. We'll often take for granted that we work in a Euclidean space where we know how to compute distances, angles, areas, and even volumes of simple geometric figures. Descartes discovered that these types of geometries could be described by what we call Cartesian space through coordinatizing the Euclidean space with Cartesian coordinates. This is the general approach we shall use, but it is still worthwhile to occasionally try to understand measurements not just algebraically or analytically, but also purely descriptively in geometric terms. For example, how does one define a circle? It can defined as a set of points given by a specific type of equation, it can be given as a parametric curve, or it can be described as the collection of points at a fixed distance from the center. Using the latter definition without referring to coordinates is often a very useful tool in solving many problems.

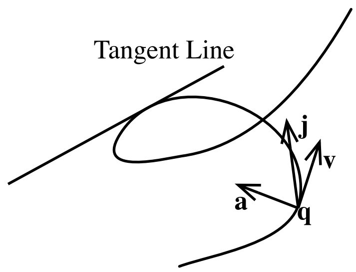
1.1. Curves

The primary goal in the geometric theory of *curves* is to measure their shapes in ways that do not take in to account how they are parametrized or how Euclidean space is coordinatized. However, it is generally hard to measure anything without coordinatizing space and parametrizing the curve. Thus the idea will be to see if some sort of canonical parametrization might exist and secondly to also show that our measurements can be defined using whatever parametrization the curve comes with. We will also try to make sure that our formulas do not necessarily refer to a specific set of Cartesian coordinates. To understand more general types of coordinates requires quite a bit of work and this will not be done until we introduce surfaces later in these notes.

Imagine traveling in a car or flying an airplane. The route traveled will trace a curve. You can easily keep track of time and distance traveled. The goal of curve theory is to decide what further measurements are needed to retrace the precise path traveled. Clearly one must also measure how one turns and that becomes the important thing to describe mathematically.

The fundamental dynamical vectors of a curve whose position is denoted by q are the *velocity* $\mathrm{v} = \frac{d\,\mathrm{q}}{dt}$, *acceleration* $\mathrm{a} = \frac{d^2\,\mathrm{q}}{dt^2}$, and *jerk* $\mathrm{j} = \frac{d^3\,\mathrm{q}}{dt^3}$.

The *tangent line* to a curve q at $\mathrm{q}(t)$ is the line through $\mathrm{q}(t)$ with direction $\mathrm{v}(t)$. The goal is to find geometric quantities that depend on velocity (or tangent lines), acceleration, and jerk that completely determine the path of the curve when we use some parameter t to travel along it.



Most of the curves we study will be given as *parametrized curves*, i.e.,

$$\mathrm{q}(t)=\begin{bmatrix} x(t) \\ y(t) \\ \vdots \end{bmatrix} : I \to \mathbb{R}^n,$$

where $I \subset \mathbb{R}$ is an interval. Such a curve might be constant, which is equivalent to its velocity vanishing everywhere.

Definition 1.1.1. A curve is called *regular* if it is never stationary. In other words, the speed is always positive, or the velocity never vanishes.

Occasionally curves are given to us in a more implicit form. They could come as solutions to first order differential equations

$$\frac{d\,\mathrm{q}}{dt}=F(\mathrm{q}(t),t).$$

In this case we obtain a unique solution (also called an integral curve) as long as we have an initial position $\mathrm{q}(t_0)=\mathrm{q}_0$ at some initial time t_0. In case the function $F(\mathrm{q})$ only depends on the position we can visualize it as a vector field as it gives a vector at each position. The solutions are then seen as curves whose velocity at each position q is the vector $\mathrm{v}=F(\mathrm{q})$.

Very often the types of differential equations are of second (or even higher order)

$$\frac{d^2\,\mathrm{q}}{dt^2}=F\left(\mathrm{q}(t),\frac{d\,\mathrm{q}}{dt},t\right).$$

In this case we have to prescribe both the initial position $\mathrm{q}(t_0)=\mathrm{q}_0$ and velocity $\mathrm{v}(t_0)=\mathrm{v}_0$ in order to obtain a unique solution curve.

The next result shows how differential equations can be used to characterize curves.

Proposition 1.1.2. *The following conditions are equivalent for a regular curve* $\mathrm{q}(t)$*:*

(1) *The curve travels along a line:* $\mathrm{q}(t)=\mathrm{q}_0+\alpha(t)\,\mathrm{v}_0$*, where* $\alpha(t)$ *is a scalar valued function and* $\mathrm{q}_0,\mathrm{v}_0$ *are fixed vectors.*
(2) *The velocities are all parallel to each other:* $\mathrm{v}(t)=\beta(t)\,\mathrm{v}_0$*, where* $\beta(t)$ *is a scalar valued function and* v_0 *is a fixed vector.*

(3) *The velocity and acceleration at each point are parallel to each other:* $\mathrm{a}(t) = \gamma(t)\,\mathrm{v}(t)$, *where* $\gamma(t)$ *is a scalar valued function.*

PROOF. (1) $\Rightarrow$ (2): Use $\beta(t) = \dot{\alpha}(t)$.

(2) $\Rightarrow$ (3): Since the curve is regular $\beta(t) \neq 0$. Thus we can use $\gamma(t) = \frac{\dot{\beta}(t)}{\beta(t)}$.

(3) $\Rightarrow$ (1): The equation $\mathrm{a}(t) = \gamma(t)\,\mathrm{v}(t)$ can be written as a differential equation

$$\frac{d\,\mathrm{v}}{dt} = \gamma(t)\,\mathrm{v}\,.$$

This shows that

$$\mathrm{v}(t) = \mathrm{v}(t_0)\exp\left(\int_{t_0}^{t}\gamma\right),$$

since the right hand side solves the equation and has the same initial value at t_0 as the left hand side. Thus we obtain a new differential equation

$$\frac{d\,\mathrm{q}}{dt} = \beta(t)\,\mathrm{v}_0,$$

which shows that $\mathrm{q}(t) = \mathrm{q}(t_0) + \mathrm{v}_0\left(\int_{t_0}^{t}\beta\right)$, since the right hand side solves the equation and has the same initial value at t_0 as the left hand side. □

PROPOSITION 1.1.3. *Let* $c(t), c^*(t) : I \to \mathbb{R}^k$ *be two vector valued curves.*

(1) $\frac{d(c\cdot c^*)}{dt} = \frac{dc}{dt}\cdot c^* + c\cdot\frac{dc^*}{dt}$.
(2) $\frac{d}{dt}\left(\frac{1}{2}|c|^2\right) = c\cdot\dot{c}$.
(3) $\frac{d}{dt}(|c|) = \frac{c\cdot\dot{c}}{|c|}$ *as long as* $c \neq 0$.
(4) $\frac{d}{dt}\left(\frac{1}{|c|}\right) = -\frac{c\cdot\dot{c}}{|c|^3}$ *as long as* $c \neq 0$.

PROOF. (1) follows from the product rule for differentiation.
(2) follows by using (1) with $c^* = c$ and that $|c|^2 = c\cdot c$.
(3) follows from (2) by observing that we also have $\frac{d}{dt}\left(\frac{1}{2}|c|^2\right) = |c|\,\frac{d|c|}{dt}$.
(4) follows from (3) by using $\frac{d}{dt}\left(\frac{1}{|c|}\right) = -\frac{\frac{d}{dt}(|c|)}{|c|^2}$. □

REMARK 1.1.4. The proposition will be used freely throughout the text. It is important to observe that the curves c or c^* could be the velocity or acceleration of a curve q. For example, if a curve q has the property that its velocity always has unit length, then $|\mathrm{v}| = 1$ and (2) shows that $\mathrm{v}\cdot\mathrm{a} = 0$.

Another very general method for generating curves is through equations. In general, one function $F(x,y) : \mathbb{R}^2 \to \mathbb{R}$ gives a collection of planar curves via the level sets

$$F(x,y) = c.$$

The implicit function theorem guarantees us that we get a unique curve as a graph over either x or y when the gradient of F doesn't vanish. The gradient is the vector

$$\nabla F = \begin{bmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \end{bmatrix}.$$

Geometrically the gradient is perpendicular to the level sets. This means that the level sets themselves have tangents that are given by the directions

$$\begin{bmatrix} -\frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial x} \end{bmatrix}$$

as this vector is orthogonal to the gradient. This in turn offers us a different way of finding these levels as they now also appear as solutions to the differential equation

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{\partial F}{\partial y}\left(x\left(t\right), y\left(t\right)\right) \\ \frac{\partial F}{\partial x}\left(x\left(t\right), y\left(t\right)\right) \end{bmatrix}.$$

In three variables we need two functions as such functions have level sets that are surfaces:

$$\begin{aligned} F_1\left(x, y, z\right) &= c_1, \\ F_2\left(x, y, z\right) &= c_2. \end{aligned}$$

In this case we also have a differential equation approach. Both of the gradients ∇F_1 and ∇F_2 are perpendicular to their level sets. Thus the cross product $\nabla F_1 \times \nabla F_2$ is tangent to the intersection of these two surfaces and we can describe the curves as solutions to

$$\frac{d\,\mathrm{q}}{dt} = \left(\nabla F_1 \times \nabla F_2\right)\left(\mathrm{q}\right).$$

It is important to realize that when we are looking for solutions to a first order system

$$\frac{d\,\mathrm{q}}{dt} = F\left(\mathrm{q}\left(t\right)\right),$$

then we geometrically obtain the same curves if we consider

$$\frac{d\,\mathrm{q}}{dt} = \lambda\left(\mathrm{q}\left(t\right)\right) F\left(\mathrm{q}\left(t\right)\right),$$

where λ is some scalar function, as the directions of the velocities stay the same. However, the parametrizations of the curves will change.

Classically curves were given descriptively in terms of geometric or even mechanical constructions. Thus a circle is the set of points in the plane that all have a fixed distance R to a fixed center. It became more common starting with Descartes to describe them by equations. Only about 1750 did Euler switch to considering parametrized curves. It is also worth mentioning that what we call curves used to be referred to as *lines*. This terminology still appears in certain concepts we introduce later, such as lines of curvature on a surface. However, when we refer to a *line* in these notes we mean a *straight line*.

We present a few classical examples of these constructions in the plane.

Example 1.1.5. Consider the equation

$$F\left(x, y\right) = x^2 + y^2 = c.$$

When $c > 0$ this describes a circle of radius $\sqrt{c}$. When $c = 0$ we only get the origin, while when $c < 0$ there are no solutions. The gradient is given by $(2x, 2y)$ and only vanishes at the origin.

The differential equation describing the level sets is

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} -2y \\ 2x \end{bmatrix}.$$

The solutions are given by $\mathrm{q}(t) = R(\cos(2(t+\varphi)), \sin(2(t+\varphi)))$ where the constants R and φ can be adjusted according to any given initial position. A more convenient parametrization happens when we scale the system to become

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}$$

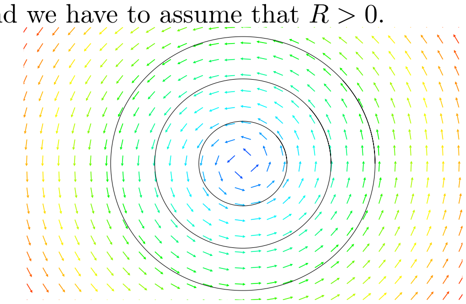
so that the solutions are $\mathrm{q}(\theta) = R(\cos(\theta+\varphi), \sin(\theta+\varphi))$ with θ being the angle to the x-axis. Yet a further scaling is possible as long as we exclude the origin

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \frac{1}{\sqrt{x^2+y^2}} \begin{bmatrix} -y \\ x \end{bmatrix}.$$

This time the solutions are given by

$$\mathrm{q}(\theta) = R\left(\cos\left(\frac{\theta+\varphi}{R}\right), \sin\left(\frac{\theta+\varphi}{R}\right)\right)$$

and we have to assume that $R > 0$.



EXAMPLE 1.1.6. Consider

$$F(x,y) = x^2 - y^2 = c.$$

When $c \neq 0$ the solution set consists of two hyperbolas. They'll be separated by the y-axis when $c > 0$ and by the x-axis when $c < 0$. When $c = 0$ the solution set consists of the two lines $y = \pm x$. A tangent direction is given by $(2y, 2x)$, which we observe only vanishes at the origin. Unlike the above example we seem to have a valid level set passing through the origin, however, it consists of two curves that pass through the point of contention.

A nicely scaled differential equation describing these curves is given by

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$$

and the solutions are given by

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ae^t + be^{-t} \\ ae^t - be^{-t} \end{bmatrix},$$

where a, b can be adjusted according to the initial values. There are five separate solutions that together give us the level set $x^2 - y^2 = 0$. We get the origin when $a = 0$, $b = 0$. The two parts of $y = x$ when $b = 0$ with the part in the first quadrant when $a > 0$ and in the third quadrant when $a < 0$. The two parts of $y = -x$ similarly come from $a = 0$.

Example 1.1.7. Consider the second order equation

$$\frac{d^2\,\mathrm{q}}{dt^2} = 0.$$

The solutions are straight lines $\mathrm{q}\,(t) = \mathrm{q}_0 + \mathrm{v}_0\,(t - t_0)$.

The next two examples show that scaling second order equations can, in contrast to first order equations, change the solutions drastically.

Example 1.1.8. The first example is given by the harmonic oscillator

$$\frac{d^2\,\mathrm{q}}{dt^2} = -\,\mathrm{q}\,.$$

This is easy to solve if we look at each coordinate separately. The solutions are:

$$\mathrm{q}\,(t) = \mathrm{q}_0 \cos\,(t - t_0) + \mathrm{v}_0 \sin\,(t - t_0)\,.$$

Example 1.1.9. A more subtle differential equation comes from Newton's inverse square law:

$$\frac{d^2\,\mathrm{q}}{dt^2} = -g\frac{\mathrm{q}}{|\mathrm{q}|^3} = -g\frac{1}{|\mathrm{q}|^2}\frac{\mathrm{q}}{|\mathrm{q}|}.$$

The solutions are conic sections. This is discussed in section 1.4 and the conic sections are defined in the next example.

Example 1.1.10 (Conic Sections). A *conic section* is the curve that results from intersecting a cone with a plane. It can be a point, two lines, circles, ellipses, parabolas, or hyperbolas. Shining a flash light at a wall at different angles will yield contour shapes that are conic sections. A point or two lines only occur when the plane goes through the vertex of the cone. In all other cases we obtain the non-degenerate conic sections that are ellipses, parabolas, or hyperbolas.

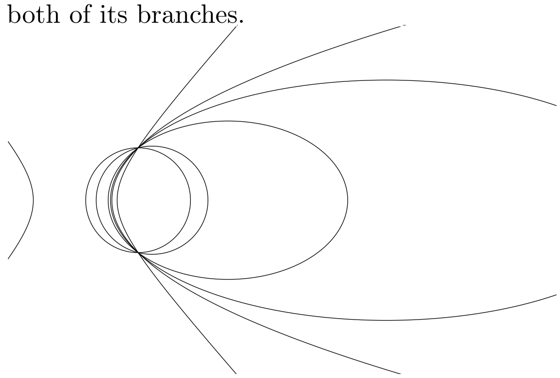
Here we offer another classical definition that is strictly planar. A conic section is determined by a focal point $\mathrm{f} \in \mathbb{R}^2$, a line l, and an eccentricity $e \geq 0$. The curve is defined as the points q whose distance to f is e times the distance to l. When $\mathrm{f} = 0$ is at the origin the curve is given by the equation

$$|\mathrm{q}| = e\,(\mathrm{q}\cdot\mathrm{n} + q)\,,$$

where n is a unit normal to l and $\mathrm{q}\cdot\mathrm{n} + q$ measures the distance from q to the line l. We can rewrite this as

$$|\mathrm{q}| = \mathrm{q}\cdot\mathrm{k} + p.$$

Below is a picture of several conic sections with the same p. Note that the hyperbola has both of its branches.



To convince ourselves that this really yields conic sections we further assume that the coordinate axes are rotated so that k $= (e, 0)$. The line is then given by $x = -\frac{p}{e}$. The equation in Cartesian coordinates now becomes

$$\sqrt{x^2 + y^2} = ex + p.$$

This can be rewritten as

$$\left(1 - e^2\right) x^2 - 2epx + y^2 = p^2.$$

When $e = 1$ this gives a sideways parabola. When $e \neq 1$ we can further rewrite it as

$$\frac{\left(x - \frac{ep}{1-e^2}\right)^2}{\left(\frac{p}{1-e^2}\right)^2} + \frac{y^2}{\frac{p^2}{1-e^2}} = 1.$$

When $e = 0$, this is the equation for a circle centered at the origin with radius p. When $0 < e < 1$ it becomes an ellipse with major axis $a = \frac{p}{1-e^2}$, minor axis $b = \frac{p}{\sqrt{1-e^2}}$, and center $\left(\frac{ep}{1-e^2}, 0\right)$. Finally when $e > 1$ it is a hyperbola as $\frac{p^2}{1-e^2} < 0$.

In polar coordinates the equation takes the simple form

$$r = er\cos\theta + p$$

or

$$r\left(1 - e\cos\theta\right) = p.$$

Example 1.1.11. Finally we mention a less well known ancient example. This is the *conchoid (shell-like) of Nicomedes*. It is given by a quartic (degree 4) equation:

$$\left(x^2 + y^2\right)(y - b)^2 - R^2y^2 = 0.$$

Descriptively it consists of two curves that are given as points (x, y) whose distance along radial lines to the line $y = b$ is R. The radial line is simply the line that passes through the origin and (x, y). So we are measuring the distance from (x, y) to the intersection of this radial line with the line $y = b$. As that intersection is $\left(\frac{x}{y}b, b\right)$ the condition is

$$\left(x - \frac{x}{y}b\right)^2 + (y - b)^2 = R^2,$$

which after multiplying both sides by y^2 reduces to the above equation.

The two parts of the curve correspond to points that are either above or below $y = b$. Note that no point on $y = b$ solves the equation as long as $b \neq 0$.

A simpler formula appears if we use polar coordinates. The line $y = b$ is described as

$$(x, y) = (b\cot\theta, b) = \frac{b}{\sin\theta}(\cos\theta, \sin\theta)$$

and the point (x, y) by

$$(x, y) = \left(\frac{b}{\sin\theta} \pm R\right)(\cos\theta, \sin\theta).$$

This gives us a natural parametrization of these curves.

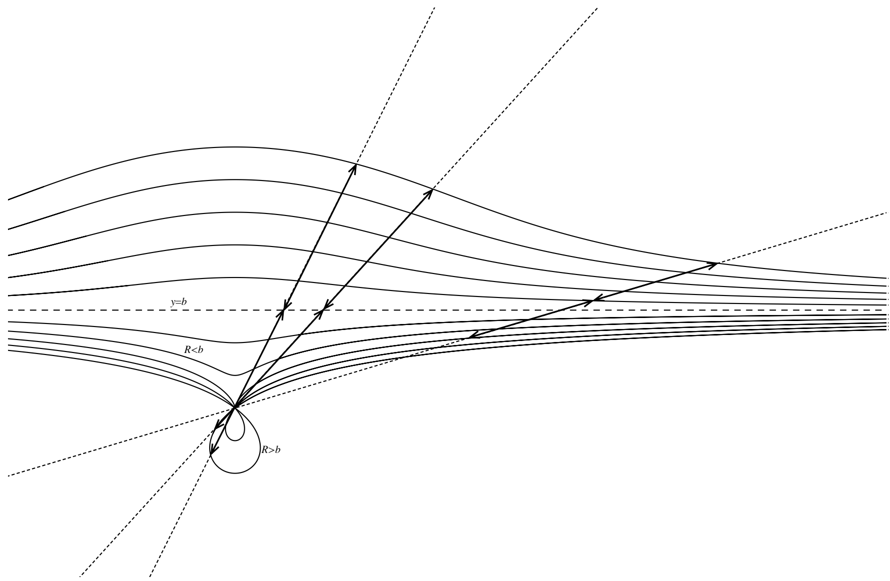


FIGURE 1.1.1. Conchoids

Another parametrization is obtained if we intersect the curve with the lines $y = tx$ and use the slope t instead of the angle θ as the parameter. This corresponds to $t = \tan\theta$ in polar coordinates. Thus we obtain the parameterized form

$$(x, y) = \left(\frac{b}{t} \pm \frac{R}{\sqrt{1+t^2}}\right)(1, t).$$

As we have seen, what we consider the same curve might have several different parametrizations.

There is also a way of characterizing curves that are radial lines. We offer two proofs that highlight some of the characterizations of curves that we have seen above.

PROPOSITION 1.1.12. *If the velocity is always radial relative to a point* c*, then the curve lies on a line through* c*.*

PROOF. The condition tells us that the curve satisfies a differential equation of the form

$$\mathrm{v}(t) = \alpha(t)(\mathrm{q}(t) - \mathrm{c}).$$

A solution to the this equation with $\mathrm{q}(t_0) = \mathrm{q}_0$ is given by

$$\mathrm{q}(t) = (\mathrm{q}_0 - \mathrm{c}) \exp\left(\int_{t_0}^{t} \alpha\right) + \mathrm{c}.$$

By uniqueness of solutions this is also the only such solution.

Alternately we can characterize radial curves as curves where

$$\frac{\mathrm{q}(t) - \mathrm{c}}{|\mathrm{q}(t) - \mathrm{c}|}$$

is constant. Note that this can be written as an equation. When we assume $\mathrm{v}(t) = \alpha(t)(\mathrm{q}(t) - \mathrm{c})$ the derivative of this ratio is

$$\begin{aligned}\frac{d}{dt}\frac{\mathrm{q}(t)-\mathrm{c}}{|\mathrm{q}(t)-\mathrm{c}|} &= \frac{\mathrm{v}(t)}{|\mathrm{q}(t)-\mathrm{c}|} - (\mathrm{q}(t)-\mathrm{c})\frac{(\mathrm{q}(t)-\mathrm{c})\cdot\mathrm{v}(t)}{|\mathrm{q}(t)-\mathrm{c}|^3} \\ &= \alpha(t)\frac{\mathrm{q}(t)-\mathrm{c}}{|\mathrm{q}(t)-\mathrm{c}|} - (\mathrm{q}(t)-\mathrm{c})\frac{\alpha(t)}{|\mathrm{q}(t)-\mathrm{c}|} \\ &= 0.\end{aligned}$$

Thus

$$\frac{\mathrm{q}(t)-\mathrm{c}}{|\mathrm{q}(t)-\mathrm{c}|}$$

is constant when the velocity is radial. □

Definition 1.1.13. Two parametrized curves $\mathrm{q}(t)$ and $\mathrm{q}^*(t^*)$ are *reparametrizations* of each other if it is possible to write $t = t(t^*)$ as a function of t^* and $t^* = t^*(t)$ such that

$$\mathrm{q}(t) = \mathrm{q}^*(t^*(t)) \text{ and } \mathrm{q}(t(t^*)) = \mathrm{q}^*(t^*).$$

If both of the functions $t(t^*)$ and $t^*(t)$ are differentiable, then it follows from the chain rule that

$$\frac{dt}{dt^*}\frac{dt^*}{dt} = 1.$$

In particular, these derivatives never vanish and have the same sign. We shall almost exclusively consider such reparametrizations. In fact we shall usually assume that these derivatives are positive so that the the direction of the curve is preserved under the reparametrization.

Lemma 1.1.14. *If $\mathrm{q}^*(t^*) = \mathrm{q}(t(t^*))$ and $t(t^*)$ is differentiable with positive derivative, then q^* is a reparametrization of q.*

Proof. The missing piece in the definition of reparametrization is to show that we can also write t^* as a differentiable function of t. However, by assumption $\frac{dt}{dt^*} > 0$ so the function $t(t^*)$ is strictly increasing. This means that for a given value of the function there is at most one point in the domain yielding this value (horizontal line test). This shows that we can find the inverse function $t^*(t)$. Graphically, simply take the graph of $t(t^*)$ and consider its mirror image reflected in the diagonal line $t = t^*$. This function is also differentiable with derivative at $t = t_0$ is given by

$$\frac{1}{\frac{dt}{dt^*}(t^*(t_0))}.$$

□

It is generally too cumbersome to use two names for curves that are reparametrizations of each other. Thus we shall simply write $\mathrm{q}(t^*)$ for a reparametrization of $\mathrm{q}(t)$ with the meaning being that

$$\mathrm{q}(t) = \mathrm{q}(t^*(t)) \text{ and } \mathrm{q}(t(t^*)) = \mathrm{q}(t^*).$$

With that in mind we shall always think of two curves as being the same if they are reparametrizations of each other.

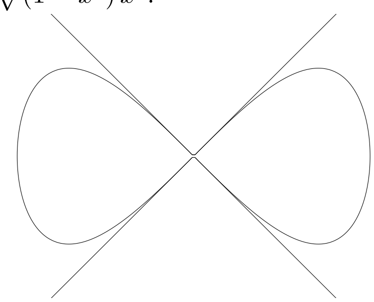
Definition 1.1.15. We say that a curve $\mathrm{q} : I \to \mathbb{R}^k$ is *closed* if there is an interval $[a,b] \subset I$ such that $\mathrm{q}(a) = \mathrm{q}(b)$ and $\mathrm{q}(I) = \mathrm{q}([a,b])$. We say that a closed curve is *simple* if it is regular and $[a,b]$ can be chosen so that $\mathrm{q} : [a,b) \to \mathbb{R}^k$ is one-to-one.

Example 1.1.16. A circle $(\cos t, \sin t)$ is a simple closed curve where we can use the interval $[a, 2\pi + a]$ for any a.

Example 1.1.17. The figure "∞" is an example of a curve that is closed, but not simple. It can be described by an equation

$$\left(1 - x^2\right) x^2 = y^2.$$

Note that as the right hand side is non-negative it follows that $x^2 \leq 1$. When $x = -1, 0, 1$ we get that $y = 0$. For other values of x there are two possibilities for $y = \pm\sqrt{(1 - x^2)\, x^2}$.



Exercises

(1) Show that if a curve q satisfies $|\mathrm{q}(t)| = R$ for all t and a constant R, then $\mathrm{q} \cdot \mathrm{v} = 0$.
(2) Show that the following properties for a regular curve are equivalent.
   (a) The curve is part of a straight line
   (b) All its tangent lines are parallel.
   (c) All its tangent lines pass through a fixed point c.
(3) Show that lines in the plane satisfy equations of the form $r\cos(\theta - \theta_0) = r_0$ in polar coordinates. Describe what the two constants θ_0, r_0 mean.
(4) Show that three points that don't lie on a line determine a unique conic section with focus at the origin.
(5) Show that a curve $\mathrm{q}(t) : I \to \mathbb{R}^2$ lies on a line if and only if there is a vector $\mathrm{n} \in \mathbb{R}^2$ such that $\mathrm{q}(t) \cdot \mathrm{n}$ is constant.
(6) Show that for a curve $\mathrm{q}(t) : I \to \mathbb{R}^3$ the following properties are equivalent:
   (a) The curve lies in a plane.
   (b) There is a vector $\mathrm{n} \in \mathbb{R}^3$ such that $\mathrm{q}(t) \cdot \mathrm{n}$ is constant.
   (c) There is a vector $\mathrm{n} \in \mathbb{R}^3$ such that $\mathrm{v}(t) \cdot \mathrm{n} = 0$ for all t.
(7) Show that a curve $\mathrm{q}(t) : I \to \mathbb{R}^3$ lies on a line if and only if there are two linearly independent vectors $\mathrm{n}_1, \mathrm{n}_2 \in \mathbb{R}^3$ such that $\mathrm{q}(t) \cdot \mathrm{n}_1$ and $\mathrm{q}(t) \cdot \mathrm{n}_2$ are constant.
(8) Show that if a curve $\mathrm{q}(t) : I \to \mathbb{R}^3$ satisfies $\dddot{\mathrm{q}} = 0$ on I, then it lies in a plane.
(9) Show that for a curve $\mathrm{q}(t) : I \to \mathbb{R}^n$ the following properties are equivalent.
   (a) The curve lies on a circle ($n = 2$) or sphere ($n > 2$.)

(b) There is a vector c such that $|\mathrm{q}-\mathrm{c}|$ is constant.
(c) There is a vector c such that $(\mathrm{q}-\mathrm{c})\cdot\mathrm{v}=0$.

(10) Consider a curve $\mathrm{q}(t): I \to \mathbb{R}^n$ and fix $t_0 \in I$. Show that the curve lies on a circle ($n=2$) or sphere ($n>2$) if and only if the curve

$$\mathrm{q}^*(t) = \frac{\mathrm{q}(t)-\mathrm{q}(t_0)}{|\mathrm{q}(t)-\mathrm{q}(t_0)|^2}$$

lies on a line ($n=2$) or hyperplane ($n>2$). Hint. The hyperplane is given by the points x that satisfy:

$$(\mathrm{q}(t_0)-\mathrm{c})\cdot\mathrm{x} = -\frac{1}{2},$$

where c is the center of the sphere.

(11) Consider a curve of the form $\mathrm{q}(\theta) = r(\theta)(\cos\theta, \sin\theta)$ where r is a function of both $\cos\theta$ and $\sin\theta$

$$r(\theta) = p(\cos\theta, \sin\theta).$$

(a) Show that this curve is closed.
(b) Show that if $r(\theta) > 0$, then it is a regular and simple curve.
(c) Let $0 \le \theta_1 < \theta_2 < 2\pi$ and $\theta_2 \ne \pi + \theta_1$. Show that if $r(\theta_1) = r(\theta_2) = 0$, $\dot r(\theta_1) \ne 0 \ne \dot r(\theta_2)$, then it is not simple. In case $\theta_2 = \pi + \theta_1$ the curve is not simple as long as $\dot r(\theta_1) \ne -\dot r(\theta_2)$.
(d) Show that if $r(\theta_0) = \dot r(\theta_0) = 0$, then its velocity vanishes at θ_0.
(e) By adjusting a in $r(\theta) = 1 + a\cos\theta$ give examples of curves that satisfy the conditions in (b), (c), and (d).

(12) Consider a curve of the form $\mathrm{q}(t) = x(t)(1,t)$.
(a) Show that $\mathrm{v} = (\dot x, x + t\dot x)$.
(b) Show that if $x(t_0) = \dot x(t_0) = 0$, then its velocity vanishes at t_0.
(c) By adjusting a in

$$x(t) = \frac{a+t^2}{1+t^2}$$

give examples of curves that are not regular.

(13) Consider a curve in $\mathbb{R}^2$ whose velocity never vanishes and intersects all the radial lines from the origin at a constant angle θ_0. These are also called *loxodromes*. Determine what this curve must be if $\theta_0 = 0$ or $\frac{\pi}{2}$. Show that *logarithmic spirals*

$$\mathrm{q}(t) = ae^{bt}(\cos t, \sin t)$$

have this property.

(14) Show that if we parametrize the sphere

$$x^2 + y^2 + z^2 = R^2$$

by

$$\begin{aligned} x &= R\sin\phi\cos\theta, \\ y &= R\sin\phi\sin\theta, \\ z &= R\cos\phi, \end{aligned}$$

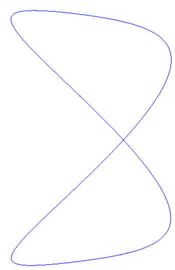
then great circles satisfy $\tan\phi\cos(\theta-\theta_0) = \tan\phi_0$. A great circle is the intersection of the sphere with a plane $ax + by + cz = 0$ through the origin.

(15) Show that the two equations

$$\begin{aligned} x^2 + y^2 + z^2 &= 4R^2, \\ (x - R)^2 + y^2 &= R^2 \end{aligned}$$

define a closed space curve that intersects itself at $x = 2R$ by showing that it can be parametrized as

$$\mathrm{q}(t) = R\left(\cos(t) + 1, \sin(t), 2\sin\left(\frac{t}{2}\right)\right).$$



(16) The *cissoid (ivy-like) of Diocles* is given by the equation

$$x\left(x^2 + y^2\right) = 2Ry^2.$$

(a) Show that this can always be parametrized by y, but that this parametrization is not smooth at $y = 0$. Hint: A cubic equation $ax^3 + bx^2 + cx + d = 0$ has a unique root if the derivative of the left hand side is positive.
(b) Show that if $y = tx$, then we obtain a parametrization

$$(x, y) = \frac{2Rt^2}{1 + t^2}(1, t).$$

(c) Show that in polar coordinates

$$r = 2R\left(\frac{1}{\cos\theta} - \cos\theta\right).$$

(17) The *folium (leaf) of Descartes* is given by the equation

$$x^3 + y^3 - 3Rxy = 0.$$

In this case the curve really does describe a leaf in the first quadrant.
(a) Show that it can not be parameterized by x or y near the origin.
(b) Show that if $y = tx$, then we obtain a parametrization

$$(x, y) = \frac{3Rt}{1 + t^3}(1, t)$$

that is valid for $t \neq -1$. What happens when $t = -1$?
(c) Show that in polar coordinates we have

$$r = \frac{3R\sin\theta\cos\theta}{\sin^3\theta + \cos^3\theta}.$$

(18) Given two planar curves q_1 and q_2 we can construct a *cissoid* q as follows: Assume that the line $y = tx$ intersects the curves in $\mathrm{q}_1 = (x_1(t), tx_1(t))$ and $\mathrm{q}_2 = (x_2(t), tx_2(t))$, then define $\mathrm{q}(t) = x(t)(1, t)$ so that $|\mathrm{q}(t)| = |\mathrm{q}_1(t) - \mathrm{q}_2(t)|$.
(a) Show that $x(t) = \pm(x_1(t) - x_2(t))$.

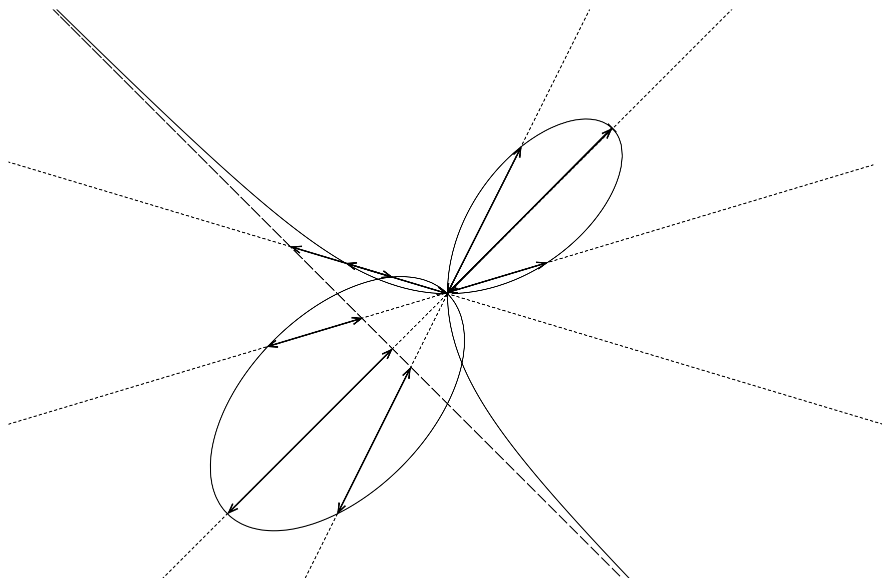
(b) Show that the conchoid of Nicomedes is a cissoid. Hint: q_1 is a circle of radius R centered at the origin and q_2 the line $y = b$. However, the parametrization of the circle is so that it is its lower half that gives the upper part of the conchoid.

(c) Show that the folium of Descartes is a cissoid. Hint: Use an ellipse

$$x^2 - xy + y^2 = -R\,(x + y)$$

and line

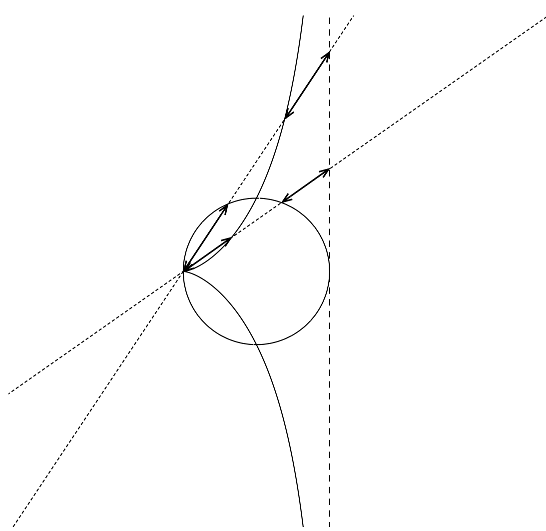
$$x + y = -R.$$



(19) Let q be a cissoid where q_1 is the circle of radius R centered at $(R, 0)$ and q_2 a vertical line $x = b$.

(a) Show that when $b = 2R$ we obtain the cissoid of Diocles

$$x\left(x^2 + y^2\right) = 2Ry^2.$$



(b) Show that when $b = \frac{R}{2}$ we obtain the *trisectrix (trisector) of Maclaurin*

$$2x\left(x^2+y^2\right) = -R\left(3x^2-y^2\right).$$

(c) Show that when $b = R$ we obtain a *strophoid*

$$y^2\left(R-x\right) = x^2\left(x+R\right).$$

(d) Show that the change of coordinates $x = u+v$, $y = \sqrt{3}\left(u-v\right)$ turns the trisectrix of Maclaurin into Descartes' folium.

1.2. Arclength and Linear Motion

The *arclength* is the distance traveled along the curve. One way of measuring the arclength geometrically is by imagining the curve as a thread that can be stretched out and measured. This, however, doesn't really help in formulating how it should be measured mathematically. Archimedes succeeded in understanding the arclength of circles by relating it to the area of the circle. The idea of measuring the length of general curves is relatively recent, going back only to about 1600. Newton was the first to give the general definition that we shall use below. As we shall quickly discover, it is generally impossible to calculate the arclength of a curve as it involves finding anti-derivatives of fairly complicated functions.

From a dynamical perspective the change in arclength measures how fast the motion is along the curve. So if there is no change in arclength, then the curve is stationary, i.e., you stopped. More precisely, if the distance traveled is denoted by s (we can't use d for distance as it is used for differentiation), then the relative change with respect to the general parameter is the speed

$$\frac{ds}{dt} = \left|\frac{d\,\mathrm{q}}{dt}\right| = \left|\mathrm{v}\right|.$$

This means that s is the anti-derivative of speed and is defined up to an additive constant. The constant is determined by where we start measuring from. This means that we should define the *length of a curve* on $[a,b]$ as follows

$$L\left(\mathrm{q}\right)_a^b = \int_a^b \left|\mathrm{v}\right| dt = s\left(b\right) - s\left(a\right).$$

Using substitution this is easily shown to be independent of the parameter t as long as the reparametrization is in the same direction. One also easily checks that a curve on $[a,b]$ is stationary if and only if its speed vanishes on $[a,b]$. We usually suppress the interval and instead simply write $L\left(\mathrm{q}\right)$.

Example 1.2.1. If $\mathrm{q}\left(t\right) = \mathrm{q}_0 + \mathrm{v}_0\, t$ is a straight line, then its speed is constant $\left|\mathrm{v}_0\right|$ and so the arclength over an interval $[a,b]$ is $\left|\mathrm{v}_0\right|\left(b-a\right)$.

Example 1.2.2. If $\mathrm{q}\left(t\right) = R\left(\cos t, \sin t\right) + \mathrm{c}$ is a circle of radius R centered at c, then the speed is the constant R and so again it becomes easy to calculate the arclength.

Example 1.2.3. Consider the hyperbola $x^2 - y^2 = 1$. It consists of two components separated by the y-axis. The component with $x > 0$ can be parametrized using hyperbolic functions $\mathrm{q}\left(t\right) = \left(\cosh t, \sinh t\right)$. The speed is

$$\frac{ds}{dt} = \sqrt{\sinh^2 t + \cosh^2 t} = \sqrt{2\sinh^2 t + 1} = \sqrt{\cosh 2t}.$$

While this is both a fairly simple curve and a not terribly difficult expression for the speed it does not appear in any way easy to find the arclength explicitly.

PROPOSITION 1.2.4. *If $F : \mathbb{R}^k \to \mathbb{R}^k$ is of the form $F(\mathrm{x}) = O\mathrm{x} + \mathrm{c}$, where O is an orthogonal transformation and $\mathrm{c} \in \mathbb{R}^k$, then $L(\mathrm{q}) = L(F \circ \mathrm{q})$ for all curves.*

PROOF. An orthogonal transformation is by definition a linear map or matrix that preserves dot products: $(Ov) \cdot (Ow) = v \cdot w$, for all $v, w \in \mathbb{R}^k$. The proof follows from the simple observation that the differential of F is given by $DF = O$. Using the chain rule

$$\left|\frac{dF \circ \mathrm{q}}{dt}\right| = |DF(\dot{\mathrm{q}})| = |O(\dot{\mathrm{q}})| = |\dot{\mathrm{q}}|$$

it follows that F preserves the speed of q and hence also the length. □

DEFINITION 1.2.5. A curve is said to be *parametrized by arclength* if its speed is always 1. Such a parametrization is also called a *unit speed parametrization.*

LEMMA 1.2.6. *A regular curve $\mathrm{q}(t)$ can be reparametrized by arclength.*

PROOF. If we have a unit speed reparametrization $\mathrm{q}(s)$ of $\mathrm{q}(t)$ with $\frac{ds}{dt} > 0$, then

$$\frac{d\,\mathrm{q}}{ds}\frac{ds}{dt} = \frac{d\,\mathrm{q}}{dt} = \mathrm{v}$$

so it follows that

$$\left|\frac{d\,\mathrm{q}}{ds}\right|\frac{ds}{dt} = \frac{ds}{dt} = \left|\frac{d\,\mathrm{q}}{dt}\right| = |\mathrm{v}|$$

must be the speed of $\mathrm{q}(t)$.

This tells us that we should define the reparametrization $s = s(t)$ as the anti-derivative of the speed:

$$s(t_1) = s(t_0) + \int_{t_0}^{t_1}\left|\frac{d\,\mathrm{q}}{dt}\right| dt.$$

It then follows that

$$\frac{ds}{dt} = \left|\frac{d\,\mathrm{q}}{dt}\right| > 0.$$

Thus it is also possible to find the inverse relationship $t = t(s)$ and we can define the reparametrized curve as $\mathrm{q}(s) = \mathrm{q}(s(t)) = \mathrm{q}(t)$.

This reparametrization depends on specifying an initial value $s(t_0)$ at some specific parameter t_0. For simplicity one often uses $s(0) = 0$ if that is at all reasonable. □

To see that arclength really is related to our usual concept of distance we show:

THEOREM 1.2.7. *The straight line is the shortest curve between any two points in Euclidean space.*

PROOF. We shall give two almost identical proofs. Without loss of generality assume that we have a curve $\mathrm{q}(t) : [a, b] \to \mathbb{R}^k$ where $\mathrm{q}(a) = 0$ and $\mathrm{q}(b) = p$. We wish to show that $L(\mathrm{q}) \geq |p|$. To that end select a unit vector field X which is also a gradient field $X = \nabla f$. Two natural choices are possible: For the first, simply let $f(x) = x \cdot \frac{p}{|p|}$, and for the second $f(x) = |x|$. In the first case the gradient is simply a parallel field and defined everywhere, in the second case we obtain the radial field which is not defined at the origin. When using the second field we need to restrict the domain of the curve to $[a_0, b]$ such that $\mathrm{q}(a_0) = 0$ but $\mathrm{q}(t) \neq 0$ for $t > a_0$. This is clearly possible as the set of points where $\mathrm{q}(t) = 0$ is a closed subset of $[a, b]$, so a_0 is just the maximum value where q vanishes.

This allows us to perform the following calculation using Cauchy-Schwarz, the chain rule, and the fundamental theorem of calculus. When we are in the second case the integrals are possibly improper at $t = a_0$, but clearly turn out to be perfectly well defined since the integrand has a continuous limit as t approaches a_0

$$\begin{aligned}
L(\mathrm{q}) &= \int_a^b |\mathrm{v}|\, dt \\
&\geq \int_{a_0}^b |\mathrm{v}|\, dt \\
&= \int_{a_0}^b |\dot{\mathrm{q}}|\, |\nabla f|\, dt \\
&\geq \int_{a_0}^b |\dot{\mathrm{q}} \cdot \nabla f|\, dt \\
&= \int_{a_0}^b \left| \frac{d(f \circ \mathrm{q})}{dt} \right| dt \\
&\geq \left| \int_{a_0}^b \frac{d(f \circ \mathrm{q})}{dt} dt \right| \\
&= |f(\mathrm{q}(b)) - f(\mathrm{q}(a_0))| \\
&= |f(p) - f(0)| \\
&= |f(p)| \\
&= |p|.
\end{aligned}$$

We can even go backwards and check what happens when $L(\mathrm{q}) = |p|$. It appears that we must have equality in the places where we had inequality. Thus we have $\frac{d(f\circ \mathrm{q})}{dt} \geq 0$ everywhere and $\dot{\mathrm{q}}$ is proportional to ∇f everywhere. This implies that q is a possibly singular reparametrization of the straight line from 0 to p. □

Corollary 1.2.8 (The Triangle Inequality). *If $p, q, r \in \mathbb{R}^k$, then $|p - q| \leq |p - r| + |r - q|$ with equality holding only when the three points lie on a line.*

Proof. Simply think of the right hand side as the length of the two line segments from p to r and r to q. □

Proposition 1.2.9. *The shortest distance from a point to a curve (if it exists) is realized by a line segment that is perpendicular to the curve.*

Proof. Let $\mathrm{q} : [a, b] \to \mathbb{R}^k$ be a curve and assume that there is a $t_0 \in (a, b)$ such that

$$|\mathrm{q}(t) - p| \geq |\mathrm{q}(t_0) - p| \text{ for all } t \in [a, b].$$

This implies that

$$\frac{1}{2} |\mathrm{q}(t) - p|^2 \geq \frac{1}{2} |\mathrm{q}(t_0) - p|^2.$$

As the left hand side reaches a minimum at an interior point its derivative must vanish at t_0, i.e.,

$$(\mathrm{q}(t_0) - p) \cdot \frac{d\,\mathrm{q}}{dt}(t_0) = 0.$$

As the vector $\mathrm{q}(t_0) - p$ represents the segment from p to $\mathrm{q}(t_0)$ we have shown that it is perpendicular to the velocity of the curve. □

The next result is another interesting geometric consequence of the above theorem.

LEMMA 1.2.10. *If $\mathrm{q} : [a, b] \to \mathbb{R}^k$ is a curve of length $2R$, then it is contained in a closed ball of radius R.*

PROOF. Consider the midpoint $\mathrm{c} = \frac{1}{2}(\mathrm{q}(a) + \mathrm{q}(b))$ on the segment between $\mathrm{q}(a)$ and $\mathrm{q}(b)$.

$$|\mathrm{q}(t) - \mathrm{c}| = \frac{1}{2}|\mathrm{q}(t) - \mathrm{q}(a) + \mathrm{q}(t) - \mathrm{q}(b)| \leq \frac{1}{2}(|\mathrm{q}(t) - \mathrm{q}(a)| + |\mathrm{q}(t) - \mathrm{q}(b)|) \leq \frac{1}{2}2R = R.$$

□

With just a little more effort one can also find the shortest curves on spheres.

THEOREM 1.2.11. *The shortest curve between two points on a round sphere $S^2(R) = \left\{q \in \mathbb{R}^3 \mid |q|^2 = R^2\right\}$ is the shortest segment of the great circle through the two points.*

PROOF. Great circles on spheres centered at the origin are given as the intersections of the sphere with 2-dimensional planes through the origin. Note that if two points are antipodal then there are infinitely many great circles passing through them and all of the corresponding segments have length πR. If the two points are not antipodal, then there is a unique great circle between them and the shortest arc on this circle joining the points has length $< \pi R$.

Let us assume for simplicity that $R = 1$. The great circle that lies in the plane $\operatorname{span}\{\mathrm{q}_0, \mathrm{v}_0\}$ where $\mathrm{q}_0 \perp \mathrm{v}_0$ and $|\mathrm{q}_0| = |\mathrm{v}_0| = 1$ can be parametrized as follows

$$\mathrm{q}(t) = \mathrm{q}_0 \cos t + \mathrm{v}_0 \sin t.$$

This curve passes through the point $\mathrm{q}_0 \in S^2(1)$ at $t = 0$ and has velocity v_0 at that point. It also passes through the antipodal point $-\mathrm{q}_0$ at time $t = \pi$. Finally, it is also parametrized by arclength.

To find the great circle that passes through two points $\mathrm{q}_0, \mathrm{q}_1 \in S^2(1)$ that are not antipodal we simply select the initial velocity v_0 to be the vector in the plane $\operatorname{span}\{\mathrm{q}_0, \mathrm{q}_1\}$ that is perpendicular to q_0 and has length 1, i.e.,

$$\begin{aligned} \mathrm{v}_0 &= \frac{\mathrm{q}_1 - (\mathrm{q}_1 \cdot \mathrm{q}_0)\,\mathrm{q}_0}{|\mathrm{q}_1 - (\mathrm{q}_1 \cdot \mathrm{q}_0)\,\mathrm{q}_0|} \\ &= \frac{\mathrm{q}_1 - (\mathrm{q}_1 \cdot \mathrm{q}_0)\,\mathrm{q}_0}{\sqrt{1 - (\mathrm{q}_1 \cdot \mathrm{q}_0)^2}}. \end{aligned}$$

Then the great circle

$$\mathrm{q}(t) = \mathrm{q}_0 \cos t + \mathrm{v}_0 \sin t$$

passes through q_1 when

$$t = \arccos(\mathrm{q}_1 \cdot \mathrm{q}_0).$$

The velocity of this great circle at q_1 is

$$\mathrm{v}_1 = \frac{-\mathrm{q}_0 + (\mathrm{q}_0 \cdot \mathrm{q}_1)\,\mathrm{q}_1}{|-\mathrm{q}_0 + (\mathrm{q}_0 \cdot \mathrm{q}_1)\,\mathrm{q}_1|}$$

since it is the initial velocity of the great circle that starts at q_1 and goes through $-\mathrm{q}_0$.

The goal now is to show that any curve $\mathrm{q}(t) : [0, L] \to S^2(1)$ between q_0 and q_1 has length $\geq \arccos(\mathrm{q}_1 \cdot \mathrm{q}_0)$. The proof of this follows the same pattern as the proof for lines. We start by assuming that $\mathrm{q}(t) \neq \mathrm{q}_0, \mathrm{q}_1$ when $t \in (0, L)$ and define

$$\mathrm{v}_1(t) = \frac{-\mathrm{q}_0 + (\mathrm{q}_0 \cdot \mathrm{q}(t))\,\mathrm{q}(t)}{|-\mathrm{q}_0 + (\mathrm{q}_0 \cdot \mathrm{q}(t))\,\mathrm{q}(t)|}.$$

Before the calculation note that since $|\mathrm{q}(t)|^2 = 1$ it follows that $\mathrm{q} \cdot \frac{d\,\mathrm{q}}{dt} = 0$. With that in mind we obtain

$$\begin{aligned}
L(\mathrm{q}) &= \int_0^L |\mathrm{v}|\,dt \\
&= \int_0^L |\mathrm{v}_1(t)|\,|\mathrm{v}|\,dt \\
&\geq \int_0^L \left|\mathrm{v}_1(t) \cdot \frac{d\,\mathrm{q}}{dt}\right| dt \\
&= \int_0^L \left|\frac{-\mathrm{q}_0 + (\mathrm{q}_0 \cdot \mathrm{q}(t))\,\mathrm{q}(t)}{|-\mathrm{q}_0 + (\mathrm{q}_0 \cdot \mathrm{q}(t))\,\mathrm{q}(t)|} \cdot \frac{d\,\mathrm{q}}{dt}\right| dt \\
&= \int_0^L \left|\frac{-\mathrm{q}_0 \cdot \frac{d\,\mathrm{q}}{dt}}{\sqrt{1 - (\mathrm{q}_0 \cdot \mathrm{q}(t))^2}}\right| dt \\
&= \int_0^L \left|\frac{d \arccos(\mathrm{q}_0 \cdot \mathrm{q}(t))}{dt}\right| dt \\
&\geq \left|\int_0^L \frac{d \arccos(\mathrm{q}_0 \cdot \mathrm{q}(t))}{dt} dt\right| \\
&= |\arccos(\mathrm{q}_0 \cdot \mathrm{q}(L)) - \arccos(\mathrm{q}_0 \cdot \mathrm{q}(0))| \\
&= |\arccos(\mathrm{q}_0 \cdot \mathrm{q}_1) - \arccos(\mathrm{q}_0 \cdot \mathrm{q}_0)| \\
&= |\arccos(\mathrm{q}_0 \cdot \mathrm{q}_1)| \\
&= \arccos(\mathrm{q}_0 \cdot \mathrm{q}_1)
\end{aligned}$$

This proves that the segment of the great circle always has the shortest length.

In case the original curve is parametrized by arclength and has minimal length we can backtrack the argument and observe that this forces $\mathrm{v} = \mathrm{v}_1$ or in other words

$$\frac{d\,\mathrm{q}}{dt} = \frac{-\mathrm{q}_0 + (\mathrm{q}_0 \cdot \mathrm{q}(t))\,\mathrm{q}(t)}{|-\mathrm{q}_0 + (\mathrm{q}_0 \cdot \mathrm{q}(t))\,\mathrm{q}(t)|}.$$

This is a differential equation for the curve and we know that great circles solve this equation as the right hand side is the velocity of the great circle at $\mathrm{q}(t)$. So it follows from uniqueness of solutions to differential equations that any curve of minimal length is part of a great circle. □

REMARK 1.2.12. The *spherical distance* between two points q_0, q_1 on the unit sphere is the angle: $\angle(q_0, q_1) = \arccos(q_0 \cdot q_1) \in [0, \pi]$ between the corresponding unit vectors in Euclidean space. The previous theorem can now be restated to say that the length of a curve on the unit sphere is always greater than the spherical distance between its end points.

Exercises

(1) Consider a curve $q(t) : I \to \mathbb{R}^k$ and let $s(t)$ be an antiderivative of the speed $|v|$, i.e., the arclength parameter. It is not assumed that the curve is regular.
   (a) Show that
   $$|v(t_0)| = \lim_{t \to t_0} \frac{|q(t) - q(t_0)|}{|t - t_0|}.$$
   (b) Show that
   $$|v(t_0)| = \lim_{t \to t_0} \frac{|s(t) - s(t_0)|}{|t - t_0|}.$$
   (c) Show that if $t_n \to t_0$ and $q(t_n) \neq q(t_0)$ for all n, then
   $$1 = \lim_{n \to \infty} \frac{|s(t_n) - s(t_0)|}{|q(t_n) - q(t_0)|}.$$
   (d) Show that if $|v(t_0)| > 0$, then $q(t) \neq q(t_0)$ for t near t_0.
   (e) Assume that $|q(t)| = 1$ for all t. Show that if $q(t) \neq q(t_0)$, then
   $$1 \leq \frac{|s(t) - s(t_0)|}{\arccos(q(t) \cdot q(t_0))} \leq \frac{|s(t) - s(t_0)|}{|q(t) - q(t_0)|}.$$

(2) Compute the arclength parameter of $y = x^{\frac{3}{2}}$.
(3) Compute the arclength parameter of the parabolas $y = \sqrt{x}$ and $y = x^2$.
(4) Redefine the concept of closed and simple curves using arclength parametrization.
(5) Compute the arclength parameter of $q(t) = R(\cosh t, \sinh t, t)$.
(6) Compute the arclength of the logarithmic spiral
$$ae^{bt}(\cos t, \sin t)$$
and explain why it is called logarithmic.
(7) Show that every regular planar curve that makes a constant angle $\theta_0 > 0$ with all radial lines can be reparametrized to be a *logarithmic spiral*
$$ae^{bt}(\cos t, \sin t)$$
for suitable constants a, b. Hint: If the curve is unit speed, then the condition gives the unit tangent at each point of the curve and thus creates a differential equation for the curve.
(8) Compute the arclength parameter of the *spiral of Archimedes*:
$$(a + bt)(\cos t, \sin t).$$
(9) Find the arclength parameter for the following *twisted cubic*
$$q(t) = \left(t, 3t^2, 3t^3\right).$$
(10) Let $q : [a, b] \to \mathbb{R}^k$ be a curve of length $2R$. Show that it is either contained in a ball of radius $< R$ or is on the line passing through $q(a)$ and $q(b)$.

(11) Let $\mathrm{q}(t) : I \to \mathbb{R}^k$ be a closed piecewise smooth planar curve. Show that if $L(\mathrm{q}) \leq 4R$, then q is contained in a ball of radius R. Hint: Cut the curve into two pieces.

(12) Let $\mathrm{q}(s) : [a,b] \to S^2$ be a piecewise smooth curve.
   (a) Show that if $L = 2R \leq \pi$, then q is contained in a cap of spherical radius R, i.e., there exists $\mathrm{c} \in S^2$ such that $\arccos(\mathrm{c} \cdot \mathrm{q}(t)) \leq R$ for all t. Hint: The proof is similar to that of lemma 1.2.10 if we let c be the midpoint on the shorter part of a great circle through $\mathrm{q}(a)$ and $\mathrm{q}(b)$ and use spherical distances instead of Euclidean distances.
   (b) Show that if q is closed and $L = 4R \leq 2\pi$, then q is contained in a cap of spherical radius R.
   (c) What goes wrong with the argument when $R > \frac{\pi}{2}$?

(13) (Spherical law of cosines) Consider three points q_i, $i = 1,2,3$ on a unit sphere centered at the origin. Join these points by great circle segments to obtain a triangle. Let the side lengths be a_{ij} and the interior angle at q_i be θ_i.
   (a) Show that
$$\cos a_{ij} = \mathrm{q}_i \cdot \mathrm{q}_j$$
   and
$$\cos\theta_1 = \left(\frac{\mathrm{q}_2 - (\mathrm{q}_2 \cdot \mathrm{q}_1)\,\mathrm{q}_1}{\sqrt{1 - (\mathrm{q}_2 \cdot \mathrm{q}_1)^2}}\right) \cdot \left(\frac{\mathrm{q}_3 - (\mathrm{q}_3 \cdot \mathrm{q}_1)\,\mathrm{q}_1}{\sqrt{1 - (\mathrm{q}_3 \cdot \mathrm{q}_1)^2}}\right).$$
   (b) Show that
$$\cos a_{23} = \cos a_{12}\cos a_{13} + \sin a_{12}\sin a_{13}\cos\theta_1.$$
   (c) Show that on a sphere of radius R the law of cosines for a triangle with sides a_{ij} and interior angle θ_R at q_1 is given by
$$\cos\frac{a_{23}}{R} = \cos\frac{a_{12}}{R}\cos\frac{a_{13}}{R} + \sin\frac{a_{12}}{R}\sin\frac{a_{13}}{R}\cos\theta_R.$$
   Hint: If the triangle is radially projected to the unit sphere then its sides are $\frac{a_{ij}}{R}$ and the angles remain the same.
   (d) Show that if we fix a_{ij}, then $\theta_R \to \theta_0$ as $R \to \infty$, where θ_0 satisfies the Euclidean law of cosines
$$a_{23}^2 = a_{12}^2 + a_{13}^2 - 2a_{12}a_{13}\cos\theta_0.$$
   One can in fact show that θ_R decreases and thus that $\theta_1 > \theta_0$.

(14) The *astroid* is given by the equation
$$\left|\frac{x}{a}\right|^{\frac{2}{3}} + \left|\frac{y}{b}\right|^{\frac{2}{3}} = 1.$$
   (a) Draw a picture of this curve and show that the velocity of the curve must vanish where it intersects the axes.
   (b) Show that the coordinate axes are tangent to the curve at the points $(\pm a, 0)$ and $(0, \pm b)$, i.e. each arc of the curve that lies in a quadrant can be given a regular parametrization, where the curve is tangent to the axes at the endpoints. The curve has *cusps* at these points.
   (c) Show that when $a = b$ the arclength of the arc in the first quadrant is $\frac{3}{2}a$.
   (d) Show that when $a = b$ the line segment between the axes that is tangent to the astroid has length a.

(e) Show that the entire curve has a smooth parametrization that is regular except at the points where the curve intersects the axes. Hint: Write the equation as

$$\left(\left|\frac{x}{a}\right|^{\frac{1}{3}}\right)^2+\left(\left|\frac{y}{b}\right|^{\frac{1}{3}}\right)^2=1.$$

(15) Show that the parametrization of the folium of Descartes given by

$$(x,y)=\frac{3Rt}{1+t^3}(1,t)$$

is regular.

(16) Show that it is not possible to parametrize the cissoid of Diocles

$$x\left(x^2+y^2\right)=2Ry^2$$

so that it is regular at the origin.

(17) Consider the *tractrix* given by

$$x=\pm\int_y^R\frac{\sqrt{R^2-t^2}}{t}dt.$$

(a) Show that

$$\begin{aligned} x &= \pm\left(R\log\frac{R+\sqrt{R^2-y^2}}{y}-\sqrt{R^2-y^2}\right)\\ &= \pm\left(R\cosh^{-1}\frac{R}{y}-\sqrt{R^2-y^2}\right), \end{aligned}$$

where $\cosh^{-1}:[R,\infty)\to[0,\infty)$ is the inverse function to cosh.

(b) Show that the segment of the tangent between the curve and the x-axis always has length R.

(c) Show that the speed is $\frac{R}{y}$ when we use y as the parameter.

(d) Show that it can be parametrized as

$$R\left(\log\cot\frac{\theta}{2}-\cos\theta,\sin\theta\right).$$

(e) Show that it can also be parametrized as

$$\left(x-R\frac{\sinh\frac{x}{R}}{\cosh\frac{x}{R}},\frac{R}{\cosh\frac{x}{R}}\right).$$

(18) A *cycloid* is a planar curve that follows a point on a circle of radius R as it rolls along a straight line without slipping.

(a) Show that

$$\mathrm{q}(t)=tR\mathrm{e}_1+R\mathrm{e}_2-R(\mathrm{e}_2\cos t+\mathrm{e}_1\sin t)$$

is a parametrization of a cycloid, when $\mathrm{e}_1,\mathrm{e}_2$ are orthonormal.

(b) Show that all cycloids can be parametrized to have the form

$$\mathrm{q}(t)=tR\mathrm{e}_1+R\mathrm{e}_2-R(\mathrm{e}_2\cos t+\mathrm{e}_1\sin t)+\mathrm{q}_0$$

where $\mathrm{q}(0)=\mathrm{q}_0$.

(c) Show that any such cycloid stays on one side of the line $\mathrm{q}_0+tR\mathrm{e}_1$ and has zero velocity cusps when it hits this line.

(d) Show that a cycloid hits the line at points that are $2\pi R$ apart.

1.3. Curvature

We saw that arclength measures how far a curve is from being stationary. Our preliminary concept of *curvature* is that it should measure how far a curve is from being a line. For a planar curve the idea used to be to find a circle that best approximates the curve at a point (just like a tangent line is the line that best approximates the curve). The radius of this circle then gives a measure of how the curve bends with larger radius implying less bending. Huygens did quite a lot to clarify this idea for fairly general curves using purely geometric considerations (no calculus) and applied it to the study involutes and evolutes. Newton seems to have been the first to take the reciprocal of this radius to create curvature as we now define it. He also generated some of the formulas in both Cartesian and polar coordinates that are still in use today.

To formalize the idea of how a curve deviates from being a line we define the *unit tangent vector* of a regular curve $\mathrm{q}(t):[a,b]\to\mathbb{R}^k$ as the direction T of the velocity:

$$\mathrm{v}=\dot{\mathrm{q}}=|\mathrm{v}|\,\mathrm{T}=\frac{ds}{dt}\mathrm{T}\,.$$

When the unit tangent vector $\mathrm{T}=\mathrm{v}\,/\,|\mathrm{v}|$ is stationary, then the curve is evidently a straight line. So the degree to which the unit tangent is stationary is a measure of how fast it changes and in turn how far the curve is from being a line. We let θ be the arclength parameter for T. The relative change between the arclength parameters for the unit tangent and the curve is by definition the *curvature*

$$\kappa=\frac{d\theta}{ds}.$$

For a general parametrization we can use the chain rule to obtain the formula

$$\kappa=\frac{dt}{ds}\frac{d\theta}{dt}.$$

We shall see that the curvature is related to the part of the acceleration that is orthogonal to the unit tangent vector. Note that $\kappa\geq 0$ as θ increases with s.

PROPOSITION 1.3.1. *A regular curve is part of a line if and only if its curvature vanishes.*

PROOF. The unit tangent of a line is clearly stationary. Conversely if the curvature vanishes, then the unit tangent is stationary. This means that when the curve is parametrized by arclength, then it will be a straight line. □

PROPOSITION 1.3.2. *If $F:\mathbb{R}^k\to\mathbb{R}^k$ is of the form $F(\mathrm{x})=O\mathrm{x}+\mathrm{c}$, where O is an orthogonal transformation and $\mathrm{c}\in\mathbb{R}^k$ and q a regular curve, then $\mathrm{q}^*=F(\mathrm{q})$ has unit tangent given by $\mathrm{T}^*=O\,\mathrm{T}$ and curvature $\kappa^*=\kappa$.*

PROOF. As in proposition 1.2.4 we use that $DF=O$. The chain rule then shows that

$$\mathrm{v}^*=O\,\mathrm{v},\ \mathrm{a}^*=O\,\mathrm{a}\,.$$

This shows that

$$\mathrm{T}^*=\frac{\mathrm{v}^*}{|\mathrm{v}^*|}=\frac{O\,\mathrm{v}}{|O\,\mathrm{v}|}=\frac{O\,\mathrm{v}}{|\mathrm{v}|}=O\left(\frac{\mathrm{v}}{|\mathrm{v}|}\right)=O\,(\mathrm{T})\,.$$

We can now use proposition 1.2.4 again to see that q and q^* have the same arclength parameter. Similarly, T and T^* have the same arclength parameter. Thus $\kappa = \kappa^*$. $\square$

Next we show how the curvature can be calculated for a general parametrization using the velocity and acceleration.

PROPOSITION 1.3.3. *The curvature of a regular curve is given by*

$$\begin{aligned} \kappa &= \frac{|\mathrm{v}|\,|\mathrm{a} - (\mathrm{a}\cdot\mathrm{T})\,\mathrm{T}|}{|\mathrm{v}|^3} \\ &= \frac{\textit{area of parallelogram}\,(\mathrm{v},\mathrm{a})}{|\mathrm{v}|^3}. \end{aligned}$$

PROOF. We calculate

$$\begin{aligned} \kappa &= \frac{d\theta}{ds} \\ &= \frac{d\theta}{dt}\frac{dt}{ds} \\ &= \left|\frac{d\,\mathrm{T}}{dt}\right| |\mathrm{v}|^{-1} \\ &= \left|\frac{d}{dt}\frac{\mathrm{v}}{|\mathrm{v}|}\right| |\mathrm{v}|^{-1} \\ &= \left|\frac{\mathrm{a}}{|\mathrm{v}|} - \frac{\mathrm{v}\,(\mathrm{a}\cdot\mathrm{v})}{|\mathrm{v}|^3}\right| \frac{1}{|\mathrm{v}|} \\ &= \frac{1}{|\mathrm{v}|^2}\left|\mathrm{a} - \frac{(\mathrm{a}\cdot\mathrm{v})\,\mathrm{v}}{|\mathrm{v}|^2}\right| \\ &= \frac{1}{|\mathrm{v}|^2}\,|\mathrm{a} - (\mathrm{a}\cdot\mathrm{T})\,\mathrm{T}|\,. \end{aligned}$$

The area of the parallelogram spanned by v and a is given by the product of the length of the base represented by v and the height represented by the component of a that is normal to the base, i.e., $\mathrm{a} - (\mathrm{a}\cdot\mathrm{T})\,\mathrm{T}$. Thus we obtain the formula

$$\begin{aligned} \kappa &= \frac{|\mathrm{v}|\,|\mathrm{a} - (\mathrm{a}\cdot\mathrm{T})\,\mathrm{T}|}{|\mathrm{v}|^3} \\ &= \frac{\text{area of parallelogram}\,(\mathrm{v},\mathrm{a})}{|\mathrm{v}|^3}. \end{aligned}$$

$\square$

REMARK 1.3.4. For 3-dimensional curves the curvature can also be written as

$$\kappa = \frac{|\mathrm{v}\times\mathrm{a}|}{|\mathrm{v}|^3}.$$

Further note that when the unit tangent vector is regular it too has a unit tangent vector called the *normal* N to the curve. Specifically

$$\frac{d\,\mathrm{T}}{d\theta} = \mathrm{N}\,.$$

The unit normal is the unit tangent to the unit tangent. This vector is in fact perpendicular to T as

$$0 = \frac{d\,|\mathrm{T}|^2}{d\theta} = 2\left(\mathrm{T}\cdot\frac{d\,\mathrm{T}}{d\theta}\right) = 2\,(\mathrm{T}\cdot\mathrm{N})\,.$$

This normal vector is also called the *principal normal* for q, when the curve is a space curve, as there are also other vectors that are normal to the curve in that case. The line through a point on a curve in the direction of the principal normal is called the *principal normal line*.

In terms of the arclength parameter s for q we obtain

$$\frac{d\,\mathrm{T}}{ds} = \frac{d\theta}{ds}\frac{d\,\mathrm{T}}{d\theta} = \kappa\,\mathrm{N}$$

and

$$\kappa = \frac{d\,\mathrm{T}}{ds}\cdot\mathrm{N} = -\,\mathrm{T}\cdot\frac{d\,\mathrm{N}}{ds},$$

where the last equality follows from

$$0 = \frac{d\,\mathrm{T}\cdot\mathrm{N}}{ds} = \frac{d\,\mathrm{T}}{ds}\cdot\mathrm{N} + \mathrm{T}\cdot\frac{d\,\mathrm{N}}{ds}.$$

PROPOSITION 1.3.5. *For a regular curve we have*

$$\mathrm{v} = (\mathrm{v}\cdot\mathrm{T})\,\mathrm{T} = |\mathrm{v}|\,\mathrm{T},$$

$$\mathrm{a} = (\mathrm{a}\cdot\mathrm{T})\,\mathrm{T} + (\mathrm{a}\cdot\mathrm{N})\,\mathrm{N} = (\mathrm{a}\cdot\mathrm{T})\,\mathrm{T} + \kappa\,|\mathrm{v}|^2\,\mathrm{N},$$

and

$$\mathrm{N} = \frac{\mathrm{a} - (\mathrm{a}\cdot\mathrm{T})\,\mathrm{T}}{|\mathrm{a} - (\mathrm{a}\cdot\mathrm{T})\,\mathrm{T}|}.$$

Thus the unit normal is the direction of the part of the acceleration that is perpendicular to the velocity.

PROOF. The first formula follows directly from the definition of . For the second we note that

$$\begin{aligned}
\mathrm{a} &= \frac{d\,\mathrm{v}}{dt} \\
&= \frac{d\theta}{dt}\frac{d\,\mathrm{v}}{d\theta} \\
&= \frac{d\theta}{dt}\left(\frac{d\,|\mathrm{v}|}{d\theta}\,\mathrm{T} + |\mathrm{v}|\,\frac{d\,\mathrm{T}}{d\theta}\right) \\
&= \frac{d\theta}{dt}\left(\frac{d\,|\mathrm{v}|}{d\theta}\,\mathrm{T} + |\mathrm{v}|\,\mathrm{N}\right) \\
&= \frac{d\,|\mathrm{v}|}{dt}\,\mathrm{T} + \frac{d\theta}{dt}\,|\mathrm{v}|\,\mathrm{N}\,.
\end{aligned}$$

This shows that a is a linear combination of T, N. It also shows that

$$\begin{aligned}
\mathrm{a}\cdot\mathrm{N} &= \frac{d\theta}{dt}\,|\mathrm{v}| \\
&= \frac{d\theta}{ds}\frac{ds}{dt}\,|\mathrm{v}| \\
&= \kappa\,|\mathrm{v}|^2\,.
\end{aligned}$$

So we obtain the second equation. The last formula then follows from the fact that N is the direction of the normal component of the acceleration. □

To get a more geometric feel for curvature we have

Proposition 1.3.6. *Consider a regular curve* $\mathrm{q}(t):(a,b)\to\mathbb{R}^n$. *If* $|\mathrm{q}(t)|\leq R$ *for all* t *and* $|\mathrm{q}(t_0)|=R$, *then* $\kappa(t_0)\geq\frac{1}{R}$.

Proof. Assume that the curve is unit speed. By assumption the function $\phi(t)=|\mathrm{q}(t)|^2$ has a maximum at t_0 thus

$$0=\frac{d\phi}{dt}(t_0)=2\,\mathrm{q}(t_0)\cdot\dot{\mathrm{q}}(t_0)$$

and

$$0\geq\frac{d^2\phi}{dt^2}(t_0)=2\,\mathrm{q}(t_0)\cdot\ddot{\mathrm{q}}(t_0)+2\,|\dot{\mathrm{q}}(t_0)|^2.$$

Thus

$$1=|\dot{\mathrm{q}}(t_0)|^2\leq-\,\mathrm{q}(t_0)\cdot\ddot{\mathrm{q}}(t_0)\leq|\mathrm{q}(t_0)|\,|\ddot{\mathrm{q}}(t_0)|=R\,|\ddot{\mathrm{q}}(t_0)|.$$

As the curve is unit speed we also have $\kappa(t)=|\ddot{\mathrm{q}}(t_0)|$. This proves the claim. □

Definition 1.3.7. An *involute* of a curve $\mathrm{q}(t)$ is a curve $\mathrm{q}^*(t)$ that lies on the corresponding tangent lines to $\mathrm{q}(t)$ and intersects these tangent lines orthogonally.

We can always construct involutes to regular curves. First of all

$$\mathrm{q}^*(t)=\mathrm{q}(t)+u(t)\,\mathrm{T}(t)$$

as it is forced to lie on the tangent lines to q. Secondly, the velocity v^* must be parallel to N. Since

$$\frac{d\,\mathrm{q}^*}{dt}=\frac{d\,\mathrm{q}}{dt}+\frac{du}{dt}\,\mathrm{T}+u\kappa\frac{ds}{dt}\,\mathrm{N}=\frac{ds}{dt}\,\mathrm{T}+\frac{du}{dt}\,\mathrm{T}+u\kappa\frac{ds}{dt}\,\mathrm{N}$$

this forces us to select u so that

$$\frac{du}{dt}=-\frac{ds}{dt}.$$

Thus

$$\mathrm{q}^*(t)=\mathrm{q}(t)-s(t)\,\mathrm{T}(t),$$

where s is any arclength parametrization of q. Note that s is only determined up to a constant so we always get infinitely many involutes to a given curve.

Example 1.3.8. If we strip a length of masking tape glued to a curve keeping it taut while doing so, then the end of the tape will trace an involute.

Assume the original curve is unit speed $\mathrm{q}(s)$. The process of stripping the tape from the curve forces the endpoint of the tape to have an equation of the form

$$\mathrm{q}^*(s)=\mathrm{q}(s)+u(s)\,\mathrm{T}(s)$$

since for each value of s the tape has two parts, the first being the curve up to $\mathrm{q}(s)$ and the second the line segment from $\mathrm{q}(s)$ to $\mathrm{q}(s)+u(s)\,\mathrm{T}(s)$. The length of this is up to a constant given by

$$s+u(s).$$

As the piece of tape doesn't change length this is constant. This shows that $u=c-s$ for some constant c and thus that the curve is an involute.

EXAMPLE 1.3.9. Huygens designed pendulums using involutes. His idea was to take two planar convex curves that are mirror images of each other in the y-axis and are tangent to the y-axis with the unit tangent at this cusp pointing downwards. Suspend a string from this cusp point of length L with a metal disc attached at the bottom end to keep the string taut. Now displace the metal disc horizontally and release it. Gravity will then force the disc to swing back and forth. The trajectory will depend on the shape of the chosen convex curve and will be an involute of that curve.

Huygens was interested in creating a pendulum with the property that its period does not depend on the amplitude of the swing. Thus the period will remain constant even though the pendulum slows down with time. A curve with this property is called *tautochronic* and Huygens showed that it has to be a cycloid that looks like

$$R(\sin t, \cos t) + R(t, 0).$$

The involute is also a cycloid (see also exercises below).

EXAMPLE 1.3.10. Consider the unit circle $\mathrm{q}(s) = (\cos s, \sin s)$. This parametrization is by arclength so we obtain the involutes

$$\mathrm{q}^*(s) = (\cos s, \sin s) + (c - s)(-\sin s, \cos s).$$

In polar coordinates we have

$$r(s) = |\mathrm{q}^*(s)| = \sqrt{1 + (c-s)^2}.$$

When $c = 0$ we see that r increases with s and that the involute looks like a spiral.

DEFINITION 1.3.11. An *evolute* of a curve $\mathrm{q}(t)$ is a curve $\mathrm{q}^*(t)$ such that the tangent lines to q^* are orthogonal to q at corresponding values of t. Thus $\mathrm{q}^*(t)$ lies on the normal line to q that goes through $\mathrm{q}(t)$ and has velocity that is tangent to this normal line.

REMARK 1.3.12. Note that if q^* is an involute to q, then conversely q is an evolute to q^*. It is however quite complicated to construct evolutes in general, but, as we shall see, there are formulas for both planar and space curves.

Evolutes must look like

$$\mathrm{q}^*(t) = \mathrm{q}(t) + \mathrm{V}(t),$$

where $\mathrm{V} \cdot \mathrm{T} = 0$ and also have the property that

$$0 = \mathrm{T} \cdot \frac{d\,\mathrm{q}^*}{dt} = \mathrm{T} \cdot \left(\frac{d\,\mathrm{q}}{dt} + \frac{d\mathrm{V}}{dt} \right),$$

which is equivalent to

$$\mathrm{T} \cdot \frac{d\mathrm{V}}{dt} = -\frac{ds}{dt}.$$

Exercises

(1) Show that a regular curve is part of a line if all its tangent lines pass through a fixed point c. Hint: Show that $\mathrm{T} = \pm \frac{\mathrm{q} - \mathrm{c}}{|\mathrm{q} - \mathrm{c}|}$, differentiate this equation, and show that $\kappa = 0$.

(2) Consider a regular curve $\mathrm{q}(t)$ with arclength parameter s. Show that if $\mathrm{T}(t_n) \neq \mathrm{T}(t_0)$ and $t_n \to t_0$, then

$$1 = \lim_{t_n \to t_0} \frac{|\theta(t_n) - \theta(t_0)|}{\arccos(\mathrm{T}(t_n) \cdot \mathrm{T}(t_0))}$$

and

$$\kappa(t_0) = \lim_{t_n \to t_0} \frac{\arccos(\mathrm{T}(t_n) \cdot \mathrm{T}(t_0))}{|s(t_n) - s(t_0)|}.$$

Hint: Use exercise 1 from section 1.2.

(3) Show that for vectors $v, w \in \mathbb{R}^n$ we have

$$\begin{aligned} \text{area of parallelogram}(v, w) &= \sqrt{|v|^2 |w|^2 - (v \cdot w)^2} \\ &= |v| |w| \sin \measuredangle(v, w). \end{aligned}$$

(4) Show that the curvature of a planar circle of radius R is $\frac{1}{R}$ by parametrizing this curve in the following way $\mathrm{q}(t) = R(\cos t, \sin t) + \mathrm{c}$.

(5) Find the curvature for the twisted cubic

$$\mathrm{q}(t) = \left(t, t^2, t^3\right).$$

(6) Let $\mathrm{q}(t)$ be a regular curve with positive curvature. Define two vector fields whose integral curves are involutes to q.

(7) Calculate the speed and curvature of the scaled curve $R\,\mathrm{q}(t)$, $R \neq 0$, in terms of the speed and curvature of $\mathrm{q}(t)$.

(8) Give examples of regular curves $\mathrm{q}(t) : (a, b) \to \mathbb{R}^n$ with $|\mathrm{q}(t)| \geq R$ for all t, $|\mathrm{q}(t_0)| = R$, and $\kappa(t_0) = c$ for any $c \geq 0$.

(9) If a curve in $\mathbb{R}^2$ is given as a graph $y = f(x)$ show that the curvature is given by

$$\kappa = \frac{|f''|}{\left(1 + (f')^2\right)^{\frac{3}{2}}}.$$

(10) Let $\mathrm{q}(t) = r(t)(\cos t, \sin t)$. Show that the speed is given by

$$\left(\frac{ds}{dt}\right)^2 = \left(\frac{dr}{dt}\right)^2 + r^2$$

and the curvature

$$\kappa = \frac{\left|2\left(\frac{dr}{dt}\right)^2 + r^2 - r\frac{d^2 r}{dt^2}\right|}{\left(\left(\frac{dr}{dt}\right)^2 + r^2\right)^{\frac{3}{2}}}.$$

(11) Let $\mathrm{q}(t) : I \to \mathbb{R}^3$ be a regular curve with speed $\frac{ds}{dt} = \left|\frac{d\,\mathrm{q}}{dt}\right|$, where s is the arclength parameter. Prove that

$$\kappa = \frac{\sqrt{\frac{d^2\,\mathrm{q}}{dt^2} \cdot \frac{d^2\,\mathrm{q}}{dt^2} - \left(\frac{d^2 s}{dt^2}\right)^2}}{\left(\frac{ds}{dt}\right)^2}.$$

(12) Compute the curvature of the logarithmic spiral

$$ae^{bt}(\cos t, \sin t).$$

(13) Compute the curvature of the spiral of Archimedes:

$$(a + bt)(\cos t, \sin t).$$

(14) Consider the tractrix from section 1.2 exercise 17.
 (a) Show that the curvature is $\kappa = \frac{1}{R}\frac{y}{\sqrt{R^2-y^2}} = \frac{1}{R}\tan\theta$.
 (b) Show that the tractrix is the involute of $y = R\cosh\frac{x}{R}$ with $c = 0$.

(15) (Huygens, 1673) Consider the cycloid

$$\mathrm{q}(t) = R(t+\sin t, 1+\cos t)$$

(see also section 1.2 exercise 18 and note that this cycloid comes with a different parametrization and initial position).
 (a) Show that the speed satisfies

$$\left|\frac{d\,\mathrm{q}}{dt}\right|^2 = 2R^2(1+\cos t) = 2R^2\frac{\sin^2 t}{1-\cos t}.$$

 (b) Show that the arclength parameter s with initial value $s(0) = 0$ satisfies

$$s^2 = 8R^2(1-\cos t).$$

 (c) Show that the curvature satisfies

$$\kappa^2 = \frac{1}{8R^2(1+\cos t)}.$$

 (d) Show that for a general cycloid it is always possible to find $a \in \mathbb{R}$ such that

$$(s-a)^2 + \frac{1}{\kappa^2} = 16R^2.$$

 (e) Show $a = 4R$ for the cycloid

$$\mathrm{q}(t) = R(t-\sin t, 1-\cos t)$$

 if we assume that $s(0) = 0$.

(16) Show that the involute to a straight line is a point.

(17) Show that a planar circle has its center as an evolute.

(18) The circular helix is given by

$$\mathrm{q}(t) = R(\cos t, \sin t, 0) + h(0,0,t).$$

Reparametrize this curve to be unit speed and show that its involutes lie in planes given by $z = c$ for some constant c.

(19) Let $\mathrm{q}(s)$ be a planar unit speed curve with positive curvature. Show that the curvature of the involute

$$\mathrm{q}^*(s) = \mathrm{q}(s) + (L-s)\,\mathrm{T}(s)$$

satisfies

$$\kappa^* = \frac{1}{|L-s|}$$

and compute the evolute of q^*.

(20) For a regular curve $\mathrm{q}(t) : I \to \mathbb{R}^n$ we say that a field X is *parallel* along q if $\mathrm{X}\cdot\mathrm{T} = 0$ and $\frac{d\mathrm{X}}{dt}$ is parallel to T, i.e.,

$$\frac{d\mathrm{X}}{dt} = \left(\frac{d\mathrm{X}}{dt}\cdot\mathrm{T}\right)\mathrm{T} = -\left(\frac{d\,\mathrm{T}}{dt}\cdot\mathrm{X}\right)\mathrm{T}.$$

 (a) Show that for a fixed t_0 and $\mathrm{X}(t_0) \perp \mathrm{T}(s_0)$ there is a unique parallel field X that has the value $\mathrm{X}(t_0)$ at t_0.
 (b) Show that if X_1 and X_2 are both parallel along q, then $\mathrm{X}_1\cdot\mathrm{X}_2$ is constant.

(c) A *Bishop frame* consists of an orthonormal frame $\mathrm{T}, \mathrm{N}_1, \mathrm{N}_2, ..., \mathrm{N}_{n-1}$ along the curve so that all N_i are parallel along q. For such a frame show that

$$\begin{aligned}
&\frac{d}{dt}\left[\begin{array}{ccccc} \mathrm{T} & \mathrm{N}_1 & \mathrm{N}_2 & \cdots & \mathrm{N}_{n-1} \end{array}\right] \\
&= \frac{ds}{dt}\left[\begin{array}{ccccc} \mathrm{T} & \mathrm{N}_1 & \mathrm{N}_2 & \cdots & \mathrm{N}_{n-1} \end{array}\right]\left[\begin{array}{ccccc} 0 & \kappa_1 & \kappa_2 & \cdots & \kappa_{n-1} \\ -\kappa_1 & 0 & 0 & \cdots & 0 \\ -\kappa_2 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\kappa_{n-1} & 0 & 0 & \cdots & 0 \end{array}\right].
\end{aligned}$$

Note that such frames always exist, even when the curve doesn't have positive curvature everywhere.

(d) Show further for such a frame that

$$\kappa^2 = \kappa_1^2 + \kappa_2^2 + \cdots + \kappa_{n-1}^2.$$

The collection $(\kappa_1, \kappa_2, ..., \kappa_{n-1})$ can in turn be thought of as a curve going into $\mathbb{R}^{n-1}$ and be investigated for higher order behavior of q. When $\kappa > 0$ one generally divides this curve by κ and considers the spherical curve into S^{n-2}.

(e) Give an example of a closed space curve where the parallel fields don't close up.

1.4. Integral Curves

In this section we try to understand the curvature of curves that are solutions to differential equations. As it is rarely possible to find explicit formulas for such solutions the goal is to use the fact that we know they exist and calculate their curvatures using only the data that the differential equation gives us. Recall that curves that are solutions to equations can also be considered as solutions to differential equations. We also explain how Kepler's laws imply Newton's gravitational law. This involves an interesting blend of geometry and calculus that is relevant for other concepts that will be developed throughout the notes.

We start by considering a solution to a first order equation

$$\mathrm{v} = \frac{d\,\mathrm{q}}{dt} = F(\mathrm{q}(t)).$$

The first observation is that the speed is given by

$$|\mathrm{v}| = \left|\frac{d\,\mathrm{q}}{dt}\right| = |F(\mathrm{q}(t))|.$$

The acceleration is computed using the chain rule

$$\mathrm{a} = \frac{d\,\mathrm{v}}{dt} = \frac{dF(\mathrm{q}(t))}{dt} = DF\left(\frac{d\,\mathrm{q}}{dt}\right) = DF(F(\mathrm{q}(t))).$$

The curvature is then given by

$$\begin{aligned}
\kappa^2(t) &= \frac{|\mathrm{v}|^2 |\mathrm{a}|^2 - (\mathrm{v}\cdot\mathrm{a})^2}{|\mathrm{v}|^6} \\
&= \frac{|F(\mathrm{q}(t))|^2 |DF(F(\mathrm{q}(t)))|^2 - (F(\mathrm{q}(t)) \cdot DF(F(\mathrm{q}(t))))^2}{|F(\mathrm{q}(t))|^6}.
\end{aligned}$$

So if we wish to calculate the curvature for a solution that passes through a fixed point q_0 at time $t = t_0$, then we have

$$\kappa^2(t_0) = \frac{|F(q_0)|^2 |DF(F(q_0))|^2 - (F(q_0) \cdot DF(F(q_0)))^2}{|F(q_0)|^6}.$$

This is a formula that does not require us to solve the equation.

For a second order equation

$$\mathrm{a} = \frac{d^2 \mathrm{q}}{dt^2} = G\left(\mathrm{q}(t), \frac{d\mathrm{q}}{dt}\right) = G(\mathrm{q}(t), \mathrm{v}(t))$$

there isn't much to compute as we now have to be given both position q_0 and velocity v_0 at time t_0. The curvature is given by

$$\begin{aligned} \kappa^2(t) &= \frac{|\mathrm{v}|^2 |\mathrm{a}|^2 - (\mathrm{v} \cdot \mathrm{a})^2}{|\mathrm{v}|^6} \\ &= \frac{|v_0|^2 |G(q_0, v_0)|^2 - (v_0 \cdot G(q_0, v_0))^2}{|v_0|^6}. \end{aligned}$$

However, note that we can also calculate the change in speed by observing that

$$\frac{d|\mathrm{v}|^2}{dt} = 2\,\mathrm{v} \cdot \mathrm{a} = 2\,\mathrm{v} \cdot G(\mathrm{q}, \mathrm{v}).$$

A few examples will hopefully clarify this a little better.

EXAMPLE 1.4.1. First an example were we know that the solutions are circles.

$$F(x, y) = (-y, x)$$

and

$$\begin{aligned} DF(F(x,y)) &= \begin{bmatrix} \frac{\partial(-y)}{\partial x} & \frac{\partial(-y)}{\partial y} \\ \frac{\partial(x)}{\partial x} & \frac{\partial(x)}{\partial y} \end{bmatrix} \begin{bmatrix} -y \\ x \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -y \\ x \end{bmatrix} \\ &= \begin{bmatrix} -x \\ -y \end{bmatrix}. \end{aligned}$$

So at $q_0 = (x_0, y_0)$ we have

$$\kappa^2 = \frac{\left(x_0^2 + y_0^2\right)^2 - (x_0 y_0 - x_0 y_0)^2}{\left(x_0^2 + y_0^2\right)^3} = \frac{1}{|q_0|^2},$$

which agrees with our knowledge that the curvature is the reciprocal of the radius.

EXAMPLE 1.4.2. Next we look at the second order equation

$$\mathrm{a} = -g \frac{\mathrm{q}}{|\mathrm{q}|^3},\ g > 0.$$

The curvature is

$$\begin{aligned} \kappa^2 &= \frac{|\mathrm{v}|^2 \left| -g \frac{\mathrm{q}}{|\mathrm{q}|^3} \right|^2 - \left(-g \frac{\mathrm{q}}{|\mathrm{q}|^3} \cdot \mathrm{v} \right)^2}{|\mathrm{v}|^6} \\ &= g^2 \frac{|\mathrm{q}|^2 |\mathrm{v}|^2 - (\mathrm{q} \cdot \mathrm{v})^2}{|\mathrm{q}|^6 |\mathrm{v}|^6}. \end{aligned}$$

Yielding

$$\kappa = g\frac{\text{area of parallelogram}\,(\mathrm{q},\mathrm{v})}{|\mathrm{q}|^3\,|\mathrm{v}|^3}.$$

So the curvature vanishes when the velocity is radial (proportional to position), this conforms with the fact that radial lines are solutions to this equation. Otherwise all other solutions must have nowhere vanishing curvature. In general the numerator is constant along solutions as

$$\begin{aligned}\frac{d}{dt}\left(|\mathrm{q}|^2\,|\mathrm{v}|^2-(\mathrm{q}\cdot\mathrm{v})^2\right) &= 2\,\mathrm{q}\cdot\mathrm{v}\,|\mathrm{v}|^2+2\,|\mathrm{q}|^2\,\mathrm{v}\cdot\mathrm{a}\\ &\quad -2\,\mathrm{q}\cdot\mathrm{v}\left(|\mathrm{v}|^2+\mathrm{q}\cdot\mathrm{a}\right)\\ &= 2\,\mathrm{q}\cdot\mathrm{v}\,|\mathrm{v}|^2-2g\frac{1}{|\mathrm{q}|}\,\mathrm{v}\cdot\mathrm{q}\\ &\quad -2\,\mathrm{q}\cdot\mathrm{v}\left(|\mathrm{v}|^2-g\frac{1}{|\mathrm{q}|}\right)\\ &= 0.\end{aligned}$$

This is better known as *Kepler's second law.* The triangle with constant area in Kepler's second law has q and v as sides. Thus its area is half the area of the parallelogram we just calculated to be constant.

Below we show how Kepler's laws imply Newton's gravitational law. This is exactly what Newton did in Principia. He also asserted that one had unique solutions to the initial value problems

$$\mathrm{a}=-g\frac{\mathrm{q}}{|\mathrm{q}|^3},\ \mathrm{q}(0)=q_0,\ \mathrm{v}(0)=v_0,$$

and then concluded that the solutions have be conic sections as asserted by Kepler's laws. This will be discussed in exercises to this section. We consider the following mathematical version of Kepler's laws. Concretely, one might think of the orbits being the planetary orbits around the sun or the moons around Jupiter.

(1) All orbits are conic sections with the origin as a focal point.
(2) A given orbit sweeps out equal areas in equal time, i.e., $A^2=|\mathrm{q}|^2\,|\mathrm{v}|^2-(\mathrm{q}\cdot\mathrm{v})^2$ is constant.
(3) The ratio $\frac{a^3}{T^2}$ is the same for all elliptical orbits, where a is the major axis and T is the period of the orbit.

To prove Newton's law it is convenient to parametrize the Cartesian coordinates using polar coordinates $\mathrm{q}(r,\theta)=(r\cos\theta,r\sin\theta)$. We then have to figure out how to calculate the velocity and acceleration of a curve

$$\mathrm{q}(t)=\begin{bmatrix} r\cos\theta\\ r\sin\theta\end{bmatrix}=r\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}=|\mathrm{q}|\,\frac{\mathrm{q}}{|\mathrm{q}|}.$$

By the chain rule the velocity becomes:

$$\mathrm{v}=\dot{\mathrm{q}}=\dot{r}\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}+\dot{\theta}\begin{bmatrix}-r\sin\theta\\ r\cos\theta\end{bmatrix}=\dot{r}\frac{\partial\,\mathrm{q}}{\partial r}+\dot{\theta}\frac{\partial\,\mathrm{q}}{\partial\theta}.$$

For the acceleration we first note that the *radial,* $\frac{\partial\,\mathrm{q}}{\partial r}$, and *angular,* $\frac{\partial\,\mathrm{q}}{\partial\theta}$, vectors are linearly independent so the second partials can be written as linear combinations of these vectors:

$$\begin{aligned}
\frac{\partial^2 \mathrm{q}}{\partial r^2} &= 0, \\
\frac{\partial^2 \mathrm{q}}{\partial r \partial \theta} &= \frac{\partial^2 \mathrm{q}}{\partial \theta \partial r} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = \frac{1}{r}\frac{\partial \mathrm{q}}{\partial \theta}, \\
\frac{\partial^2 \mathrm{q}}{\partial \theta^2} &= \begin{bmatrix} -r\cos\theta \\ -r\sin\theta \end{bmatrix} = -r\frac{\partial \mathrm{q}}{\partial r}.
\end{aligned}$$

The acceleration now has the formula:

$$\begin{aligned}
\mathrm{a} &= \ddot{\mathrm{q}} \\
&= \dot{\mathrm{v}} \\
&= \ddot{r}\frac{\partial \mathrm{q}}{\partial r} + \ddot{\theta}\frac{\partial \mathrm{q}}{\partial \theta} + \dot{r}\frac{d}{dt}\frac{\partial \mathrm{q}}{\partial r} + \dot{\theta}\frac{d}{dt}\frac{\partial \mathrm{q}}{\partial \theta} \\
&= \ddot{r}\frac{\partial \mathrm{q}}{\partial r} + \ddot{\theta}\frac{\partial \mathrm{q}}{\partial \theta} + \dot{r}\left(\frac{dr}{dt}\frac{\partial}{\partial r} + \frac{d\theta}{dt}\frac{\partial}{\partial \theta}\right)\frac{\partial \mathrm{q}}{\partial r} + \dot{\theta}\left(\frac{dr}{dt}\frac{\partial}{\partial r} + \frac{d\theta}{dt}\frac{\partial}{\partial \theta}\right)\frac{\partial \mathrm{q}}{\partial \theta} \\
&= \ddot{r}\frac{\partial \mathrm{q}}{\partial r} + \ddot{\theta}\frac{\partial \mathrm{q}}{\partial \theta} + \dot{r}^2\frac{\partial^2 \mathrm{q}}{\partial r^2} + 2\dot{r}\dot{\theta}\frac{\partial^2 \mathrm{q}}{\partial r \partial \theta} + \dot{\theta}^2\frac{\partial^2 \mathrm{q}}{\partial \theta^2} \\
&= \ddot{r}\frac{\partial \mathrm{q}}{\partial r} + \ddot{\theta}\frac{\partial \mathrm{q}}{\partial \theta} + 2\dot{r}\dot{\theta}\frac{1}{r}\frac{\partial \mathrm{q}}{\partial \theta} - \dot{\theta}^2 r\frac{\partial \mathrm{q}}{\partial r} \\
&= \left(\ddot{r} - r\dot{\theta}^2\right)\frac{\partial \mathrm{q}}{\partial r} + \left(\ddot{\theta} + \frac{2\dot{r}\dot{\theta}}{r}\right)\frac{\partial \mathrm{q}}{\partial \theta}.
\end{aligned}$$

We start by proving two results that only depend on the properties of one orbit.

Proposition 1.4.3 (Newton). *If an orbit satisfies the second law, then the acceleration is always radial, i.e.,* $\ddot{\theta} + \frac{2\dot{r}\dot{\theta}}{r} = 0$ *for all orbits.*

Proof. We start by observing that

$$\frac{d\left(r^2\dot{\theta}\right)}{dt} = r^2\ddot{\theta} + 2r\dot{r}\dot{\theta} = r^2\left(\ddot{\theta} + \frac{2\dot{r}\dot{\theta}}{r}\right).$$

Next we note that the square of the area of the parallelogram spanned by the position and velocity is

$$|\mathrm{q}|^2\,|\mathrm{v}|^2 - (\mathrm{q}\cdot\mathrm{v})^2 = r^2\dot{\theta}^2\left|\frac{\partial \mathrm{q}}{\partial \theta}\right|^2 = r^4\dot{\theta}^2.$$

Thus the second law implies that $r^2\dot{\theta}$ is constant, which in turn implies that the acceleration is radial. □

Having shown that the acceleration is radial we can now show that it must satisfy the inverse square law if it travels along a conic section.

Lemma 1.4.4 (Newton). *If an orbit satisfies the first and second law, then the acceleration satisfies an inverse square law:*

$$\mathrm{a} = -g\frac{\mathrm{q}}{|\mathrm{q}|^3}$$

for some gravitational constant g.

PROOF. We can rotate the polar coordinates so that the equation for the orbit is given by the equation

$$r\left(1 - e\cos\theta\right) = p,$$

where $e \geq 0$ and $p > 0$ (see example 1.1.10). We let $A = r^2\dot{\theta}$ which we know is a constant along the orbit. Differentiating the equation for the orbit yields:

$$0 = \dot{r}\left(1 - e\cos\theta\right) + r\dot{\theta}e\sin\theta = \frac{1}{r}\dot{r}p + \frac{1}{r}Ae\sin\theta.$$

Thus $\dot{r}p = -Ae\sin\theta$, and we can differentiate to obtain

$$\ddot{r}p = -A\dot{\theta}e\cos\theta = -\frac{A^2}{r^2}\left(e\cos\theta\right) = \frac{A^2}{r^2}\left(\frac{p}{r} - 1\right).$$

The radial part of the acceleration is then given by

$$\ddot{r} - r\dot{\theta}^2 = \frac{A^2}{r^2}\left(\frac{1}{r} - \frac{1}{p}\right) - \frac{A^2}{r^3} = -\frac{A^2}{p}\frac{1}{r^2}.$$

This establishes the inverse square law with gravitational constant $g = \frac{A^2}{p}$. □

Finally we must show that the gravitational constants $g = \frac{A^2}{p}$ are the same for all orbits.

PROPOSITION 1.4.5 (Newton). *The gravitational constant for an elliptical orbit satisfies*

$$g = \frac{A^2}{p} = 4\pi^2\frac{a^3}{T^2}.$$

PROOF. Recall from example 1.1.10 that we can write the equation of the ellipse in suitable Cartesian coordinates as

$$\frac{\left(x - \frac{ep}{1-e^2}\right)^2}{\left(\frac{p}{1-e^2}\right)^2} + \frac{y^2}{\frac{p^2}{1-e^2}} = \frac{\left(x - ea\right)^2}{a^2} + \frac{y^2}{pa} = 1.$$

If T is the period of the ellipse, then we have the formula for the area of the ellipse:

$$\pi a\sqrt{pa} = \pi ab = \frac{1}{2}\int_0^{2\pi} r^2 d\theta = \frac{1}{2}\int_0^T A dt = \frac{1}{2}AT.$$

Consequently,

$$\frac{A^2}{p} = \frac{4\pi^2a^2pa}{pT^2} = 4\pi^2\frac{a^3}{T^2}.$$

□

REMARK 1.4.6. There are similar formulas for parabolas and hyperbolas even though they don't have a period. Instead we can calculate the area of the region bounded by the curve and the y-axis and relate this to the time T it takes to travel this part of the orbit. However, it is clearly more reasonable to assume that $g = \frac{A^2}{p}$ is the same for all orbits. This can be tested on an orbit when only a small part of it is known. We already understand how to find A. To find p note that a conic section with a focus at the origin is completely determined by 3 points on the curve.

THEOREM 1.4.7 (Newton). *If all orbits are conic sections with A being constant along the orbit and $\frac{A^2}{p}$ being the same for all orbits, then the orbits satisfy*

$$\mathrm{a} = -g\frac{\mathrm{q}}{|\mathrm{q}|^3}$$

for a gravitational constant g that does not depend on the orbits.

REMARK 1.4.8. Newton took this a little further and showed that $g = GM$, where M is the mass of the central body (e.g., sun or Jupiter) and G is a universal gravitational constant that is the same for all bodies.

Exercises

(1) Assume a planar curve is given as a level set $F(x, y) = c$, where $\nabla F \neq 0$ everywhere along the curve. We orient and parametrize the curve so that $\mathrm{v} = \left(-\frac{\partial F}{\partial y}, \frac{\partial F}{\partial x}\right)$. Use the chain rule to show that the acceleration is

$$\begin{aligned}\mathrm{a} &= \begin{bmatrix} -\frac{\partial^2 F}{\partial x \partial y} & -\frac{\partial^2 F}{\partial y^2} \\ \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial y \partial x}\end{bmatrix}\begin{bmatrix} -\frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial x}\end{bmatrix} \\ &= \left[\frac{\partial \mathrm{v}}{\partial (x, y)}\right][\mathrm{v}].\end{aligned}$$

(2) Consider the equation

$$\mathrm{a} = -g\frac{\mathrm{q}}{|\mathrm{q}|^3}.$$

(a) Show that each solution lies in a plane. Hint: If n is a fixed vector, then

$$\frac{d(\mathrm{n}\cdot\mathrm{q})}{dt} = \alpha(t)(\mathrm{n}\cdot\mathrm{v}), \quad \frac{d(\mathrm{n}\cdot\mathrm{v})}{dt} = \beta(t)(\mathrm{n}\cdot\mathrm{q})$$

and use this to conclude that if n is perpendicular to $\mathrm{q}(t_0), \mathrm{v}(t_0)$, then n is perpendicular to $\mathrm{q}(t), \mathrm{v}(t)$ for all t.

(b) Show that the total energy

$$E = \frac{1}{2}|\mathrm{v}|^2 - g\frac{1}{|\mathrm{q}|}$$

is constant along solutions.

(c) Show that the tangent line to a solution can be determined by the constants

$$\begin{aligned}A^2 &= |\mathrm{q}|^2|\mathrm{v}|^2 - (\mathrm{q}\cdot\mathrm{v})^2, \\ E &= \frac{1}{2}|\mathrm{v}|^2 - g\frac{1}{|\mathrm{q}|}\end{aligned}$$

and a point on the solution.

(3) Consider an equation

$$\mathrm{a} = f(|\mathrm{q}|)\,\mathrm{q}$$

coming from a radial force field.

(a) Show that

$$A^2 = |\mathrm{q}|^2|\mathrm{v}|^2 - (\mathrm{q}\cdot\mathrm{v})^2$$

is constant along solutions.

(b) Show that each solution lies in a plane.

(4) Define the *positive perpendicular* to a planar vector as

$$\hat{X} = \widehat{\begin{bmatrix} a \\ b \end{bmatrix}} = \begin{bmatrix} -b \\ a \end{bmatrix}.$$

(a) Show that

$$\begin{aligned} \det \begin{bmatrix} X & \hat{X} \end{bmatrix} &= |X|^2, \\ |X| &= \left|\hat{X}\right|, \\ \hat{\hat{X}} &= -X, \\ \alpha\widehat{X + \beta} Y &= \alpha\hat{X} + \beta\hat{Y}. \end{aligned}$$

(b) Show that when $X = X(t)$, then

$$\widehat{\frac{dX}{dt}} = \frac{d\hat{X}}{dt}.$$

(5) Consider planar curves in $\mathbb{R}^2$ that satisfy the equation

$$\mathrm{a} = -g\frac{\mathrm{q}}{|\mathrm{q}|^3}.$$

The goal is to give a direct proof that solutions are conic sections.

(a) Show that

$$\det \begin{bmatrix} \mathrm{q} & \mathrm{v} \end{bmatrix} = -\,\mathrm{q} \cdot \hat{\mathrm{v}}$$

and

$$A^2 = |\mathrm{q}|^2 |\mathrm{v}|^2 - (\mathrm{q} \cdot \mathrm{v})^2 = (\mathrm{q} \cdot \hat{\mathrm{v}})^2.$$

The quantity $A = -\,\mathrm{q} \cdot \hat{\mathrm{v}}$ is called the signed area of the parallelogram spanned by q, v.

(b) Show that the signed area $A = -\,\mathrm{q} \cdot \hat{\mathrm{v}}$ is constant in time (see also example 1.4.2). Note that this property only uses that the acceleration is radial $\mathrm{a} = f(|\mathrm{q}|)\,\mathrm{q}$.

(c) Use $A^2 = |\mathrm{q}|^2 |\mathrm{v}|^2 - (\mathrm{q} \cdot \mathrm{v})^2 \neq 0$ to show that if $\mathrm{q} \cdot \mathrm{x} = 0$ and $\mathrm{v} \cdot \mathrm{x} = 0$, then $\mathrm{x} = 0$. Hint: Write $\mathrm{x} = \alpha\,\mathrm{q} + \beta\,\mathrm{v}$ and take dot products with the two vectors q, v.

(d) For $\mathrm{q} \neq 0$ define the vector

$$\mathrm{k} = \frac{A}{g}\hat{\mathrm{v}} + \frac{\mathrm{q}}{|\mathrm{q}|}$$

and show that $\frac{A^2}{g} = |\mathrm{q}| - \mathrm{q} \cdot \mathrm{k}$.

(e) Show that k is constant and conclude that the orbit is a conic section as in example 1.1.10. Hint: Show that

$$\mathrm{q} \cdot \frac{d\,\mathrm{k}}{dt} = 0, \ \mathrm{v} \cdot \frac{d\,\mathrm{k}}{dt} = 0.$$

(6) Consider a curve with the property that $\mathrm{q}(t)$ and $\dot{\mathrm{q}}(t)$ are linearly independent for all t. Show that for any constant $A^2 > 0$, there is a reparametrization $\mathrm{q}(s)$ such that

$$|\mathrm{q}|^2 \left|\frac{d\,\mathrm{q}}{ds}\right|^2 - \left(\mathrm{q} \cdot \frac{d\,\mathrm{q}}{ds}\right)^2 = A^2$$

for all s.

(7) Fix $g > 0$. In this exercise you'll see how Newton indicated that the solutions to the inverse square law are conic sections. He was criticized for not solving the equations directly.

(a) Show that any conic section that is an ellipse, parabola, or hyperbola with focus at the origin can be parametrized so that $|\mathrm{q}|^2 \left|\frac{d\,\mathrm{q}}{ds}\right|^2 - \left(\mathrm{q} \cdot \frac{d\,\mathrm{q}}{ds}\right)^2 = pg$.

(b) Show that any conic section can be parametrized so that it solves

$$\mathrm{a} = -g \frac{\mathrm{q}}{|\mathrm{q}|^3}.$$

(c) Show that for each set of initial values $\mathrm{q}(0)$ and $\mathrm{v}(0)$, there is a conic section with these initial values that solves

$$\mathrm{a} = -g \frac{\mathrm{q}}{|\mathrm{q}|^3}.$$

CHAPTER 2

Planar Curves

2.1. The Fundamental Equations

Our approach to planar curves follows very closely the concepts that we shall use for space curves. This is certainly not the way the subject developed historically, but it has shown itself to be a very useful and general strategy.

Before delving into the theory the keen reader might be interested in a few generalities about taking derivatives of a basis $U(t), V(t)$ that depends on t, and viewed as a choice of basis at $\mathrm{q}(t)$. We often use $U(t) = \mathrm{T}(t)$. Given any choice for $U(t)$, a natural choice for $V(t)$ would be the unit vector orthogonal to $U(t)$ such that $\det \begin{bmatrix} U & V \end{bmatrix} = 1$, i.e., $V = \hat{U}$ (see also exercise 4 in section 1.4). The goal is to identify the matrix $[D]$ that appears in

$$\frac{d}{dt}\begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} \frac{d}{dt}U & \frac{d}{dt}V \end{bmatrix} = \begin{bmatrix} U & V \end{bmatrix}[D]$$

There is a complicated formula (see theorem A.1.1)

$$[D] = \left(\begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} U & V \end{bmatrix}\right)^{-1} \begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} \frac{d}{dt}U & \frac{d}{dt}V \end{bmatrix}$$

that can be simplified to

THEOREM 2.1.1. *Let $U(t), V(t)$ be an orthonormal frame that depends on a parameter t, then*

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} U & V \end{bmatrix} &= \begin{bmatrix} U & V \end{bmatrix}\begin{bmatrix} 0 & \lambda \\ -\lambda & 0 \end{bmatrix}, \\ \lambda &= U \cdot \frac{d}{dt}V = -V \cdot \frac{d}{dt}U \end{aligned}$$

or

$$\begin{aligned} \frac{d}{dt}U &= \lambda V, \\ \frac{d}{dt}V &= -\lambda U \end{aligned}$$

PROOF. We use that

$$\begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} U \cdot U & U \cdot V \\ V \cdot U & V \cdot V \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The derivative of this then gives

$$\begin{aligned} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} \frac{d}{dt}U & \frac{d}{dt}V \end{bmatrix}^t \begin{bmatrix} U & V \end{bmatrix} + \begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} \frac{d}{dt}U & \frac{d}{dt}V \end{bmatrix} \\ &= \begin{bmatrix} \left(\frac{d}{dt}U\right) \cdot U & \left(\frac{d}{dt}U\right) \cdot V \\ \left(\frac{d}{dt}V\right) \cdot U & \left(\frac{d}{dt}V\right) \cdot V \end{bmatrix} + \begin{bmatrix} U \cdot \frac{d}{dt}U & U \cdot \frac{d}{dt}V \\ V \cdot \frac{d}{dt}U & V \cdot \frac{d}{dt}V \end{bmatrix}. \end{aligned}$$

Thus

$$\begin{aligned}\left(\frac{d}{dt}U\right)\cdot U &= 0=\left(\frac{d}{dt}V\right)\cdot V,\\ \left(\frac{d}{dt}V\right)\cdot U &= -V\cdot\frac{d}{dt}U.\end{aligned}$$

Our formula for $[D]$ then becomes

$$\begin{aligned}[D] &= \left[\begin{array}{cc} U & V\end{array}\right]^t\left[\begin{array}{cc}\frac{d}{dt}U & \frac{d}{dt}V\end{array}\right]\\ &= \left[\begin{array}{cc} U\cdot\frac{d}{dt}U & U\cdot\frac{d}{dt}V\\ V\cdot\frac{d}{dt}U & V\cdot\frac{d}{dt}V\end{array}\right]\\ &= \left[\begin{array}{cc} 0 & \lambda\\ -\lambda & 0\end{array}\right].\end{aligned}$$

□

Occasionally we need one more derivative

$$\begin{aligned}\frac{d^2}{dt^2}\left[\begin{array}{cc} U & V\end{array}\right] &= \left[\begin{array}{cc} U & V\end{array}\right]\left[\begin{array}{cc} -\lambda^2 & \frac{d\lambda}{dt}\\ -\frac{d\lambda}{dt} & -\lambda^2\end{array}\right],\\ \frac{d^2U}{dt^2} &= -\lambda^2U-\frac{d\lambda}{dt}V,\\ \frac{d^2V}{dt^2} &= \frac{d\lambda}{dt}U-\lambda^2V.\end{aligned}$$

For a planar regular curve $\mathrm{q}(t):[a,b]\to\mathbb{R}^2$ we have as for general curves

$$\frac{d\,\mathrm{q}}{dt}=|\mathrm{v}|\frac{\mathrm{v}}{|\mathrm{v}|}=\frac{ds}{dt}\,\mathrm{T}.$$

Instead of the choice of normal that depended on the acceleration (see section 1.3) we select an *oriented normal* $\mathrm{N}_\pm$ such that T and $\mathrm{N}_\pm$ are positively oriented, i.e., if $\mathrm{T}=(a,b)$, then $\mathrm{N}_\pm=(-b,a)$. This orientation is set up so that $\mathrm{N}_\pm$ points to the left when facing in the direction of T. Note that $\mathrm{N}_\pm$ can be either N or $-\mathrm{N}$.

Definition 2.1.2. The *signed curvature* is defined by

$$\kappa_\pm=\mathrm{N}_\pm\cdot\frac{d\,\mathrm{T}}{ds}.$$

Proposition 2.1.3. *If $F:\mathbb{R}^2\to\mathbb{R}^2$ is of the form $F(\mathrm{x})=O\mathrm{x}+\mathrm{c}$, where O is an orthogonal transformation, $\mathrm{c}\in\mathbb{R}^2$, and q a regular curve, then $\mathrm{q}^*=F(\mathrm{q})$ has unit tangent given by $\mathrm{T}^*=O\,\mathrm{T}$, signed normal $\mathrm{N}_\pm^*=(\det O)\,O\,\mathrm{N}_\pm$ and curvature $\kappa_\pm^*=(\det O)\,\kappa_\pm$.*

Proof. We saw in proposition 1.3.2 that $\mathrm{T}^*=O\,\mathrm{T}$. When $\det O=1$, then O preserves the orientation of being on the lefthand side, so $\mathrm{N}_\pm^*=O\,\mathrm{N}_\pm$. While if $\det O=-1$, then $\mathrm{N}_\pm^*=-O\,\mathrm{N}_\pm$ as this transformation reverses left and right. We can now also calculate the curvature:

$$\kappa_\pm^*=(\det O)\,(O\,\mathrm{N})\cdot\left(\frac{dO\,\mathrm{T}}{ds}\right)=(\det O)\,(\mathrm{N})\cdot\left(\frac{d\,\mathrm{T}}{ds}\right)=(\det O)\,\kappa.$$

□

PROPOSITION 2.1.4. *The signed curvature can be calculated using the formula*

$$\kappa_{\pm} = \frac{\textit{signed area of parallelogram}\,(\mathrm{v},\mathrm{a})}{|\mathrm{v}|^3} = \frac{\det\begin{bmatrix} \mathrm{v} & \mathrm{a} \end{bmatrix}}{|\mathrm{v}|^3}.$$

THEOREM 2.1.5. *(Euler, 1736) The fundamental equations that govern planar curves are*

$$\begin{aligned} \frac{d\,\mathrm{q}}{dt} &= \frac{ds}{dt}\,\mathrm{T}, \\ \frac{d\,\mathrm{T}}{dt} &= \kappa_{\pm}\frac{ds}{dt}\,\mathrm{N}_{\pm}, \\ \frac{d\,\mathrm{N}_{\pm}}{dt} &= -\kappa_{\pm}\frac{ds}{dt}\,\mathrm{T} = -\kappa_{\pm}\frac{d\,\mathrm{q}}{dt}. \end{aligned}$$

Moreover, given an initial position $\mathrm{q}(0)$ *and unit direction* $\mathrm{T}(0)$ *the curve* $\mathrm{q}(t)$ *is uniquely determined by its speed and signed curvature.*

PROOF. The three equations are simple to check as $\mathrm{T}, \mathrm{N}_{\pm}$ form an orthonormal basis. For fixed speed and signed curvature functions these equations form a differential equation that has a unique solution given the initial values $\mathrm{q}(0)$, $\mathrm{T}(0)$ and $\mathrm{N}_{\pm}(0)$. The normal vector is determined by the unit tangent so we have all of that data. □

Geometrically we say that the planar curve $\mathrm{q}(t)$ is determined by the planar curve $\left(\frac{ds}{dt}, \kappa_{\pm}\right)$. If it is possible to find the arc-length parametrization, then the data $(s(t), \kappa_{\pm}(t))$ can equally well be used to describe the geometry of a planar curve. A more explicit relationship between a curve and its curvature can be found in exercise 9 in this section.

We offer a combined characterization of lines and circles as the curves that are horizontal lines in $(s, \kappa_{\pm})$ coordinates, i.e., they have constant curvature.

THEOREM 2.1.6. *A planar curve is part of a line if and only if its signed curvature vanishes. A planar curve is part of a circle if and only if its signed curvature is non-zero and constant.*

PROOF. If the curvature vanishes, then we already know that it has to be a straight line.

If the curve is a circle of radius R with center c, then

$$|\mathrm{q}(s) - \mathrm{c}|^2 = R^2.$$

Differentiating this yields

$$\mathrm{T} \cdot (\mathrm{q}(s) - \mathrm{c}) = 0.$$

Thus the unit tangent is perpendicular to the radius vector $\mathrm{q}(s) - \mathrm{c}$. Differentiating again yields

$$\kappa_{\pm}\,\mathrm{N}_{\pm} \cdot (\mathrm{q}(s) - \mathrm{c}) + 1 = 0.$$

However, the normal and radius vectors must be parallel so their inner product is $\pm R$. This shows that the curvature is constant. We also obtain the equation

$$\mathrm{q} = \mathrm{c} - \frac{1}{\kappa_{\pm}}\,\mathrm{N}_{\pm}.$$

This indicates that, if a curve has constant curvature, then we should attempt to show that

$$\mathrm{c} = \mathrm{q} + \frac{1}{\kappa_\pm}\mathrm{N}_\pm$$

is constant. Since $\kappa_\pm$ is constant the derivative of this curve is

$$\frac{d\,\mathrm{c}}{ds} = \mathrm{T} + \frac{1}{\kappa_\pm}\left(-\kappa_\pm \mathrm{T}\right) = 0.$$

So c is constant and

$$\left|\mathrm{q}(s) - \mathrm{c}\right|^2 = \left|\frac{1}{\kappa_\pm}\mathrm{N}_\pm\right|^2 = \frac{1}{\kappa_\pm^2}.$$

thus showing that q is a circle of radius $\frac{1}{|\kappa_\pm|}$ centered at c. □

Proposition 2.1.7. *The evolute of a regular planar curve* $\mathrm{q}(t)$ *with non-zero curvature is given by*

$$\mathrm{q}^* = \mathrm{q} + \frac{1}{\kappa_\pm}\mathrm{N}_\pm = \mathrm{q} + \frac{1}{\kappa}\mathrm{N}.$$

Proof. This follows from remark 1.3.12 and

$$\frac{d\,\mathrm{q}^*}{dt} = \frac{d\,\mathrm{q}}{dt} + \frac{1}{\kappa_\pm}\left(-\kappa_\pm\frac{d\,\mathrm{q}}{dt}\right) + \frac{d}{dt}\left(\frac{1}{\kappa_\pm}\right)\mathrm{N}_\pm = \frac{d}{dt}\left(\frac{1}{\kappa_\pm}\right)\mathrm{N}_\pm.$$

□

Exercises

(1) Compute the signed curvature of $\mathrm{q}(t) = \left(t, t^3\right)$ and show that it vanishes at $t = 0$, is negative for $t < 0$, and positive for $t > 0$.

(2) Let $\mathrm{q}(s) = (x(s), y(s)) : [0, L] \to \mathbb{R}^2$ be a unit speed planar curve with signed curvature $\kappa_\pm(s)$ and $\mathrm{q}^*(s) = x(s)e_1 + y(s)e_2 + \mathrm{x}$ another planar curve where e_1, e_2 is a positively oriented orthonormal basis and x a point.
   (a) Show that q^* is a unit speed curve with curvature $\kappa_\pm^*(s) = \kappa_\pm(s)$.
   (b) Show that a planar unit speed curve with the same curvature as q is of the form q^*.

(3) Compute the signed curvature of the logarithmic spiral

$$ae^{bt}(\cos t, \sin t).$$

(4) Compute the signed curvature of the spiral of Archimedes:

$$(a + bt)(\cos t, \sin t).$$

(5) Show that if a planar unit speed curve $\mathrm{q}(s)$ satisfies:

$$\kappa_\pm(s) = \frac{1}{es + f}$$

for constants $e, f > 0$, then it is a logarithmic spiral.

(6) Show that a planar curve is part of a circle if all its normal lines pass through a fixed point.

(7) Show that $\kappa_\pm \frac{ds}{dt} = \det\begin{bmatrix} \mathrm{T} & \frac{d\,\mathrm{T}}{dt} \end{bmatrix}$.

(8) Show that

$$\mathrm{q}(t)=\left(\int_0^t \cos\left(\frac{u^2}{2}\right)du, \int_0^t \sin\left(\frac{u^2}{2}\right)du\right),\ t\in\mathbb{R}$$

is parametrized by arclength and that $\kappa_\pm(t)=t$.

(9) Show that

$$\mathrm{q}(s)=\left(\int_{s_0}^s \cos(\phi(u))\,du, \int_{s_0}^s \sin(\phi(u))\,du\right),$$

is a unit speed curve with $\kappa_\pm=\frac{d\phi}{ds}$.

(10) Let $\mathrm{q}(t)=r(t)(\cos t,\sin t)$. Show that the speed satisfies

$$\left(\frac{ds}{dt}\right)^2=\left(\frac{dr}{dt}\right)^2+r^2$$

and the curvature

$$\kappa_\pm=\frac{2\left(\frac{dr}{dt}\right)^2+r^2-r\frac{d^2r}{dt^2}}{\left(\left(\frac{dr}{dt}\right)^2+r^2\right)^{\frac{3}{2}}}.$$

Parametrize the curve $\left(1-x^2\right)x^2=y^2$ in this way and compute its curvature. Note that such a parametrization won't be valid for all t.

(11) For a planar unit speed curve $\mathrm{q}(s)$ consider the parallel curve

$$\mathrm{q}_\epsilon=\mathrm{q}+\epsilon\,\mathrm{N}_\pm$$

for some fixed ϵ.

(a) Show that this curve is regular as long as $\epsilon\kappa_\pm\neq 1$.

(b) Show that the curvature is

$$\frac{\kappa_\pm}{|1-\epsilon\kappa_\pm|}.$$

(12) If a curve in $\mathbb{R}^2$ is given as a graph $y=f(x)$ show that the curvature is given by

$$\kappa_\pm=\frac{f''}{\left(1+(f')^2\right)^{\frac{3}{2}}}.$$

(13) Assume a planar curve is given as a level set $F(x,y)=c$ where $\nabla F\neq 0$ everywhere along the curve. We orient and parametrize the curve so that $\mathrm{v}=\left(-\frac{\partial F}{\partial y},\frac{\partial F}{\partial x}\right)$.

(a) Show that the signed normal is given by

$$\mathrm{N}_\pm=-\frac{\nabla F}{|\nabla F|}.$$

(b) Use the chain rule to show that the acceleration is

$$\begin{aligned}\mathrm{a} &= \begin{bmatrix} -\frac{\partial^2 F}{\partial x\partial y} & -\frac{\partial^2 F}{\partial y^2} \\ \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial y\partial x}\end{bmatrix}\begin{bmatrix} -\frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial x}\end{bmatrix} \\ &= \left[\frac{\partial\,\mathrm{v}}{\partial(x,y)}\right][\mathrm{v}].\end{aligned}$$

(c) Show that

$$\begin{aligned}
\kappa_{\pm} &= \frac{1}{|\nabla F|^3}\begin{bmatrix} -\frac{\partial F}{\partial x} & -\frac{\partial F}{\partial y}\end{bmatrix}\begin{bmatrix} -\frac{\partial^2 F}{\partial x \partial y} & -\frac{\partial^2 F}{\partial y^2} \\ \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial y \partial x}\end{bmatrix}\begin{bmatrix} -\frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial x}\end{bmatrix} \\
&= \frac{1}{|\nabla F|^3}\begin{bmatrix} -\frac{\partial F}{\partial y} & \frac{\partial F}{\partial x}\end{bmatrix}\begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} \\ \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial y^2}\end{bmatrix}\begin{bmatrix} -\frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial x}\end{bmatrix} \\
&= -\frac{1}{|\nabla F|^3}\det\begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial F}{\partial x} \\ \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} & 0\end{bmatrix}.
\end{aligned}$$

(14) (Jerrard, 1961) With notation as in the previous exercise show that

$$\kappa_{\pm} = \operatorname{div}\frac{\nabla F}{|\nabla F|}.$$

(15) Compute the curvature of $\left(1-x^2\right)x^2 = y^2$ at the points where the above formula works. What can you say about the curvature at the origin where the curve intersects itself.

(16) Consider a unit speed curve q. Show that if $\kappa_{\pm}(t_0) = 0$ and $\kappa'_{\pm}(t_0) \neq 0$, then the curve crosses the tangent line at t_0, i.e., the curve has an inflection point. Hint: Calculate the first three derivatives of $f(s) = \mathrm{q}(s) \cdot \mathrm{N}_{\pm}(t_0)$.

(17) Consider a unit speed curve q. Show that if $\kappa_{\pm}(t_0) \neq 0$ and $\kappa'_{\pm}(t_0) \neq 0$, then the curve crosses *osculating circle*, i.e., the circle that at t_0 has the same unit tangent and signed curvature. Thus the curve is on the inside on one side and the outside on the other. Hint: Calculate the first three derivatives of $f(s) = |\mathrm{q}(s) - \mathrm{c}|^2$, where c is the center of the circle.

(18) Compute the curvature of the cissoid of Diocles $x\left(x^2+y^2\right) = 2Ry^2$.

(19) Compute the curvature of the conchoid of Nicomedes $\left(x^2+y^2\right)(y-b)^2 - R^2y^2 = 0$.

(20) Consider a unit speed curve $\mathrm{q}(s)$ with non-vanishing curvature and use the notation $\frac{df}{ds} = f'$. Show that q satisfies the third order equation

$$\mathrm{q}''' - \frac{\kappa'_{\pm}}{\kappa_{\pm}}\mathrm{q}'' + \kappa_{\pm}^2\mathrm{q}' = 0.$$

(21) Show that the curvature of the evolute q^* of a unit speed curve $\mathrm{q}(s)$ satisfies

$$\frac{1}{\kappa_{\pm}^*} = \frac{1}{2}\frac{d}{ds}\left(\frac{1}{\kappa_{\pm}^2}\right).$$

(22) (Huygens, 1673) Consider the cycloid

$$\mathrm{q}(t) = R(t + \sin t, 1 + \cos t).$$

It traces a point on a circle of radius R that rolls along the x-axis. Any curve that is constructed by tracing a point on a circle rolling along a line is called a cycloid (see also section 1.3 exercise 15).

(a) Show that the signed curvature is given by

$$\kappa_{\pm} = \frac{-1}{2R\sqrt{2(1+\cos t)}}.$$

(b) Show that the evolute is also a cycloid.

(c) Show that any curve that satisfies

$$(s-a)^2 + \frac{1}{\kappa^2} = 16R^2$$

for a constant $a \in \mathbb{R}$ is a cycloid. In other words cycloids are circles centered on the first axis in $\left(s, \frac{1}{\kappa_\pm}\right)$ coordinates.

(d) Show that any cycloid is the involute of a cycloid.

(23) (Newton, 1671 and Huygens, 1673) Consider a regular planar curve $\mathrm{q}(t)$ with $\kappa_\pm(t_0) \neq 0$. Let $l(t)$ denote the normal line to q at $\mathrm{q}(t)$.

(a) Show that $l(t)$ and $l(t_0)$ are not parallel for t near t_0.

(b) Let $\mathrm{x}(t)$ denote the intersection of $l(t)$ and $l(t_0)$. Show that $\lim_{t\to t_0}\mathrm{x}(t)$ exists and denote this limit $\mathrm{c}(t_0)$.

(c) Show that

$$(\mathrm{c}(t_0) - \mathrm{q}(t_0)) \cdot \mathrm{N}(t_0) = \frac{1}{\kappa_\pm(t_0)}.$$

Note that the left hand side is the signed distance from $\mathrm{c}(t_0)$ to $\mathrm{q}(t_0)$ along the normal through $\mathrm{q}(t_0)$. The circle of radius $\left|\frac{1}{\kappa_\pm(t_0)}\right|$ centered at $\mathrm{c}(t_0)$ is the circle that best approximates the curve at $\mathrm{q}(t_0)$, it is called the *osculating* circle.

(d) Show that the curve $\mathrm{c}(t)$ is the evolute of $\mathrm{q}(t)$.

(24) Find the evolute of $y = x^2$ and show that when $x = 0$ it is tangent to y-axis.

(25) Find the evolute of $y = x^3$ and show that it is asymptotic to the y-axis as $x \to 0^\pm$.

(26) Show that the evolute of $y^2 = 2px$ is $27py^2 = 8(x-p)^3$.

(27) Show that evolute of the astroid (see also section 1.2 exercise 14) $|x|^{\frac{2}{3}} + |y|^{\frac{2}{3}} = 1$ is the astroid $|x+y|^{\frac{2}{3}} + |x-y|^{\frac{2}{3}} = 2$.

(28) Show that the evolute of the ellipse $(a\cos t, b\sin t)$ is the astroid $\left(\frac{a^2-b^2}{a}\cos^3 t, \frac{b^2-a^2}{b}\sin^3 t\right)$.

(29) (Newton, 1671, but the idea is much older for specific curves. Kepler considered it well-known.) Consider a regular planar curve $\mathrm{q}(t)$. For 3 "consecutive" values $t-\epsilon < t < t+\epsilon$ let $\mathrm{c}(t,\epsilon)$ denote the center of the unique circle that goes through the three points $\mathrm{q}(t-\epsilon)$, $\mathrm{q}(t)$, $\mathrm{q}(t+\epsilon)$ with $\mathrm{c}(t,\epsilon) = \infty$ if the points lie on a line.

(a) Show that $\mathrm{c}(t,\epsilon)$ is the point of intersection between the two normal lines to the segments between $\mathrm{q}(t)$ and $\mathrm{q}(t\pm\epsilon)$ that pass through the midpoint of these segments.

(b) Show that $\mathrm{q}(t-\epsilon)$, $\mathrm{q}(t)$, $\mathrm{q}(t+\epsilon)$ do not lie on a line for small ϵ if $\kappa_\pm(t) \neq 0$.

(c) Show that $\mathrm{c}(t,\epsilon)$ lies on the normal line through some point $\mathrm{q}(t_0)$ where $t_0 \in (t-\epsilon, t+\epsilon)$. Hint: Show that there is a point $\mathrm{q}(t_0)$ on the curve closest to $\mathrm{c}(t,\epsilon)$ for some $t_0 \in (t-\epsilon, t+\epsilon)$ and use that as the desired point.

(d) Show that

$$\lim_{\epsilon\to 0}(\mathrm{c}(t,\epsilon) - \mathrm{q}(t)) = \lim_{\epsilon\to 0}(\mathrm{c}(t,\epsilon) - \mathrm{q}(t_0)) = \frac{1}{\kappa_\pm(t)}\mathrm{N}_\pm(t).$$

(30) (Normal curves) Consider a family of lines in the (x,y)-plane parametrized by t:

$$F(x,y,t) = a(t)x + b(t)y + c(t) = 0.$$

A *normal curve* or *envelope* to this family is a curve $(x(t), y(t))$ such that its tangent lines are precisely the lines of this family.

(a) Show that such a curve exists and can be determined by the equations:

$$\begin{aligned} F = a(t)x + b(t)y + c(t) &= 0, \\ \frac{\partial F}{\partial t} = \dot{a}(t)x + \dot{b}(t)y + \dot{c}(t) &= 0, \end{aligned}$$

when the Wronskian

$$\det \begin{bmatrix} a & b \\ \dot{a} & \dot{b} \end{bmatrix} \neq 0.$$

(b) Show that for fixed x_0, y_0 the number of solutions or roots to the equation $F(x_0, y_0, t) = 0$ corresponds to the number of tangent lines to the normal curve that pass through (x_0, y_0).

(c) Consider the cases where $a = 1$, $b = t$, $c = t^n$, $n = 2, 3, 4, \ldots$

(i) Show that the normal curve satisfies:

$$D = (-1)^n (n-1)^{n-1} y^n - n^n x^{n-1} = 0,$$

here D is the *discriminant* of the equation.

(ii) Determine the number of roots in relation to how (x_0, y_0) is placed relative to the normal curve.

(iii) Show that roots with multiplicity only occur when (x_0, y_0) is on the normal curve.

2.2. The Rotation Index

We now turn to a more geometric interpretation of the signed curvature.

THEOREM 2.2.1. *For a regular curve the angle between the unit tangent and the x-axis is an anti-derivative of the signed curvature with respect to arclength.*

PROOF. We start with an analysis of the problem. Assume that we have a parametrization of the unit tangent by using the angle to the first axis (we don't know yet that it is possible to select such a parametrization):

$$\begin{aligned} \mathrm{T}(t) &= (\cos\theta(t), \sin\theta(t)), \\ \mathrm{N}_{\pm}(t) &= (-\sin\theta(t), \cos\theta(t)). \end{aligned}$$

In this case

$$\frac{ds}{dt}\kappa_{\pm}\,\mathrm{N}_{\pm} = \frac{d\,\mathrm{T}}{dt} = \frac{d\theta}{dt}\,\mathrm{N}_{\pm}.$$

So we should be able to declare that θ is an antiderivative of $\frac{ds}{dt}\kappa_{\pm}$. Note that as long as the signed curvature is non-negative this is consistent with the interpretation of θ as an arclength parameter for T.

To verify that such a choice works, define

$$\begin{aligned} \theta(t_1) &= \theta_0 + \int_a^{t_1} \frac{ds}{dt}\kappa_{\pm} dt, \text{ where} \\ \mathrm{T}(a) &= (\cos\theta_0, \sin\theta_0) \end{aligned}$$

and consider the orthonormal unit fields

$$\begin{aligned} U &= (\cos\theta(t), \sin\theta(t)), \\ V &= (-\sin\theta(t), \cos\theta(t)). \end{aligned}$$

They are clearly related by

$$\frac{dU}{dt} = \frac{d\theta}{dt} V.$$

If we can show that $\mathrm{T} \cdot U \equiv 1$, then it follows that $\mathrm{T} = U$. Our choice of θ_0 forces the dot product to be 1 at $t = a$. To show that it is constant we show that the derivative vanishes

$$\begin{aligned}
\frac{d}{dt}(\mathrm{T} \cdot U) &= \frac{d\,\mathrm{T}}{dt} \cdot U + \mathrm{T} \cdot \frac{dU}{dt} \\
&= \frac{ds}{dt} \kappa_{\pm} \mathrm{N}_{\pm} \cdot U + \frac{d\theta}{dt} \mathrm{T} \cdot V \\
&= \frac{ds}{dt} \kappa_{\pm} (\mathrm{N}_{\pm} \cdot U + \mathrm{T} \cdot V) \\
&= 0,
\end{aligned}$$

where the last equality follows by noting that if $\mathrm{T} = (f, g)$, then $\mathrm{N}_{\pm}(-g, f)$ so

$$\mathrm{N}_{\pm} \cdot U + \mathrm{T} \cdot V = -g \cos + f \sin + - f \sin + g \cos = 0.$$

In other words, the two inner products define complementary angles. $\square$

Definition 2.2.2. The *total curvature* of a curve $\mathrm{q} : [a, b] \to \mathbb{R}^2$ is defined as

$$\int_a^b \kappa_{\pm} \frac{ds}{dt} dt.$$

When we reparametrize the curve by arclength this simplifies to

$$\int_0^L \kappa_{\pm} ds.$$

The total curvature measures the total change in the tangent since the curvature measures the infinitesimal change of the tangent.

The ancient Greeks actually used a similar idea to calculate the angle sum in a convex polygon. Specifically, the sum of the exterior angles in a polygon adds up to 2π. This is because we can imagine the tangent line at each vertex jumping from one side to the next and while turning measuring the angle it is turning. When we return to the side we started with we have completed a full circle. When the polygon has n vertices this gives us the formula $(n-2)\pi$ for the sum of the interior angles.

A similar result holds for closed planar curves as $\mathrm{T}(a) = \mathrm{T}(b)$ for such a curve.

Proposition 2.2.3. *The total curvature of a planar closed curve is an integer multiple of* 2π.

The integer is called the *rotation index* of the curve:

$$i_{\mathrm{q}} = \frac{1}{2\pi} \int_a^b \kappa_{\pm} \frac{ds}{dt} dt.$$

We can more generally define the *winding number* of a closed unit curve $\mathrm{t} : [a, b] \to S^1 \subset \mathbb{R}^2$. Being closed now simply means that $\mathrm{t}(a) = \mathrm{t}(b)$. The idea is to measure the number of times such a curve winds or rotates around the circle. The specific formula is very similar. First construct the positively oriented normal $\mathrm{n}(t)$ to $\mathrm{t}(t)$, i.e. the unit vector perpendicular to $\mathrm{t}(t)$ such that $\det \begin{bmatrix} \mathrm{t}(t) & \mathrm{n}(t) \end{bmatrix} = 1$

and then check the change of t against n. Note that as t is a unit vector its derivative is proportional to n. The winding number is given by

$$w_{\mathrm{t}} = \frac{1}{2\pi}\int_a^b \frac{d\mathrm{t}}{dt}\cdot \mathrm{n}dt.$$

With this definition

$$i_{\mathrm{q}} = w_{\mathrm{T}}.$$

Proposition 2.2.4. *The winding number of a closed unit curve is an integer. Moreover, it doesn't change under small changes in* t.

Proof. The results holds for all continuous curves, but as we've used derivatives to define the winding number we have to assume that it is smooth. However, the proof works equally well if we assume that the curve is piecewise smooth.

As above define

$$\theta(t_0) = \theta_0 + \int_a^{t_0} \frac{d\mathrm{t}}{dt}\cdot \mathrm{n}dt,$$

where

$$\mathrm{t}(a) = (\cos\theta_0, \sin\theta_0).$$

We can now use the same argument to conclude that

$$\mathrm{t}(t) = (\cos\theta(t), \sin\theta(t)).$$

Consequently,

$$w_{\mathrm{t}} = \frac{1}{2\pi}\int_a^b \frac{d\mathrm{t}}{dt}\cdot \mathrm{n}dt = \frac{\theta(b)-\theta(a)}{2\pi}.$$

So when the curve is closed it follows that $\theta(b) = \theta(a) + 2\pi n$ for some $n \in \mathbb{Z}$ and that $w_{\mathrm{t}} = n$ is an integer.

Next suppose that we have two $\mathrm{t}_1, \mathrm{t}_2$ parametrized on the same interval $[a,b]$ such that

$$|\mathrm{t}_1 - \mathrm{t}_2| \leq \epsilon < 2.$$

If in addition their derivatives are also close and bounded then it is not hard to see directly that the winding numbers are close. However, as they are integers, the only way in which they can be close is if they agree.

To prove the result without assumptions about derivatives we start with the crucial observation that if

$$|\theta_1 - \theta_2| < \pi,$$

then

$$|\theta_1 - \theta_2| < \frac{\pi}{2}\left|(\cos\theta_1, \sin\theta_1) - (\cos\theta_2, \sin\theta_2)\right|.$$

In other words, if the difference in angles between two points on the circle is less that π then the difference in angles is bounded by a uniform multiple of the cord length between the points.

Now assume that we have

$$\begin{aligned}
\mathrm{t}_1(t) &= (\cos\theta_1(t), \sin\theta_1(t)),\\
\mathrm{t}_2(t) &= (\cos\theta_2(t), \sin\theta_2(t)),
\end{aligned}$$

with

$$|\theta_1(a) - \theta_2(a)| < \pi,$$

then we claim that
$$\left|\theta_1(t)-\theta_2(t)\right|<\frac{\pi}{2}\left|\mathrm{t}_1(t)-\mathrm{t}_2(t)\right|$$
for all t.

We know the claim holds for $t=a$ and as all the functions are continuous the set of parameters t that satisfy this condition is open (it is a strict inequality). Next we can show that this set is also closed. To see this assume that the inequality holds for t_n and that $t_n\to t$. We have
$$\left|\theta_1(t_n)-\theta_2(t_n)\right|<\frac{\pi}{2}\left|\mathrm{t}_1(t_n)-\mathrm{t}_2(t_n)\right|\le\frac{\pi}{2}\epsilon,$$
so it follows from continuity that
$$\left|\theta_1(t)-\theta_2(t)\right|\le\frac{\pi}{2}\epsilon<\pi.$$
This shows that
$$\left|\theta_1(t)-\theta_2(t)\right|<\frac{\pi}{2}\left|\mathrm{t}_1(t)-\mathrm{t}_2(t)\right|.$$

It now follows that
$$\begin{aligned}\left|w_{\mathrm{t}_1}-w_{\mathrm{t}_2}\right| &\le \frac{1}{2\pi}\left|(\theta_1(b)-\theta_1(a))-(\theta_2(b)-\theta_2(a))\right|\\ &\le \frac{1}{2\pi}\left|(\theta_1(b)-\theta_2(b))-(\theta_1(a)-\theta_2(a))\right|\\ &\le \frac{1}{2\pi}\left|(\theta_1(b)-\theta_2(b))\right|+\frac{1}{2\pi}\left|(\theta_1(a)-\theta_2(a))\right|\\ &\le \frac{1}{2}\epsilon<1.\end{aligned}$$
This shows that the winding numbers are equal. □

The next theorem is often called the Umlaufsatz (going around theorem). It is universally credited to H. Hopf, however, the name and theorem is due to A. Ostrowski. Ostrowski's papers were in fact published in the same journal in the same year as Hopf's paper. Hopf's proof was meant as a shorter more elegant version of Ostrowski's far longer version. Ostrowski himself also refers to the theorem as Rolle's theorem.

Theorem 2.2.5 (Ostrowski, 1935). *A simple closed curve has rotation index* ±1.

Proof. (Hopf, 1935) We assume that we have a simple closed curve $\mathrm{q}(s):[0,l]\to\mathbb{R}^2$ that is parametrized by arclength. Moreover, after possibly rotating and translating the curve we'll assume that $\mathrm{q}(0)=(0,0)$, $\mathrm{T}(0)=(\pm1,0)$, and $y(s)\ge0$ for all s. The idea is to create a family of unit vectors on a triangle where $0\le s\le t\le l$.
$$\mathrm{T}(s,t)=\begin{cases}\mathrm{T}(s) & s=t,\\ -\mathrm{T}(0) & s=0,\ t=l,\\ \frac{\mathrm{q}(t)-\mathrm{q}(s)}{|\mathrm{q}(t)-\mathrm{q}(s)|} & \text{for all other } s<t.\end{cases}$$
Since the curve is simple, closed, and smooth this is a well-defined function whose values are aways unit vectors. If we select any simple path in this triangle that passes from $(0,0)$ to (l,l) then T will wind around the unit circle and end up where it began as $\mathrm{T}(0,0)=\mathrm{T}(l,l)$. Moreover, if we make a slight change in this path it will wind around the same number of times. Along the diagonal the number of

windings is the rotation index of the curve. However, if we move up the y-axis and then along the upper edge of the triangle, then we are first following $\mathrm{T}(0,t)=\frac{\mathrm{q}(t)}{|\mathrm{q}(t)|}$ and then $\mathrm{T}(s,l)=\frac{\mathrm{q}(l)-\mathrm{q}(s)}{|\mathrm{q}(l)-\mathrm{q}(s)|}$. Assume that $\mathrm{T}(0)=(1,0)$, then $\mathrm{T}(0,t)$ rotates precisely π from right to left while it points upwards as q lies in the upper half plane, and $\mathrm{T}(s,l)$ rotates π from left to right while pointing downwards. Thus this rotation is precisely 2π. This shows that q also has rotation index 1. When instead $\mathrm{T}(0)=(-1,0)$ the rotation index is -1. □

Definition 2.2.6. The *total absolute curvature* is defined as

$$\int_a^b \kappa \frac{ds}{dt}dt=\int_a^b |\kappa_\pm| \frac{ds}{dt}dt.$$

Exercises

(1) Let $\mathrm{q}(t)=r(t)(\cos(nt),\sin(nt))$ where is $t\in[0,2\pi]$, $n\in\mathbb{Z}$, and $r(t)>0$ is 2π-periodic. Show that $i_{\mathrm{q}}=n$, by showing that its unit tangent has the same winding number as the curve $(-\sin(nt),\cos(nt))$.
(2) Draw a picture of the curve $\left(1-x^2\right)x^2=y^2$. Use this to show that the index is zero and that the total absolute curvature is $>2\pi$.
(3) Let $\mathrm{q}(s):[0,L]\to\mathbb{R}^2$ be a closed planar curve parametrized by arclength and consider the parallel curves $\mathrm{q}_\epsilon(s)=\mathrm{q}(s)-\epsilon\,\mathrm{N}_\pm(s)$.
   (a) Show that $\frac{ds_\epsilon}{ds}=1+\epsilon\kappa_\pm$, where s_ϵ is the arclength parameter for q_ϵ.
   (b) Show that $L(\mathrm{q}_\epsilon)=L(\mathrm{q})+\epsilon 2\pi i_{\mathrm{q}}$.
   (c) When q is simple and runs counterclockwise show that the area A_ϵ enclosed by q_ϵ is related the the area A enclosed by q by

$$A_\epsilon=A+\epsilon L+\epsilon^2\pi.$$

   Hint: Use Green's theorem as in exercise 12 from section 2.3.
(4) Let $\mathrm{q}(s):[0,L]\to\mathbb{R}^2$ be a unit speed curve that is piecewise smooth, i.e., the domain can be subdivided

$$0=a_1<a_2<\cdots<a_{k+1}=L$$

such that the curve is smooth on each interval $[a_i,a_{i+1}]$, $i=1,...,k$. The *exterior angle* $\theta_i\in[-\pi,\pi]$ at a_i is defined by

$$\begin{aligned}\cos\theta_i &= \mathrm{T}\left(a_i^-\right)\cdot\mathrm{T}\left(a_i^+\right),\\ \sin\theta_i &= \mathrm{N}_\pm\left(a_i^-\right)\cdot\mathrm{T}\left(a_i^+\right),\end{aligned}$$

where

$$\mathrm{T}\left(a_i^\pm\right)=\frac{d\,\mathrm{q}}{ds^\pm}(a_i)=\lim_{h\to0}\frac{\mathrm{q}(a_i\pm h)-\mathrm{q}(a_i)}{\pm h}$$

and $\mathrm{N}_\pm$ defined as the corresponding signed normal.
   (a) If q is closed show that

$$\int_0^L \kappa_\pm ds+\sum_{i=1}^k\theta_i=i_{\mathrm{q}}2\pi$$

   for some $i_{\mathrm{q}}\in\mathbb{Z}$.

(b) If q is both closed and simple show that $i_{\mathrm{q}} = \pm 1$. Hint: For each i replace q on some small interval $[a_i - \epsilon, a_i + \epsilon]$ with a smooth curve q^* such that

$$\int_{a_i-\epsilon}^{a_i+\epsilon} \kappa_\pm^* ds^* = \theta_i + \int_{a_i-\epsilon}^{a_i+\epsilon} \kappa_\pm ds.$$

(c) Show that the sum of the exterior angles in a polygon is 2π if the polygon is oriented appropriately.

(5) Let

$$\mathrm{q}(t) = (1 + a\cos t)(\cos t, \sin t),\ t \in [0, 2\pi].$$

(a) Show that this is a simple curve when $|a| < 1$ and intersects it self once when $|a| > 1$. Hint: Show that if $r(t) > 0$, then $r(t)(\cos t, \sin t)$ defines a simple curve. When $r(t)$ changes sign investigate what happens when it vanishes.

(b) Show that

$$\frac{d\theta}{dt} = 1 + \frac{a(a + \cos t)}{1 + a^2 + 2a\cos t}$$

and conclude that

$$\int_0^{2\pi} \frac{a(a+\cos t)}{1 + a^2 + 2a\cos t} dt = \begin{cases} 0 & |a| < 1, \\ 2\pi & |a| > 1. \end{cases}$$

(6) Show that any closed planar curve satisfies

$$\int_a^b \kappa \frac{ds}{dt} dt \geq 2\pi.$$

(7) Show that by selecting a very flat ∞ shape where the tangents at the intersection are close to the x-axis we obtain examples with rotation index 0 and total absolute curvature close to 2π.

(8) Let $\mathrm{q} : [0, L] \to \mathbb{R}^2$ be a closed curve parametrized by arclength. Show that if $\int_0^L \kappa ds = 2\pi$, then $\kappa_\pm$ cannot change sign and the rotation index is ± 1. In section 2.4 we will show that this implies that the curve is simple as well. Hint: Show that for a general curve $|i_{\mathrm{q}}| \leq \frac{1}{2\pi}\int_0^L \kappa ds$ with equality only when $\kappa = \kappa_\pm$ or $\kappa = -\kappa_\pm$ everywhere.

(9) Let $\mathrm{q}(t)$, $t \in [a, b]$ be a regular planar curve and $\theta(t) \in [\theta_0, \theta_1]$ an arclength parameter for T. Define $v(t)$ as the distance from the origin to the tangent line through $\mathrm{q}(t)$.

(a) Show that

$$v(t) = -\mathrm{q}(t) \cdot \mathrm{N}_\pm(t).$$

(b) Show by an example (e.g., a straight line) that q is not necessarily a function of θ.

(c) Define the curve

$$\mathrm{q}^*(\theta) = \frac{dv}{d\theta}\mathrm{T} - v\,\mathrm{N}_\pm = \frac{dv}{d\theta}(\cos\theta, \sin\theta) - v(-\sin\theta, \cos\theta)$$

and show that

$$\frac{d\,\mathrm{q}^*}{d\theta} = \left(\frac{d^2 v}{d\theta^2} + v\right)(\cos\theta, \sin\theta).$$

(d) Show that when q^* is a regular curve then it is a reparametrization of q.

(e) Under that assumption show further that

$$v+\frac{d^2v}{d\theta^2}=\frac{1}{\kappa},$$

$$L(\mathrm{q})=\int_{\theta_0}^{\theta_1} v(\theta)\,d\theta.$$

2.3. Three Interesting Results

In this section we establish three interesting results for closed planar curves. The only result that will be used again from time to time is the Jordan curve theorem 2.3.1.

Theorem 2.3.1 (Jordan Curve Theorem). *A simple, closed planar curve divides the plane into two regions one that is bounded and one that is unbounded.*

2.3.1. The Four Vertex Theorem.

Definition 2.3.2. A *vertex* of a curve is a point on the curve where the curvature has a local maximum or a local minimum.

Theorem 2.3.3 (Mukhopadhyaya, 1909 and Kneser, 1912). *A simple closed curve has at least 4 vertices.*

We start with the following observation.

Proposition 2.3.4. *If a curve* q *is tangent to a circle with their unit tangents being the same and lies inside (resp. outside) the circle, then its curvature is larger (resp. smaller) than or equal to the curvature of the circle at the points where they are tangent.*

Proof. Compare also proposition 1.3.6 for the case when the curve q is tangent to the circle of radius R centered at c at $s=s_0$ and satisfies

$$|\mathrm{q}(s)-\mathrm{c}|^2\leq R^2 \text{ and } |\mathrm{q}(s_0)-\mathrm{c}|^2=R^2.$$

Here we focus on the case where the curve lies outside the circle as it is unique to planar curves: The function $s\mapsto|\mathrm{q}(s)-\mathrm{c}|^2$ has a (local) minimum at $s=s_0$. Thus its derivative at s_0 vanishes. This is simply the fact that the curve is tangent to the circle. Moreover, the second derivative is nonnegative. Both the circle and the curve are parametrized to have the same unit tangents where they touch and we can further assume that the circle is parametrized to run counterclockwise. Consequently, their signed normals are equal and point inward. This normal is

$$\mathrm{N}_\pm(s_0)=-\frac{\mathrm{q}(s_0)-\mathrm{c}}{|\mathrm{q}(s_0)-\mathrm{c}|}=-\frac{\mathrm{q}(s_0)-\mathrm{c}}{R}.$$

The second derivative of $s\mapsto|\mathrm{q}(s)-\mathrm{c}|^2$ is

$$2+2(\mathrm{q}-\mathrm{c})\cdot\ddot{\mathrm{q}}=2+2\kappa_\pm\,\mathrm{N}_\pm\cdot(\mathrm{q}-\mathrm{c}).$$

Therefore, at s_0 we have

$$\begin{aligned}0&\leq 2+2\kappa_\pm(s_0)\,\mathrm{N}_\pm(s_0)\cdot(\mathrm{q}(s_0)-\mathrm{c})\\&= 2+2\kappa_\pm(s_0)\,\mathrm{N}_\pm(s_0)\cdot(-R\,\mathrm{N}_\pm(s_0))\\&= 2-2R\kappa_\pm(s_0).\end{aligned}$$

This implies our claim. □

We are now ready to prove the four vertex theorem. Mukhopadhyaya proved this result for simple planar curves with strictly positive curvature and a few years later Kneser proved the general version, apparently without knowledge of Mukhopadhyaya's earlier contribution. An excellent account of the history of this fascinating result can be found here: Four Vertex Theorem

Proof. (Osserman, 1985) Select the circle of smallest radius R circumscribing the simple closed curve. The points of contact between this circle and the curve cannot lie on one side of a diagonal of the circle. If they did, then it'd be possible to slide the circle in the orthogonal direction to the diagonal until it doesn't hit the curve. We could then find a circle of smaller radius that contains the curve. This means that we can find points $q_1, ..., q_{k+1}$ of contact where $q_{k+1} = q_1$ and either $k = 2$ and q_1 and q_2 are antipodal, or $k > 2$ and any two consecutive points q_i and q_{i+1} lie one one side of a diagonal. Note there might be more points of contact.

Now orient both circle and curve so that their normals always point inside. At points of contact where the tangent lines are equal, the normal vectors must then also be equal, as the curve is inside the circle. This forces the unit tangent vectors to be equal.

First observe that the curvature at these k points is $\geq R^{-1}$.

If the curve coincides with the circle between two consecutive points of contact q_i and q_{i+1}, then the curvature is constant and we have infinitely many vertices. Otherwise there will be a point q on the curve between q_i and q_{i+1} that is inside the circle. Then we can select a circle of radius $> R$ that passes through q_i and q_{i+1} and still contains q in its interior. Now slide this new circle orthogonally to the cord between q_i and q_{i+1} until the part of the curve between q_i and q_{i+1} lies outside the circle but still touches it somewhere. At this place the curvature will be $< R^{-1}$.

This shows that we can find k points where the curvature is $\geq R^{-1}$ and k points between these where the curvature is $< R^{-1}$. This implies that there must be at least k local maxima for κ where the curvature is $\geq R^{-1}$ and between each two consecutive local maxima a minimum where the curvature is $< R^{-1}$. Note that the maxima and minima don't have to be at the points of contact. Thus we have found $2k$ vertices. $\square$

2.3.2. The Isoperimetric Inequality. The *isoperimetric ratio* of a simple closed planar curve q is L^2/A where L is the perimeter, i.e., length of q, and A is the area of the interior. We say that q *minimizes the isoperimetric ratio* if L^2/A is as small as it can be.

The isomerimetric inequality asserts that the isoperimetric ratio always exceeds 4π and is only minimal for circles. This will be established in the next theorem using a very elegant proof that does not assume the existence of a curve that realizes this ratio. Steiner in the 1830s devised several descriptive proofs of the isoperimetric inequality assuming that such minimizers exist. It is, however, not so simple to show that such curves exist as Dirichlet repeatedly pointed out to Steiner. Some of Steiner's ideas will be explored in the exercises.

The isoperimetric inequality would seem almost obvious and has been investigated for millennia. In fact a closely related problem, known as Dido's problem, appears in ancient legends. Dido founded Carthage and was faced with the problem of enclosing the largest possible area for the city with a long string (called a length of hide as the string had to be cut from a cow hide). However, the city was to be placed along the shoreline and so it was only necessary to enclose the city on the land side. In mathematical terms we can let the shore line be a line, and the curve that will enclose the city on the land side is a curve that begins and ends on the line and otherwise stays on one side of the line. It is not hard to imagine that a semicircle whose diameter is on the given line yields the largest area for a curve of fixed length.

THEOREM 2.3.5. *The isoperimetric inequality states that if a simple closed curve bounds an area A and has circumference L, then*

$$L^2 \geq 4\pi A.$$

Moreover, equality can only happen when the curve is a circle.

PROOF. (Knothe, 1957) We give a very direct proof using Green's theorem in the form of the divergence theorem. Unlike many other proofs, this one also easily generalizes to higher dimensions.

Consider a simple closed curve q of length L that can be parametrized by arclength. The domain of area A is then the interior of this curve. Let the domain be denoted Ω. We wish to select a (Knothe) map $F : \Omega \to B(0,R)$ where $B(0,R)$ also has area A. More specifically we seek a map with the properties

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

and

$$\det DF = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} = 1.$$

Such a map can be constructed if we select $x(u_0)$ and $y(u_0,v_0)$ for a specific $(u_0,v_0) \in \Omega$ to satisfy

$$\text{area}(\{u < u_0\} \cap \Omega) = \text{area}(\{x < x(u_0)\} \cap B(0,R))$$

and

$$\text{area}(\{u < u_0\} \cap \{v < v_0\} \cap \Omega) = \text{area}(\{x < x(u_0)\} \cap \{y < y(u_0,v_0)\} \cap B(0,R)).$$

The choice of $B(0,R)$ together with the intermediate value theorem guarantee that we can construct this map. This map is area preserving as it is forced to map any rectangle in Ω to a region of equal area in $B(0,R)$. To see this note that it preserves the area of sets $\{u_0 \leq u < u_1\} \cap \Omega$ as they can be written as a difference of sets whose areas are preserved by definition of the map:

$$\{u_0 \leq u < u_1\} \cap \Omega = \{u < u_1\} \cap \Omega - (\{u < u_0\} \cap \Omega).$$

We then obtain the rectangle $[u_0,u_1) \times [v_0,v_1)$ by intersecting this strip with the set $\{v_0 \leq v < v_1\} \cap \Omega$. This rectangle is in turn written as a difference between two sets whose areas are preserved by the map:

$$\begin{aligned}
[u_0,u_1) \times [v_0,v_1) &= (\{u < u_1\} \cap \Omega - (\{u < u_0\} \cap \Omega)) \cap (\{v < v_1\} \cap \Omega - (\{v < v_0\} \cap \Omega)) \\
&= (\{u < u_1\} \cap \Omega - (\{u < u_0\} \cap \Omega)) \cap (\{v < v_1\} \cap \Omega) \\
&\quad - (\{u < u_1\} \cap \Omega - (\{u < u_0\} \cap \Omega)) \cap (\{v < v_0\} \cap \Omega).
\end{aligned}$$

The two conditions additionally force $\frac{\partial x}{\partial u} > 0$, $\frac{\partial y}{\partial v} > 0$. To prove the isoperimetric inequality we use Green's theorem in the form of the divergence theorem in the plane. The vector field is given by the map F. Note that the outward unit normal for Ω is the vector $-\mathrm{N}_{\pm}$ if the curve q runs counter clockwise. Using that $|F| \leq R$ we obtain:

$$\begin{aligned}\int_{\Omega} \operatorname{div} F du dv &= -\int_{0}^{L} F \cdot \mathrm{N}_{\pm}\, ds \\ &\leq RL.\end{aligned}$$

On the other hand the geometric mean $\sqrt{ab}$ is always smaller than the arithmetic mean $\frac{1}{2}(a+b)$ so we also have:

$$\begin{aligned}\operatorname{div} F &= \frac{\partial x}{\partial u} + \frac{\partial y}{\partial v} \\ &\geq 2\sqrt{\frac{\partial x}{\partial u}\frac{\partial y}{\partial v}} \\ &= 2.\end{aligned}$$

Consequently

$$2A \leq RL,$$

which implies

$$4A^2 \leq R^2 L^2.$$

Now we constructed $B(0,R)$ so that $A = \pi R^2$. So we obtain the isoperimetric inequality

$$4\pi A \leq L^2.$$

The equality case can only occur when we have equality in all of the above inequalities. In particular

$$\frac{\partial x}{\partial u} = \frac{\partial y}{\partial v}$$

everywhere, showing that

$$\frac{\partial x}{\partial u} = \frac{\partial y}{\partial v} = 1.$$

This tells us that the function takes the form: $F(u,v) = (u + u_0, v + g(u))$. We also used that $|F \cdot \mathrm{N}_{\pm}| \leq |F| \leq R$ when the function is restricted to the boundary curve. Thus we also have $F \circ \mathrm{q} = -R\,\mathrm{N}_{\pm}$, i.e.,

$$\mathrm{q} + (u_0, g(u(s))) = -R\,\mathrm{N}_{\pm},$$

where $\mathrm{q}(s) = (u(s), v(s))$. Differentiating with respect to s then implies that

$$(1 - R\kappa)\,\mathrm{T} = \left(0, \frac{\partial g}{\partial u}\frac{du}{ds}\right).$$

This means either that $1 = R\kappa$ or that q is constant in the first coordinate. In the latter case $\frac{du}{ds} = 0$, so it still follows that $1 = R\kappa$. Thus the curve has constant non-zero curvature which shows that it must be a circle. □

Remark 2.3.6. We've used without justification that the Knothe map is smooth so that we can take its divergence. This may however not be the case. The partial derivative $\frac{\partial x}{\partial u}$, when it exists, is equal to the sum of lengths of the intervals that make up the set $\{u = u_0\} \cap \Omega$. So if we assume that part of the boundary is a vertical line at $u = u_0$ and that the domain contains points both to the right and left of this line, then $\frac{\partial x}{\partial u}$ is not continuous at $u = u_0$.

To get around this issue one can assume that the domain is convex. Or in general that the boundary curve has the property that its tangent lines at points where the curvature vanishes are not parallel to the axes. The latter condition can generally be achieved by rotating the curve and appealing to Sard's theorem. Specifically, we wish to ensure that the normal $\mathrm{N}_{\pm}$ is never parallel to the axes at places where $\frac{d\,\mathrm{N}_{\pm}}{dt}=0$.

Alternately it is also possible to prove the divergence theorem under fairly weak assumptions about the derivatives of the function.

2.3.3. Counting Self Intersections.

Definition 2.3.7. For a line and a curve consider the points on the line where the curve is tangent to the line. This set will generally be empty. We say that the line is a *double tangent* if this set is not empty and not a segment of the line. Thus the curve will have contact with the line in at least two places but will not have contact with the line at all of the points in between these two points of contact.

When a curve is not too wild it is possible to relate double tangents and self intersections.

A *generic curve* is defined as a regular curve such that:

(1) Tangent lines cannot be tangent to the curve at more than 2 points.
(2) At self-intersection points the curve intersects itself twice.
(3) The curve only has a finite number of inflection points where the curvature changes sign.
(4) Finally, no point on the curve can belong to more than one of these categories of points.

For a generic curve T_+ is the number of tangent lines that are tangent to the curve in two places such that the curve lies on the same side of the tangent line at the points of contact. T_- is the number of tangent lines that are tangent to the curve in two places such that the curve lies on opposite sides of the tangent line at the points of contact. I is the number of inflection points, i.e., points where the curvature changes sign. D is the number of self-intersections (double points).

Theorem 2.3.8 (Fabricius-Bjerre, 1962). *For a generic closed curve we have*

$$2T_+-2T_--2D-I=0.$$

Proof. The proof proceeds by checking the number of intersections between the positive tangent lines $\mathrm{q}(t)+r\,\mathrm{v}(t)$, $r\geq 0$ and the curve as we move forwards along the curve. As we move along the curve this number will change but ultimately return to its initial value.

When we pass through an inflection point or a self-intersection this number will decrease by 1. When we pass a point that corresponds to a double tangent $T_{\pm}$ the change will be 0 or ± 2 with the sign being consistent with the type of tangent.

To keep track of what happens we subdivide the two types of double tangents into three categories denoted $\overset{\rightarrow\rightarrow}{T}_{\pm}, \overset{\rightarrow\leftarrow}{T}_{\pm}, \overset{\leftarrow\rightarrow}{T}_{\pm}$. Here $\overset{\rightarrow\rightarrow}{T}_{\pm}$ indicates that the tangent vectors at the double points have the same directions, $\overset{\rightarrow\leftarrow}{T}_{\pm}$ indicate that the tangent vectors at the double points have opposite directions but towards each other, and $\overset{\leftarrow\rightarrow}{T}_{\pm}$ indicate that the tangent vectors at the double points have opposite directions but away from each other.

For double tangents of the type $\overset{\leftarrow\rightarrow}{T}_{\pm}$ no intersections will be gained or lost as we pass through points of that type. For $\overset{\rightarrow\leftarrow}{T}_{\pm}$ the change is always ± 2 at both of the points of contact. For $\overset{\rightarrow\rightarrow}{T}_{\pm}$ the change is ± 2 for one of the points and 0 for the other. Thus as we complete one turn of the curve we must have

$$2\overset{\rightarrow\rightarrow}{T}_{+} + 4\overset{\rightarrow\leftarrow}{T}_{+} - 2\overset{\rightarrow\rightarrow}{T}_{-} - 4\overset{\rightarrow\leftarrow}{T}_{-} - I - 2D = 0.$$

We now reverse the direction of the curve and repeat the counting procedure. The points of type $I, D, \overset{\rightarrow\rightarrow}{T}_{\pm}$ remain the same, while the points of types $\overset{\rightarrow\leftarrow}{T}_{\pm}$ and $\overset{\leftarrow\rightarrow}{T}_{\pm}$ are interchanged. Thus we also have

$$2\overset{\rightarrow\rightarrow}{T}_{+} + 4\overset{\leftarrow\rightarrow}{T}_{+} - 2\overset{\rightarrow\rightarrow}{T}_{-} - 4\overset{\leftarrow\rightarrow}{T}_{-} - I - 2D = 0.$$

Adding these two equations and dividing by 2 now gives us the formula. □

Exercises

(1) Show that a vertex is a critical point for the curvature. Draw an example where a critical point for the curvature does not correspond to a local maximum/minimum.
(2) Show that a simple closed planar curve $\mathrm{q}(t)$ has the property that its unit tangent T is parallel to $\frac{d^2\mathrm{T}}{ds^2}$ at at least four points.
(3) Show that the concept of a vertex does not depend on the parametrization of the curve.
(4) Show that an ellipse that is not a circle has 4 vertices.
(5) Find the vertices of the curve
$$x^4 + y^4 = 1.$$
(6) Show that the closed curve
$$(1 - 2\sin\theta)(\cos\theta, \sin\theta), \ \theta \in [0, 2\pi]$$
is not simple and has exactly two vertices.
(7) Show that a vertex for a curve given by a graph $y = f(x)$ satisfies
$$\left(1 + \left(\frac{df}{dx}\right)^2\right)\frac{d^3 f}{dx^3} = 3\frac{df}{dx}\left(\frac{d^2 f}{dx^2}\right)^2.$$
(8) Consider a curve
$$\mathrm{q}(t) = r(t)(\cos t, \sin t)$$
where $r > 0$ and is 2π-periodic. Draw pictures where maxima/minima for r correspond to vertices. Is it possible to find an example where the minimum for r corresponds to a local maximum for $\kappa_{\pm}$ and the maximum for r corresponds to a local minimum for $\kappa_{\pm}$?
(9) Consider a unit speed curve $\mathrm{q} : [0, b] \to \mathbb{R}^2$ with $\kappa \geq 1/R$ and assume that it is tangent to a circle of radius $1/R$ at $t = 0$.
   (a) Show that the curve lies inside the circle on a sufficiently small interval $[0, \epsilon]$.
   (b) Draw an example where the curve does not lie inside the circle on all of $[0, b]$. Hint: The curve will not be simple.
(10) Consider a curve $\mathrm{q} : (a, b) \to \mathbb{R}^2$ that lies inside a circle of radius R. Show that if q touches the circle at $t_0 \in (a, b)$, then either $\kappa(t_0) > 1/R$ or κ has a critical point at t_0. Hint: Look at section 2.1 exercise 17.

(11) Show that for a domain $\Omega \subset \mathbb{R}^2$ with smooth boundary curve q, the divergence theorem
$$\int_\Omega \operatorname{div} F\, dudv = -\int_0^L F \cdot \mathrm{N}_\pm \, ds$$
follows from Green's theorem.

(12) Show that
$$A = \int_\Omega dudv = -\frac{1}{2}\int_0^L \mathrm{q} \cdot \mathrm{N}_\pm \, ds.$$

(13) Compute the area in the leaf of the folium of Descartes (see section 1.1 exercise 17).

(14) We say that a simple closed planar curve q has *convex interior* if the domain Ω bounded by q has the property that for any two points in Ω the line segments between the points also lie in Ω.
   (a) Show that if q minimizes the isoperimetric ratio, then its interior must be convex.
   (b) Show that if q minimizes the isoperimetric ratio and has perimeter L, then any section of q that has length $L/2$ solves Dido's problem.
   (c) Show that the isoperimetric problem is equivalent to Dido's problem.

(15) Consider all triangles where two side lengths a, b are fixed. Show that the triangle of largest area is the right triangle where a and b are perpendicular. Note that this triangle can be inscribed in a semicircle where the diameter is the hypotenuse. Use this to solve Dido's problem if we assume that there is a curve that solves Dido's problem.

(16) Show that among all quadrilaterals that have the same four side lengths $a, b, c, d > 0$ in order, the one with the largest area is the one that can be inscribed in a circle so that all four vertices are on the circle. Use this to solve the isoperimetric problem assuming that there is a curve that minimizes the isoperimetric ratio.

(17) Try to prove that the regular $2n$-gon maximizes the area among all $2n$-gons with the same perimeter.

2.4. Convex Curves

DEFINITION 2.4.1. We say that a regular planar curve is *convex* if it always lies on one side of its tangent lines. We say that it is *strictly convex* if it only intersects its tangent lines at the point of contact. A closed strictly convex curve is also called an *oval*.

Note that we do not need to assume that the curve is closed for this definition to make sense, but for convenience we do assume that it is defined on a closed interval. The definition can also be extended to piecewise smooth curves by requiring that it lies to one side of one or both tangent lines at points where the unit tangents don't agree.

THEOREM 2.4.2. *A planar convex curve is simple and the signed curvature cannot change sign.*

PROOF. First we show that the curvature can't change sign. Assume that the curve $\mathrm{q} : [0, L] \to \mathbb{R}^2$ has unit speed. Since the curve lies on one side of its tangent at any point $\mathrm{q}(s_0)$ it follows that
$$(\mathrm{q}(s) - \mathrm{q}(s_0)) \cdot \mathrm{N}_\pm (s_0)$$

is either nonnegative or nonpositive for all s. If it vanishes, then the curve must be part of the tangent line through $\mathrm{q}(s_0)$. In this case it is clearly simple and the curvature vanishes. Otherwise we have two disjoint sets $I_\pm \subset [0, L]$, where

$$\begin{aligned} I_+ &= \{s_0 \in [0, L] \mid (\mathrm{q}(s) - \mathrm{q}(s_0)) \cdot \mathrm{N}_\pm(s_0) \geq 0 \text{ for all } s \in [0, L]\} \\ I_- &= \{s_0 \in [0, L] \mid (\mathrm{q}(s) - \mathrm{q}(s_0)) \cdot \mathrm{N}_\pm(s_0) \leq 0 \text{ for all } s \in [0, L]\} \end{aligned}$$

Both of these sets must be closed by the continuity of $(\mathrm{q}(s) - \mathrm{q}(s_0)) \cdot \mathrm{N}_\pm(s_0)$. However, it is not possible to write an interval as the disjoint union of two closed sets unless one of these sets is empty.

Now assume that $I_+ = [0, L]$ so that $(\mathrm{q}(s) - \mathrm{q}(s_0)) \cdot \mathrm{N}_\pm(s_0) \geq 0$ for all s, s_0 with equality for $s = s_0$. Thus the second derivative with respect to s is also non-negative at s_0:

$$\begin{aligned} 0 &\leq \frac{d^2\mathrm{q}}{ds^2}(s_0) \cdot \mathrm{N}_\pm(s_0) \\ &= \frac{d\mathrm{T}}{ds}(s_0) \cdot \mathrm{N}_\pm(s_0) \\ &= \kappa_\pm(s_0). \end{aligned}$$

This shows that the signed curvature is always non-negative.

Assume for the remainder of the proof that the curve always lies to the left of its oriented tangent lines so that $\kappa_\pm \geq 0$. In case the curve is one-to-one when restricted to $(0, L)$ one of two things can happen: Either $\mathrm{q}(0) \neq \mathrm{q}(L)$ or $\mathrm{q}(0) = \mathrm{q}(L)$ in which case it is a convex loop that might not be smooth at $\mathrm{q}(0)$. In either case, the curve is simple. We can therefore assume below that $\mathrm{q} : (0, L) \to \mathbb{R}^2$ is not one-to-one.

First we claim that for every point $\mathrm{q}(a)$ of the curve there is a neighborhood $U \ni \mathrm{q}(a)$ such that if $\mathrm{q}(s) \in U$ for some other parameter value, then there is a $b \in [0, L]$ close to s such that $\mathrm{q}(b) = \mathrm{q}(a)$. This uses compactness of $[0, L]$. Assume that $\mathrm{q}(s_i) \to \mathrm{q}(a)$. By continuity $\mathrm{q}(b) = \mathrm{q}(a)$ for any b that is an accumulation point for the sequence s_i. Moreover, by compactness the sequence will have accumulation points. The existence of U now follows from a simple contradiction argument.

Next we show that if $a < b$ and $\mathrm{q}(a) = \mathrm{q}(b)$, then $\mathrm{T}(a) = \mathrm{T}(b)$. We exclude the case where $a = 0$ and $b = L$. If the tangent lines do not agree, then the curve crosses itself and cannot lie on one side of both tangent lines. The oriented tangent lines must now also agree as the curve would otherwise be forced to lie on both sides of the tangent line. This forces the curve to lie on the tangent line and that prevents it from intersecting itself as it is has no stationary points.

Since we assumed that $\kappa_\pm \geq 0$ it follows that $\theta(t)$ is increasing. If we combine that with the previous claim, then we see that if $a_0 < \cdots < a_k$ and $\mathrm{q}(a_0) = \cdots = \mathrm{q}(a_k)$, then

$$2\pi k \leq \sum_{i=1}^{k} \theta(a_i) - \theta(a_{i-1}) = \theta(a_k) - \theta(a_0) \leq \theta(L) - \theta(0).$$

In particular, $k \leq \frac{1}{2\pi}(\theta(L) - \theta(0))$ showing that the curve can only return to the same point a finite number of times.

After translating and rotating the curve assume that $\mathrm{q}(a) = 0$ and $\mathrm{T}(a) = (1, 0)$ and consider any $b > a$ such that $\mathrm{q}(b) = \mathrm{q}(a)$. Near a and b the parts of the curve will be graphs over a small interval on the x-axis. These graphs lie above the x-axis and are tangent to it at the origin. If the graphs do not coincide, then there will be values on one that are below the other. But then the points that are below must be to the right of the tangent lines through the points that are above. In case $a = 0$ or $b = L$, respectively, only the parts of the curve that are defined for $x \geq 0$ or $x \leq 0$, respectively, are shown to agree.

If we assemble all of these claims, it follows that for each $\mathrm{q}(a)$, $a \in (0, L)$, there is a neighborhood $U \ni \mathrm{q}(a)$ such that near a the part of the curve inside U is a graph over the tangent line. Moreover, if there are other parameter values that get mapped to U, then there are nearby parameter values that get mapped to $\mathrm{q}(a)$ and near these parameter values the curve is a graph over the same tangent line that coincides with the part of the curve near a. Finally as only a finite number of parameter values get mapped to $\mathrm{q}(a)$ we can ensure that any part of the curve that lies in a possibly smaller U is simply a reparameterized part of the curve near a. This shows that the curve is simple. □

Lemma 2.4.3. *If a curve has non-negative signed curvature and total curvature $\leq \pi$, then it is convex.*

Proof. Any curve with non-negative curvature always locally lies on the left of its tangent lines. So if it comes back to intersect a tangent l after having travelled to the left of l, then there will be a point of locally maximal distance to the left of l. At this local maximum the tangent line l^* must be parallel but not equal to l. If they are oriented in the same direction, then the curve will locally be on the right of l^*. As that does not happen they have opposite direction. This shows that the total curvature is $\geq \pi$. However, the curve will have strictly larger total curvature as it still has to make its way back to intersect l. □

Theorem 2.4.4. *If a closed curve has non-negative signed curvature and total curvature $\leq 2\pi$, then it is convex.*

Proof. The argument is similar to the one above. Assume that we have a tangent line l such that the curve lies on both sides of this line. As the curve is closed there'll be points one both sides of this tangent at maximal distance from the tangent. The tangent lines l^* and l^{**} at these points are then parallel to l. Thus we have three parallel tangent lines that are not equal. Two of these must correspond to unit tangents that point in the same direction. As the curvature does not change sign this implies that the total curvature of part of the curve is 2π. The total curvature must then be $> 2\pi$ as these two tangent lines are different and the curve still has to return to both of the points of contact. □

Exercises

(1) Consider a convex curve $\mathrm{q}(s) : [0, L] \to \mathbb{R}^2$ and fix a tangent line l through $\mathrm{q}(a)$. Show that

$$\{s \in [0, L] \mid \mathrm{T}(s) = \mathrm{T}(a)\} = \{s \in [0, L] \mid \mathrm{q}(s) \in l\}.$$

(2) Let $\mathrm{q}(\theta)$ be a simple closed planar curve with $\kappa > 0$ parametrized by θ, where θ is defined as the arclength parameter of the unit tangent field T. Show that

$$\begin{aligned}
\frac{d\,\mathrm{q}}{d\theta} &= \frac{1}{\kappa}\mathrm{T}, \\
\frac{d\,\mathrm{T}}{d\theta} &= \mathrm{N}_{\pm}, \\
\frac{d\,\mathrm{N}_{\pm}}{d\theta} &= -\mathrm{T}, \\
\mathrm{T}(\theta+\pi) &= -\mathrm{T}(\theta).
\end{aligned}$$

(3) Let $\mathrm{q}(\theta)$ be a simple closed planar curve with $\kappa > 0$ parametrized by θ, where θ is defined as the arclength parameter of the unit tangent field T. Define $v(\theta)$ as the distance from the origin to the tangent line through $\mathrm{q}(\theta)$.

(a) Show that

$$v(\theta) = -\mathrm{q}(\theta)\cdot\mathrm{N}_{\pm}(\theta).$$

(b) Show that the *width* (distance) between the parallel tangent lines through $\mathrm{q}(\theta)$ and $\mathrm{q}(\theta+\pi)$ is

$$w(\theta) = v(\theta) + v(\theta+\pi) = \mathrm{N}_{\pm}(\theta)\cdot(\mathrm{q}(\theta+\pi) - \mathrm{q}(\theta)).$$

(c) Show that:

$$L(\mathrm{q}) = \int_0^{2\pi} v(\theta)\,d\theta.$$

(d) Show that

$$\frac{1}{\kappa} = v + \frac{d^2 v}{d\theta^2}.$$

(e) Let A denote the area enclosed by the curve. Establish the following formulas for A

$$A = \frac{1}{2}\int_0^L v\,ds = \frac{1}{2}\int_0^{2\pi}\left(v^2 + v\frac{d^2 v}{d\theta^2}\right)d\theta = \frac{1}{2}\int_0^{2\pi}\left(v^2 - \left(\frac{dv}{d\theta}\right)^2\right)d\theta.$$

(4) Let $\mathrm{q}(\theta)$ be a simple closed planar curve with $\kappa > 0$ parametrized by θ, where θ is defined as the arclength parameter of the unit tangent field T. Show that the width from the previous problem satisfies:

$$\frac{d^2 w}{d\theta^2} + w = \frac{1}{\kappa(\theta)} + \frac{1}{\kappa(\theta+\pi)}.$$

(5) Let $\mathrm{q}(\theta)$ be a simple closed planar curve with $\kappa > 0$ parametrized by θ, where θ is defined as the arclength parameter of the unit tangent field T. With the width defined as in the previous exercises show that:

$$\int_0^{2\pi} w\,d\theta = 2L(\mathrm{q}).$$

(6) Let $\mathrm{q}(\theta)$ be a simple closed planar curve of constant width with $\kappa > 0$. The curve is parametrized by θ, where θ is defined as the arclength parameter of the unit tangent field T.

(a) Show that if θ corresponds to a local maximum for κ, then the opposite point $\theta+\pi$ corresponds to a local minimum.

(b) Assume for the remainder of the exercise that κ has a finite number of critical points and that they are all local maxima or minima. Show that the number of vertices is even and ≥ 6.
(c) Show that each point on the evolute corresponds to two points on the curve.
(d) Show that the evolute consists of n convex curves that are joined at n cusps that correspond to pairs of vertices on the curve.
(e) Show that the evolute has no double tangents.

(7) (Euler) Reverse the construction in the previous exercise to create curves of constant width by taking involutes of suitable curves.

(8) Let q be a closed convex curve and l a line.
   (a) Show that l can only intersect q in one point, two points, or a line segment.
   (b) Show that if l is also a tangent line then it cannot intersect q in only two points.
   (c) Show that the interior of q is convex, i.e., the segment between any two points in the interior also lies in the interior.

(9) Let q be a planar curve with non-negative signed curvature. Show that if q has a double tangent, then its total curvature is $\geq 2\pi$. Note that it is possible for the double tangent to have opposite directions at the points of tangency.

(10) Give an example of a planar curve (not closed) with positive curvature and no double tangents that is not convex.

(11) Let q be a closed planar curve without double tangents. Show that q is convex. Hint: Consider the set A of points (parameter values) on q where q lies on one side of the tangent line. Show that A is closed and not empty. Show that boundary points of A (i.e., points in A that are limit points of sequences in the complement of A) correspond to double tangents.

CHAPTER 3

Space Curves

3.1. The Fundamental Equations

The theory of space curves dates back to Clairaut in 1731. He considered them as the intersection of two surfaces given by equations. Clairaut showed that space curves have two curvatures, but they did not corresponds exactly to the curvature and torsion we introduce below. The subject was later taken up by Euler who was the first to work with parametrized curves and use arclength as a parameter. Lancret in 1806 introduced the concepts of unit tangent, principal normal and bi-normal and with those curvature and torsion as we now understand them. It is possible that Monge had some inklings of what torsion was, but he never presented an explicit formula. Cauchy in 1826 considerably modernized the subject and formulated some of the relations that later became part of the Serret and Frenet equations that we shall introduce below.

In order to create a set of equations for space curves $\mathrm{q}(t) : [a, b] \to \mathbb{R}^3$ we need to not only assume that the curve is regular but also that the velocity and acceleration are linearly independent. In this case it is possible to define a suitable positively oriented orthonormal frame T, N, and B by declaring

$$
\begin{aligned}
\mathrm{T} &= \frac{\mathrm{v}}{|\mathrm{v}|}, \\
\mathrm{N} &= \frac{\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}, \\
\mathrm{B} &= \mathrm{T} \times \mathrm{N}.
\end{aligned}
$$

The new normal vector B is called the *bi-normal*. We define the *curvature* and *torsion* by

$$
\begin{aligned}
\kappa &= \mathrm{N} \cdot \frac{d\,\mathrm{T}}{ds}, \\
\tau &= \mathrm{B} \cdot \frac{d\,\mathrm{N}}{ds}.
\end{aligned}
$$

PROPOSITION 3.1.1. *If $F : \mathbb{R}^3 \to \mathbb{R}^3$ is of the form $F(\mathrm{x}) = O\mathrm{x} + \mathrm{c}$, where O is a rotation and $\mathrm{c} \in \mathbb{R}^3$ and q a regular curve, then $\mathrm{q}^* = F(\mathrm{q})$ has unit tangent given by $\mathrm{T}^* = O\,\mathrm{T}$, normal $\mathrm{N}^* = O\,\mathrm{N}$, binormal $\mathrm{B} = (\det O)\, O\,\mathrm{B}$, curvature $\kappa^* = \kappa$, and torsion $\tau^* = (\det O)\,\tau$.*

PROOF. We saw in proposition 1.3.2 that $\mathrm{T}^* = O\,\mathrm{T}$ and $\mathrm{a}^* = O\,\mathrm{a}$. This shows that

$$
\mathrm{N}^* = \frac{O\,\mathrm{a} - (O\,\mathrm{a} \cdot O\,\mathrm{T})\,O\,\mathrm{T}}{|O\,\mathrm{a} - (O\,\mathrm{a} \cdot O\,\mathrm{T})\,O\,\mathrm{T}|} = O\left(\frac{\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}\right) = O\,\mathrm{N}.
$$

Thus the curvature is preserved just as in the proof of proposition 2.1.3. The binormal is created using the cross product and consequently uses the righthand rule. Orthogonal transformations therefore preserve cross products when $\det O = 1$, while it reverses the sign when $\det O = -1$, i.e., makes the righthand rule a lefthand rule. This establishes the formulas for the binormal and torsion as in proposition 2.1.3. $\square$

In section 1.3 we started by defining θ as arclength parameter for T and then proceeded to show that the above formulas for κ and N hold. After the next theorem we use the above definitions of κ and N to derive the old definitions.

THEOREM 3.1.2 (Serret, 1851 and Frenet, 1852). *If* $\mathrm{q}(t)$ *is a regular space curve with linearly independent velocity and acceleration, then*

$$\begin{aligned}
\frac{d\,\mathrm{q}}{dt} &= \frac{ds}{dt}\mathrm{T},\\
\frac{d\,\mathrm{T}}{dt} &= \kappa\frac{ds}{dt}\mathrm{N},\\
\frac{d\,\mathrm{N}}{dt} &= -\kappa\frac{ds}{dt}\mathrm{T}+\tau\frac{ds}{dt}\mathrm{B},\\
\frac{d\,\mathrm{B}}{dt} &= -\tau\frac{ds}{dt}\mathrm{N},
\end{aligned}$$

or

$$\frac{d}{dt}\begin{bmatrix} \mathrm{q} & \mathrm{T} & \mathrm{N} & \mathrm{B}\end{bmatrix} = \frac{ds}{dt}\begin{bmatrix} \mathrm{q} & \mathrm{T} & \mathrm{N} & \mathrm{B}\end{bmatrix}\begin{bmatrix} 0 & 0 & 0 & 0\\ 1 & 0 & -\kappa & 0\\ 0 & \kappa & 0 & -\tau\\ 0 & 0 & \tau & 0\end{bmatrix}.$$

Moreover,

$$\begin{aligned}
\kappa &= \frac{|\mathrm{v}\times\mathrm{a}|}{|\mathrm{v}|^3} = \frac{|\mathrm{a}-(\mathrm{a}\cdot\mathrm{T})\,\mathrm{T}|}{|\mathrm{v}|^2},\\
\tau &= \frac{\det\begin{bmatrix}\mathrm{v} & \mathrm{a} & \mathrm{j}\end{bmatrix}}{|\mathrm{v}\times\mathrm{a}|^2} = \frac{(\mathrm{v}\times\mathrm{a})\cdot\mathrm{j}}{|\mathrm{v}\times\mathrm{a}|^2},\\
\mathrm{N} &= \frac{\mathrm{a}-(\mathrm{a}\cdot\mathrm{T})\,\mathrm{T}}{|\mathrm{a}-(\mathrm{a}\cdot\mathrm{T})\,\mathrm{T}|},\\
\mathrm{B} &= \frac{\mathrm{v}\times\mathrm{a}}{|\mathrm{v}\times\mathrm{a}|}.
\end{aligned}$$

PROOF. The explicit formula for N is our explicit formula for the principal normal. As $\mathrm{T},\mathrm{N},\mathrm{B}$ form an orthonormal basis we have

$$\frac{d\,\mathrm{T}}{dt} = \left(\frac{d\,\mathrm{T}}{dt}\cdot\mathrm{T}\right)\mathrm{T}+\left(\frac{d\,\mathrm{T}}{dt}\cdot\mathrm{N}\right)\mathrm{N}+\left(\frac{d\,\mathrm{T}}{dt}\cdot\mathrm{B}\right)\mathrm{B}.$$

Here

$$\frac{d\,\mathrm{T}}{dt}\cdot\mathrm{T} = \frac{1}{2}\frac{d\,|\mathrm{T}|^2}{dt} = 0,$$

$$\frac{d\,\mathrm{T}}{dt}\cdot\mathrm{N} = \frac{ds}{dt}\frac{d\,\mathrm{T}}{ds}\cdot\mathrm{N} = \frac{ds}{dt}\kappa,$$

$$\frac{d\,\mathrm{T}}{dt}\cdot\mathrm{B} = \frac{d}{dt}\left(\frac{\mathrm{v}}{|\mathrm{v}|}\right)\cdot\mathrm{B} = \left(\frac{\mathrm{a}}{|\mathrm{v}|}-\frac{\mathrm{v}}{|\mathrm{v}|^2}\frac{d\,|\mathrm{v}|}{dt}\right)\cdot\mathrm{B} = 0,$$

as B is perpendicular to T, N and thus also to v, a. This establishes

$$\frac{d\,\mathrm{T}}{dt} = \kappa \frac{ds}{dt}\,\mathrm{N}\,.$$

Next note that

$$0 = \mathrm{B} \cdot \frac{d\,\mathrm{T}}{dt} = -\frac{d\,\mathrm{B}}{dt} \cdot \mathrm{T}\,.$$

This together with

$$\mathrm{B} \cdot \frac{d\,\mathrm{B}}{dt} = 0$$

shows that

$$\frac{d\,\mathrm{B}}{dt} = \left(\frac{d\,\mathrm{B}}{dt} \cdot \mathrm{N}\right)\mathrm{N}\,.$$

However, we also have

$$0 = \frac{d\,\mathrm{B}}{dt} \cdot \mathrm{N} + \mathrm{B} \cdot \frac{d\,\mathrm{N}}{dt} = \frac{d\,\mathrm{B}}{dt} \cdot \mathrm{N} + \frac{ds}{dt}\tau.$$

This implies

$$\frac{d\,\mathrm{B}}{dt} = -\tau \frac{ds}{dt}\,\mathrm{N}\,.$$

Finally the equation

$$\frac{d\,\mathrm{N}}{dt} = -\kappa \frac{ds}{dt}\,\mathrm{T} + \tau \frac{ds}{dt}\,\mathrm{B}\,.$$

is a direct consequence of the other two equations.

The formula for the curvature follows from observing that

$$\begin{aligned}
\frac{d\,\mathrm{T}}{ds} \cdot \mathrm{N} &= \left(\frac{d}{ds}\frac{\mathrm{v}}{|\mathrm{v}|}\right) \cdot \mathrm{N} \\
&= \frac{dt}{ds}\left(\frac{d}{dt}\frac{\mathrm{v}}{|\mathrm{v}|}\right) \cdot \mathrm{N} \\
&= \frac{1}{|\mathrm{v}|}\left(\frac{\mathrm{a}}{|\mathrm{v}|} - \frac{\mathrm{v}}{|\mathrm{v}|^2}\frac{d\,|\mathrm{v}|}{dt}\right) \cdot \mathrm{N} \\
&= \frac{\mathrm{a}}{|\mathrm{v}|^2} \cdot \frac{(\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T})}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|} \\
&= \frac{\mathrm{a} \cdot \mathrm{a} - (\mathrm{a} \cdot \mathrm{T})^2}{|\mathrm{v}|^2\,|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|} \\
&= \frac{|\mathrm{a}|^2\,|\mathrm{v}|^2 - (\mathrm{a} \cdot \mathrm{v})^2}{|\mathrm{v}|^4\,|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|} \\
&= \frac{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}{|\mathrm{v}|^2},
\end{aligned}$$

where $|\mathrm{v}|\,|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}| = \sqrt{|\mathrm{a}|^2\,|\mathrm{v}|^2 - (\mathrm{a} \cdot \mathrm{v})^2}$.

The formula for the binormal B follows directly from the calculation

$$\begin{aligned}
\mathrm{T} \times \mathrm{N} &= \frac{1}{|\mathrm{v}|}\,\mathrm{v} \times \left(\frac{\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}\right) \\
&= \frac{1}{|\mathrm{v}|}\,\mathrm{v} \times \left(\frac{\mathrm{a}}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}\right) \\
&= \frac{\mathrm{v} \times \mathrm{a}}{|\mathrm{v}|\,|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|} \\
&= \frac{\mathrm{v} \times \mathrm{a}}{|\mathrm{v} \times \mathrm{a}|}.
\end{aligned}$$

In the last equality recall that the denominators are the areas of the same parallelogram spanned by v and a.

To establish the general formula for τ we note

$$\begin{aligned}
\mathrm{B} \cdot \frac{d\,\mathrm{N}}{dt} &= \frac{\mathrm{v} \times \mathrm{a}}{|\mathrm{v} \times \mathrm{a}|} \cdot \frac{d}{dt}\left(\frac{\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}\right) \\
&= \frac{\mathrm{v} \times \mathrm{a}}{|\mathrm{v} \times \mathrm{a}|} \cdot \frac{\mathrm{j}}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|} \\
&\quad \frac{\mathrm{v} \times \mathrm{a}}{|\mathrm{v} \times \mathrm{a}|} \cdot \mathrm{a}\left(\frac{d}{dt}\frac{1}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}\right) \\
&\quad - \frac{\mathrm{v} \times \mathrm{a}}{|\mathrm{v} \times \mathrm{a}|} \cdot \mathrm{T}\left(\frac{d}{dt}\frac{(\mathrm{a} \cdot \mathrm{T})}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}\right) \\
&\quad - \frac{\mathrm{v} \times \mathrm{a}}{|\mathrm{v} \times \mathrm{a}|} \cdot \frac{d\,\mathrm{T}}{dt}\left(\frac{(\mathrm{a} \cdot \mathrm{T})}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}\right) \\
&= \frac{(\mathrm{v} \times \mathrm{a}) \cdot \mathrm{j}}{|\mathrm{v} \times \mathrm{a}|^2}\,|\mathrm{v}|\,.
\end{aligned}$$

The last line follows from our formulas for the area of the parallelogram spanned by v and a. A different strategy works by first noticing that

$$\begin{aligned}
\mathrm{v} &= (\mathrm{v} \cdot \mathrm{T})\,\mathrm{T}, \\
\mathrm{a} &= (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T} + (\mathrm{a} \cdot \mathrm{N})\,\mathrm{N}, \\
\mathrm{j} &= (\mathrm{j} \cdot \mathrm{T})\,\mathrm{T} + (\mathrm{j} \cdot \mathrm{N})\,\mathrm{N} + (\mathrm{j} \cdot \mathrm{B})\,\mathrm{B}\,.
\end{aligned}$$

Thus

$$\det \begin{bmatrix} \mathrm{v} & \mathrm{a} & \mathrm{j} \end{bmatrix} = (\mathrm{v} \cdot \mathrm{T})\,(\mathrm{a} \cdot \mathrm{N})\,(\mathrm{j} \cdot \mathrm{B})\,.$$

Next we recall that

$$\begin{aligned}
\mathrm{v} \cdot \mathrm{T} &= |\mathrm{v}|\,, \\
\mathrm{a} \cdot \mathrm{N} &= |\mathrm{v}|^2\,\kappa.
\end{aligned}$$

So we have to calculate $\mathrm{j} \cdot \mathrm{B}$. Keeping in mind that $\mathrm{a} \cdot \mathrm{B} = 0$ we obtain

$$\begin{aligned}
\mathrm{j} \cdot \mathrm{B} &= -\,\mathrm{a} \cdot \frac{d\,\mathrm{B}}{dt} \\
&= \tau\,|\mathrm{v}|\,\mathrm{a} \cdot \mathrm{N} \\
&= \tau\kappa\,|\mathrm{v}|^3
\end{aligned}$$

and finally combine this with

$$\kappa = \frac{|\mathrm{v} \times \mathrm{a}|}{|\mathrm{v}|^3}$$

to obtain the desired identity. □

The curvature and torsion can also be described by the formulas

$$\kappa = \frac{\text{area of parallelogram } (\mathrm{v}, \mathrm{a})}{|\mathrm{v}|^3},$$

$$\tau = \frac{\text{signed volume of the parallepiped } (\mathrm{v}, \mathrm{a}, \mathrm{j})}{(\text{area of the parallelogram } (\mathrm{v}, \mathrm{a}))^2}.$$

Corollary 3.1.3. *If* $\mathrm{q}(t)$ *is a regular space curve with linearly independent velocity and acceleration, then* T *is regular and if* θ *is its arclength parameter, then*

$$\frac{d\theta}{ds} = \frac{d\,\mathrm{T}}{ds} \cdot \mathrm{N}$$

and

$$\frac{d\,\mathrm{T}}{d\theta} = \frac{\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}}{|\mathrm{a} - (\mathrm{a} \cdot \mathrm{T})\,\mathrm{T}|}.$$

Proof. By assumption

$$0 < \kappa = \frac{d\,\mathrm{T}}{ds} \cdot \mathrm{N}\,.$$

This implies in particular that T is regular. We know from the chain rule that

$$\frac{d\,\mathrm{T}}{d\theta} = \frac{ds}{d\theta}\frac{d\,\mathrm{T}}{ds} = \frac{ds}{d\theta}\kappa\,\mathrm{N}\,.$$

Here both sides are unit vectors that are perpendicular to T and by definition $\frac{ds}{d\theta} > 0$ and $\kappa > 0$. This forces

$$\frac{d\theta}{ds} = \kappa$$

and

$$\frac{d\,\mathrm{T}}{d\theta} = \mathrm{N}\,.$$

This establishes the formulas. □

There is a very elegant way of collecting the Serret-Frenet formulas.

Corollary 3.1.4 (Darboux). *For a space curve as above define the Darboux vector*

$$\mathrm{D} = \tau\,\mathrm{T} + \kappa\,\mathrm{B}\,.$$

The Darboux vector has the property that

$$\frac{d}{dt}\begin{bmatrix} \mathrm{T} & \mathrm{N} & \mathrm{B} \end{bmatrix} = \frac{ds}{dt}\mathrm{D} \times \begin{bmatrix} \mathrm{T} & \mathrm{N} & \mathrm{B} \end{bmatrix}.$$

Proof. We have

$$\begin{aligned} \mathrm{D} \times \mathrm{T} &= \kappa\,\mathrm{N}, \\ \mathrm{D} \times \mathrm{N} &= \tau\,\mathrm{B} - \kappa\,\mathrm{T}, \\ \mathrm{D} \times \mathrm{B} &= -\tau\,\mathrm{N}, \end{aligned}$$

so the equation follows directly from the Serret-Frenet formulas. □

Exercises

(1) Find the curvature, torsion, normal, and binormal for the twisted cubic

$$\mathrm{q}(t) = \left(at, bt^2, ct^3\right),$$

where $a, b, c > 0$.

(2) Show that for

$$\mathrm{q}(t) = \left(t, \frac{1+t}{t}, \frac{1-t^2}{t}\right)$$

we have

$$\kappa^2 = \frac{3t^6}{2\left(t^4+t^2+1\right)^3},\ \tau = 0.$$

Show that this curve lies in the plane $x - y + z = 1$.

(3) Consider a cylindrical curve of the form

$$\mathrm{q}(\theta) = (\cos\theta, \sin\theta, z(\theta)).$$

Show that

$$\kappa = \frac{\left(1 + (z')^2 + (z'')^2\right)^{\frac{1}{2}}}{\left(1 + (z')^2\right)^{\frac{3}{2}}},$$

$$\tau = \frac{z' + z'''}{1 + (z')^2 + (z'')^2}.$$

(4) For a unit speed curve $\mathrm{q}(s)$ with positive curvature and torsion define $\mathrm{q}^*(s) = \int \mathrm{B}(s)\,ds$. Show that q^* is also unit speed and that $\mathrm{T}^* = \mathrm{B}$, $\mathrm{N}^* = -\mathrm{N}$, $\mathrm{B}^* = \mathrm{T}$, $\kappa^* = \tau$, and $\tau^* = \kappa$.

(5) Let $\mathrm{q}(t) : I \to \mathbb{R}^3$ be a regular curve such that its tangent field $\mathrm{T}(t)$ is also regular. Let s be the arclength parameter for q and θ the arclength parameter for T.

(a) Show that

$$\det\left[\begin{array}{ccc} \mathrm{T} & \frac{d\,\mathrm{T}}{ds} & \frac{d^2\,\mathrm{T}}{ds^2} \end{array}\right] = \kappa^2\tau.$$

(b) Show that

$$\det\left[\begin{array}{ccc} \mathrm{T} & \frac{d\,\mathrm{T}}{d\theta} & \frac{d^2\,\mathrm{T}}{d\theta^2} \end{array}\right] = \frac{\tau}{\kappa}.$$

(c) Show that $\mathrm{N} \times \frac{d\,\mathrm{N}}{ds} = \mathrm{D}$ and $\frac{d\,\mathrm{T}}{ds} \times \frac{d^2\,\mathrm{T}}{ds^2} = \kappa^2\mathrm{D}$.

(6) Show that for a unit speed curve $\mathrm{q}(s)$ with positive curvature

$$\det\left[\begin{array}{ccc} \frac{d^2\,\mathrm{q}}{ds^2} & \frac{d^3\,\mathrm{q}}{ds^3} & \frac{d^4\,\mathrm{q}}{ds^4} \end{array}\right] = \kappa^5 \frac{d}{ds}\left(\frac{\tau}{\kappa}\right).$$

(7) Let $\mathrm{q}(s)$ be a unit speed curve and define functions $a_n(s)$, $b_n(s)$, and $c_n(s)$ such that

$$\mathrm{q}^{(n)} = \frac{d^n\,\mathrm{q}}{ds^n} = a_n\,\mathrm{T} + b_n\,\mathrm{N} + c_n\,\mathrm{B}.$$

Show that

$$\begin{aligned} a_{n+1} &= \frac{da_n}{ds} - \kappa b_n \\ b_{n+1} &= \frac{db_n}{ds} + \kappa a_n - \tau c_n \\ c_{n+1} &= \frac{dc_n}{ds} + \tau b_n. \end{aligned}$$

(8) Show that T is regular when $\kappa > 0$ and that in this case the curvature of T is given by

$$\sqrt{1+\left(\frac{\tau}{\kappa}\right)^2}$$

and the torsion by

$$\frac{1}{\kappa\left(1+\left(\frac{\tau}{\kappa}\right)^2\right)}\frac{d}{ds}\left(\frac{\tau}{\kappa}\right).$$

(9) Show that the circular helix

$$(R\cos t, R\sin t, ht)$$

has constant curvature and torsion. Compute R, h in terms of the curvature and torsion. Conversely show that any unit speed space curve with constant curvature and torsion must look like

$$\mathrm{q}(s) = R\cos\left(\frac{s}{\sqrt{R^2+h^2}}\right)e_1 + R\sin\left(\frac{s}{\sqrt{R^2+h^2}}\right)e_2 + \frac{h}{\sqrt{R^2+h^2}}se_3 + \mathrm{q}(0),$$

where e_1, e_2, e_3 is an orthonormal basis.

(10) Let $\mathrm{q}(s) = (x(s), y(s), z(s)) : [0, L] \to \mathbb{R}^3$ be a unit speed space curve with curvature $\kappa(s)$ and torsion $\tau(s)$. Construct another space curve $\mathrm{q}^*(s) = x(s)e_1 + y(s)e_2 + z(s)e_3 + \mathrm{x}$, where e_1, e_2, e_3 is a positively oriented orthonormal basis and x and point.

(a) Show that q^* is a unit speed curve with curvature $\kappa^*(s) = \kappa(s)$ and torsion $\tau^*(s) = \tau(s)$.

(b) Show that a unit speed space curve with the same curvature and torsion as q is of the form q^*.

(11) Show that B is regular when $|\tau| > 0$ and that in this case the curvature of B is given by

$$\sqrt{1+\left(\frac{\kappa}{\tau}\right)^2}.$$

(12) Show that for a unit speed curve $\mathrm{q}(s)$ with positive curvature and non-zero torsion

$$\det\left[\begin{array}{ccc}\frac{d\mathrm{B}}{ds} & \frac{d^2\mathrm{B}}{ds^2} & \frac{d^3\mathrm{B}}{ds^3}\end{array}\right] = \tau^5\frac{d}{ds}\left(\frac{\kappa}{\tau}\right).$$

(13) Show that N is regular when $\kappa^2+\tau^2 > 0$ and that in this case the curvature of N is given by

$$\sqrt{1+\frac{\left(\kappa\frac{d\tau}{ds}-\tau\frac{d\kappa}{ds}\right)^2}{\left(\kappa^2+\tau^2\right)^3}}.$$

(14) Show that any unit speed curve $\mathrm{q}(s)$ with constant nonzero torsion τ satisfies

$$\mathrm{q}(s) = \mathrm{q}(s_0) + \int_{s_0}^{s}\frac{1}{\tau}\mathrm{B}\times\frac{d\mathrm{B}}{dt}dt.$$

(15) Show a closed unit speed curve $\mathrm{q}(s) : [0, L] \to \mathbb{R}^3$ with positive curvature satisfies

$$\int_0^L(\dot{\kappa}\,\mathrm{q}+\tau\,\mathrm{B})\,ds = 0.$$

(16) Define $\rho = \sqrt{\kappa^2+\tau^2}$ and ϕ by

$$\kappa = \rho\cos\phi,\ \tau = \rho\sin\phi.$$

(a) Show that $\rho = |\mathrm{D}|$ and that ϕ is the natural arclength parameter for the unit field $\frac{1}{\rho}\mathrm{D}$.

(b) Show that if

$$P(s) = \mathrm{q}(s) + \frac{\kappa}{\rho^2}\mathrm{N},$$

then

$$\frac{dP}{ds} = \frac{\tau}{\rho^2}\mathrm{D} + \frac{d}{ds}\left(\frac{\kappa}{\rho^2}\right)\mathrm{N}.$$

(17) Show that a space curve is part of a line if all its tangent lines pass through a fixed point.

(18) Let $\mathrm{Q}(t)$ be a vector associated to a curve $\mathrm{q}(t)$ such that

$$\frac{d}{dt}\begin{bmatrix} \mathrm{T} & \mathrm{N} & \mathrm{B} \end{bmatrix} = \frac{ds}{dt}\mathrm{Q} \times \begin{bmatrix} \mathrm{T} & \mathrm{N} & \mathrm{B} \end{bmatrix}.$$

Show that $\mathrm{Q} = \mathrm{D}$.

(19) Let $\mathrm{q}(s)$ be a unit speed space curve with non-vanishing curvature and torsion. Show that

$$\frac{d}{ds}\left(\frac{1}{\tau}\frac{d}{ds}\left(\frac{1}{\kappa}\frac{d^2\mathrm{q}}{ds^2}\right)\right) + \frac{d}{ds}\left(\frac{\kappa}{\tau}\frac{d\mathrm{q}}{ds}\right) + \frac{\tau}{\kappa}\frac{d^2\mathrm{q}}{ds^2} = 0.$$

(20) Consider a unit speed space curve $\mathrm{q}(s)$ with non-vanishing curvature and torsion. Let k be a fixed vector and denote by ϕ_{T}, ϕ_{N}, ϕ_{B} the angles between T, N, B and k. Show that

$$\begin{aligned} \kappa\cos\phi_{\mathrm{N}} &= -\frac{d\phi_{\mathrm{T}}}{ds}\sin\phi_{\mathrm{T}}, \\ \kappa\cos\phi_{\mathrm{T}} - \tau\cos\phi_{\mathrm{B}} &= \frac{d\phi_{\mathrm{N}}}{ds}\sin\phi_{\mathrm{N}}, \\ \tau\cos\phi_{\mathrm{N}} &= \frac{d\phi_{\mathrm{B}}}{ds}\sin\phi_{\mathrm{B}}, \end{aligned}$$

and

$$\frac{d\phi_{\mathrm{B}}}{ds}\sin\phi_{\mathrm{B}} = -\frac{\tau}{\kappa}\frac{d\phi_{\mathrm{T}}}{ds}\sin\phi_{\mathrm{T}}.$$

(21) For a regular space curve $\mathrm{q}(t)$ we say that a normal field X is *parallel* along q if $\mathrm{X}\cdot\mathrm{T} = 0$ and $\frac{d\mathrm{X}}{dt}$ is parallel to T.

(a) Show that for a fixed t_0 and $\mathrm{X}(t_0) \perp \mathrm{T}(s_0)$ there is a unique parallel field X that is $\mathrm{X}(t_0)$ at t_0.

(b) A *Bishop frame* consists of an orthonormal frame $\mathrm{T}, \mathrm{N}_1, \mathrm{N}_2$ along the curve so that $\mathrm{N}_1, \mathrm{N}_2$ are both parallel along q. For such a frame show that

$$\frac{d}{dt}\begin{bmatrix} \mathrm{T} & \mathrm{N}_1 & \mathrm{N}_2 \end{bmatrix} = \frac{ds}{dt}\begin{bmatrix} \mathrm{T} & \mathrm{N}_1 & \mathrm{N}_2 \end{bmatrix}\begin{bmatrix} 0 & \kappa_1 & \kappa_2 \\ -\kappa_1 & 0 & 0 \\ -\kappa_2 & 0 & 0 \end{bmatrix}.$$

Note that such frames always exist, even when the space curve doesn't have positive curvature everywhere.

(c) Show further that for such a frame

$$\kappa^2 = \kappa_1^2 + \kappa_2^2.$$

(d) Show that if q has positive curvature so that N is well-defined, then
$$\mathrm{N} = \cos\phi\,\mathrm{N}_1 + \sin\phi\,\mathrm{N}_2,$$
where
$$\frac{d\phi}{dt} = \frac{ds}{dt}\tau,$$
$$\kappa_1 = \kappa\cos\phi,\ \kappa_2 = \kappa\sin\phi.$$

(e) Give an example of a closed space curve where the parallel curves don't close up.

3.2. Characterizations of Space Curves

We show that the tangent lines determine a space curve, but that the (principal) normal lines do not necessarily characterize the curve. Unless otherwise stated we assume that all curves are regular and have positive curvature.

THEOREM 3.2.1. *If* $\mathrm{q}(t)$ *and* $\mathrm{q}^*(t)$ *are two regular curves that admit a common parametrization such that their tangent lines agree at corresponding points, then* $\mathrm{q}(t) = \mathrm{q}^*(t)$ *for all* t *where either* $\kappa(t) \neq 0$ *or* $\kappa^*(t) \neq 0$.

PROOF. Note that the common parametrization is not necessarily the arclength parametrization for either curve. These arclength parametrizations are denoted s, s^*. The assumption implies that corresponding velocity vectors are always parallel and that
$$\mathrm{q}^*(t) = \mathrm{q}(t) + u(t)\,\mathrm{T}(t)$$
for some function $u(t)$. We obtain by differentiation
$$\frac{d\,\mathrm{q}^*}{dt} = \frac{d\,\mathrm{q}}{dt} + \frac{du}{dt}\,\mathrm{T} + u\frac{ds}{dt}\kappa\,\mathrm{N}\,.$$
This forces
$$u\frac{ds}{dt}\kappa = 0$$
as N is perpendicular to the other vectors. So whenever $\kappa \neq 0$ it follows that $u = 0$. This means that the curves agree on the set where $\kappa \neq 0$. Reversing the roles of the curves we similarly obtain that the curves agree when $\kappa^* \neq 0$. □

The analogous question for principal normal lines requires that these normal lines are defined and thus that the curvatures never vanish. Nevertheless it is easy to find examples of pairs of curves that have the same normal lines without being the same curve. The double helix is in fact a great example of this. This corresponds to the two pairs of circular helices
$$\mathrm{q} = (R\cos t, R\sin t, ht) \text{ and } \mathrm{q}^* = (-R\cos t, -R\sin t, ht)$$
More generally for fixed $h > 0$ all of the curves
$$(R\cos t, R\sin t, ht)$$
have the same normal lines for all $R \in \mathbb{R}$.

DEFINITION 3.2.2. We say that two curves q and q^* are *Bertrand mates* if it is possible to find a common parametrization of both curves such that their principal normal lines agree at corresponding points.

THEOREM 3.2.3. *Let* q *and* q^* *be Bertrand mates with non-zero curvatures and torsion. Either the curves agree or there are linear relationships*

$$a\kappa + b\tau = 1,\ a\kappa^* - b\tau^* = 1$$

between curvature and torsion. Conversely, any curve with non-zero curvature and torsion such that $a\kappa + b\tau = 1$ *for constants* a, b *has a Bertrand mate.*

PROOF. We'll use s, s^* for the arclength of the two curves. That two curves are Bertrand mates is equivalent to

$$\mathrm{N}(t) = \pm\,\mathrm{N}^*(t)$$

and

$$\mathrm{q}^*(t) = \mathrm{q}(t) + r(t)\,\mathrm{N}(t)$$

for some function $r(t)$.

The first condition implies

$$\frac{d}{dt}(\mathrm{T}\cdot\mathrm{T}^*) = \frac{ds}{dt}\kappa\,\mathrm{N}\cdot\mathrm{T}^* + \frac{ds^*}{dt}\kappa^*\,\mathrm{T}\cdot\mathrm{N}^* = 0.$$

Thus $\mathrm{T}\cdot\mathrm{T}^* = \cos\theta$ for a fixed angle θ and

$$\mathrm{T}^*(t) = \mathrm{T}(t)\cos\theta + \mathrm{B}(t)\sin\theta.$$

Differentiating the second condition implies

$$\begin{aligned}\frac{ds^*}{dt}\mathrm{T}^* &= \frac{ds}{dt}\mathrm{T} + \frac{dr}{dt}\mathrm{N} + r\frac{ds}{dt}(-\kappa\,\mathrm{T} + \tau\,\mathrm{B})\\ &= \left(\frac{ds}{dt} - r\frac{ds}{dt}\kappa\right)\mathrm{T} + \frac{dr}{dt}\mathrm{N} + r\frac{ds}{dt}\tau\,\mathrm{B}.\end{aligned}$$

If we combine this with $\mathrm{T}^*(t) = \mathrm{T}(t)\cos\theta + \mathrm{B}(t)\sin\theta$, then it follows that

$$\begin{aligned}\frac{dr}{dt} &= 0,\\ \frac{ds^*}{dt}\cos\theta &= \frac{ds}{dt} - r\frac{ds}{dt}\kappa,\\ \frac{ds^*}{dt}\sin\theta &= r\frac{ds}{dt}\tau.\end{aligned}$$

In particular, r is constant,

$$\frac{ds^*}{ds}\cos\theta = (1 - r\kappa),$$

and

$$\frac{ds^*}{ds}\sin\theta = r\tau.$$

When $r \neq 0$ the fact that $\tau \neq 0$ implies

$$-(1 - r\kappa)\sin\theta + r\tau\cos\theta = 0,$$

which shows

$$\kappa r + \tau r\cot\theta = 1.$$

Switching the roles of the curve forces us to change the sign of θ. Thus

$$\mathrm{T}(t) = \mathrm{T}^*(t)\cos\theta - \mathrm{B}^*(t)\sin\theta$$

and

$$\kappa^* r - \tau^* r\cot\theta = 1.$$

Conversely, assume that we have a regular curve $\mathrm{q}(s)$ parametrized by arclength so that

$$\kappa r + \tau r \cot\theta = 1.$$

Inspired by our conclusions from the first part of the proof we define

$$\mathrm{q}^*(s) = \mathrm{q}(s) + r\,\mathrm{N}(s)$$

and note that

$$\frac{d\,\mathrm{q}^*}{ds} = \mathrm{T} + r(-\kappa\,\mathrm{T} + \tau\,\mathrm{B}) = \tau r(\cot\theta\,\mathrm{T} + \mathrm{B}).$$

Thus $\mathrm{T}^* = \pm(\cos\theta\,\mathrm{T} + \sin\theta\,\mathrm{B})$. This shows that

$$\frac{d\,\mathrm{T}^*}{ds} = \pm(\kappa\cos\theta - \tau\sin\theta)\,\mathrm{N}$$

and in particular that $\mathrm{N}^* = \pm\mathrm{N}$. □

Exercises

(1) A curve is *planar* if there is a vector k such that $\mathrm{q}(t)\cdot\mathrm{k}$ is constant.
  (a) Show that this is equivalent to saying that the tangent T is always perpendicular to k and implies that all derivatives $\frac{d^k\mathrm{q}}{dt^k}$, $k \geq 1$ are perpendicular to k.
  (b) Show that a curve is planar if and only if τ vanishes.

(2) Show that a curve $\mathrm{q}(t)$ is planar if and only if $\mathrm{j} \in \operatorname{span}\{\mathrm{v},\mathrm{a}\}$ for all t.

(3) Consider solutions to the second order equation

$$\mathrm{a} = F(\mathrm{q},\mathrm{v}).$$

Show that all solutions are planar if $F(\mathrm{q},\mathrm{v}) \in \operatorname{span}\{\mathrm{q},\mathrm{v}\}$ for all vectors q, v. This happens, e.g., when the force field F is radial, i.e., F is proportional to position q.

(4) Show that a curve q is planar if and only if there is a point c such that $\mathrm{q}(s) - \mathrm{c} \in \operatorname{span}\{\mathrm{T}(s), \mathrm{N}(s)\}$ for all s.

(5) Let $\mathrm{q}(t)$ and $\mathrm{q}^*(t)$ be two regular curves that admit a common parametrization such that their unit tangents are equal at corresponding points.
  (a) Show that their normals and binormals are also equal.
  (b) Show that

$$\frac{\kappa^*}{\kappa} = \frac{ds}{ds^*} = \frac{\tau^*}{\tau}.$$

(6) (Lancret, 1806) A *generalized helix* is a curve such that $\mathrm{T}\cdot\mathrm{k}$ is constant for some fixed vector k, i.e., T is planar. Note that since the unit tangent traces a curve on the sphere it has to lie in the intersection of the unit sphere and a plane, i.e., a latitude, and must in particular be a circle.
  (a) Show that this is equivalent to the normal N always being perpendicular to k.
  (b) Show that a curve is a generalized helix if and only if the ratio τ/κ is constant.
  (c) Show that this is equivalent to assuming that the curvature of the unit tangent T is constant.
  (d) Show that this is equivalent to the torsion of T vanishing.

(7) Show that
$$r(t)(\cos t, \sin t, 0) + h(t)(0,0,1)$$
is a generalized helix if and only if
$$\frac{\dot{h}^2}{r^2 + \dot{r}^2 + \dot{h}^2}$$
is constant.

(8) Show that (at, bt^2, ct^3) is a generalized helix when $3ac = \pm 2b^2$.

(9) Show that $(3t - t^3, 3t^2, 3t + t^3)$ is a generalized helix.

(10) Show that $\left(\frac{(1+s)^{\frac{3}{2}}}{3}, \frac{(1-s)^{\frac{3}{2}}}{3}, \frac{s}{\sqrt{2}}\right)$ is unit speed and a generalized helix.

(11) Show that $(\cosh t, \sinh t, t)$ is a generalized helix.

(12) Let q be a unit speed curve with positive curvature.
   (a) Show that the Darboux vector D is constant if and only if κ and τ are constant.
   (b) Show that $\frac{\mathrm{D}}{|\mathrm{D}|}$ is constant if and only if $\frac{\tau}{\kappa}$ is constant.

(13) Show that a unit speed curve q has the property that $\frac{\tau}{\kappa} = as + b$ for constants a, b if and only if there is a point c such that $\mathrm{q}(s) - \mathrm{c} \in \operatorname{span}\{\mathrm{T}(s), \mathrm{B}(s)\}$ for all s. Hint: Try $\mathrm{c} = \mathrm{q} - (s + c)\,\mathrm{T} - d\,\mathrm{B}$ where $as + b = \frac{s+c}{d}$.

(14) A curve is *spherical* if it lies on some sphere. Show that a curve is spherical if and only if its normal planes all pass through a fixed point.

(15) Show that a curve q is spherical if and only if there is a point c such that $\mathrm{q}(s) - \mathrm{c} \in \operatorname{span}\{\mathrm{N}(s), \mathrm{B}(s)\}$ for all s.

(16) Show that a unit speed curve on a sphere of radius R satisfies
$$\kappa \geq \frac{1}{R}.$$

(17) Show that if a curve with constant curvature lies on a sphere, then it is part of a circle, i.e., it is forced to be planar.

(18) Assume we have a unit speed spherical curve. If the center of the sphere is c and the radius R, then the curve must satisfy
$$|\mathrm{q}(s) - \mathrm{c}|^2 = R^2.$$
   (a) Show that if a spherical curve has nowhere vanishing curvature, then
$$\begin{aligned}(\mathrm{q} - \mathrm{c}) \cdot \mathrm{N} &= -\frac{1}{\kappa}, \\ \tau(\mathrm{q} - \mathrm{c}) \cdot \mathrm{B} &= \frac{d}{ds}\left(\frac{1}{\kappa}\right).\end{aligned}$$
   (b) Show that if both curvature and torsion are nowhere vanishing, then
$$\frac{1}{\kappa^2} + \left(\frac{1}{\tau}\frac{d}{ds}\left(\frac{1}{\kappa}\right)\right)^2 = R^2$$
   and
$$\frac{\tau}{\kappa} + \frac{d}{ds}\left(\frac{1}{\tau}\frac{d}{ds}\left(\frac{1}{\kappa}\right)\right) = 0.$$

(19) Conversely, show that if a curve q with nowhere vanishing curvature and torsion satisfies
$$\frac{\tau}{\kappa} + \frac{d}{ds}\left(\frac{1}{\tau}\frac{d}{ds}\left(\frac{1}{\kappa}\right)\right) = 0,$$

then

$$\frac{1}{\kappa^2}+\left(\frac{1}{\tau}\frac{d}{ds}\left(\frac{1}{\kappa}\right)\right)^2=R^2$$

for some constant R. Furthermore show that

$$\mathrm{c}(s)=\mathrm{q}+\frac{1}{\kappa}\mathrm{N}+\frac{1}{\tau}\frac{d}{ds}\left(\frac{1}{\kappa}\right)\mathrm{B}$$

is constant and conclude that q lies on the sphere with center c and radius R.

(20) Prove that a unit speed curve q with non-zero curvature and torsion lies on a sphere if there are constants a, b such that

$$\kappa\left(a\cos\left(\int\tau ds\right)+b\sin\left(\int\tau ds\right)\right)=1.$$

Hint: Show

$$\frac{1}{\tau}\frac{d}{ds}\left(\frac{1}{\kappa}\right)=-a\sin\left(\int\tau ds\right)+b\cos\left(\int\tau ds\right)$$

and

$$\frac{\tau}{\kappa}=-\frac{d}{ds}\left(-a\sin\left(\int\tau ds\right)+b\cos\left(\int\tau ds\right)\right)$$

and use the previous exercise.

(21) Let $\mathrm{q}(t)=(x(t), y(t), z(t))$ be a generalized helix that lies on the cylinder $x^2+y^2=1$.

(a) Show that as long as $(x(t), y(t))$ is not stationary, then the curve can be parametrized as

$$\mathrm{q}(\phi)=(\cos\phi, \sin\phi, z(\phi)).$$

(b) Use that parametrization to compute the normal component of the acceleration

$$\mathrm{a}-\frac{\mathrm{a}\cdot\mathrm{v}}{|\mathrm{v}|^2}\mathrm{v}$$

and show that this vector can only stay perpendicular to vectors $\mathrm{k}=(0,0,c)$ and in this case only when $z''=0$.

(c) Show that $(x(t), y(t))$ is never stationary. Hint: First show that it can't be stationary everywhere as it can't be a line parallel to the z-axis.

(d) Conclude that the original curve is a circular helix.

(22) Let $\mathrm{q}(t)=(x(t), y(t), z(t))$ be a generalized helix that lies on the cone $x^2+y^2=z^2$ with $z>0$. Show that the planar curve $(x(t), y(t))$ forms a constant angle with the radial lines and conclude that it is either a radial line or can be reparametrized as a logarithmic spiral

$$(x(\phi), y(\phi))=ae^{b\phi}(\cos\phi, \sin\phi).$$

Hint: Look at the previous exercise, but the calculations are more involved.

(23) Show that

$$\mathrm{q}^*(s)=\mathrm{q}+\frac{1}{\kappa}\mathrm{N}+\frac{1}{\kappa}\cot\left(\int\tau ds\right)\mathrm{B}$$

defines an evolute for q. Hint: See remark 1.3.12.

(24) Show that a planar curve has infinitely many Bertrand mates.

(25) Let q, q^* be two Bertrand mates.

(a) (Schell) Show that

$$\tau\tau^* = \frac{\sin^2\theta}{r^2}.$$

(b) (Mannheim) Show that

$$(1 - r\kappa)(1 + r\kappa^*) = \cos^2\theta.$$

(26) Consider a curve $\mathrm{q}(s)$ parametrized by arclength with positive curvature and non-vanishing torsion such that

$$\kappa r + \tau r \cot\theta = 1,$$

i.e., there is a Bertrand mate.

(a) Show that the Bertrand mate is uniquely determined by r.
(b) Show that if q has two different Bertrand mates then it must be a generalized helix.
(c) Show that if a generalized helix has a Bertrand mate, then its curvature and torsion are constant, consequently it is a circular helix.

(27) Investigate properties of a pair of curves that have the same normal planes at corresponding points, i.e., their tangent lines are parallel.

(28) Investigate properties of a pair of curves that have the same binormal lines at corresponding points.

3.3. Closed Space Curves

This section discusses various aspects of surface theory. It can be used as motivation for what is to come in chapters 4 and 5, or it can be covered later and treated as a culmination of those same chapters.

We start by studying spherical curves. In fact any regular space curve generates a natural spherical curve, the unit tangent. This was studied for planar curves in section 2.2 where the unit tangent became a curve on a circle. In that case the length of the unit tangent curve could be interpreted as an integral of the curvature that measured how much the curve turns. When the planar curve was closed this turning necessarily had to be a multiple of 2π.

A regular spherical curve $\mathrm{q}(t) : I \to S^2(1)$ has an alternate set of equations that describe its properties. Instead of the principal normal it has a signed normal that is tangent to the sphere. If we note that q is also normal to the sphere, then the signed normal can be defined as the vector

$$\mathrm{S} = \mathrm{q} \times \mathrm{T}\,.$$

This leads to the a new set of equations

$$\begin{aligned}
\frac{d\mathrm{T}}{dt} &= \frac{ds}{dt}\left(\kappa_g \mathrm{S} - \mathrm{q}\right), \\
\frac{d\mathrm{S}}{dt} &= \frac{ds}{dt}\left(-\kappa_g \mathrm{T}\right), \\
\frac{d\mathrm{q}}{dt} &= \frac{ds}{dt}\mathrm{T},
\end{aligned}$$

where the *geodesic curvature* κ_g is defined as $\kappa_g = \frac{d\mathrm{T}}{ds} \cdot \mathrm{S}$. The geodesic curvature measures how far a curve is from being a great circle as those curves have the property that $\frac{d\mathrm{T}}{ds} \cdot \mathrm{S} = 0$. The last equation is obvious by now. The first follows from the definition of κ_g and the second from the other two.

It'll be convenient to develop a new formula that calculates the length of a spherical curve by counting the number of great circles that it intersects. There is a similar formula for planar curves, Crofton's formula, that uses how many times the curve intersects lines. An oriented great circle thought of as an equator is uniquely determined by its corresponding North pole if we think in terms of the right hand rule. Thus intersections with the oriented great circle with *pole* x can be counted as

$$n_{\mathrm{q}}(\mathrm{x}) = |\{t \mid \mathrm{x} \cdot \mathrm{q}(t) = 0\}|$$

and Crofton's formula becomes

$$\frac{1}{4}\int_{S^2} n_{\mathrm{q}}(\mathrm{x})\, d\mathrm{x} = L(\mathrm{q}),$$

where the integral on the left is a surface integral over the unit sphere.

A point $\mathrm{q}(t)$ on the curve intersects the great circles going through that point. These great circles are described by their poles which in turn lie on the great circle with pole $\mathrm{q}(t)$. This great circle can be parametrized by

$$\mathrm{x}(\theta, t) = \cos(\theta)\,\mathrm{T}(t) + \sin(\theta)\,\mathrm{S}(t).$$

Thus the surface integral becomes

$$\begin{aligned}
&\int_a^b \int_0^{2\pi} \left| \frac{d\mathrm{x}}{d\theta} \times \frac{d\mathrm{x}}{dt} \right| d\theta dt \\
&= \int_a^b \int_0^{2\pi} \left|(-\sin(\theta)\,\mathrm{T} + \cos(\theta)\,\mathrm{S}) \times ((\kappa_g \mathrm{S} - \mathrm{q})\cos(\theta) - \kappa_g\,\mathrm{T}\sin(\theta))\right| \frac{ds}{dt} d\theta dt \\
&= \int_a^b \int_0^{2\pi} \left|(-\sin(\theta)\,\mathrm{T} + \cos(\theta)\,\mathrm{S}) \times (\kappa_g(-\sin(\theta)\,\mathrm{T} + \cos(\theta)\,\mathrm{S}) - \cos(\theta)\,\mathrm{q})\right| \frac{ds}{dt} d\theta dt \\
&= \int_a^b \int_0^{2\pi} \left|(-\sin(\theta)\,\mathrm{T} + \cos(\theta)\,\mathrm{S}) \times (-\cos(\theta)\,\mathrm{q})\right| \frac{ds}{dt} d\theta dt \\
&= \int_0^{L(\mathrm{q})} \int_0^{2\pi} \left|(-\sin(\theta)\,\mathrm{T} + \cos(\theta)\,\mathrm{S}) \times (-\cos(\theta)\,\mathrm{q})\right| d\theta ds \\
&= \int_0^{L(\mathrm{q})} \int_0^{2\pi} \left|\left(\sin(\theta)\cos(\theta)\,\mathrm{T}(t) \times \mathrm{q} - \cos^2(\theta)\,\mathrm{S} \times \mathrm{q}\right)\right| d\theta ds \\
&= \int_0^{L(\mathrm{q})} \int_0^{2\pi} \sqrt{\sin^2(\theta)\cos^2(\theta) + \cos^4(\theta)}\, d\theta ds \\
&= \int_0^{L(\mathrm{q})} \int_0^{2\pi} |\cos(\theta)|\, d\theta ds \\
&= \int_0^{L(\mathrm{q})} 4 ds \\
&= 4L(\mathrm{q}).
\end{aligned}$$

THEOREM 3.3.1 (Fenchel, 1929). *If* q *is a closed space curve with nonvanishing curvature, then*

$$\int \kappa ds \geq 2\pi.$$

Proof. If the unit tangent field lies in a hemisphere with pole x, i.e., $\mathrm{T}\cdot\mathrm{x} \geq 0$ for all s, then after integration we obtain

$$(\mathrm{q}(L) - \mathrm{q}(0)) \cdot \mathrm{x} \geq 0.$$

However, $\mathrm{q}(L) = \mathrm{q}(0)$ as the curve is closed. So it follows that $\mathrm{T}\cdot\mathrm{x} = 0$ for all s, i.e., the unit tangent is always perpendicular to x and hence the curve is planar.

This in turn implies that the unit tangent must intersect all great circles in at least two points. In fact if it does not intersect a certain great circle, then it must lie in an open hemisphere. If it intersects a great circle exactly once, then it must lie on one side of it and be tangent to the great circle. By moving the great circle slightly away from the point of tangency we obtain a new great circle that does not intersect the unit tangent, another contradiction. Having shown that T intersects all great circles at least twice it follows from Crofton's formula that

$$\int \kappa ds = L(\mathrm{T}) = \frac{1}{4}\int_{S^2} n_{\mathrm{T}}(\mathrm{x})\, d\mathrm{x} \geq \frac{2}{4}\cdot 4\pi = 2\pi.$$

□

Remark 3.3.2. There is an alternate proof of this theorem which also addresses what happens when the total curvature is 2π. Exercise 12 in section 1.2 shows that if $L(\mathrm{T}) \leq 2\pi$, then T must lie in a closed hemisphere. Consequently, it is a great circle. This shows that q is planar with nonvanishing curvature and total curvature 2π. The results in section 2.4 can then be used to show that the curve is convex.

Definition 3.3.3. A simple closed curve q is called an *unknot* or said to be *unknotted* if there is a one-to-one map from the disc to $\mathbb{R}^3$ such that boundary of the disc is q.

Theorem 3.3.4 (Fary, 1949 and Milnor, 1950). *If a simple closed space curve is knotted, then*

$$\int \kappa ds \geq 4\pi.$$

Proof. We assume that $\int \kappa ds < 4\pi$ and show that the curve is not knotted. Crofton's formula implies that

$$\frac{1}{4}\int_{S^2} n_{\mathrm{T}}(\mathrm{x})\, d\mathrm{x} = \int \kappa ds < 4\pi.$$

As the sphere has area 4π this can only happen if we can find x such that $n_{\mathrm{T}}(\mathrm{x}) \leq 3$. Now observe that

$$\frac{d(\mathrm{q}\cdot\mathrm{x})}{ds} = \mathrm{T}(s)\cdot\mathrm{x}.$$

So the function $\mathrm{q}\cdot\mathrm{x}$ has at most three critical points. Since q is closed there will be a maximum and a minimum. The third critical point, should it exist, can consequently only be an inflection point. Assume that the minimum is obtained at $s = 0$ and the maximum at $s_0 \in (0, L)$. The third critical point can be assumed to be in $(0, s_0)$. This implies that the function $\mathrm{q}(s)\cdot x$ is strictly increasing on $(0, s_0)$ and strictly decreasing on (s_0, L). For each $t \in (0, s_0)$ we can then find a unique $s(t) \in (s_0, L)$ such that $\mathrm{q}(t)\cdot\mathrm{x} = \mathrm{q}(s(t))\cdot\mathrm{x}$. Join the two points $\mathrm{q}(t)$ and $\mathrm{q}(s(t))$ by a segment. These segments will sweep out an area whose boundary is the curve and no two of the segments intersect as they belong to parallel planes orthogonal to x. This shows that the curve is the unknot. □

Exercises

(1) Let q be a unit speed spherical curve.

(a) Show that

$$\begin{aligned}\kappa^2 &= 1+\kappa_g^2,\\ \mathrm{N} &= \frac{1}{\kappa}\left(-\mathrm{q}+\kappa_g \mathrm{S}\right),\\ \mathrm{B} &= \frac{1}{\kappa}\left(\kappa_g \mathrm{q}+\mathrm{S}\right),\\ \tau &= \frac{1}{1+\kappa_g^2}\frac{d\kappa_g}{ds}.\end{aligned}$$

(b) Show that q is planar if and only if the curvature is constant.

(c) Show that $L(\mathrm{T}) \geq L(\mathrm{q})$ with equality only holding for great circles.

(2) Show that for a regular spherical curve $\mathrm{q}(t)$ we have

$$\kappa_g = \frac{\det\begin{bmatrix} \mathrm{q} & \frac{d\,\mathrm{q}}{dt} & \frac{d^2\,\mathrm{q}}{dt^2}\end{bmatrix}}{\left(\frac{ds}{dt}\right)^3}.$$

(3) (Jacobi) Let $\mathrm{q}(s) : [0,L] \to \mathbb{R}^3$ be a closed unit speed curve with positive curvature and consider the unit normal N as a closed curve on S^2.

(a) Show that if s_{N} denotes the arclength parameter of N, then

$$\left(\frac{ds_{\mathrm{N}}}{ds}\right)^2 = \kappa_{\mathrm{q}}^2+\tau_{\mathrm{q}}^2,$$

where κ_{q} and τ_{q} are the curvature and torsion of q.

(b) Show that the geodesic curvature κ_g of N is given by

$$\kappa_g = \frac{\kappa_{\mathrm{q}}\frac{d\tau_{\mathrm{q}}}{ds} - \tau_{\mathrm{q}}\frac{d\kappa_{\mathrm{q}}}{ds}}{\left(\kappa_{\mathrm{q}}^2+\tau_{\mathrm{q}}^2\right)^{\frac{3}{2}}}.$$

(c) Show that

$$\int_0^L \kappa_g(s)\frac{ds_{\mathrm{N}}}{ds}ds = 0.$$

(4) Let $\mathrm{q}(t)$ be a regular closed space curve with positive curvature. Show that if its curvature is $\leq R^{-1}$, then its length is $\geq 2\pi R$.

(5) (Curvature characterization of great circles) Show that great circles $\mathrm{q}(t) = a\cos(t) + b\sin(t)$, where a, b are orthonormal, are unit speed spherical curves with vanishing geodesic curvature. Conversely show that any spherical unit speed curve with vanishing geodesic curvature is a great circle.

(6) Show that the curve

$$\mathrm{q}(t) = (\cos(t)\cos(at), \sin(t)\cos(at), \sin(at))$$

lies on the unit sphere. Compute its curvature.

(7) Show that a simple closed planar curve is unknotted. Show similarly that a simple closed spherical curve is unknotted. Hint: Use the Jordan curve theorem and note that it also holds for spherical curves.

(8) The *trefoil curve* is given by

$$\mathrm{q}(t) = ((a+R\cos(3t))\cos(2t), (a+R\cos(3t))\sin(2t), R\sin(3t)),$$

where $a > b > 0$ and $t \in [0, 2\pi]$. Sketch this curve (it lies on a torus which is created by rotating the circle in the x,z-plane of radius R centered at $(a,0,0)$ around the z-axis) and try to prove that it is knotted.

(9) (Segre, 1947) Let $\mathrm{q}(s) : [0, L] \to \mathbb{R}^3$ be a closed unit speed curve. Show that if $\int |\tau|\, ds = 4R \leq 2\pi$, then the binormal is contained in a spherical cap of radius $\leq R$.

(10) Show that if $\mathrm{C}(\sigma)$ is a unit speed curve on the unit sphere, then for all r, θ the curve

$$\mathrm{q} = r \int \mathrm{C} d\sigma + r \cot\theta \int \mathrm{C} \times \frac{d\mathrm{C}}{d\sigma} d\sigma$$

has a Bertrand mate. Hint: Start by establishing the formulas

$$\begin{aligned}
\mathrm{T} &= \sin\theta \mathrm{C} + \cos\theta \mathrm{C} \times \frac{d\mathrm{C}}{d\sigma}, \\
\mathrm{N} &= \pm \frac{d\mathrm{C}}{d\sigma}, \\
\mathrm{B} &= \pm \left(-\cos\theta \mathrm{C} + \sin\theta \mathrm{C} \times \frac{d\mathrm{C}}{d\sigma} \right).
\end{aligned}$$

Conversely show that any curve that has a Bertrand mate can be written in this way.

(11) Let $\mathrm{q}(s) : [0, L] \to S^2$ be a closed unit speed spherical curve. Show that

$$\int_0^L \frac{\tau}{\kappa} ds = 0.$$

Hint: Use that for a spherical curve $\frac{\tau}{\kappa} = \frac{df}{ds}$ for a suitable function f.

(12) Let $\mathrm{q}(s) : [0, L] \to S^2$ be a unit speed spherical curve and write

$$\mathrm{q} = \alpha(s) \mathrm{T} + \beta(s) \mathrm{N} + \gamma(s) \mathrm{B}.$$

(a) Show that

$$\alpha = 0,\ \beta = -\frac{1}{\kappa},\ \frac{d\beta}{ds} = \tau\gamma.$$

(b) When $\kappa > 1$ show that

$$\tau = \frac{df}{ds}$$

for a suitable function $f(s)$ that only depends on κ and κ'.

(c) When $\kappa > 1$ and q is closed show that

$$\int_0^L \tau ds = 0.$$

(d) When $\kappa(s) = 1$ for $s = 0, L$ and $\kappa > 1$ for $s \in (0, L)$ show that

$$\int_0^L \tau ds = 0.$$

Note this does not rely on q being closed.

(e) Show that if $\kappa = 1$ at only finitely many points and q is closed then

$$\int_0^L \tau ds = 0.$$

This result holds for all closed spherical curves. Segre has also shown that a closed space curve with $\int_0^L \tau ds = 0$ must be spherical.

CHAPTER 4

Basic Surface Theory

In this chapter we define some of the fundamental concepts for surfaces, such as parametrizations, tangents spaces, the first fundamental form, and maps.

4.1. Surfaces

DEFINITION 4.1.1. A *parametrized surface* is defined as a map $\mathrm{q}(u,v) : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ where $\frac{\partial \mathrm{q}}{\partial u}$ and $\frac{\partial \mathrm{q}}{\partial v}$ are linearly independent everywhere on U.

For parametrized surfaces we generally do not worry about self-intersections or other topological pathologies. For example one can parametrize all but the North and South pole of a sphere $S^2(R) = \left\{q \in \mathbb{R}^3 \mid |q| = R > 0\right\}$ using latitudes and meridians:

$$\mathrm{q}(\mu,\phi) = R\begin{bmatrix} \cos\mu\cos\phi \\ \sin\mu\cos\phi \\ \sin\phi \end{bmatrix},$$

where $\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ denotes the latitude and μ the meridian/longitude. This is a valid parametrization of a surface as long as $\cos\phi \neq 0$. This parametrization is called the *equi-rectangular* parametrization and is the most common way of coordinatizing Earth and the sky. Curiously, it predates Cartesian coordinates by about 1500 years and is very likely the oldest parametrization of a surface that is still in use.

DEFINITION 4.1.2. A *reparametrization of a parametrized surface* $\mathrm{q}(u,v) : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ is a parametrized surface $\mathrm{q}(s,t) : O \subset \mathbb{R}^2 \to \mathbb{R}^3$ such that the parameters are smooth functions of each other on their respective domains: $(u,v) = (u(s,t), v(s,t))$ for all $(s,t) \in O$, $(s,t) = (s(u,v), t(u,v))$ for $u,v \in U$, and finally that with these changes we still obtain the same surface $\mathrm{q}(u,v) = \mathrm{q}(s,t)$.

DEFINITION 4.1.3. A map $F : O \to U$ between open sets $O, U \subset \mathbb{R}^2$ is called a *diffeomorphism* if it is one-to-one, onto and both F and the inverse map $F^{-1} : U \to O$ are smooth. Thus a reparametrization is a diffeomorphism between the domains.

When we wish to avoid self-intersections on the surface, then we resort to a more restrictive class of surfaces that come from the next two general constructions. For curves this corresponds to the notion of being simple and in that case we could have used the approach we shall take for surfaces.

The first construction is to use a particularly nice way of parametrizing surfaces without self-intersections or other nasty topological problems. These are the three different types of parametrizations where the surface is represented as a smooth graph:

$$\begin{aligned}
\mathrm{q}(u,v) &= (u,v,f(u,v)),\\
\mathrm{q}(u,v) &= (u,f(u,v),v),\\
\mathrm{q}(u,v) &= (f(u,v),u,v).
\end{aligned}$$

They are also known as *Monge patches*.

EXAMPLE 4.1.4. The western hemisphere on $S^2(1)$ can be parametrized using the y,z coordinates

$$\mathrm{q}(u,v)=\begin{bmatrix} -\sqrt{1-u^2-v^2} \\ u \\ v \end{bmatrix},$$

where $(u,v)\in U=\left\{u^2+v^2<1\right\}$. Using latitudes/meridians the parametrization is instead

$$\mathrm{q}(\mu,\phi)=\begin{bmatrix} \cos\mu\cos\phi \\ \sin\mu\cos\phi \\ \sin\phi \end{bmatrix},$$

with $(\mu,\phi)\in\left(\frac{\pi}{2},\frac{3\pi}{2}\right)\times\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$. Setting these two expressions equal to each other tells us that

$$\begin{bmatrix} u \\ v \end{bmatrix}=G(\mu,\phi)=\begin{bmatrix} \sin\mu\cos\phi \\ \sin\phi \end{bmatrix}.$$

This map is smooth and it is not hard to check that as a map from $\left(\frac{\pi}{2},\frac{3\pi}{2}\right)\times\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ to U it is one-to-one and onto. The differential is

$$DG=\frac{\partial(u,v)}{\partial(\mu,\phi)}=\begin{bmatrix} \frac{\partial u}{\partial\mu} & \frac{\partial u}{\partial\phi} \\ \frac{\partial v}{\partial\mu} & \frac{\partial v}{\partial\phi} \end{bmatrix}=\begin{bmatrix} \cos\mu\cos\phi & -\sin\mu\sin\phi \\ 0 & \cos\phi \end{bmatrix}.$$

The determinant is $\cos\mu\cos^2\phi$ which is always negative on our domain. The inverse function theorem then guarantees us that G is indeed a diffeomorphism. In this case it is also possible to construct the inverse using inverse trigonometric functions.

THEOREM 4.1.5. *Let $\mathrm{q}(u,v):U\subset\mathbb{R}^2\to\mathbb{R}^3$ be a parametrized surface. For every $(u_0,v_0)\in U$ there exists a neighborhood $(u_0,v_0)\in V\subset U$ such that the smaller parametrized surface $\mathrm{q}(u,v):V\to\mathbb{R}^3$ can be represented as a Monge patch.*

PROOF. By assumption the matrix

$$\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix}=\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{bmatrix}$$

always has rank 2. Assume for the sake of argument that at (u_0,v_0) the middle row is a linear combination of the other two rows. Then the matrix

$$\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{bmatrix}$$

is nonsingular at (u_0,v_0). Thus the map $(x,z)=(x(u,v),z(u,v)):U\to\mathbb{R}^2$ has nonsingular differential at (u_0,v_0). The inverse function theorem then tells us that there must exist neighborhoods $(u_0,v_0)\in V\subset U$ and $(x(u_0,v_0),x(u_0,v_0))\in O\subset\mathbb{R}^2$ such that the function $(x,z)=(x(u,v),z(u,v)):V\to O$ can be smoothly inverted, i.e., there is a smooth inverse $(u,v)=(u(x,z),v(x,z)):O\to V$ that allows us to smoothly solve for (u,v) in terms of (x,z). This gives us the desired reparametrization to a Monge patch

$$\begin{bmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{bmatrix} = \mathrm{q}(u,v) = \mathrm{q}(x,z) = \begin{bmatrix} x \\ y(u(x,z), v(x,z)) \\ z \end{bmatrix}.$$

□

Definition 4.1.6. A *surface* is a subset $M \subset \mathbb{R}^3$ with the property that any $q \in M$ is contained in an open set $O \subset \mathbb{R}^3$ such that $O \cap M$ can be represented as a Monge patch, i.e., it is locally a smooth graph over one of the three coordinate planes.

A parametrization $\mathrm{q}(u,v) : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ is called a *coordinate system* if the map is one-to-one and the image $\mathrm{q}(U)$ is a surface.

Example 4.1.7. Despite the above theorem not all parametrized surfaces are surfaces in this restrictive sense. Let $\mathrm{q}(u) = (x(u), y(u))$ be a regular planar curve and consider the parametrized surface $\mathrm{q}(u,v) = (x(u), y(u), v)$. This might not be a surface if the planar curve looks like a figure 8. We could also take something like a figure 6 but parametrize it so that the loop gets arbitrarily close without intersecting. In the latter case we simply parametrize the figure 6 using an open interval $(0,1)$.

The second construction comes from considering level sets. A *level set* is a set of the form

$$\{(x,y,z) \in O \mid F(x,y,x) = c\},$$

where c is a fixed constant and $O \subset \mathbb{R}^3$ is an open set.

Example 4.1.8. For example

$$x^2 + y^2 + z^2 = R^2$$

describes the sphere as a level set. Depending on where we are on the sphere different parametrizations are possible. At points where, say, $y < 0$ we can use

$$\mathrm{q}(u,v) = \left(u, -\sqrt{R^2 - u^2 - v^2}, v\right)$$

This will in fact parametrize all points where $y < 0$ if we use all (u,v) with $u^2+v^2 < R^2$.

The implicit function theorem tells us when level sets are surfaces.

Theorem 4.1.9. *Let $F : O \to \mathbb{R}$ be a smooth function and $c \in \mathbb{R}$ a constant. The level set*

$$M = \{(x,y,z) \in O \mid F(x,y,x) = c\}$$

is a smooth surface if it is not empty and for all $q \in M$ the gradient

$$\nabla F(q) = \begin{bmatrix} \frac{\partial F}{\partial x}(q) \\ \frac{\partial F}{\partial y}(q) \\ \frac{\partial F}{\partial z}(q) \end{bmatrix} \neq 0.$$

Proof. Fix $q = (x_0, y_0, z_0) \in M$ and assume for the sake of argument that $\frac{\partial F}{\partial y}(q) \neq 0$. The implicit function theorem tells us that there are neighborhoods

$q \in O_1 \subset O$ and $(x_0, z_0) \in U \subset \mathbb{R}^2$ as well as a smooth function $f(u,v) : U \to \mathbb{R}$ such that for all $(u,v) \in U$ we have $(u, f(u,v), v) \in O_1$ and

$$M \cap O_1 = \{(x,y,z) \in O_1 \mid F(x,y,x) = c\} = \{(u, f(u,v), v) \mid (u,v) \in U\}.$$

Thus $M \cap O_1$ can be written as a graph over the (x,z)-plane. □

EXAMPLE 4.1.10. A *generalized helicoid* is a surface of the form

$$\mathrm{q}(u,v) = (u\cos v, u\sin v, f(u) + cv).$$

Note that the v-curves given by holding u constant are helices. In case $c = 0$ we obtain *surfaces of revolution* as the v-curves revolve in circles around the z-axis. To check when this is a parametrized surface we calculate

$$\frac{\partial \mathrm{q}}{\partial u} = \begin{bmatrix} \cos v \\ \sin v \\ f'(u) \end{bmatrix}, \frac{\partial \mathrm{q}}{\partial v} = \begin{bmatrix} -u\sin v \\ u\cos v \\ c \end{bmatrix}.$$

When $c \neq 0$ these two vectors are always non-zero and linearly independent. Moreover, if $\mathrm{q}(u_1, v_1) = \mathrm{q}(u,v)$, then we see by looking at the x- and y-coordinates that either $u_1 = u$ and $v_1 = v + 2n\pi$ or $u_1 = -u$ and $v_1 = v + (2n+1)\pi$. In the first case the z-coordinates will be different. In the second case the z-coordinates can be equal if there are u values where

$$f(u) = f(-u) + (2n+1)c.$$

When $c = 0$ we obtain a parametrized surface as long as $u \neq 0$. Thus we might as well assume that $u > 0$ in this case. The parametrization is never one-to-one as the v-curves form circles. In the special case where $c = 0$ and $f' = 0$ the surface is a plane perpendicular to the z-axis that is parametrized using polar coordinates.

Exercises

(1) A *generalized cylinder* is determined by a regular curve $c(t)$ and a vector X that is never tangent to the curve. It consists of the lines that are parallel to the vector and pass through the curve.
   (a) Show that
   $$\mathrm{q}(s,t) = c(t) + sX$$
   is a natural parametrization and show that it gives a parametrized surface.
   (b) Show that we can reconstruct the cylinder so that the curve lies in the plane perpendicular to the vector X. Hint: Try the case where $X = (0,0,1)$ and the plane is the (x,y)-plane and make sure your new parametrization is a valid parametrization precisely when the old parametrization was valid.
   (c) Find the equation for a generalized cylinder when the curve c in the (x,y)-plane is given by $F(x,y) = C$ and $X = (0,0,1)$.
(2) A *generalized cone* is generated by a regular curve $c(t)$ and a point p not on the curve. It consists of the lines that pass through the point and the curve.
   (a) Show that
   $$\mathrm{q}(s,t) = s(c(t) - p) + p$$
   is a natural parametrization and determine when/where it yields a parametrized surface.
   (b) Show that we can replace $c(t)$ by a curve $c^*(t)$ that lies on a unit sphere centered at the vertex p of the cone.

(c) Show that the level set $F(x, y, z) = 0$ is a cone through the origin when F is homogeneous, i.e., there is an $\alpha \neq 0$ such that $F(\lambda x, \lambda y, \lambda z) = \lambda^\alpha F(x, y, z)$.

(d) Further show that the condition for the cone in (c) to be a smooth surface away from the origin is that $\nabla F(q)$ is not proportional to q, when $F(q) = 0$.

(3) A *ruled surface* is given by a parametrization of the form

$$\mathrm{q}(s, t) = c(t) + sX(t).$$

It is evidently a surface that is a union of lines (rulers) and generalizes the constructions in the previous exercises. Give conditions on c, X and the parameter s that guarantee we get a parametrized surface. A special case occurs when X is tangent to c. These are also called *tangent developables*.

(4) A *surface of revolution* is determined by a planar regular curve and a line in the same plane. The surface is generated by rotating the curve around the line.

(a) Show that for a regular curve $(r(t), z(t))$ in the (x, z)- plane that is rotated around the z-axis the parametrization is

$$\mathrm{q}(t, \mu) = (r(t)\cos\mu, r(t)\sin\mu, z(t))$$

and show that it is a parametrized surface.

(b) Show that the equation for the surface is $F\left(\sqrt{x^2 + y^2}, z\right) = c$ when the curve is given by $F(r, z) = c$ with $x > 0$.

(c) Show that the equation for the surface is given by $R^2(z) = x^2 + y^2$, when the curve can be written as $r = R(z)$.

(5) Let $\mathrm{q}(z, \mu) = \left(\sqrt{1 - z^2}\cos\mu, \sqrt{1 - z^2}\sin\mu, z\right)$ with $-1 < z < 1$ and $-\pi < \mu < \pi$. Show that q defines a surface. What is the surface?

(6) Consider a regular curve $c(t)$ with non-vanishing curvature and construct the *tube* of radius R around it

$$\mathrm{q}(t, \phi) = c(t) + R(\mathrm{N}_c \cos\phi + \mathrm{B}_c \sin\phi),$$

where N_c, B_c are the normal and binormal to the curve.

(a) Show that this defines a parametrized surface as long as $\kappa_c < R^{-1}$.

(b) Show by example that this surface might intersect itself if there is a cord of length $< 2R$ that is normal to the curve at both end points.

(c) Show that when c is a circle, then we obtain a surface of revolution that looks like a torus.

(7) Consider the level set:

$$z\left(x^2 + y^2\right) - 2xy = R.$$

(a) Show that this defines a surface when $R \neq 0$.

(b) When $R = 0$ show that we get a surface as long as $(x, y) \neq (0, 0)$. This is called *Plücker's conoid*.

(c) When $R = 0$ show that it is a ruled surface. Hint: Use polar coordinates $x = t\cos\theta$ and $y = t\sin\theta$.

(8) Show that

$$\left(x^2 + y^2 + z^2 + R^2 - r^2\right)^2 = 4R^2\left(x^2 + y^2\right)$$

defines a surface when $R > r > 0$. Show that it is rotationally symmetric and a torus, i.e., it is a parametrized circle rotated around the z-axis.

(9) The *helicoid* is given by the equation

$$\tan \frac{z}{h} = \frac{y}{x}$$

where $h \neq 0$ is a fixed constant.
(a) Show that this defines a surface for suitable (x, y, z).
(b) Show that the surface can be parametrized by

$$\mathrm{q}\,(r, \theta) = (r \cos \theta, r \sin \theta, h\theta)$$

and determine for which r, θ this defines a parametrized surface. Note that for fixed r we obtain helices.

(10) *Enneper's surface* is defined by the parametrization

$$\mathrm{q}\,(u, v) = \left(u - \frac{1}{3}u^3 + uv^2, v - \frac{1}{3}v^3 + vu^2, u^2 - v^2 \right).$$

(a) For which u, v does this define a parametrization?
(b) Show that Enneper's surface satisfies the equation

$$\left(\frac{y^2 - x^2}{2z} + \frac{2z^2}{9} + \frac{2}{3} \right)^3 = 6 \left(\frac{y^2 - x^2}{4z} - \frac{1}{4} \left(x^2 + y^2 + \frac{8}{9}z^2 \right) + \frac{2}{9} \right)^2.$$

(11) Show that there exists a parametrization $\mathrm{q} : U \to S^2 \subset \mathbb{R}^3$ that covers the entire sphere. Hint: If U is the disjoint union of two discs, then this can be done using two maps that each cover a region slightly larger than a hemisphere. Now connect these discs by a band in $\mathbb{R}^2$ and similarly in the sphere to obtain a parametrization from a domain U that is diffeomorphic to $\mathbb{R}^2$.

(12) Many classical surfaces are of the form

$$F\,(x, y, z) = ax^2 + by^2 + cz^2 + dx + ey + fz + g = 0.$$

These are called *quadratic surfaces* if one of a, b, or $c \neq 0$.
(a) Show that it is either empty or defines a surface at points $(x, y, z) \in \mathbb{R}^3 - C$, where C is the closed set defined by

$$C = \{(x, y, z) \mid 2ax = -d,\ 2by = -e,\ 2cz = -f\}.$$

(b) Show that C is either empty, a point, a line, or a plane.
(c) Under what conditions does it become a surface of revolution around the z-axis?
(d) Show that when the equation does not depend on one of the coordinates, then we obtain a generalized cylinder.
(e) When, say $c = 0$, but $abf \neq 0$ we obtain a *paraboloid*. It is *elliptic* when a, b have the same sign and otherwise *hyperbolic*. Draw pictures of these two situations.
(f) When $abc \neq 0$ show that it can be rewritten in the form

$$F\,(x, y, z) = a\,(x - x_0)^2 + b\,(y - y_0)^2 + c\,(z - z_0)^2 + h = 0.$$

(g) When all three a, b, c have the same sign show that it is either empty or an *ellipsoid*.
(h) When not all of a, b, c have the same sign and $h \neq 0$ we obtain a *hyperboloid*. Show that it might be connected or have two components (called sheets) depending of the signs of all four constants.
(i) When not all of a, b, c have the same sign and $h = 0$ we obtain a cone.

(j) Given constants a_x, a_y, a_z determine when

$$\begin{aligned}
\mathrm{q}(u,v) &= (a_x \cos u \cos v, a_y \cos u \sin v, a_z \sin u),\\
\mathrm{q}(u,v) &= (a_x \sinh u \cos v, a_y \sinh u \sin v, a_z \cosh u),\\
\mathrm{q}(u,v) &= (a_x \cosh u \cos v, a_y \cosh u \sin v, a_z \sinh u),\\
\mathrm{q}(u,v) &= \left(a_x u \cos v, a_y u \sin v, a_z u^2\right),\\
\mathrm{q}(u,v) &= \left(a_x u \cosh v, a_y u \sinh v, a_z u^2\right),
\end{aligned}$$

yield parametrizations and identify them with the appropriate quadratics.

4.2. Tangent Spaces and Maps

DEFINITION 4.2.1. The *tangent space* at $q \in M$ of a (parametrized) surface is defined as

$$T_q M = \operatorname{span}\left\{\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}\right\},$$

and *normal space* is the orthogonal complement

$$N_q M = (T_q M)^{\perp}.$$

REMARK 4.2.2. For a parametrized surface with self-intersections this is a bit ambivalent as the tangent space in that case depends on the parameter values (u,v) and not just the point $q = \mathrm{q}(u,v)$. This is just as for curves where the tangent line at a point really is the tangent line at a point with respect to a specific parameter value.

REMARK 4.2.3. Note that the tangent and normal spaces are subspaces. We can also define the *tangent plane* at $q \in M$, as the plane parallel to $T_q M$ that contains q. The tangent plane is then similar to the tangent line for a curve. Similarly, the *normal line* to $q \in M$ is the line through q that is parallel to $N_q M$. A *normal* to M at q is a choice of a unit vector in $\mathrm{n} \in N_q M$. There are two normals at each point. In the rest of these notes n will always denote a unit normal to a surface. The principal normal to a curve c will be denoted N_c.

EXAMPLE 4.2.4. A parametrized surface $\mathrm{q}(u,v) : U \to \mathbb{R}^3$ always has a natural normal $\mathrm{n}(u,v)$ defined by

$$\mathrm{n}(u,v) = \frac{\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}}{\left|\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right|}$$

However, it is possible (as well shall see in the exercises) that there are parameter values that give the same points and tangent spaces to the surface without giving the same normal vectors.

EXAMPLE 4.2.5 (Example 4.1.10 continued). The normal of the generalized helicoid is given by

$$\mathrm{n}(u,v) = \frac{\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}}{\left|\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right|} = \frac{1}{\sqrt{c^2 + u^2\left(1 + (f')^2\right)}} \begin{bmatrix} c \sin v - u f' \cos v \\ -c \cos v - u f' \sin v \\ u \end{bmatrix}.$$

PROPOSITION 4.2.6. *Both tangent and normal spaces are subspaces that do not change under reparametrization.*

PROOF. This would seem intuitively clear, just as with curves, where the tangent line does not depend on parametrizations. For curves it boils down to the simple fact that velocities for different parametrizations are proportional. With surfaces something similar happens, but it is a bit more involved. Suppose we have two different parametrizations of the same surface:

$$\mathrm{q}(s,t) = \mathrm{q}(u,v).$$

This tells us that the parameters are functions of each other

$$\begin{aligned} u &= u(s,t),\ v = v(s,t), \\ s &= s(u,v),\ t = t(u,v). \end{aligned}$$

The chain rule then gives us

$$\frac{\partial \mathrm{q}}{\partial u} = \frac{\partial \mathrm{q}}{\partial s}\frac{\partial s}{\partial u} + \frac{\partial \mathrm{q}}{\partial t}\frac{\partial t}{\partial u} \in \operatorname{span}\left\{\frac{\partial \mathrm{q}}{\partial s}, \frac{\partial \mathrm{q}}{\partial t}\right\},$$

and

$$\frac{\partial \mathrm{q}}{\partial v} \in \operatorname{span}\left\{\frac{\partial \mathrm{q}}{\partial s}, \frac{\partial \mathrm{q}}{\partial t}\right\}.$$

In the other direction we similarly have

$$\frac{\partial \mathrm{q}}{\partial s}, \frac{\partial \mathrm{q}}{\partial t} \in \operatorname{span}\left\{\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}\right\}.$$

This shows that at a fixed point q on a surface, the tangent space does not depend on how the surface is parametrized. The normal space is then also well-defined. □

Note that the chain rule shows in matrix notation that

$$\begin{aligned} \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{bmatrix} \begin{bmatrix} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{bmatrix}, \\ \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{bmatrix} &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial s} & \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial t} \end{bmatrix} \end{aligned}$$

with

$$\begin{bmatrix} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{\partial u}{\partial s} & \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial t} \end{bmatrix}.$$

A better way of defining the tangent space that also shows that it is independent of parametrizations comes from the next result.

PROPOSITION 4.2.7. *The tangent space at $q = \mathrm{q}(u_0, v_0)$ for a (parametrized) surface is given by*

$$T_qM = \left\{\mathrm{v} \in \mathbb{R}^3 \mid \mathrm{v} = \frac{d\,\mathrm{q}}{dt}(0) \text{ for a smooth curve } \mathrm{q}(t) : I \to M \text{ with } \mathrm{q}(0) = q\right\}.$$

PROOF. Any curve $\mathrm{q}(t)$ on the surface that passes through q at $t = 0$ can be written as

$$\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$$

for smooth functions $u(t)$ and $v(t)$ with $u(0) = u_0$ and $v(0) = v_0$ as long as t is sufficiently small. This is because the parametrization is locally one-to-one. If we write the curve this way, then

$$\frac{d\,\mathrm{q}}{dt} = \frac{\partial \mathrm{q}}{\partial u}\frac{du}{dt} + \frac{\partial \mathrm{q}}{\partial v}\frac{dv}{dt}.$$

Showing that velocities of curves on the surface are always tangent vectors. Conversely by using $u(t) = at + u_0$ and $v(t) = bt + v_0$ we obtain all possible linear combinations of tangent vectors as

$$\frac{d\,\mathrm{q}}{dt}(0) = \frac{\partial\,\mathrm{q}}{\partial u}a + \frac{\partial\,\mathrm{q}}{\partial v}b.$$

$\square$

COROLLARY 4.2.8. *Let $M = \{(x, y, z) \in O \mid F(x, y, z) = c\}$ be a level set as in theorem 4.1.9. The normal space is spanned by*

$$\nabla F(q) = \begin{bmatrix} \frac{\partial F}{\partial x}(q) \\ \frac{\partial F}{\partial y}(q) \\ \frac{\partial F}{\partial z}(q) \end{bmatrix}.$$

PROOF. We saw in proposition 4.2.7 that any tangent vector in T_qM can be represented as a velocity vector $\dot{\mathrm{q}}(0)$. Since $\mathrm{q}(t) \in M$ it follows that $F(\mathrm{q}(t)) = c$ for all t. The chain rule then implies that

$$0 = \nabla F(\mathrm{q}(0)) \cdot \dot{\mathrm{q}}(0) = \nabla F(q) \cdot \dot{\mathrm{q}}(0).$$

This shows that the gradient is perpendicular to all tangent vectors and hence a normal vector. $\square$

EXAMPLE 4.2.9. The sphere of radius R centered at the origin has a unit normal given by the unit radial vector at $q = (x, y, z) \in S^2(R)$

$$\mathrm{n} = \frac{1}{R}\mathrm{q} = \frac{1}{R}\begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

The basis for the tangent space with respect to the meridian/latitude parametrization is

$$\frac{\partial\,\mathrm{q}}{\partial \mu} = R\begin{bmatrix} -\sin\mu\cos\phi \\ \cos\mu\cos\phi \\ 0 \end{bmatrix}, \quad \frac{\partial\,\mathrm{q}}{\partial \phi} = R\begin{bmatrix} -\cos\mu\sin\phi \\ -\sin\mu\sin\phi \\ \cos\phi \end{bmatrix}.$$

It is often useful to find coordinates suited to a particular situation. However, unlike for curves, it isn't always possible to parametrize a surface such that the coordinate curves are unit speed and orthogonal to each other. But there is one general construction we can do.

THEOREM 4.2.10. *Assume that we have linearly independent tangent vector fields X, Y defined on a surface M. Then it is possible to find a parametrization $\mathrm{q}(u, v)$ in a neighborhood of any point such that $\frac{\partial\,\mathrm{q}}{\partial u}$ is proportional to X and $\frac{\partial\,\mathrm{q}}{\partial v}$ is proportional to Y.*

PROOF. The vector fields have integral curves forming a net on the surface. Apparently the goal is to reparametrize the curves in this net in some fashion. The difficulty lies in ensuring that the levels where u is constant correspond to the v-curves, and vice versa. We proceed as with the classical construction of Cartesian coordinates. Select a point p and let the u-axis be the integral curve for X through p, similarly let the v-axis be the integral curve for Y through p. Both of these curves retain the parametrizations that make them integral curves for X and Y. Thus p will naturally correspond to $(u, v) = (0, 0)$. We now wish to assign (u, v) coordinates to a point q near p. There are also unique integral curves for X and Y through

q. These will be our way of projecting onto the chosen axes and will in this way yield the desired coordinates. Specifically, $u(q)$ is the parameter where the integral curve for Y through q intersects the u-axis, and similarly with $v(q)$. In general, integral curves can intersect axes in several places or might not intersect them at all. However, a continuity argument offers some justification when we consider that the axes themselves are the proper integral curves for the qs that lie on these axes and so when q sufficiently close to both axes it should have a well-defined set of coordinates. We also note that as the projection happens along integral curves we have ensured that coordinate curves are simply reparametrizations of integral curves. To completely justify this construction we need to know quite a bit about the existence, uniqueness and smoothness of solutions to differential equations and the inverse function theorem. □

REMARK 4.2.11. Note that this proof gives us a little more information. Specifically, we obtain a parametrization where the parameter curves through $(0,0)$ are the integral curves for X and Y.

DEFINITION 4.2.12. A *map* between surfaces $F : M_1 \to M_2$ is an assignment of points in the first surface to points in the second. The map is *smooth* if around every point $q \in M_1$ we can find a parametrization $\mathrm{q}_1(u,v)$ where $q = \mathrm{q}_1(u_0,v_0)$ such that the composition $F \circ \mathrm{q}_1 : U \to \mathbb{R}^3$ is a smooth map as a map from the space of parameters to the ambient space that contains the target M_2.

We can also define maps between parametrized surfaces in a similar way. Clearly parametrizations are themselves smooth maps. It is also often the case that the compositions $F \circ \mathrm{q}_1$ are themselves parametrizations for M_2.

EXAMPLE 4.2.13. Two classical examples of maps are the *Archimedes* and *Mercator projections* from the sphere to the cylinder of the same radius placed to touch the sphere at the equator. We give the formulas for the unit sphere and note that neither map is defined at the poles.

The Archimedes map is simply a horizontal projection that preserves the z-coordinate

$$A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} \\ \frac{y}{\sqrt{x^2+y^2}} \\ z \end{bmatrix}.$$

In the meridian/latitude parametrization it looks particularly nice:

$$A\begin{bmatrix} \cos\mu\cos\phi \\ \sin\mu\cos\phi \\ \sin\phi \end{bmatrix} = \begin{bmatrix} \cos\mu \\ \sin\mu \\ \sin\phi \end{bmatrix}.$$

Note that what is here referred to as the Archimedes map is often called the Lambert projection. However, Archimedes was the first to discover that the areas of the sphere and cylinder are equal. This will be discussed in greater detail in section 4.4.

The Mercator projection (1569) differs in that the z-coordinate is not preserved:

$$M\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} \\ \frac{y}{\sqrt{x^2+y^2}} \\ \frac{1}{2}\log\frac{1+z}{1-z} \end{bmatrix}$$

or

$$M\begin{bmatrix}\cos\mu\cos\phi\\ \sin\mu\cos\phi\\ \sin\phi\end{bmatrix}=\begin{bmatrix}\cos\mu\\ \sin\mu\\ \frac{1}{2}\log\frac{1+\sin\phi}{1-\sin\phi}\end{bmatrix}.$$

Both of these maps really are maps in the traditional sense that they can be used to picture the Earth on a flat piece of paper by cutting the cylinder vertically and unfolding it. This unfolding is done along a meridian. For Eurocentric people it is along the date line. In the Americas one also sees maps cut along a meridian that bisects Asia so as to place the Americas in the center.

Definition 4.2.14. The *differential* of a smooth map $F : M_1 \to M_2$ at $q \in M_1$ is the map

$$DF_q : T_qM_1 \to T_{F(q)}M_2$$

defined by

$$DF_q(\mathrm{v}) = \frac{dF\circ \mathrm{q}}{dt}(0)$$

if $\mathrm{q}(t)$ is a curve (in M_1) with $q = \mathrm{q}(0)$ and $\mathrm{v} = \frac{d\,\mathrm{q}}{dt}(0)$.

Proposition 4.2.15. *When* $\mathrm{v} = \frac{d\,\mathrm{q}}{dt}(0) = \frac{\partial\,\mathrm{q}}{\partial u}\mathrm{v}^u + \frac{\partial\,\mathrm{q}}{\partial v}\mathrm{v}^v$ *we have*

$$DF_q(\mathrm{v}) = \begin{bmatrix}\frac{\partial F\circ \mathrm{q}}{\partial u} & \frac{\partial F\circ \mathrm{q}}{\partial v}\end{bmatrix}\begin{bmatrix}\mathrm{v}^u\\ \mathrm{v}^v\end{bmatrix}.$$

In particular, the differential is a linear map and is completely determined by the two partial derivatives $\frac{\partial F\circ \mathrm{q}}{\partial u}, \frac{\partial F\circ \mathrm{q}}{\partial v}$.

Proof. This follows from the chain rule:

$$\begin{aligned}\frac{dF\circ \mathrm{q}}{dt}(t) &= \frac{dF(\mathrm{q}(t))}{dt}\\ &= \frac{dF(\mathrm{q}(u(t),v(t)))}{dt}\\ &= \frac{\partial F\circ \mathrm{q}}{\partial u}\frac{du}{dt}+\frac{\partial F\circ \mathrm{q}}{\partial v}\frac{dv}{dt}\\ &= \begin{bmatrix}\frac{\partial F\circ \mathrm{q}}{\partial u} & \frac{\partial F\circ \mathrm{q}}{\partial v}\end{bmatrix}\begin{bmatrix}\frac{du}{dt}\\ \frac{dv}{dt}\end{bmatrix}.\end{aligned}$$

□

Example 4.2.16. The Archimedes map satisfies

$$\frac{\partial(A\circ \mathrm{q})}{\partial\mu}=\begin{bmatrix}-\sin\mu\\ \cos\mu\\ 0\end{bmatrix}, \quad \frac{\partial(A\circ \mathrm{q})}{\partial\phi}=\begin{bmatrix}0\\ 0\\ -\cos\phi\end{bmatrix}$$

and the Mercator map

$$\frac{\partial(M\circ \mathrm{q})}{\partial\mu}=\begin{bmatrix}-\sin\mu\\ \cos\mu\\ 0\end{bmatrix}, \quad \frac{\partial(M\circ \mathrm{q})}{\partial\phi}=\begin{bmatrix}0\\ 0\\ \frac{1}{\cos\phi}\end{bmatrix}.$$

Definition 4.2.17. A surface M is *orientable* if we can select a smooth normal field. Thus we require a smooth function

$$\mathrm{n} : M \to S^2(1) \subset \mathbb{R}^3$$

such that for all $q \in M$ the vector $\mathrm{n}(q)$ is perpendicular to the tangent space $T_q M$. The map $\mathrm{n} : M \to S^2(1)$ is called the *Gauss map*.

PROPOSITION 4.2.18. *A surface which is given as a level set is orientable.*

PROOF. Form corollary 4.2.8 we know that the normal can be given by

$$\mathrm{n} = \frac{\nabla F}{|\nabla F|}$$

if $M = \{q \in O \mid F(q) = c\}$. $\square$

DEFINITION 4.2.19. The parameters u, v on a parameterized surface $\mathrm{q}(u, v)$ define two differentials du and dv. These are not mysterious infinitesimals, but linear functions on tangent vectors to the surface that compute the coefficients of the vector with respect to the basis $\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}$. Thus

$$\begin{aligned} du(\mathrm{v}) &= du\left(\frac{\partial \mathrm{q}}{\partial u}\mathrm{v}^u + \frac{\partial \mathrm{q}}{\partial v}\mathrm{v}^v\right) = \mathrm{v}^u, \\ dv(\mathrm{v}) &= dv\left(\frac{\partial \mathrm{q}}{\partial u}\mathrm{v}^u + \frac{\partial \mathrm{q}}{\partial v}\mathrm{v}^v\right) = \mathrm{v}^v, \end{aligned}$$

and

$$\mathrm{v} = \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix} (\mathrm{v}) = \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} \mathrm{v}^u \\ \mathrm{v}^v \end{bmatrix}.$$

From the chain rule we obtain the very natural transformation laws for differentials

$$\begin{aligned} du &= \frac{\partial u}{\partial s}ds + \frac{\partial u}{\partial t}dt, \\ dv &= \frac{\partial v}{\partial s}ds + \frac{\partial v}{\partial t}dt, \end{aligned}$$

or

$$\begin{bmatrix} du \\ dv \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial s} & \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial t} \end{bmatrix} \begin{bmatrix} ds \\ dt \end{bmatrix}.$$

Exercises

(1) Show that the following conditions for a surface are equivalent:
  (a) It is part of a plane.
  (b) The normal vector is constant.
  (c) All the tangent planes are parallel.
(2) Show that the following conditions for a surface are equivalent:
  (a) It is part of a sphere.
  (b) All normal lines pass through a fixed point.
  (c) $n = \lambda(\mathrm{q} - \mathrm{c})$ for some function λ and point c. Hint: Select a parametrization, show that $\frac{\partial n}{\partial w}$ is a tangent vector for $w = u, v$, and use this to show that λ must be constant.
(3) Show that the ruled surface

$$\mathrm{q}(t, \phi) = (\cos\phi, \sin\phi, 0) + t\left(\sin\frac{\phi}{2}\cos\phi, \sin\frac{\phi}{2}\sin\phi, \cos\frac{\phi}{2}\right)$$

defines a parametrized surface. It is called the *Möbius band*. Show that it is not orientable by showing that when $t = 0$ and $\phi = \pm\pi$ we obtain the same point and tangent space on the surface, but the normals

$$\mathrm{n}(t,\phi) = \frac{\frac{\partial \mathrm{q}}{\partial t} \times \frac{\partial \mathrm{q}}{\partial \phi}}{\left|\frac{\partial \mathrm{q}}{\partial t} \times \frac{\partial \mathrm{q}}{\partial \phi}\right|}$$

are not the same.

(4) Show that $\mathrm{q}(t,\phi) = t(\cos\phi, \sin\phi, 1)$ defines a parametrization for $(t,\phi) \in (0,\infty)\times\mathbb{R}$. Show that the corresponding surface is $x^2+y^2-z^2 = 0$, $z > 0$. Show that this parametrization is not one-to-one. Find a different parametrization of the entire surface that is one-to-one.

(5) Consider the two surfaces M_1 and M_2 defined by the parametrizations:

$$\begin{aligned} \mathrm{q}_1(t,\phi) &= (\sinh\phi\cos t, \sinh\phi\sin t, t) \\ &= (0,0,t) + \sinh\phi(\cos t, \sin t, 0), \\ \mathrm{q}_2(t,\phi) &= (\cosh t\cos\phi, \cosh t\sin\phi, t). \end{aligned}$$

(a) Show that $\mathrm{q}_1 : \mathbb{R}\times\mathbb{R} \to M_1$ is a one-to-one parametrization of a helicoid (see section 4.1 exercise 9).

(b) Show that q_2 is a parametrization that is not one-to-one. Show that M_2 is rotationally symmetric (see section 4.1 exercise 4) and can also be described by the equation

$$x^2 + y^2 = \cosh^2 z.$$

Show further that this equation defines a surface. It is called the *catenoid*.

(c) Define a map $F : M_1 \to M_2$ by $F \circ \mathrm{q}_1(t,\phi) = \mathrm{q}_2(t,\phi)$. Show that this map is smooth, not one-to-one, but locally a diffeomorphism.

(6) Show that a parametrized surface

$$\mathrm{q}(z,\theta) = \begin{bmatrix} r(z,\theta)\cos\theta \\ r(z,\theta)\sin\theta \\ z \end{bmatrix}$$

is rotationally symmetric, i.e., $\frac{\partial r}{\partial \theta} = 0$, if all its normal lines pass through the z-axis.

(7) The *inversion in the unit sphere or circle* is defined as

$$F(q) = \frac{q}{|q|^2}.$$

(a) Show that this is a diffeomorphism of $\mathbb{R}^n - 0$ to it self with the property that $q \cdot F(q) = 1$.

(b) Show that F preserves the unit sphere, but reverses the unit normal directions.

(c) Let M be a surface. Show that $M^* = F(M)$ defines another surface. Show that $DF : T_qM \to T_{q^*}M^*$ satisfies

$$DF(v) = \frac{|q|^2 v - 2(q\cdot v)q}{|q|^4}.$$

(d) Show that if n is a unit normal to M, then the unit normal to M^* is given by

$$\mathrm{n}^* = \pm\left(\mathrm{n} - q\frac{2q\cdot\mathrm{n}}{|q|^2}\right).$$

(8) A *perspective projection* is defined as a radial projection along lines emanating from a fixed point $\mathrm{c}\in\mathbb{R}^n$ to a hyper-plane $H\subset\mathbb{R}^n$.

(a) Let $\mathrm{c}=(0,0,c)\in\mathbb{R}^3$ and H be the (x,y)-plane. Show that the projection is given by $(x,y,z)\mapsto\left(\frac{cx}{c-z},\frac{cy}{c-z},0\right)$.

(b) Let $\mathrm{c}=(0,0,0)\in\mathbb{R}^3$ and H be the $\{z=1\}$-plane. Show that the projection is given by $(x,y,z)\mapsto\left(\frac{x}{z},\frac{y}{z},1\right)$.

(c) Let $\mathrm{c}=(0,0,1)\in\mathbb{R}^3$ and H be the $\{z=-1\}$-plane. Show that the projection is given by $(x,y,z)\mapsto\left(\frac{2x}{1-z},\frac{2y}{1-z},-1\right)$.

(9) Consider the two maps $\mathrm{q}^\pm:\mathbb{R}^n\to\mathbb{R}^n\times\mathbb{R}=\mathbb{R}^{n+1}$

$$\mathrm{q}^\pm(q)=(q,0)+\frac{1-|q|^2}{1+|q|^2}(q,\pm1).$$

These two maps are inverses of perspective projections to the unit sphere. They are also called *stereographic projections*.

(a) Show that these maps are one-to-one, map into the unit sphere, and that together they cover the unit sphere.

(b) Show that they are the inverse maps of the perspective projections from $(0,\mp1)\in\mathbb{R}^n\times\mathbb{R}$ to the $\mathbb{R}^n\subset\mathbb{R}^n\times\mathbb{R}$ plane where the last coordinate vanishes.

(c) Show that $\mathrm{q}^+\left(\frac{q}{|q|^2}\right)=\mathrm{q}^-(q)$ and $\mathrm{q}^+(q)=\mathrm{q}^-\left(\frac{q}{|q|^2}\right)$.

4.3. The First Fundamental Form

Let $\mathrm{q}(u,v):U\to\mathbb{R}^3$ be a parametrized surface. At each point of this surface we have a basis

$$\begin{aligned}&\frac{\partial\,\mathrm{q}}{\partial u}(u,v),\\ &\frac{\partial\,\mathrm{q}}{\partial v}(u,v),\\ \mathrm{n}(u,v)\quad=\quad&\frac{\frac{\partial\,\mathrm{q}}{\partial u}\times\frac{\partial\,\mathrm{q}}{\partial v}}{\left|\frac{\partial\,\mathrm{q}}{\partial u}\times\frac{\partial\,\mathrm{q}}{\partial v}\right|}.\end{aligned}$$

These vectors are again parametrized by u,v. The first two vectors are tangent to the surface and give us an unnormalized version of the tangent vector for a curve, while the third is the normal and is naturally normalized just as the normal vector is for a curve. One of the issues that make surface theory more difficult than curve theory is that there is no canonical parametrization such as the arclength parametrization for curves.

The *first fundamental form* is the symmetric positive definite form that comes from the matrix

$$\begin{aligned}
[\mathrm{I}] &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix}^t \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial \mathrm{q}}{\partial v} \\ \frac{\partial \mathrm{q}}{\partial v} \cdot \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \cdot \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \\
&= \begin{bmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{bmatrix}.
\end{aligned}$$

For a curve the analogous term would simply be the square of the speed

$$\left(\frac{d\,\mathrm{q}}{dt}\right)^t \frac{d\,\mathrm{q}}{dt} = \frac{d\,\mathrm{q}}{dt} \cdot \frac{d\,\mathrm{q}}{dt}.$$

The first fundamental form dictates how one computes dot products of vectors tangent to the surface assuming they are expanded according to the basis $\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}$. If

$$\begin{aligned}
X &= X^u \frac{\partial \mathrm{q}}{\partial u} + X^v \frac{\partial \mathrm{q}}{\partial v} = \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} X^u \\ X^v \end{bmatrix}, \\
Y &= Y^u \frac{\partial \mathrm{q}}{\partial u} + Y^v \frac{\partial \mathrm{q}}{\partial v} = \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} Y^u \\ Y^v \end{bmatrix},
\end{aligned}$$

then

$$\begin{aligned}
\mathrm{I}(X,Y) &= \begin{bmatrix} X^u & X^v \end{bmatrix} \begin{bmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{bmatrix} \begin{bmatrix} Y^u \\ Y^v \end{bmatrix} \\
&= \begin{bmatrix} X^u & X^v \end{bmatrix} \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix}^t \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} Y^u \\ Y^v \end{bmatrix} \\
&= \left(\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} X^u \\ X^v \end{bmatrix} \right)^t \left(\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} Y^u \\ Y^v \end{bmatrix} \right) \\
&= X^t Y \\
&= X \cdot Y.
\end{aligned}$$

In particular, we see that while the *metric coefficients* $g_{w_1 w_2}$ depend on our parametrization, the dot product $\mathrm{I}(X,Y)$ of two tangent vectors remains the same if we change parameters. Note that I stands for the bilinear form $\mathrm{I}(X,Y)$ which does not depend on parametrizations, while [I] is the matrix representation with respect to a parametrization.

Our first observation is that the normalization factor $\left|\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right|$ can be computed from $[\mathrm{I}]$.

Definition 4.3.1. The *area form* of a parametrized surface is given by

$$\sqrt{\det[\mathrm{I}]}.$$

The next lemma shows that this is given by the area of the parallelogram spanned by $\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}$.

Lemma 4.3.2. *We have*

$$\left|\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right|^2 = \det[\mathrm{I}] = g_{uu} g_{vv} - (g_{uv})^2.$$

PROOF. This is simply the observation that both sides of the equation are formulas for the square of the area of the parallelogram spanned by $\frac{\partial \mathrm{q}}{\partial u}$, $\frac{\partial \mathrm{q}}{\partial v}$, i.e.,

$$\left|\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right|^2 = \left|\frac{\partial \mathrm{q}}{\partial u}\right|^2 \left|\frac{\partial \mathrm{q}}{\partial v}\right|^2 - \left(\frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial \mathrm{q}}{\partial v}\right)^2.$$

$\square$

EXAMPLE 4.3.3 (Example 4.1.10 continued). The first fundamental form and area form of the generalized helicoid are given by

$$[\mathrm{I}] = \begin{bmatrix} 1+(f')^2 & cf' \\ cf' & u^2+c^2 \end{bmatrix}, \ \det[\mathrm{I}] = c^2 + u^2\left(1+(f')^2\right).$$

EXAMPLE 4.3.4. We also need to know how the first fundamental form changes under a reparametrization. Consider for example $(u,v) = e^r(\cos\theta, \sin\theta)$ so that

$$\begin{aligned}
\frac{\partial \mathrm{q}}{\partial r} &= \frac{\partial \mathrm{q}}{\partial u}\frac{\partial u}{\partial r} + \frac{\partial \mathrm{q}}{\partial v}\frac{\partial v}{\partial r} = \frac{\partial \mathrm{q}}{\partial u}e^r\cos\theta + \frac{\partial \mathrm{q}}{\partial v}e^r\sin\theta,\\
\frac{\partial \mathrm{q}}{\partial \theta} &= \frac{\partial \mathrm{q}}{\partial u}\frac{\partial u}{\partial \theta} + \frac{\partial \mathrm{q}}{\partial v}\frac{\partial v}{\partial \theta} = -\frac{\partial \mathrm{q}}{\partial u}e^r\sin\theta + \frac{\partial \mathrm{q}}{\partial v}e^r\cos\theta.
\end{aligned}$$

Thus

$$\begin{aligned}
g_{rr} &= \left(\frac{\partial \mathrm{q}}{\partial u}\frac{\partial u}{\partial r} + \frac{\partial \mathrm{q}}{\partial v}\frac{\partial v}{\partial r}\right)\cdot\left(\frac{\partial \mathrm{q}}{\partial u}\frac{\partial u}{\partial r} + \frac{\partial \mathrm{q}}{\partial v}\frac{\partial v}{\partial r}\right)\\
&= g_{uu}\frac{\partial u}{\partial r}\frac{\partial u}{\partial r} + 2g_{uv}\frac{\partial u}{\partial r}\frac{\partial v}{\partial r} + g_{vv}\frac{\partial v}{\partial r}\frac{\partial v}{\partial r}\\
&= e^{2r}\left(g_{uu}\cos^2\theta + 2g_{uv}\cos\theta\sin\theta + g_{vv}\sin^2\theta\right),
\end{aligned}$$

$$\begin{aligned}
g_{r\theta} &= \left(\frac{\partial \mathrm{q}}{\partial u}\frac{\partial u}{\partial r} + \frac{\partial \mathrm{q}}{\partial v}\frac{\partial v}{\partial r}\right)\cdot\left(\frac{\partial \mathrm{q}}{\partial u}\frac{\partial u}{\partial \theta} + \frac{\partial \mathrm{q}}{\partial v}\frac{\partial v}{\partial \theta}\right)\\
&= g_{uu}\frac{\partial u}{\partial r}\frac{\partial u}{\partial \theta} + g_{uv}\left(\frac{\partial u}{\partial r}\frac{\partial v}{\partial \theta} + \frac{\partial u}{\partial \theta}\frac{\partial v}{\partial r}\right) + g_{vv}\frac{\partial v}{\partial r}\frac{\partial v}{\partial \theta}\\
&= e^{2r}\left(-g_{uu}\cos\theta\sin\theta + g_{uv}\left(\cos^2\theta - \sin^2\theta\right) + g_{vv}\cos\theta\sin\theta\right),
\end{aligned}$$

$$\begin{aligned}
g_{\theta\theta} &= \left(\frac{\partial \mathrm{q}}{\partial u}\frac{\partial u}{\partial \theta} + \frac{\partial \mathrm{q}}{\partial v}\frac{\partial v}{\partial \theta}\right)\cdot\left(\frac{\partial \mathrm{q}}{\partial u}\frac{\partial u}{\partial \theta} + \frac{\partial \mathrm{q}}{\partial v}\frac{\partial v}{\partial \theta}\right)\\
&= g_{uu}\frac{\partial u}{\partial \theta}\frac{\partial u}{\partial \theta} + 2g_{uv}\frac{\partial u}{\partial \theta}\frac{\partial v}{\partial \theta} + g_{vv}\frac{\partial v}{\partial \theta}\frac{\partial v}{\partial \theta}\\
&= e^{2r}\left(g_{uu}\sin^2\theta - 2g_{uv}\sin\theta\cos\theta + g_{vv}\cos^2\theta\right).
\end{aligned}$$

In matrix notation

$$\begin{aligned}
\begin{bmatrix} g_{rr} & g_{r\theta} \\ g_{\theta r} & g_{\theta\theta} \end{bmatrix} &= \begin{bmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial \theta} \end{bmatrix}^t \begin{bmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial \theta} \end{bmatrix}\\
&= \begin{bmatrix} \frac{\partial u}{\partial r} & \frac{\partial v}{\partial r} \\ \frac{\partial u}{\partial \theta} & \frac{\partial v}{\partial \theta} \end{bmatrix} \begin{bmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial \theta} \end{bmatrix}.
\end{aligned}$$

The inverse matrix

$$[\mathrm{I}]^{-1} = \begin{bmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{bmatrix}^{-1} = \begin{bmatrix} g^{uu} & g^{uv} \\ g^{vu} & g^{vv} \end{bmatrix}$$

can be used to find the expansion of a tangent vector by computing its dot products with the basis:

PROPOSITION 4.3.5. *If $X \in T_qM$, then*

$$\begin{aligned} X &= \left(g^{uu}\left(X\cdot\frac{\partial\,\mathrm{q}}{\partial u}\right)+g^{uv}\left(X\cdot\frac{\partial\,\mathrm{q}}{\partial v}\right)\right)\frac{\partial\,\mathrm{q}}{\partial u} \\ &\quad+\left(g^{vu}\left(X\cdot\frac{\partial\,\mathrm{q}}{\partial u}\right)+g^{vv}\left(X\cdot\frac{\partial\,\mathrm{q}}{\partial v}\right)\right)\frac{\partial\,\mathrm{q}}{\partial v} \\ &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}[\mathrm{I}]^{-1}\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}^t X. \end{aligned}$$

More generally, for any $Z \in \mathbb{R}^3$

$$\begin{aligned} Z &= \left(g^{uu}\left(Z\cdot\frac{\partial\,\mathrm{q}}{\partial u}\right)+g^{uv}\left(Z\cdot\frac{\partial\,\mathrm{q}}{\partial v}\right)\right)\frac{\partial\,\mathrm{q}}{\partial u} \\ &\quad+\left(g^{vu}\left(Z\cdot\frac{\partial\,\mathrm{q}}{\partial u}\right)+g^{vv}\left(Z\cdot\frac{\partial\,\mathrm{q}}{\partial v}\right)\right)\frac{\partial\,\mathrm{q}}{\partial v}+(Z\cdot\mathrm{n})\,\mathrm{n} \\ &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}[\mathrm{I}]^{-1}\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}^t Z+(Z\cdot\mathrm{n})\,\mathrm{n}\,. \end{aligned}$$

PROOF. This formula works for $X \in T_qM$ by writing

$$X = \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}\begin{bmatrix} X^u \\ X^v \end{bmatrix} = X^u\frac{\partial\,\mathrm{q}}{\partial u}+X^v\frac{\partial\,\mathrm{q}}{\partial v}$$

and then observing that

$$\begin{aligned} \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}[\mathrm{I}]^{-1}\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}^t X &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}[\mathrm{I}]^{-1}\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}^t\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}\begin{bmatrix} X^u \\ X^v \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}[\mathrm{I}]^{-1}[\mathrm{I}]\begin{bmatrix} X^u \\ X^v \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}\begin{bmatrix} X^u \\ X^v \end{bmatrix} \\ &= X. \end{aligned}$$

For a general vector $Z \in \mathbb{R}^3$ the result follows by using the orthogonal decomposition

$$Z = X + (Z\cdot\mathrm{n})\,\mathrm{n},$$

where $X = Z - (Z\cdot\mathrm{n})\,\mathrm{n} \in T_qM$ and observing that the operation

$$Z \mapsto \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}[\mathrm{I}]^{-1}\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v}\end{bmatrix}^t Z$$

is a linear map defined for all $Z \in \mathbb{R}^3$ with kernel spanned by n. In fact, it orthogonally projects Z to T_qM. □

Defining the *gradient* of a function is another important use of the first fundamental form as well as its inverse. Let $f(u,v)$ be viewed as a function on the surface $\mathrm{q}(u,v)$. Our definition of the gradient should definitely be so that it conforms with the chain rule for a curve $\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$. Thus on one hand we want

$$\begin{aligned} \frac{d(f\circ\mathrm{q})}{dt} &= \nabla f\cdot\dot{\mathrm{q}} \\ &= \begin{bmatrix} (\nabla f)^u & (\nabla f)^v \end{bmatrix}[\mathrm{I}]\begin{bmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{bmatrix} \end{aligned}$$

while the chain rule also dictates

$$\frac{d(f\circ\mathrm{q})}{dt} = \begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v}\end{bmatrix}\begin{bmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{bmatrix}.$$

This indicates that

$$\left[\begin{array}{cc} (\nabla f)^u & (\nabla f)^v \end{array}\right] = \left[\begin{array}{cc} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{array}\right] [\mathrm{I}]^{-1}$$

or

$$\begin{aligned}
\nabla f &= \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{array}\right] \left[\begin{array}{c} (\nabla f)^u \\ (\nabla f)^v \end{array}\right] \\
&= \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{array}\right] \left(\left[\begin{array}{cc} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{array}\right] [\mathrm{I}]^{-1}\right)^t \\
&= \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{array}\right] [\mathrm{I}]^{-1} \left[\begin{array}{cc} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{array}\right]^t \\
&= \left(g^{uu}\frac{\partial f}{\partial u} + g^{uv}\frac{\partial f}{\partial v}\right)\frac{\partial \mathrm{q}}{\partial u} + \left(g^{vu}\frac{\partial f}{\partial u} + g^{vv}\frac{\partial f}{\partial v}\right)\frac{\partial \mathrm{q}}{\partial v}.
\end{aligned}$$

In particular, we see that changing coordinates changes the gradient in such a way that it isn't simply the vector corresponding to the partial derivatives! The other nice feature is that we now have a concept of the gradient that gives a vector field independently of parametrizations. The defining equation

$$\frac{d(f \circ \mathrm{q})}{dt} = \nabla f \cdot \dot{\mathrm{q}} = \mathrm{I}(\nabla f, \dot{\mathrm{q}})$$

gives an implicit definition of ∇f that makes sense without reference to parametrizations of the surface.

Exercises

(1) For a surface of revolution $\mathrm{q}(t,\mu) = (r(t)\cos\mu, r(t)\sin\mu, z(t))$ (see section 4.1 exercise 4) show that the first fundamental form is given by

$$\left[\begin{array}{cc} g_{tt} & g_{t\mu} \\ g_{\mu t} & g_{\mu\mu} \end{array}\right] = \left[\begin{array}{cc} \dot{r}^2 + \dot{z}^2 & 0 \\ 0 & r^2 \end{array}\right].$$

A special and important case of this occurs when $z = 0$ and $r = t$ as that corresponds to polar coordinates in the (x,y)-plane.

(2) Assume that we have a cone (see section 4.1 exercise 2) given by

$$\mathrm{q}(r,\phi) = rc(\phi),$$

where c is a space curve with $|c| = 1$ and $\left|\frac{dc}{d\phi}\right| = 1$. Show that the first fundamental form is given by

$$\left[\begin{array}{cc} g_{rr} & g_{r\phi} \\ g_{\phi r} & g_{\phi\phi} \end{array}\right] = \left[\begin{array}{cc} 1 & 0 \\ 0 & r^2 \end{array}\right]$$

and compare this to polar coordinates in the plane.

(3) Assume that we have a generalized cylinder (see section 4.1 exercise 1) given by

$$\mathrm{q}(s,t) = (x(s), y(s), t),$$

where $(x(s), y(s))$ is unit speed. Show that the first fundamental form is given by

$$\left[\begin{array}{cc} g_{ss} & g_{st} \\ g_{ts} & g_{tt} \end{array}\right] = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right].$$

(4) Assume that we have a ruled surface (see section 4.1 exercise 3) given by

$$\mathrm{q}\left(s,t\right)=c\left(t\right)+sX\left(t\right),$$

where c is a space curve and X is a unit vector for each t. Show that the first fundamental form is given by

$$\begin{bmatrix} g_{ss} & g_{st} \\ g_{ts} & g_{tt} \end{bmatrix}=\begin{bmatrix} 1 & \frac{dc}{dt}\cdot X \\ \frac{dc}{dt}\cdot X & \left|\frac{dc}{dt}+s\frac{dX}{dt}\right|^2 \end{bmatrix}.$$

(5) Show that if we have a parametrized surface $\mathrm{q}\left(r,\theta\right)$ such that the first fundamental form is given by

$$\begin{bmatrix} g_{rr} & g_{r\theta} \\ g_{\theta r} & g_{\theta\theta} \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix},$$

then we can locally reparametrize the surface to $\mathrm{q}\left(u,v\right)$ where the new first fundamental form is

$$\begin{bmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Hint: Let $u=r\cos\theta$ and $v=r\sin\theta$.

(6) Let $c\left(s\right)$ be a unit speed curve with non-zero curvature, binormal B_c and torsion τ. Show that the first fundamental form for the ruled surface

$$\mathrm{q}\left(s,t\right)=c\left(s\right)+t\,\mathrm{B}_c\left(s\right)$$

is given by

$$\begin{bmatrix} g_{ss} & g_{st} \\ g_{ts} & g_{tt} \end{bmatrix}=\begin{bmatrix} 1+t^2\tau^2 & 0 \\ 0 & 1 \end{bmatrix}.$$

(7) Consider a unit speed curve $c\left(s\right)$ with non-vanishing curvature and the tube (see section 4.1 exercise 6) of radius R around it

$$\mathrm{q}\left(s,\phi\right)=c\left(s\right)+R\left(\mathrm{N}_c\cos\phi+\mathrm{B}_c\sin\phi\right),$$

where $\mathrm{T}_c,\mathrm{N}_c,\mathrm{B}_c$ are the unit tangent, normal, and binormal to the curve.

(a) Show that T_c and $-\mathrm{N}_c\sin\phi+\mathrm{B}_c\cos\phi$ are an orthonormal basis for the tangent space and that the normal to the tube is $\mathrm{n}=-\left(\mathrm{N}_c\cos\phi+\mathrm{B}_c\sin\phi\right)$.

(b) Show that

$$\begin{bmatrix} g_{ss} & g_{s\phi} \\ g_{\phi s} & g_{\phi\phi} \end{bmatrix}=\begin{bmatrix} \left(1-\kappa R\cos\phi\right)^2+\left(\tau R\right)^2 & \tau R^2 \\ \tau R^2 & R^2 \end{bmatrix}.$$

(8) Compute the first fundamental form of the Möbius band

$$\mathrm{q}\left(t,\phi\right)=\left(\cos\phi,\sin\phi,0\right)+t\left(\sin\frac{\phi}{2}\cos\phi,\sin\frac{\phi}{2}\sin\phi,\cos\frac{\phi}{2}\right).$$

(9) For a parametrized surface $\mathrm{q}\left(u,v\right)$ show that

$$\mathrm{n}\times\frac{\partial\,\mathrm{q}}{\partial u}=\frac{g_{uu}\frac{\partial\,\mathrm{q}}{\partial v}-g_{uv}\frac{\partial\,\mathrm{q}}{\partial u}}{\left|\frac{\partial\,\mathrm{q}}{\partial u}\times\frac{\partial\,\mathrm{q}}{\partial v}\right|},$$

$$\mathrm{n}\times\frac{\partial\,\mathrm{q}}{\partial v}=\frac{g_{uv}\frac{\partial\,\mathrm{q}}{\partial v}-g_{vv}\frac{\partial\,\mathrm{q}}{\partial u}}{\left|\frac{\partial\,\mathrm{q}}{\partial u}\times\frac{\partial\,\mathrm{q}}{\partial v}\right|}.$$

(10) Assume a surface has a parametrization $\mathrm{q}(s,\mu)$ where

$$\begin{bmatrix} g_{ss} & g_{s\mu} \\ g_{\mu s} & g_{\mu\mu} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix},$$

where $r(s)$ is only a function of s.

(a) Show that if $0 < \frac{dr}{ds} < 1$, then there is a function $z(s)$ so that $(r(s), 0, z(s))$ is a unit speed curve.

(b) Conclude that there is a surface of revolution with the same first fundamental form.

(11) Assume a surface has a parametrization $\mathrm{q}(u,v)$ where

$$\begin{bmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{bmatrix} = \begin{bmatrix} r^2 & 0 \\ 0 & r^2 \end{bmatrix},$$

where $r(u) > 0$ is only a function of u. Show that there is a reparametrization $u = u(s)$ such that the first fundamental form becomes

$$\begin{bmatrix} g_{ss} & g_{sv} \\ g_{vs} & g_{vv} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix}.$$

(12) Show that if we have a parametrization where

$$[\mathrm{I}] = \begin{bmatrix} 1 & 0 \\ 0 & g_{vv} \end{bmatrix},$$

then the coordinate function $f(u,v) = u$ has

$$\nabla u = \frac{\partial \mathrm{q}}{\partial u}.$$

(13) Show that it is always possible to find an orthogonal parametrization, i.e., g_{uv} vanishes. Hint: Use theorem 4.2.10.

(14) Show that if

$$\frac{\partial g_{uu}}{\partial v} = \frac{\partial g_{vv}}{\partial u} = g_{uv} = 0,$$

then we can reparametrize u and v separately, i.e., $u = u(s)$ and $v = v(t)$, in such a way that we obtain *Cartesian coordinates*:

$$\begin{aligned} g_{ss} &= g_{tt} = 1, \\ g_{st} &= 0. \end{aligned}$$

(15) Show that if

$$\frac{\partial^2 \mathrm{q}}{\partial u \partial v} = 0,$$

then

$$\mathrm{q}(u,v) = F(u) + G(v),$$

and conclude that

$$\frac{\partial g_{uu}}{\partial v} = \frac{\partial g_{vv}}{\partial u} = 0.$$

Give an example where $g_{uv} \neq 0$.

4.4. Special Maps and Parametrizations

DEFINITION 4.4.1. We call a map $F : M_1 \to M_2$ between surfaces an *isometry* if its differential preserves the first fundamental form

$$\mathrm{I}_1(X,Y) = \mathrm{I}_2(DF(X), DF(Y)).$$

We call the map *area preserving* if it preserves the areas of parallelograms spanned by vectors:

$$\det \begin{bmatrix} \mathrm{I}_1(X,X) & \mathrm{I}_1(X,Y) \\ \mathrm{I}_1(X,Y) & \mathrm{I}_1(Y,Y) \end{bmatrix} = \det \begin{bmatrix} \mathrm{I}_1(DF(X), DF(X)) & \mathrm{I}_1(DF(X), DF(Y)) \\ \mathrm{I}_1(DF(X), DF(Y)) & \mathrm{I}_1(DF(Y), DF(Y)) \end{bmatrix}.$$

We call the map *conformal* if it preserves angles between vectors:

$$\cos\angle(X,Y) = \frac{\mathrm{I}_1(X,Y)}{|X|_1 |Y|_1} = \frac{\mathrm{I}_2(DF(X), DF(Y))}{|DF(X)|_2 |DF(Y)|_2}.$$

When the first surface is given as a parametrized surface these conditions can be checked as follows.

PROPOSITION 4.4.2. *Let* $\mathrm{q} : U \to M_1$ *be a parametrization and* $F : M_1 \to M_2$ *a map. The map is an isometry if*

$$[\mathrm{I}_1] = [\mathrm{I}_2],$$

area preserving if

$$\det[\mathrm{I}_1] = \det[\mathrm{I}_2],$$

and conformal if

$$[\mathrm{I}_1] = \lambda^2 [\mathrm{I}_2]$$

for some non-zero function λ.

PROOF. Note that it is not necessary to first check that $F \circ \mathrm{q}$ is also a parametrization as that will be a consequence of any one of the three conditions if we define

$$[\mathrm{I}_2] = \begin{bmatrix} \frac{\partial F\circ \mathrm{q}}{\partial u} \cdot \frac{\partial F\circ \mathrm{q}}{\partial u} & \frac{\partial F\circ \mathrm{q}}{\partial u} \cdot \frac{\partial F\circ \mathrm{q}}{\partial v} \\ \frac{\partial F\circ \mathrm{q}}{\partial v} \cdot \frac{\partial F\circ \mathrm{q}}{\partial u} & \frac{\partial F\circ \mathrm{q}}{\partial v} \cdot \frac{\partial F\circ \mathrm{q}}{\partial v} \end{bmatrix}$$

and observe that $\frac{\partial F\circ \mathrm{q}}{\partial v}$, $\frac{\partial F\circ \mathrm{q}}{\partial u}$ are linearly independent if and only if the matrix $[\mathrm{I}_2]$ has nonzero determinant.

Next note that the chain rule implies that

$$DF\left(\frac{\partial \mathrm{q}}{\partial u}\right) = \frac{\partial F\circ \mathrm{q}}{\partial u},\ DF\left(\frac{\partial \mathrm{q}}{\partial v}\right) = \frac{\partial F\circ \mathrm{q}}{\partial v}$$

So the three conditions are necessarily true if the map is an isometry, area preserving, or conformal respectively. More generally, we see that

$$DF(X) = DF\left(X^u \frac{\partial \mathrm{q}}{\partial u} + X^v \frac{\partial \mathrm{q}}{\partial v}\right) = X^u \frac{\partial F\circ \mathrm{q}}{\partial u} + X^v \frac{\partial F\circ \mathrm{q}}{\partial v}.$$

So if $[\mathrm{I}_1] = [\mathrm{I}_2]$, then

$$\begin{aligned}
\mathrm{I}_2\left(DF\left(X\right), DF\left(Y\right)\right) &= \begin{bmatrix} X^u & X^v \end{bmatrix} \begin{bmatrix} \frac{\partial F\circ \mathrm{q}}{\partial u} & \frac{\partial F\circ \mathrm{q}}{\partial v} \end{bmatrix}^t \begin{bmatrix} \frac{\partial F\circ \mathrm{q}}{\partial u} & \frac{\partial F\circ \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} Y^u \\ Y^v \end{bmatrix} \\
&= \begin{bmatrix} X^u & X^v \end{bmatrix} [\mathrm{I}_2] \begin{bmatrix} Y^u \\ Y^v \end{bmatrix} \\
&= \begin{bmatrix} X^u & X^v \end{bmatrix} [\mathrm{I}_1] \begin{bmatrix} Y^u \\ Y^v \end{bmatrix} \\
&= \mathrm{I}_1\left(X, Y\right).
\end{aligned}$$

A similar calculation with the assumption that $[\mathrm{I}_1] = \lambda^2 [\mathrm{I}_2]$ gives us

$$\begin{aligned}
\mathrm{I}_1\left(X, Y\right) &= \lambda^2\, \mathrm{I}_2\left(X, Y\right), \\
|X|^2 = \mathrm{I}_1\left(X, X\right) &= \lambda^2\, \mathrm{I}_2\left(X, X\right), \\
|Y|^2 = \mathrm{I}_1\left(Y, Y\right) &= \lambda^2\, \mathrm{I}_2\left(Y, Y\right).
\end{aligned}$$

As angles are given by

$$\cos\angle\left(X, Y\right) = \frac{X \cdot Y}{|X|\,|Y|}$$

this establishes the last claim.

Finally, if

$$\left|\frac{\partial\,\mathrm{q}}{\partial u} \times \frac{\partial\,\mathrm{q}}{\partial v}\right|^2 = \det\left[\mathrm{I}_1\right] = \det\left[\mathrm{I}_2\right] = \left|\frac{\partial F\circ \mathrm{q}}{\partial u} \times \frac{\partial F\circ \mathrm{q}}{\partial v}\right|^2,$$

then the observation that

$$\begin{aligned}
X \times Y &= \left(X^u \frac{\partial\,\mathrm{q}}{\partial u} + X^v \frac{\partial\,\mathrm{q}}{\partial v}\right) \times \left(Y^u \frac{\partial\,\mathrm{q}}{\partial u} + Y^v \frac{\partial\,\mathrm{q}}{\partial v}\right) \\
&= \left(X^u Y^v - X^v Y^u\right) \frac{\partial\,\mathrm{q}}{\partial u} \times \frac{\partial\,\mathrm{q}}{\partial v}, \\
DF\left(X\right) \times DF\left(Y\right) &= \left(X^u \frac{\partial F\circ \mathrm{q}}{\partial u} + X^v \frac{\partial F\circ \mathrm{q}}{\partial v}\right) \times \left(Y^u \frac{\partial F\circ \mathrm{q}}{\partial u} + Y^v \frac{\partial F\circ \mathrm{q}}{\partial v}\right) \\
&= \left(X^u Y^v - X^v Y^u\right) \frac{\partial F\circ \mathrm{q}}{\partial u} \times \frac{\partial F\circ \mathrm{q}}{\partial v},
\end{aligned}$$

shows that

$$|X \times Y|^2 = |DF\left(X\right) \times DF\left(Y\right)|^2.$$

The last statement can also be rephrased without the use of λ by checking that

$$\frac{\frac{\partial F\circ \mathrm{q}}{\partial u} \cdot \frac{\partial F\circ \mathrm{q}}{\partial u}}{\frac{\partial\,\mathrm{q}}{\partial u} \cdot \frac{\partial\,\mathrm{q}}{\partial u}} = \frac{\frac{\partial F\circ \mathrm{q}}{\partial v} \cdot \frac{\partial F\circ \mathrm{q}}{\partial v}}{\frac{\partial\,\mathrm{q}}{\partial v} \cdot \frac{\partial\,\mathrm{q}}{\partial v}}$$

and

$$\frac{\partial F\circ \mathrm{q}}{\partial v} \cdot \frac{\partial F\circ \mathrm{q}}{\partial u} = \frac{\frac{\partial F\circ \mathrm{q}}{\partial u} \cdot \frac{\partial F\circ \mathrm{q}}{\partial u}}{\frac{\partial\,\mathrm{q}}{\partial u} \cdot \frac{\partial\,\mathrm{q}}{\partial u}} \frac{\partial\,\mathrm{q}}{\partial v} \cdot \frac{\partial\,\mathrm{q}}{\partial u}$$

□

Definition 4.4.3. In case the map is a parametrization $\mathrm{q} : U \to M$ then we always use the *Cartesian metric* on U given by

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

So the parametrization is an isometry or *Cartesian* when

$$[\mathrm{I}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

area preserving when

$$\det [\mathrm{I}] = 1,$$

and conformal or *isothermal* when

$$g_{uu} = g_{vv},$$

$$g_{uv} = 0.$$

EXAMPLE 4.4.4. It follows from proposition 4.4.2 and example 4.2.16 that the Archimedes map is area preserving and the Mercator map is conformal.

EXAMPLE 4.4.5. Consider the reparametrization $(u, v) = e^r (\cos\theta, \sin\theta)$ from example 4.3.4. If $\mathrm{q}(u, v)$ is conformal, then $g_{uu} = g_{vv} = \lambda^2$ and $g_{uv} = 0$. Thus the reparametrized surface $\mathrm{q}(r, \theta) = \mathrm{q}(e^r \cos\theta, e^r \sin\theta)$ has

$$\begin{aligned} g_{rr} &= e^{2r}\lambda^2, \\ g_{r\theta} &= 0, \\ g_{\theta\theta} &= e^{2r}\lambda^2. \end{aligned}$$

In particular, this gives a conformal reparametrization of the Cartesian plane. This example is part of a much broader class of conformal maps. If we write $w = u + \sqrt{-1}v$ and $z = x + \sqrt{-1}y$, then the transformation $w = F(z)$ is conformal when F is holomorphic. This follows from the fact that the linear map given by the matrix

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

preserves angles and that $F(z)$ is holomorphic when it satisfies the Cauchy-Riemann equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ and } \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

i.e.,

$$DF = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x} & -\frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial x} & \frac{\partial u}{\partial x} \end{bmatrix}.$$

The above reparametrization is simply complex exponentiation: $w = e^z$.

DEFINITION 4.4.6. The *area* of a parametrized surface $\mathrm{q}(u, v) : U \to M$ over a region $R \subset U$ where q is one-to-one is defined by the integral

$$\mathrm{Area}(\mathrm{q}(R)) = \int_R \sqrt{\det [\mathrm{I}]}\, du dv.$$

PROPOSITION 4.4.7. *The area is independent under reparametrization.*

PROOF. Assume we have a different parametrization $\mathrm{q}(s, t) : V \to M$ and a new region $T \subset V$ with $\mathrm{q}(R) = \mathrm{q}(T)$ and the property that the reparametrization

$(u(s,t), v(s,t)) : T \to R$ is a diffeomorphism. Then

$$\begin{aligned}
\operatorname{Area}(\mathrm{q}(R)) &= \int_R \sqrt{\det[\mathrm{I}]}\,dudv \\
&= \int_R \sqrt{\det\left(\left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{array}\right]\right)}\,dudv \\
&= \int_R \sqrt{\det\left(\left(\left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]\left[\begin{array}{cc} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{array}\right]\right)^t \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]\left[\begin{array}{cc} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{array}\right]\right)}\,dudv \\
&= \int_R \sqrt{\det\left(\left[\begin{array}{cc} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]\left[\begin{array}{cc} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{array}\right]\right)}\,dudv \\
&= \int_R \sqrt{\det\left[\begin{array}{cc} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{array}\right]^t}\det\sqrt{\left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]}\sqrt{\det\left[\begin{array}{cc} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{array}\right]}\,dudv \\
&= \int_R \det\sqrt{\left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]}\left|\det\left[\begin{array}{cc} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{array}\right]\right|\,dudv \\
&= \int_R \det\sqrt{\left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial s} & \frac{\partial \mathrm{q}}{\partial t} \end{array}\right]}\,dsdt,
\end{aligned}$$

where the last equality follows from the change of variables formula for integrals. □

Finally we show that general maps that are not conformal are related to certain nice parametrizations. This depends on a more general result that we will use in several situations.

Definition 4.4.8. A symmetric bilinear form Q on a surface, is a symmetric bilinear form $Q(X,Y)$ on each of the tangent spaces that varies smoothly, i.e., $Q(X,Y)$ is linear in each of the two variables separately, $Q(X,Y) = Q(Y,X)$, and when X and Y are smooth vector fields then $Q(X,Y)$ is also smooth.

The first fundamental form is an example of a symmetric bilinear form on a surface.

Theorem 4.4.9. *Let Q be a symmetric bilinear form on a surface. If Q is not a multiple of* I *at p, then there is a parametrization* $\mathrm{q}(u,v)$ *around p such that*

$$\mathrm{I}\left(\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}\right) = Q\left(\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}\right) = 0.$$

Proof. By theorem 4.2.10 it suffices to find orthogonal unit vector fields E_1 and E_2 near p such that $Q(E_1, E_2) = 0$.

At a point q consider $Q(E,E)$ for all unit vectors $E \in T_qM$. This function will have a maximum at some vector $E_1 \in T_qM$. Let $E_2 \in T_qM$ be a unit vector orthogonal to E_1. It follows that $E(\theta) = \cos\theta E_1 + \sin\theta E_2 \in T_qM$ is also a unit vector. Now consider

$$Q(E(\theta), E(\theta)) = \cos^2\theta Q(E_1,E_1) + 2\cos\theta\sin\theta Q(E_1,E_2) + \sin^2\theta Q(E_2,E_2).$$

By construction this is a function of θ that has a maximum at $\theta = 0$. The derivative at $\theta = 0$ is $2Q(E_1,E_2)$. Therefore, $Q(E_1,E_2) = 0$ and

$$Q(E(\theta), E(\theta)) = \cos^2\theta Q(E_1,E_1) + \sin^2\theta Q(E_2,E_2).$$

If $Q = \lambda \mathrm{I}$, then $Q(E(\theta), E(\theta)) = \lambda$ for all θ. Otherwise,

$$Q(E_1, E_1) > Q(E_2, E_2)$$

and $Q(E(\theta), E(\theta))$ will only have a maximum when $\theta = 0, \pi$ and a minimum when $\theta = \pm\frac{\pi}{2}$.

Since Q is not a multiple of I at p it follows by continuity that it won't be a multiple of I for q near p. This means that E_1 is well-defined up to a choice of sign. If we fix a choice at p, then we can uniquely extend this to a unit vector field E_1 in a neighborhood of p. Similarly for E_2. This finishes the construction of the orthonormal frame E_1, E_2. □

Corollary 4.4.10. *Let $F : M_1 \to M_2$ be a map between surfaces. If F is not conformal near $p \in M_1$, then there is a parametrization $\mathrm{q}(u, v)$ of a neighborhood of p such that*

$$0 = \mathrm{I}_1\left(\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}\right) = \mathrm{I}_2\left(\frac{\partial F \circ \mathrm{q}}{\partial u}, \frac{\partial F \circ \mathrm{q}}{\partial v}\right) = 0.$$

Proof. We can simply use $Q(X, Y) = \mathrm{I}_2(DF(X), DF(Y))$ as our symmetric bilinear form on M_1. The fact that F is not conformal at p shows that Q is not a multiple of I_1. □

Exercises

(1) Check if the parameterization $\mathrm{q}(t, \phi) = t(\cos\phi, \sin\phi, 1)$ for the cone is an isometry, area preserving, or conformal? Can the surface be reparametrized to have any of these properties? Hint: See section 4.3 exercise 2.

(2) Show that the two surfaces defined by $z = x^2 - y^2$ and $z = 2xy$ are isometric.

(3) Compute the areas of the following surfaces by integrating the area form for a suitable parametrization.
   (a) Show that the sphere of radius R has area $4\pi R^2$.
   (b) Show that the circular cylinder of radius R and height h has area $2\pi Rh$.
   (c) Show that the torus from section 4.1 exercise 8 has area $4\pi^2 Rr$.

(4) Consider a ruled surface

$$\mathrm{q}(s, t) = c(s) + tX(s)$$

where c is unit speed and X is a unit field. Show that it is conformal if and only if it is Cartesian (in which case X is constant and normal to c for all s.) Hint: See section 4.2 exercise 4.

(5) Show that there is a map from a surface of revolution $\mathrm{q}_1(r, \mu) = (r\cos\mu, r\sin\mu, z_1(r))$ to a circular cylinder $\mathrm{q}_2(r, \mu) = (\cos\mu, \sin\mu, z_2(r))$ that is either
   (a) conformal or
   (b) area preserving.

(6) Show that the curve $(r(u), z(u))$ can be reparametrized so that the new parametrization

$$\mathrm{q}(t, \mu) = (r(t)\cos\mu, r(t)\sin\mu, z(t))$$

is either
   (a) conformal or
   (b) area preserving.

(7) Show that a Monge patch $z = f(x, y)$ becomes:

(a) area preserving if and only if f is constant;
(b) conformal if and only if f is constant.

(8) Show that the equation
$$ax + by + cz = d$$
defines a surface if and only if $(a, b, c) \neq (0, 0, 0)$. Show that this surface has a parametrization that is Cartesian.

(9) The *conoid* is a special type of ruled surface where c is a straight line and X always lies in a fixed plane. The simplest case is when c is the z-axis and X lies in the (x, y)-plane
$$\begin{aligned} \mathrm{q}(s,t) &= (tx(s), ty(s), z(s)) \\ &= (0,0,z(s)) + t(x(s), y(s), 0) \end{aligned}$$
(a) Compute its first fundamental form when $|X| = 1$.
(b) Show that this parametrization is conformal (or area preserving) if and only if the surface is a plane.
(c) Show that this surface is a helicoid when both X and z have constant speed.
(d) Show that such a helicoid can be reparametrized using $t = t(v)$ so as to obtain either a conformal or an area preserving parametrization.

(10) Consider the two parametrized surfaces given by
$$\begin{aligned} \mathrm{q}_1(\phi, t) &= (\sinh\phi\cos t, \sinh\phi\sin t, t) \\ &= (0,0,t) + \sinh\phi(\cos t, \sin t, 0) \\ \mathrm{q}_2(s, \mu) &= (\cosh s\cos\mu, \cosh s\sin\mu, s) \end{aligned}$$
Compute the first fundamental forms for both surfaces and construct a local isometry from the first surface to the second. (The first surface is a ruled surface with a one-to-one parametrization called the helicoid, the second surface is a surface of revolution called the catenoid.) Hint: See section 4.2 exercise 5.

(11) Consider the tube from section 4.3 exercise 7 with $s \in [0, L]$ and $\phi \in [0, 2\pi]$.
(a) Show that the area is given by $2\pi RL$.
(b) Find an area preserving map from this tube to a cylinder of the form
$$F(\mathrm{q}(s,\phi)) = (R\cos\phi, R\sin\phi, h(s,\phi)).$$

(12) Consider a generalized cylinder parametrized as in section 4.3 exercise 3 with $s \in [0, L]$ and $t \in [a, b]$. Show that its area is $L(b - a)$.

(13) Consider a generalized cone parametrized as in section 4.3 exercise 2 with $\phi \in [0, L]$ and $r \in [a, b]$. Show that its area is $\frac{1}{2}L\left(b^2 - a^2\right)$.

(14) Show that Enneper's surface
$$\mathrm{q}(u,v) = \begin{bmatrix} u - \frac{1}{3}u^3 + uv^2 \\ v - \frac{1}{3}v^3 + vu^2 \\ u^2 - v^2 \end{bmatrix}$$
defines a conformal parametrization.

(15) Show that Catalan's surface
$$\mathrm{q}(u,v) = \begin{bmatrix} u - \sin u\cosh v \\ 1 - \cos u\cosh v \\ 4\sin\frac{u}{2}\sinh\frac{v}{2} \end{bmatrix}$$

defines a conformal parametrization. Hint: Start by showing that: $2\sin^2\frac{u}{2} = 1 - \cos u$ etc.

(16) Show that the following two parametrizations of the unit sphere are area preserving:

(a) (Lambert, 1772)

$$\mathrm{q}(\mu, z) = \begin{bmatrix} \sqrt{1-z^2}\cos\mu \\ \sqrt{1-z^2}\sin\mu \\ z \end{bmatrix}, \ |\mu| < \pi, \ |z| < 1.$$

(b) (Sinusoidal projection, Cossin, 1570)

$$\mathrm{q}(s, t) = \begin{bmatrix} \cos s \cos\left(\frac{t}{\cos s}\right) \\ \cos s \sin\left(\frac{t}{\cos s}\right) \\ \sin s \end{bmatrix}, \ |s| < \frac{\pi}{2}, \ t < \pi \cos s.$$

(c) Relate the Lambert parametrization to the Archimedes map.

(17) (Stabius-Werner, c. 1500, Sylvanus, 1511, Bonne, c. 1780) Show that the Bonne parametrizations

$$\mathrm{q}(r, \theta) = \begin{bmatrix} \cos(r - r_0)\cos\left(\frac{r(\theta - \pi/2)}{\cos(r - r_0)}\right) \\ \cos(r - r_0)\sin\left(\frac{r(\theta - \pi/2)}{\cos(r - r_0)}\right) \\ \sin(r - r_0) \end{bmatrix},$$

have the property that $\det[\mathrm{I}] = r^2$. Conclude that they are area preserving when (r, θ) correspond to polar coordinates

$$x = r\cos\theta, \ y = r\sin\theta.$$

For $r_0 = 0$ this is a sinusoidal projection, for $r_0 = \pi/2$ the Stabius-Werner projection, and for $0 < r_0 < \pi/2$ the Sylvanus projection. The planar shape of these maps is bordered on the outside by an implicitly given curve

$$|r|\left|\theta - \frac{\pi}{2}\right| = \cos(r - r_0)$$

as $r \to \pi/2$ this looks like a heart shaped region.

(18) Show that the inversion map

$$F(q) = \frac{q}{|q|^2}.$$

is a conformal map of $\mathbb{R}^n - 0$ to it self. Hint: See section 4.2 exercise 7.

(19) Show that the inverse stereographic projections $\mathrm{q}^{\pm} : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$ defined by

$$\mathrm{q}^{\pm}(q) = (q, 0) + \frac{1 - |q|^2}{1 + |q|^2}(q, \pm 1)$$

are conformal parametrizations of the unit sphere. Hint: See section 4.2 exercise 9. More specifically when $n = 2$ it is given by

$$\mathrm{q}^{\pm}(u, v) = \left(\frac{2u}{u^2 + v^2 + 1}, \frac{2v}{u^2 + v^2 + 1}, \mp\frac{u^2 + v^2 - 1}{u^2 + v^2 + 1}\right).$$

(20) Consider the map $F : H \to \mathbb{R}^2$ defined by

$$\begin{aligned} F(x,y) &= \frac{1}{x^2+(y+1)^2}\left(2x, 2(y+1)\right) + (0,-1) \\ &= \frac{1}{x^2+(y+1)^2}\left(2x, 1-x^2-y^2\right), \end{aligned}$$

where $H = \{(x,y) \mid y > 0\}$.

(a) Show that F is one-to-one and that the image is $D = \left\{(x,y) \mid x^2+y^2 < 1\right\}$. Hint: Show that

$$|F(x,y)|^2 = 1 - \frac{4y}{x^2+(y+1)^2}.$$

(b) Show that the inverse is given by

$$\begin{aligned} F^{-1}(u,v) &= \frac{1}{u^2+(v+1)^2}\left(2u, 2(v+1)\right) + (0,-1) \\ &= \frac{1}{u^2+(v+1)^2}\left(2u, 1-u^2-v^2\right). \end{aligned}$$

(c) Show that F and F^{-1} are conformal.

(d) Show that F can be interpreted as an inversion in the circle of radius $\sqrt{2}$ centered at $(0,-1)$.

(21) Consider a map $F : S^2 \to P$, where $P = \{z = 1\}$ is the plane tangent to the North Pole, that takes each meridian to the radial line that is tangent to the meridian at the North Pole. Sometimes the map might only be defined on part of the sphere such as the upper hemisphere.

(a) Show that such a map has a parametrization of the form

$$F\left(\begin{bmatrix} \cos\mu\cos\phi \\ \sin\mu\cos\phi \\ \sin\phi \end{bmatrix}\right) = \begin{bmatrix} r(\phi)\cos\mu \\ r(\phi)\sin\mu \\ 1 \end{bmatrix}$$

for some function r, where $r\left(\frac{\pi}{2}\right) = 0$.

(b) Show that when $r = \sqrt{2(1-\sin\phi)}$, then we obtain an area preserving map on the upper hemisphere.

(c) Show that when the map projects a point on the upper hemisphere along the radial line through the origin, then $r = \cot\phi$. Show that this map takes all great circles (not just meridians) to straight lines. This is also called the Beltrami projection and is an example of a perspective projection (see section 4.2 exercise 8).

(d) Show that the inverse of the Beltrami projection from (c) onto the upper hemisphere is given by

$$B^{-1}(s,t,1) = \left(\frac{s}{\sqrt{1+s^2+t^2}}, \frac{t}{\sqrt{1+s^2+t^2}}, \frac{1}{\sqrt{1+s^2+t^2}}\right).$$

(22) Show that a map $F : M \to M^*$ that is both conformal and area preserving is an isometry.

(23) (Girard, 1626) A hemisphere on the unit sphere S^2 is the part that lies on one side of a great circle. A *lune* is the intersection of two hemispheres. It has two antipodal vertices. A spherical triangle is the region bounded by three hemispheres.

(a) Show that the area of a hemisphere is 2π.
(b) Use the Archimedes map to show that the area of a lune where the great circles meet at an angle of α is 2α.
(c) If $A(H)$ denotes the area of a region on S^2 use a Venn type diagram to show that

$$\begin{aligned}A(H_1 \cup H_2 \cup H_3) &= A(H_1) + A(H_2) + A(H_3) \\ &\quad - A(H_1 \cap H_3) - A(H_2 \cap H_3) - A(H_1 \cap H_2) \\ &\quad + A(H_1 \cap H_2 \cap H_3).\end{aligned}$$

(d) Let H_1, H_2, H_3 be hemispheres and $H_i' = S^2 - H_i$ the complementary hemispheres. Show that

$$H_1' \cap H_2' \cap H_3' = S^2 - H_1 \cup H_2 \cup H_3.$$

And further show that the spherical triangle $H_1 \cap H_2 \cap H_3$ is congruent to the spherical triangle $H_1' \cap H_2' \cap H_3'$ via the antipodal map.
(e) Show that the area A of a spherical triangle is given by

$$A = \alpha + \beta + \gamma - \pi,$$

where α, β, γ are the interior angles at the vertices of the triangle.

CHAPTER 5

Curvature of Surfaces

The goal of this chapter is to understand curvature of surfaces. This is quite a bit more complicated than for curves. There are two curvatures and they are defined as *extrinsic* invariants, i.e., they depend on how the normal to the surface changes. This is analogous to the curvature of curves. One of the surprising discoveries by Gauss was that one of these curvatures is an *intrinsic* invariant. This means that it can be calculated knowing only the first fundamental form. Another old problem we investigate is that of understanding which surfaces admit Cartesian coordinates.

5.1. Curves on Surfaces

In this section the second fundamental form is introduced as the normal part of the acceleration of a curve. This is used to find its matrix representation. In section 5.2 a more algebraic definition is offered.

We start with the observation that for a surface $M \subset \mathbb{R}^3$ and a point $p \in M$ the tangent space T_pM and normal space $N_pM = (T_pM)^{\perp}$ are defined independently of parametrizations (see proposition 4.2.6). Thus the projections of a vector $Z \in \mathbb{R}^3$ onto both the tangent space and the normal space are well-defined without reference to parametrizations.

Consider a curve $\mathrm{q}(t)$ on the surface. The velocity $\dot{\mathrm{q}}$ and acceleration $\ddot{\mathrm{q}}$ can be calculated in $\mathbb{R}^3$ without reference to the surface. The velocity will be tangent to the surface, but the acceleration rarely is. The projections of $\ddot{\mathrm{q}}$ onto the normal space, $\ddot{\mathrm{q}}^{\mathrm{II}} = (\ddot{\mathrm{q}} \cdot \mathrm{n})\,\mathrm{n}$, and the tangent space, $\ddot{\mathrm{q}}^{\mathrm{I}} = \ddot{\mathrm{q}} - (\ddot{\mathrm{q}} \cdot \mathrm{n})\,\mathrm{n}$, can be computed without parametrizing the surface. This shows that tangential and normal accelerations are well-defined.

Theorem 5.1.1 ((Euler, 1760 and Meusnier, 1776)). *The normal component of the acceleration only depends on the surface and the velocity of the curve. In particular, two curves with the same velocity at a point have the same normal acceleration components at that point.*

Proof. Select a parametrization and write $\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$. Then

$$\dot{\mathrm{q}} = \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix}$$

and

$$\begin{aligned} \ddot{\mathrm{q}} &= \left(\frac{d}{dt} \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \right) \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} \ddot{u} \\ \ddot{v} \end{bmatrix} \\ &= \begin{bmatrix} \dot{u} & \dot{v} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \mathrm{q}}{\partial u^2} & \frac{\partial^2 \mathrm{q}}{\partial u \partial v} \\ \frac{\partial^2 \mathrm{q}}{\partial v \partial u} & \frac{\partial^2 \mathrm{q}}{\partial v^2} \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} \ddot{u} \\ \ddot{v} \end{bmatrix}. \end{aligned}$$

Taking inner products with the normal will eliminate the second term as it is a tangent vector. So we obtain

$$\ddot{\mathrm{q}} \cdot \mathrm{n} = \begin{bmatrix} \dot{u} & \dot{v} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \mathrm{q}}{\partial u^2} \cdot \mathrm{n} & \frac{\partial^2 \mathrm{q}}{\partial u \partial v} \cdot \mathrm{n} \\ \frac{\partial^2 \mathrm{q}}{\partial v \partial u} \cdot \mathrm{n} & \frac{\partial^2 \mathrm{q}}{\partial v^2} \cdot \mathrm{n} \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix}.$$

This establishes the result since the velocity of a curve is determined by $(\dot{u}, \dot{v})$ and the parametrization of the surface. □

To define the second fundamental form we use the velocity characterization of the tangent space from proposition 4.2.7.

DEFINITION 5.1.2. The *second fundamental form* $\mathrm{II}(Z, Z)$ is defined as the normal component of $\ddot{\mathrm{q}}$, $\mathrm{II}(\dot{\mathrm{q}}, \dot{\mathrm{q}}) = \ddot{\mathrm{q}} \cdot \mathrm{n}$, where $\mathrm{q}(t)$ is a curve with $\dot{\mathrm{q}}(0) = Z$. To compute $\mathrm{II}(X, Y)$ we can use *polarization*:

$$\mathrm{II}(X, Y) = \frac{1}{2}\left(\mathrm{II}(X+Y, X+Y) - \mathrm{II}(X, X) - \mathrm{II}(Y, Y)\right).$$

The general matrix representation is given by

$$\begin{aligned} \mathrm{II}(X, Y) &= \begin{bmatrix} X^u & X^v \end{bmatrix} [\mathrm{II}] \begin{bmatrix} Y^u \\ Y^v \end{bmatrix} \\ &= \begin{bmatrix} X^u & X^v \end{bmatrix} \begin{bmatrix} L_{uu} & L_{uv} \\ L_{vu} & L_{vv} \end{bmatrix} \begin{bmatrix} Y^u \\ Y^v \end{bmatrix}, \\ &= \begin{bmatrix} X^u & X^v \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \mathrm{q}}{\partial u^2} \cdot \mathrm{n} & \frac{\partial^2 \mathrm{q}}{\partial u \partial v} \cdot \mathrm{n} \\ \frac{\partial^2 \mathrm{q}}{\partial v \partial u} \cdot \mathrm{n} & \frac{\partial^2 \mathrm{q}}{\partial v^2} \cdot \mathrm{n} \end{bmatrix} \begin{bmatrix} Y^u \\ Y^v \end{bmatrix}. \end{aligned}$$

Since there are two choices for the normal we also write II^{n} should we wish to specify the normal.

EXAMPLE 5.1.3 (Examples 4.1.10 and 4.2.5 continued). The second fundamental form of the generalized helicoid is computed by first noting that

$$\frac{\partial^2 \mathrm{q}}{\partial u^2} = \begin{bmatrix} 0 \\ 0 \\ f''(u) \end{bmatrix}, \quad \frac{\partial^2 \mathrm{q}}{\partial v \partial u} = \begin{bmatrix} -\sin v \\ \cos v \\ 0 \end{bmatrix}, \quad \frac{\partial^2 \mathrm{q}}{\partial v^2} = \begin{bmatrix} -u \cos v \\ -u \sin v \\ 0 \end{bmatrix}.$$

We then take inner products with the normal

$$\mathrm{n}(u, v) = \frac{1}{\sqrt{c^2 + u^2\left(1 + (f')^2\right)}} \begin{bmatrix} c \sin v - u f' \cos v \\ -c \cos v - u f' \sin v \\ u \end{bmatrix}$$

to obtain

$$[\mathrm{II}] = \frac{1}{\sqrt{c^2 + u^2\left(1 + (f')^2\right)}} \begin{bmatrix} u f'' & -c \\ -c & u^2 f' \end{bmatrix}.$$

As with space curves a regular curve $\mathrm{q}(t)$ on a surface has a unit tangent T. To use that the curve is on the surface we choose the normal n to the surface instead of the principal normal component $\mathrm{N}_{\mathrm{q}(t)}$ to the curve. From these two vectors we can define $\mathrm{S} = \mathrm{n} \times \mathrm{T}$ as the (oriented) normal to the curve that is tangent to the surface. In this way curve theory on surfaces is closer to the theory of planar curves, as we can think of S as the signed normal to the curve in the surface (see also section

3.3 for the special case of curves on spheres). Using an arclength parameter s we define the *normal curvature*

$$\kappa_n = \mathrm{II}\left(\mathrm{T},\mathrm{T}\right) = \mathrm{n}\cdot\frac{d\,\mathrm{T}}{ds} = -\frac{d\,\mathrm{n}}{ds}\cdot\mathrm{T},$$

the *geodesic curvature*

$$\kappa_g = \mathrm{S}\cdot\frac{d\,\mathrm{T}}{ds} = -\frac{d\mathrm{S}}{ds}\cdot\mathrm{T},$$

and the *geodesic torsion*

$$\tau_g = \mathrm{n}\cdot\frac{d\mathrm{S}}{ds} = -\frac{d\,\mathrm{n}}{ds}\cdot\mathrm{S}.$$

Note that the geodesic curvature of curves on the sphere from section 3.3 is consistent with the above definition.

Example 5.1.4. A plane always has vanishing second fundamental form as its normal is constant

$$\begin{aligned}\mathrm{II}\left(\dot{\mathrm{q}},\dot{\mathrm{q}}\right) &= \ddot{\mathrm{q}}\cdot\mathrm{n}\\ &= \frac{d}{dt}\left(\dot{\mathrm{q}}\cdot\mathrm{n}\right)-\dot{\mathrm{q}}\cdot\dot{\mathrm{n}}\\ &= -\dot{\mathrm{q}}\cdot\dot{\mathrm{n}}\\ &= 0.\end{aligned}$$

This means that any curve in this plane has vanishing normal curvature and geodesic torsion. The geodesic curvature is the signed curvature $\kappa_\pm$.

Example 5.1.5. A sphere of radius R centered at c is given by the equation

$$F\left(x,y,z\right) = \left|\mathrm{q}-\mathrm{c}\right|^2 = R^2 > 0.$$

The gradient is

$$\nabla F = 2\left(\mathrm{q}-\mathrm{c}\right) = 2\left(x-a,y-b,z-c\right),$$

which cannot vanish unless $\mathrm{q}=\mathrm{c}$. This shows that the sphere is a surface and also computes the two normals

$$\mathrm{n} = \pm\frac{1}{R}\left(\mathrm{q}-\mathrm{c}\right).$$

as $\left|\mathrm{q}-\mathrm{c}\right| = R$ (see theorem 4.1.9 and corollary 4.2.8). The $+$ gives us an outward pointing normal. Since n is perpendicular to all tangent vectors this shows that for a curve we have

$$\begin{aligned}\mathrm{II}\left(\dot{\mathrm{q}},\dot{\mathrm{q}}\right) &= \ddot{\mathrm{q}}\cdot\mathrm{n}\\ &= \frac{d}{dt}\left(\dot{\mathrm{q}}\cdot\mathrm{n}\right)-\dot{\mathrm{q}}\cdot\dot{\mathrm{n}}\\ &= -\dot{\mathrm{q}}\cdot\frac{d}{dt}\left(\pm\frac{1}{R}\left(\mathrm{q}-\mathrm{c}\right)\right)\\ &= \mp\frac{1}{R}\dot{\mathrm{q}}\cdot\dot{\mathrm{q}}\\ &= \mp\frac{1}{R}\mathrm{I}\left(\dot{\mathrm{q}},\dot{\mathrm{q}}\right).\end{aligned}$$

Thus $\mathrm{II} = \mp\frac{1}{R}\mathrm{I}$ and the normal curvature of any curve on the sphere is $\mp\frac{1}{R}$.

EXAMPLE 5.1.6. We can also relate normal curvature of certain special curves to the curvature of the curve as follows: For each unit vector $X \in T_pM$ to a surface with normal n consider the plane through p that is spanned by X, $\mathrm{n}(p)$. This plane has $Y = X \times \mathrm{n}(p)$ as a unit normal and intersects the surface in a unit speed curve $\mathrm{q}(s)$ with velocity $\dot{\mathrm{q}} = \mathrm{n} \times Y$, i.e., it is the integral curve for $\mathrm{n} \times Y$ that passes through $p = \mathrm{q}(0)$ (see section 1.1). Note that at p we have $\dot{q}(0) = X$, while at other points it changes with the change in the normal to the surface. The principal normal as well as the acceleration of this curve at $s = 0$ must be parallel to $\mathrm{n}(p)$ as it is a unit speed curve lies in the plane spanned by X, n. It now follows that

$$\kappa(0) = |\ddot{\mathrm{q}}(0)| = \pm\ddot{\mathrm{q}}(0) \cdot \mathrm{n}(p) = \pm\kappa_n(0) = \pm\,\mathrm{II}(X, X).$$

We shall show below that only planes and spheres have the property that the normal curvature is the same for all curves on a surface. Another interesting consequence is the important theorem, first noted by Euler and later in greater generality by Gauss (see theorem 5.3.6), that it is not possible to draws maps of the Earth with the property that all distances and angles are preserved.

We start by showing that if a surface admits a Cartesian parametrization, then the tangential part of the acceleration is calculated as in the plane.

PROPOSITION 5.1.7. *Consider a Cartesian parametrization* $\mathrm{q}(u, v)$ *and a curve* $\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$. *The tangential and normal components of the acceleration are given by*

$$\ddot{\mathrm{q}}^{\mathrm{I}} = \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} \ddot{u} \\ \ddot{v} \end{bmatrix} = \ddot{u}\frac{\partial \mathrm{q}}{\partial u} + \ddot{v}\frac{\partial \mathrm{q}}{\partial v},$$

$$\ddot{\mathrm{q}}^{\mathrm{II}} = \begin{bmatrix} \dot{u} & \dot{v} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \mathrm{q}}{\partial u^2} & \frac{\partial^2 \mathrm{q}}{\partial v \partial u} \\ \frac{\partial^2 \mathrm{q}}{\partial u \partial v} & \frac{\partial^2 \mathrm{q}}{\partial v^2} \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \dot{u}^2\frac{\partial^2 \mathrm{q}}{\partial u^2} + 2\dot{u}\dot{v}\frac{\partial^2 \mathrm{q}}{\partial u \partial v} + \dot{v}^2\frac{\partial^2 \mathrm{q}}{\partial v^2}.$$

PROOF. We saw above that

$$\ddot{\mathrm{q}} = \begin{bmatrix} \dot{u} & \dot{v} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \mathrm{q}}{\partial u^2} & \frac{\partial^2 \mathrm{q}}{\partial v \partial u} \\ \frac{\partial^2 \mathrm{q}}{\partial u \partial v} & \frac{\partial^2 \mathrm{q}}{\partial v^2} \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} \ddot{u} \\ \ddot{v} \end{bmatrix}.$$

To prove the proposition we need to show that the three vectors $\frac{\partial^2 \mathrm{q}}{\partial u^2}$, $\frac{\partial^2 \mathrm{q}}{\partial v^2}$, and $\frac{\partial^2 \mathrm{q}}{\partial v \partial u}$ are normal to the surface. This is equivalent to showing that

$$\begin{aligned} \frac{\partial^2 \mathrm{q}}{\partial u^2} \cdot \frac{\partial \mathrm{q}}{\partial w} &= 0, \\ \frac{\partial^2 \mathrm{q}}{\partial v \partial u} \cdot \frac{\partial \mathrm{q}}{\partial w} &= 0, \\ \frac{\partial^2 \mathrm{q}}{\partial v^2} \cdot \frac{\partial \mathrm{q}}{\partial w} &= 0. \end{aligned}$$

Note that as $w = u, v$ there are 6 identities. Using that $\frac{\partial \mathrm{q}}{\partial u}$, $\frac{\partial \mathrm{q}}{\partial v}$ are unit vectors we obtain

$$0 = \frac{\partial}{\partial w}\left(\frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial \mathrm{q}}{\partial u}\right) = 2\frac{\partial^2 \mathrm{q}}{\partial w \partial u} \cdot \frac{\partial \mathrm{q}}{\partial u} = \frac{\partial^2 \mathrm{q}}{\partial u \partial w} \cdot \frac{\partial \mathrm{q}}{\partial u}$$

and similarly

$$0 = \frac{\partial}{\partial w}\left(\frac{\partial \mathrm{q}}{\partial v} \cdot \frac{\partial \mathrm{q}}{\partial v}\right) = 2\frac{\partial^2 \mathrm{q}}{\partial w \partial v} \cdot \frac{\partial \mathrm{q}}{\partial v} = \frac{\partial^2 \mathrm{q}}{\partial v \partial w} \cdot \frac{\partial \mathrm{q}}{\partial v}.$$

This shows that four of the identities hold. Next we use that $\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}$ are perpendicular to conclude

$$0 = \frac{\partial}{\partial w}\left(\frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial \mathrm{q}}{\partial v}\right) = \frac{\partial^2 \mathrm{q}}{\partial w \partial u} \cdot \frac{\partial \mathrm{q}}{\partial v} + \frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial^2 \mathrm{q}}{\partial w \partial v}.$$

Depending on whether $w = u$ or v the second or first term on the right vanishes from what we just did. Thus the remaining term also vanishes. This takes care of the last two identities. □

THEOREM 5.1.8 (Euler, 1775). *A sphere does not admit a Cartesian parametrization.*

PROOF. Assume that the surface is a sphere of radius $R > 0$ centered at the origin and that part of the sphere admits a Cartesian parametrization $\mathrm{q}(u, v)$. Recall that great circles can be characterized as curves with acceleration normal to the sphere, i.e., the tangential acceleration vanishes $\ddot{\mathrm{q}}^{\mathrm{I}} = 0$ (see section 3.3 exercise 5 or exercise 12 in this section). Thus lines $(u(t), v(t))$ in the parameter domain with vanishing acceleration become parts of great circles $\mathrm{q}(u(t), v(t))$ on the sphere. It then simply remains to observe that if we select a small triangle in the u, v plane, then it is mapped to a congruent spherical triangle whose sides are parts of great circles. This, however, violates the spherical law of cosines as well as Girard's theorem (see section 4.4 exercise 23). To give a self contained argument here, select an equilateral triangle in the plane with side lengths ϵ. Then we obtain an equilateral triangle on the sphere with side lengths ϵ and interior angles $\frac{\pi}{3}$. As the sides are parts of great circles we can check explicitly if this is possible. Let the vertices be q_i, $i = 1, 2, 3$, then $\mathrm{q}_i \cdot \mathrm{q}_j = \cos\epsilon$ when $i \neq j$. The unit directions of the great circles at q_1 are given by

$$\begin{aligned}
\mathrm{v}_{12} &= \frac{\mathrm{q}_2 - (\mathrm{q}_2 \cdot \mathrm{q}_1)\,\mathrm{q}_1}{\sqrt{1 - (\mathrm{q}_2 \cdot \mathrm{q}_1)^2}} = \frac{\mathrm{q}_2 - \cos\epsilon\,\mathrm{q}_1}{\sqrt{1 - \cos^2\epsilon}}, \\
\mathrm{v}_{13} &= \frac{\mathrm{q}_3 - \cos\epsilon\,\mathrm{q}_1}{\sqrt{1 - \cos^2\epsilon}}
\end{aligned}$$

and consequently,

$$\begin{aligned}
\frac{1}{2} &= \mathrm{v}_{12} \cdot \mathrm{v}_{13} \\
&= \left(\frac{\mathrm{q}_2 - \cos\epsilon\,\mathrm{q}_1}{\sqrt{1 - \cos^2\epsilon}}\right) \cdot \left(\frac{\mathrm{q}_3 - \cos\epsilon\,\mathrm{q}_1}{\sqrt{1 - \cos^2\epsilon}}\right) \\
&= \frac{\cos\epsilon - 2\cos^2\epsilon + \cos^2\epsilon}{1 - \cos^2\epsilon} \\
&= \frac{\cos\epsilon - \cos^2\epsilon}{1 - \cos^2\epsilon} \\
&= \frac{\cos\epsilon}{1 + \cos\epsilon} \\
&< \frac{1}{2}
\end{aligned}$$

since $\cos\epsilon < 1$. So we have arrived at a contradiction. □

Exercises

(1) Show that $\frac{\partial \mathrm{n}}{\partial w}$ is always tangent to the surface.

(2) Show that

$$L_{w_1 w_2} = \frac{\partial^2 \mathrm{q}}{\partial w_1 \partial w_2} \cdot \mathrm{n} = -\frac{\partial \mathrm{q}}{\partial w_2} \cdot \frac{\partial \mathrm{n}}{\partial w_1}.$$

This shows that the derivatives of the normal can be computed knowing the first and second fundamental forms.

(3) Show that the unit normal is constant if and only if the surface is part of a plane.

(4) Show that [II] vanishes if and only if the normal vector is constant. (Hint: use exercise (2))

(5) Consider a parametrized surface $\mathrm{q}(u, v)$.
   (a) Show that $\frac{\partial^2 \mathrm{q}}{\partial u \partial v}$ is normal to the surface if and only if $\frac{\partial g_{uu}}{\partial v} = \frac{\partial g_{vv}}{\partial u} = 0$.
   (b) Show that $\frac{\partial^2 \mathrm{q}}{\partial u^2}$ is normal to the surface when $\frac{\partial g_{uu}}{\partial u} = \frac{\partial g_{uu}}{\partial v} = \frac{\partial g_{uv}}{\partial u} = 0$.
   (c) Show that $\frac{\partial^2 \mathrm{q}}{\partial u^2} + \frac{\partial^2 \mathrm{q}}{\partial v^2}$ is normal to the surface when $g_{uu} = g_{vv}$ and $g_{uv} = 0$.

(6) A curve $\mathrm{q}(t)$ on a surface is called an *asymptotic curve* if $\mathrm{II}(\dot{\mathrm{q}}, \dot{\mathrm{q}}) = 0$, i.e., κ_n vanishes.
   (a) Show that a curve is asymptotic if and only if its acceleration is tangent to the surface.
   (b) Show that the binormal to an asymptotic curve is normal to the surface.

(7) Let $c(s)$ be a unit speed curve with non-vanishing curvature. Show that c is an asymptotic curve on the ruled surface

$$\mathrm{q}(s, t) = c(s) + t\,\mathrm{N}_c(s),$$

where N_c is the principal normal to c as a space curve.

(8) Let $\mathrm{q}(s)$ be a unit speed curve on a surface with normal n. Show that $\kappa_g = 0$ if and only if

$$\det[\dot{\mathrm{q}}, \ddot{\mathrm{q}}, \mathrm{n}] = 0.$$

(9) Consider the parabolic surface $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ where $a, b > 0$.
   (a) Show that $q(t) = \left(at\cos\theta, bt\sin\theta, t^2\right)$ is a regular curve on this surface.
   (b) Show that when $a = b$ (surface of revolution), then $\kappa_g = 0$ for all θ.
   (c) Show that when $a \neq b$, then $\kappa_g = 0$ if and only if $\sin 2\theta = 0$.

(10) Show that latitudes on a sphere have constant κ_g.

(11) Let $\mathrm{q}(s)$ be a unit speed curve on a surface and let n be the normal to the surface. Show that

$$\frac{d}{ds}\begin{bmatrix} \mathrm{T} & \mathrm{S} & \mathrm{n} \end{bmatrix} = \begin{bmatrix} \mathrm{T} & \mathrm{S} & \mathrm{n} \end{bmatrix} \begin{bmatrix} 0 & -\kappa_g & -\kappa_n \\ \kappa_g & 0 & -\tau_g \\ \kappa_n & \tau_g & 0 \end{bmatrix}.$$

(12) For a curve on the unit sphere show that
   (a) $\tau_g = 0$.
   (b) $\kappa_g = 0$ if and only if it is a great circle.
   (c) κ_g is constant if and only if it is a circle.

(13) Let $\mathrm{q}(t)$ be a regular curve on a surface with n being the normal to the surface. Show that

$$\kappa_n = \frac{\mathrm{II}(\dot{\mathrm{q}}, \dot{\mathrm{q}})}{\mathrm{I}(\dot{\mathrm{q}}, \dot{\mathrm{q}})}, \ \kappa_g = \frac{\det(\dot{\mathrm{q}}, \ddot{\mathrm{q}}, \mathrm{n})}{(\mathrm{I}(\dot{\mathrm{q}}, \dot{\mathrm{q}}))^{3/2}}.$$

(14) Let $\mathrm{q}(u, v)$ be a parametrization such that $g_{uu} = 1$ and $g_{uv} = 0$. Prove that the u-curves are unit speed with acceleration that is perpendicular to the surface. Hint: The u-curves are given by $\mathrm{q}(u) = \mathrm{q}(u, v)$ where v is fixed.

(15) Consider a surface of revolution

$$\mathrm{q}(s, \theta) = (r(s)\cos(\theta), r(s)\sin(\theta), z(s)),$$

where, $r > 0$, $\dot{z} > 0$, and $(r(s), 0, z(s))$ is unit speed.

(a) Compute the second fundamental form.

(b) Compute $\kappa_g,\ \kappa_n,\ \tau_g$ for the meridians $\mathrm{q}(s) = \mathrm{q}(s, \theta)$. Conclude that their acceleration is perpendicular to the surface

(c) Compute $\kappa_g,\ \kappa_n,\ \tau_g$ for the latitudes $\mathrm{q}(\theta) = \mathrm{q}(s, \theta)$. Hint: The latitudes are not necessarily unit speed, but they do have constant speed.

(16) Let M be a surface with normal n and $X, Y \in T_pM$. Show that if $\mathrm{q}(t)$ is a curve with velocity X at $t = 0$ and $Y(t)$ is an extension of the vector Y to a vector field along q, then

$$\mathrm{II}(X, Y) = \mathrm{n} \cdot \frac{dY}{dt}(0).$$

(17) Let $\mathrm{q}(s)$ be a unit speed curve on a surface with normal n. Show that the space curvature κ is related to the geodesic and normal curvatures as follows

$$\kappa^2 = \kappa_g^2 + \kappa_n^2$$

and that the torsion is given by

$$\tau = \tau_g + \frac{\kappa_g\dot{\kappa}_n - \kappa_n\dot{\kappa}_g}{\kappa_g^2 + \kappa_n^2}.$$

Hint: Start by showing that

$$\begin{aligned}
\ddot{\mathrm{q}} &= \kappa_g \mathrm{S} + \kappa_n \mathrm{n}, \\
\dddot{\mathrm{q}} &= -\left(\kappa_g^2 + \kappa_n^2\right)\mathrm{T} + \left(\dot{\kappa}_g - \kappa_n\tau_g\right)\mathrm{S} + \left(\dot{\kappa}_n + \kappa_g\tau_g\right)\mathrm{n}.
\end{aligned}$$

(18) Let M be a surface given by an equation $F(x, y, z) = R$.

(a) If $\mathrm{q}(t)$ is a curve on M show that

$$\begin{aligned}
\ddot{\mathrm{q}} \cdot \nabla F &= -\dot{\mathrm{q}}^t \begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial x \partial z} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial y \partial z} \\ \frac{\partial^2 F}{\partial z \partial x} & \frac{\partial^2 F}{\partial z \partial y} & \frac{\partial^2 F}{\partial z^2} \end{bmatrix} \dot{\mathrm{q}} \\
&= -\dot{\mathrm{q}}^t \left[\frac{\partial \nabla F}{\partial(x, y, z)}\right] \dot{\mathrm{q}}.
\end{aligned}$$

(b) Show that

$$\mathrm{II}(X, Y) = -\frac{X^t \left[\frac{\partial \nabla F}{\partial(x,y,z)}\right] Y}{|\nabla F|}.$$

(19) Assume that a unit speed curve $\mathrm{q}(s) = \mathrm{q}(u(s), v(s))$ on a parametrized surface satisfies an equation $F(u, v) = R$.

(a) If we use $\partial_w F = \frac{\partial F}{\partial w}$ show that $\dot{u}\partial_u F + \dot{v}\partial_v F = 0$.

(b) Show that

$$\begin{aligned}\dot{\mathrm{q}} &= \dot{u}\frac{\partial\,\mathrm{q}}{\partial u}+\dot{v}\frac{\partial\,\mathrm{q}}{\partial v} \\ &= \frac{\pm 1}{\sqrt{g_{uu}\partial_v F\partial_v F-2g_{uv}\partial_u F\partial_v F+g_{vv}\partial_u F\partial_u F}}\left(-\partial_v F\frac{\partial\,\mathrm{q}}{\partial u}+\partial_u F\frac{\partial\,\mathrm{q}}{\partial v}\right).\end{aligned}$$

This means that the unit tangent can be calculated without reference to the parametrization of the curve.

(c) Show that if we use this formula for the velocity, then the geodesic curvature can be computed as

$$\kappa_g = \frac{\frac{\partial}{\partial u}\left(\dot{\mathrm{q}}\cdot\frac{\partial\,\mathrm{q}}{\partial v}\right)-\frac{\partial}{\partial v}\left(\dot{\mathrm{q}}\cdot\frac{\partial\,\mathrm{q}}{\partial u}\right)}{\sqrt{\det\left[\mathrm{I}\right]}}.$$

(d) Generalize this to the situation where a unit speed curve satisfies a differential relation

$$P\dot{u}+Q\dot{v}=0,$$

where $P=P(u,v)$ and $Q=Q(u,v)$.

5.2. The Gauss and Weingarten Maps and Equations

In the last section we calculated the normal part of the acceleration of a curve. To gain a better understanding of the tangential component we need to further analyze the second partial derivatives of a parametrized surface. We use proposition 4.3.5 to decompose these derivatives in to tangential and normal components

$$\frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2} = \begin{bmatrix}\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix}\left[\mathrm{I}\right]^{-1}\begin{bmatrix}\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix}^t\frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2}+\left(\frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2}\cdot\mathrm{n}\right)\mathrm{n},$$

where w_1, w_2 can be either u or v. In the previous section the normal component was identified as an entry in the matrix representation of the second fundamental form. The tangential part is denoted by

$$\Gamma_{w_1w_2} = \begin{bmatrix}\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix}\left[\mathrm{I}\right]^{-1}\begin{bmatrix}\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix}^t\frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2}.$$

DEFINITION 5.2.1. The *Christoffel symbols of the first kind* are defined as

$$\begin{aligned}\Gamma_{w_1w_2w} &= \frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2}\cdot\frac{\partial\,\mathrm{q}}{\partial w}, \\ \begin{bmatrix}\Gamma_{w_1w_2u} & \Gamma_{w_1w_2v}\end{bmatrix} &= \begin{bmatrix}\frac{\partial\,\mathrm{q}}{\partial u}\cdot\frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2} & \frac{\partial\,\mathrm{q}}{\partial v}\cdot\frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2}\end{bmatrix} \\ &= \begin{bmatrix}\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix}^t\frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2}\end{aligned}$$

and the *Christoffel symbols of the second kind* are defined as

$$\begin{aligned}\Gamma^w_{w_1w_2} &= g^{wu}\Gamma_{w_1w_2u}+g^{wv}\Gamma_{w_1w_2v}, \\ \begin{bmatrix}\Gamma^u_{w_1w_2} \\ \Gamma^v_{w_1w_2}\end{bmatrix} &= \begin{bmatrix}g^{uu} & g^{uv} \\ g^{vu} & g^{vv}\end{bmatrix}\begin{bmatrix}\Gamma_{w_1w_2u} \\ \Gamma_{w_1w_2v}\end{bmatrix} \\ &= \left[\mathrm{I}\right]^{-1}\begin{bmatrix}\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix}^t\frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2},\end{aligned}$$

where the entries in $[\mathrm{I}]^{-1}$ are denoted by

$$\begin{bmatrix} g^{uu} & g^{uv} \\ g^{vu} & g^{vv} \end{bmatrix} = \begin{bmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{bmatrix}^{-1} = \frac{1}{g_{uu}g_{vv} - g_{uv}^2} \begin{bmatrix} g_{vv} & -g_{uv} \\ -g_{vu} & g_{uu} \end{bmatrix}.$$

This gives us the tangential component as

$$\begin{aligned} \Gamma_{w_1 w_2} &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} [\mathrm{I}]^{-1} \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix}^t \frac{\partial^2 \mathrm{q}}{\partial w_1 \partial w_2} \\ &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} [\mathrm{I}]^{-1} \begin{bmatrix} \Gamma_{w_1 w_2 u} \\ \Gamma_{w_1 w_2 v} \end{bmatrix} \\ &= \Gamma^u_{w_1 w_2} \frac{\partial \mathrm{q}}{\partial u} + \Gamma^v_{w_1 w_2} \frac{\partial \mathrm{q}}{\partial v}. \end{aligned}$$

The second derivatives of $\mathrm{q}(u,v)$ can now be expressed as follows in terms of the Christoffel symbols of the second kind and the second fundamental form. These are often called the *Gauss formulas* or *equations*:

$$\begin{aligned} \frac{\partial^2 \mathrm{q}}{\partial u^2} &= \Gamma^u_{uu} \frac{\partial \mathrm{q}}{\partial u} + \Gamma^v_{uu} \frac{\partial \mathrm{q}}{\partial v} + L_{uu}\,\mathrm{n}, \\ \frac{\partial^2 \mathrm{q}}{\partial u \partial v} &= \Gamma^u_{uv} \frac{\partial \mathrm{q}}{\partial u} + \Gamma^v_{uv} \frac{\partial \mathrm{q}}{\partial v} + L_{uv}\,\mathrm{n} = \frac{\partial^2 \mathrm{q}}{\partial v \partial u}, \\ \frac{\partial^2 \mathrm{q}}{\partial v^2} &= \Gamma^u_{vv} \frac{\partial \mathrm{q}}{\partial u} + \Gamma^v_{vv} \frac{\partial \mathrm{q}}{\partial v} + L_{vv}\,\mathrm{n}, \end{aligned}$$

or

$$\frac{\partial^2 \mathrm{q}}{\partial w_1 \partial w_2} = \Gamma^u_{w_1 w_2} \frac{\partial \mathrm{q}}{\partial u} + \Gamma^v_{w_1 w_2} \frac{\partial \mathrm{q}}{\partial v} + L_{w_1 w_2}\,\mathrm{n},$$

or

$$\frac{\partial}{\partial w} \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} & \mathrm{n} \end{bmatrix} \begin{bmatrix} \Gamma^u_{wu} & \Gamma^u_{wv} \\ \Gamma^v_{wu} & \Gamma^v_{wv} \\ L_{wu} & L_{wv} \end{bmatrix}.$$

EXAMPLE 5.2.2. Consider a Cartesian parametrization. We saw in the proof of proposition 5.1.7 that the second derivatives $\frac{\partial^2 \mathrm{q}}{\partial w_1 \partial w_2}$ are normal to the surface. This implies that the Christoffel symbols vanish.

As we shall see, and indeed already saw in section 1.4 when considering polar coordinates in the plane, these formulas are important for defining accelerations of curves. They are also important for giving a proper definition of the Hessian or second derivative matrix of a function on a surface. This will be explored in an exercise later.

For now we note that this gives us a formula for the acceleration of a curve:

COROLLARY 5.2.3. *The acceleration of a curve can be calculated as*

$$\begin{aligned} \ddot{\mathrm{q}} &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} & \mathrm{n} \end{bmatrix} \begin{bmatrix} \ddot{u} + \Gamma^u(\dot{\mathrm{q}}, \dot{\mathrm{q}}) \\ \ddot{v} + \Gamma^v(\dot{\mathrm{q}}, \dot{\mathrm{q}}) \\ \mathrm{II}(\dot{\mathrm{q}}, \dot{\mathrm{q}}) \end{bmatrix} \\ &= (\ddot{u} + \Gamma^u(\dot{\mathrm{q}}, \dot{\mathrm{q}})) \frac{\partial \mathrm{q}}{\partial u} + (\ddot{v} + \Gamma^v(\dot{\mathrm{q}}, \dot{\mathrm{q}})) \frac{\partial \mathrm{q}}{\partial v} + \mathrm{n}\,\mathrm{II}(\dot{\mathrm{q}}, \dot{\mathrm{q}}), \end{aligned}$$

where

$$\Gamma^w(\dot{\mathrm{q}}, \dot{\mathrm{q}}) = \sum_{w_1, w_2 = u, v} \Gamma^w_{w_1 w_2} \dot{w}_1 \dot{w}_2 = \begin{bmatrix} \dot{u} & \dot{v} \end{bmatrix} \begin{bmatrix} \Gamma^w_{uu} & \Gamma^w_{uv} \\ \Gamma^w_{vu} & \Gamma^w_{vv} \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix}.$$

PROOF. This follows directly from

$$\ddot{\mathrm{q}} = \begin{bmatrix} \dot{u} & \dot{v} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \mathrm{q}}{\partial u^2} & \frac{\partial^2 \mathrm{q}}{\partial u \partial v} \\ \frac{\partial^2 \mathrm{q}}{\partial v \partial u} & \frac{\partial^2 \mathrm{q}}{\partial v^2} \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} \begin{bmatrix} \ddot{u} \\ \ddot{v} \end{bmatrix}$$

and the Gauss formulas above. □

Note that the tangential component is quite complicated

$$\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix} [\mathrm{I}]^{-1} \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{bmatrix}^t \ddot{\mathrm{q}} = \ddot{\mathrm{q}}^{\mathrm{I}} = \frac{\partial \mathrm{q}}{\partial u} \left(\ddot{u} + \Gamma^u \left(\dot{\mathrm{q}}, \dot{\mathrm{q}} \right) \right) + \frac{\partial \mathrm{q}}{\partial v} \left(\ddot{v} + \Gamma^v \left(\dot{\mathrm{q}}, \dot{\mathrm{q}} \right) \right).$$

But it seems to be a more genuine acceleration as it includes second derivatives. It tells us what acceleration we feel on the surface.

To complete the Gauss formulas it is natural to also include the derivatives of the normal vector.

DEFINITION 5.2.4. The *Gauss map* for an orientable surface M with normal n is the map $\mathrm{n} : M \to S^2(1)$ that takes each point to the chosen normal at that point. The *Weingarten map* at a point $p \in M$ is the linear map $L : T_pM \to T_pM$ defined as the negative of the differential of n:

$$L = -D\,\mathrm{n}.$$

REMARK 5.2.5. The definition of the Weingarten map requires some explanation as the differential should be a linear map

$$D\,\mathrm{n} : T_pM \to T_{\mathrm{n}(p)}S^2(1).$$

However, the normal vector to any point $\mathrm{x} \in S^2(1)$ is simply $\mathrm{n} = \pm\mathrm{x}$. As the tangent space is the orthogonal complement to the normal vector it follows that

$$T_pM = T_{\mathrm{n}(p)}S^2(1).$$

For a parametrized surface this tells us.

PROPOSITION 5.2.6 (The Weingarten Equations). *For a parametrized surface* $\mathrm{q}(u,v)$ *we have*

$$\begin{aligned} -\frac{\partial \mathrm{n}}{\partial u} &= L\left(\frac{\partial \mathrm{q}}{\partial u}\right), \\ -\frac{\partial \mathrm{n}}{\partial v} &= L\left(\frac{\partial \mathrm{q}}{\partial v}\right). \end{aligned}$$

More generally, for a curve $\mathrm{q}(t)$ *on the surface*

$$-\frac{d\,\mathrm{n}\circ\mathrm{q}}{dt} = L\left(\frac{d\,\mathrm{q}}{dt}\right).$$

PROOF. The equations simply follow from the chain rule and the first two are special cases of the last. If we write the curve $\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$, then

$$\begin{aligned} L\left(\frac{d\,\mathrm{q}}{dt}\right) &= -D\,\mathrm{n}\left(\frac{d\,\mathrm{q}}{dt}\right) \\ &= -\frac{d\,\mathrm{n}\circ\mathrm{q}}{dt} \\ &= -\left(\frac{\partial \mathrm{n}}{\partial u}\frac{du}{dt} + \frac{\partial \mathrm{n}}{\partial v}\frac{dv}{dt}\right). \end{aligned}$$

This proves the claim □

Next we show that the Weingarten map L is a self-adjoint map with respect to the first fundamental form.

PROPOSITION 5.2.7. *The Weingarten map is abstractly related to the second fundamental form through the first fundamental form by the formula:*

$$\mathrm{I}\left(L\left(X\right),Y\right)=\mathrm{II}\left(X,Y\right)=\mathrm{I}\left(X,L\left(Y\right)\right)$$

In particular, L is self-adjoint as II *is symmetric.*

PROOF. Since the second fundamental form is symmetric $\mathrm{II}\left(X,Y\right)=\mathrm{II}\left(Y,X\right)$, it follows that we only need to show that $\mathrm{I}\left(L\left(X\right),Y\right)=\mathrm{II}\left(X,Y\right)$, as we have

$$\begin{aligned}\mathrm{I}\left(X,L\left(Y\right)\right) &= \mathrm{I}\left(L\left(Y\right),X\right)\\ &= \mathrm{II}\left(Y,X\right)\\ &= \mathrm{II}\left(X,Y\right).\end{aligned}$$

Next observe that since L is linear it suffices to prove that

$$\mathrm{II}\left(\frac{\partial\,\mathrm{q}}{\partial w_1},\frac{\partial\,\mathrm{q}}{\partial w_2}\right)=\mathrm{I}\left(L\left(\frac{\partial\,\mathrm{q}}{\partial w_1}\right),\frac{\partial\,\mathrm{q}}{\partial w_2}\right)$$

for all choices $w_1,w_2\in\{u,v\}$ where $\mathrm{q}\left(u,v\right)$ is a parametrization. Using that $\frac{\partial\,\mathrm{q}}{\partial w_2}$ and n are perpendicular it follows that

$$\begin{aligned}0 &= \frac{\partial}{\partial w_1}\left(\frac{\partial\,\mathrm{q}}{\partial w_2}\cdot\mathrm{n}\right)\\ &= \frac{\partial^2\,\mathrm{q}}{\partial w_1\partial w_2}\cdot\mathrm{n}+\frac{\partial\,\mathrm{q}}{\partial w_2}\cdot\frac{\partial\,\mathrm{n}\circ\mathrm{q}}{\partial w_1}\\ &= \mathrm{II}\left(\frac{\partial\,\mathrm{q}}{\partial w_1},\frac{\partial\,\mathrm{q}}{\partial w_2}\right)-\frac{\partial\,\mathrm{q}}{\partial w_2}\cdot L\left(\frac{\partial\,\mathrm{q}}{\partial w_1}\right)\\ &= \mathrm{II}\left(\frac{\partial\,\mathrm{q}}{\partial w_1},\frac{\partial\,\mathrm{q}}{\partial w_2}\right)-\mathrm{I}\left(L\left(\frac{\partial\,\mathrm{q}}{\partial w_1}\right),\frac{\partial\,\mathrm{q}}{\partial w_2}\right).\end{aligned}$$

This proves the claim. □

All in all this is still a bit abstract, but the relationship between the Weingarten map and the first and second fundamental forms allow us to obtain explicit formulas for a parametrized surface.

Given a parametrized surface $\mathrm{q}\left(u,v\right)$ the entries in the matrix representation of the Weingarten map are defined as

$$\begin{aligned}\begin{bmatrix} L\left(\frac{\partial\,\mathrm{q}}{\partial u}\right) & L\left(\frac{\partial\,\mathrm{q}}{\partial v}\right)\end{bmatrix} &= \begin{bmatrix}\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix}[L]\\ &= \begin{bmatrix}\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix}\begin{bmatrix}L_u^u & L_v^u\\ L_u^v & L_v^v\end{bmatrix}.\end{aligned}$$

This matrix representation can be calculated as follows.

PROPOSITION 5.2.8. *The matrix representations of the Weingarten map and the second fundamental form satisfy:*

$$[L]=[\mathrm{I}]^{-1}[\mathrm{II}]$$

and

$$
\begin{aligned}
[\mathrm{II}] &= -\left[\begin{array}{cc} \frac{\partial\,\mathrm{n}}{\partial u} & \frac{\partial\,\mathrm{n}}{\partial v} \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v} \end{array}\right] \\
&= -\left[\begin{array}{cc} \frac{\partial\,\mathrm{n}}{\partial u}\cdot\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{n}}{\partial u}\cdot\frac{\partial\,\mathrm{q}}{\partial v} \\ \frac{\partial\,\mathrm{n}}{\partial v}\cdot\frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{n}}{\partial v}\cdot\frac{\partial\,\mathrm{q}}{\partial v} \end{array}\right] \\
&= -\left[\begin{array}{cc} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v} \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial\,\mathrm{n}}{\partial u} & \frac{\partial\,\mathrm{n}}{\partial v} \end{array}\right].
\end{aligned}
$$

PROOF. To establish the formula for [II] use that n is perpendicular to $\frac{\partial\,\mathrm{q}}{\partial w_2}$ and note that

$$
\begin{aligned}
L_{w_1 w_2} &= \frac{\partial^2\,\mathrm{q}}{\partial w_1 \partial w_2}\cdot\mathrm{n} \\
&= \left(\frac{\partial}{\partial w_1}\left(\frac{\partial\,\mathrm{q}}{\partial w_2}\right)\right)\cdot\mathrm{n} \\
&= \frac{\partial}{\partial w_1}\left(\frac{\partial\,\mathrm{q}}{\partial w_2}\cdot\mathrm{n}\right) - \frac{\partial\,\mathrm{q}}{\partial w_2}\cdot\frac{\partial\,\mathrm{n}}{\partial w_1} \\
&= -\frac{\partial\,\mathrm{q}}{\partial w_2}\cdot\frac{\partial\,\mathrm{n}}{\partial w_1}.
\end{aligned}
$$

It now follows that

$$
\begin{aligned}
[\mathrm{II}] &= \left[\begin{array}{cc} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v} \end{array}\right]^t \left[\begin{array}{cc} -\frac{\partial\,\mathrm{n}}{\partial u} & -\frac{\partial\,\mathrm{n}}{\partial v} \end{array}\right] \\
&= \left[\begin{array}{cc} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v} \end{array}\right]^t \left[\begin{array}{cc} L\left(\frac{\partial\,\mathrm{q}}{\partial u}\right) & L\left(\frac{\partial\,\mathrm{q}}{\partial v}\right) \end{array}\right] \\
&= \left[\begin{array}{cc} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v} \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v} \end{array}\right][L] \\
&= [\mathrm{I}]\,[L]\,.
\end{aligned}
$$

□

REMARK 5.2.9. It is important to realize that while L is self-adjoint its matrix representation

$$[L] = [\mathrm{I}]^{-1}\,[\mathrm{II}]$$

need not be symmetric. In fact, as [I] and [II] are symmetric it follows that

$$[L]^t = [\mathrm{II}]\,[\mathrm{I}]^{-1}\,.$$

So $[L]$ is only symmetric if [I] and [II] commute.

The *Weingarten equations* are the formulas for the derivatives of the normal:

$$\frac{\partial\,\mathrm{n}}{\partial w} = -L_w^u\frac{\partial\,\mathrm{q}}{\partial u} - L_w^v\frac{\partial\,\mathrm{q}}{\partial v} = -L\left(\frac{\partial\,\mathrm{q}}{\partial w}\right).$$

Together the Gauss formulas and Weingarten equations tell us how the derivatives of our basis $\frac{\partial\,\mathrm{q}}{\partial u}, \frac{\partial\,\mathrm{q}}{\partial v}, \mathrm{n}$ relate back to the basis. They can be collected as follows:

COROLLARY 5.2.10 (The Gauss and Weingarten Formulas).

$$
\begin{aligned}
\frac{\partial}{\partial w}\left[\begin{array}{ccc} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v} & \mathrm{n} \end{array}\right] &= \left[\begin{array}{ccc} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v} & \mathrm{n} \end{array}\right][D_w] \\
&= \left[\begin{array}{ccc} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v} & \mathrm{n} \end{array}\right] \left[\begin{array}{ccc} \Gamma_{wu}^u & \Gamma_{wv}^u & -L_w^u \\ \Gamma_{wu}^v & \Gamma_{wv}^v & -L_w^v \\ L_{wu} & L_{wv} & 0 \end{array}\right].
\end{aligned}
$$

Finally we show how the Christoffel symbols can be calculated directly from the first fundamental form without knowing the second derivatives $\frac{\partial^2 \mathrm{q}}{\partial w_1 \partial w_2}$.

PROPOSITION 5.2.11. *The Christoffel symbols of the first kind satisfy*

$$\begin{aligned}
\Gamma_{uuu} &= \frac{1}{2}\frac{\partial g_{uu}}{\partial u}, \\
\Gamma_{uvu} &= \frac{1}{2}\frac{\partial g_{uu}}{\partial v} = \Gamma_{vuu}, \\
\Gamma_{vvv} &= \frac{1}{2}\frac{\partial g_{vv}}{\partial v}, \\
\Gamma_{uvv} &= \frac{1}{2}\frac{\partial g_{vv}}{\partial u} = \Gamma_{vuv}, \\
\Gamma_{uuv} &= \frac{\partial g_{uv}}{\partial u} - \frac{1}{2}\frac{\partial g_{uu}}{\partial v}, \\
\Gamma_{vvu} &= \frac{\partial g_{uv}}{\partial v} - \frac{1}{2}\frac{\partial g_{vv}}{\partial u}.
\end{aligned}$$

PROOF. We prove only two of these as the proofs are all similar. First use the product rule to see

$$\Gamma_{uvu} = \frac{\partial^2 \mathrm{q}}{\partial u \partial v} \cdot \frac{\partial \mathrm{q}}{\partial u} = \left(\frac{\partial}{\partial v}\left(\frac{\partial \mathrm{q}}{\partial u}\right)\right) \cdot \frac{\partial \mathrm{q}}{\partial u} = \frac{1}{2}\frac{\partial}{\partial v}\left(\frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial \mathrm{q}}{\partial u}\right) = \frac{1}{2}\frac{\partial g_{uu}}{\partial v}.$$

Now use this together with the product rule to find

$$\begin{aligned}
\Gamma_{uuv} &= \frac{\partial^2 \mathrm{q}}{\partial u \partial u} \cdot \frac{\partial \mathrm{q}}{\partial v} \\
&= \left(\frac{\partial}{\partial u}\left(\frac{\partial \mathrm{q}}{\partial u}\right)\right) \cdot \frac{\partial \mathrm{q}}{\partial v} \\
&= \frac{\partial}{\partial u}\left(\frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial \mathrm{q}}{\partial v}\right) - \left(\frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial^2 \mathrm{q}}{\partial u \partial v}\right) \\
&= \frac{\partial g_{uv}}{\partial u} - \frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial^2 \mathrm{q}}{\partial v \partial u} \\
&= \frac{\partial g_{uv}}{\partial u} - \frac{1}{2}\frac{\partial g_{uu}}{\partial v}.
\end{aligned}$$

□

EXAMPLE 5.2.12. While these formulas for the Christoffel symbols can't be made simpler as such, it is possible to be a bit more efficient in several concrete situations. Specifically, we often do calculations in orthogonal coordinates, i.e., $g_{uv} \equiv 0$. In such coordinates

$$\begin{aligned}
g^{uv} &= 0, \\
g^{uu} &= (g_{uu})^{-1}, \\
g^{vv} &= (g_{vv})^{-1},
\end{aligned}$$

$$\begin{aligned}
\Gamma_{uuu} &= \frac{1}{2}\frac{\partial g_{uu}}{\partial u}, \\
\Gamma_{uvu} &= \frac{1}{2}\frac{\partial g_{uu}}{\partial v} = \Gamma_{vuu}, \\
\Gamma_{vvv} &= \frac{1}{2}\frac{\partial g_{vv}}{\partial v}, \\
\Gamma_{uvv} &= \frac{1}{2}\frac{\partial g_{vv}}{\partial u} = \Gamma_{vuv}, \\
\Gamma_{uuv} &= -\frac{1}{2}\frac{\partial g_{uu}}{\partial v}, \\
\Gamma_{vvu} &= -\frac{1}{2}\frac{\partial g_{vv}}{\partial u},
\end{aligned}$$

and

$$\begin{aligned}
\Gamma_{uu}^{u} &= \frac{1}{2}g^{uu}\frac{\partial g_{uu}}{\partial u} = \frac{1}{2}\frac{1}{g_{uu}}\frac{\partial g_{uu}}{\partial u} = \frac{\partial \ln\sqrt{g_{uu}}}{\partial u}, \\
\Gamma_{uu}^{v} &= -\frac{1}{2}g^{vv}\frac{\partial g_{uu}}{\partial v} = -\frac{1}{2}\frac{1}{g_{vv}}\frac{\partial g_{uu}}{\partial v}, \\
\Gamma_{vv}^{v} &= \frac{1}{2}g^{vv}\frac{\partial g_{vv}}{\partial v} = \frac{1}{2}\frac{1}{g_{vv}}\frac{\partial g_{vv}}{\partial v} = \frac{\partial \ln\sqrt{g_{vv}}}{\partial v}, \\
\Gamma_{vv}^{u} &= -\frac{1}{2}g^{uu}\frac{\partial g_{vv}}{\partial u} = -\frac{1}{2}\frac{1}{g_{uu}}\frac{\partial g_{vv}}{\partial u}, \\
\Gamma_{uv}^{u} &= \frac{1}{2}g^{uu}\frac{\partial g_{uu}}{\partial v} = \frac{1}{2}\frac{1}{g_{uu}}\frac{\partial g_{uu}}{\partial v} = \frac{\partial \ln\sqrt{g_{uu}}}{\partial v}, \\
\Gamma_{uv}^{v} &= \frac{1}{2}g^{vv}\frac{\partial g_{vv}}{\partial u} = \frac{1}{2}\frac{1}{g_{vv}}\frac{\partial g_{vv}}{\partial u} = \frac{\partial \ln\sqrt{g_{vv}}}{\partial u}.
\end{aligned}$$

Often there might be even more specific information. This could be that the metric coefficients only depend on one of the parameters, or that $g_{uu} = 1$. In such circumstances it is quite manageable to calculate the Christoffel symbols. What is more, we show in proposition 7.4.1 that it is always possible to find parametrizations where $g_{uu} \equiv 1$ and $g_{uv} \equiv 0$. In this case:

$$\Gamma_{uu}^{u} = \Gamma_{uu}^{v} = \Gamma_{uv}^{u} = \Gamma_{vu}^{u} = \Gamma_{uuu} = \Gamma_{uuv} = \Gamma_{uvu} = \Gamma_{vuu} = 0.$$

Example 5.2.13. Consider a Cartesian parametrization. We saw in the proof of proposition 5.1.7 that the second derivatives $\frac{\partial^2 \mathrm{q}}{\partial w_1 \partial w_2}$ are normal to the surface. This fact now also follows from the fact that the Christoffel symbols vanish.

Exercises

(1) For a surface of revolution

$$\mathrm{q}(t,\theta) = (r(t)\cos(\theta), r(t)\sin(\theta), z(t))$$

compute the first and second fundamental forms and the Weingarten map.

(2) Compute the matrix representation of the Weingarten map for a Monge patch $\mathrm{q}(x,y) = (x, y, f(x,y))$ with respect to the basis $\frac{\partial \mathrm{q}}{\partial x}, \frac{\partial \mathrm{q}}{\partial y}$.

(3) Show that if a surface satisfies $\mathrm{II} = \pm\frac{1}{R}\mathrm{I}$, then it is part of a sphere of radius R. Hint: Show that $\mathrm{n} \pm \frac{1}{R}\mathrm{q}$ is constant and use that to find the center of the sphere.

(4) Let M be a surface with normal n and $X, Y \in T_pM$. Show that if $\mathrm{q}(t)$ is a curve with velocity X at $t = 0$, then

$$\mathrm{II}(X,Y) = -Y \cdot \frac{d\,\mathrm{n}\circ\mathrm{q}}{dt}(0).$$

(5) Show that for a curve on a surface the geodesic torsion satisfies

$$\tau_g = \mathrm{II}(\mathrm{T},\mathrm{S}).$$

(6) Show that g_{uu}, g_{uv}, and g_{vv} are constant if and only if the Christoffel symbols of the first kind vanish.

(7) Show that

$$\frac{\partial g_{w_1w_2}}{\partial w} = \Gamma_{ww_1w_2} + \Gamma_{ww_2w_1}$$

and use these equations to show that

$$2\Gamma_{w_1w_2w} = \frac{\partial g_{ww_2}}{\partial w_1} + \frac{\partial g_{ww_1}}{\partial w_2} - \frac{\partial g_{w_1w_2}}{\partial w}.$$

(8) Show that

$$\Gamma_{w_1w_2w} = g_{wu}\Gamma^u_{w_1w_2} + g_{wv}\Gamma^v_{w_1w_2}.$$

(9) Show that

$$\frac{\partial \det[\mathrm{I}]}{\partial w} = \frac{2}{\det[\mathrm{I}]}\left(\Gamma^u_{wu} + \Gamma^v_{wv}\right),$$

$$\frac{\partial\sqrt{\det[\mathrm{I}]}}{\partial w} = \sqrt{\det[\mathrm{I}]}\left(\Gamma^u_{wu} + \Gamma^v_{wv}\right),$$

and

$$\frac{\partial \log\sqrt{\det[\mathrm{I}]}}{\partial w} = \Gamma^u_{wu} + \Gamma^v_{wv}.$$

(10) Let θ be the angle between $\frac{\partial\,\mathrm{q}}{\partial u}$ and $\frac{\partial\,\mathrm{q}}{\partial v}$. Show that

$$\log\sin\theta = \log\sqrt{\det[\mathrm{I}]} - \frac{1}{2}\log g_{uu} - \frac{1}{2}\log g_{vv},$$

$$\cot\theta = \frac{g_{uv}}{\sqrt{\det[\mathrm{I}]}},$$

and

$$\frac{\partial\theta}{\partial w} = -\frac{\sqrt{\det[\mathrm{I}]}}{g_{uu}}\Gamma^v_{wu} - \frac{\sqrt{\det[\mathrm{I}]}}{g_{vv}}\Gamma^u_{wv}.$$

(11) Show that

$$\begin{aligned}
\mathrm{I}\left(\Gamma_{w_1w_2},\Gamma_{w_3w_4}\right) &= \begin{bmatrix} \Gamma^u_{w_1w_2} & \Gamma^v_{w_1w_2} \end{bmatrix}[\mathrm{I}]\begin{bmatrix} \Gamma^u_{w_3w_4} \\ \Gamma^v_{w_3w_4} \end{bmatrix} \\
&= \begin{bmatrix} \Gamma_{w_1w_2u} & \Gamma_{w_1w_2v} \end{bmatrix}\begin{bmatrix} \Gamma^u_{w_3w_4} \\ \Gamma^v_{w_3w_4} \end{bmatrix} \\
&= \begin{bmatrix} \Gamma_{w_1w_2u} & \Gamma_{w_1w_2v} \end{bmatrix}[\mathrm{I}]^{-1}\begin{bmatrix} \Gamma_{w_3w_4u} \\ \Gamma_{w_3w_4v} \end{bmatrix}.
\end{aligned}$$

(12) Show directly from the formulas for the Christoffel symbols in terms of the first fundamental form that

$$\frac{\partial}{\partial v}\Gamma_{uuv} - \frac{\partial}{\partial u}\Gamma_{uvv} = \frac{\partial}{\partial u}\Gamma_{vvu} - \frac{\partial}{\partial v}\Gamma_{uvu}$$

and

$$\frac{\partial}{\partial v}\Gamma_{uuv} - \frac{\partial}{\partial u}\Gamma_{uvv} = -\frac{1}{2}\frac{\partial^2 g_{uu}}{\partial v^2} + \frac{\partial^2 g_{uv}}{\partial u \partial v} - \frac{1}{2}\frac{\partial^2 g_{vv}}{\partial u^2}.$$

(13) (Surface of revolution) Find the Christoffel symbols of the first and second kind when the first fundamental form is given by

$$[\mathrm{I}] = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix},$$

where $r = r(u) > 0$.

(14) (Polar and Fermi coordinates) Find the Christoffel symbols of the first and second kind when the first fundamental form is given by

$$[\mathrm{I}] = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix},$$

where $r = r(u, v) > 0$. Gauss showed that such coordinates exist around any point in a surface with r denoting the "intrinsic" distance to the point. Fermi created such coordinates in a neighborhood of a geodesic with r denoting the "intrinsic" distance to the geodesic. The terminology will be explained later.

(15) Find the Christoffel symbols of the first and second kind when the first fundamental form is given by

$$[\mathrm{I}] = \begin{bmatrix} r^2 & 0 \\ 0 & r^2 \end{bmatrix},$$

where $r = r(u) > 0$.

(16) (Isothermal coordinates) Find the Christoffel symbols of the first and second kind when the first fundamental form is given by

$$[\mathrm{I}] = \begin{bmatrix} r^2 & 0 \\ 0 & r^2 \end{bmatrix},$$

where $r = r(u, v) > 0$.

(17) (Liouville surfaces) Find the Christoffel symbols of the first and second kind when the first fundamental form is given by

$$[\mathrm{I}] = \begin{bmatrix} r^2 & 0 \\ 0 & r^2 \end{bmatrix},$$

where $r^2 = f(u) + g(v) > 0$.

(18) (Monge patch) Find the Christoffel symbols of the first and second kind when the first fundamental form is given by

$$[\mathrm{I}] = \begin{bmatrix} 1 + p^2 & pq \\ pq & 1 + q^2 \end{bmatrix},$$

where $p = \frac{\partial F}{\partial u}$, $q = \frac{\partial F}{\partial v}$ and $F = F(u, v)$.

5.3. The Gauss and Mean Curvatures

We are now finally ready to define the curvatures of a surface. Historically the principal curvatures defined in section 5.4 came first, but it seems equally natural to start with the Gauss and mean curvatures.

Definition 5.3.1. The *Gauss curvature* is defined as the determinant of the Weingarten map

$$K = \det L.$$

The *mean curvature* is related to the trace as follows

$$H = \frac{1}{2}\operatorname{tr} L.$$

To calculate these quantities we have:

Proposition 5.3.2. *The Gauss and mean curvatures of a parametrized surface* $\mathrm{q}(u,v)$ *can be computed as*

$$K = \frac{\det[\mathrm{II}]}{\det[\mathrm{I}]} = \frac{L_{uu}L_{vv} - (L_{uv})^2}{g_{uu}g_{vv} - (g_{uv})^2}$$

and

$$H = \frac{1}{2}\frac{g_{vv}L_{uu} + g_{uu}L_{vv} - 2g_{uv}L_{uv}}{g_{uu}g_{vv} - (g_{uv})^2}.$$

Proof. To calculate the Gauss and mean curvatures we use the formulas for determinant and trace of a matrix representation:

$$K = \det[L] = L_u^u L_v^v - L_u^v L_v^u,$$

$$H = \frac{1}{2}\operatorname{tr}[L] = \frac{1}{2}\left(L_u^u + L_v^v\right),$$

and $[L] = [\mathrm{I}]^{-1}[\mathrm{II}]$ (see proposition 5.2.8). The formula for K now follows from standard determinant rules.

For H we use that

$$L_u^u = g^{uu}L_{uu} + g^{uv}L_{vu} \text{ and } L_v^v = g^{vu}L_{uv} + g^{vv}L_{vv}$$

together with

$$\begin{bmatrix} g^{uu} & g^{uv} \\ g^{vu} & g^{vv} \end{bmatrix} = \frac{1}{\det[\mathrm{I}]}\begin{bmatrix} g_{vv} & -g_{uv} \\ -g_{vu} & g_{uu} \end{bmatrix}$$

to get the desired formula. □

Example 5.3.3. For a sphere of radius R we have that $\mathrm{II} = \pm\frac{1}{R}\mathrm{I}$. Thus $K = \frac{1}{R^2}$ and $H = \pm\frac{1}{R}$. For a plane $\mathrm{II} = 0$ and $K = H = 0$.

Example 5.3.4 (Examples 4.3.3 and 5.1.3 continued). We can now calculate the Gauss and mean curvatures of the generalized helicoids

$$K = \frac{\det[\mathrm{II}]}{\det[\mathrm{I}]} = \frac{u^3 f' f'' - c^2}{\left(c^2 + u^2\left(1 + (f')^2\right)\right)^2},$$

$$H = \frac{1}{2}\frac{\left(1 + (f')^2\right)u^2 f' + \left(u^2 + c^2\right)u f'' + 2c^2 f'}{\left(c^2 + u^2\left(1 + (f')^2\right)\right)^{\frac{3}{2}}}.$$

REMARK 5.3.5. It is often convenient to calculate the mean curvature using the formula

$$H = \frac{1}{2}\frac{\left(g_{vv}\frac{\partial^2 \mathrm{q}}{\partial u^2} + g_{uu}\frac{\partial^2 \mathrm{q}}{\partial v^2} - 2g_{uv}\frac{\partial^2 \mathrm{q}}{\partial u \partial v}\right)\cdot \mathrm{n}}{g_{uu}g_{vv} - (g_{uv})^2},$$

which follows directly from the above proposition and the definition of the entries in the second fundamental form.

We can now significantly improve theorem 5.1.8 as was first done by Gauss. This result is also a corollary of the next theorem.

THEOREM 5.3.6 (Gauss). *If a surface in $\mathbb{R}^3$ admits Cartesian coordinates, then the Gauss curvature vanishes.*

PROOF. We saw in the proof of proposition 5.1.7 that when a parametrization of a surface $\mathrm{q}(u,v)$ is Cartesian, then the second derivatives $\frac{\partial^2 \mathrm{q}}{\partial u \partial v}$, $\frac{\partial^2 \mathrm{q}}{\partial u^2}$, and $\frac{\partial^2 \mathrm{q}}{\partial v^2}$ are all normal to the surface. This shows with explanations below that

$$\begin{aligned}
\det[\mathrm{II}] &= \left(\frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot \mathrm{n}\right)\left(\frac{\partial^2 \mathrm{q}}{\partial v^2}\cdot \mathrm{n}\right) - \left(\frac{\partial^2 \mathrm{q}}{\partial u \partial v}\cdot \mathrm{n}\right)^2 \\
&= \frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial^2 \mathrm{q}}{\partial v^2} - \frac{\partial^2 \mathrm{q}}{\partial u \partial v}\cdot\frac{\partial^2 \mathrm{q}}{\partial u \partial v} \\
&= \frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial^2 \mathrm{q}}{\partial v^2} + \frac{\partial^3 \mathrm{q}}{\partial u^2 \partial v}\cdot\frac{\partial \mathrm{q}}{\partial v} \\
&= \frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial^2 \mathrm{q}}{\partial v^2} + \frac{\partial^3 \mathrm{q}}{\partial v \partial u^2}\cdot\frac{\partial \mathrm{q}}{\partial v} \\
&= \frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial^2 \mathrm{q}}{\partial v^2} - \frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial^2 \mathrm{q}}{\partial v^2} \\
&= 0,
\end{aligned}$$

where lines 3 and 5 follow from

$$0 = \frac{\partial}{\partial u}\left(\frac{\partial^2 \mathrm{q}}{\partial u \partial v}\cdot\frac{\partial \mathrm{q}}{\partial v}\right) = \frac{\partial^3 \mathrm{q}}{\partial u^2 \partial v}\cdot\frac{\partial \mathrm{q}}{\partial v} + \frac{\partial^2 \mathrm{q}}{\partial u \partial v}\cdot\frac{\partial^2 \mathrm{q}}{\partial u \partial v}$$

and

$$0 = \frac{\partial}{\partial v}\left(\frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial \mathrm{q}}{\partial v}\right) = \frac{\partial^3 \mathrm{q}}{\partial v \partial u^2}\cdot\frac{\partial \mathrm{q}}{\partial v} + \frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial^2 \mathrm{q}}{\partial v^2}.$$

This shows that the Gauss curvature vanishes. □

The converse is also true and is covered in section 6.3 exercise 9 and theorem 7.7.1. Section 5.5 contains a more detailed discussion of what surfaces with zero Gauss curvature look like.

REMARK 5.3.7. Given that planes, generalized cylinders, and generalized cones all admit Cartesian coordinates it is easy to come up with examples showing that the mean curvature cannot be calculated from the first fundamental form. In fact only planes have the property that the Gauss and mean curvatures both vanish.

We can use our knowledge of Christoffel symbols to improve the theorem for Cartesian coordinates to the general result that the Gauss curvature can always be computed from the first fundamental form. Given the definition of K this is certainly a big surprise.

THEOREM 5.3.8 (*Theorema Egregium, Gauss, 1827*). *The Gauss curvature can be computed knowing only the first fundamental form.*

PROOF. Assume that we have a parametrized surface $\mathrm{q}(u,v)$. The calculations are similar to what we just did for a Cartesian parametrization. First we observe that it suffices to show that $\det[\mathrm{II}]$ can be calculated from the first fundamental form since

$$\begin{aligned} K &= \det L = \det[\mathrm{I}]^{-1}\det[\mathrm{II}], \\ \det[\mathrm{I}] &= g_{uu}g_{vv} - (g_{uv})^2. \end{aligned}$$

We use the Gauss formulas to compute

$$\begin{aligned} \det[\mathrm{II}] &= \det\begin{bmatrix} L_{uu} & L_{uv} \\ L_{vu} & L_{vv} \end{bmatrix} \\ &= \det\begin{bmatrix} \frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot \mathrm{n} & \frac{\partial^2 \mathrm{q}}{\partial u\partial v}\cdot \mathrm{n} \\ \frac{\partial^2 \mathrm{q}}{\partial v\partial u}\cdot \mathrm{n} & \frac{\partial^2 \mathrm{q}}{\partial v^2}\cdot \mathrm{n} \end{bmatrix} \\ &= \left(\frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot \mathrm{n}\right)\left(\frac{\partial^2 \mathrm{q}}{\partial v^2}\cdot \mathrm{n}\right) - \left(\frac{\partial^2 \mathrm{q}}{\partial u\partial v}\cdot \mathrm{n}\right)\left(\frac{\partial^2 \mathrm{q}}{\partial u\partial v}\cdot \mathrm{n}\right) \\ &= \frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial^2 \mathrm{q}}{\partial v^2} - \left(\Gamma_{uu}^u\frac{\partial \mathrm{q}}{\partial u} + \Gamma_{uu}^v\frac{\partial \mathrm{q}}{\partial v}\right)\cdot\left(\Gamma_{vv}^u\frac{\partial \mathrm{q}}{\partial u} + \Gamma_{vv}^v\frac{\partial \mathrm{q}}{\partial v}\right) \\ &\quad -\frac{\partial^2 \mathrm{q}}{\partial u\partial v}\cdot\frac{\partial^2 \mathrm{q}}{\partial u\partial v} + \left|\Gamma_{uv}^u\frac{\partial \mathrm{q}}{\partial u} + \Gamma_{uv}^v\frac{\partial \mathrm{q}}{\partial v}\right|^2 \end{aligned}$$

Here the inner products

$$\left(\Gamma_{uu}^u\frac{\partial \mathrm{q}}{\partial u} + \Gamma_{uu}^v\frac{\partial \mathrm{q}}{\partial v}\right)\cdot\left(\Gamma_{vv}^u\frac{\partial \mathrm{q}}{\partial u} + \Gamma_{vv}^v\frac{\partial \mathrm{q}}{\partial v}\right)$$

and

$$\left|\Gamma_{uv}^u\frac{\partial \mathrm{q}}{\partial u} + \Gamma_{uv}^v\frac{\partial \mathrm{q}}{\partial v}\right|^2$$

can be calculated from the first fundamental form as we proved in proposition 5.2.11 that the inner products

$$\Gamma_{w_1w_2w} = \frac{\partial \mathrm{q}}{\partial w}\cdot\frac{\partial^2 \mathrm{q}}{\partial w_1\partial w_2}$$

have formulas that only use the derivatives of g_{uu}, g_{uv}, and g_{vv} (see also section 5.2 exercise 11).

To finish the proof it simply remains to observe that

$$\begin{aligned} &\frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial^2 \mathrm{q}}{\partial v^2} - \frac{\partial^2 \mathrm{q}}{\partial u\partial v}\cdot\frac{\partial^2 \mathrm{q}}{\partial u\partial v} \\ &= \frac{\partial}{\partial v}\left(\frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\frac{\partial \mathrm{q}}{\partial v}\right) - \frac{\partial^3 \mathrm{q}}{\partial v\partial u^2}\cdot\frac{\partial \mathrm{q}}{\partial v} \\ &\quad -\frac{\partial}{\partial u}\left(\frac{\partial^2 \mathrm{q}}{\partial u\partial v}\cdot\frac{\partial \mathrm{q}}{\partial v}\right) + \frac{\partial^3 \mathrm{q}}{\partial^2 u\partial v}\cdot\frac{\partial \mathrm{q}}{\partial v} \\ &= \frac{\partial}{\partial v}\Gamma_{uuv} - \frac{\partial}{\partial u}\Gamma_{vuv}. \end{aligned}$$

(See also section 5.2 exercise 12 for a nice formula of this combination of derivatives.) The complete formula for the Gauss curvature in terms of the first fundamental form and the Christoffel symbols of the first kind is given in exercise 13 to this section. □

EXAMPLE 5.3.9. Assume that $g_{uu} = 1$ and $g_{uv} = 0$ as in the end of example 5.2.12. In this case the above proof reduces the Gauss curvature to:

$$\begin{aligned} K &= \frac{1}{g_{vv}}\left(-\frac{\partial}{\partial u}\Gamma_{vuv} + \left(\Gamma_{uv}^{v}\right)^2 g_{vv}\right) \\ &= \frac{1}{g_{vv}}\left(-\frac{1}{2}\frac{\partial^2 g_{vv}}{\partial u^2} + \left(\frac{1}{g_{vv}}\frac{1}{2}\frac{\partial g_{vv}}{\partial u}\right)^2 g_{vv}\right) \\ &= -\frac{1}{2}\frac{1}{g_{vv}}\frac{\partial^2 g_{vv}}{\partial u^2} + \frac{1}{4}\left(\frac{1}{g_{vv}}\frac{\partial g_{vv}}{\partial u}\right)^2 \\ &= -\frac{1}{\sqrt{g_{vv}}}\frac{\partial^2\sqrt{g_{vv}}}{\partial u^2}. \end{aligned}$$

The Gauss curvature can also be expressed more directly in terms of the unit normal.

PROPOSITION 5.3.10 (Gauss). *The Gauss curvature satisfies*

$$K = \frac{\left(\frac{\partial \mathrm{n}}{\partial u} \times \frac{\partial \mathrm{n}}{\partial v}\right)\cdot \mathrm{n}}{\left(\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right)\cdot \mathrm{n}}.$$

PROOF. Simply use the Weingarten equations to calculate

$$\begin{aligned} \frac{\partial \mathrm{n}}{\partial u} \times \frac{\partial \mathrm{n}}{\partial v} &= \left(-L_u^u\frac{\partial \mathrm{q}}{\partial u} - L_u^v\frac{\partial \mathrm{q}}{\partial v}\right) \times \left(-L_v^u\frac{\partial \mathrm{q}}{\partial u} - L_v^v\frac{\partial \mathrm{q}}{\partial v}\right) \\ &= L_u^u L_v^v \frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v} + L_u^v L_v^u \frac{\partial \mathrm{q}}{\partial v} \times \frac{\partial \mathrm{q}}{\partial u} \\ &= \left(L_u^u L_v^v - L_u^v L_v^u\right)\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v} \\ &= K\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}. \end{aligned}$$

□

Note that the denominator in

$$K = \frac{\left(\frac{\partial \mathrm{n}}{\partial u} \times \frac{\partial \mathrm{n}}{\partial v}\right)\cdot \mathrm{n}}{\left(\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right)\cdot \mathrm{n}}$$

is already computed in terms of the first fundamental form

$$\left(\left(\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right)\cdot \mathrm{n}\right)^2 = \left|\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right|^2 = g_{uu}g_{vv} - \left(g_{uv}\right)^2.$$

The numerator is the signed volume of the parallelepiped $\frac{\partial \mathrm{n}}{\partial u}, \frac{\partial \mathrm{n}}{\partial v}, \mathrm{n}$ corresponding to the Gauss map $\mathrm{n}(u,v) : U \to S^2(1) \subset \mathbb{R}^3$ of the surface. Thus it can be computed from the first fundamental form of $\mathrm{n}(u,v)$. However, there is a sign that depends on whether n and $\frac{\partial \mathrm{n}}{\partial u} \times \frac{\partial \mathrm{n}}{\partial v}$ point in the same direction or not. Recall from curve theory that the tangent spherical image was also related to curvature in a similar way. Here the formulas are a bit more complicated as we use arbitrary parameters.

DEFINITION 5.3.11. The *third fundamental form* III on T_pM is defined as the first fundamental form for $S^2(1)$ on $T_{\mathrm{n}(p)}S^2(1)$. If we use the Gauss map $\mathrm{n}(u,v) = \mathrm{n}\circ\mathrm{q}(u,v)$ as the parametrization, then the matrix representation is given by

$$\begin{aligned}[\mathrm{III}] &= \begin{bmatrix} \frac{\partial\,\mathrm{n}}{\partial u} & \frac{\partial\,\mathrm{n}}{\partial v}\end{bmatrix}^t \begin{bmatrix} \frac{\partial\,\mathrm{n}}{\partial u} & \frac{\partial\,\mathrm{n}}{\partial v}\end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial\,\mathrm{n}}{\partial u}\cdot\frac{\partial\,\mathrm{n}}{\partial u} & \frac{\partial\,\mathrm{n}}{\partial u}\cdot\frac{\partial\,\mathrm{n}}{\partial v} \\ \frac{\partial\,\mathrm{n}}{\partial v}\cdot\frac{\partial\,\mathrm{n}}{\partial u} & \frac{\partial\,\mathrm{n}}{\partial v}\cdot\frac{\partial\,\mathrm{n}}{\partial v}\end{bmatrix}.\end{aligned}$$

This always defines a quadratic from, but n might not be a genuine parametrization if the Gauss curvature vanishes. Nevertheless, we alway have the relationship

$$\frac{\partial\,\mathrm{n}}{\partial u}\times\frac{\partial\,\mathrm{n}}{\partial v} = K\left(\frac{\partial\,\mathrm{q}}{\partial u}\times\frac{\partial\,\mathrm{q}}{\partial v}\right).$$

The three fundamental forms and two curvatures are related by a very interesting formula which also shows that the third fundamental form is almost redundant.

THEOREM 5.3.12. *All three fundamental forms are related by*

$$\mathrm{III} - 2H\,\mathrm{II} + K\,\mathrm{I} = 0.$$

PROOF. We prove this for the matrix representations

$$[\mathrm{III}] - 2H\,[\mathrm{II}] + K\,[\mathrm{I}] = 0$$

by reducing it to the Cayley-Hamilton theorem for $[L]$

$$[L]^2 - (\operatorname{tr} L)\,[L] + (\det L)\,I_2 = 0,$$

where I_2 is the 2×2 identity matrix. This relies on showing: $[\mathrm{I}]\,[L] = [\mathrm{II}]$ and $[\mathrm{I}]\,[L]^2 = [\mathrm{III}]$. The first identity has already been established. The second likewise follows from the Weingarten equations:

$$\begin{aligned}[\mathrm{III}] &= \begin{bmatrix} L\left(\frac{\partial\,\mathrm{q}}{\partial u}\right) & L\left(\frac{\partial\,\mathrm{q}}{\partial v}\right)\end{bmatrix}^t \begin{bmatrix} L\left(\frac{\partial\,\mathrm{q}}{\partial u}\right) & L\left(\frac{\partial\,\mathrm{q}}{\partial v}\right)\end{bmatrix} \\ &= [L]^t \begin{bmatrix} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix}^t \begin{bmatrix} \frac{\partial\,\mathrm{q}}{\partial u} & \frac{\partial\,\mathrm{q}}{\partial v}\end{bmatrix} [L] \\ &= [L]^t\,[\mathrm{I}]\,[L] \\ &= [\mathrm{II}]^t\left([\mathrm{I}]^{-1}\right)^t [\mathrm{I}]\,[L] \\ &= [\mathrm{II}]\,[\mathrm{I}]^{-1}\,[\mathrm{I}]\,[L] \\ &= [\mathrm{II}]\,[L] \\ &= [\mathrm{I}]\,[L]\,[L] \\ &= [\mathrm{I}]\,[L]^2.\end{aligned}$$

Finally, if $[L] = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$, then the Cayley-Hamilton theorem follows by a direct calculation:

$$\begin{aligned} & [L]^2 - (\operatorname{tr} L)\,[L] + (\det L)\, I_2 \\ = \;& \begin{bmatrix} a & c \\ b & d \end{bmatrix}^2 - (a+d) \begin{bmatrix} a & c \\ b & d \end{bmatrix} + (ad - bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ = \;& \begin{bmatrix} a^2 + bc & ac + dc \\ ab + db & bc + d^2 \end{bmatrix} - (a+d) \begin{bmatrix} a & c \\ b & d \end{bmatrix} + (ad - bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ = \;& \begin{bmatrix} bc - ad & 0 \\ 0 & bc - ad \end{bmatrix} + (ad - bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ = \;& 0. \end{aligned}$$

□

Definition 5.3.13. A surface is called *minimal* if its mean curvature vanishes.

Proposition 5.3.14. *A minimal surface has conformal Gauss map.*

Proof. Let $\mathrm{q}\,(u, v)$ be a parametrization of the surface, then $\mathrm{n}\,(u, v)$ is a potential parametrization of the unit sphere via the Gauss map. The first fundamental form with respect to this parametrization is the third fundamental form. Using $H = 0$ we obtain

$$[\mathrm{III}] + K\,[\mathrm{I}] = 0,$$

which implies that the Gauss map is conformal. □

Example 5.3.15. Note that the Gauss map for the unit sphere centered at the origin is simply the identity map on the sphere. Thus its Gauss map is an isometry and in particular conformal. However, the sphere is not a minimal surface. More generally, the Gauss map

$$\mathrm{n}\,(\mathrm{q}) = \pm \frac{\mathrm{q} - \mathrm{c}}{R}$$

for a sphere of radius R centered at c is also conformal as its derivative is given by $D\,\mathrm{n} = \pm \frac{1}{R} I$, where I is the identity map/matrix.

The name for minimal surfaces is justified by the next result. Meusnier in 1785 was the first to consider such surfaces and he also indicated with a geometric argument that their areas should be minimal. In fact Lagrange in 1761 had already come up with an (Euler-Lagrange) equation for surfaces that minimize area, but it was not until the mid 19th century with Bonnet and Beltrami that this was definitively connected to the condition that the mean curvature should vanish.

Example 5.3.16 (Example 5.3.4 continued). We note that the generalized helicoids are minimal when $f' = 0$, i.e., when they are regular helicoids. In the rotationally symmetric case where $c = 0$, they are minimal when $f = a \cosh^{-1}\left(\frac{u}{a}\right) + b$, for constants a, b. These are all catenoids.

Finally, we also obtain a more complicated family by using

$$f' = \frac{c}{u} \sqrt{\frac{u^2 + c^2}{u^2 - c^2}}.$$

To see this first note that

$$1+(f')^2=\frac{u^4+c^4}{u^2\left(u^2-c^2\right)},\ f''=-f'\left(\frac{1}{u}+\frac{2c^2u}{\left(u^2-c^2\right)\left(u^2+c^2\right)}\right).$$

The numerator in the formula for H then becomes

$$\frac{u^4+c^4}{u^2\left(u^2-c^2\right)}u^2f'-\left(u^2+c^2\right)uf'\left(\frac{1}{u}+\frac{2c^2u}{\left(u^2-c^2\right)\left(u^2+c^2\right)}\right)+2c^2f'$$

after eliminating f' and multiplying through by u^2-c^2 this expression becomes

$$u^4+c^4-\left(u^4-c^4+2c^2u^2\right)+2c^2\left(u^2-c^2\right)=0.$$

THEOREM 5.3.17. *A surface whose area is minimal among nearby surfaces is a minimal surface.*

PROOF. We assume that the surface is given by a parametrization $\mathrm{q}(u,v)$ and only consider nearby surfaces that are graphs over the given surface, i.e.,

$$\mathrm{q}^*=\mathrm{q}+\phi\,\mathrm{n}$$

for some function $\phi(u,v)$. From such a surface we can then create a family of surfaces

$$\mathrm{q}_\epsilon=\mathrm{q}+\epsilon\phi\,\mathrm{n}$$

that interpolates between these two surfaces. To calculate the area density as a function of ϵ we first note that

$$\frac{\partial\,\mathrm{q}_\epsilon}{\partial w}=\frac{\partial\,\mathrm{q}}{\partial w}+\epsilon\left(\frac{\partial\phi}{\partial w}\mathrm{n}+\phi\frac{\partial\,\mathrm{n}}{\partial w}\right).$$

Then the first fundamental form becomes

$$\begin{aligned}
g^\epsilon_{ww} &= g_{ww}+2\epsilon\phi\frac{\partial\,\mathrm{q}}{\partial w}\cdot\frac{\partial\,\mathrm{n}}{\partial w}+\epsilon^2\left(\left(\frac{\partial\phi}{\partial w}\right)^2+\phi^2\left|\frac{\partial\,\mathrm{n}}{\partial w}\right|^2\right)\\
&= g_{ww}-2\epsilon\phi L_{ww}+\epsilon^2\left(\left(\frac{\partial\phi}{\partial w}\right)^2-K\phi^2g_{ww}+2H\phi^2L_{ww}\right),\\
g^\epsilon_{uv} &= g_{uv}+\epsilon\phi\left(\frac{\partial\,\mathrm{q}}{\partial u}\cdot\frac{\partial\,\mathrm{n}}{\partial v}+\frac{\partial\,\mathrm{q}}{\partial v}\cdot\frac{\partial\,\mathrm{n}}{\partial u}\right)+\epsilon^2\left(\frac{\partial\phi}{\partial u}\cdot\frac{\partial\phi}{\partial v}+\phi^2\frac{\partial\,\mathrm{n}}{\partial u}\cdot\frac{\partial\,\mathrm{n}}{\partial v}\right)\\
&= g_{uv}-2\epsilon\phi L_{uv}+\epsilon^2\left(\frac{\partial\phi}{\partial u}\cdot\frac{\partial\phi}{\partial v}-K\phi^2g_{uv}+2H\phi^2L_{uv}\right),
\end{aligned}$$

and the square of the area density

$$\begin{aligned}
g^\epsilon_{uu}g^\epsilon_{vv}-\left(g^\epsilon_{uv}\right)^2 &= g_{uu}g_{vv}-\left(g_{uv}\right)^2-2\epsilon\left(g_{uu}L_{vv}+g_{vv}L_{uu}-2g_{uv}L_{uv}\right)+O\left(\epsilon^2\right)\\
&= \left(g_{uu}g_{vv}-\left(g_{uv}\right)^2\right)(1-4\epsilon\phi H)+O\left(\epsilon^2\right).
\end{aligned}$$

This shows that if $H\neq 0$ somewhere then we can select ϕ such that the area density will decrease for nearby surfaces. □

REMARK 5.3.18. Conversely note that when $H = 0$ everywhere, then the area density is critical. The term that involves ϵ^2 has a coefficient that looks like

$$\begin{aligned}
& g_{uu}\left(\left(\frac{\partial\phi}{\partial v}\right)^2 - \phi^2 K g_{vv}\right) + g_{vv}\left(\left(\frac{\partial\phi}{\partial u}\right)^2 - \phi^2 K g_{uu}\right) \\
& -2g_{uv}\left(\frac{\partial\phi}{\partial u}\cdot\frac{\partial\phi}{\partial v} - \phi^2 K g_{uv}\right) + 4\phi^2\left(L_{uu}L_{vv} - L_{uv}^2\right) \\
= & \left|-\frac{\partial\phi}{\partial v}\frac{\partial\,\mathrm{q}}{\partial u} + \frac{\partial\,\mathrm{q}}{\partial v}\frac{\partial\phi}{\partial u}\right|^2 - 2\phi^2 K\left(g_{uu}g_{vv} - g_{uv}^2\right) + 4\phi^2\left(L_{uu}L_{vv} - L_{uv}^2\right) \\
= & \left|-\frac{\partial\phi}{\partial v}\frac{\partial\,\mathrm{q}}{\partial u} + \frac{\partial\,\mathrm{q}}{\partial v}\frac{\partial\phi}{\partial u}\right|^2 + 2\phi^2 K\left(g_{uu}g_{vv} - g_{uv}^2\right)
\end{aligned}$$

and it is not clear that this is positive. In fact, minimal surfaces have $K \leq 0$ so when ϕ is constant the area decreases!

Exercises

(1) Let $X, Y \in T_pM$ be an orthonormal basis for the tangent space at p to the surface M. Prove that the mean and Gauss curvatures can be computed as follows:

$$\begin{aligned}
H &= \frac{1}{2}\left(\mathrm{II}(X,X) + \mathrm{II}(Y,Y)\right), \\
K &= \mathrm{II}(X,X)\,\mathrm{II}(Y,Y) - \left(\mathrm{II}(X,Y)\right)^2.
\end{aligned}$$

(2) Show that if $K = 0$ and $H = 0$, then the Weingarten map $L = 0$ and the normal is constant. Hint: First show that the third fundamental form vanishes. Give an example of a 2×2 matrix $A \neq 0$ such that $A^2 = 0$ and $\operatorname{tr} A = 0 = \det A$.

(3) Assume that $K = H = 0$ and use the equations

$$\begin{aligned}
L_{uu}L_{vv} - \left(L_{uv}\right)^2 &= 0, \\
g_{vv}L_{uu} + g_{uu}L_{vv} - 2g_{uv}L_{uv} &= 0,
\end{aligned}$$

to show that $\mathrm{II} = 0$. Hint: First show that if $L_{uv} = 0$, then $L_{uu} = L_{vv} = 0$. Second, if $L_{uv} \neq 0$, then use $g_{uu}g_{vv} > \left(g_{uv}\right)^2$ and $L_{uu}L_{vv} = \left(L_{uv}\right)^2$ to show that the last equation can't be satisfied.

(4) Show that if $K = \frac{1}{R^2}$ and $H = \pm\frac{1}{R}$, then the Weingarten map $L = \mp\frac{1}{R}I$, where I is the identity operator. Use this to show that the surface is part of a sphere on radius R.

(5) For a surface of revolution

$$\mathrm{q}(t,\mu) = \left(r(t)\cos\mu, r(t)\sin\mu, z(t)\right)$$

compute the first and second fundamental forms as well as the Gauss and mean curvatures. Show that if $\left(r(t), z(t)\right)$ is unit speed, then $K = -\frac{\ddot{r}}{r}$.

(6) Compute the second fundamental form of a tangent developable $\mathrm{q}(s,t) = c(t) + s\frac{dc}{dt}$ of a unit speed curve $c(t)$. Show that the Gauss curvature vanishes. Show that the mean curvature vanishes if and only if the second fundamental form vanishes.

(7) Show that the surface of revolution

$$\mathrm{q}(s,\theta)=\left(R\cos(as)\cos\theta,R\cos(as)\sin\theta,\int_0^s\sqrt{1-a^2R^2\sin^2(at)}dt\right)$$

has constant Gauss curvature a^2. Show that this is a sphere centered at the origin if and only if $R=\frac{1}{a}$. Hint: When $s=0$ this is a circle of radius R.

(8) Show that a Monge patch $z=F(x,y)$ is minimal if and only if Lagrange's equation holds:

$$\left(1+\left(\frac{\partial F}{\partial y}\right)^2\right)\frac{\partial^2 F}{\partial x^2}-2\frac{\partial F}{\partial x}\frac{\partial F}{\partial y}\frac{\partial^2 F}{\partial x\partial y}+\left(1+\left(\frac{\partial F}{\partial x}\right)^2\right)\frac{\partial^2 F}{\partial y^2}=0.$$

Use this to show that Scherk's surface $e^{cz}\cos cx=\cos cy$ is minimal. In fact Scherk's surface is the only minimal surface of the form $z=F(x,y)=f(x)+h(y)$.

(9) Let $\mathrm{q}(t)$ be a curve on a surface with normal n. Denote the Gauss image of the curve by $\mathrm{n}(t)=\mathrm{n}\circ\mathrm{q}(t)$. Show that the velocities of these curves are related by

$$\left|\frac{d\,\mathrm{n}}{dt}\right|^2+2H\frac{d\,\mathrm{n}}{dt}\cdot\frac{d\,\mathrm{q}}{dt}+K\left|\frac{d\,\mathrm{q}}{dt}\right|^2=0.$$

(10) Let $\mathrm{q}(t)=\mathrm{q}(u(t),v(t))$ be an asymptotic curve on a surface, i.e., $\kappa_n=0$.
  (a) Show that $K\leq 0$ along the curve.
  (b) (Beltrami-Enneper) If τ is the torsion of the curve as a space curve, then

$$\tau^2=-K.$$

  Hint: Use the previous exercise.

(11) Show that a minimal surface satisfies $K\leq 0$.

(12) Show that if a parametrized surface has the property that g_{uu}, g_{vv}, and g_{uv} are constant, then the second derivatives $\frac{\partial^2\mathrm{q}}{\partial u\partial v}$, $\frac{\partial^2\mathrm{q}}{\partial u^2}$, and $\frac{\partial^2\mathrm{q}}{\partial v^2}$ are all normal to the surface. Use this to conclude that the Gauss curvature vanishes.

(13) Show that

$$\begin{aligned}K &= \frac{1}{\det[\mathrm{I}]}\left(\frac{\partial}{\partial v}\Gamma_{uuv}-\frac{\partial}{\partial u}\Gamma_{uvv}\right)\\ &\quad-\frac{1}{\det[\mathrm{I}]}\begin{bmatrix}\Gamma_{uuu} & \Gamma_{uuv}\end{bmatrix}[\mathrm{I}]^{-1}\begin{bmatrix}\Gamma_{vvu}\\ \Gamma_{vvv}\end{bmatrix}\\ &\quad+\frac{1}{\det[\mathrm{I}]}\begin{bmatrix}\Gamma_{uvu} & \Gamma_{uvv}\end{bmatrix}[\mathrm{I}]^{-1}\begin{bmatrix}\Gamma_{uvu}\\ \Gamma_{uvv}\end{bmatrix}.\end{aligned}$$

Hint: See section 5.2 exercise 11.

(14) Compute the Gauss curvatures of the generalized cones (section 4.1 exercise 2), generalized cylinders (section 4.1 exercise 1), and tangent developables (section 4.1 exercise 3). We shall offer several proofs below that these are essentially the only surfaces with vanishing Gauss curvature (see 29 below and section 5.5). Hint: In each case the normal vector is constant along the lines in the ruled surface.

(15) Show that

$$\frac{\partial\,\mathrm{q}}{\partial u}\times\frac{\partial\,\mathrm{n}}{\partial v}+\frac{\partial\,\mathrm{n}}{\partial u}\times\frac{\partial\,\mathrm{q}}{\partial v}=-2H\frac{\partial\,\mathrm{q}}{\partial u}\times\frac{\partial\,\mathrm{q}}{\partial v},$$

and more generally that

$$\begin{aligned}
\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{n}}{\partial w} &= -L_w^v \frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}, \\
\frac{\partial \mathrm{n}}{\partial w} \times \frac{\partial \mathrm{q}}{\partial v} &= -L_w^u \frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}.
\end{aligned}$$

(16) Show that

$$\begin{aligned}
2K\sqrt{\det[\mathrm{I}]}\,\mathrm{n} &= \frac{\partial}{\partial u}\left(n \times \frac{\partial \mathrm{n}}{\partial v}\right) - \frac{\partial}{\partial v}\left(n \times \frac{\partial \mathrm{n}}{\partial u}\right), \\
\sqrt{\det[\mathrm{I}]}\,\mathrm{n} \times \frac{\partial \mathrm{n}}{\partial u} &= L_{uv}\frac{\partial \mathrm{q}}{\partial u} - L_{uu}\frac{\partial \mathrm{q}}{\partial v}, \\
\sqrt{\det[\mathrm{I}]}\,\mathrm{n} \times \frac{\partial \mathrm{n}}{\partial v} &= L_{vv}\frac{\partial \mathrm{q}}{\partial u} - L_{uv}\frac{\partial \mathrm{q}}{\partial v}.
\end{aligned}$$

Hint: For the last two formulas it might be useful to use section 4.3 exercise 9 and $[\mathrm{II}] = [\mathrm{I}]\,[L]$.

(17) Compute the first and second fundamental forms as well as the Gauss and mean curvatures for the conoid

$$\begin{aligned}
\mathrm{q}(s,t) &= (sx(t), sy(t), z(t)) \\
&= (0,0,z(t)) + s(x(t), y(t), 0)
\end{aligned}$$

when $X = (x(t), y(t), 0)$ is a unit field.

(18) Show that a conformally parametrized (isothermal) surface $\mathrm{q}(u,v)$ is minimal if and only if

$$\Delta\,\mathrm{q} = \frac{\partial^2 \mathrm{q}}{\partial u^2} + \frac{\partial^2 \mathrm{q}}{\partial v^2} = 0.$$

Hint: Use section 5.1 exercise 5c.

(19) Show that Enneper's surface

$$\mathrm{q}(u,v) = \left(u - \frac{1}{3}u^3 + uv^2, v - \frac{1}{3}v^3 + vu^2, u^2 - v^2\right)$$

is minimal. Hint: Do section 4.4 exercise 14 first.

(20) Show that Catalan's surface

$$\mathrm{q}(u,v) = \left(u - \sin u \cosh v, 1 - \cos u \cosh v, -4\sin\frac{u}{2}\sinh\frac{v}{2}\right)$$

is minimal. Hint: Do section 4.4 exercise 15 first.

(21) Show that for a fixed $\theta \in \mathbb{R}$ the parametrized surface

$$\mathrm{q}(u,v) = (u\cos\theta \pm \sin u \cosh v, v \pm \cos\theta \cos u \sinh v, \pm\sin\theta\cos u\cosh v)$$

is isothermal and minimal.

(22) Consider a unit speed curve $c(s) : [0,L] \to \mathbb{R}^3$ with non-vanishing curvature and the tube of radius R around it

$$\mathrm{q}(s,\phi) = c(s) + R(\mathrm{N}_c \cos\phi + \mathrm{B}_c \sin\phi)$$

(see section 4.1 exercise 6 and section 4.3 exercise 7).

(a) Use the formula for n together with Gauss' formula for K to show that

$$K = \frac{-\kappa\cos\phi}{R(1 - \kappa R\cos\phi)}.$$

(b) Show that

$$\int_0^{2\pi}\int_0^L K\sqrt{\det[\mathrm{I}]}\,ds\,d\phi = 0$$

and

$$\int_0^{2\pi}\int_0^L |K|\sqrt{\det[\mathrm{I}]}\,ds\,d\phi = 4\int_0^L \kappa ds.$$

(23) Consider a surface with negative Gauss curvature.

(a) Show that locally it admits a parametrization $\mathrm{q}(s,t)$ where the parameter curves are asymptotic curves, i.e., the second fundamental form looks like

$$[\mathrm{II}] = \begin{bmatrix} 0 & L_{st} \\ L_{st} & 0 \end{bmatrix}.$$

(b) Show that this implies

$$[\mathrm{III}] = -K\begin{bmatrix} g_{ss} & -g_{st} \\ -g_{st} & g_{tt} \end{bmatrix}.$$

(24) (Meusnier, 1785) Consider a surface of revolution of the form

$$\mathrm{q}(t,\mu) = (r(t)\cos\mu, r(t)\sin\mu, t).$$

(a) Show that if the surface of revolution is minimal then

$$\frac{\ddot r}{\dot r^2+1} = \frac{1}{r}.$$

(b) Show that the catenoids

$$\mathrm{q}(t,\mu) = \left(\frac{1}{a}\cosh(at+b)\cos\mu, \frac{1}{a}\cosh(at+b)\sin\mu, t\right),$$

$a>0$ and $b\in\mathbb{R}$ are minimal.

(c) Show that the functions

$$r(t) = \frac{1}{a}\cosh(at+b)$$

solve the initial value problems:

$$\frac{\ddot r}{\dot r^2+1} = \frac{1}{r},\ r(0) = r_0 > 0,\ \dot r(0) = \dot r_0 \in \mathbb{R}.$$

(d) Conclude that the catenoids are the only surfaces of revolution that are minimal.

(25) (Meusnier, 1785) Show that the helicoid

$$\mathrm{q}(r,\theta) = (r\cos\theta, r\sin\theta, \theta)$$

is minimal. Conversely, show that if a conoid

$$\mathrm{q}(r,\theta) = (r\cos\theta, r\sin\theta, z(\theta))$$

is minimal, then $z = a\theta + b$, for constants a, b.

(26) Consider a parametrized surface $\mathrm{q}(u,v)$ with normal $\mathrm{n}(u,v)$ and let $f(u,v)$ be a function.

(a) Show that

$$K = \frac{\det\left(\frac{\partial(f\,\mathrm{n})}{\partial u}, \frac{\partial(f\,\mathrm{n})}{\partial v}, f\,\mathrm{n}\right)}{f^2\det\left(\frac{\partial\mathrm{q}}{\partial u}, \frac{\partial\mathrm{q}}{\partial v}, f\,\mathrm{n}\right)}.$$

(b) Show that

$$H=-\frac{1}{2}\frac{\det\left(\frac{\partial \mathrm{q}}{\partial u},\frac{\partial (f\,\mathrm{n})}{\partial v},f\,\mathrm{n}\right)+\det\left(\frac{\partial (f\,\mathrm{n})}{\partial u},\frac{\partial \mathrm{q}}{\partial v},f\,\mathrm{n}\right)}{f\det\left(\frac{\partial \mathrm{q}}{\partial u},\frac{\partial \mathrm{q}}{\partial v},f\,\mathrm{n}\right)}.$$

Hint: When $f=1$ this follows from 15.

(c) Show that if the surface satisfies $F(x,y,z)=C$, then there is a function f such that $\nabla F=f\,\mathrm{n}$.

(27) Consider a surface that satisfies the equation $F(x,y,z)=C$ and use $\mathrm{n}=\frac{\nabla F}{|\nabla F|}$. (In section 2.1 exercise 13 there is a similar problem for planar curves given by equations.)

(a) Assume that $\frac{\partial F}{\partial z}\neq 0$ and use x,y as parameters for a Monge patch. Show that

$$\frac{\partial \mathrm{q}}{\partial x}=\frac{1}{\frac{\partial F}{\partial z}}\begin{bmatrix}\frac{\partial F}{\partial z}\\ 0\\ -\frac{\partial F}{\partial x}\end{bmatrix},\ \frac{\partial \mathrm{q}}{\partial y}=\frac{1}{\frac{\partial F}{\partial z}}\begin{bmatrix}0\\ \frac{\partial F}{\partial z}\\ -\frac{\partial F}{\partial y}\end{bmatrix}$$

and

$$\frac{\partial \nabla F}{\partial x}=\begin{bmatrix}\frac{\partial^2 F}{\partial x^2}\\ \frac{\partial^2 F}{\partial x\partial y}\\ \frac{\partial^2 F}{\partial x\partial z}\end{bmatrix}-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}}\begin{bmatrix}\frac{\partial^2 F}{\partial z\partial x}\\ \frac{\partial^2 F}{\partial z\partial y}\\ \frac{\partial^2 F}{\partial z^2}\end{bmatrix},\ \frac{\partial \nabla F}{\partial y}=\begin{bmatrix}\frac{\partial^2 F}{\partial y\partial x}\\ \frac{\partial^2 F}{\partial y^2}\\ \frac{\partial^2 F}{\partial y\partial z}\end{bmatrix}-\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}\begin{bmatrix}\frac{\partial^2 F}{\partial z\partial x}\\ \frac{\partial^2 F}{\partial z\partial y}\\ \frac{\partial^2 F}{\partial z^2}\end{bmatrix}.$$

Hint: Keep in mind that $z=z(x,y)$ and that its derivatives can be calculated using implicit differentiation.

(b) Use (a) and exercise 26 to show that

$$K=-\frac{1}{|\nabla F|^4}\det\begin{bmatrix}\frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial y\partial x} & \frac{\partial^2 F}{\partial z\partial x} & \frac{\partial F}{\partial x}\\ \frac{\partial^2 F}{\partial x\partial y} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial z\partial y} & \frac{\partial F}{\partial y}\\ \frac{\partial^2 F}{\partial x\partial z} & \frac{\partial^2 F}{\partial y\partial z} & \frac{\partial^2 F}{\partial z^2} & \frac{\partial F}{\partial z}\\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} & \frac{\partial F}{\partial z} & 0\end{bmatrix}.$$

Hint: Use a Laplace expansion along the bottom row.

(c) Why is the formula in (b) valid at all points where $\nabla F\neq 0$?

(d) Show that the surfaces

$$\begin{aligned}\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}&=1,\\ \frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}&=-1\end{aligned}$$

have

$$K=\frac{1}{a^2b^2c^2}\frac{1}{\left(\frac{x^2}{a^4}+\frac{y^2}{b^4}+\frac{z^2}{c^4}\right)^2}.$$

(e) Show that the surface

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

has

$$K=-\frac{1}{a^2b^2c^2}\frac{1}{\left(\frac{x^2}{a^4}+\frac{y^2}{b^4}+\frac{z^2}{c^4}\right)^2}.$$

(f) Show that the surface

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=z,$$

has

$$K=\frac{1}{4a^2b^2}\frac{1}{\left(\frac{x^2}{a^4}+\frac{y^2}{b^4}+\frac{1}{4}\right)^2}.$$

(g) Show that the surface

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=z$$

has

$$K=-\frac{1}{4a^2b^2}\frac{1}{\left(\frac{x^2}{a^4}+\frac{y^2}{b^4}+\frac{1}{4}\right)^2}.$$

(28) Consider a surface that satisfies the equation $F(x,y,z)=C$ and use $\mathrm{n}=\frac{\nabla F}{|\nabla F|}$. (In section 2.1 exercise 14 a similar problem for planar curves given by equations.)

(a) Assume that $\frac{\partial F}{\partial z}\neq 0$ and use x,y as parameters for a Monge patch. Show that

$$\frac{\partial\,\mathrm{q}}{\partial x}=\frac{1}{\frac{\partial F}{\partial z}}\begin{bmatrix}\frac{\partial F}{\partial z}\\ 0\\ -\frac{\partial F}{\partial x}\end{bmatrix},\quad \frac{\partial\,\mathrm{q}}{\partial y}=\frac{1}{\frac{\partial F}{\partial z}}\begin{bmatrix}0\\ \frac{\partial F}{\partial z}\\ -\frac{\partial F}{\partial y}\end{bmatrix}$$

and

$$\frac{\partial\nabla F}{\partial x}=\begin{bmatrix}\frac{\partial^2F}{\partial x^2}\\ \frac{\partial^2F}{\partial x\partial y}\\ \frac{\partial^2F}{\partial x\partial z}\end{bmatrix}-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}}\begin{bmatrix}\frac{\partial^2F}{\partial z\partial x}\\ \frac{\partial^2F}{\partial z\partial y}\\ \frac{\partial^2F}{\partial z^2}\end{bmatrix},\quad \frac{\partial\nabla F}{\partial y}=\begin{bmatrix}\frac{\partial^2F}{\partial y\partial x}\\ \frac{\partial^2F}{\partial y^2}\\ \frac{\partial^2F}{\partial y\partial z}\end{bmatrix}-\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}\begin{bmatrix}\frac{\partial^2F}{\partial z\partial x}\\ \frac{\partial^2F}{\partial z\partial y}\\ \frac{\partial^2F}{\partial z^2}\end{bmatrix}.$$

Hint: Keep in mind that $z=z(x,y)$ and that its derivatives can be calculated using implicit differentiation.

(b) Using part (a) and exercise 26 show that

$$H=-\frac{1}{2}\operatorname{div}\frac{\nabla F}{|\nabla F|},$$

where

$$\operatorname{div}\begin{bmatrix}P\\ Q\\ R\end{bmatrix}=\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}.$$

Hint: It might help to first show that

$$\operatorname{div}\frac{\nabla F}{|\nabla F|}=\frac{\Delta F}{|\nabla F|}-\frac{(\nabla F)^t D^2F\nabla F}{|\nabla F|^3},$$

where

$$\Delta F=\frac{\partial^2F}{\partial x^2}+\frac{\partial^2F}{\partial y^2}+\frac{\partial^2F}{\partial z^2}$$

and

$$(\nabla F)^t D^2F\nabla F=\begin{bmatrix}\frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} & \frac{\partial F}{\partial z}\end{bmatrix}\begin{bmatrix}\frac{\partial^2F}{\partial x^2} & \frac{\partial^2F}{\partial y\partial x} & \frac{\partial^2F}{\partial z\partial x}\\ \frac{\partial^2F}{\partial x\partial y} & \frac{\partial^2F}{\partial y^2} & \frac{\partial^2F}{\partial z\partial y}\\ \frac{\partial^2F}{\partial x\partial z} & \frac{\partial^2F}{\partial y\partial z} & \frac{\partial^2F}{\partial z^2}\end{bmatrix}\begin{bmatrix}\frac{\partial F}{\partial x}\\ \frac{\partial F}{\partial y}\\ \frac{\partial F}{\partial z}\end{bmatrix}.$$

(c) Show that Scherk's surface $e^z \cos x = \cos y$ is minimal.
(d) Show that $\sin az = \sinh ax \sinh ay$ defines a minimal surface.
(e) Can you find other minimal surfaces of the form $F(x)\,G(y)\,H(z) = 1$?

(29) (Monge 1775) Consider a Monge patch $z = F(x, y)$. Define the two functions $p = \frac{\partial F}{\partial x}$ and $q = \frac{\partial F}{\partial y}$.

(a) Show that the Gauss curvature vanishes if and only if

$$\frac{\partial^2 F}{\partial x^2}\frac{\partial^2 F}{\partial y^2} - \left(\frac{\partial^2 F}{\partial x \partial y}\right)^2 = 0.$$

(b) Assume that $\frac{\partial^2 F}{\partial x \partial y} = 0$ on an open set.
  (i) Show that $F = f(x) + h(y)$.
  (ii) Show that the Gauss curvature vanishes if and only if $f'' = 0$ or $h'' = 0$.
  (iii) Show that if, say, $h'' = 0$, then it is a generalized cylinder.

(c) Assume that $\frac{\partial^2 F}{\partial x \partial y} \neq 0$ and that the Gauss curvature vanishes.
  (i) Show that we can locally reparametrize the surface using the reparametrization $(u, q) = (x, q(x, y))$.
  (ii) Show that $p = f(q)$ for some function f. Hint: In the (u, q)-coordinates $\frac{\partial p}{\partial u} = 0$. When doing this calculation keep in mind that y depends on u in (u, q)-coordinates as q depends on both x and y.
  (iii) Show in the same way that $F(x, y) - (px + qy) = h(q)$.
  (iv) Show that in the new parametrization:

$$y = -h'(q) - uf'(q)$$

and

$$\begin{aligned} z &= xp + qy + h(q) \\ &= h(q) - qh'(q) + (f(q) - qf'(q))\,u. \end{aligned}$$

  (v) Show that this is a ruled surface.
  (vi) Show that this ruled surface is a generalized cylinder when f'' vanishes.
  (vii) Show that it is a generalized cone when $h'' = af''$ for some constant a.
  (viii) Show that otherwise it is a tangent developable by showing that the lines in the ruling are all tangent to the curve that corresponds to

$$u = -\frac{h''}{f''}.$$

(30) Show that

$$\begin{aligned}
K &= \frac{\det\begin{bmatrix} \frac{\partial^2 \mathrm{q}}{\partial u^2}\cdot\left(\frac{\partial \mathrm{q}}{\partial u}\times\frac{\partial \mathrm{q}}{\partial v}\right) & \frac{\partial^2 \mathrm{q}}{\partial u\partial v}\cdot\left(\frac{\partial \mathrm{q}}{\partial u}\times\frac{\partial \mathrm{q}}{\partial v}\right) \\ \frac{\partial^2 \mathrm{q}}{\partial v\partial u}\cdot\left(\frac{\partial \mathrm{q}}{\partial u}\times\frac{\partial \mathrm{q}}{\partial v}\right) & \frac{\partial^2 \mathrm{q}}{\partial v^2}\cdot\left(\frac{\partial \mathrm{q}}{\partial u}\times\frac{\partial \mathrm{q}}{\partial v}\right)\end{bmatrix}}{\left(\det[\mathrm{I}]\right)^2} \\
&= \frac{\left(\frac{\partial}{\partial v}\Gamma_{uuv}-\frac{\partial}{\partial u}\Gamma_{uvv}\right)}{\det[\mathrm{I}]} \\
&\quad+\frac{\det\begin{bmatrix} 0 & \Gamma_{vvu} & \Gamma_{vvv} \\ \Gamma_{uuu} & g_{uu} & g_{uv} \\ \Gamma_{uuv} & g_{vu} & g_{vv}\end{bmatrix}-\det\begin{bmatrix} 0 & \Gamma_{uvu} & \Gamma_{uvv} \\ \Gamma_{uvu} & g_{uu} & g_{uv} \\ \Gamma_{uvv} & g_{vu} & g_{vv}\end{bmatrix}}{\left(\det[\mathrm{I}]\right)^2}.
\end{aligned}$$

(31) (Gauss) Show that if we define $|g|^2=\det[\mathrm{I}]$, then

$$\begin{aligned}
4|g|^4 K &= g_{uu}\left(\frac{\partial g_{uu}}{\partial v}\frac{\partial g_{vv}}{\partial v}-2\frac{\partial g_{uv}}{\partial u}\frac{\partial g_{vv}}{\partial v}+\left(\frac{\partial g_{vv}}{\partial u}\right)^2\right) \\
&\quad+g_{uv}\left(\frac{\partial g_{uu}}{\partial u}\frac{\partial g_{vv}}{\partial v}-\frac{\partial g_{vv}}{\partial u}\frac{\partial g_{uu}}{\partial v}-2\frac{\partial g_{uu}}{\partial v}\frac{\partial g_{uv}}{\partial v}-2\frac{\partial g_{vv}}{\partial u}\frac{\partial g_{uv}}{\partial u}+4\frac{\partial g_{uv}}{\partial u}\frac{\partial g_{uv}}{\partial v}\right) \\
&\quad+g_{vv}\left(\frac{\partial g_{uu}}{\partial u}\frac{\partial g_{vv}}{\partial u}-2\frac{\partial g_{uu}}{\partial u}\frac{\partial g_{uv}}{\partial v}+\left(\frac{\partial g_{uu}}{\partial v}\right)^2\right) \\
&\quad-2|g|^2\left(\frac{\partial^2 g_{uu}}{\partial v^2}-2\frac{\partial^2 g_{uv}}{\partial u\partial v}+\frac{\partial^2 g_{vv}}{\partial u^2}\right).
\end{aligned}$$

(32) (Frobenius) Show that if we define $|g|^2=\det[\mathrm{I}]$, then

$$\begin{aligned}
K &= -\frac{1}{4|g|^2}\det\begin{bmatrix} g_{uu} & g_{uv} & g_{vv} \\ \frac{\partial g_{uu}}{\partial u} & \frac{\partial g_{uv}}{\partial u} & \frac{\partial g_{vv}}{\partial u} \\ \frac{\partial g_{uu}}{\partial v} & \frac{\partial g_{uv}}{\partial v} & \frac{\partial g_{vv}}{\partial v}\end{bmatrix} \\
&\quad-\frac{1}{2|g|}\left(\frac{\partial}{\partial u}\left(\frac{\frac{\partial g_{vv}}{\partial u}-\frac{\partial g_{uv}}{\partial v}}{|g|}\right)+\frac{\partial}{\partial v}\left(\frac{\frac{\partial g_{uu}}{\partial v}-\frac{\partial g_{uv}}{\partial u}}{|g|}\right)\right).
\end{aligned}$$

(33) (Liouville) Show that if we define $|g|^2=\det[\mathrm{I}]$, then

$$\begin{aligned}
K &= \frac{1}{|g|}\left(\frac{\partial}{\partial v}\left(\frac{|g|}{g_{uu}}\Gamma_{uu}^{v}\right)-\frac{\partial}{\partial u}\left(\frac{|g|}{g_{uu}}\Gamma_{uv}^{v}\right)\right) \\
&= \frac{1}{|g|}\left(\frac{\partial}{\partial v}\left(\frac{|g|}{g_{vv}}\Gamma_{vv}^{u}\right)+\frac{\partial}{\partial u}\left(\frac{|g|}{g_{vv}}\Gamma_{uv}^{u}\right)\right).
\end{aligned}$$

5.4. Principal Curvatures

DEFINITION 5.4.1. The *principal curvatures* at a point q on a surface are

$$\begin{aligned}
\kappa_1 &= \max\{\mathrm{II}(X,X)\mid X\in T_qM \text{ and } |X|=1\}, \\
\kappa_2 &= \min\{\mathrm{II}(X,X)\mid X\in T_qM \text{ and } |X|=1\}.
\end{aligned}$$

We say that q is umbilic if the principal curvatures coincide, i.e., II is a multiple of I at q.

THEOREM 5.4.2 (Euler, 1760). *Let $E \in T_qM$ be a unit vector and κ_1, κ_2 the principal curvatures, then*

$$\mathrm{II}(E,E) = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta.$$

Moreover, the principal curvatures are eigenvalues for the Weingarten map.

PROOF. We argue as in the proof of theorem 4.4.9 with $Q = \mathrm{II}$. Let $\mathrm{II}(E,E)$ have a maximum at E_1 with $E_2 \in T_qM$ a unit vector orthogonal to E_1. It follows that all unit vectors at q have the form $E(\theta) = \cos\theta E_1 + \sin\theta E_2 \in T_qM$. Now consider

$$\mathrm{II}(E(\theta), E(\theta)) = \cos^2\theta\, \mathrm{II}(E_1,E_1) + 2\cos\theta\sin\theta\, \mathrm{II}(E_1,E_2) + \sin^2\theta\, \mathrm{II}(E_2,E_2).$$

By construction this is a function of θ that has a maximum at $\theta = 0$. The derivative at $\theta = 0$ is $2\,\mathrm{II}(E_1,E_2)$. Therefore, $\mathrm{II}(E_1,E_2) = 0$ and

$$\mathrm{II}(E(\theta), E(\theta)) = \cos^2\theta\, \mathrm{II}(E_1,E_1) + \sin^2\theta\, \mathrm{II}(E_2,E_2).$$

We claim that $L(E_i) = \kappa_i E_i$ for $i = 1,2$. To see this note that

$$\begin{aligned} L(E_1) &= \mathrm{I}(L(E_1), E_1) E_1 + \mathrm{I}(L(E_1), E_2) E_2 \\ &= \mathrm{II}(E_1,E_1) E_1 + \mathrm{II}(E_1,E_2) E_2 \\ &= \kappa_1 E_1. \end{aligned}$$

A similar argument works for E_2. □

DEFINITION 5.4.3. A vector that is an eigenvector for the Weingarten map is called a *principal direction*. A curve on a surface with the property that its velocity is always an eigenvector for the Weingarten map, i.e., a principal direction, is called a *line of curvature*.

By using $\mathrm{II} = Q$ in theorem 4.4.9 we obtain:

COROLLARY 5.4.4. *If a point q on a surface is not umbilic, then there is a parametrization* $\mathrm{q}(u,v)$ *such that the coordinate curves are lines of curvature:*

$$L\left(\frac{\partial \mathrm{q}}{\partial u}\right) = \kappa_1 \frac{\partial \mathrm{q}}{\partial u},\ L\left(\frac{\partial \mathrm{q}}{\partial v}\right) = \kappa_2 \frac{\partial \mathrm{q}}{\partial v},$$

and

$$\mathrm{I}\left(\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}\right) = Q\left(\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}\right) = 0.$$

EXAMPLE 5.4.5. The height function that measures the distance from a point on the surface to the tangent space T_qM is given by

$$f(\mathrm{q}) = (\mathrm{q} - q)\cdot \mathrm{n}(q).$$

Its partial derivatives with respect to a parametrization of the surface are

$$\begin{aligned} \frac{\partial f}{\partial w} &= \frac{\partial \mathrm{q}}{\partial w}\cdot \mathrm{n}(q), \\ \frac{\partial^2 f}{\partial w_1 \partial w_2} &= \frac{\partial^2 \mathrm{q}}{\partial w_1 \partial w_2}\cdot \mathrm{n}(q). \end{aligned}$$

Thus f has a critical point at q, and the second derivative matrix at q is simply $[\mathrm{II}]$. The second derivative test then tells us something about how the surface is placed in relation to T_qM. Specifically we see that if both principal curvatures have the same sign, or $K > 0$, then the surface must locally be on one side of the tangent plane, while if the principal curvatures have opposite signs, or $K < 0$, then the surface lies on both sides. In that case it'll look like a saddle.

We can now give a rather surprising characterization of planes and spheres.

THEOREM 5.4.6 (Meusnier, 1776). *If a surface has the property that $\kappa_1 = \kappa_2$ at all points, then $\kappa_1 = \kappa_2 = H$ is constant and the surface is part of a plane or sphere.*

PROOF. Since the principal curvatures are the eigenvalues of L it follows that $H = \frac{\kappa_1+\kappa_2}{2} = \kappa_1 = \kappa_2$.

Assume we have a parametrization $\mathrm{q}(u,v)$ of part of the surface. Since the principal curvatures agree at all points it follows that all directions are principal directions. In particular:

$$-\frac{\partial \mathrm{n}}{\partial w} = L\left(\frac{\partial \mathrm{q}}{\partial w}\right) = H\frac{\partial \mathrm{q}}{\partial w}.$$

By letting $w = u, v$ and taking partial derivatives of this equation we obtain

$$\begin{aligned} -\frac{\partial^2 \mathrm{n}}{\partial u \partial v} &= \frac{\partial H}{\partial u}\frac{\partial \mathrm{q}}{\partial v} + H\frac{\partial^2 \mathrm{q}}{\partial u \partial v}, \\ -\frac{\partial^2 \mathrm{n}}{\partial v \partial u} &= \frac{\partial H}{\partial v}\frac{\partial \mathrm{q}}{\partial u} + H\frac{\partial^2 \mathrm{q}}{\partial v \partial u}. \end{aligned}$$

As partial derivatives commute it follows that

$$\frac{\partial H}{\partial u}\frac{\partial \mathrm{q}}{\partial v} = \frac{\partial H}{\partial v}\frac{\partial \mathrm{q}}{\partial u}.$$

Since $\frac{\partial \mathrm{q}}{\partial u}$, $\frac{\partial \mathrm{q}}{\partial v}$ are linearly independent this forces $\frac{\partial H}{\partial u} = \frac{\partial H}{\partial v} = 0$. Thus H is constant.

Returning to the equation

$$-\frac{\partial \mathrm{n}}{\partial w} = H\frac{\partial \mathrm{q}}{\partial w}$$

we see that

$$\frac{\partial\left(\mathrm{n} + H\,\mathrm{q}\right)}{\partial w} = 0.$$

This implies that $\mathrm{n} + H\,\mathrm{q}$ is constant. When $H = 0$ this shows that n is constant and consequently the surface lies in a plane orthogonal to n. When H does not vanish we can assume that $H = \pm\frac{1}{R}$, $R > 0$. We then have that

$$\pm R\,\mathrm{n} + \mathrm{q} = \mathrm{c}$$

for some $\mathrm{c} \in \mathbb{R}^3$. This shows that

$$\left|\mathrm{q} - \mathrm{c}\right|^2 = R^2.$$

Hence q lies on the sphere of radius R centered at c. $\square$

Exercises

(1) Show that if two nonzero tangent vectors X, Y to a surface satisfy $\mathrm{I}(X,Y) = 0 = \mathrm{II}(X,Y)$, then they are principal directions.

(2) Show that the principal curvatures for a parametrized surface are the roots to the equation

$$\det([\mathrm{II}] - \kappa[\mathrm{I}]) = 0.$$

(3) Show that the principal curvatures are given by

$$\kappa_1 = H + \sqrt{H^2 - K} \text{ and } \kappa_2 = H - \sqrt{H^2 - K}.$$

(4) Show that following conditions are equivalent:
   (a) The principal curvatures at a point are equal.
   (b) The mean and Gauss curvatures at the point are related by $H^2 = K$.
   (c) $L = HI$, where I is the identity map on the tangent spaces.
   (d) $(L - HI)^2 = 0$.
   (e) The characteristic polynomial for L is a perfect square.

(5) Let $\mathrm{q}(u,v)$ be a parametrized surface without umbilics. Show that $\frac{\partial \mathrm{q}}{\partial u}$ and $\frac{\partial \mathrm{q}}{\partial v}$ are the principal directions if and only if $g_{uv} = 0 = L_{uv}$.

(6) Consider

$$z\left(x^2 + y^2\right) = \kappa_1 x^2 + \kappa_2 y^2.$$

   (a) Show that this defines a surface when $x^2 + y^2 > 0$.
   (b) Show that it is a ruled surface where the lines go through the z-axis and are perpendicular to the z-axis.
   (c) Show that if a general surface has principal curvatures κ_1, κ_2 at a point, then z corresponds to the possible values of the normal curvature at that point.

(7) (Rodrigues) Show that a curve $\mathrm{q}(t)$ on a surface with normal n is a line of curvature if and only if there is a function $\lambda(t)$ such that

$$-\lambda(t)\frac{d\,\mathrm{q}}{dt} = \frac{d(\mathrm{n} \circ \mathrm{q})}{dt}.$$

(8) Show that the principal curvatures are constant if and only if the Gauss and mean curvatures are constant.

(9) Consider the *pseudo-sphere*

$$\mathrm{q}(t,\mu) = \left(\frac{\cos\mu}{\cosh t}, \frac{\sin\mu}{\cosh t}, t - \tanh t\right).$$

This is a model for a surface with constant negative Gauss curvature. Note that the surface

$$\mathrm{q}(t,\mu) = \left(\frac{\cos\mu}{\cosh t}, \frac{\sin\mu}{\cosh t}, \tanh t\right)$$

is the sphere with a conformal (Mercator) parametrization.
   (a) Compute the first and second fundamental forms
   (b) Compute the principal curvatures, Gauss curvature, and mean curvature.

(10) A ruled surface $\mathrm{q}(u,v) = c(v) + uX(v)$ is called *developable* if all of the u-curves $\mathrm{q}(u) = \mathrm{q}(u,v)$ for fixed v are lines of curvature with $\kappa = 0$, i.e., $\frac{\partial \mathrm{n}}{\partial u} = 0$. Show that ruled surfaces are developable if and only if they have vanishing Gauss curvature.

(11) (Monge) Show that a curve $\mathrm{q}(t)$ on a surface with normal n is a line of curvature if and only if the ruled surface $\mathrm{q}^*(s,t) = \mathrm{q}(t) + s\,\mathrm{n}\circ\mathrm{q}(t)$ is developable. Hint: Note that the normal to $\mathrm{q}^*(s,t) = \mathrm{q}(t) + s\,\mathrm{n}\circ\mathrm{q}(t)$ at $s=0$ is $\mathrm{S}(t)$. So this surface is developable if and only if $S(t_0)$ is the normal to $\mathrm{q}^*(s,t_0)$ for all s.

(12) Show that if a surface has conformal Gauss map, then it is either minimal or part of a sphere.

(13) Show that if $\mathrm{III} = \lambda\,\mathrm{II}$ for some function λ on the surface, then either $K=0$ or the surface is part of a sphere.

(14) Show that all curves on a sphere or plane are lines of curvature. Use this to show that if two spheres; a plane and a sphere; or two planes intersect in a curve, then they intersect at a constant angle along this curve.

(15) Consider a surface of revolution

$$\mathrm{q}(r,\mu) = (r\cos\mu, r\sin\mu, h(r)).$$

Show that $\kappa_1 = \frac{d}{dr}(r\kappa_2)$.

(16) Consider a unit speed curve $c(s) : [0,L] \to \mathbb{R}^3$ with non-vanishing curvature and the tube of radius R around it

$$\mathrm{q}(s,\phi) = c(s) + R(\mathrm{N}_c\cos\phi + \mathrm{B}_c\sin\phi)$$

(see section 4.1 exercise 6 and section 4.3 exercise 7). Show that the principal directions are $-\mathrm{N}_c\sin\phi + \mathrm{B}_c\cos\phi$ and T_c with principal curvatures $1/R$ and $-\frac{\kappa\cos\phi}{1-\kappa R\cos\phi}$.

(17) Show that Enneper's surface

$$\mathrm{q}(u,v) = \left(u - \frac{1}{3}u^3 + uv^2, v - \frac{1}{3}v^3 + vu^2, u^2 - v^2\right)$$

has the property that $g_{uv} = L_{uv} = 0$ and that the u- and v-curves are planar. Hint: A curve is planar if $\det\begin{bmatrix} \mathrm{v} & \mathrm{a} & \mathrm{j} \end{bmatrix} = 0$.

(18) Show that

$$\mathrm{q}(u,v) = (u\cos\theta \pm \sin u\cosh v, v \pm \cos\theta\cos u\sinh v, \pm\sin\theta\cos u\cosh v)$$

has the property that $g_{uv} = L_{uv} = 0$ and that the u- and v-curves are planar. Hint: A curve is planar if $\det\begin{bmatrix} \mathrm{v} & \mathrm{a} & \mathrm{j} \end{bmatrix} = 0$.

(19) Consider a parametrized surface $\mathrm{q}(t,\phi)$ where the t- and ϕ-curves correspond to the principal directions. Assume that the principal curvatures are $\kappa_2 < \kappa_1$ and $\kappa_1 = 1/R$ is constant.
   (a) Consider $\mathrm{c}(t,\phi) = \mathrm{q}(t,\phi) + R\,\mathrm{n}(t,\phi)$ and show that

$$\frac{\partial\,\mathrm{c}}{\partial\phi} = 0,\ \frac{\partial\,\mathrm{c}}{\partial t} \neq 0.$$

   (b) Conclude that q is a tube of radius R (see section 4.1 exercise 6).
   (c) Show that a surface without umbilics where one of the principal curvatures is a positive constant is a tube.
   (d) Is it necessary to assume that the surface has no umbilics?

(20) Show that the geodesic torsion of a curve on a surface satisfies

$$\tau_g = (\kappa_2 - \kappa_1)\sin\phi\cos\phi,$$

where ϕ is the angle between the tangent to the curve and the principal direction corresponding to κ_1.

(21) (Rodrigues) Show that a unit speed curve on a surface is a line of curvature if and only if its geodesic torsion vanishes.

(22) (Joachimsthal) Let $\mathrm{q}(t)$ be a curve that lies on two surfaces M_1 and M_2 that have normals n_1 and n_2 respectively. Define

$$\theta(t) = \angle(\mathrm{n}_1 \circ \mathrm{q}(t), \mathrm{n}_2 \circ \mathrm{q}(t))$$

and assume that $0 < \theta(t) < \pi$, in other words the surfaces are not tangent to each other along the curve.

(a) Show that if $\mathrm{q}(t)$ is a line of curvature on both surfaces, then $\theta(t)$ is constant.

(b) Show that if $\mathrm{q}(t)$ is a line of curvature on one of the surfaces and $\theta(t)$ is constant, then $\mathrm{q}(t)$ is also a line of curvature on the other surface.

(23) Let $\mathrm{q}(u, v)$ be a parametrized surface and $\mathrm{q}^R = \mathrm{q} + R\,\mathrm{n}$ the parallel surface at distance R from q.

(a) Show that

$$\frac{\partial\,\mathrm{q}^R}{\partial w} = \frac{\partial\,\mathrm{q}}{\partial w} + R\frac{\partial\,\mathrm{n}}{\partial w} = (I - RL)\left(\frac{\partial\,\mathrm{q}}{\partial w}\right),$$

where I is the identity map $I(v) = v$.

(b) Show that q^R is a parametrized surface with normal n provided $R \neq \frac{1}{\kappa_1}, \frac{1}{\kappa_2}$.

(c) Show that

$$L^R = L \circ (I - RL)^{-1}$$

by using that

$$L\left(\frac{\partial\,\mathrm{q}}{\partial w}\right) = -\frac{\partial\,\mathrm{n}}{\partial w} = L^R\left(\frac{\partial\,\mathrm{q}^R}{\partial w}\right).$$

(d) Show that these surfaces all have the same principal directions with principal curvatures

$$\kappa_i^R = \frac{\kappa_i}{1 - R\kappa_i}.$$

(24) Let $\mathrm{q}(u, v)$ be a parametrized surface and $\mathrm{q}^R = \mathrm{q} + R\,\mathrm{n}$ the parallel surface at distance ϵ from q.

(a) Show that

$$\mathrm{I}^R = \mathrm{I} - 2R\,\mathrm{II} + R^2\,\mathrm{III}.$$

(b) Show that

$$\mathrm{II}^R = \mathrm{II} - R\,\mathrm{III}.$$

(c) Show that

$$\mathrm{III}^R = \mathrm{III}.$$

(d) How do you reconcile these relations with the formula

$$L^R = L \circ (I - RL)^{-1}$$

from the previous exercise?

(e) Show that

$$K^R = \frac{K}{1 - 2RH + R^2K}$$

and

$$H^R = \frac{H - RK}{1 - 2RH + R^2K}.$$

(f) Show that if $K = \frac{1}{R^2}$, then $H^{\pm R} = \frac{1}{2R}$. Conversely, if $H = \frac{1}{2R}$, then $K^R = \frac{1}{R^2}$.

(g) Show that

$$\det\left[\mathrm{I}^R\right] = \left(1 - 2RH + R^2K\right)^2 \det\left[\mathrm{I}\right].$$

5.5. Ruled Surfaces

A ruled surface comes about by selecting a curve $c(v)$ and then considering the surface one obtains by adding a line through each of the points on the curve. If the directions of those lines are given by $X(v)$, then the surface can be parametrized by $\mathrm{q}(u,v) = c(v) + uX(v)$. We can without loss of generality assume that X is a unit field, however, in many concrete examples throughout the exercises X might not be given as a unit vector. The condition for obtaining a parametrized surface is that $\frac{\partial \mathrm{q}}{\partial u} = X$ and $\frac{\partial \mathrm{q}}{\partial v} = \frac{dc}{dv} + u\frac{dX}{dv}$ are linearly independent. Even though we don't always obtain a surface for all parameter values it is important to consider the extended lines in the rulings for all values of u.

Example 5.5.1. A *generalized cylinder* is a ruled surface where X is constant, i.e., $\frac{dX}{dv} = 0$. This will be a parametrized surface everywhere if X is never tangent to c.

Example 5.5.2. A *generalized cone* is a ruled surface where c can be chosen to be constant, i.e., $\frac{dc}{dv} = 0$. This will clearly not be a parametrized surface when $u = 0$.

Example 5.5.3. A *tangent developable*, is a ruled surface where X is always tangent to c, i.e., X and $\frac{dc}{dv}$ are always proportional. This is also not a surface when $u = 0$. Note that generalized cones can be considered a special case of tangent developables. It is not unusual to also assume that that a tangent developable has the property that c is regular so as to avoid this overlap in definitions.

Example 5.5.4. An example of a cone that is not rotationally symmetric is the elliptic cone

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z^2.$$

The elliptic hyperboloid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z^2 + 1$$

is an example of a surface that is ruled in two different ways, but which does not have zero Gauss curvature. We can let

$$c(t) = (a\cos(t), b\sin(t), 0)$$

be the ellipse where $z = 0$. The fields generating the lines are given by

$$X = \frac{dc}{dt} + (0, 0, \pm 1)$$

and it is not difficult to check that

$$\mathrm{q}(s,t) = c(t) + s\left(\frac{dc}{dt} + (0, 0, \pm 1)\right)$$

are both rulings of the elliptic hyperboloid.

PROPOSITION 5.5.5. *Ruled surfaces have non-positive Gauss curvature and the Gauss curvature vanishes if and only if*

$$\left(X \times \frac{dc}{dv}\right) \cdot \frac{dX}{dv} = 0.$$

In particular, generalized cylinders, generalized cones, and tangent developables have vanishing Gauss curvature.

PROOF. Since $\frac{\partial^2 \mathrm{q}}{\partial u^2} = 0$ it follows that $L_{uu} = 0$ and

$$K = \frac{-L_{uv}^2}{g_{uu}g_{vv} - g_{uv}^2} \leq 0.$$

Moreover, K vanishes precisely when

$$L_{uv} = \frac{\partial^2 \mathrm{q}}{\partial u \partial v} \cdot \mathrm{n} = \frac{dX}{dv} \cdot \mathrm{n} = 0.$$

Since the normal is given by

$$\mathrm{n} = \frac{X \times \left(\frac{dc}{dv} + u\frac{dX}{dv}\right)}{\left|X \times \left(\frac{dc}{dv} + u\frac{dX}{dv}\right)\right|}$$

this translates to

$$\begin{aligned} 0 &= \left(X \times \left(\frac{dc}{dv} + u\frac{dX}{dv}\right)\right) \cdot \frac{dX}{dv} \\ &= \left(X \times \frac{dc}{dv}\right) \cdot \frac{dX}{dv}, \end{aligned}$$

which is what we wanted to prove. □

DEFINITION 5.5.6. A ruled surface with the property that the normal is constant in the direction of the ruling, i.e.,

$$\frac{\partial \mathrm{n}}{\partial u} = 0,$$

is called a *developable* or *developable surface*.

EXAMPLE 5.5.7. Generalized cylinders, cones, and tangent developables are all developables. In general the normal is given by

$$\mathrm{n} = \frac{X \times \left(\frac{dc}{dv} + u\frac{dX}{dv}\right)}{\left|X \times \left(\frac{dc}{dv} + u\frac{dX}{dv}\right)\right|}.$$

For a generalized cylinder this specializes to

$$\mathrm{n} = \frac{X \times \frac{dc}{dv}}{\left|X \times \frac{dc}{dv}\right|}$$

while for cones and tangent developables

$$\mathrm{n} = \frac{X \times u\frac{dX}{dv}}{\left|X \times u\frac{dX}{dv}\right|} = \pm\frac{X \times \frac{dX}{dv}}{\left|X \times \frac{dX}{dv}\right|}.$$

In each case the normal is independent of u.

We start with a characterization of developables in terms of Gauss curvature. Euler was the the first to suggest this result and used it to show that spheres can't admit Cartesian parametrizations. Monge clarified the statement and gave the first proof.

LEMMA 5.5.8 (Monge, 1775). *A surface with vanishing Gauss curvature and no umbilics is a developable surface. Conversely any developable has vanishing Gauss curvature.*

PROOF. First note that a developable has the property that the lines in the ruling are lines of curvature and that the principal value vanishes in the direction of the lines. This establishes the second claim and also guides us as to how to find the lines in a ruling.

Assume now that the surface has zero Gauss curvature. We shall show that the principal directions that correspond to the principal value 0 generate lines of curvature that are straight lines. This will create a ruling. The normal is by definition constant along these lines as they are lines of curvature for the principal value 0.

Since the surface has no umbilics we can use corollary 5.4.4 to select a parametrization where $\frac{\partial \mathrm{q}}{\partial u}$, $\frac{\partial \mathrm{q}}{\partial v}$ are principal directions and $g_{uv} = L_{uv} = 0$. Using that the Gauss curvature vanishes allows us to assume that

$$\begin{aligned} -\frac{\partial \mathrm{n}}{\partial u} &= L\left(\frac{\partial \mathrm{q}}{\partial u}\right) = 0, \\ -\frac{\partial \mathrm{n}}{\partial v} &= L\left(\frac{\partial \mathrm{q}}{\partial v}\right) = \kappa \frac{\partial \mathrm{q}}{\partial v},\ \kappa \neq 0. \end{aligned}$$

Combining these two equations we obtain

$$0 = \frac{\partial^2 \mathrm{n}}{\partial u \partial v} = \kappa \frac{\partial^2 \mathrm{q}}{\partial u \partial v} + \frac{\partial \kappa}{\partial u} \frac{\partial \mathrm{q}}{\partial v}.$$

This shows that

$$\begin{aligned} \frac{\partial^2 \mathrm{q}}{\partial u^2} \cdot \frac{\partial \mathrm{q}}{\partial v} &= \frac{\partial}{\partial u}\left(\frac{\partial \mathrm{q}}{\partial u} \cdot \frac{\partial \mathrm{q}}{\partial v}\right) - \frac{\partial^2 \mathrm{q}}{\partial u \partial v} \cdot \frac{\partial \mathrm{q}}{\partial u} \\ &= 0 + \frac{1}{\kappa} \frac{\partial \kappa}{\partial u} \frac{\partial \mathrm{q}}{\partial v} \cdot \frac{\partial \mathrm{q}}{\partial u} \\ &= 0. \end{aligned}$$

By assumption we also have

$$\frac{\partial^2 \mathrm{q}}{\partial u^2} \cdot \mathrm{n} = L_{uu} = 0.$$

Thus $\frac{\partial^2 \mathrm{q}}{\partial u^2}$ must be parallel to $\frac{\partial \mathrm{q}}{\partial u}$ and consequently the u-curves on the surface have zero curvature as curves in $\mathbb{R}^3$. This shows that they are straight lines. □

The next result shows that ruled surfaces admit a standard set of parameters that make it easier to recognize the three different types of developables.

PROPOSITION 5.5.9. *A ruled surface* $\mathrm{q}(u,v) = c(v) + uX(v)$ *can be reparametrized as* $\mathrm{q}(s,v) = c^*(v) + sX(v)$, *where* $\frac{dc^*}{dv} \perp \frac{dX}{dv}$.

The ruled surface is a generalized cone if and only if c^* *is constant. The ruled surface is a tangent developable if and only if* $\frac{dc^*}{dv}$ *and* X *are proportional for all* v.

PROOF. Note that no change in the parametrization is necessary if X is constant. When $\frac{dX}{dv} \neq 0$ define

$$c^* = c - \frac{\frac{dc}{dv} \cdot \frac{dX}{dv}}{\left|\frac{dX}{dv}\right|^2} X$$

and

$$s = u + \frac{\frac{dc}{dv} \cdot \frac{dX}{dv}}{\left|\frac{dX}{dv}\right|^2}.$$

Here $\frac{\partial s}{\partial u} = 1$, so $(s(u,v), v)$ is locally a valid reparametrization of the surface. It is also clear that $\mathrm{q}(u,v) = c^*(v) + sX(v) = \mathrm{q}(s,v)$. Moreover, as X is a unit field it is perpendicular to its derivative so we have

$$\begin{aligned}
\frac{dc^*}{dv} \cdot \frac{dX}{dv} &= \left(\frac{dc}{dv} - \frac{d}{dv}\left(\frac{\frac{dc}{dv} \cdot \frac{dX}{dv}}{\left|\frac{dX}{dv}\right|^2}\right) X - \left(\frac{\frac{dc}{dv} \cdot \frac{dX}{dv}}{\left|\frac{dX}{dv}\right|^2}\right)\frac{dX}{dv}\right) \cdot \frac{dX}{dv} \\
&= \frac{dc}{dv} \cdot \frac{dX}{dv} - \left(\frac{\frac{dc}{dv} \cdot \frac{dX}{dv}}{\left|\frac{dX}{dv}\right|^2}\right)\frac{dX}{dv} \cdot \frac{dX}{dv} \\
&= \frac{dc}{dv} \cdot \frac{dX}{dv} - \left(\frac{dc}{dv} \cdot \frac{dX}{dv}\right) \\
&= 0.
\end{aligned}$$

It is clear that we obtain a generalized cone when c^* is constant and a tangent developable if $\frac{dc^*}{dv}$ and X are parallel to each other.

Conversely if the ruled surface $\mathrm{q}(u,v)$ is a generalized cone, then there is a unique function $u = u(v)$ such that $\mathrm{q}(u(v), v)$ is constant. Thus

$$0 = \frac{dc}{dv} + u(v)\frac{dX}{dv} + \frac{du(v)}{dv} X.$$

If we multiply by $\frac{dX}{dv}$, then we obtain

$$u(v) = -\frac{\frac{dc}{dv} \cdot \frac{dX}{dv}}{\left|\frac{dX}{dv}\right|^2}.$$

This corresponds exactly to $s = 0$ in the parametrization $\mathrm{q}(s,v) = c^*(v) + sX(v)$. So it follows that $c^*(v)$ is constant.

When the ruled surface is a tangent developable it is possible to find $u = u(v)$ such that the curve $\beta(v) = \mathrm{q}(u(v), v)$ is tangent to the extended lines in the ruling, i.e., $\frac{d\beta}{dv}$ and X are proportional. In particular,

$$\begin{aligned}
0 &= \frac{d\beta}{dv} \cdot \frac{dX}{dv} \\
&= \left(\frac{dc}{dv} + u(v)\frac{dX}{dv} + \frac{du(v)}{dv} X\right) \cdot \frac{dX}{dv} \\
&= \frac{dc}{dv} \cdot \frac{dX}{dv} + u(v)\left|\frac{dX}{dv}\right|^2.
\end{aligned}$$

It follows again that $u(v)$ corresponds exactly to $s = 0$, which forces β to be c^*. $\square$

We are now ready to explain the possible shapes of surfaces with zero Gauss curvature. This gives us a partial answer to the converse of theorem 5.3.6, where it was shown that a surface with a Cartesian parametrization has vanishing Gauss curvature.

THEOREM 5.5.10 (Monge, 1775). *A developable surface is a generalized cylinder, generalized cone, or a tangent developable at almost all points of the surface.*

PROOF. We can assume that the surface is given by

$$\mathrm{q}(s,v) = c(v) + sX(v),$$

where $\frac{dc}{dv} \perp \frac{dX}{dv}$. The Gauss curvature vanishes precisely when

$$\left(X \times \frac{dc}{dv}\right) \cdot \frac{dX}{dv} = 0.$$

If $\frac{dX}{dv} = 0$ on an interval, then the surface is a generalized cylinder. So we can assume that $\frac{dX}{dv} \neq 0$. This implies that X and $\frac{dX}{dv}$ are linearly independent as they are orthogonal. The condition

$$\left(X \times \frac{dc}{dv}\right) \cdot \frac{dX}{dv} = 0$$

on the other hand implies that the three vectors are linearly dependent. We already know that $\frac{dc}{dv} \perp \frac{dX}{dv}$, so this forces

$$\frac{dc}{dv} = \left(\frac{dc}{dv} \cdot X\right) X.$$

When $\frac{dc}{dv} \neq 0$, then X is tangent to c and so we have a tangent developable. On the other hand, if $\frac{dc}{ds} = 0$ on an interval, then the surface must be a generalized cone on that interval.

Thus the surface is divided into regions each of which can be identified with our three basic types of ruled surfaces and then glued together along lines that go through parameter values where either $\frac{dX}{dv} = 0$ or $\frac{dc}{dv} = 0$. □

There is also a similar and very interesting result for ruled minimal surfaces.

THEOREM 5.5.11 (Catalan). *Any ruled surface that is minimal is planar or a helicoid at almost all points of the surface.*

PROOF. Assume that we have a parametrization $\mathrm{q}(s,v) = c(v) + sX(v)$, where $\frac{dc}{dv} \cdot \frac{dX}{dv} = 0$. In case the surface also has vanishing Gauss curvature it follows that it is planar as the second fundamental form vanishes. Therefore, we can assume that both c and X are regular curves and additionally that $\frac{dc}{dv}$ is not parallel to X.

The mean curvature is given by the general formula

$$H = \frac{L_{ss}g_{vv} - 2L_{sv}g_{sv} + L_{vv}g_{ss}}{2\left(g_{ss}g_{vv} - g_{sv}^2\right)},$$

where

$$\begin{aligned}
g_{ss} &= 1,\\
g_{sv} &= \frac{dc}{dv} \cdot X,\\
g_{vv} &= \left|\frac{dc}{dv}\right|^2 + s^2,
\end{aligned}$$

$$
\begin{aligned}
\mathrm{n} &= \frac{X \times \left(\frac{dc}{dv} + s\frac{dX}{dv}\right)}{\left|X \times \left(\frac{dc}{dv} + s\frac{dX}{dv}\right)\right|}, \\
L_{ss} &= 0, \\
L_{sv} &= -\frac{dX}{dv} \cdot \mathrm{n}, \\
L_{vv} &= -\left(\frac{d^2c}{dv^2} + s\frac{d^2X}{dv^2}\right) \cdot \mathrm{n}\,.
\end{aligned}
$$

Thus $H = 0$ precisely when

$$
-2\left(\frac{dc}{dv} \cdot X\right)\left(\frac{dX}{dv} \cdot \mathrm{n}\right) = -\left(\frac{d^2c}{dv^2} + s\frac{d^2X}{dv^2}\right) \cdot \mathrm{n},
$$

which implies

$$
2\left(\frac{dc}{dv} \cdot X\right)\left(\frac{dX}{dv} \cdot \left(X \times \left(\frac{dc}{dv} + s\frac{dX}{dv}\right)\right)\right) = \left(\frac{d^2c}{dv^2} + s\frac{d^2X}{dv^2}\right) \cdot \left(X \times \left(\frac{dc}{dv} + s\frac{dX}{dv}\right)\right).
$$

The left hand side can be simplified to be independent of s:

$$
2\left(\frac{dc}{dv} \cdot X\right)\left(\frac{dX}{dv} \cdot \left(X \times \left(\frac{dc}{dv} + s\frac{dX}{dv}\right)\right)\right) = 2\left(\frac{dc}{dv} \cdot X\right)\left(\frac{dX}{dv} \cdot \left(X \times \frac{dc}{dv}\right)\right).
$$

The right hand side can be expanded in terms of s as follows

$$
\begin{aligned}
\left(\frac{d^2c}{dv^2} + s\frac{d^2X}{dv^2}\right) \cdot \left(X \times \left(\frac{dc}{dv} + s\frac{dX}{dv}\right)\right) &= \frac{d^2c}{dv^2} \cdot \left(\frac{dc}{dv} \times X\right) \\
&\quad + s\left(\frac{d^2c}{dv^2} \cdot \left(X \times \frac{dX}{dv}\right) + \frac{d^2X}{dv^2} \cdot \left(X \times \frac{dc}{dv}\right)\right) \\
&\quad + s^2\frac{d^2X}{dv^2} \cdot \left(X \times \frac{dX}{dv}\right).
\end{aligned}
$$

This leads us to 3 identities depending on the powers of s. From the s^2-term we have

$$
\frac{d^2X}{dv^2} \cdot \left(X \times \frac{dX}{dv}\right) = 0.
$$

In other words:

$$
\frac{d^2X}{dv^2} \in \operatorname{span}\left\{X, \frac{dX}{dv}\right\}.
$$

At this point it is convenient to assume that v is the arclength parameter for X. With that in mind we have

$$
\begin{aligned}
\frac{d^2X}{dv^2} &= \left(\frac{d^2X}{dv^2} \cdot X\right) X + \left(\frac{d^2X}{dv^2} \cdot \frac{dX}{dv}\right)\frac{dX}{dv} \\
&= -\left(\frac{dX}{dv} \cdot \frac{dX}{dv}\right) X \\
&= -X.
\end{aligned}
$$

This implies that X is in fact a planar circle of radius 1. For simplicity let us further assume that it is the unit circle in the (x, y)-plane, i.e.,

$$
X(v) = (\cos v, \sin v, 0)\,.
$$

From the s-term we obtain

$$\begin{aligned}
0 &= \frac{d^2c}{dv^2}\cdot\left(X\times\frac{dX}{dv}\right)+\frac{d^2X}{dv^2}\cdot\left(X\times\frac{dc}{dv}\right)\\
&= \frac{d^2c}{dv^2}\cdot\left(X\times\frac{dX}{dv}\right)-X\cdot\left(X\times\frac{dc}{dv}\right)\\
&= \frac{d^2c}{dv^2}\cdot\left(X\times\frac{dX}{dv}\right)\\
&= \frac{d^2c}{dv^2}\cdot\begin{bmatrix}0\\0\\1\end{bmatrix}
\end{aligned}$$

showing that $\frac{d^2c}{dv^2}$ also lies in the (x,y)-plane. In particular,

$$\frac{dc}{dv}\cdot\begin{bmatrix}0\\0\\1\end{bmatrix}=h$$

is constant. Since $\frac{dc}{dv}\perp\frac{dX}{dv}$ we obtain

$$\frac{dc}{dv}=\left(\frac{dc}{dv}\cdot X\right)X+\begin{bmatrix}0\\0\\h\end{bmatrix}$$

and

$$\frac{dc}{dv}\times X=\begin{bmatrix}0\\0\\h\end{bmatrix}\times X=h\frac{dX}{dv}.$$

This considerably simplifies the terms that are independent of s in the mean curvature equation

$$2\left(\frac{dc}{dv}\cdot X\right)\left(\frac{dX}{dv}\cdot\left(X\times\frac{dc}{dv}\right)\right)=\frac{d^2c}{dv^2}\cdot\left(X\times\frac{dc}{dv}\right)$$

as we now obtain

$$\begin{aligned}
2h\frac{dc}{dv}\cdot X &= h\frac{d^2c}{dv^2}\cdot\frac{dX}{dv}\\
&= -h\frac{dc}{dv}\cdot\frac{d^2X}{dv^2}\\
&= h\frac{dc}{dv}\cdot X.
\end{aligned}$$

When $h=0$ the curve c also lies in the (x,y)-plane and the surface is planar. Otherwise $\frac{dc}{dv}\cdot X=0$ which implies that

$$\frac{dc}{dv}=\left(\frac{dc}{dv}\cdot X\right)X+\begin{bmatrix}0\\0\\h\end{bmatrix}=\begin{bmatrix}0\\0\\h\end{bmatrix}$$

and

$$c=\begin{bmatrix}0\\0\\hv+v_0\end{bmatrix}$$

for a constant v_0.

The surface is then given by

$$\mathrm{q}(s,v) = \begin{bmatrix} s\cos v \\ s\sin v \\ hv + v_0 \end{bmatrix},$$

which shows explicitly that it is a helicoid. □

Exercises

(1) Show that a generalized cylinder $\mathrm{q}(u,v) = c(v) + uX$ where X is a fixed unit vector admits a parametrization $\mathrm{q}(s,t) = c^*(t) + sX$, where c^* is parametrized by arclength and lies a plane orthogonal to X.

(2) Does the equation

$$\left(\alpha^2 - y^2\right)(\beta - z)^2 = \alpha\beta^2 x^2$$

define a ruled surface? Hint: A ruled surface contains a straight line through every point.

(3) Let $\mathrm{q}(u,v)$ be a parametrized surface and $\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$ a curve with $\dot{\mathrm{q}}(0) \neq 0$ and $\ddot{\mathrm{q}}(0) = \dddot{\mathrm{q}}(0) = 0$ as a curve in $\mathbb{R}^3$. Show that $\dot{\mathrm{q}}(0)$ is a principal direction with principal curvature 0.

(4) Consider the surface $\mathrm{q}(u,v) = c(v) + u\mathrm{D}_c(v)$, where $\mathrm{D}_c = \tau\,\mathrm{T}_c + \kappa\,\mathrm{B}_c$ is the Darboux vector for the unit speed curve c. Keep in mind that the Darboux vector is not necessarily a unit vector so the properties developed in this section don't apply directly.
 (a) Show that this is a ruled surface that is developable.
 (b) Show that this is a generalized cylinder precisely when $\frac{d}{dv}\frac{\tau}{\kappa} = 0$. Hint. See section 3.2 exercise 12.
 (c) Show that this is a generalized cone precisely when $\frac{d^2}{dv^2}\frac{\tau}{\kappa} = 0$ and $\frac{d}{dv}\frac{\tau}{\kappa} \neq 0$. Hint: If it is a generalized cone then there is a function $u(v)$ so that $\mathrm{q}(u(v), v)$ is constant. Show that this implies that $u(v)\kappa(v)$ is constant and $u(v)\tau(v)$ has constant derivative.
 (d) Show that this is a tangent developable when $\frac{d^2}{dv^2}\frac{\tau}{\kappa} \neq 0$.

(5) Consider a parameterized surface $\mathrm{q}(u,v)$. Show that the Gauss curvature vanishes if and only if $\frac{\partial\,\mathrm{n}}{\partial u}$, $\frac{\partial\,\mathrm{n}}{\partial v}$ are linearly dependent everywhere.

(6) Consider

$$\mathrm{q}(u,v) = \left(u+v, u^2 + 2uv, u^3 + 3u^2 v\right).$$

 (a) Determine where it defines a surface.
 (b) Show that the Gauss curvature vanishes.
 (c) What type of ruled surface is it?

(7) Consider the Monge patch

$$z = \sum_{k=2}^{n} (ax+by)^k + cx + dy + f.$$

 (a) Show that the Gauss curvature vanishes.
 (b) Show that it defines a generalized cylinder.

(8) Consider the equation

$$xy = (z-\alpha)^2.$$

 (a) Show that this defines a surface when $(x,y,z) \neq (0,0,\alpha)$.

(b) Show that it defines a generalized cone.

(9) Consider the equation

$$4\left(y-x^{2}\right)\left(xz-y^{2}\right)=\left(xy-z\right)^{2}.$$

(a) Show that this defines a surface when $(x,y,z)\neq\left(x,x^{2},x^{3}\right)$.

(b) Show that it defines a tangent developable.

(10) (Euler, 1775) Let $c(t)$ be a unit speed space curve with curvature $\kappa(t)>0$. Show that the tangent developable

$$\mathrm{q}(s,t)=c(t)+s\frac{dc}{dt}$$

admits Cartesian coordinates. Hint: There is a unit speed planar curve $c^{*}(t)$ whose curvature is $\kappa(t)$. Show that there is a natural isometry between the part of the plane parametrized by

$$\mathrm{q}^{*}(s,t)=c^{*}(t)+s\frac{dc^{*}}{dt}$$

and the tangent developable $\mathrm{q}(s,t)$.

(11) Use the previous exercise to show that a surface with $K=0$ and no umbilics locally admits Cartesian coordinates at almost all points.

(12) Show that a surface given by an equation

$$F(x,y,z)=R$$

has vanishing Gauss curvature if and only if

$$\det\begin{bmatrix}\frac{\partial^{2}F}{\partial x^{2}} & \frac{\partial^{2}F}{\partial y\partial x} & \frac{\partial^{2}F}{\partial z\partial x} & \frac{\partial F}{\partial x}\\ \frac{\partial^{2}F}{\partial x\partial y} & \frac{\partial^{2}F}{\partial y^{2}} & \frac{\partial^{2}F}{\partial z\partial y} & \frac{\partial F}{\partial y}\\ \frac{\partial^{2}F}{\partial x\partial z} & \frac{\partial^{2}F}{\partial y\partial z} & \frac{\partial^{2}F}{\partial z^{2}} & \frac{\partial F}{\partial z}\\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} & \frac{\partial F}{\partial z} & 0\end{bmatrix}=0.$$

Hint: Locally represent this surface as a Monge patch and use implicit differentiation.

(13) Show that a ruled surface with constant and non-zero mean curvature is a generalized cylinder.

(14) Show directly that if a minimal surface has vanishing Gauss curvature, then it is part of a plane.

(15) Assume that we have a ruled surface

$$\mathrm{q}(u,v)=c(v)+uX(v)$$

where $|X|=1$.

(a) Show that if we use

$$c^{*}=c+\left(a-\int\frac{dc}{dv}\cdot Xdv\right)X$$

and

$$w=u+\int\frac{dc}{dv}\cdot Xdv-a$$

for some constant a, then

$$\mathrm{q}(w,v)=c^{*}(v)+wX(v)$$

parametrizes the same surface and has the property that all v-curves are orthogonal to X and thus to the lines in the ruling.

(b) Show that if $\frac{dX}{dv} \neq 0$, then we can reparametrize X by arclength and thus obtain a parametrization

$$\mathrm{q}(w,t) = c^*(t) + wX(t),$$

where the t-curves are orthogonal to the ruling and X is a unit field parametrized by arclength.

(c) Show that if c^* is regular and has positive curvature and s denotes the arclength parameter for c^* we obtain $X(s) = \cos(\phi(s))\,\mathrm{N}_{c^*} + \sin(\phi(s))\,\mathrm{B}_{c^*}$ for some function $\phi(s)$.

(16) Assume that we have a minimal ruled surface

$$\mathrm{q}(w,t) = c(t) + wX(t)$$

parametrized as in the previous exercise with t-curves perpendicular to X and X a unit field parametrized by arclength. Reprove Catalan's theorem using this parametrization. Hint: One strategy is to first show that X is a unit circle, then show that $\ddot{c}$ is proportional to X, and finally conclude that the t-curves are all Bertrand mates to each other (see section 3.2).

(17) Let $\mathrm{q}(s)$ be a unit speed asymptotic line (see section 5.1 exercise 6) of a ruled surface $\mathrm{q}(u,v) = c(v) + uX(v)$. Note that u-curves are asymptotic lines.

(a) Show that

$$\det\begin{bmatrix} \ddot{\mathrm{q}}, & X, & \frac{dc}{dv} + u\frac{dX}{dv} \end{bmatrix} = 0.$$

(b) Assume for the remainder of the exercise that $K < 0$. Show that there is a unique asymptotic line through every point that is not tangent to X.

(c) Show that this asymptotic line can locally be reparametrized as

$$c(v) + u(v)X(v),$$

where

$$\frac{du}{dv} = \frac{\det\begin{bmatrix} X, & \frac{dc}{dv} + u(v)\frac{dX}{dv}, & \frac{d^2c}{dv^2} + u(v)\frac{d^2X}{dv^2} \end{bmatrix}}{2\det\begin{bmatrix} \frac{dc}{dv}, & X, & \frac{dX}{dv} \end{bmatrix}}.$$

(18) Consider the cubic equation with variable t:

$$x + yt + zt^2 + t^3 = 0$$

and discriminant:

$$D = z^2y^2 - 4z^3x + 18xyz - 4y^3 - 27x^2.$$

Show that $D = 0$ corresponds to the tangent developable of the curve $\left(t^3, 3t^2, 3t\right)$. Hint: Show that if

$$(x,y,z) = \left(t^3, 3t^2, 3t\right) + s\left(3t^2, 6t, 3\right),$$

then

$$\begin{aligned} z^2y^2 - 4z^3x &= -27t^3(t+s)\left(t^2 + 5st + 4s^2\right), \\ 18xyz &= 27t^3(t+s)\left(6t^2 + 30st + 36s^2\right), \\ -4y^3 - 27x^2 &= -27t^3(t+s)\left(5t^2 + 25st + 32s^2\right). \end{aligned}$$

(19) Consider the reduced quartic equation with variable t:

$$x + yt + zt^2 + t^4 = 0$$

and discriminant:

$$D = \left(x + \frac{z^2}{12}\right)^3 - 3^3\left(\frac{xz}{6} - \frac{y^2}{16} - \frac{z^3}{6^3}\right)^2.$$

Show that $D = 0$ corresponds to the tangent developable of the curve $\left(-3t^4, 8t^3, -6t^2\right)$. Hint: Show that if

$$(x, y, z) = \left(-3t^4, 8t^3, -6t^2\right) + s\left(-12t^3, 24t^2, -12t\right),$$

then

$$\begin{aligned} x + \frac{z^2}{12} &= 12s^2t^2, \\ \frac{xz}{6} - \frac{y^2}{16} - \frac{z^3}{6^3} &= 8s^3t^3. \end{aligned}$$

(20) Consider a family of planes in (x, y, z)-space parametrized by t:

$$F(x, y, z, t) = a(t)x + b(t)y + c(t)z + d(t) = 0.$$

An *envelope* to this family is a surface with the property that these planes are precisely the tangent planes to the surface.

(a) Show that an envelope exists and can be determined by the equations:

$$\begin{aligned} F = a(t)x + b(t)y + c(t)z + d(t) &= 0 \\ \frac{\partial F}{\partial t} = \dot{a}(t)x + \dot{b}(t)y + \dot{c}(t)z + \dot{d}(t) &= 0 \end{aligned}$$

when

$$\begin{bmatrix} a & b & c \\ \dot{a} & \dot{b} & \dot{c} \end{bmatrix}$$

has rank 2. Hint: use t and one of the coordinates x, y, z as parameters. The parametrization might be singular for some parameter values. Specifically assume that

$$\det\begin{bmatrix} a & b \\ \dot{a} & \dot{b} \end{bmatrix} \neq 0,$$

so that the surface can be parametrized as $\mathrm{q}(t, z) = (x(t, z), y(t, z), z)$ and show that the tangent vectors

$$X = \frac{\partial\,\mathrm{q}}{\partial t}, \frac{\partial\,\mathrm{q}}{\partial z}$$

satisfy the equation:

$$\begin{bmatrix} a & b & c \end{bmatrix} X = 0.$$

Finally show that the points in the tangent plane all have the form

$$\mathrm{q} + \alpha\frac{\partial\,\mathrm{q}}{\partial t} + \beta\frac{\partial\,\mathrm{q}}{\partial z},\ \alpha, \beta \in \mathbb{R}^2$$

and satisfy

$$F\left(\mathrm{q} + \alpha\frac{\partial\,\mathrm{q}}{\partial t} + \beta\frac{\partial\,\mathrm{q}}{\partial z}, t\right) = 0.$$

(b) Show that the envelope is a ruled surface. Hint: Assume that

$$\det \begin{bmatrix} a & b \\ \dot{a} & \dot{b} \end{bmatrix} \neq 0,$$

and show that $\frac{\partial^2 \mathrm{q}}{\partial z^2} = 0$.

(c) Show that the envelope is a generalized cylinder when the three functions a, b, and c are linearly dependent. Hint: show that the tangent vectors $\frac{\partial \mathrm{q}}{\partial z}$ are all parallel.

(d) Show that the envelope is a generalized cone when the function d is a linear combination of a, b, and c and the Wronskian

$$\det \begin{bmatrix} a & b & c \\ \dot{a} & \dot{b} & \dot{c} \\ \ddot{a} & \ddot{b} & \ddot{c} \end{bmatrix} \neq 0.$$

Hint: If $d + \alpha_0 a + \beta_0 b + \gamma_0 c = 0$, then $(\alpha_0, \beta_0, \gamma_0)$ is the vertex of the cone.

(e) Show that the envelope is a tangent developable when the Wronskian

$$\det \begin{bmatrix} a & b & c & d \\ \dot{a} & \dot{b} & \dot{c} & \dot{d} \\ \ddot{a} & \ddot{b} & \ddot{c} & \ddot{d} \\ \dddot{a} & \dddot{b} & \dddot{c} & \dddot{d} \end{bmatrix} \neq 0.$$

Note that the two previous exercises are concrete examples of this. Hint: Show that the equations:

$$\begin{aligned} F = a(t)\,x + b(t)\,y + c(t)\,z + d(t) &= 0 \\ \frac{\partial F}{\partial t} = \dot{a}(t)\,x + \dot{b}(t)\,y + \dot{c}(t)\,z + \dot{d}(t) &= 0 \\ \frac{\partial^2 F}{\partial t^2} = \ddot{a}(t)\,x + \ddot{b}(t)\,y + \ddot{c}(t)\,z + \ddot{d}(t) &= 0 \end{aligned}$$

determine the curve that generates the tangent developable.

(f) Show that for fixed (x_0, y_0, z_0) the solutions or roots to the equation $F(x_0, y_0, z_0, t) = 0$ correspond to the tangent planes to the envelope that pass through (x_0, y_0, z_0).

(21) Let $\mathrm{q}(u, v) = c(v) + uX(v)$ be a developable surface. Show that there exist functions $\alpha(v)$, $\beta(v)$, $\gamma(v)$, and $\delta(v)$ such that the surface is an envelope of the planes

$$\alpha(v)\,x + \beta(v)\,y + \gamma(v)\,z + \delta(v) = 0.$$

CHAPTER 6

Surface Theory

In this chapter we continue the study of curvature with the aim of proving several profound results for surfaces that also involve more global considerations. The highlight being the local and global Gauss-Bonnet theorems. This chapter also introduces abstract surfaces that might not come with a suitable second fundamental form. We also explain the Codazzi equations and establish the fundamental theorem for surfaces.

6.1. Generalized and Abstract Surfaces

It is possible to work with *generalized surfaces* in Euclidean spaces of arbitrary dimension: $\mathrm{q}(u,v) : U \to \mathbb{R}^k$ for any $k \geq 2$. What changes is that we no longer have a single normal vector n. In fact for $k \geq 4$ there will be a whole family of normal vectors, not unlike what happened for space curves. What all of these surfaces do have in common is that we can define the first fundamental form. Thus we can also calculate the Christoffel symbols using the formulas in terms of derivatives of the first fundamental form. This leads us to the possibility of an abstract definition of a surface that is independent of a particular map into a coordinate space $\mathbb{R}^k$.

One of the simplest examples of a generalized surface is the *flat* torus in $\mathbb{R}^4$. It is parametrized by

$$\mathrm{q}(u,v) = (\cos u, \sin u, \cos v, \sin v)$$

and its first fundamental form is

$$\mathrm{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

So this yields a Cartesian parametrization of the entire torus. This is why it is called the flat torus. It is in fact not possible for a closed surface in $\mathbb{R}^3$ to be flat everywhere (see section 6.6).

An *abstract parametrized surface* consists of a domain $U \subset \mathbb{R}^2$ and a first fundamental form

$$[\mathrm{I}] = \begin{bmatrix} g_{uu} & g_{uv} \\ g_{uv} & g_{vv} \end{bmatrix},$$

where g_{uu}, g_{vv}, and g_{uv} are functions on U. The inner product of vectors $X = (X^u, X^v)$ and $Y = (Y^u, Y^v)$ thought of as having the same base point $p \in U$ is defined as

$$\mathrm{I}(X,Y) = \begin{bmatrix} X^u & X^v \end{bmatrix} \begin{bmatrix} g_{uu}(p) & g_{uv}(p) \\ g_{vu}(p) & g_{vv}(p) \end{bmatrix} \begin{bmatrix} Y^u \\ Y^v \end{bmatrix}.$$

For this to give us an inner product we also have to make sure that it is positive definite, i.e., for $X \neq 0$

$$\begin{aligned} 0 &< \mathrm{I}(X,X) \\ &= \begin{bmatrix} X^u & X^v \end{bmatrix} \begin{bmatrix} g_{uu} & g_{uv} \\ g_{uv} & g_{vv} \end{bmatrix} \begin{bmatrix} X^u \\ X^v \end{bmatrix} \\ &= X^u X^u g_{uu} + 2X^u X^v g_{uv} + X^v X^v g_{vv}. \end{aligned}$$

PROPOSITION 6.1.1. I *is positive definite if and only if* $\mathrm{tr}\,[\mathrm{I}] = g_{uu} + g_{vv} > 0$ *and* $\det[\mathrm{I}] = g_{uu}g_{vv} - (g_{uv})^2 > 0$.

PROOF. If I is positive definite, then it follows that g_{uu} and g_{vv} are positive by letting $X = (1,0)$ and $(0,1)$. Next use $X = \left(\sqrt{g_{vv}}, \pm\sqrt{g_{uu}}\right)$ to get

$$0 < \mathrm{I}(X,X) = 2g_{uu}g_{vv} \pm 2\sqrt{g_{uu}}\sqrt{g_{vv}}g_{uv}.$$

Thus

$$\pm g_{uv} < \sqrt{g_{uu}}\sqrt{g_{vv}}$$

showing that

$$g_{uu}g_{vv} > (g_{uv})^2.$$

To check that I is positive definite when $g_{uu} + g_{vv}$, and $g_{uu}g_{vv} - (g_{uv})^2$ are positive we start by observing that

$$g_{uu}g_{vv} > g_{uv}^2 \geq 0.$$

Thus g_{uu} and g_{vv} have the same sign. As their sum is positive both terms are positive. It then follows that

$$\begin{aligned} \mathrm{I}(X,X) &= X^u X^u g_{uu} + 2X^u X^v g_{uv} + X^v X^v g_{vv} \\ &\geq X^u X^u g_{uu} - 2|X^u||X^v|\sqrt{g_{uu}g_{vv}} + X^v X^v g_{vv} \\ &= \left(|X^u|\sqrt{g_{uu}} - |X^v|\sqrt{g_{vv}}\right)^2 \\ &\geq 0. \end{aligned}$$

Here first inequality is in fact $>$ unless $X^u = 0$ or $X^v = 0$. In case $X^u = 0$ we obtain

$$\mathrm{I}(X,X) = (X^v)^2 g_{vv} > 0$$

unless also $X^v = 0$. □

EXAMPLE 6.1.2. The hyperbolic space $H \subset \mathbb{R}^{2,1}$ is defined as the imaginary unit sphere with $z > 0$, specifically it is the rotationally symmetric surface

$$x^2 + y^2 - z^2 = -1,\ z \geq 1$$

or equivalently the Monge patch

$$z = \sqrt{1 + x^2 + y^2}.$$

The metric on this surface, however, is inherited from a different inner product structure on $\mathbb{R}^3$ which is why we use the notation $\mathbb{R}^{2,1}$. Specifically:

$$X \cdot Y = X^x Y^x + X^y Y^y - X^z Y^z.$$

The x- and y-coordinates are the "space" part and the z-coordinate the "time" part. We say that a vector is *space-like*, *null*, or *time-like* if $|X|^2 = X \cdot X$ is positive, zero, or negative. Thus $(x, y, 0)$ is space-like while $(0, 0, z)$ is time-like. Null vectors satisfy the equation

$$X^x X^x + X^y X^y - X^z X^z = 0.$$

This describes a cone. The two insides of this cone consist of the time-like vectors, while the outside contains the space-like vectors.

Our surface H given by the equation

$$F(x, y, z) = x^2 + y^2 - z^2 = -1, \ z \geq 1$$

therefore consists of time-like points. However, all of the tangent spaces consist of space-like vectors. This means that we obtain a surface with a valid first fundamental form. In the Monge patch representation we have

$$\frac{\partial z}{\partial x} = \frac{x}{\sqrt{1 + x^2 + y^2}} = \frac{x}{z}, \ \frac{\partial z}{\partial y} = \frac{y}{\sqrt{1 + x^2 + y^2}} = \frac{y}{z}.$$

Thus the tangent space at $q = (x, y, z) = \left(x, y, \sqrt{1 + x^2 + y^2}\right)$ is given by

$$\begin{aligned} T_q H &= \operatorname{span}\left\{\left(1, 0, \frac{x}{z}\right), \left(0, 1, \frac{y}{z}\right)\right\} \\ &= \left\{X^x \left(1, 0, \frac{x}{z}\right) + X^y \left(0, 1, \frac{y}{z}\right) \mid X^x, X^y \in \mathbb{R}\right\}. \end{aligned}$$

Consequently

$$\begin{aligned} X \cdot X &= (X^x)^2 + (X^y)^2 - \left(X^x \frac{x}{z} + X^y \frac{y}{z}\right)^2 \\ &= (X^x)^2 \left(1 - \frac{x^2}{z^2}\right) + (X^y)^2 \left(1 - \frac{y^2}{z^2}\right) \\ &\quad -2X^x X^y \frac{xy}{z^2} \\ &= (X^x)^2 \frac{1 + y^2}{z^2} + (X^y)^2 \frac{1 + x^2}{z^2} - 2X^x X^y \frac{xy}{z^2} \\ &= \frac{1}{z^2}\left((X^x)^2 + (X^y)^2 + (yX^x - xX^y)^2\right). \end{aligned}$$

This is clearly positive unless $X = 0$. The first fundamental form is

$$\begin{bmatrix} 1 - \frac{x^2}{z^2} & -\frac{xy}{z^2} \\ -\frac{xy}{z^2} & 1 - \frac{y^2}{z^2} \end{bmatrix}$$

which is also easily checked to be positive using proposition 6.1.1.

In order to find a nicer expression of the first fundamental form we switch to a surface of revolution parametrization

$$\mathrm{q}(\phi, \mu) = \begin{bmatrix} \cos\mu \sinh\phi \\ \sin\mu \sinh\phi \\ \cosh\phi \end{bmatrix}, \ \mu \in \mathbb{R}, \ \phi > 0,$$

where $\phi = 0$ corresponds to the point $(0, 0, 1)$ which we can think of as a pole. In this parametrization we obtain

$$\partial_\phi \mathrm{q} = \frac{\partial \mathrm{q}}{\partial \phi} = \begin{bmatrix} \cos\mu \cosh\phi \\ \sin\mu \cosh\phi \\ \sinh\phi \end{bmatrix}, \ \partial_\mu \mathrm{q} = \frac{\partial \mathrm{q}}{\partial \mu} = \begin{bmatrix} \sin\mu \sinh\phi \\ -\cos\mu \sinh\phi \\ 0 \end{bmatrix}$$

which gives us the first fundamental form

$$\begin{bmatrix} \partial_\phi \mathrm{q} \cdot \partial_\phi \mathrm{q} & \partial_\phi \mathrm{q} \cdot \partial_\mu \mathrm{q} \\ \partial_\mu \mathrm{q} \cdot \partial_\phi \mathrm{q} & \partial_\mu \mathrm{q} \cdot \partial_\mu \mathrm{q} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sinh^2 \phi \end{bmatrix}.$$

REMARK 6.1.3. It is not possible for a surface of revolution to have this first fundamental form in $\mathbb{R}^3$. But we shall see later that the pseudo-sphere from section 5.4 exercise 9 is a local Euclidean model that is locally isometric to H.

On the other hand a theorem of Hilbert (see theorem 6.3.6) shows that one cannot represent the entire surface H in $\mathbb{R}^3$, i.e., there is no parametrization $\mathrm{q}(x, y) : H \to \mathbb{R}^3$ defined for all $(x, y) \in \mathbb{R}^2$ such that

$$\begin{bmatrix} \partial_x \mathrm{q} & \partial_y \mathrm{q} \end{bmatrix}^t \begin{bmatrix} \partial_x \mathrm{q} & \partial_y \mathrm{q} \end{bmatrix} = \begin{bmatrix} \partial_x \mathrm{q} \cdot \partial_x \mathrm{q} & \partial_x \mathrm{q} \cdot \partial_y \mathrm{q} \\ \partial_y \mathrm{q} \cdot \partial_x \mathrm{q} & \partial_y \mathrm{q} \cdot \partial_y \mathrm{q} \end{bmatrix} = \begin{bmatrix} 1 - \frac{x^2}{z^2} & -\frac{xy}{z^2} \\ -\frac{xy}{z^2} & 1 - \frac{y^2}{z^2} \end{bmatrix}.$$

Janet showed that if the metric coefficients of an abstract surface are analytic, then one can always locally represent the abstract surface in $\mathbb{R}^3$. Nash showed that any abstract surface can be represented by a map $\mathrm{q}(u, v) : U \to \mathbb{R}^k$ on the entire domain, but only at the expense of making k very large. Based in part on Nash's work Greene and Gromov both showed that one can always locally represent an abstract surface in $\mathbb{R}^5$. It is still unknown if there exists an abstract surface that cannot be locally realized as a surface in $\mathbb{R}^3$.

DEFINITION 6.1.4. We say that a surface $M \subset \mathbb{R}^{2,1}$ is *space-like* if all tangent vectors are space-like. This means that if we use the first fundamental form that comes from the inner product in $\mathbb{R}^{2,1}$, then we obtain an abstract surface.

REMARK 6.1.5. Space-like surfaces $\mathrm{q}(u, v) : U \to \mathbb{R}^{2,1}$ also have a normal n, but it has the property that $|\mathrm{n}|^2 = \mathrm{n} \cdot \mathrm{n} = -1$ as well as the usual conditions: $\mathrm{n} \cdot \frac{\partial \mathrm{q}}{\partial u} = 0 = \mathrm{n} \cdot \frac{\partial \mathrm{q}}{\partial v}$. However, n cannot be calculated as easily from the standard vector calculus cross product $\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}$. The projection formulas will also look a little different. If we focus on a curve $\mathrm{q}(t)$ in this surface, then we still have

$$\dot{\mathrm{q}} = \frac{d\,\mathrm{q}}{dt} = \frac{d\,\mathrm{q}}{dt} = \partial_u \mathrm{q} \frac{du}{dt} + \partial_v \mathrm{q} \frac{dv}{dt} = \begin{bmatrix} \partial_u \mathrm{q} & \partial_v \mathrm{q} \end{bmatrix} \begin{bmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{bmatrix}$$

since this doesn't depend on any geometric structure. The acceleration however, now decomposes as

$$\begin{aligned} \ddot{\mathrm{q}} &= \ddot{\mathrm{q}}^{\mathrm{I}} + \ddot{\mathrm{q}}^{\mathrm{II}} \\ &= \begin{bmatrix} \partial_u \mathrm{q} & \partial_v \mathrm{q} \end{bmatrix} [\mathrm{I}]^{-1} \begin{bmatrix} \partial_u \mathrm{q} & \partial_v \mathrm{q} \end{bmatrix}^t \ddot{\mathrm{q}} - (\ddot{\mathrm{q}} \cdot \mathrm{n})\,\mathrm{n}, \end{aligned}$$

where $\ddot{\mathrm{q}}^{\mathrm{I}}$ is tangent to the surface and $\ddot{\mathrm{q}}^{\mathrm{II}}$ proportional to n. Note that all products are space-time inner products. The negative sign on the normal component is easier to understand if we remember that the formula for projecting a vector X onto another vector N is given by

$$\frac{X \cdot N}{N \cdot N} N.$$

This formula remains valid in space-time. The tangential part of the acceleration can also be calculated intrinsically with the same formula as before:

$$\ddot{\mathrm{q}}^{\mathrm{I}} = \begin{bmatrix} \partial_u \mathrm{q} & \partial_v \mathrm{q} \end{bmatrix} [\mathrm{I}]^{-1} \begin{bmatrix} \partial_u \mathrm{q} & \partial_v \mathrm{q} \end{bmatrix}^t \ddot{\mathrm{q}} = \partial_u \mathrm{q} \left(\frac{d^2 u}{dt^2} + \Gamma^u(\dot{\mathrm{q}}, \dot{\mathrm{q}}) \right) + \partial_v \mathrm{q} \left(\frac{d^2 v}{dt^2} + \Gamma^v(\dot{\mathrm{q}}, \dot{\mathrm{q}}) \right).$$

Finally, we also have to define what we mean by an *abstract surface*. There are several competing definitions. The more general and abstract ones unfortunately also have a very steep learning curve before a metric can be introduced. So we stay with the more classical context. Essentially we define a surface as a set of points where we can use the language of first fundamental form, convergence etc. This is generally too vague for modern mathematicians but at least allows us to move on to the issues that are relevant in differential geometry. There are several other standard concepts included in this definition so as to have everything in one place

DEFINITION 6.1.6. A *surface* with a first fundamental form is a space M where we can work locally as if it is an abstract parametrized surface, i.e., every point is included in a parametrization $\mathrm{q} : U \subset \mathbb{R}^2 \to M$. When a point $q \in M$ is covered by more than one parametrization, then they are pairwise reparametrizations of each other near q and the first fundamental forms are the same via this reparametrization. Globally we are allowed to talk about convergence of sequences as we do in $\mathbb{R}^2$. A sequence converges to q if eventually it lies in a parametrization around q and converges to q in that parametrization. Moreover, if the sequence eventually lies in more than one parametrization then its limit will be q in each of these parametrizations. This allows us to talk about continuous maps $F : M \to \mathbb{R}^k$ and $F : \mathbb{R}^l \to M$. Such a map is *smooth* if it smooth within the given parametrizations. Finally we want the surface to be path connected in the sense that any two points are joined by a piecewise smooth curve.

A surface is said to be *closed* if it is compact, i.e., any sequence has a convergent subsequence.

A surface M is said to be *orientable* if the parametrizations can be chosen so that the differentials of all the reparametrizations have positive determinant, e.g., if $\mathrm{q}(u, v) = \mathrm{q}(u(s, t), v(s, t)) = \mathrm{q}(s, t)$, then

$$\det \begin{bmatrix} \frac{\partial u}{\partial s} & \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial t} \end{bmatrix} > 0, \ \det \begin{bmatrix} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{bmatrix} > 0.$$

Such a choice of parametrizations that cover all of M will be called an *orientation* for M. Note that this tells us that if we have tangent vectors $v, w \in T_pM$ that are not proportional, then w either lies to the right or left of v.

The *tangent space* T_qM at a point $q \in M$ in a parametrization is defined as $T_qM = \operatorname{span}\{\partial_u \mathrm{q}, \partial_v \mathrm{q}\}$. In a different parametrization the two bases are related by

$$\begin{aligned} \begin{bmatrix} \partial_u \mathrm{q} & \partial_v \mathrm{q} \end{bmatrix} &= \begin{bmatrix} \partial_s \mathrm{q} & \partial_t \mathrm{q} \end{bmatrix} \begin{bmatrix} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{bmatrix}, \\ \begin{bmatrix} \partial_s \mathrm{q} & \partial_t \mathrm{q} \end{bmatrix} &= \begin{bmatrix} \partial_u \mathrm{q} & \partial_v \mathrm{q} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial s} & \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial t} \end{bmatrix}. \end{aligned}$$

A tangent vector $X \in T_qM$ can thus be written

$$\begin{aligned}
X &= X^u \partial_u \mathrm{q} + X^v \partial_v \mathrm{q} \\
&= \begin{bmatrix} \partial_u \mathrm{q} & \partial_v \mathrm{q} \end{bmatrix} \begin{bmatrix} X^u \\ X^v \end{bmatrix} \\
&= \begin{bmatrix} \partial_s \mathrm{q} & \partial_t \mathrm{q} \end{bmatrix} \begin{bmatrix} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{bmatrix} \begin{bmatrix} X^u \\ X^v \end{bmatrix} \\
&= \begin{bmatrix} \partial_s \mathrm{q} & \partial_t \mathrm{q} \end{bmatrix} \begin{bmatrix} X^u \frac{\partial s}{\partial u} + X^v \frac{\partial s}{\partial v} \\ X^u \frac{\partial t}{\partial u} + X^v \frac{\partial t}{\partial v} \end{bmatrix} \\
&= \left(X^u \frac{\partial s}{\partial u} + X^v \frac{\partial s}{\partial v} \right) \partial_s \mathrm{q} + \left(X^u \frac{\partial t}{\partial u} + X^v \frac{\partial t}{\partial v} \right) \partial_t \mathrm{q} \\
&= X^s \partial_s \mathrm{q} + X^t \partial_t \mathrm{q}\,.
\end{aligned}$$

A surface is said to be *isometrically embedded* in $\mathbb{R}^3$ if it can be represented as a surface $M \subset \mathbb{R}^3$ in such a way that that the induced first fundamental form agrees with the abstract one on M. Specifically, we seek a map $F : M \to F(M) \subset \mathbb{R}^3$ such that F is a diffeomorphism from M to $F(M)$ and $\mathrm{I}_M(X,Y) = \mathrm{I}_{F(M)}(DF(X), DF(Y))$.

A surface is said to be *isometrically immersed* in $\mathbb{R}^3$ if there is a map $F : M \to \mathbb{R}^3$ such that $\mathrm{I}_M(X,Y) = \mathrm{I}_{F(M)}(DF(X), DF(Y))$. In this case F will be a local diffeomorphism onto its image, but globally it might not be one-to-one (see also 4.1.5).

Remark 6.1.7. In modern usage a surface does not necessarily come with a first fundamental form. We could have called our surfaces *Riemannian surfaces* (*Riemannian manifolds* are their higher dimensional analogues), but that too can be confused with *Riemann surfaces* which are surfaces where the reparametrizations are holomorphic, i.e., satisfy the Cauchy-Riemann equations.

Exercises

(1) Assume that $g_{uu} = g_{vv} = 1$ on a domain $U \subset \mathbb{R}^2$. Show that the corresponding first fundamental form represents an abstract surface if $|g_{uv}| < 1$ on U.
(2) Show that it is possible to define S and κ_g for unit speed curves in oriented abstract surfaces.
(3) Consider a regular space-like curve $\mathrm{q}(t) = (r(t), 0, z(t)) : I \to \mathbb{R}^{2,1}$, i.e., $\dot{\mathrm{q}}$ is non-zero and space-like everywhere.
    (a) Show that $\dot{r}^2 > \dot{z}^2$ along q.
    (b) Show that
    $$\mathrm{q}(t,\mu) = (r(t)\cos\mu, r(t)\sin\mu, z(t))$$
    defines a space-like surface of revolution with
    $$[\mathrm{I}] = \begin{bmatrix} \dot{r}^2 - \dot{z}^2 & 0 \\ 0 & r^2 \end{bmatrix}.$$
    (c) Show that when $r = R\cosh t$ and $z = R\sinh t$, then we obtain the surface described in exercise 3.

(d) Show that $\mathrm{q}(t) = (r(t), 0, z(t)) : I \to \mathbb{R}^{2,1}$ can be reparametrized to have unit speed, i.e.,

$$\left(\frac{dr}{ds}\right)^2 - \left(\frac{dz}{ds}\right)^2 = 1.$$

(4) Consider the space-like surface of revolution from exercise 3.

(a) Show that

$$\mathrm{n} = \pm \frac{(\dot{z}\cos\mu, \dot{z}\sin\mu, \dot{r})}{\sqrt{\dot{r}^2 - \dot{z}^2}}.$$

(b) Show that a curve $\mathrm{q}(t(s), \mu(s))$ on this surface has constant speed when $\left(\dot{r}^2 - \dot{z}^2\right)\left(\frac{dt}{ds}\right)^2 + r^2\left(\frac{d\mu}{ds}\right)^2$ is constant.

(c) Show that if $\dot{r}^2 - \dot{z}^2 = 1$, then a curve $\mathrm{q}(t(s), \mu(s))$ satisfies:

$$\ddot{\mathrm{q}}^{\mathrm{I}} = \left(\frac{d^2t}{ds^2} - r\dot{r}\left(\frac{d\mu}{ds}\right)^2\right)\begin{bmatrix} \dot{r}\cos\mu \\ \dot{r}\sin\mu \\ \dot{z} \end{bmatrix} + \left(\frac{d^2\mu}{ds^2} + 2\frac{\dot{r}}{r}\frac{dt}{ds}\frac{d\mu}{ds}\right)\begin{bmatrix} -r\sin\mu \\ r\cos\mu \\ 0 \end{bmatrix}.$$

6.2. Curvature on Abstract Surfaces

The goal of this section is to define Christoffel symbols and curvature on abstract surfaces. We assume throughout that we have an abstract parametrized surface on a domain $U \subset \mathbb{R}^2$ with first fundamental form:

$$[\mathrm{I}] = \begin{bmatrix} g_{uu} & g_{vu} \\ g_{uv} & g_{vv} \end{bmatrix}.$$

It will be convenient to use indices i, j, k, l, s, t, r to denote the two specific indices u, v. Thus we can write g_{ij} for a generic entry in $[\mathrm{I}]$. This notation can conveniently be extended to partial derivatives

$$\partial_i F = \frac{\partial F}{\partial i},$$

where again i can be u or v.

Definition 6.2.1. The Christoffel symbols of the first and second kind are defined as follows:

$$\begin{aligned} \Gamma_{ijk} &= \frac{1}{2}\left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right), \\ \Gamma_{ij}^k &= \Gamma_{iju}g^{uk} + \Gamma_{ijv}g^{vk} = \sum_{s=u,v} \Gamma_{ijs}g^{sk}. \end{aligned}$$

It is not immediately clear that this definition agrees with with proposition 5.2.11.

PROPOSITION 6.2.2. *The Christoffel symbols of the first kind satisfy*

$$\begin{aligned}
\Gamma_{uuu} &= \frac{1}{2}\frac{\partial g_{uu}}{\partial u},\\
\Gamma_{uvu} &= \frac{1}{2}\frac{\partial g_{uu}}{\partial v} = \Gamma_{vuu},\\
\Gamma_{vvv} &= \frac{1}{2}\frac{\partial g_{vv}}{\partial v},\\
\Gamma_{uvv} &= \frac{1}{2}\frac{\partial g_{vv}}{\partial u} = \Gamma_{vuv},\\
\Gamma_{uuv} &= \frac{\partial g_{uv}}{\partial u} - \frac{1}{2}\frac{\partial g_{uu}}{\partial v},\\
\Gamma_{vvu} &= \frac{\partial g_{uv}}{\partial v} - \frac{1}{2}\frac{\partial g_{vv}}{\partial u}
\end{aligned}$$

as well as the "product" rule

$$\partial_k g_{ij} = \Gamma_{kij} + \Gamma_{kji}.$$

PROOF. First observe that as $g_{ij} = g_{ji}$ we have that

$$\Gamma_{ijk} = \frac{1}{2}\left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right) = \frac{1}{2}\left(\partial_j g_{ik} + \partial_i g_{jk} - \partial_k g_{ji}\right) = \Gamma_{jik}.$$

Next note that, e.g.,

$$\Gamma_{uuu} = \frac{1}{2}\left(\partial_u g_{uu} + \partial_u g_{uu} - \partial_u g_{uu}\right) = \frac{1}{2}\partial_u g_{uu},$$

$$\Gamma_{uvu} = \Gamma_{vuu} = \frac{1}{2}\left(\partial_v g_{uu} + \partial_u g_{uv} - \partial_u g_{vu}\right) = \frac{1}{2}\partial_v g_{uu},$$

and

$$\Gamma_{uuv} = \frac{1}{2}\left(\partial_u g_{uv} + \partial_u g_{uv} - \partial_v g_{uu}\right) = \partial_u g_{uv} - \frac{1}{2}\partial_v g_{uu}.$$

The proofs of the other 4 equations are identical if we replace u by v and v by u.

Finally, note that the equations that define the Christoffel symbols are linear combinations of the derivatives of the entries in the first fundamental form. The last claim says that we can solve for these derivatives in terms of the Christoffel symbols:

$$\begin{aligned}
\Gamma_{kij} + \Gamma_{kji} &= \frac{1}{2}\left(\partial_k g_{ij} + \partial_i g_{kj} - \partial_j g_{ik}\right) + \frac{1}{2}\left(\partial_k g_{ji} + \partial_j g_{ki} - \partial_i g_{jk}\right)\\
&= \frac{1}{2}\left(\partial_k g_{ij} + \partial_i g_{kj} - \partial_j g_{ik}\right) + \frac{1}{2}\left(\partial_k g_{ij} + \partial_j g_{ik} - \partial_i g_{kj}\right)\\
&= \partial_k g_{ij}.
\end{aligned}$$

□

The next goal is to define the Gauss curvature. Recall from the proof of theorem 5.3.8 that

$$K = \frac{\partial_v \Gamma_{uuv} - \partial_u \Gamma_{vuv} + \left(\Gamma_{uu}^u \frac{\partial \mathrm{q}}{\partial u} + \Gamma_{uu}^v \frac{\partial \mathrm{q}}{\partial v}\right) \cdot \left(\Gamma_{vv}^u \frac{\partial \mathrm{q}}{\partial u} + \Gamma_{vv}^v \frac{\partial \mathrm{q}}{\partial v}\right) - \left|\Gamma_{uv}^u \frac{\partial \mathrm{q}}{\partial u} + \Gamma_{uv}^v \frac{\partial \mathrm{q}}{\partial v}\right|^2}{\det[\mathrm{I}]}.$$

By using that $g_{ij} = \partial_i \mathrm{q} \cdot \partial_j \mathrm{q}$ this can be compressed to the formula

$$K = \frac{\partial_v \Gamma_{uuv} - \partial_u \Gamma_{vuv} + \sum_{s,t=u,v} g_{st}\left(\Gamma_{uu}^s \Gamma_{vv}^t - \Gamma_{uv}^s \Gamma_{uv}^t\right)}{\det[\mathrm{I}]}.$$

Definition 6.2.3. The Riemann curvature tensors are defined as

$$R_{ijkl} = \partial_i \Gamma_{jkl} - \partial_j \Gamma_{ikl} + \sum_{s,t=u,v} g_{st} \left(\Gamma^s_{ik} \Gamma^t_{jl} - \Gamma^s_{il} \Gamma^t_{jk} \right)$$

and

$$R^l_{ijk} = \sum_{s=u,v} R_{ijks} g^{sl}.$$

In particular,

$$K = \frac{R_{vuuv}}{\det[\mathrm{I}]}.$$

Proposition 6.2.4. *The Riemann curvature tensors satisfy the symmetry properties*

$$R_{ijkl} = -R_{jikl} = R_{jilk},$$

in particular

$$R_{iikl} = R_{ijkk} = 0,$$

and the possibly nontrivial terms

$$R_{uvvu} = -R_{vuvu} = R_{vuuv} = -R_{uvuv} = K \det[\mathrm{I}].$$

All in all

$$R_{ijkl} = K \left(g_{il} g_{jk} - g_{ik} g_{jl} \right),$$

and

$$R^l_{ijk} = K \left(\delta^l_i g_{jk} - \delta^l_j g_{ik} \right).$$

Proof. First note that

$$\begin{aligned} R_{ijkl} &= \partial_i \Gamma_{jkl} - \partial_j \Gamma_{ikl} + \sum_{s,t=u,v} g_{st} \left(\Gamma^s_{ik} \Gamma^t_{jl} - \Gamma^s_{il} \Gamma^t_{jk} \right) \\ &= -\left(\partial_j \Gamma_{ikl} - \partial_i \Gamma_{jkl} \right) - \sum_{s,t=u,v} g_{st} \left(\Gamma^s_{jk} \Gamma^t_{il} - \Gamma^s_{jl} \Gamma^t_{ik} \right) \\ &= -R_{jikl}. \end{aligned}$$

Next we have

$$\begin{aligned} R_{ijkl} + R_{ijlk} &= \partial_i \Gamma_{jkl} - \partial_j \Gamma_{ikl} + \partial_i \Gamma_{jlk} - \partial_j \Gamma_{ilk} \\ &\quad + \sum_{s,t=u,v} g_{st} \left(\Gamma^s_{ik} \Gamma^t_{jl} - \Gamma^s_{il} \Gamma^t_{jk} \right) - \sum_{s,t=u,v} g_{st} \left(\Gamma^s_{il} \Gamma^t_{jk} - \Gamma^s_{ik} \Gamma^t_{jl} \right) \\ &= \partial_i \Gamma_{jkl} - \partial_j \Gamma_{ikl} + \partial_i \Gamma_{jlk} - \partial_j \Gamma_{ilk} \\ &= \partial_i \left(\Gamma_{jkl} + \Gamma_{jlk} \right) - \partial_j \left(\Gamma_{ikl} + \Gamma_{ilk} \right) \\ &= \partial_i \partial_j g_{kl} - \partial_j \partial_i g_{kl} \\ &= 0. \end{aligned}$$

To establish the penultimate claim note that the expression on the right hand side has the same skew-symmetry properties we just established for R_{ijkl}. Thus it suffices to check these equations in case, say, $v = i = l$ and $u = j = k$. In this case this simply becomes our new definition for the Gauss curvature.

To prove the last claim first note that the matrix identity

$$\begin{bmatrix} g_{uu} & g_{vu} \\ g_{uv} & g_{vv} \end{bmatrix} \begin{bmatrix} g^{uu} & g^{vu} \\ g^{uv} & g^{vv} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is equivalent to saying

$$\sum g_{is}g^{sj} = \delta_i^j = \begin{cases} 1 & \text{when } i = j, \\ 0 & \text{when } i \neq j. \end{cases}$$

With that in mind we obtain

$$\begin{aligned} R_{ijk}^l &= \sum_{s=u,v} K\left(g_{is}g_{jk} - g_{ik}g_{js}\right) g^{sl} \\ &= K\left(\delta_i^l g_{jk} - g_{ik}\delta_j^l\right). \end{aligned}$$

□

The curvature terms R_{ijk}^l will appear again in the next section in the form presented in the next proposition.

PROPOSITION 6.2.5. *We have that*

$$\begin{aligned} R_{ijk}^l &= \partial_i\Gamma_{jk}^l - \partial_j\Gamma_{ik}^l + \begin{bmatrix} \Gamma_{iu}^l & \Gamma_{iv}^l \end{bmatrix} \begin{bmatrix} \Gamma_{jk}^u \\ \Gamma_{jk}^v \end{bmatrix} - \begin{bmatrix} \Gamma_{ju}^l & \Gamma_{jv}^l \end{bmatrix} \begin{bmatrix} \Gamma_{ik}^u \\ \Gamma_{ik}^v \end{bmatrix} \\ &= \partial_i\Gamma_{jk}^l - \partial_j\Gamma_{ik}^l + \sum_{s=u,v}\left(\Gamma_{is}^l\Gamma_{jk}^s - \Gamma_{js}^l\Gamma_{ik}^s\right). \end{aligned}$$

PROOF. We first differentiate the equation $\sum_{s=u,v} g_{is}g^{sj} = \delta_i^j$ to obtain

$$\sum_{s=u,v}\left(g^{sj}\partial_k g_{is} + g_{is}\partial_k g^{sj}\right) = 0.$$

This is used to obtain the fifth equality below:

$$
\begin{aligned}
R_{ijk}^{l} &= \sum_{r=u,v} g^{rl} R_{ijkr} \\
&= \sum_{r=u,v} g^{rl} \left(\partial_i \Gamma_{jkr} - \partial_j \Gamma_{ikr}\right) + \sum_{r,s,t=u,v} g^{rl} g_{st} \left(\Gamma_{ik}^{s} \Gamma_{jr}^{t} - \Gamma_{jk}^{s} \Gamma_{ir}^{t}\right) \\
&= \sum_{r=u,v} \left(\partial_i \left(g^{rl} \Gamma_{jkr}\right) - \partial_j \left(g^{rl} \Gamma_{ikr}\right)\right) + \sum_{r,s=u,v} g^{rl} \left(\Gamma_{ik}^{s} \Gamma_{jrs} - \Gamma_{jk}^{s} \Gamma_{irs}\right) \\
&\quad - \sum_{r=u,v} \left(\Gamma_{jkr} \partial_i \left(g^{rl}\right) - \Gamma_{ikr} \partial_j \left(g^{rl}\right)\right) \\
&= \partial_i \Gamma_{jk}^{l} - \partial_j \Gamma_{ik}^{l} + \sum_{r,s=u,v} g^{rl} \left(\Gamma_{ik}^{s} \Gamma_{jrs} - \Gamma_{jk}^{s} \Gamma_{irs}\right) \\
&\quad - \sum_{r,s=u,v} \left(\Gamma_{jk}^{s} g_{sr} \partial_i \left(g^{rl}\right) - \Gamma_{ik}^{s} g_{sr} \partial_j \left(g^{rl}\right)\right) \\
&= \partial_i \Gamma_{jk}^{l} - \partial_j \Gamma_{ik}^{l} + \sum_{r,s=u,v} g^{rl} \left(\Gamma_{ik}^{s} \Gamma_{jrs} - \Gamma_{jk}^{s} \Gamma_{irs}\right) \\
&\quad + \sum_{r,s=u,v} g^{rl} \left(\Gamma_{jk}^{s} \partial_i g_{sr} - \Gamma_{ik}^{s} \partial_j g_{sr}\right) \\
&= \partial_i \Gamma_{jk}^{l} - \partial_j \Gamma_{ik}^{l} + \sum_{r,s=u,v} g^{rl} \left(\Gamma_{ik}^{s} \Gamma_{jrs} - \Gamma_{jk}^{s} \Gamma_{irs}\right) \\
&\quad + \sum_{r,s=u,v} g^{rl} \left(\Gamma_{jk}^{s} \left(\Gamma_{isr} + \Gamma_{irs}\right) - \Gamma_{ik}^{s} \left(\Gamma_{jsr} + \Gamma_{jrs}\right)\right) \\
&= \partial_i \Gamma_{jk}^{l} - \partial_j \Gamma_{ik}^{l} + \sum_{r,s=u,v} g^{rl} \left(\Gamma_{jk}^{s} \Gamma_{isr} - \Gamma_{ik}^{s} \Gamma_{jsr}\right) \\
&= \partial_i \Gamma_{jk}^{l} - \partial_j \Gamma_{ik}^{l} + \sum_{s=u,v} \left(\Gamma_{jk}^{s} \Gamma_{is}^{l} - \Gamma_{ik}^{s} \Gamma_{js}^{l}\right).
\end{aligned}
$$

$\square$

Exercises

(1) Show that

$$
\begin{aligned}
R_{ijkl} &= \sum_{s=u,v} R_{ijk}^{s} g_{sl}, \\
\Gamma_{ijk} &= \sum_{s=u,v} \Gamma_{ij}^{s} g_{sk}.
\end{aligned}
$$

(2) Show that

$$
\begin{aligned}
\partial_k g^{ij} &= -\sum_{s,t=u,v} g^{si} g^{tj} \partial_k g_{st} \\
&= -\sum_{s=u,v} g^{si} \Gamma_{ks}^{j} - \sum_{t=u,v} g^{tj} \Gamma_{kt}^{i}.
\end{aligned}
$$

(3) Show that the surfaces in $\mathbb{R}^{2,1}$ given by the equation

$$x^2 + y^2 - z^2 = -R^2$$

have constant Gauss curvature $-R^{-2}$.

(4) This is an extension of exercise 3 from section 6.1. Show that when $\mathrm{q}(t)$ has been reparametrized to have unit speed then the Gauss curvature is given by

$$K = -\frac{\frac{d^2 r}{ds^2}}{r}.$$

(5) Assume the first fundamental form is given by the conditions that $g_{uu} = g_{vv} = 1$ and $g_{uv} = \cos\theta$, where $\theta : U \to \mathbb{R}$. Show that

$$\begin{aligned}
\Gamma_{uvw} &= \Gamma_{uuu} = \Gamma_{vvv} = 0, \\
\Gamma_{uuv} &= -\frac{\partial\theta}{\partial u}\sin\theta, \\
\Gamma_{vvu} &= -\frac{\partial\theta}{\partial v}\sin\theta, \\
\frac{\partial^2\theta}{\partial u\partial v} &= -K\sin\theta.
\end{aligned}$$

(6) Show for a generalized parametrized surface $\mathrm{q}(u,v) : U \to \mathbb{R}^n$ the Christoffel symbols can be defined as in section5.2

$$\Gamma_{ijk} = \partial^2_{ij}\,\mathrm{q}\cdot\partial_k\,\mathrm{q}.$$

(7) Assume that a parametrized surface $\mathrm{q} : U \to \mathbb{R}^n$ has a first fundamental form where $g_{uu} = g_{vv} = 1$ on U. Show that $\partial^2_{uv}\,\mathrm{q} = \frac{\partial^2\mathrm{q}}{\partial u\partial v}$ is perpendicular to the surface. Hint: Use the previous exercise.

(8) Assume that an abstract parametrized surface $\mathrm{q}(u,v)$ has first fundamental form

$$[\mathrm{I}] = \begin{bmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{bmatrix}.$$

(a) Show that the Gauss curvature satisfies

$$K = -\frac{\Delta\ln\lambda}{\lambda^2} = -\frac{\frac{\partial^2\ln\lambda}{\partial u^2} + \frac{\partial^2\ln\lambda}{\partial v^2}}{\lambda^2}.$$

(b) Show that if

$$\lambda = \frac{1}{a\left(u^2+v^2\right) + b_u u + b_v v + c}$$

for constants a, b_u, b_v, c, then

$$K = 4ac - b_u^2 - b_v^2.$$

(c) Consider the reparametrization

$$(u,v) = e^r(\cos\theta, \sin\theta)$$

and show that in (r,θ) parameters the first fundamental form looks like

$$[\mathrm{I}] = \begin{bmatrix} e^{2r}\lambda^2 & 0 \\ 0 & e^{2r}\lambda^2 \end{bmatrix}.$$

Thus this is a conformal reparametrization.

(d) With λ as in (b) show that in (r,θ) parameters the conformal factor is given by

$$e^r\lambda = \frac{1}{ae^r + b_u\cos\theta + b_v\sin\theta + e^{-r}c}.$$

6.3. The Gauss and Codazzi Equations

The goal in this section is to establish the classical Gauss equation and the accompanying Codazzi equations from the Gauss formulas and Weingarten equations. The Codazzi equations were historically first discovered by K.M. Peterson in 1853, then rediscovered by G. Mainardi in 1856, and then finally by D. Codazzi in 1867.

Recall from section 5.2 the Gauss formulas and Weingarten equations in combined form:

$$\frac{\partial}{\partial w}\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right]=\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right][D_w].$$

Taking one more derivative on both sides yields

$$\begin{aligned}\frac{\partial^2}{\partial u\partial v}\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right] &= \left(\frac{\partial}{\partial u}\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right]\right)[D_v]\\ &\quad+\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right]\left(\frac{\partial}{\partial u}[D_v]\right)\\ &= \left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right][D_u][D_v]\\ &\quad+\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right]\left(\frac{\partial}{\partial u}[D_v]\right)\end{aligned}$$

and similarly

$$\frac{\partial^2}{\partial v\partial u}\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right]=\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right]\left([D_v][D_u]+\frac{\partial}{\partial v}[D_u]\right).$$

Using that

$$\frac{\partial^2}{\partial u\partial v}\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right]=\frac{\partial^2}{\partial v\partial u}\left[\begin{array}{ccc}\frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} & n\end{array}\right]$$

we obtain

$$\frac{\partial}{\partial u}[D_v]+[D_u][D_v]=\frac{\partial}{\partial v}[D_u]+[D_v][D_u].$$

Writing out the entries in the matrices this becomes

$$\begin{aligned}&\left[\begin{array}{ccc}\partial_u\Gamma_{vu}^u & \partial_u\Gamma_{vv}^u & -\partial_uL_v^u\\ \partial_u\Gamma_{vu}^v & \partial_u\Gamma_{vv}^v & -\partial_uL_v^v\\ \partial_uL_{vu} & \partial_uL_{vv} & 0\end{array}\right]+\left[\begin{array}{ccc}\Gamma_{uu}^u & \Gamma_{uv}^u & -L_u^u\\ \Gamma_{uu}^v & \Gamma_{uv}^v & -L_u^v\\ L_{uu} & L_{uv} & 0\end{array}\right]\left[\begin{array}{ccc}\Gamma_{vu}^u & \Gamma_{vv}^u & -L_v^u\\ \Gamma_{vu}^v & \Gamma_{vv}^v & -L_v^v\\ L_{vu} & L_{vv} & 0\end{array}\right]\\ &=\left[\begin{array}{ccc}\partial_v\Gamma_{uu}^u & \partial_v\Gamma_{uv}^u & -\partial_vL_u^u\\ \partial_v\Gamma_{uu}^v & \partial_v\Gamma_{uv}^v & -\partial_vL_u^v\\ \partial_vL_{uu} & \partial_vL_{uv} & 0\end{array}\right]+\left[\begin{array}{ccc}\Gamma_{vu}^u & \Gamma_{vv}^u & -L_v^u\\ \Gamma_{vu}^v & \Gamma_{vv}^v & -L_v^v\\ L_{vu} & L_{vv} & 0\end{array}\right]\left[\begin{array}{ccc}\Gamma_{uu}^u & \Gamma_{uv}^u & -L_u^u\\ \Gamma_{uu}^v & \Gamma_{uv}^v & -L_u^v\\ L_{uu} & L_{uv} & 0\end{array}\right].\end{aligned}$$

When restricting attention to the general terms of the entries in the first two columns and rows we obtain 4 equations for the partial derivatives of Γ_{ij}^k where each i,j,k can be u,v:

$$\partial_i\Gamma_{jk}^l+\left[\begin{array}{ccc}\Gamma_{iu}^l & \Gamma_{iv}^l & -L_i^l\end{array}\right]\left[\begin{array}{c}\Gamma_{jk}^u\\ \Gamma_{jk}^v\\ L_{jk}\end{array}\right]=\partial_j\Gamma_{ik}^l+\left[\begin{array}{ccc}\Gamma_{ju}^l & \Gamma_{jv}^l & -L_j^l\end{array}\right]\left[\begin{array}{c}\Gamma_{ik}^u\\ \Gamma_{ik}^v\\ L_{ik}\end{array}\right]$$

which can further be rearranged by isolating Γs on one side:

$$\partial_i\Gamma_{jk}^l-\partial_j\Gamma_{ik}^l+\left[\begin{array}{cc}\Gamma_{iu}^l & \Gamma_{iv}^l\end{array}\right]\left[\begin{array}{c}\Gamma_{jk}^u\\ \Gamma_{jk}^v\end{array}\right]-\left[\begin{array}{cc}\Gamma_{ju}^l & \Gamma_{jv}^l\end{array}\right]\left[\begin{array}{c}\Gamma_{ik}^u\\ \Gamma_{ik}^v\end{array}\right]=R_{ijk}^l=L_i^lL_{jk}-L_j^lL_{ik}.$$

These are called the *Gauss Equations*. Note that we only established these equations when $i = u$ and $j = v$. Clearly they also hold when $u = j$ and $v = i$ as both sides just change sign. They also hold trivially when $i = j$ as both sides vanish in that case. This means that the 4 original equations can be expanded to 16 equations where each of the 4 indices i, j, k, l can be both u, v.

EXAMPLE 6.3.1. It might be instructive to see what happens when $\mathrm{q}(u, v) : U \to \mathbb{R}^2$ is simply a reparametrization of the plane. In this case the derivatives have no normal component and we obtain

$$\frac{\partial}{\partial w}\left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{array}\right] = \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{array}\right]\left[\begin{array}{cc} \Gamma^u_{wu} & \Gamma^u_{wv} \\ \Gamma^v_{wu} & \Gamma^v_{wv} \end{array}\right] = \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{array}\right][D_w].$$

The Christoffel symbols tell us how the tangent fields change with respect to themselves. A good example comes from considering polar coordinates $\mathrm{q}(r, \theta) = (r\cos\theta, r\sin\theta)$ as in section 1.4. We have

$$\begin{aligned}
\frac{\partial}{\partial r}\left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial r} & \frac{\partial \mathrm{q}}{\partial \theta} \end{array}\right] &= \left[\begin{array}{cc} \frac{\partial^2 \mathrm{q}}{\partial r \partial r} & \frac{\partial^2 \mathrm{q}}{\partial r \partial \theta} \end{array}\right] = \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial r} & \frac{\partial \mathrm{q}}{\partial \theta} \end{array}\right]\left[\begin{array}{cc} 0 & 0 \\ 0 & \frac{1}{r} \end{array}\right], \\
\frac{\partial}{\partial \theta}\left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial r} & \frac{\partial \mathrm{q}}{\partial \theta} \end{array}\right] &= \left[\begin{array}{cc} \frac{\partial^2 \mathrm{q}}{\partial \theta \partial r} & \frac{\partial^2 \mathrm{q}}{\partial \theta \partial \theta} \end{array}\right] = \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial r} & \frac{\partial \mathrm{q}}{\partial \theta} \end{array}\right]\left[\begin{array}{cc} 0 & -r \\ \frac{1}{r} & 0 \end{array}\right].
\end{aligned}$$

Taking one more derivative in the Gauss formulas shows that:

$$[D_v][D_u] + \left[\frac{\partial D_u}{\partial v}\right] = [D_u][D_v] + \left[\frac{\partial D_v}{\partial u}\right]$$

or

$$\left[\frac{\partial D_v}{\partial u}\right] - \left[\frac{\partial D_u}{\partial v}\right] + [D_u][D_v] - [D_v][D_u] = 0.$$

These are in fact the integrability conditions for admitting Cartesian coordinates and, as we shall see, equivalent to saying that the the Gauss curvature vanishes. For polar coordinates the integrability conditions can be verified directly:

$$\begin{aligned}
\left[\frac{\partial D_r}{\partial \theta}\right] - \left[\frac{\partial D_\theta}{\partial r}\right] &= 0 - \left[\begin{array}{cc} 0 & -1 \\ -\frac{1}{r^2} & 0 \end{array}\right] = \left[\begin{array}{cc} 0 & 1 \\ \frac{1}{r^2} & 0 \end{array}\right], \\
[D_r][D_\theta] - [D_\theta][D_r] &= \left[\begin{array}{cc} 0 & 0 \\ 0 & \frac{1}{r} \end{array}\right]\left[\begin{array}{cc} 0 & -r \\ \frac{1}{r} & 0 \end{array}\right] - \left[\begin{array}{cc} 0 & -r \\ \frac{1}{r} & 0 \end{array}\right]\left[\begin{array}{cc} 0 & 0 \\ 0 & \frac{1}{r} \end{array}\right] \\
&= \left[\begin{array}{cc} 0 & 1 \\ \frac{1}{r^2} & 0 \end{array}\right].
\end{aligned}$$

All of the 16 Gauss equations can be reduced to a single equation.

THEOREM 6.3.2 (*Theorema Egregium*). *The Gauss equations are equivalent to the single equation:*

$$K \det[\mathrm{I}] = R_{uvvu}.$$

PROOF. We start by showing that the Gauss equations

$$R^l_{ijk} = L^l_i L_{jk} - L^l_j L_{ik}$$

are equivalent to the equations

$$R_{ijkl} = L_{il} L_{jk} - L_{jl} L_{ik}.$$

To see we simply use the relations

$$L_i^l = \sum_{s=u,v} g^{ls} L_{si} \text{ and } L_{ij} = \sum_{s=u,v} g_{is} L_j^s$$

$$R_{ijk}^l = \sum_{s=u,v} g^{ls} R_{ijks} \text{ and } R_{ijkl} = \sum_{s=u,v} g_{sl} R_{ijk}^l$$

to note that first of all

$$\begin{aligned}
\sum_{s=u,v} g^{sl} R_{ijks} &= R_{ijk}^l \\
&= L_i^l L_{jk} - L_j^l L_{ik} \\
&= \sum_{s=u,v} g^{ls} \left(L_{si} L_{jk} - L_{sj} L_{ik} \right)
\end{aligned}$$

and secondly

$$\begin{aligned}
L_{il} L_{jk} - L_{jl} L_{ik} &= \sum_{s=u,v} g_{ls} \left(L_i^s L_{jk} - L_j^s L_{ik} \right) \\
&= \sum_{s=u,v} g_{ls} R_{ijk}^s \\
&= R_{ijkl}.
\end{aligned}$$

Next we observe that the expressions $L_{il} L_{jk} - L_{jl} L_{ik}$ are, like the curvature tensors, skew-symmetric in i, j and k, l. Thus it suffices to check the single equation

$$R_{vuuv} = K \det [\mathrm{I}] = L_{vv} L_{uu} - L_{uv}^2 = \det [\mathrm{II}]$$

which is simply the definition of the Gauss curvature. □

The two entries in the bottom row in the matrices above reduce to the *Codazzi Equations*

$$\partial_i L_{jk} + \begin{bmatrix} L_{iu} & L_{iv} & 0 \end{bmatrix} \begin{bmatrix} \Gamma_{jk}^u \\ \Gamma_{jk}^v \\ L_{jk} \end{bmatrix} = \partial_j L_{ik} + \begin{bmatrix} L_{ju} & L_{jv} & 0 \end{bmatrix} \begin{bmatrix} \Gamma_{ik}^u \\ \Gamma_{ik}^v \\ L_{ik} \end{bmatrix}$$

or rearranged

$$\partial_i L_{jk} - \partial_j L_{ik} = \begin{bmatrix} L_{ju} & L_{jv} \end{bmatrix} \begin{bmatrix} \Gamma_{ik}^u \\ \Gamma_{ik}^v \end{bmatrix} - \begin{bmatrix} L_{iu} & L_{iv} \end{bmatrix} \begin{bmatrix} \Gamma_{jk}^u \\ \Gamma_{jk}^v \end{bmatrix}.$$

Note again that while we only established these for $u = i$ and $j = v$, they also hold when u, v are switched and that both sides vanish when $i = j$. These 8 equations can be reduced to only two Codazzi equations:

$$\begin{aligned}
\partial_u L_{vu} - \partial_v L_{uu} &= L_{vu} \Gamma_{uu}^u + L_{vv} \Gamma_{uu}^v - L_{uu} \Gamma_{vu}^u - L_{uv} \Gamma_{vu}^v, \\
\partial_u L_{vv} - \partial_v L_{uv} &= L_{vu} \Gamma_{uv}^u + L_{vv} \Gamma_{uv}^v - L_{uu} \Gamma_{vv}^u - L_{uv} \Gamma_{vv}^v.
\end{aligned}$$

The last column yields a similar set of equations

$$\partial_i L_j^k - \partial_j L_i^k = \begin{bmatrix} \Gamma_{ju}^k & \Gamma_{jv}^k \end{bmatrix} \begin{bmatrix} L_i^u \\ L_i^v \end{bmatrix} - \begin{bmatrix} \Gamma_{iu}^k & \Gamma_{iv}^k \end{bmatrix} \begin{bmatrix} L_j^u \\ L_j^v \end{bmatrix}$$

These can, however, be derived from the above Codazzi equations using the relationship between the Weingarten map and the second fundamental form.

We are now ready to present the fundamental theorem of surface theory. It is analogous to theorem 2.1.5 for planar curves.

THEOREM 6.3.3. (Fundamental Theorem of Surface Theory, Bonnet, 1848) *A surface is uniquely determined by its first and second fundamental forms if its position and tangent space space are known at just one point. Conversely, any choice of abstract first and second fundamental forms that are related by the Gauss and Codazzi equations are locally the first and second fundamental forms of a surface.*

PROOF. We start by observing that the matrices $[D_w]$ can be defined as long as we are given [I] and [II]. The problem then depends on understanding the solutions to the following big system:

$$\begin{aligned} \frac{\partial \mathrm{q}}{\partial u} &= U, \\ \frac{\partial \mathrm{q}}{\partial v} &= V, \\ \frac{\partial}{\partial u}\begin{bmatrix} U & V & \mathrm{n} \end{bmatrix} &= \begin{bmatrix} U & V & \mathrm{n} \end{bmatrix}[D_u], \\ \frac{\partial}{\partial v}\begin{bmatrix} U & V & \mathrm{n} \end{bmatrix} &= \begin{bmatrix} U & V & \mathrm{n} \end{bmatrix}[D_v], \end{aligned}$$

with initial conditions

$$\begin{aligned} \mathrm{q}(0,0) &= \mathrm{q}_0 \in \mathbb{R}^3, \\ U(0,0) &= U_0 \in \mathbb{R}^3, \\ V(0,0) &= V_0 \in \mathbb{R}^3, \\ \mathrm{n}(0,0) &= \mathrm{n}_0 \in \mathbb{R}^3, \end{aligned}$$

where we additionally require that

$$\begin{aligned} U_0 \cdot U_0 &= g_{uu}(0,0), \\ U_0 \cdot V_0 &= g_{uv}(0,0), \\ V_0 \cdot V_0 &= g_{vv}(0,0), \\ \mathrm{n}_0 &= \frac{U_0 \times V_0}{|U_0 \times V_0|}. \end{aligned}$$

It is clear that this big system has a unique solution given the initial values. To solve it we must check that the necessary integrability conditions are satisfied. We can separate the problem into first solving

$$\begin{aligned} \frac{\partial}{\partial u}\begin{bmatrix} U & V & \mathrm{n} \end{bmatrix} &= \begin{bmatrix} U & V & \mathrm{n} \end{bmatrix}[D_u], \\ \frac{\partial}{\partial v}\begin{bmatrix} U & V & \mathrm{n} \end{bmatrix} &= \begin{bmatrix} U & V & \mathrm{n} \end{bmatrix}[D_v]. \end{aligned}$$

Here the integrability conditions are satisfied as we assumed that

$$[D_u][D_v] + \frac{\partial}{\partial u}[D_v] = [D_v][D_u] + \frac{\partial}{\partial v}[D_u].$$

Having solved this system with the given initial values it remains to find the surface by solving

$$\begin{aligned} \frac{\partial \mathrm{q}}{\partial u} &= U, \\ \frac{\partial \mathrm{q}}{\partial v} &= V. \end{aligned}$$

Here the right hand side does not depend on q so the integrability conditions are simply

$$\frac{\partial U}{\partial v}=\frac{\partial V}{\partial u}.$$

However, we know that

$$\begin{aligned}\frac{\partial U}{\partial v} &= \begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}\begin{bmatrix}\Gamma_{vu}^{u}\\ \Gamma_{vu}^{v}\\ L_{vu}\end{bmatrix},\\ \frac{\partial V}{\partial u} &= \begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}\begin{bmatrix}\Gamma_{uv}^{u}\\ \Gamma_{uv}^{v}\\ L_{uv}\end{bmatrix}.\end{aligned}$$

Here the right-hand sides are equal as $L_{uv}=L_{vu}$ and $\Gamma_{uv}^{w}=\Gamma_{vu}^{w}$.

Having solved the equations it then remains to show that the surface we have constructed has the correct first and second fundamental forms. This will of course depend on the extra conditions that we imposed:

$$\begin{aligned}U_0\cdot U_0 &= g_{uu}(0,0),\\ U_0\cdot V_0 &= g_{uv}(0,0),\\ V_0\cdot V_0 &= g_{vv}(0,0),\\ \mathrm{n}_0 &= \frac{U_0\times V_0}{|U_0\times V_0|}.\end{aligned}$$

In fact they show that at $(0,0)$ the surface has the correct first fundamental form and normal vector. More generally consider the 3×3 matrix of inner products

$$\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}^t\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix},$$

where the block consisting of

$$\begin{bmatrix} U & V\end{bmatrix}^t\begin{bmatrix} U & V\end{bmatrix}$$

corresponds to the first fundamental form of the surface we have constructed. The derivative of this 3×3 matrix satisfies

$$\begin{aligned}&\frac{\partial}{\partial w}\left(\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}^t\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}\right)\\ &= \left(\frac{\partial}{\partial w}\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}\right)^t\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}+\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}^t\frac{\partial}{\partial w}\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}\\ &= \left(\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}[D_w]\right)^t\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}+\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}^t\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}[D_w]\\ &= [D_w]^t\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}^t\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}+\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}^t\begin{bmatrix} U & V & \mathrm{n}\end{bmatrix}[D_w].\end{aligned}$$

This is a differential equation of the type

$$\frac{\partial X}{\partial w}=[D_w]^t X+X[D_w],$$

where X is a 3×3 matrix. Now

$$X=\begin{bmatrix} g_{uu} & g_{uv} & 0\\ g_{vu} & g_{vv} & 0\\ 0 & 0 & 1\end{bmatrix}$$

also satisfies this equation as we constructed $[D_w]$ directly from the given first and second fundamental forms. However, these two solutions have the same initial value at $(0,0)$ so they must be equal. This shows that our surface has the correct first fundamental form and also that n is a unit normal to the surface. This in turn implies that we also obtain the correct second fundamental form since we now also know that

$$\begin{bmatrix} U & V & \mathrm{n} \end{bmatrix}^t \frac{\partial}{\partial w} \begin{bmatrix} U & V & \mathrm{n} \end{bmatrix} = \begin{bmatrix} U & V & \mathrm{n} \end{bmatrix}^t \begin{bmatrix} U & V & \mathrm{n} \end{bmatrix} [D_w] .$$

Here the right hand side is known and the left hand side contains all of the terms we need for calculating the second fundamental form of the constructed surface. □

This theorem allows us to give a complete local characterization of abstract surfaces with constant non-negative Gauss curvature. In fact such surfaces are forced to be locally isometric to the plane or a sphere.

THEOREM 6.3.4. *An abstract surface of constant Gauss curvature $K \geq 0$, can locally be represented as part of a plane when $K = 0$ and part of a sphere of radius $1/\sqrt{K}$ when $K > 0$.*

PROOF. We are given I and have to guess II. The natural choice is $\mathrm{II} = \sqrt{K}\,\mathrm{I}$, i.e., $L^i_j = \sqrt{K}\delta^i_j$ and $L_{ij} = \sqrt{K}g_{ij}$. This allows us to calculate $[D_i]$ in a specific parametrization. We are then left with the goal of checking the integrability conditions, i.e., the Gauss and Codazzi equations. The Codazzi equations are obviously satisfied when $\mathrm{II} = 0$, and follow from the formula for the Christoffel symbols when $K > 0$. More precisely we start with the right hand side of the Codazzi equations and use the intrinsic formulas for the Christoffel symbols to show that they hold:

$$\begin{aligned}
& \begin{bmatrix} L_{ju} & L_{jv} \end{bmatrix} \begin{bmatrix} \Gamma^u_{ik} \\ \Gamma^v_{ik} \end{bmatrix} - \begin{bmatrix} L_{iu} & L_{iv} \end{bmatrix} \begin{bmatrix} \Gamma^u_{jk} \\ \Gamma^v_{jk} \end{bmatrix} \\
= \; & \sqrt{K} \begin{bmatrix} g_{ju} & g_{jv} \end{bmatrix} \begin{bmatrix} \Gamma^u_{ik} \\ \Gamma^v_{ik} \end{bmatrix} - \sqrt{K} \begin{bmatrix} g_{iu} & g_{iv} \end{bmatrix} \begin{bmatrix} \Gamma^u_{jk} \\ \Gamma^v_{jk} \end{bmatrix} \\
= \; & \sqrt{K} \left(\Gamma^u_{ik} g_{uj} + \Gamma^v_{ik} g_{vj} \right) - \sqrt{K} \left(\Gamma^u_{jk} g_{ui} + \Gamma^v_{jk} g_{vi} \right) \\
= \; & \sqrt{K} \left(\Gamma_{ikj} - \Gamma_{jki} \right) \\
= \; & \frac{\sqrt{K}}{2} \left((\partial_i g_{kj} + \partial_k g_{ij} - \partial_j g_{ik}) - (\partial_j g_{ik} + \partial_k g_{ji} - \partial_i g_{jk}) \right) \\
= \; & \frac{\sqrt{K}}{2} \left((\partial_i g_{kj} - \partial_j g_{ik}) - (\partial_j g_{ik} - \partial_i g_{jk}) \right) \\
= \; & \sqrt{K} \left(\partial_i g_{kj} - \partial_j g_{ik} \right) \\
= \; & \partial_i L_{jk} - \partial_j L_{ik}.
\end{aligned}$$

Our assumptions about the the second fundamental form imply

$$L^l_i L_{jk} - L^l_j L_{ik} = K\delta^l_i g_{jk} - K\delta^l_j g_{ik}$$

and proposition 6.2.4 shows that this gives us the Gauss equations:

$$R^l_{ijk} = L^l_i L_{jk} - L^l_j L_{ik}.$$

Now that we have a local representation of the abstract surface as a parametrized surface in $\mathbb{R}^3$ with $\mathrm{II} = \sqrt{K}\,\mathrm{I}$ we can finish the proof in the way we finished the proof of theorem 5.4.6. □

REMARK 6.3.5. It is possible to develop a theory for space-like surfaces in $\mathbb{R}^{2,1}$ that mirrors the theory for surfaces in $\mathbb{R}^3$. This includes new versions of the Gauss and Codazzi equations that also lead to exact analogies of theorems 6.3.3 and 6.3.4. Thus abstract surfaces of constant negative curvature $-R^{-2}$ can locally be represented as part of the surface in $\mathbb{R}^{2,1}$ given by the equation

$$x^2 + y^2 - z^2 = -R^2.$$

We end this long section with a profound theorem that relates to the concepts discussed here. In essence it shows that while it is occasionally possible to choose a second fundamental form locally so that it satisfies the Gauss and Codazzi equations, it might not be possible to extend it to be defined on the entire abstract surface. The result also indicates that in order to characterize hyperbolic space in a way that is similar to theorem 6.3.4 it is more natural to use $\mathbb{R}^{2,1}$ as the ambient space.

THEOREM 6.3.6. *(Hilbert, 1901) It is not possible to select a second fundamental form* II *on all of hyperbolic space H such that* I *and* II *satisfy the Gauss and Codazzi equations.*

PROOF. We argue by contradiction and assume that such a second fundamental form exists. The Gauss equations imply that at each point there is a positive and negative principal direction for II. This in turn implies that there are two linearly independent asymptotic directions at each point, i.e., directions where $\mathrm{II}(X,X) = 0$. Specifically, if $L(E_1) = \kappa E_1$ and $L(E_2) = -\frac{1}{\kappa}E_2$, where $\kappa > 0$, then we can use $X = \frac{1}{\sqrt{\kappa}}E_1 \pm \sqrt{\kappa}E_2$. Fix a parametrization (x,y) of H, e.g., the one that makes hyperbolic space a Monge patch. At $(0,0)$ make a choice of unit asymptotic directions P, Q. Extend this choice to be consistent along the x-axis, and then finally along vertical lines to obtain a consistent choice on all of H. This gives us two unit vector fields P, Q that form an angle $\theta \in (0,\pi)$ with $\mathrm{II}(P,P) = \mathrm{II}(Q,Q) = 0$. We claim that there is a global parametrization where these are the coordinate vector fields. This would follow directly from the global version of theorem A.5.3 if we could check the integrability conditions

$$\frac{\partial P}{\partial x}Q^x + \frac{\partial P}{\partial y}Q^y = \frac{\partial Q}{\partial x}P^x + \frac{\partial Q}{\partial y}P^y$$

and find M, C such that

$$\sqrt{(P^x)^2 + (P^y)^2},\ \sqrt{(Q^x)^2 + (Q^y)^2} \leq M + C\sqrt{x^2 + y^2}.$$

Note that in this case P, Q are independent of (u,v). To prove the bounds for P, Q we note that if $|X|^2 = 1$, then example 6.1.2 implies

$$\begin{aligned} 1 &= X \cdot X \\ &= \frac{1}{z^2}\left((X^x)^2 + (X^y)^2 + (yX^x - xX^y)^2\right). \end{aligned}$$

Since $z^2 = 1 + x^2 + y^2$ this shows that

$$(X^x)^2 + (X^y)^2 \leq 1 + x^2 + y^2$$

which implies the desired bounds on P and Q.

The integrability conditions are bit more tricky. They are in fact a consequence of the Codazzi equations. We do the calculation by an indirect method where we show that there are local parametrizations $\mathrm{q}(u,v)$ of H where $\partial_u \mathrm{q} = P$ and $\partial_v \mathrm{q} = Q$, i.e., it is possible to locally find the desired parametrizations.

By appealing to remark 4.2.11 we can for any $q \in H$ find a local parametrization $\mathrm{q}(u,v)$ with $\mathrm{q}(0,0) = q$, where $\partial_u \mathrm{q} = \lambda P$ and $\partial_v \mathrm{q} = \mu Q$ for some functions λ, μ with $\lambda(u,0) = 1$ and $\mu(0,v) = 1$. The Codazzi equations for a parametrization with $L_{uu} = L_{vv} = 0$ reduce to:

$$\begin{aligned} \frac{\partial L_{vu}}{\partial u} + L_{uv}\left(\Gamma^v_{vu} - \Gamma^u_{uu}\right) &= 0, \\ -\frac{\partial L_{uv}}{\partial v} + L_{uv}\left(\Gamma^v_{vv} - \Gamma^u_{uv}\right) &= 0. \end{aligned}$$

If we combine these equations with the formula from section 5.2 exercise 9:

$$\frac{\partial \sqrt{\det[\mathrm{I}]}}{\partial w} = \sqrt{\det[\mathrm{I}]}\left(\Gamma^u_{uw} + \Gamma^v_{vw}\right)$$

and the curvature assumption:

$$K = -\frac{L^2_{uv}}{\det[\mathrm{I}]} = -1,$$

then it follows that $L_{uv} = \pm\sqrt{\det[\mathrm{I}]}$ and

$$\begin{aligned} \Gamma^u_{uu} - \Gamma^v_{vu} &= \Gamma^u_{uu} + \Gamma^v_{vu}, \\ \Gamma^v_{vv} - \Gamma^u_{uv} &= \Gamma^u_{uv} + \Gamma^v_{vv}. \end{aligned}$$

This in turn implies

$$\Gamma^u_{uv} = 0 = \Gamma^v_{uv}$$

which is equivalent to

$$\frac{1}{2}\partial_v g_{uu} = \Gamma_{uvu} = 0 \text{ and } \frac{1}{2}\partial_u g_{vv} = \Gamma_{uvv} = 0.$$

In our case $g_{uu} = \lambda^2$ and $g_{vv} = \mu^2$ so that the conditions $\lambda(u,0) = 1$ and $\mu(0,v) = 1$ now imply that $g_{uu} = g_{vv} = 1$.

All in all this gives us the desired global parametrization $\mathrm{q}(u,v)$ where

$$[\mathrm{I}] = \begin{bmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{bmatrix}, \ \theta \in (0,\pi).$$

The formula for the Gauss curvature in such coordinates reduces to (see section 6.1 exercise 5)

$$\partial^2_{st}\theta = -K\sin\theta = \sin\theta.$$

Here $\sin\theta = \sqrt{g_{uu}g_{vv} - g^2_{uv}}$ is also the area element so it follows that

$$\begin{aligned} \left(\theta(b,t) - \theta(a,t)\right)\Big|_{t=c}^{t=d} &= \int_c^d \int_a^b \partial_{st}\theta \, ds\, dt \\ &= \int_c^d \int_a^b \sin\theta \, ds\, dt \\ &= \operatorname{area} \mathrm{q}\left([a,b]\times[c,d]\right). \end{aligned}$$

In particular $\operatorname{area} \mathrm{q} \leq 2\pi$. On the other hand the area of H can be calculated in the Monge patch representation to be

$$\int_{\mathbb{R}^2} \frac{dxdy}{\sqrt{1+x^2+y^2}} = \infty.$$

Thus the parametrization $\mathrm{q}(s,t)$ cannot cover all of H. $\square$

COROLLARY 6.3.7. *There is no Riemannian immersion from hyperbolic space H into $\mathbb{R}^3$.*

REMARK 6.3.8. Elie Cartan developed an approach to the Gauss and Codazzi equations that uses orthonormal bases. Thus he chose an orthonormal frame E_1, E_2, E_3 along part of a surface with the property that $E_3 = \mathrm{n}$ is normal to the surface. Consequently, E_1, E_2 form an orthonormal basis for the tangent space. The goal is again to take derivatives. For that purpose we can still use parameters

$$\frac{\partial}{\partial w}\begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix} = \begin{bmatrix} \frac{\partial E_1}{\partial w} & \frac{\partial E_2}{\partial w} & \frac{\partial E_3}{\partial w} \end{bmatrix} = \begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix}[D_w].$$

The first observation is that $[D_w]$ is skew-symmetric since we used an orthonormal basis:

$$\begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix}^t \begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

so

$$\begin{aligned}
0 &= \frac{\partial}{\partial w}\left(\begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix}^t \begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix}\right) \\
&= \left(\frac{\partial}{\partial w}\begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix}\right)^t \begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix} \\
&\quad + \begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix}^t \frac{\partial}{\partial w}\begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix} \\
&= \left(\begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix}[D_w]\right)^t \begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix} \\
&\quad + \begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix}^t \begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix}[D_w] \\
&= [D_w]^t + [D_w].
\end{aligned}$$

In particular, there will only be 3 entries to sort out. This is a significant reduction from what we had to deal with above. What is more, the entries can easily be found by computing the dot products

$$E_i \cdot \frac{\partial E_j}{\partial w}.$$

This is also in sharp contrast to what happens in the above situation as we shall see. Taking one more derivative will again yield a formula

$$\left[\frac{\partial D_{w_2}}{\partial w_1}\right] - \left[\frac{\partial D_{w_1}}{\partial w_2}\right] = [D_{w_2}][D_{w_1}] - [D_{w_1}][D_{w_2}],$$

where both sides are skew symmetric.

Given the simplicity of using orthonormal frames it is perhaps puzzling why one would bother developing the more cumbersome approach that uses coordinate fields. The answer lies, as with curves, in the unfortunate fact that it is often easier to find coordinate fields than orthonormal bases that are easy to work with.

Monge patches are prime examples. For specific examples and many theoretical developments, however, Cartan's approach has many advantages.

Exercises

(1) Is it possible for a parametrized surface $\mathrm{q}(u,v) : U \to \mathbb{R}^3$ to have:

(a) $\mathrm{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathrm{II} = \begin{bmatrix} 0 & 0 \\ 0 & f(u) \end{bmatrix}$?

(b) $\mathrm{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathrm{II} = \begin{bmatrix} 0 & 0 \\ 0 & f(v) \end{bmatrix}$?

(c) $\mathrm{I} = \begin{bmatrix} 1 & 0 \\ 0 & \cos^2 u \end{bmatrix}$ and $\mathrm{II} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2 u \end{bmatrix}$?

(d) $\mathrm{I} = \begin{bmatrix} 1 & 0 \\ 0 & f(u) \end{bmatrix}$ and $\mathrm{II} = \begin{bmatrix} 1 & 0 \\ 0 & f(u) \end{bmatrix}$?

(e) $\mathrm{I} = \begin{bmatrix} 1 & 0 \\ 0 & f(u) \end{bmatrix}$ and $\mathrm{II} = \begin{bmatrix} f(u) & 0 \\ 0 & 1 \end{bmatrix}$?

(2) If the principal curvatures are not equal on some part of the surface, then we can use corollary 5.4.4 to find an orthogonal parametrization $\mathrm{q}(u,v)$ where the tangent fields are principal directions, i.e., $L_{uv} = g_{uv} = 0$. Show that in this case the Codazzi equations can be written as

$$\begin{aligned} \frac{\partial L_{uu}}{\partial v} &= H\frac{\partial g_{uu}}{\partial v}, \\ \frac{\partial L_{vv}}{\partial u} &= H\frac{\partial g_{vv}}{\partial u}. \end{aligned}$$

(3) Show that a minimal surface admits an orthogonal parametrization where $L_{uu} = -L_{vv} = 1$ and $L_{uv} = 0$. Show further that such a parametrization is conformal (isothermal).

(4) Use the Codazzi equations to show that if the principal curvatures $\kappa_1 = \kappa_2$ are equal on a surface, then they are constant. Hint: In this case $L_{ij} = Hg_{ij}$, where H is the mean curvature.

(5) Show that the equations

$$\partial_i L_j^k - \partial_j L_i^k = \begin{bmatrix} \Gamma_{ju}^k & \Gamma_{jv}^k \end{bmatrix} \begin{bmatrix} L_i^u \\ L_i^v \end{bmatrix} - \begin{bmatrix} \Gamma_{iu}^k & \Gamma_{iv}^k \end{bmatrix} \begin{bmatrix} L_j^u \\ L_j^v \end{bmatrix}$$

follow from the Codazzi equations.

(6) If the principal curvatures κ_1 and κ_2 are not equal on some part of the surface, then we can use corollary 5.4.4 to find an orthogonal parametrization where the tangent fields are principal directions. Show that in this case the Codazzi equations can be written as

$$\begin{aligned} \frac{\partial \kappa_1}{\partial v} &= \frac{1}{2}(\kappa_2 - \kappa_1)\frac{\partial \ln g_{uu}}{\partial v}, \\ \frac{\partial \kappa_2}{\partial u} &= \frac{1}{2}(\kappa_1 - \kappa_2)\frac{\partial \ln g_{vv}}{\partial u}. \end{aligned}$$

(7) (Hilbert, 1901) The goal is to show: If there is a point p on a surface, where K is positive, κ_1 has a maximum, and κ_2 a minimum, then the principal curvatures are equal and constant. More specifically, we show that if

$$\sup \kappa_1 = \kappa_1(p) > \kappa_2(p) = \inf \kappa_2,$$

then $K \leq 0$. Select a parametrization around p, as in exercise 6, where the coordinate curves are lines of curvature.

(a) Show that at p

$$\begin{aligned}
\frac{\partial \kappa_1}{\partial u} &= \frac{\partial \kappa_1}{\partial v} = 0,\ \frac{\partial^2 \kappa_1}{\partial v^2} \leq 0, \\
\frac{\partial \kappa_2}{\partial u} &= \frac{\partial \kappa_2}{\partial v} = 0,\ \frac{\partial^2 \kappa_2}{\partial u^2} \geq 0.
\end{aligned}$$

(b) Using the Codazzi equations from the previous exercise to show that at p

$$\frac{\partial \ln g_{uu}}{\partial v} = 0 = \frac{\partial \ln g_{vv}}{\partial u}$$

and after differentiation also at p that

$$\frac{\partial^2 \ln g_{uu}}{\partial v^2} \geq 0,\ \frac{\partial^2 \ln g_{vv}}{\partial u^2} \geq 0.$$

(c) Next show that at p

$$K = -\frac{1}{2}\left(\frac{1}{g_{vv}}\frac{\partial^2 \ln g_{uu}}{\partial v^2} + \frac{1}{g_{uu}}\frac{\partial^2 \ln g_{vv}}{\partial u^2}\right) \leq 0.$$

(d) Finally establish the first statement.

(8) Show that a surface with constant principal curvatures must be part of a plane, sphere, or right circular cylinder. Note that the two former cases happen when the principal curvatures are equal.

(9) Show that having zero Gauss curvature is the integrability condition for admitting Cartesian coordinates on an abstract surface. Hint: Think of U, V as 2-dimensional vectors and consider the system

$$\begin{aligned}
\partial_w \begin{bmatrix} U & V \end{bmatrix} &= \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} \Gamma^u_{wu} & \Gamma^u_{wv} \\ \Gamma^v_{wu} & \Gamma^v_{wv} \end{bmatrix}, \\
\partial_u \mathrm{q} &= U, \\
\partial_v \mathrm{q} &= V.
\end{aligned}$$

These are simply the Gauss formulas with the last columns and rows erased. Show that the integrability equations for U, V are $K = 0$. Then use the last equations to find $\mathrm{q} : U \to \mathbb{R}^2$ after having checked the integrability conditions are satisfied. Finally, show that $(x, y) = \mathrm{q}(u, v)$ is a Cartesian parametrization provided the correct initial conditions for U, V have be specified.

(10) Consider potential surfaces $\mathrm{q}(u, v)$ where

$$[\mathrm{I}] = \begin{bmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{bmatrix},\ [\mathrm{II}] = \begin{bmatrix} \lambda^2 \kappa & 0 \\ 0 & -\frac{\lambda^2}{\kappa} \end{bmatrix}.$$

See also section 6.1 exercise 8.

(a) Show that

$$K = -1$$

and

$$\Delta \ln \lambda = \frac{\partial^2 \ln \lambda}{\partial u^2} + \frac{\partial^2 \ln \lambda}{\partial v^2} = \lambda^2.$$

Hint: The first is easy and to do the second use section 5.3 exercise 32.

(b) Show that if we choose

$$\lambda = \frac{1}{a\left(u^2+v^2\right)+b_u u+b_v v+c},$$

where a, b_u, b_v, c are constants such that

$$4ac - b_u^2 - b_v^2 = -1,$$

then the first fundamental form

$$[\mathrm{I}] = \begin{bmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{bmatrix}$$

has $K = -1$.

(c) Show that when $a = 0$ and either $b_u = 0$ or $b_v = 0$, i.e., λ must be a function of either u or v, then it is possible to find κ so that the Codazzi equations are satisfied. Thus we obtain a surface in space.

(d) Show that the pseudo-sphere (see section 5.4 exercise 9) is an example of such a surface with $\lambda = \frac{1}{v}$, $v > 0$.

(e) When $\lambda = \frac{1}{v}$ show that $\kappa^2 + 1 = ev^2$ for some constant $e > 0$ and conclude that κ is not defined for all v.

(f) Show that if $a \neq 0$ or $b_u b_v \neq 0$, then it is not possible to find κ so that the Codazzi equations are satisfied. Hint: The Codazzi equations yield formulas

$$\frac{\partial \kappa}{\partial u} = P\left(\kappa, \frac{\partial \ln \lambda}{\partial u}\right) \text{ and } \frac{\partial \kappa}{\partial v} = Q\left(\kappa, \frac{\partial \ln \lambda}{\partial v}\right).$$

So it comes down to checking the integrability conditions when both $\frac{\partial \ln \lambda}{\partial u} \neq 0$ and $\frac{\partial \ln \lambda}{\partial v} \neq 0$.

(11) Consider a space-like parametrized surface $\mathrm{q}(u,v) : U \to \mathbb{R}^{2,1}$ and define the Weingarten map

$$L\left(\frac{d\,\mathrm{q}}{dt}\right) = -\frac{d\,\mathrm{n}\circ\mathrm{q}}{dt}$$

and second fundamental form $\mathrm{II}(X,Y) = \mathrm{I}(L(X),Y)$ as for surfaces in $\mathbb{R}^3$.

(a) Show that the Gauss-Weingarten equations are given by

$$\begin{aligned} \frac{\partial}{\partial w}\begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} & \mathrm{n} \end{bmatrix} &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} & \mathrm{n} \end{bmatrix}[D_w] \\ &= \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} & \mathrm{n} \end{bmatrix}\begin{bmatrix} \Gamma^u_{wu} & \Gamma^u_{wv} & -L^u_w \\ \Gamma^v_{wu} & \Gamma^v_{wv} & -L^v_w \\ -L_{wu} & -L_{wv} & 0 \end{bmatrix}. \end{aligned}$$

Note the change in signs in the last column when compared to surfaces in $\mathbb{R}^3$.

(b) Show that the Codazzi equations are the same as above.

(c) Show that the Gauss equations are now given by

$$R^l_{ijk} = -\left(L^l_i L_{jk} - L^l_j L_{ik}\right).$$

(d) Show that these Codazzi and Gauss equations are the integrability equations for a space-like surface in $\mathbb{R}^{2,1}$.

(12) Let $\mathrm{q}(u,v)$ be a parametrized surface in $\mathbb{R}^3$. Assume E_1 and E_2 are tangent vector fields forming an orthonormal basis for the tangent space everywhere and

$$E_1 \times E_2 = \mathrm{n} = \frac{\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}}{\left|\frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v}\right|}.$$

(a) Show that

$$\begin{aligned}
\frac{\partial}{\partial w}\left[\begin{array}{ccc} E_1 & E_2 & \mathrm{n} \end{array}\right] &= \left[\begin{array}{ccc} E_1 & E_2 & \mathrm{n} \end{array}\right][D_w],\\
[D_w] &= \left[\begin{array}{ccc} 0 & -\phi_w & -\phi_{w1} \\ \phi_w & 0 & -\phi_{w2} \\ \phi_{w1} & \phi_{w2} & 0 \end{array}\right],
\end{aligned}$$

where

$$\begin{aligned}
\phi_w &= \frac{\partial E_1}{\partial w} \cdot E_2 = -\frac{\partial E_2}{\partial w} \cdot E_1,\\
\phi_{w1} &= \frac{\partial E_1}{\partial w} \cdot \mathrm{n} = -\mathrm{I}\left(\frac{\partial \mathrm{n}}{\partial w}, E_1\right),\\
\phi_{w2} &= \frac{\partial E_2}{\partial w} \cdot \mathrm{n} = -\mathrm{I}\left(\frac{\partial \mathrm{n}}{\partial w}, E_2\right).
\end{aligned}$$

(b) Show that

$$\begin{aligned}
\phi_{w1} &= \mathrm{II}\left(\partial_w \mathrm{q}, E_1\right) = \mathrm{I}\left(\partial_w \mathrm{q}, L\left(E_1\right)\right),\\
\phi_{w2} &= \mathrm{II}\left(\partial_w \mathrm{q}, E_2\right) = \mathrm{I}\left(\partial_w \mathrm{q}, L\left(E_2\right)\right).
\end{aligned}$$

(c) Use the Weingarten equations and $[L]$ as the matrix of the Weingarten map with respect to E_1, E_2 to show that

$$[L]\left[\begin{array}{cc} E_1 & E_2 \end{array}\right]^t \left[\begin{array}{cc} \frac{\partial \mathrm{q}}{\partial u} & \frac{\partial \mathrm{q}}{\partial v} \end{array}\right] = \left[\begin{array}{cc} \phi_{u1} & \phi_{v1} \\ \phi_{u2} & \phi_{v2} \end{array}\right]$$

and

$$K\sqrt{\det[\mathrm{I}]} = \phi_{u1}\phi_{v2} - \phi_{u2}\phi_{v1}.$$

(d) Show that the integrability conditions

$$\frac{\partial}{\partial u}[D_v] - \frac{\partial}{\partial v}[D_u] + [D_u][D_v] - [D_v][D_u] = 0$$

can be reduced to the three equations:

$$\begin{aligned}
\frac{\partial \phi_v}{\partial u} - \frac{\partial \phi_u}{\partial v} &= \phi_{u2}\phi_{v1} - \phi_{v2}\phi_{u1},\\
\frac{\partial \phi_{v1}}{\partial u} - \frac{\partial \phi_{u1}}{\partial v} &= \phi_{v2}\phi_u - \phi_{u2}\phi_v,\\
\frac{\partial \phi_{v2}}{\partial u} - \frac{\partial \phi_{u2}}{\partial v} &= -\phi_{v1}\phi_u + \phi_{u1}\phi_v.
\end{aligned}$$

(e) Show that

$$\frac{\partial \phi_v}{\partial u} - \frac{\partial \phi_u}{\partial v} = \phi_{u2}\phi_{v1} - \phi_{v2}\phi_{u1}$$

corresponds to the Gauss equation.

(f) Show that

$$\frac{\partial \phi_{v1}}{\partial u} - \frac{\partial \phi_{u1}}{\partial v} = \phi_{v2}\phi_u - \phi_{u2}\phi_v,$$
$$\frac{\partial \phi_{v2}}{\partial u} - \frac{\partial \phi_{u2}}{\partial v} = -\phi_{v1}\phi_u + \phi_{u1}\phi_v$$

correspond to the Codazzi equations.

6.4. The Gauss-Bonnet Theorem

Recall from section 2.2 that the integral of the curvature of a planar curve is related to how the tangent moves. In this section we shall prove a similar result for curves on abstract surfaces. To check that we cannot expect the same statement to hold, consider the equator on a sphere. This curve has acceleration normal to itself and lies in the (x, y)-plane, in particular, the acceleration is also normal to the sphere and so has no geodesic curvature. On the other hand the tangent field clearly turns around 360 degrees.

Throughout this section we assume that a parametrized abstract surface is given with a rectangular parameter domain $U = (a_u, b_u) \times (a_v, b_v)$. The key is that the domain should not have any holes in it. We further assume that we have a smaller domain $R \subset U$ that is bounded by a piecewise smooth curve

$$(u(s), v(s)) : [0, L] \to (a_u, b_u) \times (a_v, b_v)$$

running counter clockwise in the plane and such that $\mathrm{q}(s) = \mathrm{q}(u(s), v(s))$ is unit speed with respect to the given first fundamental form $[\mathrm{I}]$ on the abstract surface.

Integration of functions on the surface is done by defining a suitable integral using the parametrization. To make this invariant under parametrizations we define

$$\int_{\mathrm{q}(R)} f dA = \int_R f(u, v) \sqrt{\det [\mathrm{I}]} du dv = \int_R f(u, v) \left| \frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v} \right| du dv.$$

This ensures that if we use a different parametrization (s, t) where $\mathrm{q}(Q) = \mathrm{q}(R)$, then

$$\int_R f(u, v) \sqrt{\det [\mathrm{I}]} du dv = \int_Q f(s, t) \sqrt{\det [\mathrm{I}]} ds dt.$$

We start by calculating the geodesic curvature of q assuming further that

$$[\mathrm{I}] = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix}.$$

The existence of such coordinate systems will be established in proposition 7.4.1. The formulas for the Christoffel symbols and Gauss curvature in such coordinates are given in section B.5.

Lemma 6.4.1. *Let θ be the angle between* q *and the u-curves, then*

$$\kappa_g = \frac{d\theta}{ds} + \frac{\partial r}{\partial u} \frac{1}{r} \sin \theta.$$

Proof. Given the specific form of $[\mathrm{I}]$ and that the velocity has unit length we have

$$\begin{aligned} \frac{d\,\mathrm{q}}{ds} &= \frac{du}{ds}\frac{\partial \mathrm{q}}{\partial u} + \frac{dv}{ds}\frac{\partial \mathrm{q}}{\partial v} \\ &= \cos\theta \frac{\partial \mathrm{q}}{\partial u} + \frac{1}{r}\sin\theta \frac{\partial \mathrm{q}}{\partial v}. \end{aligned}$$

The natural unit normal field to q in the surface is

$$\mathrm{S} = -\sin\theta \frac{\partial\,\mathrm{q}}{\partial u} + \frac{1}{r}\cos\theta \frac{\partial\,\mathrm{q}}{\partial v}.$$

The geodesic curvature is then given by

$$\begin{aligned}
\kappa_g &= \mathrm{I}\left(\mathrm{S}, \ddot{\mathrm{q}}^{\mathrm{I}}\right) \\
&= \mathrm{S}\cdot\left(\left(\frac{d^2u}{ds^2} + \Gamma^u\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right)\right)\frac{\partial\,\mathrm{q}}{\partial u} + \left(\frac{d^2v}{ds^2} + \Gamma^v\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right)\right)\frac{\partial\,\mathrm{q}}{\partial v}\right) \\
&= -\sin\theta\left(\frac{d^2u}{ds^2} + \Gamma^u\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right)\right) + r^2\frac{1}{r}\cos\theta\left(\frac{d^2v}{ds^2} + \Gamma^v\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right)\right) \\
&= -\sin\theta\left(\frac{d^2u}{ds^2} + \Gamma^u\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right)\right) + r\cos\theta\left(\frac{d^2v}{ds^2} + \Gamma^v\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right)\right).
\end{aligned}$$

We further have

$$\begin{aligned}
\frac{d^2u}{ds^2} &= \frac{d\cos\theta}{ds} = -\sin\theta\frac{d\theta}{ds}, \\
\frac{d^2v}{ds^2} &= \frac{d\frac{1}{r}\sin\theta}{ds} \\
&= \frac{-1}{r^2}\frac{dr}{ds}\sin\theta + \frac{1}{r}\cos\theta\frac{d\theta}{ds} \\
&= \frac{-1}{r^2}\left(\frac{\partial r}{\partial u}\frac{du}{ds} + \frac{\partial r}{\partial v}\frac{dv}{ds}\right)\sin\theta + \frac{1}{r}\cos\theta\frac{d\theta}{ds} \\
&= \frac{-1}{r^2}\frac{\partial r}{\partial u}\cos\theta\sin\theta + \frac{-1}{r^3}\frac{\partial r}{\partial v}\sin^2\theta + \frac{1}{r}\cos\theta\frac{d\theta}{ds}.
\end{aligned}$$

And using example 5.2.12 the Christoffel symbols are

$$\begin{aligned}
\Gamma^u\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right) &= \Gamma^u_{vv}\left(\frac{dv}{ds}\right)^2 \\
&= -r\frac{\partial r}{\partial u}\frac{1}{r^2}\sin^2\theta \\
&= \frac{-1}{r}\frac{\partial r}{\partial u}\sin^2\theta,
\end{aligned}$$

$$\begin{aligned}
\Gamma^v\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right) &= 2\Gamma^v_{uv}\frac{du}{ds}\frac{dv}{ds} + \Gamma^v_{vv}\left(\frac{dv}{ds}\right)^2 \\
&= \frac{2}{r}\frac{\partial r}{\partial u}\frac{du}{ds}\frac{dv}{ds} + \frac{1}{r}\frac{\partial r}{\partial v}\left(\frac{dv}{ds}\right)^2 \\
&= \frac{2}{r^2}\frac{\partial r}{\partial u}\sin\theta\cos\theta + \frac{1}{r^3}\frac{\partial r}{\partial v}\sin^2\theta.
\end{aligned}$$

Thus

$$\begin{aligned}
\kappa_g &= -\sin\theta\left(-\sin\theta\frac{d\theta}{ds} - \frac{1}{r}\frac{\partial r}{\partial u}\sin^2\theta\right) + r\cos\theta\left(\frac{1}{r}\cos\theta\frac{d\theta}{ds} + \frac{1}{r^2}\frac{\partial r}{\partial u}\sin\theta\cos\theta\right) \\
&= \frac{d\theta}{ds} + \frac{1}{r}\frac{\partial r}{\partial u}\sin^3\theta + \frac{1}{r}\frac{\partial r}{\partial u}\sin\theta\cos^2\theta \\
&= \frac{d\theta}{ds} + \frac{\partial r}{\partial u}\frac{1}{r}\sin\theta.
\end{aligned}$$

□

We first prove the local Gauss-Bonnet theorem. It is stated in the way that Gauss and Bonnet proved it. Gauss considered regions bounded by geodesics thus eliminating the geodesic curvature, while Bonnet presented the version given below.

THEOREM 6.4.2 (*Gauss, 1825 and Bonnet, 1848*). *The surface and curve are as above. Let θ_i be the exterior angles at the points where* $\mathrm{q}(s)$ *has vertices. Then*

$$\int_{\mathrm{q}(R)} KdA + \int_0^L \kappa_g ds = 2\pi - \sum \theta_i.$$

PROOF. It follows from example 5.3.9 that

$$\begin{aligned}\int_{\mathrm{q}(R)} KdA &= \int_R K\sqrt{\det[\mathrm{I}]}dudv \\ &= -\int_R \frac{\frac{\partial^2 r}{\partial u^2}}{r} rdudv \\ &= -\int_R \frac{\partial^2 r}{\partial u^2}dudv.\end{aligned}$$

The last integral can be turned into a line integral if we use Green's theorem

$$\int_R \frac{\partial^2 r}{\partial u^2}dudv = \int_{\partial R}\frac{\partial r}{\partial u}dv.$$

This line integral can now be recognized as one of the terms in the formula for the geodesic curvature

$$\begin{aligned}\int_{\partial R}\frac{\partial r}{\partial u}dv &= \int_0^L \frac{\partial r}{\partial u}\frac{dv}{ds}ds \\ &= \int_0^L \frac{\partial r}{\partial u}\frac{1}{r}\sin\theta ds \\ &= \int_0^L \left(\kappa_g - \frac{d\theta}{ds}\right)ds \\ &= \int_0^L \kappa_g ds - \int_0^L \frac{d\theta}{ds}ds.\end{aligned}$$

Thus we obtain

$$\begin{aligned}\int_{\mathrm{q}(R)} KdA + \int_0^L \kappa_g ds &= -\int_R \frac{\partial^2 r}{\partial u^2}dudv + \int_{\partial R}\frac{\partial r}{\partial u}dv + \int_0^L \frac{d\theta}{ds}ds \\ &= \int_0^L \frac{d\theta}{ds}ds.\end{aligned}$$

Finally we must show that

$$\int_0^L \frac{d\theta}{ds}ds + \sum\theta_i = 2\pi.$$

For a planar simple closed curve this is a consequence of knowing that the rotation index is 1 when parametrized to run counterclockwise (see theorem 2.2.5 and section

2.2 exercise 4). In this case we know that $\int_0^L \frac{d\theta}{ds} ds + \sum \theta_i$ must be a multiple of 2π. Consider the abstract metrics

$$[\mathrm{I}_\epsilon] = \begin{bmatrix} 1 & 0 \\ 0 & 1 - \epsilon + \epsilon r^2 \end{bmatrix}.$$

For each $\epsilon \in [0,1]$ this defines a metric on U and the rotation index for our curve has to be a multiple 2π. Moreover, when $\epsilon = 0$ the rotation index is 1 as the metric is the standard Euclidean metric. It is easy to see that the angles θ_ϵ between the curve and the u-curves is continuous in ϵ. Thus $\int_0^L \frac{d\theta}{ds} ds + \sum \theta_i$ also varies continuously. However, as it is always a multiple of 2π and is 2π in case $\epsilon = 0$ it follows that it is always 2π. $\square$

Clearly there are subtle things about the regions R we are allowed to use. Aside from the topological restriction on R there is also an orientation choice (counter clockwise) for ∂R in Green's theorem. If we reverse that orientation there is a sign change, and the geodesic curvature also changes sign when we run backwards.

We used rather special coordinates as well, but it is possible to extend the proof to work for all coordinate systems. The same strategy even works, but is complicated by the nasty formula we have for the Gauss curvature in general coordinates. If we use a conformal or isothermal parametrization, then the argument about the winding number is much simpler as angles would be the same in the plane and on the surface. Thus the winding number is clearly 1.

Cartan's approach using orthonormal frames rather than special coordinates makes for a fairly simple proof that works within all coordinate systems. This is exploited in an exercise below. To keep things in line with what we have already covered we still restrict attention to how this works in relation to a parametrization.

Let us now return to our examples from above. Without geodesic curvature and exterior angles we expect to end up with the formula

$$\int_{\mathrm{q}(R)} K dA = 2\pi.$$

But there has to be a region R bounding the closed curve. On the sphere we can clearly use the upper hemisphere. As $K = 1$ we end up with the well known fact that the upper hemisphere has area 2π. If we consider a cylinder, then there are lots of closed curves without geodesic curvature. However, there is no reasonable region bounding these curves despite the fact that we have a valid geodesic coordinate system. The issue is that the bounding curve cannot be set up to be a closed curve in a parametrization where there is a rectangle containing the curve.

It is possible to modify the Gauss-Bonnet formula so that more general regions can be used in the statement, but it requires topological information about the region R. This will be studied in detail later and also in some interesting cases in the exercises below.

In case the surface lies in $\mathbb{R}^3$ it is possible to reinterpret the integral of the Gauss curvature. Recall that

$$K dA = K \left| \frac{\partial \mathrm{q}}{\partial u} \times \frac{\partial \mathrm{q}}{\partial v} \right| = \left(\frac{\partial \mathrm{n}}{\partial u} \times \frac{\partial \mathrm{n}}{\partial v} \right) \cdot \mathrm{n} = \pm \left| \frac{\partial \mathrm{n}}{\partial u} \times \frac{\partial \mathrm{n}}{\partial v} \right|.$$

Thus $\int_R K dA$ measures the signed area of the spherical image traced by the normal vector, or the image of the Gauss map.

Exercises

(1) Consider a surface of revolution and two latitudes q_1 and q_2 on it. These curves bound a band or annular region $\mathrm{q}(R)$. By subdividing the region and using proper orientations and parametrizations on the curves show that

$$\int_{\mathrm{q}(R)} KdA = \int_{\mathrm{q}_1} \kappa_g ds_1 - \int_{\mathrm{q}_2} \kappa_g ds_2.$$

(2) Generalize the previous exercise to suitable regions on general surfaces that are bounded both on the inside and outside by smooth (or even piecewise smooth) closed curves.

(3) Consider a geodesic triangle (curve with three vertices and vanishing geodesic curvature) with interior angles α, β, γ inside a parametrization. Show that

$$\int_{\mathrm{q}(R)} KdA = \alpha + \beta + \gamma - \pi.$$

(4) For a surface with $K \le 0$ and a geodesic polygon (sides have vanishing geodesic curvature) in a parametrization as above show that the number of vertices must be ≥ 3.

(5) Let $\mathrm{q}(u,v)$ be a parametrized surface without special assumptions about the parametrization. Create tangent vector fields E_1 and E_2 forming an orthonormal basis for the tangent space everywhere with the further property that E_1 is proportional to the first tangent field $\frac{\partial \mathrm{q}}{\partial u}$.

(a) Use section 6.3 exercise 12 to conclude that

$$\begin{aligned}\int_{\mathrm{q}(R)} KdA &= -\int_R \left(\frac{\partial \phi_v}{\partial u} - \frac{\partial \phi_u}{\partial v}\right) dudv \\ &= -\int_{\mathrm{q}} \phi_u du + \phi_v dv.\end{aligned}$$

(b) Finally prove the Gauss-Bonnet theorem by establishing

$$\int_{\mathrm{q}} \phi_u du + \phi_v dv = \int \left(k_g - \frac{d\theta}{ds}\right) ds,$$

where θ is the angle with E_1 or $\frac{\partial \mathrm{q}}{\partial u}$. To aid the last calculation show that

$$\frac{dE_1}{ds} = \frac{\partial E_1}{\partial u}\frac{du}{ds} + \frac{\partial E_1}{\partial v}\frac{dv}{ds} = (\cos\theta\phi_u + \sin\theta\phi_v) E_2,$$

$$\frac{dE_2}{ds} = \frac{\partial E_2}{\partial u}\frac{du}{ds} + \frac{\partial E_2}{\partial v}\frac{dv}{ds} = -(\cos\theta\phi_u + \sin\theta\phi_v) E_1,$$

$$\begin{aligned}\dot{\mathrm{q}} &= \cos\theta E_1 + \sin\theta E_2, \\ \mathrm{S} &= -\sin\theta E_1 + \cos\theta E_2, \\ \ddot{\mathrm{q}}^{\mathrm{I}} &= \mathrm{S}\frac{d\theta}{ds} - \sin\theta(\cos\theta\phi_u + \sin\theta\phi_v) E_1 \\ &\quad + \cos\theta(\cos\theta\phi_u + \sin\theta\phi_v) E_2.\end{aligned}$$

6.5. Topology of Surfaces

So far we have only worked with the geometry of surfaces. When studying the global behavior of closed surfaces there are also some interesting topological concepts that are important in our geometric understanding of these surfaces.

Definition 6.5.1. A *polygon*, or *n-gon* is a piecewise smooth simple closed curve inside a rectangular parameterization as in the previous section. The number $n = 0, 1, 2, ...$ refers to the number of points where the curve is not differentiable. We call these points *vertices* of the closed curve and the connecting arcs the *edges*.

Remark 6.5.2. The edges will not include the two boundary points, thus they are simple smooth curves defined on an open interval. A 0-gon is a smooth simple closed curve, it has 0 vertices and 0 edges. A 1-gon is a loop with one vertex and one edge etc. The inside is well defined by the Jordan curve theorem (see theorem 2.3.1) and is called the *face*. Thus the face of an n-gon, $n > 0$, has a boundary that consists of n vertices and n edges, each of the edges in turn has vertices as boundary points. In the special case of a 0-gon the boundary is a smooth circle. Note that the inside of a circle or regular n-gon in the plane is homeomorphic to an open disc. Thus any face is homeomorphic to an open disc.

Definition 6.5.3. A *polygonal subdivision* of an abstract surface is a disjoint decomposition of the surface into faces and their boundaries. More specifically, if a point lies inside a face, then it cannot be in any other face or on the boundary of any face. If a point is a vertex for one face, then it cannot lie on an edge of any other face, but it can be a vertex for several other faces.

Remark 6.5.4. If a point lies on an edge of one face, then it can only lie on edges of other faces. In fact it can only lie on the edge for one other face since such a point has a neighborhood that is homeomorphic to a disc. More precisely, if 3 or more faces meet in a common edge, then no point on that edge has a neighborhood that is homeomorphic to a disc.

Definition 6.5.5. A *triangulation* is a polygonal subdivision into triangles (3-gons) with the added condition that two faces can have at most one edge in common. Note that one can subdivide the sphere into two triangles as in a triangular pillow, but this is not a triangulation. The tetrahedron is a triangulation of the sphere and in fact the triangulation with the smallest number of vertices, edges and faces.

In any given concrete situation it is not hard to find a triangulation, but for an abstract surface this is much less easy to see. We will take it for granted that our surfaces have polygonal subdivisions and triangulations. A polygonal subdivision in fact creates a triangulation if each n-gon is broken up into $2n$ triangles that all have a common vertex in the face and where the edges connect this common vertex to to the original vertices and midpoints of the original edges. A polygonal subdivision can be created using some of the geometric developments in the next chapter. One has to find a finite covering of small sets B_j that have boundaries with positive geodesic curvature. These form a big Venn-type subdivision of the surface. If the sets are chosen appropriately this will also be a polygonal subdivision.

Definition 6.5.6. The *Euler characteristic* of a polygonal subdivision is defined as the alternating sum: $\chi = V - E + F$ where F is the number of faces, E the number of edges, and V the number of vertices. Here E and V are not counted with multiplicity.

Example 6.5.7. If we take a smooth simple closed curve on a sphere, then we obtain a polygonal subdivision where $F = 2$, $E = 0$, and $V = 0$. If we triangulate the sphere using the tetrahedron then $F = 4$, $E = 6$, and $V = 4$. In either case $\chi = 2$.

We will first use geometry to show that the Euler characteristic does not depend on the metric. Below we indicate how closed oriented surfaces are classified and how the Euler characteristic is constrained to be ≤ 2.

Theorem 6.5.8. *Let M be an oriented closed surface, then*

$$\int_M K dA = 2\pi\chi$$

for any polygonal subdivision of M. In particular, χ does not depend on the polygonal subdivision.

Proof. The orientation is used to ensure that integration has a consistent sign when we switch parametrizations.

We consider a polygonal subdivision with F polygons. Each n_j-gon is denoted by P_j. The local version of Gauss-Bonnet for each polygon can be written:

$$\begin{aligned}\int_{P^j} K dA &= -\int_0^{L_j} \kappa_g ds + 2\pi - \sum_{i_j=1}^{n_j} \theta_{i_j} \\ &= -\int_0^{L_j} \kappa_g ds + 2\pi - \pi n_j + \sum_{i_j=1}^{n_j} \alpha_{i_j},\end{aligned}$$

where α_{i_j} is the interior angle. The global formula is now gotten by adding up these contributions. When doing this it is important to orient each polygon so that the winding number is 1. Each edge occurs in exactly two adjacent polygons, but the edge will have opposite orientations in each of the polygons when we insist that they both have winding number 1. Thus the geodesic curvature changes sign and those terms cancel each other in the sum.

$$\begin{aligned}\int_M K dA &= \sum_{j=1}^{F} \int_{P_j} K dA \\ &= 2\pi F - \sum_{j=1}^{F} \pi n_j + \sum_{j=1}^{F} \sum_{i_j=1}^{n_j} \alpha_{i_j} \\ &= 2\pi (F - E + V).\end{aligned}$$

Here the last equality follows from the fact that at each vertex the interior angles add up to 2π, while $n_1 + \cdots + n_F = 2E$ since each edge gets counted twice in that sum.

This shows that $F - E + V$ does not depend on what subdivision we picked. Given that information we observe that $\int_M K dA$ does not vary if we change the first fundamental form on a given abstract surface as we can always use the same subdivision regardless of what the first fundamental form is. $\square$

Definition 6.5.9. The *genus* g of an orientable closed surface is the maximum number of disjoint simple closed curves with the property that the complement is connected. Orientability is used to guarantee that any simple closed curve has a well-defined right and and left hand side, i.e., it locally divides the surface in two. Globally, however, the complement might still be connected. Note that the Jordan curve theorem implies that $g = 0$ for the sphere.

Using g surgeries (see proof below) one can obtain a closed surface with $g = 0$. We shall be concerned with the opposite question: What can we say about a closed oriented surface with $g = 0$ and more generally about a surface with genus g?

It is easy to construct surfaces of genus g by adding g "handles" to a sphere. The next theorem explains why there are no other orientable surfaces.

THEOREM 6.5.10. *An oriented surface with genus g has $\chi = 2 - 2g$ and is a sphere with g handles attached.*

PROOF. We will fix a triangulation for a closed oriented surface. A *simple cycle* in a triangulation is a simple closed loop of edges and vertices, i.e. each vertex and edge only appears once as we run around in the loop. For a fixed triangulation we can redefine the genus as the maximum number of simple cycles with connected complement.

Surgery for a triangulation is defined by cutting along a simple cycle whose complement is connected and adding two *pyramids* to create a new surface with a triangulation. This reduces the genus and increases χ by 2. The latter is because the simple cycle has the same number of edges and vertices and thus does not contribute to χ. For each pyramid we add the same number of faces and edges and 1 vertex. Thus χ is increased by 1 for each of the two pyramids. Thus g such surgeries will result in a triangulated surface with $g = 0$ and where χ has been increased by $2g$. We show below that a triangulated surface with $g = 0$ has $\chi = 2$ and is a sphere. Given this, we can reverse the surgeries, also known as adding *handles* to the sphere, and conclude that the original surface is a sphere with g handles.

Recall that faces and edges do not include their boundary points. Consider a collection of faces, edges, and vertices whose union is homeomorphic to an open disc and has $\chi = V - E + F = 1$. The boundary consists of the edges and vertices that meet the faces in the collection, but are excluded from being part of the union. Since the collection is an open set it can't contain a vertex without also including all edges and faces that have the vertex on their boundaries. On the other hand, it is possible for it to contain two adjacent faces without the common edge. In particular, such a collection could contain all faces in the triangulation and still have nonempty boundary. Note that in defining χ for such a collection we only count the vertices and edges included, not the remaining vertices and edges on the boundary. The simplest example of such a collection is a single face.

The claim is that any surface contains a collection that forms an open disc with $\chi = 1$; includes all faces in the triangulation; and such that the boundary graph is connected and has no branches, i.e., there are no vertices that are met by just one edge.

Consider any collection whose union is an open disc with $\chi = 1$ and whose boundary is connected.

Faces outside this collection that meet the boundary either do so in one, two, or three edges. Regardless of which situation occurs we can add the face and exactly one of the edges that is also an edge for a face in the collection. This preserves the properties that the collection forms an open disc with $\chi = 1$. To be specific, note that the open half-disc

$$H = \left\{ (x, y) \in (-\infty, 0) \times \mathbb{R} \mid \sqrt{x^2 + y^2} < 1 \right\}$$

and with a wedge added

$$H \cup \{(x, y) \in [0, 1] \times \mathbb{R} \mid |y| < a\,(1 - x)\,,\, 0 < a \leq 1\}$$

are both homeomorphic to open discs. Note that the boundary still has all of the original vertices, one edge is deleted, and the other two edges and vertex of the added face are added. Thus the boundary stays connected. Now continue this process until all faces in the triangulation of the surface have been included.

Next we eliminate branches from the boundary. If the boundary contains a vertex that is met by exactly one edge, then add the vertex and edge to the collection. This keeps the properties that the collection forms an open disc with $\chi = 1$. To be specific, note that all of the open sets $\left\{ (x, y) \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} < 1,\ x \geq a > -1 \right\}$ are homeomorphic to open discs. When we delete a vertex and edge, we are essentially just increasing a. Clearly the boundary stays connected. Continue this until there are no branches on the boundary.

We can now characterize the sphere as the only surface with $g = 0$. In this case the boundary of the open disc with $\chi = 1$ that includes all faces can't contain any simple cycles since the complement of the boundary is the open disc and hence connected. If there are no branches, then it can only be a single vertex. This implies that $\chi = 2$ and that the surface is a sphere. □

Exercises

(1) Show that for a triangulation of a closed surface of genus g we have:
   (a) $E \leq \binom{V}{2}$,
   (b) $E \leq \binom{F}{2}$,
   (c) $2E = 3F$,
   (d) $E = 3\,(V - \chi)$,
   (e) $F \geq V$,
   (f) $V \geq \frac{1}{2}\left(7 + \sqrt{1 + 48g}\right)$,
   (g) When $g = 0$ show that at least one vertex has degree ≤ 5. The degree of a vertex is the number of edges that meet the vertex.
   (h) When $g \geq 1$ show that at least one vertex has degree $\leq \frac{1}{2}\left(7 + \sqrt{1 + 48g}\right) - 1$.

   The number $\frac{1}{2}\left(7 + \sqrt{1 + 48g}\right)$ is also known as the *coloring number* of the of the surface. The fact that any map on a surface can be colored with at most that many colors is the famous 4-coloring conjecture/theorem for the sphere. Heawood established the result for surfaces of genus $g \geq 1$ by showing that (h) holds. The same proof shows that (g) implies that maps on the sphere can be 6 colored. Heawood also showed that maps on the sphere can be 5 colored. It was not until 1968 that Ringel and Youngs showed that this is the correct coloring number for all $g \geq 1$. The 4 color problem ($g = 0$) was solved by Appel and Haken in 1977.

(2) Show that a closed surface with constant curvature has the property that the curvature and the Euler characteristic have the same sign.

(3) Show that a closed surface must have a point where K has the same sign as χ.

(4) Show that if a closed orientable surface has $K \geq R^{-2} > 0$, then its area is $\leq 4\pi R^2$. Show that if the area is $4\pi R^2$, then $K = R^{-2}$.

(5) A polyhedron consists of a collection of planar polygons that are glued together along edges of equal length. It should look like a polygonal subdivision of a surface. Thus each edge is met by exactly to planar polygons. The fact that the polygons are planar implies that the sum of the interior angles is $\pi(p-2)$, where p is the number or edges for the polygon. Let Θ_v be the sum of the interior angles of faces that meet the vertex v. The angular defect $2\pi - \Theta_v$ measures of far the vertex is from being flat in analogy with exterior angles for vertices measuring how far the edges are from being straight. Show Descartes theorem
$$2\pi\chi = \sum (2\pi - \Theta_v),$$
where $\chi = V - E + F$ as for a polygonal subdivision of a surface.

(6) A polygonal subdivision of a closed surface is said to be *cubical* if each vertex is met by exactly three edges, just as the vertices on a cube.
   (a) Show that a cubical subdivision satisfies: $2E = 3V$ and $F = \sum_n p_n$, where p_n is the number of n-gons.
   (b) Show that a cubical subdivision satisfies:
$$6\chi = \sum_n (6-n) p_n.$$
   (c) Show that a cubical subdivision of the sphere into 4-gons consists of exactly 6 quadrilaterals and looks like a cube.
   (d) Show that a cubical subdivision of the sphere into quadrilaterals and hexagons contains exactly 6 quadrilaterals. Give an example that contains hexagons.
   (e) A soccer ball is a cubical subdivision of a sphere into hexagons and pentagons. Show that it contains 12 pentagons.
   (f) If a surface has a cubical subdivision into hexagons, then $\chi = 0$. Does the torus admit a cubical subdivision into hexagons?

6.6. Closed and Convex Surfaces

First we need the equivalent of the Jordan curve theorem for closed surfaces.

PROPOSITION 6.6.1. *A closed surface $M \subset \mathbb{R}^3$ has the property that the Gauss map is onto. There are no closed space-like surfaces $M \subset \mathbb{R}^{2,1}$.*

PROOF. The proof in either case uses the function $f(p) = p \cdot n$ for a fixed $n \in \mathbb{R}^3$. Select a parametrization $\mathrm{q}(u,v)$ such that $f(\mathrm{q}(u,v))$ has a critical point at $(u,v) = (0,0)$. The first partials of this function are given by
$$\frac{\partial f \circ \mathrm{q}}{\partial w} = \frac{\partial \mathrm{q}}{\partial w} \cdot n.$$
At a critical point they have to vanish. As $\frac{\partial \mathrm{q}}{\partial u}, \frac{\partial \mathrm{q}}{\partial v}$ span the tangent space it follows that n must be a unit normal at all critical points for $f(\mathrm{q}(u,v))$. At a maximum f is positive so if the normal to the surface is outward pointing it follows that n becomes the normal to the surface.

In the case of $M \subset \mathbb{R}^{2,1}$ it follows that there will be both time-like and space-like normal vectors. That's impossible if all tangent spaces are space-like as that forces the normals to be time-like. $\square$

PROPOSITION 6.6.2. *A closed surface $M \subset \mathbb{R}^3$ has points where both principal curvatures are positive.*

PROOF. Consider the function $f(p) = \frac{1}{2}|p|^2$ restricted to M. Select a parametrization $\mathrm{q}(u,v)$ such that $f(\mathrm{q}(u,v))$ has a maximum at $(u,v) = (0,0)$. The first and second partials of this function are given by

$$\begin{aligned}\frac{\partial f \circ \mathrm{q}}{\partial w} &= \mathrm{q} \cdot \frac{\partial \mathrm{q}}{\partial w}, \\ \frac{\partial^2 f \circ \mathrm{q}}{\partial w_1 \partial w_2} &= \frac{\partial \mathrm{q}}{\partial w_1} \cdot \frac{\partial \mathrm{q}}{\partial w_2} + \mathrm{q} \cdot \frac{\partial^2 \mathrm{q}}{\partial w_1 \partial w_2} = g_{w_1 w_2} + \mathrm{q} \cdot \frac{\partial^2 \mathrm{q}}{\partial w_1 \partial w_2}.\end{aligned}$$

Since there is a maximum at $(0,0)$ the first partials vanish and we can assume that the normal to the surface is given by

$$\mathrm{n}(0,0) = -\frac{\mathrm{q}(0,0)}{|\mathrm{q}(0,0)|}.$$

The second partials at $(0,0)$ are then given by:

$$\frac{\partial^2 f \circ \mathrm{q}}{\partial w_1 \partial w_2} = g_{w_1 w_2} - |\mathrm{q}(0,0)|\, L_{w_1 w_2}.$$

The second derivative test tells us that the Hessian of $f(\mathrm{q}(u,v))$ must be nonpositive at $(0,0)$. This is equivalent to

$$\mathrm{I}(X,X) - |\mathrm{q}(0,0)|\, \mathrm{II}(X,X) \leq 0$$

for all $X \in T_{q(0,0)}M$. This in turn implies that both principal curvatures at $\mathrm{q}(0,0)$ are $\geq \frac{1}{|\mathrm{q}(0,0)|}$. □

An abstract surface with $g = 0$ also has points with $K > 0$ by the Gauss-Bonnet theorem.

THEOREM 6.6.3. *(Liebmann, 1900) If $M \subset \mathbb{R}^3$ is closed and has constant Gauss curvature, then it is a constant curvature sphere.*

PROOF. First note that the surface must have positive curvature. Since the surface is closed and $K = \kappa_1 \cdot \kappa_2$ is constant it follows that when κ_1 has a maximum, then κ_2 has a minimum. Hilbert's lemma (see section 6.3 exercise 7) then tells us that the principal curvatures must be equal and constant. □

THEOREM 6.6.4. *If $M \subset \mathbb{R}^3$ is closed, has constant mean curvature, and positive Gauss curvature, then it is a constant curvature sphere.*

PROOF. Same proof as above. □

THEOREM 6.6.5. *(Hadamard, 1897) Let $M \subset \mathbb{R}^3$ be a closed surface with $K > 0$, then the Gauss map is a diffeomorphism and M is convex.*

PROOF. Consider the signed height function to the tangent plane at a point $p \in M$:

$$f(x) = (x - p) \cdot \mathrm{n}_p\,.$$

Note that the critical points q for f are the points where $\mathrm{n}_q = \pm\,\mathrm{n}_p$. Example 5.4.5 shows that every critical point is either a strict local maximum or minimum. Assume there are two local minima at p, q and consider the min/max

$$\inf_{c \in \Omega_{p,q}} \max f \circ c.$$

If a curve achieves this min/max, then the maximum value for $f \circ c$ is also a critical point for f and consequently a local maximum for f. But this violates that there are no curves with smaller max $f \circ c$. We conclude that there is a unique minimum and maximum for f. This shows that the Gauss map is injective. The issue is to show that such a curve exists.

First observe that $\mathrm{n} : M \to S^2(1)$ is has nonsingular differential everywhere as $\det DN = K > 0$. The global Gauss-Bonnet theorem tells us that

$$0 < \int_M K dA = 2\pi\chi(M).$$

This implies that $\chi(M) = 2$.

Note that n is onto by proposition 6.6.1.

If n is not one-to-one, then we can find a small open set $O \subset M$ such that n is onto when restricted to $M - O$. This leads to the following contradiction

$$4\pi = \int_{M-O} K dA + \int_O K dA = 4\pi + \int_O K dA > 4\pi.$$

Consider the signed height function to the tangent plane at a point $p \in M$:

$$f(x) = (x - p) \cdot \mathrm{n}_p$$

This has exactly two critical points where $\mathrm{n}_x = \pm\mathrm{n}_p$. These correspond to the maximum and minimum. Assume p is the minimum. Then $f(x) > 0$ for all $x \neq p$ and the surface lies on one side of the tangent plane at p. □

Theorem 6.6.6. *Any two simple closed geodesics on a closed surface with $K > 0$ intersect.*

Proof. If they don't intersect then there is an annular region with $K > 0$ where the boundary curves have no geodesic curvature. This violates the Gauss-Bonnet theorem. See also theorem 7.8.4 for a different proof. □

Remark 6.6.7. One can reprove the results in this section for isometric immersions $F : M \to \mathbb{R}^3$ when M is oriented. In particular, it will follow that all such immersions are embeddings when $K > 0$.

Definition 6.6.8. A *Weingarten surface* is a surface where the principal curvatures depend on each other, i.e., $W(\kappa_1, \kappa_2) = 0$ for some function W. Surfaces of constant Gauss or mean curvature are examples of such surfaces, as are surfaces where all points are umbilics.

Exercises

(1) Show that surfaces of revolution are Weingarten surfaces. Hint: The principal curvatures are constant along latitudes.
(2) Show that tubes are Weingarten surfaces where one principal curvature is constant and that
$$aH + bK + c = 0$$
for suitable constants a, b, c.
(3) Give an example of a closed Weingarten surface $M \subset \mathbb{R}^3$ with $K > 0$ that does not have constant curvature.

(4) Consider a closed surface $M \subset \mathbb{R}^3$ that satisfies $H = RK$, for some constant $R > 0$. Let $O = \{p \in M \mid K(p) > 0\}$.
   (a) Show that $O \neq \emptyset$.
   (b) Show that $M = O$ by using that on O we have the relationship:
$$\frac{1}{\kappa_1} + \frac{1}{\kappa_2} = 2R.$$
   (c) Show that M is a sphere of radius R.

(5) Let $M \subset \mathbb{R}^3$ be a closed surface.
   (a) Show that $\int_M \max\{K, 0\}\, dA \geq 4\pi$.
   (b) Show that $\int_M H^2 dA \geq 4\pi$.
   (c) Show that if $\int_M H^2 dA = 4\pi$, then M is a round sphere. Hint: Show that the principal curvatures are equal and positive wherever $K > 0$. Use this to conclude that they are constant. Finally show that it is not possible to have $K \leq 0$ anywhere.

CHAPTER 7

Geodesics and Metric Geometry

This chapter covers the basics of geodesics and their properties as shortest curves. We also give models for constant curvature spaces and calculate the geodesics in these models. We discuss isometries and the local/global classification of surfaces with constant Gauss curvature. The chapter ends with a treatment of a few classical comparison theorems. Virtually all results have analogues for higher dimensional Riemannian manifolds, but certain proofs are a bit easier for surfaces. It will be noted that there is no mention of parallel translation although we do introduce second partial derivatives for 2-parameter maps in to an abstract surface. This is more or less in line with the classical treatment, as parallel translation was not introduced until the early part of the 20th century. It also eases the treatment quite a bit.

Throughout we study abstract surfaces, but note that many calculations are much easier if we think of the surfaces as sitting in $\mathbb{R}^3$.

7.1. Geodesics

Definition 7.1.1. A curve q on a surface M is called a *geodesic* if the tangential part of the acceleration vanishes, $\ddot{\mathrm{q}}^{\mathrm{I}} = 0$, specifically

$$\begin{aligned} \frac{d^2 u}{dt^2} + \Gamma^u (\dot{\mathrm{q}}, \dot{\mathrm{q}}) &= 0, \\ \frac{d^2 v}{dt^2} + \Gamma^v (\dot{\mathrm{q}}, \dot{\mathrm{q}}) &= 0. \end{aligned}$$

When $M \subset \mathbb{R}^3$ this is equivalent to saying that $\ddot{\mathrm{q}}$ is normal to the surface or that $\ddot{\mathrm{q}} = \ddot{\mathrm{q}}^{\mathrm{II}} = \mathrm{n}\,\mathrm{II}(\dot{\mathrm{q}}, \dot{\mathrm{q}})$.

Proposition 7.1.2. *A geodesic has constant speed.*

Proof. Let $\mathrm{q}(t)$ be a geodesic. We compute the derivative of the square of the speed:

$$\frac{d}{dt}\,\mathrm{I}(\dot{\mathrm{q}}, \dot{\mathrm{q}}) = \frac{d}{dt}(\dot{\mathrm{q}} \cdot \dot{\mathrm{q}}) = 2\ddot{\mathrm{q}} \cdot \dot{\mathrm{q}} = 2\,\mathrm{II}(\dot{\mathrm{q}}, \dot{\mathrm{q}})\,\mathrm{n} \cdot \dot{\mathrm{q}} = 0$$

since n and $\dot{\mathrm{q}}$ are perpendicular. Thus q has constant speed.

There is also a purely intrinsic proof that works for abstract surfaces. Since it is convenient to do this proof in a more general context it will be delayed until the end of the next section. □

Next we address existence of geodesics.

Theorem 7.1.3. *Given a point* $p = \mathrm{q}(u_0, v_0)$ *and a tangent vector* $V = V^u \frac{\partial \mathrm{q}}{\partial u}(u_0, v_0) + V^v \frac{\partial \mathrm{q}}{\partial v}(u_0, v_0) \in T_p M$ *there is a unique geodesic* $\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$

defined on some small interval $t \in (-\varepsilon, \varepsilon)$ with the initial values

$$\begin{aligned} \mathrm{q}(0) &= p, \\ \dot{\mathrm{q}}(0) &= V. \end{aligned}$$

PROOF. The existence and uniqueness part is a very general statement about solutions to differential equations (see theorem A.5.1). In this case we note that in the (u, v) parameters we must solve a system of second order equations

$$\begin{aligned} \frac{d^2u}{dt^2} &= -\begin{bmatrix} \frac{du}{dt} & \frac{dv}{dt} \end{bmatrix} \begin{bmatrix} \Gamma^u_{uu} & \Gamma^u_{uv} \\ \Gamma^u_{vu} & \Gamma^u_{vv} \end{bmatrix} \begin{bmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{bmatrix}, \\ \frac{d^2v}{dt^2} &= -\begin{bmatrix} \frac{du}{dt} & \frac{dv}{dt} \end{bmatrix} \begin{bmatrix} \Gamma^v_{uu} & \Gamma^v_{uv} \\ \Gamma^v_{vu} & \Gamma^v_{vv} \end{bmatrix} \begin{bmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{bmatrix}, \end{aligned}$$

with the initial values

$$\begin{aligned} (u(0), v(0)) &= (u_0, v_0), \\ (\dot{u}(0), \dot{v}(0)) &= (V^u, V^v). \end{aligned}$$

As long as the Christoffel symbols are sufficiently smooth there is a unique solution to such a system of equations given the initial values. The domain $(-\varepsilon, \varepsilon)$ on which such a solution exists is quite hard to determine. It'll depend on the domain of parameters U, the initial values, and Christoffel symbols. □

This theorem allows us to find all geodesics on spheres and in the plane without calculation.

EXAMPLE 7.1.4. In $\mathbb{R}^2$ straight lines $\mathrm{q}(t) = p + vt$ are clearly geodesics. Since these solve all possible initial problems there are no other geodesics.

EXAMPLE 7.1.5. On S^2 we claim that the great circles

$$\begin{aligned} \mathrm{q}(t) &= q\cos(|v|\,t) + \frac{v}{|v|}\sin(|v|\,t), \\ q &\in S^2, \\ q \cdot v &= 0 \end{aligned}$$

are geodesics. Note that this is a curve on S^2, and that $\mathrm{q}(0) = q$, $\dot{\mathrm{q}}(0) = v$. The acceleration as computed in $\mathbb{R}^3$ is given by

$$\ddot{\mathrm{q}}(t) = -q\,|v|^2\cos(|v|\,t) - v\,|v|\sin(|v|\,t) = -|v|^2\,\mathrm{q}(t)$$

and is consequently normal to the sphere. In particular $\ddot{\mathrm{q}}^{\mathrm{I}} = 0$. This means that we have also solved all initial value problems on the sphere.

Depending on our parametrization (u, v)-geodesics can be pictured in many ways. We'll study a few models or parametrizations of the sphere where geodesics take on some familiar shapes and can be described directly by equations rather than in parametrized form.

Unit Sphere Model: Consider the unit sphere. Great circles and hence geodesics are described by the two equations:

$$\begin{aligned} ax + by + cz &= 0, \\ x^2 + y^2 + z^2 &= 1. \end{aligned}$$

Given a specific geodesic $\mathrm{q}(t) = q\cos(|v|\,t) + \frac{v}{|v|}\sin(|v|\,t)$ we can use $(a, b, c) = q \times v$.

Elliptic Model: If we use the Monge patch $\left(u, v, \sqrt{1-u^2-v^2}\right)$ on the upper hemisphere, i.e., project to the (x, y)-plane along the z-axis, then the equations of the geodesics become

$$\left(a^2+c^2\right) u^2+2abuv+\left(b^2+c^2\right) v^2=c^2.$$

These are the equations of ellipses whose axes go through the origin and are inscribed in the unit circle. This is how you draw great circles on the sphere!

Recall that the level sets to quadratic equations:

$$\alpha x^2+2\beta xy+\gamma y^2=R^2$$

are ellipses centered at the origin when $\alpha+\gamma>0$ and $\alpha\gamma-\beta^2>0$.

Beltrami Model: If we use the parametrization

$$\frac{1}{\sqrt{1+u^2+v^2}}(u, v, 1)$$

on the upper hemisphere, i.e., $\frac{x}{z}=u$, $\frac{y}{z}=v$, then these equations simply become straight lines in (u, v) coordinates:

$$au+bv+c=0.$$

This reparametrization was also discussed in section 4.4 exercise 21, where it was called the Beltrami projection. It is simply the projection of the upper hemisphere along radial lines to the tangent plane $\{z=1\}$ at the North pole.

Conformal Model: The radial projection that was used for the Beltrami model is an example of a perspective projection, i.e., a projection along radial lines from a point to a plane that does not pass through this point. The stereographic parametrization from section 4.4 exercise 19 is projection along lines through $(0,0,-1)$ to the (x, y)-plane. In this model the upper hemisphere is parametrized as

$$\mathrm{q}^{+}(u, v)=\frac{1}{1+u^2+v^2}\left(2u, 2v, 1-u^2+v^2\right).$$

One can show that this is a conformal or isothermal parametrization. In case $c=0$ the geodesics are straight lines through the origin:

$$au+bv=0.$$

When $c \neq 0$ we can normalize so that $c=1$ in which case the geodesics become circles

$$(u+a)^2+(v+b)^2=1+a^2+b^2.$$

Next we consider hyperbolic space.

Imaginary Unit Sphere Model: We defined hyperbolic space $H \subset \mathbb{R}^{2,1}$ in example 6.1.2 as the the imaginary unit sphere with $z>0$. More precisely, it is the rotationally symmetric surface

$$x^2+y^2-z^2=-1,\ z \geq 1$$

with a metric that is inherited from the space-time inner product structure. Observe that the tangent space can be characterized as

$$T_q M=\left\{v \in \mathbb{R}^{2,1} \mid v \cdot q=0\right\}.$$

This means that the normal is be given by $\mathrm{n}(q) = q$. In analogy with the sphere we consider the curves

$$\begin{aligned} \mathrm{q}(t) &= q\cosh(|v|\,t) + \frac{v}{|v|}\sinh(|v|\,t), \\ q &\in H, \\ v &\in T_qH. \end{aligned}$$

Since $q \cdot v = 0$ this is a curve on H with $\mathrm{q}(0) = q$, $\dot{\mathrm{q}}(0) = v$. Note also that it lies in the plane spanned by q and v.

The acceleration as computed in $\mathbb{R}^{2,1}$ is given by

$$\ddot{\mathrm{q}}(t) = q\,|v|^2\cosh(|v|\,t) + v\,|v|\sinh(|v|\,t) = |v|^2\,\mathrm{q}(t).$$

In particular, it has no tangential component and thus has vanishing intrinsic acceleration (see also remark 6.1.5).

If we use $(a, b, c) = q \times v$, then we also obtain the equation form:

$$\begin{aligned} ax + by + cz &= 0, \\ x^2 + y^2 - z^2 &= -1,\ z \geq 1. \end{aligned}$$

Note that for these planes to intersect the surface it is necessary to assume that:

$$c^2 < a^2 + b^2.$$

Hyperbolic Model: This is the orthogonal projection onto the (x, y)-plane. The parametrization is a Monge patch and is given by $\left(u, v, \sqrt{1 + u^2 + v^2}\right)$. The geodesics will be straight lines through the origin when $c = 0$ and hyperbolas whose asymptotes are lines through the origin when $0 < c^2 < a^2 + b^2$:

$$\left(a^2 - c^2\right)u^2 + 2abuv + \left(b^2 - c^2\right)v^2 = c^2.$$

Recall that the level sets to quadratic equations:

$$\alpha x^2 + 2\beta xy + \gamma y^2 = R^2$$

are hyperbolas with asymptotes that pass through the origin when $\alpha\gamma - \beta^2 < 0$.

Beltrami Model: The Beltrami model comes from a perspective projection along radial lines through the origin to the plane $z = 1$. It gives us the parametrization

$$\frac{1}{\sqrt{1 - u^2 - v^2}}(u, v, 1),\ u^2 + v^2 < 1.$$

And the geodesics are straight lines:

$$au + bv + c = 0.$$

Conformal Models: Stereographic projection along radial lines through $(0, 0, -1)$ to the (x, y)-plane gives the *Poincaré model*. The parametrization is given by:

$$\frac{1}{1 - u^2 - v^2}\left(2u, 2v, 1 + u^2 + v^2\right),\ u^2 + v^2 < 1.$$

It is also called the unit disc model since the open disc is the domain for the parameters. One can show that this parametrization is conformal or isothermal. The geodesics are either straight lines through the origin

$$au + bv = 0$$

or when $c \neq 0$ and we scale so that $c = 1$ circles centered outside the unit disc:

$$(u-a)^2 + (v-b)^2 = a^2 + b^2 - 1, \ a^2 + b^2 > 1.$$

The upper half plane model comes from a conformal transformation of the upper half plane to the unit disc (see section 4.4 exercise 20). This map is given by

$$F(x,y) = \frac{1}{x^2 + (y+1)^2} \left(2x, 1 - x^2 - y^2\right).$$

The geodesics will again be lines and circles but F does not necessarily take lines to lines. The lines are all vertical:

$$x = 0, \text{ when } c = 0, \, b = 0,$$

or

$$x = {}^1\!/\!_a, \text{ when } c = 1, \, b = -1,$$

and the circles have centers along the x-axis

$$\left(x - \frac{a}{b}\right)^2 + y^2 = 1 + \frac{a^2}{b^2}, \text{ when } c = 0,$$

or

$$\left(x - \frac{a}{b+1}\right)^2 + y^2 = \frac{a^2 + b^2 - 1}{(b+1)^2}, \text{ when } c = 1.$$

It is interesting to note that for the sphere only the unit sphere model actually covers the entire sphere. In contrast, all of the models for hyperbolic space are equivalent in the sense that they are models for all of hyperbolic space, not just part of it.

Definition 7.1.6. An abstract surface is said to be *geodesically complete* if all geodesics exist for all time $t \in \mathbb{R}$. It is said to be *geodesically complete at a point*, if all geodesics through that point are defined for all time.

Example 7.1.7. The unit sphere, all of the above models for hyperbolic space, and all planes are geodesically complete.

As we have seen, it is often simpler to find the unparametrized form of the geodesics, i.e., in a given parametrization they are easier to find as an equation or as functions $u(v)$ or $v(u)$. There is in fact a tricky characterization of geodesics that does not refer to the arc-length parameter. The idea is that a regular curve can be reparametrized to be a geodesic if and only if its tangential acceleration $\ddot{\mathrm{q}}^{\mathrm{I}}$ is tangent to the curve.

Lemma 7.1.8. *A regular curve* $\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$ *can be reparametrized as a geodesic if and only if*

$$\frac{dv}{dt}\left(\frac{d^2u}{dt^2} + \Gamma^u(\dot{\mathrm{q}}, \dot{\mathrm{q}})\right) = \frac{du}{dt}\left(\frac{d^2v}{dt^2} + \Gamma^v(\dot{\mathrm{q}}, \dot{\mathrm{q}})\right).$$

Proof. First observe that this formula holds iff $\lambda(t)\,\dot{\mathrm{q}}(t) = \ddot{\mathrm{q}}^{\mathrm{I}}(t)$ for some function λ.

If we reparametrize the curve, then the velocity satisfies: $\dot{\mathrm{q}}(t) = \frac{ds}{dt}\dot{\mathrm{q}}(s)$. For the acceleration we calculate in coordinates:

$$\begin{aligned}
\frac{d^2u}{dt^2} + \Gamma^u(\dot{\mathrm{q}}, \dot{\mathrm{q}}) &= \frac{d^2s}{dt^2}\frac{du}{ds} + \left(\frac{ds}{dt}\right)^2 \frac{d^2u}{ds^2} + \Gamma^u\left(\frac{ds}{dt}\frac{d\,\mathrm{q}}{ds}, \frac{ds}{dt}\frac{d\,\mathrm{q}}{ds}\right) \\
&= \frac{d^2s}{dt^2}\frac{du}{ds} + \left(\frac{ds}{dt}\right)^2 \frac{d^2u}{ds^2} + \left(\frac{ds}{dt}\right)^2 \Gamma^u\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right) \\
&= \frac{d^2s}{dt^2}\frac{du}{ds} + \left(\frac{ds}{dt}\right)^2 \left(\frac{d^2u}{ds^2} + \Gamma^u\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right)\right).
\end{aligned}$$

Similarly

$$\frac{d^2v}{dt^2} + \Gamma^v(\dot{\mathrm{q}}, \dot{\mathrm{q}}) = \frac{d^2s}{dt^2}\frac{dv}{ds} + \left(\frac{ds}{dt}\right)^2 \left(\frac{d^2v}{ds^2} + \Gamma^v\left(\frac{d\,\mathrm{q}}{ds}, \frac{d\,\mathrm{q}}{ds}\right)\right).$$

It follows that

$$\ddot{\mathrm{q}}^{\mathrm{I}}(t) = \frac{d^2s}{dt^2}\dot{\mathrm{q}}(s) + \left(\frac{ds}{dt}\right)^2 \ddot{\mathrm{q}}^{\mathrm{I}}(s).$$

This shows first of all that, if $\mathrm{q}(s)$ is a geodesic, then $\ddot{\mathrm{q}}^{\mathrm{I}}(t) = \frac{d^2s}{dt^2}\dot{\mathrm{q}}(s)$ as claimed. Conversely assume that $\lambda(t)\dot{\mathrm{q}}(t) = \ddot{\mathrm{q}}^{\mathrm{I}}(t)$. Then

$$\lambda(s)\frac{ds}{dt}\dot{\mathrm{q}}(s) = \frac{d^2s}{dt^2}\dot{\mathrm{q}}(s) + \left(\frac{ds}{dt}\right)^2 \ddot{\mathrm{q}}^{\mathrm{I}}(s).$$

So $\ddot{\mathrm{q}}^{\mathrm{I}}(s) = \mu(s)\dot{\mathrm{q}}(s)$ for some function μ. If we assume that s is the arclength parameter, then we also know that

$$\begin{aligned}
0 &= \mathrm{I}\left(\ddot{\mathrm{q}}^{\mathrm{I}}(s), \dot{\mathrm{q}}(s)\right) \\
&= \mu(s).
\end{aligned}$$

This shows that $\ddot{\mathrm{q}}^{\mathrm{I}}(s) = 0$. □

Exercises

(1) Let $\mathrm{q}(t)$ be a regular curve on a surface with normal n. Show that it can be reparametrized to become a geodesic if and only if

$$\det[\dot{\mathrm{q}}, \ddot{\mathrm{q}}, \mathrm{n}] = 0.$$

(2) Let $\mathrm{q}(t)$ be a unit speed curve on a surface. Show that

$$|\kappa_g| = \left|\ddot{\mathrm{q}}^{\mathrm{I}}\right|.$$

(3) Consider a unit speed curve $\mathrm{q}(t)$ on a surface of revolution

$$\mathrm{q}(u, \mu) = (r(u)\cos\mu, r(u)\sin\mu, z(u)),$$

where the profile curve $(r(u), 0, z(u))$ is unit speed. Let $\theta(t)$ denote the angle with the meridians.

(a) Show that if $\dot{\mathrm{q}}(t) = \dot{u}\partial_u \mathrm{q} + \dot{v}\partial_v \mathrm{q}$, then

$$\dot{u} = \cos\theta \text{ and } \dot{v} = \frac{\sin\theta}{r}.$$

(b) Show that if $\mathrm{q}(t)$ is a geodesic, then

$$\begin{aligned}\ddot{u} &= (\dot{v})^2\, r\frac{\partial r}{\partial u},\\ \ddot{v} &= -2\dot{u}\dot{v}\frac{\partial \log r}{\partial u}.\end{aligned}$$

(c) (Clairaut) Show that $\mathrm{I}(\dot{\mathrm{q}}, \partial_v) = r\sin\theta$ is constant along a geodesic.

(d) We say that $\mathrm{q}(t)$ is a *loxodrome* if θ is constant. Show that if all geodesics are loxodromes, then the surface is a cylinder.

(4) Let $\mathrm{q}(t)$ be a unit speed geodesic on a surface in $\mathbb{R}^3$. Show that

$$\begin{aligned}0 &= \kappa_g,\\ \kappa &= \kappa_n,\\ \tau &= \tau_g,\end{aligned}$$

where κ and τ are the curvature and torsion of $\mathrm{q}(t)$ as a space curve.

(5) Consider the two parametrized surfaces

$$\begin{aligned}\mathrm{q}(r,\mu) &= (r\cos\mu, r\sin\mu, \log r),\\ \mathrm{q}^*(r,\mu) &= (r\cos\mu, r\sin\mu, \mu).\end{aligned}$$

(a) Show that their first fundamental forms are not equal.

(b) Show that they have the same Gauss curvature $K(r,\mu) = K^*(r,\mu)$.

(c) Show that the surfaces are not isometric.

(6) Let $\mathrm{q}(t)$ be a unit speed geodesic on a surface in space with curvature κ and torsion τ as in the previous exercise.

(a) Show that $\kappa = \kappa_1\cos^2\theta + \kappa_2\sin^2\theta$, where $\kappa_{1,2}$ are the principal curvatures and θ the angles between $\dot{\mathrm{q}}$ and the first principal direction.

(b) Show that $\tau^2 = (\kappa - \kappa_1)(\kappa_2 - \kappa)$.

(c) Show that if $\kappa_1 = 0$, then $\kappa = \tau\tan\theta$.

(7) Show that in the conformal model of the unit sphere the geodesics that pass through $(u,v) = (1,0)$ all have center on the v-axis. Show that all initial value problems can be solved.

(8) Show that the ϕ curves on a tube

$$\mathrm{q}(t,\phi) = c(t) + R(\mathrm{N}_c\cos\phi + \mathrm{B}_c\sin\phi),$$

(see section 4.1 exercise 6 and section 4.3 exercise 7) are geodesics.

(9) Show that if a unit speed curve on a surface also lies in a plane that is perpendicular to the surface, then it is a geodesic.

(10) Show that geodesics satisfy a second order equation

$$\frac{d^2v}{du^2} = \Gamma^u_{vv}\left(\frac{dv}{du}\right)^3 + (2\Gamma^u_{uv} - \Gamma^v_{vv})\left(\frac{dv}{du}\right)^2 + (\Gamma^u_{uu} - 2\Gamma^v_{uv})\frac{dv}{du} - \Gamma^v_{uu}.$$

(11) (Beltrami) Assume that $\mathrm{q}(u,v)$ is a parametrized surface with the property that all geodesics are lines in the domain U, i.e., each geodesic satisfies an equation of the form

$$au + bv + c = 0,\ (a,b) \neq (0,0).$$

Note that all formulas below remain the same when u and v are interchanged. This reduces the number of calculations that need to be done.

(a) Show using exercise 10 that

$$\begin{aligned} \Gamma_{uu}^{v} &= \Gamma_{vv}^{u} = 0, \\ \Gamma_{uu}^{u} &= 2\Gamma_{uv}^{v}, \\ \Gamma_{vv}^{v} &= 2\Gamma_{uv}^{u}. \end{aligned}$$

Hint: Use lemma 7.1.8 and parametrize the curve by u or v.

(b) Show that

$$\begin{bmatrix} \partial_v g_{uu} - \partial_u g_{uv} \\ \partial_u g_{vv} - \partial_v g_{uv} \end{bmatrix} = \begin{bmatrix} g_{uu} & -g_{uv} \\ -g_{vu} & g_{vv} \end{bmatrix} \begin{bmatrix} \Gamma_{uv}^{u} \\ \Gamma_{uv}^{v} \end{bmatrix}.$$

(c) Recall from proposition 6.2.4 that

$$K \begin{bmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{bmatrix} = \begin{bmatrix} R_{vuu}^{v} & R_{vuv}^{v} \\ R_{uvu}^{u} & R_{uvv}^{u} \end{bmatrix}.$$

Use the definition of R_{ijk}^{l}, part (a), and for the second equality, $R_{uvu}^{u} = R_{vuv}^{v}$, to show that

$$\begin{aligned} \begin{bmatrix} R_{vuu}^{v} & R_{vuv}^{v} \\ R_{uvu}^{u} & R_{uvv}^{u} \end{bmatrix} &= \begin{bmatrix} -\partial_u \Gamma_{vu}^{v} & \partial_v \Gamma_{vu}^{v} - 2\partial_u \Gamma_{uv}^{u} \\ \partial_u \Gamma_{uv}^{u} - 2\partial_v \Gamma_{uv}^{v} & -\partial_v \Gamma_{uv}^{u} \end{bmatrix} + \begin{bmatrix} \Gamma_{vu}^{v}\Gamma_{vu}^{v} & \Gamma_{uv}^{u}\Gamma_{vu}^{v} \\ \Gamma_{vu}^{v}\Gamma_{uv}^{u} & \Gamma_{uv}^{u}\Gamma_{uv}^{u} \end{bmatrix} \\ &= -\begin{bmatrix} \partial_u \Gamma_{vu}^{v} & \partial_v \Gamma_{vu}^{v} \\ \partial_u \Gamma_{uv}^{u} & \partial_v \Gamma_{uv}^{u} \end{bmatrix} + \begin{bmatrix} \Gamma_{vu}^{v}\Gamma_{vu}^{v} & \Gamma_{uv}^{u}\Gamma_{vu}^{v} \\ \Gamma_{vu}^{v}\Gamma_{uv}^{u} & \Gamma_{uv}^{u}\Gamma_{uv}^{u} \end{bmatrix}. \end{aligned}$$

(d) Use (c) to show that

$$\begin{aligned} \begin{bmatrix} \partial_v (K g_{uu}) - \partial_u (K g_{uv}) \\ \partial_u (K g_{vv}) - \partial_v (K g_{vu}) \end{bmatrix} &= \begin{bmatrix} \partial_v R_{vuu}^{v} - \partial_u R_{vuv}^{v} \\ \partial_u R_{uvv}^{u} - \partial_v R_{uvu}^{u} \end{bmatrix} \\ &= \begin{bmatrix} R_{vuu}^{v} - R_{uvu}^{u} \\ R_{uvv}^{u} - R_{vuv}^{v} \end{bmatrix} \begin{bmatrix} \Gamma_{uv}^{u} \\ \Gamma_{uv}^{v} \end{bmatrix} \\ &= K \begin{bmatrix} g_{uu} & -g_{uv} \\ -g_{vu} & g_{vv} \end{bmatrix} \begin{bmatrix} \Gamma_{uv}^{u} \\ \Gamma_{uv}^{v} \end{bmatrix}. \end{aligned}$$

(e) Use (b) and (d) to show that

$$0 = \begin{bmatrix} \partial_v K & \partial_u K \\ \partial_v K & \partial_u K \end{bmatrix} \begin{bmatrix} g_{uu} & -g_{uv} \\ -g_{vu} & g_{vv} \end{bmatrix}.$$

Conclude that the Gauss curvature is constant.

7.2. Mixed Partials

We need to generalize the intrinsic acceleration to also include mixed partial derivatives. The formulas obtained in section 5.2 will guide us.

Instead of just having a curve $\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$ within a parametrization we assume that we have a family of curves $\mathrm{q}(s,t) = \mathrm{q}(u(s,t), v(s,t))$ such that for each s there is a curve parametrized by t. We shall generally assume that $(s,t) \in (-\epsilon, \epsilon) \times [a,b]$. In this case such a family of curves is called a *variation* of the base curve $\mathrm{q}(t) = \mathrm{q}(0,t)$. Note that $\mathrm{q}(s,t)$ does not have to be a valid parametrization of the surface.

To ease the notation we will use the conventions $\mathrm{q}^w(s,t) = w(s,t)$ so that we can write $\partial_s \mathrm{q}^w = \frac{\partial w}{\partial s}$, $\partial_t \partial_s \mathrm{q}^w = \frac{\partial^2 w}{\partial t \partial s}$, etc, and also use $\partial_t \partial_s \mathrm{q}^i$ with i in place of

w. We also define

$$\Gamma^{w}(X,Y)=\sum_{i,j=u,v}\Gamma_{ij}^{w}X^{i}Y^{j}=\left[\begin{array}{cc}X^{u} & X^{v}\end{array}\right]\left[\begin{array}{cc}\Gamma_{uu}^{w} & \Gamma_{uv}^{w}\\ \Gamma_{vu}^{w} & \Gamma_{vv}^{w}\end{array}\right]\left[\begin{array}{c}Y^{u}\\ Y^{v}\end{array}\right].$$

Keeping t or s fixed we already have that

$$\begin{aligned}\left(\frac{\partial^{2}\mathrm{q}}{\partial s^{2}}\right)^{\mathrm{I}}(s,t) &= \left(\partial_{s}^{2}u+\Gamma^{u}(\partial_{s}\mathrm{q},\partial_{s}\mathrm{q})\right)\partial_{u}\mathrm{q}+\left(\partial_{s}^{2}v+\Gamma^{v}(\partial_{s}\mathrm{q},\partial_{s}\mathrm{q})\right)\partial_{v}\mathrm{q}\\ &= \sum_{i=u,v}\left(\partial_{s}^{2}\mathrm{q}^{i}+\Gamma^{i}(\partial_{s}\mathrm{q},\partial_{s}\mathrm{q})\right)\partial_{i}\mathrm{q}.\end{aligned}$$

and

$$\left(\frac{\partial^{2}\mathrm{q}}{\partial t^{2}}\right)^{\mathrm{I}}(s,t)=\sum_{i=u,v}\left(\partial_{t}^{2}\mathrm{q}^{i}+\Gamma^{i}(\partial_{t}\mathrm{q},\partial_{t}\mathrm{q})\right)\partial_{i}\mathrm{q}.$$

Moreover, when the surface lies in $\mathbb{R}^{3}$, then these intrinsic second partials are in fact the tangential components of the second partials in $\mathbb{R}^{3}$.

The intrinsic mixed partial is similarly defined as

$$\left(\frac{\partial^{2}\mathrm{q}}{\partial s\partial t}\right)^{\mathrm{I}}(s,t)=\sum_{i=u,v}\left(\partial_{s}\partial_{t}\mathrm{q}^{i}+\Gamma^{i}(\partial_{s}\mathrm{q},\partial_{t}\mathrm{q})\right)\partial_{i}\mathrm{q}.$$

This mixed partial also commutes commutes since

$$\frac{\partial^{2}w}{\partial s\partial t}=\frac{\partial^{2}w}{\partial t\partial s}$$

and

$$\Gamma^{w}\left(\frac{\partial\mathrm{q}}{\partial s},\frac{\partial\mathrm{q}}{\partial t}\right)=\Gamma^{w}\left(\frac{\partial\mathrm{q}}{\partial t},\frac{\partial\mathrm{q}}{\partial s}\right).$$

We can also show that all possible product formulas for taking derivatives hold:

$$\begin{aligned}\partial_{s}\mathrm{I}(\partial_{s}\mathrm{q},\partial_{t}\mathrm{q}) &= \mathrm{I}\left(\left(\partial_{s}^{2}\mathrm{q}\right)^{\mathrm{I}},\partial_{t}\mathrm{q}\right)+\mathrm{I}\left(\partial_{s}\mathrm{q},(\partial_{s}\partial_{t}\mathrm{q})^{\mathrm{I}}\right),\\ \partial_{s}\mathrm{I}(\partial_{t}\mathrm{q},\partial_{t}\mathrm{q}) &= 2\,\mathrm{I}\left(\partial_{t}\mathrm{q},(\partial_{s}\partial_{t}\mathrm{q})^{\mathrm{I}}\right),\\ \partial_{s}\mathrm{I}(\partial_{s}\mathrm{q},\partial_{s}\mathrm{q}) &= 2\,\mathrm{I}\left(\left(\partial_{s}^{2}\mathrm{q}\right)^{\mathrm{I}},\partial_{s}\mathrm{q}\right),\\ \partial_{t}\mathrm{I}(\partial_{s}\mathrm{q},\partial_{t}\mathrm{q}) &= \mathrm{I}\left((\partial_{t}\partial_{s}\mathrm{q})^{\mathrm{I}},\partial_{t}\mathrm{q}\right)+\mathrm{I}\left(\partial_{s}\mathrm{q},\left(\partial_{t}^{2}\mathrm{q}\right)^{\mathrm{I}}\right),\\ \partial_{t}\mathrm{I}(\partial_{s}\mathrm{q},\partial_{s}\mathrm{q}) &= 2\,\mathrm{I}\left(\partial_{s}\mathrm{q},(\partial_{t}\partial_{s}\mathrm{q})^{\mathrm{I}}\right),\\ \partial_{t}\mathrm{I}(\partial_{t}\mathrm{q},\partial_{t}\mathrm{q}) &= 2\,\mathrm{I}\left(\partial_{t}\mathrm{q},\left(\partial_{t}^{2}\mathrm{q}\right)^{\mathrm{I}}\right).\end{aligned}$$

The proofs are all similar so we concentrate on the first. The essential idea is that we have the product formula

$$\partial_{s}g_{ij}=\Gamma_{sij}+\Gamma_{sji}$$

directly from the abstract definition of the Christoffel symbols as in section 6.1.

$$
\begin{aligned}
&\partial_s \mathrm{I}(\partial_s \mathrm{q}, \partial_t \mathrm{q}) \\
&= \partial_s \left(g_{ij} \partial_s \mathrm{q}^i \partial_t \mathrm{q}^j \right) \\
&= \partial_s (g_{ij}) \partial_s \mathrm{q}^i \partial_t \mathrm{q}^j + g_{ij} \partial_s^2 \mathrm{q}^i \partial_t \mathrm{q}^j + g_{ij} \partial_s \mathrm{q}^i \partial_s \partial_t \mathrm{q}^j \\
&= (\partial_k g_{ij}) \partial_s \mathrm{q}^k \partial_s \mathrm{q}^i \partial_t \mathrm{q}^j + g_{ij} \partial_s^2 \mathrm{q}^i \partial_t \mathrm{q}^j + g_{ij} \partial_s \mathrm{q}^i \partial_s \partial_t \mathrm{q}^j \\
&= (\Gamma_{kij} + \Gamma_{kji}) \partial_s \mathrm{q}^k \partial_s \mathrm{q}^i \partial_t \mathrm{q}^j + g_{ij} \partial_s^2 \mathrm{q}^i \partial_t \mathrm{q}^j + g_{ij} \partial_s \mathrm{q}^i \partial_s \partial_t \mathrm{q}^j \\
&= \Gamma_{kij} \partial_s \mathrm{q}^k \partial_s \mathrm{q}^i \partial_t \mathrm{q}^j + g_{ij} \partial_s^2 \mathrm{q}^i \partial_t \mathrm{q}^j \\
&\quad + \partial_s \mathrm{q}^i \Gamma_{kji} \partial_s \mathrm{q}^k \partial_t \mathrm{q}^j + g_{ij} \partial_s \mathrm{q}^i \partial_s \partial_t \mathrm{q}^j \\
&= g_{lj} \Gamma_{ki}^l \partial_s \mathrm{q}^k \partial_s \mathrm{q}^i \partial_t \mathrm{q}^j + g_{ij} \partial_s^2 \mathrm{q}^i \partial_t \mathrm{q}^j \\
&\quad + g_{il} \partial_s \mathrm{q}^i \Gamma_{kj}^l \partial_s \mathrm{q}^k \partial_t \mathrm{q}^j + g_{ij} \partial_s \mathrm{q}^i \partial_s \partial_t \mathrm{q}^j \\
&= g_{ij} \Gamma_{kl}^i \partial_s \mathrm{q}^k \partial_s \mathrm{q}^l \partial_t \mathrm{q}^j + g_{ij} \partial_s^2 \mathrm{q}^i \partial_t \mathrm{q}^j \\
&\quad + g_{ij} \partial_s \mathrm{q}^i \Gamma_{kl}^j \partial_s \mathrm{q}^k \partial_t \mathrm{q}^l + g_{ij} \partial_s \mathrm{q}^i \partial_s \partial_t \mathrm{q}^j \\
&= g_{ij} \left(\Gamma^i (\partial_s \mathrm{q}, \partial_s \mathrm{q}) + \partial_s^2 \mathrm{q}^i \right) \partial_t \mathrm{q}^j \\
&\quad + g_{ij} \partial_s \mathrm{q}^i \left(\Gamma_{kl}^j \partial_s \mathrm{q}^k \partial_t \mathrm{q}^l + \partial_s \partial_t \mathrm{q}^j \right) \\
&= \mathrm{I}\left(\left(\partial_s^2 \mathrm{q}\right)^{\mathrm{I}}, \partial_t \mathrm{q} \right) + \mathrm{I}\left(\partial_s \mathrm{q}, \left(\partial_s \partial_t \mathrm{q}\right)^{\mathrm{I}} \right).
\end{aligned}
$$

Finally we should also justify why these second partial derivatives do not depend on the initial (u, v)-parametrization. This could be done via a notationally nasty change of parameters or by a more general formula that doesn't depend a parametrization. This general formula, however, also has a defect in that it involves a new variable r so that $w = w(r, s, t)$:

$$
2\,\mathrm{I}\left(\left(\partial_s \partial_t \mathrm{q}\right)^{\mathrm{I}}, \partial_r \mathrm{q} \right) = \partial_s \mathrm{I}(\partial_t \mathrm{q}, \partial_r \mathrm{q}) + \partial_t \mathrm{I}(\partial_s \mathrm{q}, \partial_r \mathrm{q}) - \partial_r \mathrm{I}(\partial_s \mathrm{q}, \partial_t \mathrm{q}).
$$

Here the right hand side can be calculated independently of a (u, v)-parametrization. Since we can think of the r-variable as being anything we please, this implicitly calculates $(\partial_s \partial_t \mathrm{q})^{\mathrm{I}}$. The proof of this identity comes from from using the product rule on each on the terms on the right hand side and using that the intrinsic mixed partials commute:

$$
\begin{aligned}
&\partial_s \mathrm{I}(\partial_t \mathrm{q}, \partial_r \mathrm{q}) + \partial_t \mathrm{I}(\partial_s \mathrm{q}, \partial_r \mathrm{q}) - \partial_r \mathrm{I}(\partial_s \mathrm{q}, \partial_t \mathrm{q}) \\
&= \mathrm{I}\left(\left(\partial_s \partial_t \mathrm{q}\right)^{\mathrm{I}}, \partial_r \mathrm{q} \right) + \mathrm{I}\left(\partial_t \mathrm{q}, \left(\partial_s \partial_r \mathrm{q}\right)^{\mathrm{I}} \right) \\
&\quad + \mathrm{I}\left(\left(\partial_t \partial_s \mathrm{q}\right)^{\mathrm{I}}, \partial_r \mathrm{q} \right) + \mathrm{I}\left(\partial_s \mathrm{q}, \left(\partial_t \partial_r \mathrm{q}\right)^{\mathrm{I}} \right) \\
&\quad - \left(\mathrm{I}\left(\left(\partial_r \partial_s \mathrm{q}\right)^{\mathrm{I}}, \partial_t \mathrm{q} \right) + \mathrm{I}\left(\partial_s \mathrm{q}, \left(\partial_r \partial_t \mathrm{q}\right)^{\mathrm{I}} \right) \right) \\
&= 2\,\mathrm{I}\left(\left(\partial_s \partial_t \mathrm{q}\right)^{\mathrm{I}}, \partial_r \mathrm{q} \right).
\end{aligned}
$$

Proposition 7.2.1. *Let* $\mathrm{q}(t)$ *be a curve on a surface* M. q *has constant speed if and only if its intrinsic acceleration is perpendicular to the speed.*

Proof. The proof is now a simple calculation using the product rule for intrinsic second derivatives:

$$
\frac{d}{dt} \mathrm{I}(\dot{\mathrm{q}}, \dot{\mathrm{q}}) = 2\,\mathrm{I}\left(\ddot{\mathrm{q}}^{\mathrm{I}}, \dot{\mathrm{q}}\right).
$$

□

Exercises

(1) Consider a curve $\mathrm{q}(t) = \mathrm{q}(u(t), v(t))$ on an abstract parameterized surface.

(a) Show that $\mathrm{q}(t)$ is a geodesic if and only if

$$\begin{aligned}
\frac{d}{dt}\mathrm{I}(\dot{\mathrm{q}}, \partial_u) &= \mathrm{I}\left(\dot{\mathrm{q}}, \frac{d}{dt}\partial_u\right), \\
\frac{d}{dt}\mathrm{I}(\dot{\mathrm{q}}, \partial_v) &= \mathrm{I}\left(\dot{\mathrm{q}}, \frac{d}{dt}\partial_v\right).
\end{aligned}$$

(b) Show that $\mathrm{q}(t)$ is a geodesic if and only if

$$\begin{aligned}
\frac{d}{dt}(g_{uu}\dot{u} + g_{uv}\dot{v}) &= \frac{1}{2}\left(\partial_u g_{uu}\dot{u}^2 + 2\partial_u g_{uv}\dot{u}\dot{v} + \partial_u g_{vv}\dot{v}^2\right), \\
\frac{d}{dt}(g_{vu}\dot{u} + g_{vv}\dot{v}) &= \frac{1}{2}\left(\partial_v g_{uu}\dot{u}^2 + 2\partial_v g_{uv}\dot{u}\dot{v} + \partial_v g_{vv}\dot{v}^2\right).
\end{aligned}$$

These formulas are often quite convenient as they do not explicitly involve Christoffel symbols.

(2) A *Liouville surface* has a first fundamental form where

$$g_{uu} = g_{vv} = U - V \text{ and } g_{uv} = 0$$

and U is a function of u and V a function of v. Consider a unit speed geodesic $\mathrm{q}(t)$ on such as surface and let $\theta(t)$ be the angle the geodesic forms with the u-curves.

(a) Show that if we write $\dot{\mathrm{q}} = \dot{u}\frac{\partial \mathrm{q}}{\partial u} + \dot{v}\frac{\partial \mathrm{q}}{\partial v}$, then

$$\dot{u}^2 = \frac{\cos^2\theta}{U - V} \text{ and } \dot{v}^2 = \frac{\sin^2\theta}{U - V}.$$

(b) Show that $\mathrm{q}(t)$ satisfies

$$\begin{aligned}
\frac{d}{dt}((U - V)\dot{u}) &= \frac{1}{2}\partial_u U\left(\dot{u}^2 + \dot{v}^2\right), \\
\frac{d}{dt}((U - V)\dot{v}) &= -\frac{1}{2}\partial_v V\left(\dot{u}^2 + \dot{v}^2\right).
\end{aligned}$$

(c) Show that

$$\frac{d}{dt}\left(\frac{V\dot{u}^2 + U\dot{v}^2}{\dot{u}^2 + \dot{v}^2}\right) = 0.$$

(d) Conclude that $U\sin^2\theta + V\cos^2\theta$ is constant along geodesics.

(3) Consider two Liouville surfaces

$$g_{uu} = U - V = g_{vv} \text{ and } g_{uv} = 0$$

and

$$g'_{uu} = \left(\frac{1}{V} - \frac{1}{U}\right)\frac{1}{U},\ g'_{vv} = \left(\frac{1}{V} - \frac{1}{U}\right)\frac{1}{V},\ g'_{uv} = 0,$$

where U is a function of u and V a function of v. Show that the equations of the geodesics on these two surfaces, in the sense of section 7.1 exercise 10, are the same.

(4) Consider an abstract surface with first fundamental form I and the conformally related first fundamental form $e^{2f}\,\mathrm{I}$ for some function f. Show that

$$(\partial_s\partial_t\,\mathrm{q})^{e^{2f}\,\mathrm{I}} = (\partial_s\partial_t\,\mathrm{q})^{\mathrm{I}} + \partial_s\,\mathrm{q}\,df\,(\partial_t\,\mathrm{q}) + \partial_t\,\mathrm{q}\,df\,(\partial_s\,\mathrm{q}) - \mathrm{I}\,(\partial_t\,\mathrm{q},\partial_s\,\mathrm{q})\,\nabla f.$$

Here $\nabla f = g^{ij}\partial_i f\partial_j$ and $df\,(X) = \partial_i f X^i$.

(5) Use the previous exercise to show that if all geodesics for I can be reparametrized to also be geodesics for $e^{2f}\,\mathrm{I}$, then f is constant.

7.3. Shortest Curves

The goal is to show that the shortest curves are geodesics.

In the last section we considered variations $\mathrm{q}\,(s,t) = \mathrm{q}\,(u\,(s,t)\,,v\,(s,t))$ where $(s,t) \in (-\epsilon,\epsilon)\times[a,b]$. The variational field of $\mathrm{q}\,(t) = \mathrm{q}\,(0,t)$ is given by the tangent vectors $V\,(t) = \frac{\partial\,\mathrm{q}}{\partial s}\,(0,t)$ along the curve. The first proposition shows that any such field $V\,(t) \in T_{\mathrm{q}(t)}M$ comes from a variation.

Proposition 7.3.1. *For any curve* $\mathrm{q}\,(t)$*,* $t \in [a,b]$ *and tangent field* $V\,(t) \in T_{\mathrm{q}(t)}M$*, there is a variation whose variational field is* $V\,(t)$*.*

Proof. For each $V\,(t)$ let $s \mapsto \mathrm{q}\,(s,t)$ be the unique geodesic with $\mathrm{q}\,(0,t) = \mathrm{q}\,(t)$ and $\frac{\partial\,\mathrm{q}}{\partial s}\,(0,t) = V\,(t)$. The fact that $[a,b]$ is compact shows that we can find $\epsilon > 0$ so that $\mathrm{q}\,(s,t)$ is defined on $(-\epsilon,\epsilon)\times[a,b]$.

The fact that the geodesics depend smoothly on the initial values shows that the variation is a smooth as $\mathrm{q}\,(t)$ and $V\,(t)$. In particular, if $\mathrm{q}\,(t)$ is only piecewise smooth, then the variation will also consist of piecewise smooth curves that break at exactly the same points. □

Definition 7.3.2. The *length of a curve* is defined as

$$L\,(\mathrm{q}) = \int_a^b |\dot{\mathrm{q}}|\,dt$$

and the (kinetic) *energy* as

$$E\,(\mathrm{q}) = \frac{1}{2}\int_a^b |\dot{\mathrm{q}}|^2\,dt.$$

We know that the length of a curve does not change if we parametrize it. This is very far from true for the energy. You might even have noticed this yourself in terms of gas consumption when driving. Stop and go city driving consumes far more gas, than the more steady driving on an empty stretch of road on the country side. On the other hand this feature of the energy has the advantage that minima or stationary points for the energy functional come with a fixed parametrization.

Lemma 7.3.3 (*First Variation Formula*). *Consider a smooth variation* $\mathrm{q}\,(s,t)$*,* $(s,t) \in (-\epsilon,\epsilon)\times[0,1]$*, with base curve* $\mathrm{q}\,(t) = \mathrm{q}\,(0,t)$*, then*

$$\frac{d}{ds}\frac{1}{2}\int_0^1 \mathrm{I}\,(\dot{\mathrm{q}},\dot{\mathrm{q}})\,dt = \mathrm{I}\,(\partial_s\,\mathrm{q},\partial_t\,\mathrm{q})|_0^1 - \int_0^1 \mathrm{I}\left(\partial_s\,\mathrm{q},\ddot{\mathrm{q}}^{\mathrm{I}}\right)dt.$$

If $0 = a_0 < a_1 < \cdots < a_n = 1$ *and the variation is smooth when restricted to* $(-\epsilon,\epsilon)\times[a_{i-1},a_i]$, then

$$\frac{d}{ds}\frac{1}{2}\int_0^1 \mathrm{I}\,(\dot{\mathrm{q}},\dot{\mathrm{q}})\,dt = \sum_{i=1}^n \mathrm{I}\,(\partial_s\,\mathrm{q},\partial_t\,\mathrm{q})|_{a_{i-1}}^{a_i} - \int_0^1 \mathrm{I}\left(\partial_s\,\mathrm{q},\ddot{\mathrm{q}}^{\mathrm{I}}\right)dt.$$

PROOF. The calculation is straightforward in the smooth case:

$$\begin{aligned}
\frac{d}{ds}\frac{1}{2}\int_0^1 \mathrm{I}(\dot{\mathrm{q}},\dot{\mathrm{q}})\,dt &= \int_0^1 \mathrm{I}\left((\partial_s\partial_t\,\mathrm{q})^{\mathrm{I}},\partial_t\,\mathrm{q}\right)dt \\
&= \int_0^1 \left(\partial_t\,\mathrm{I}(\partial_s\,\mathrm{q},\partial_t\,\mathrm{q}) - \mathrm{I}\left(\partial_s\,\mathrm{q},\left(\partial_t^2\,\mathrm{q}\right)^{\mathrm{I}}\right)\right)dt \\
&= \mathrm{I}(\partial_s\,\mathrm{q},\partial_t\,\mathrm{q})|_0^1 - \int_0^1 \mathrm{I}\left(\partial_s\,\mathrm{q},\left(\partial_t^2\,\mathrm{q}\right)^{\mathrm{I}}\right)dt \\
&= \mathrm{I}(\partial_s\,\mathrm{q},\partial_t\,\mathrm{q})|_0^1 - \int_0^1 \mathrm{I}\left(\partial_s\,\mathrm{q},\ddot{\mathrm{q}}^{\mathrm{I}}\right)dt.
\end{aligned}$$

When the variation is only piecewise smooth, then we can break it up into smooth parts and add the contributions. □

We define $\Omega_{p,q}$ as the space of piecewise smooth curves between points $p,q\in M$ parametrized on $[0,1]$.

THEOREM 7.3.4. *If a piecewise curve on a surface is stationary for the energy functional on $\Omega_{p,q}$, then it is a geodesic.*

PROOF. We consider a piecewise smooth variation $\mathrm{q}(s,t)$ where the base curve $\mathrm{q}(t)=\mathrm{q}(0,t)$ corresponds to $s=0$. For simplicity assume that there is only one break point at a. Computing the energy of the curves $t\to\mathrm{q}(s,t)$ gives a function of s. The derivative with respect to s can be calculated as

$$\frac{d}{ds}\frac{1}{2}\int_0^1 \mathrm{I}(\dot{\mathrm{q}},\dot{\mathrm{q}})\,dt = \mathrm{I}(\partial_s\,\mathrm{q},\partial_t\,\mathrm{q})|_0^a + \mathrm{I}(\partial_s\,\mathrm{q},\partial_t\,\mathrm{q})|_a^1 - \int_0^1 \mathrm{I}\left(\partial_s\,\mathrm{q},\ddot{\mathrm{q}}^{\mathrm{I}}\right)dt.$$

When all the curves lie in $\Omega_{p,q}$ they have the same end points at $t=0,1$, i.e., $\mathrm{q}(s,0)=p$ and $\mathrm{q}(s,1)=q$ for all s. Such a variation is also called a *proper variation*. Thus, $\frac{\partial\,\mathrm{q}}{\partial s}(0,t)=0$ at $t=0,1$ and the formula simplifies to

$$\frac{d}{ds}\frac{1}{2}\int_0^1 \mathrm{I}(\dot{\mathrm{q}},\dot{\mathrm{q}})\,dt = \mathrm{I}\left(\partial_s\,\mathrm{q}(a),\frac{\partial\,\mathrm{q}}{\partial t^-}(a)-\frac{\partial\,\mathrm{q}}{\partial t^+}(a)\right) - \int_0^1 \mathrm{I}\left(\partial_s\,\mathrm{q},\ddot{\mathrm{q}}^{\mathrm{I}}\right)dt.$$

By assumption $s=0$ is a stationary point for $\frac{1}{2}\int_0^1 \mathrm{I}(\dot{\mathrm{q}},\dot{\mathrm{q}})\,dt$ so

$$0 = \mathrm{I}\left(\partial_s\,\mathrm{q}(a),\frac{\partial\,\mathrm{q}}{\partial t^-}(a)-\frac{\partial\,\mathrm{q}}{\partial t^+}(a)\right) - \int_0^1 \mathrm{I}\left(\partial_s\,\mathrm{q},\ddot{\mathrm{q}}^{\mathrm{I}}\right)dt.$$

First select the variation so that $\partial_s\,\mathrm{q}(0,t)$ is proportional to the tangential acceleration $\ddot{\mathrm{q}}^{\mathrm{I}}$, i.e., $\partial_s\,\mathrm{q}(0,t)=\mu(t)\,\ddot{\mathrm{q}}^{\mathrm{I}}$, where $\mu(a)=0$. Then we obtain

$$0 = -\int_0^1 \mu(t)\left|\ddot{\mathrm{q}}^{\mathrm{I}}\right|^2 dt.$$

Since μ can be chosen to be positive on $(0,a)\cup(a,1)$ this shows that $\ddot{\mathrm{q}}^{\mathrm{I}}=0$ on $(0,a)\cup(a,1)$. This shows that each of the two parts of $\mathrm{q}(t)$ on $[0,a]$ and $[a,1]$ are geodesics.

Next select a variation where

$$\partial_s\,\mathrm{q}(0,a) = \frac{\partial\,\mathrm{q}}{\partial t^-}(a) - \frac{\partial\,\mathrm{q}}{\partial t^+}(a).$$

In this case

$$0 = \mathrm{I}\left(\partial_s \mathrm{q}(a), \frac{\partial \mathrm{q}}{\partial t^-}(a) - \frac{\partial \mathrm{q}}{\partial t^+}(a)\right) = \mathrm{I}\left(\frac{\partial \mathrm{q}}{\partial t^-}(a) - \frac{\partial \mathrm{q}}{\partial t^+}(a), \frac{\partial \mathrm{q}}{\partial t^-}(a) - \frac{\partial \mathrm{q}}{\partial t^+}(a)\right)$$

so it follows that

$$\frac{\partial \mathrm{q}}{\partial t^-}(a) = \frac{\partial \mathrm{q}}{\partial t^+}(a).$$

Uniqueness of geodesics, then shows that the two parts of $\mathrm{q}(t)$ fit together to form a smooth geodesic on $[0,1]$.

Finally any curve of minimal energy is necessarily stationary since the derivative always vanishes at a minimum for a function. □

Now that we have identified the minima for the energy functional we show that they are also minima for the length functional.

LEMMA 7.3.5. *A minimizing curve for the energy functional is also a minimizing curve for the length functional.*

PROOF. We start by observing that the Cauchy-Schwarz inequality for the inner product of functions defined by

$$(f,g) = \int_a^b f(t)\, g(t)\, dt$$

implies that:

$$\int_a^b |\dot{\mathrm{q}}|\, dt \leq \sqrt{\int_a^b 1^2 dt}\sqrt{\int_a^b |\dot{\mathrm{q}}|^2\, dt} = \sqrt{b-a}\sqrt{\int_a^b |\dot{\mathrm{q}}|^2\, dt},$$

where equality occurs if $|\dot{\mathrm{q}}|$ is constant multiple of 1, i.e., q has constant speed. When the right hand side is minimized we just saw that q has zero acceleration and consequently constant speed. Let $\mathrm{q}_{\min}$ be a minimum for the energy in $\Omega_{p,q}$ and q any other curve in $\Omega_{p,q}$. We further assume that q has constant speed as reparametrizing the curve won't change its length. We now have

$$\begin{aligned}
\int_0^1 |\dot{\mathrm{q}}_{\min}|\, dt &\leq \sqrt{\int_0^1 |\dot{\mathrm{q}}_{\min}|^2\, dt} \\
&\leq \sqrt{\int_0^1 |\dot{\mathrm{q}}|^2\, dt} \\
&= \int_0^1 |\dot{\mathrm{q}}|\, dt
\end{aligned}$$

which shows the claim. □

COROLLARY 7.3.6. *If a piecewise smooth curve has constant speed and is a minimizer for the length functional, then it is a minimum for the energy and a geodesic.*

PROOF. If $\mathrm{q}_{\min}$ is a constant speed minimum for the length functional and $\mathrm{q} \in \Omega_{p,q}$, then

$$\begin{aligned} \int_0^1 \left|\dot{\mathrm{q}}_{\min}\right|^2 dt &= \left(\int_0^1 \left|\dot{\mathrm{q}}_{\min}\right| dt\right)^2 \\ &\leq \left(\int_0^1 \left|\dot{\mathrm{q}}\right| dt\right)^2 \\ &\leq \int_0^1 \left|\dot{\mathrm{q}}\right|^2 dt. \end{aligned}$$

This shows that $\mathrm{q}_{\min}$ also minimizes the energy functional and by theorem 7.3.4 that it must be a geodesic. □

REMARK 7.3.7. Note that minima for the length functional are not forced to be geodesics unless they are assumed to have constant speed!

Exercises

(1) Consider the curves $\mathrm{q}(t) = \left(at\cos(\theta), bt\sin(\theta), t^2\right)$ on $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$, $a, b > 0$.
   (a) Show that this is a geodesic only when $\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$.
   (b) Assume $a < b$ and $t \in [0, d]$ show that
   $$\int_0^d \sqrt{a^2 + 4t^2} dt \leq L(\mathrm{q}) \leq \int_0^d \sqrt{b^2 + 4t^2} dt$$
   with the lower bound holding for $\theta = 0, \pi$ and the upper bound for $\theta = \frac{\pi}{2}, \frac{3\pi}{2}$.

(2) Consider the curves $\mathrm{q}(t) = (a\cos(t)\cos(\theta), b\cos(t)\sin(\theta), c\sin(t))$ on $1 = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$, $a, b, c > 0$.
   (a) Show that this is a geodesic only when $\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$.
   (b) Assume $a < b < c$ and $t \in [0, d]$ show that
   $$\int_0^d \sqrt{a^2 + (c^2 - a^2)\cos^2 t} dt \leq L(\mathrm{q}) \leq \int_0^d \sqrt{b^2 + (c^2 - b^2)\cos^2 t} dt$$
   with the lower bound holding for $\theta = 0, \pi$ and the upper bound for $\theta = \frac{\pi}{2}, \frac{3\pi}{2}$.

7.4. Short Geodesics

We start by introducing geodesic coordinates along a curve. We then proceed to do the same construction around a point. This construction is similar but complicated by the fact that our base curve is a fixed point. In Euclidean space this corresponds to the singularity at the origin when switching from Cartesian to polar coordinates.

PROPOSITION 7.4.1. *Every surface admits geodesic coordinates around every point.*

PROOF. Start by choosing a unit speed curve $\mathrm{q}(v)$, $v \in [a, b]$ such that the specified point $q = \mathrm{q}(v_0)$ for some $v_0 \in (a, b)$. Next select a consistent choice of unit normal vector $\mathrm{S}(v)$ to this curve inside the surface as a variational field. Then let $u \mapsto \mathrm{q}(u, v)$ be the unique unit speed geodesic with $\mathrm{q}(0, v) = \mathrm{q}(v)$ and $\partial_u \mathrm{q}(0, v) = \mathrm{S}(v)$ to obtain a variation on $(-\epsilon, \epsilon) \times [a, b]$.

Since $u \mapsto \mathrm{q}(u, v)$ is unit speed we have $\mathrm{I}(\partial_u \mathrm{q}, \partial_u \mathrm{q}) = 1$. Next consider the inner product $\mathrm{I}(\partial_u \mathrm{q}, \partial_v \mathrm{q})$. Since $\partial_u \mathrm{q}(0, v) = \mathrm{S}(v)$ is perpendicular to $\partial_v \mathrm{q}(0, v) = \partial_v \mathrm{q}(v)$ this inner product vanishes for all parameters $(0, v)$. If we differentiate the inner product with respect to u and use the product rule twice we obtain

$$\begin{aligned}
\partial_u \mathrm{I}(\partial_u \mathrm{q}, \partial_v \mathrm{q}) &= \mathrm{I}\left(\left(\partial_u^2 \mathrm{q}\right)^{\mathrm{I}}, \partial_v \mathrm{q}\right) + \mathrm{I}\left(\partial_u \mathrm{q}, \left(\partial_u \partial_v \mathrm{q}\right)^{\mathrm{I}}\right) \\
&= \mathrm{I}\left(\partial_u \mathrm{q}, \left(\partial_u \partial_v \mathrm{q}\right)^{\mathrm{I}}\right) \\
&= \mathrm{I}\left(\partial_u \mathrm{q}, \left(\partial_v \partial_u \mathrm{q}\right)^{\mathrm{I}}\right) \\
&= \frac{1}{2}\left(\mathrm{I}\left(\partial_u \mathrm{q}, \left(\partial_v \partial_u \mathrm{q}\right)^{\mathrm{I}}\right) + \mathrm{I}\left(\left(\partial_v \partial_u \mathrm{q}\right)^{\mathrm{I}}, \partial_u \mathrm{q}\right)\right) \\
&= \frac{1}{2}\left(\partial_v \mathrm{I}(\partial_u \mathrm{q}, \partial_u \mathrm{q})\right) \\
&= \frac{1}{2}\left(\partial_v 1\right) \\
&= 0.
\end{aligned}$$

This shows that $\mathrm{I}(\partial_u \mathrm{q}, \partial_v \mathrm{q})$ is constant along u-curves and vanishes at $u = 0$. Thus it vanishes everywhere.

Finally define $g_{vv} = \mathrm{I}(\partial_v \mathrm{q}, \partial_v \mathrm{q})$. It now just remains to note that $g_{vv}(0, v) = 1$ and $g_{vv}(u, v)$ is continuous. Thus we can, after possibly decreasing ϵ, assume that $g_{vv} > 0$ on all of the region $(-\epsilon, \epsilon) \times [a, b]$. This shows that the velocity fields $\partial_u \mathrm{q}$ and $\partial_v \mathrm{q}$ never vanish and are always orthogonal. Thus they give the desired parametrization. We can then further restrict the domain around $(0, v_0)$ if we wish to obtain a coordinate system where the parametrization is a local diffeomorphism. □

We now fix a point $p \in M$. For a tangent vector $X \in T_pM$, let q_X be the unique geodesic with $\mathrm{q}(0) = p$ and $\dot{\mathrm{q}}(0) = X$, and $[0, b_X)$ the non-negative part of the maximal interval on which q is defined. Notice that uniqueness of geodesics implies the *homogeneity property*: $\mathrm{q}_{\alpha X}(t) = \mathrm{q}_X(\alpha t)$ for all $\alpha > 0$ and $t < b_{\alpha X}$. In particular, $b_{\alpha X} = \alpha^{-1} b_X$. Let $O_p \subset T_pM$ be the set of vectors X such that $1 < b_X$. In other words $\mathrm{q}_X(t)$ is defined on $[0, 1]$.

Definition 7.4.2. The *exponential map* at p, $\exp_p : O_p \to M$, is defined by

$$\exp_p(X) = \mathrm{q}_X(1).$$

The homogeneity property $\mathrm{q}_X(t) = \mathrm{q}_{tX}(1)$ shows that $\exp_p(tX) = \mathrm{q}_X(t)$. Therefore, it is natural to think of $\exp_p(X)$ in a polar coordinate representation, where from p one goes "distance" $|X|$ in the direction of $\frac{X}{|X|}$. This gives the point $\exp_p(X)$, since $\mathrm{q}_{\frac{X}{|X|}}(|X|) = \mathrm{q}_X(1)$.

It is an important property that $\exp_p$ is in fact a local diffeomorphism around $0 \in T_pM$.

Proposition 7.4.3. *For each $p \in M$ there exists $\epsilon > 0$ so that $B(0, \epsilon) \subset O_p \subset T_pM$ and the differential $D\exp_p$ is nonsingular at the origin. Consequently, $\exp_p$ is a local diffeomorphism.*

Proof. By theorem A.5.1 there exists $\epsilon > 0$ such that $\mathrm{q}_X(t)$ is defined on $[0, 2\epsilon)$ for all unit vectors $X \in T_pM$. The homogeneity property shows that $B(0, \epsilon) \subset O_p$.

That the differential is non-singular also follows from the homogeneity property of geodesics. For a fixed vector $X \in T_pM$ we just saw that

$$\exp_p(tX) = \mathrm{q}_X(t)$$

and thus

$$\begin{aligned} \left(D\exp_p\right)(X) &= \frac{d}{dt}|_{t=0}\exp_p(tX) \\ &= \dot{\mathrm{q}}_X(0) \\ &= X. \end{aligned}$$

This shows that the differential is the identity map and in particular non-singular. The second statement follows from the inverse function theorem. □

We can now introduce Gauss's version of geodesic polar coordinates.

Lemma 7.4.4 (Gauss Lemma). *Around any point $p \in M$ it is possible to introduce polar geodesic coordinate parameters $\mathrm{q}(r,\theta)$ where the r-parameter curves are the unit speed geodesics emanating from p and*

$$[\mathrm{I}] = \begin{bmatrix} 1 & 0 \\ 0 & g_{\theta\theta} \end{bmatrix}.$$

Proof. Pick $\epsilon > 0$ such that $\exp_p : B(0,\varepsilon) \to B = \exp_p(B(0,\varepsilon))$ is a diffeomorphism. Then $r(q) = |\exp_p^{-1}(q)|$ is well-defined for all $q \in B$. Note that r is simply the Euclidean distance function from the origin on $B(0,\varepsilon) \subset T_pM$ in exponential coordinates. This function can be continuously extended to $\bar{B}$ by defining $r(\partial B) = \varepsilon$. Select an orthonormal basis E_1, E_2 for T_pM and introduce Cartesian coordinates (x,y) on T_pM. These parameters are then also used on B via the exponential map $\mathrm{q}(x,y) = \exp_p(xE_1 + yE_2)$. We define the polar coordinates by

$$x = r\cos\theta,\ y = r\sin\theta$$

and note that

$$\begin{aligned} r &= \sqrt{x^2+y^2}, \\ \partial_r \mathrm{q} &= \frac{x}{r}\partial_x \mathrm{q} + \frac{y}{r}\partial_y \mathrm{q}, \\ \partial_\theta \mathrm{q} &= -y\partial_x \mathrm{q} + x\partial_y \mathrm{q}. \end{aligned}$$

Observe that $\partial_r \mathrm{q}$ is not defined at p, while $\partial_\theta \mathrm{q}$ is defined on all of B even though the angle θ is not. We now need to check what the first fundamental form looks like in polar coordinates. First note that the r-parameter curves by definition have velocity $\partial_r \mathrm{q}$. On the other hand via the exponential map they correspond to unit speed radial lines rX, where $|X| = 1$. This means that they are of the form $\exp_p(rX) = \mathrm{q}_X(r)$ and are unit speed geodesics. This shows that $g_{rr} = \mathrm{I}(\partial_r \mathrm{q}, \partial_r \mathrm{q}) = 1$. To show that $g_{r\theta} = 0$ we first calculate its derivative

$$\begin{aligned} \partial_r \mathrm{I}(\partial_r \mathrm{q}, \partial_\theta \mathrm{q}) &= \mathrm{I}\left(\left(\partial_r^2 \mathrm{q}\right)^{\mathrm{I}}, \partial_\theta \mathrm{q}\right) + \mathrm{I}\left(\partial_r \mathrm{q}, \left(\partial_r\partial_\theta \mathrm{q}\right)^{\mathrm{I}}\right) \\ &= 0 + \mathrm{I}\left(\partial_r \mathrm{q}, \left(\partial_\theta\partial_r \mathrm{q}\right)^{\mathrm{I}}\right) \\ &= \frac{1}{2}\partial_\theta \mathrm{I}(\partial_r \mathrm{q}, \partial_r \mathrm{q}) \\ &= 0. \end{aligned}$$

Thus $\mathrm{I}(\partial_r \mathrm{q}, \partial_\theta \mathrm{q})$ is constant along geodesics emanating from p. To show that it vanishes it is tempting to simply evaluate at p since $\partial_\theta \mathrm{q}$ vanishes there. However, $\partial_r \mathrm{q}$ is undefined so we use a limit argument. First observe that

$$\begin{aligned}
|\mathrm{I}(\partial_r \mathrm{q}, \partial_\theta \mathrm{q})| &\leq |\partial_r \mathrm{q}|\,|\partial_\theta \mathrm{q}| \\
&= |\partial_\theta \mathrm{q}| \\
&\leq |x|\,|\partial_y \mathrm{q}| + |y|\,|\partial_x \mathrm{q}| \\
&\leq r\left(|\partial_x \mathrm{q}| + |\partial_y \mathrm{q}|\right).
\end{aligned}$$

Continuity of $D\exp_p$ shows that $\partial_x \mathrm{q}, \partial_y \mathrm{q}$ are bounded near p. Thus $\mathrm{I}(\partial_r \mathrm{q}, \partial_\theta \mathrm{q}) \to 0$ as $r \to 0$. This forces $\mathrm{I}(\partial_r \mathrm{q}, \partial_\theta \mathrm{q}) = 0$.

Finally we can just define $g_{\theta\theta} = \mathrm{I}(\partial_\theta \mathrm{q}, \partial_\theta \mathrm{q})$ and note that it is positive as $\partial_\theta \mathrm{q}$ only vanishes at p. □

THEOREM 7.4.5. *Let M be a surface, $p \in M$, and $\varepsilon > 0$ chosen such that*

$$\exp_p : B(0, \varepsilon) \to B \subset M$$

is a diffeomorphism onto its image $B \subset M$. It follows that the geodesic $\mathrm{q}_X(t) = \exp_p(tX)$, $t \in [0,1]$ is the one and only minimal geodesic in M from p to $q = \exp_p X$.

PROOF. The proof is analogous to the specific situation on the round sphere covered in example 1.2.11.

To see that $\mathrm{q}_X(t)$ is the one and only shortest curve in M, we must show that any other curve from p to q has length $> |X|$. Suppose we have a curve $\mathrm{q} : [0,b] \to M$ from p to q. If $a \in [0,b]$ is the largest value so that $\mathrm{q}(a) = p$, then $\mathrm{q}|_{[a,b]}$ is a shorter curve from p to q. Next let $b_0 \in (a,b)$ be the first value for which $\mathrm{q}(t_0) \notin B$ if such points exist, otherwise $b_0 = b$. The curve $\mathrm{q}|_{(a,b_0)}$ now lies entirely in $B - \{p\}$ and is shorter than the original curve. Its length is easily estimated from below

$$\begin{aligned}
L\left(\mathrm{q}|_{(a,b_0)}\right) &= \int_a^{b_0} |\dot{\mathrm{q}}|\,dt \\
&= \int_a^{b_0} |\partial_r \mathrm{q}| \cdot |\dot{\mathrm{q}}|\,dt \\
&\geq \int_a^{b_0} \mathrm{I}(\partial_r \mathrm{q}, \dot{\mathrm{q}})\,dt \\
&= \int_a^{b_0} \mathrm{I}\left(\partial_r \mathrm{q}, \frac{dr(\mathrm{q}(t))}{dt}\partial_r \mathrm{q} + \frac{d\theta(\mathrm{q}(t))}{dt}\partial_\theta \mathrm{q}\right) dt \\
&= \int_a^{b_0} \frac{dr(\mathrm{q}(t))}{dt}\,dt \\
&= r(\mathrm{q}(b_0)) - r(\mathrm{q}(a)) \\
&= r(\mathrm{q}(b_0)),
\end{aligned}$$

where we used that $r(p) = 0$. If $\mathrm{q}(b_0) \in \partial B$, then q is not a segment from p to q as it has length $\geq \varepsilon > |X|$. If $b = b_0$, then $L\left(\mathrm{q}|_{(a,b)}\right) \geq r(\mathrm{q}(b)) = |X|$ and equality can only hold if $\dot{\mathrm{q}}(t)$ is proportional to $\partial_r \mathrm{q}$ for all $t \in (a,b]$. This shows the short geodesic is a minimal geodesic and that any other curve of the same length must be a reparametrization of this short geodesic. □

7.5. Distance and Completeness

Definition 7.5.1. The distance between two points in a surface M is defined by attempting to minimize the length of curves between the points:

$$|pq| = \inf\{L(\mathrm{q}) \mid \mathrm{q} \in \Omega_{pq}\}.$$

This distance satisfies the usual properties of a distance:

(1) $|pq| > 0$ unless $p = q$,
(2) $|pq| = |qp|$,
(3) $|pq| \leq |px| + |xq|$.

2 and 3 are also immediate from the definition. It is also clear that $|pq| \geq 0$. Finally, if $|pq| = 0$, then $q \in B = \exp_p(B(0,\epsilon))$ as in theorem 7.4.5. In this case $|pq|$ is a minimum realized by the short geodesic in B joining p and q. Thus $p = q$.

Definition 7.5.2. We define the open ball, closed ball and distance sphere around a point $p \in M$ as:

$$\begin{aligned} B(p,r) &= \{x \in M \mid |px| < r\}, \\ \bar{B}(p,r) &= \{x \in M \mid |px| \leq r\}, \\ S(p,r) &= \{x \in M \mid |px| = r\}. \end{aligned}$$

The next corollary is almost an immediate consequence of theorem 7.4.5 and its proof now that we have introduced the concept of distance.

Corollary 7.5.3. *If $p \in M$ and $\varepsilon > 0$ is such that $\exp_p : B(0,\varepsilon) \to B$ is defined and a diffeomorphism, then for each $\delta \leq \varepsilon$,*

$$\exp_p(B(0,\delta)) = B(p,\delta),$$

and for each $\delta < \epsilon$

$$\exp_p(\bar{B}(0,\delta)) = \bar{B}(p,\delta).$$

In particular, it follows that $p_i \to p$ if and only if $|pp_i| \to 0$.

Proof. We first have to show that $B(p,\varepsilon) = B$. We already have $B \subset B(p,\varepsilon)$. Conversely if $q \in B(p,\varepsilon)$, then it is joined to p by a curve $\mathrm{q}(t) \in \Omega_{pq}$ of length $< \varepsilon$. The proof of theorem 7.4.5 now shows that any curve starting at p that leaves B has length $\geq \epsilon$. This means that $\mathrm{q}(t)$ lies in O and $q \in O$. This argument can now be repeated for each $\delta < \epsilon$. This in turn also shows that $\exp_p(\bar{B}(0,\delta)) = \bar{B}(p,\delta)$ when $\delta < \epsilon$.

Finally, note that by our definition of convergence any sequence p_i that converges to p eventually must lie within the exponential parametrization of $B(p,\delta)$. The same clearly also holds if $|pp_i| \to 0$. Since this is true for all $\delta > 0$ the claim follows. □

We are now ready to connect the concept of geodesic completeness with the existence of shortest curves on a larger scale.

Theorem 7.5.4. *(Hopf-Rinow, 1931) If a surface M is geodesically complete at p, then any point $q \in M$ is joined to p by a minimal geodesic of length $|pq|$.*

Proof. Consider p, q and choose $\epsilon > 0$ such that any point in $\bar{B}(p,\epsilon)$ can be joined to p by a unique minimal geodesic (see corollary 7.5.3). This shows that $\bar{B}(p,\epsilon)$ is homeomorphic to a disc with boundary $S(p,\epsilon)$. In particular $S(p,\epsilon)$ is compact. This shows that there exists a $q_0 \in S(p,\epsilon)$ closest to q. For this

q_0 we claim that $|pq_0| + |q_0q| = |pq|$. Otherwise there would be a unit speed curve $\gamma \in \Omega_{p,q}$ with $L(\gamma) < |pq_0| + |q_0q|$. Choose t so that $\gamma(t) \in S(p,\epsilon)$. Since $t + |\gamma(t)\, q| \leq L(\gamma) < |pq_0| + |q_0q|$ it follows that $|\gamma(t)\, q| < |q_0q|$ contradicting the choice of q_0. Now let $\mathrm{q}(t)$ be the unit speed geodesic with $\mathrm{q}(0) = p$, $\mathrm{q}(\epsilon) = q_0$, and

$$A = \{t \in [0, |pq|] \mid |pq| = t + |\mathrm{q}(t)\, q|\}.$$

Clearly $0 \in A$. Also $\epsilon \in A$ since $\mathrm{q}(\epsilon) = q_0$. Note that if $t \in A$, then

$$|pq| = t + |\mathrm{q}(t)\, q| \geq |p\, \mathrm{q}(t)| + |\mathrm{q}(t)\, q| \geq |pq|,$$

which implies that $t = |p\, \mathrm{q}(t)|$. We first claim that if $t_0 \in A$, then $[0, t_0] \subset A$. Let $t < t_0$ and note that

$$\begin{aligned}
|pq| &\leq |p\, \mathrm{q}(t)| + |\mathrm{q}(t)\, q| \\
&\leq |p\, \mathrm{q}(t)| + |\mathrm{q}(t)\, \mathrm{q}(t_0)| + |\mathrm{q}(t_0)\, q| \\
&\leq t + t_0 - t + |\mathrm{q}(t_0)\, q| \\
&\leq t_0 + |\mathrm{q}(t_0)\, q| \\
&= |pq|.
\end{aligned}$$

This implies that $|p\, \mathrm{q}(t)| + |\mathrm{q}(t)\, q| = |pq|$ and $t = |p\, \mathrm{q}(t)|$, showing together that $t \in A$.

Since $t \mapsto |\mathrm{q}(t)\, q|$ is continuous it follows that A is closed.

Finally if $t_0 \in A$, then $t_0 + \delta \in A$ for sufficiently small $\delta > 0$. Select $\delta > 0$ so that any point in $\bar{B}(\mathrm{q}(t_0), \delta)$ can be joined to $\mathrm{q}(t_0)$ by a minimal geodesic. Then select $q_1 \in S(\mathrm{q}(t_0), \delta)$ closest to q. We now have

$$\begin{aligned}
|pq| &= t_0 + |\mathrm{q}(t_0)\, q| \\
&= t_0 + |\mathrm{q}(t_0)\, q_1| + |q_1q| \\
&= t_0 + \delta + |q_1q| \\
&\geq |pq_1| + |q_1q| \\
&\geq |pq|.
\end{aligned}$$

It follows that $|pq_1| = t_0 + \delta$ from which we conclude that the piecewise smooth geodesic that goes from p to $\mathrm{q}(t_0)$ and then from $\mathrm{q}(t_0)$ to q_1 has length $|pq_1|$. Consequently it is a smooth geodesic and $q_1 = \mathrm{q}(t_0 + \delta)$. It then follows from $|pq| = t_0 + \delta + |q_1q|$ that $\mathrm{q}(t_0 + \delta) \in A$. □

This in turns shows that several different completeness criteria are all equivalent.

THEOREM 7.5.5. *(Hopf-Rinow, 1931) The following statements are equivalent for a surface M:*

(1) *M is geodesically complete, i.e., all geodesics are defined for all time.*
(2) *M is geodesically complete at p, i.e., all geodesics through p are defined for all time.*
(3) *M satisfies the Heine-Borel property, i.e., every closed bounded set is compact.*
(4) *M is metrically complete.*

PROOF. (1)⇒(2) is trivial. (3)⇒(4) follows from the fact that Cauchy sequences are bounded.

For (4)⇒(1): If we have a unit speed geodesic $\mathrm{q} : [0, b) \to M$, then $|\mathrm{q}(t)\,\mathrm{q}(s)| \leq |t - s|$. So if $b < \infty$, it follows that $|\mathrm{q}(t)\,\mathrm{q}(s)| \to 0$ as $t, s \to b$. This shows that $\mathrm{q}(t)$ is a Cauchy sequence as $t \to b$ and by (4) must converge to a point p. In particular, $\mathrm{q}(t)$ lies in a compact set $\bar{B}(p, \delta)$ as $t \to b$. The derivative is also bounded, so it follows from theorem A.5.1 that starting at any time t_0 where $\mathrm{q}(t_0) \in \bar{B}(p, \delta)$ the geodesic exists on an interval $(-\epsilon + t_0, t_0 + \epsilon)$ where ϵ is independent of t_0. When $t_0 + \epsilon > b$ we'll have found an extension of the geodesic. This shows that the any geodesic must be defined on $[0, \infty)$.

Finally the traditionally difficult part (2)⇒(3) is an easy consequence of theorem 7.5.4. We show that $\exp_p\left(\bar{B}(0, r)\right) = \bar{B}(p, r)$ for all $r > 0$. It is clear that any point in $\exp_p\left(\bar{B}(0, r)\right)$ is joined to p by a geodesic of length $\leq r$. Thus $\exp_p\left(\bar{B}(0, r)\right) \subset \bar{B}(p, r)$. Conversely we just proved in theorem 7.5.4 that any point in $\bar{B}(p, r)$ is joined to p by a geodesic of length $\leq r$. But any such geodesic is of the form $\mathrm{q}_X(t)$ with $\mathrm{q}_X(0) = p$, $t \in [0, 1]$, and $|X| \leq r$. This shows that $\mathrm{q}_X(1) \in \exp_p\left(\bar{B}(0, r)\right)$. We now have that all of the closed balls $\bar{B}(p, r)$ are compact as they are the image of a closed ball in $\mathbb{R}^2$. Since any bounded subset of M lies in such a ball $\bar{B}(p, r)$ the Heine-Borel property follows. □

Exercises

(1) Show that hyperbolic space H (see 6.1.2) is complete.
(2) Show that generalized cones and tangent developables are never complete.
(3) Consider a generalized cylinder $\mathrm{q}(s, t) = c(t) + sX$, $t \in I$, where c is parametrized by arclength.
   (a) Show that the surface is complete if c is closed.
   (b) Show that the surface is complete if $I = \mathbb{R}$.
   (c) Show the surface is not complete if $I \neq \mathbb{R}$ and c is not closed.
(4) Give an example of an abstract surface (i.e., first fundamental form) defined on all of $\mathbb{R}^2$ that is not complete.

7.6. Isometries

So far we've mostly discussed how quantities remain invariant if we change parameters at a given point. Here we shall exploit more systematically what isometries can do to help us find and calculate geometric invariants. Recall that an isometry is simply a map that preserves the first fundamental forms. Thus isometries preserve all intrinsic notions. Isometries are also often referred to as symmetries, especially when they are maps from a surface to it self.

Corollary 7.6.1. *An isometry maps geodesics to geodesics, preserves Gauss curvature, and preserves the length of curves.*

Proof. Let $\mathrm{q}(t)$ be a geodesic and F an isometry. The geodesic equation depends only on the first fundamental form. By definition isometries preserve the first fundamental form, thus $F(\mathrm{q}(t))$ must also be a geodesic.

Next assume that F is an isometry such that $F(p) = q$. Again F preserves the first fundamental form so the Gauss curvatures must again be the same.

Finally when $\mathrm{q}(t)$ is a curve we have

$$\begin{aligned} L(F\circ \mathrm{q}) &= \int_a^b \left|\frac{d}{dt}\left(F\left(\mathrm{q}(t)\right)\right)\right| dt \\ &= \int_a^b \left|DF\left(\dot{\mathrm{q}}(t)\right)\right| dt \\ &= \int_a^b \left|\dot{\mathrm{q}}(t)\right| dt \\ &= L(\mathrm{q}). \end{aligned}$$

□

COROLLARY 7.6.2. *An isometry is distance decreasing. Moreover, if it is a bijection then it is distance preserving.*

PROOF. Since isometries preserve length of curves it is clear from the definition of distance that they are distance decreasing. In case F is also a bijection it follows that F^{-1} exists and is also an isometry. Thus both F and F^{-1} are distance decreasing. This shows that they are distance preserving. □

Basic examples of isometries are rotations around the z axis for surfaces of revolution around the z axis, or mirror symmetries in meridians on a surface of revolution. The sphere has an even larger number of isometries as it is a surface of revolution around any line through the origin. The plane also has rotational and mirror symmetries, but in addition translations.

It is possible to construct isometries that do not preserve the second fundamental form. The simplest example is to imagine a flat tarp or blanket, here all points have vanishing second fundamental form and also there are isometries between all points. Now lift one side of the tarp. Part of it will still be flat on the ground, while the part that's lifted off the ground is curved. The first fundamental form has not changed but the curved part will now have nonzero entries in the second fundamental form.

It is not always possible to directly determine all isometries. But as with geodesics there are some uniqueness results that will help.

THEOREM 7.6.3. *If F and G are isometries that satisfy $F(p)=G(p)$ and $DF(p)=DG(p)$, then $F=G$ in a neighborhood of p.*

PROOF. We just saw that isometries preserve geodesics. So if $\mathrm{q}(t)$ is a geodesic with $\mathrm{q}(0)=p$, then $F(\mathrm{q}(t))$ and $G(\mathrm{q}(t))$ are both geodesics. Moreover they have the same initial values

$$\begin{aligned} F(\mathrm{q}(0)) &= F(p), \\ G(\mathrm{q}(0)) &= G(p), \\ \frac{d}{dt}F(\mathrm{q}(t))\,|_{t=0} &= DF(\dot{\mathrm{q}}(0)), \\ \frac{d}{dt}G(\mathrm{q}(t))\,|_{t=0} &= DG(\dot{\mathrm{q}}(0)). \end{aligned}$$

This means that $F(\mathrm{q}(t))=G(\mathrm{q}(t))$. By varying the initial velocity of $\dot{\mathrm{q}}(0)$ we can reach all points in a neighborhood of p. □

Often the best method for finding isometries is to make educated guesses based on what the metric looks like. One general guideline for creating isometries is the observation that if the first fundamental form doesn't depend on a specific variable such as v, then translations in that variable will generate isometries. This is exemplified by surfaces of revolution where the metric doesn't depend on μ. Translations in μ are the same as rotations by a fixed angle and we know that such transformations are isometries. Note that reflections in such a parameter where v is mapped to $v_0 - v$ will also be isometries in such a case.

EXAMPLE 7.6.4. The linear orthogonal transformations $O(3)$ of $\mathbb{R}^3$ preserve the spheres centered at the origin. Moreover, with these transformations it is possible to solve all possible initial value problems as in theorem 7.6.3. To see this last statement we concentrate on the unit sphere. An orthonormal basis e_1, e_2 for T_pS^2 will give us an orthonormal basis e_1, e_2, p for $\mathbb{R}^3$. Let f_1, f_2, q be another orthonormal basis, i.e., f_1, f_2 is an orthonormal basis for T_qS^2. We then have two orthogonal matrices

$$\begin{bmatrix} f_1 & f_2 & q \end{bmatrix}, \begin{bmatrix} e_1 & e_2 & p \end{bmatrix} \in O(3).$$

We define $O \in O(3)$ by

$$O = \begin{bmatrix} f_1 & f_2 & q \end{bmatrix}\begin{bmatrix} e_1 & e_2 & p \end{bmatrix}^{-1}.$$

Thus

$$\begin{aligned}
\begin{bmatrix} O(e_1) & O(e_2) & O(p) \end{bmatrix} &= O\begin{bmatrix} e_1 & e_2 & p \end{bmatrix} \\
&= \begin{bmatrix} f_1 & f_2 & q \end{bmatrix}\begin{bmatrix} e_1 & e_2 & p \end{bmatrix}^{-1}\begin{bmatrix} e_1 & e_2 & p \end{bmatrix} \\
&= \begin{bmatrix} f_1 & f_2 & q \end{bmatrix}.
\end{aligned}$$

In other words $O(p) = q$, $O(e_1) = f_1$, and $O(e_2) = f_2$. This shows that we can solve all initial value problems.

EXAMPLE 7.6.5. The isometries of $\mathbb{R}^2$ are all of the form $F(x) = Ox + q$, where $O \in O(2)$ represents the differential $O = DF(0)$ and $q \in \mathbb{R}^2$ the initial point $q = F(0)$. Theorem 7.6.3 again shows that there are no more isometries.

EXAMPLE 7.6.6. The linear transformations that preserve the space-time inner product on $\mathbb{R}^{2,1}$ are denoted $O(2,1)$. They are characterized by being of the form $O = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix}$, where $e_i \cdot e_j = 0$ when $i \neq j$, $|e_1|^2 = |e_2|^2 = 1$, and $|e_3|^2 = -1$. Note that

$$O\begin{bmatrix} x \\ y \\ z \end{bmatrix} = xe_1 + ye_2 + ze_3$$

and that

$$|xe_1 + ye_2 + ze_3|^2 = x^2 + y^2 - z^2.$$

This means that these transformations preserve the two sheeted hyperboloid $x^2 + y^2 - z^2 = -1$. Any given O either preserves each of the two sheets or flips the two sheets. The first case happens when O preserves H and the set of these transformations is denoted $O^+(2,1)$. We can determine when $O \in O^+(2,1)$ by checking that the 33 entry in O is positive as that means that $(0,0,1)$ is mapped to a point in H. The key observation is that any orthonormal basis e_1, e_2 for T_pH will give us an element $\begin{bmatrix} e_1 & e_2 & p \end{bmatrix} \in O^+(2,1)$. Consequently, we can, as in the sphere case, create the desired transformation using

$$O = \begin{bmatrix} f_1 & f_2 & q \end{bmatrix} \begin{bmatrix} e_1 & e_2 & p \end{bmatrix}^{-1}.$$

Here is a slightly more surprising relationship between geodesics and isometries.

THEOREM 7.6.7. *Let F be a nontrivial isometry and $\mathrm{q}(t)$ a unit speed curve such that $F(\mathrm{q}(t)) = \mathrm{q}(t)$ for all t, then $\mathrm{q}(t)$ is a geodesic.*

PROOF. Since F is an isometry and it preserves q we must also have that it preserves its velocity and tangential acceleration

$$\begin{aligned} DF(\dot{\mathrm{q}}(t)) &= \dot{\mathrm{q}}(t), \\ DF\left(\ddot{\mathrm{q}}^{\mathrm{I}}(t)\right) &= \ddot{\mathrm{q}}^{\mathrm{I}}(t). \end{aligned}$$

As q is unit speed we have $\dot{\mathrm{q}} \cdot \ddot{\mathrm{q}}^{\mathrm{I}} = 0$. If $\ddot{\mathrm{q}}^{\mathrm{I}}(t) \neq 0$, then DF preserves $\mathrm{q}(t)$ as well as the basis $\dot{\mathrm{q}}(t), \ddot{\mathrm{q}}^{\mathrm{I}}(t)$ for the tangent space at $\mathrm{q}(t)$. By the uniqueness result above this shows that F is the identity map as that map is always an isometry that fixes any point and basis. But this contradicts that F is nontrivial. □

Note that circles in the plane are preserved by rotations, but they are not fixed, nor are they geodesics. The picture we should have in mind for such an isometry and geodesic is a mirror symmetry in a line, or a mirror symmetry in a great circle on the sphere.

Exercises

(1) Show that the set of bijective isometries (or symmetries) of a surface M form a group if they product structure is composition of isometries.
(2) Consider three distinct points $p_1, p_2, p_3 \in M$ with the property that each pair is joined by a unique geodesic segment. Show that if an isometry fixes all three points then it is the identity map.
(3) Consider the ellipsoid
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$
  (a) Show that the eight maps $(x, y, z) \mapsto (\pm x, \pm y, \pm z)$ are isometries.
  (b) Show that when $a > b > c > 0$ then these are the only isometries: Hint: The Gauss curvature is calculated in section 5.3 exercise 27. Use that isometries preserve both curvature as well as critical points for the curvature to show that the three sets of points where two coordinates vanish are preserved.
(4) Consider the parabolic surface $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$, $a, b > 0$.
  (a) Show that all isometries fix the origin. Hint: Calculate the Gauss curvature.
  (b) Show that the four maps $(x, y) \mapsto (\pm x, \pm y)$ are isometries.
  (c) Show that when $a > b$, then these are the only isometries. Hint: Use that isometries preserve both distance and curvature to show that an isometry must preserve the curves where either $x = 0$ or $y = 0$. Specifically, on the level set where $z = R$ the curvature and distance to the origin is maximal when $y = 0$. Moreover, when $z > R$ the curvature is strictly smaller than this value.

(5) Consider a unit speed curve $c(s) : [0, L] \to \mathbb{R}^3$ with non-vanishing curvature and the tube of radius R around it

$$\mathrm{q}(s, \phi) = c(s) + R(\mathrm{N}_c \cos\phi + \mathrm{B}_c \sin\phi)$$

(see section 4.3 exercise 7 and section 5.3 exercise 22).

(a) Show that the map $(s, \phi) \mapsto (s, -\phi)$ is an isometry.

(b) Assume that κ has a maximum at s_0. Show that any isometry must fix $(s_0, 0)$.

7.7. Constant Curvature

We've already seen many models of surfaces with constant curvature and in some cases we explicitly showed how they could be reparametrized to be isometric. This is no accident and can be done more abstractly. The goal will be to give a canonical local structure for surfaces with constant Gauss curvature. This will be done in the form of a canonical parametrization.

Theorem 7.7.1 (Gauss, 1827). *If an abstract surface has vanishing Gauss curvature, then it admits Cartesian coordinates.*

Proof. We use geodesic coordinates along a unit speed geodesic as in proposition 7.4.1. Thus $v \mapsto \mathrm{q}(0, v)$ is a unit speed geodesic and all of the u-curves are unit speed geodesics. The first fundamental form is

$$[\mathrm{I}] = \begin{bmatrix} 1 & 0 \\ 0 & g_{vv} \end{bmatrix}.$$

Assuming $K = 0$, the formula for the Gauss curvature

$$K = -\frac{\partial_u^2 \sqrt{g_{vv}}}{\sqrt{g_{vv}}}$$

from example 5.3.9 shows that

$$\sqrt{g_{vv}(u, v)} = \sqrt{g_{vv}(0, v)} + u \cdot \left(\partial_u \sqrt{g_{vv}}\right)(0, v).$$

We also have the initial condition:

$$\sqrt{g_{vv}(0, v)} = \left|\frac{\partial \mathrm{q}}{\partial v}\right| = 1.$$

Note that

$$\begin{aligned}
\partial_u g_{vv} &= 2\,\mathrm{I}\left((\partial_u \partial_v \mathrm{q})^{\mathrm{I}}, \partial_v \mathrm{q}\right) \\
&= 2\,\mathrm{I}\left((\partial_v \partial_u \mathrm{q})^{\mathrm{I}}, \partial_v \mathrm{q}\right) \\
&= 2\partial_v \mathrm{I}(\partial_u \mathrm{q}, \partial_v \mathrm{q}) - 2\,\mathrm{I}\left(\partial_u \mathrm{q}, (\partial_v^2 \mathrm{q})^{\mathrm{I}}\right) \\
&= 2\partial_v g_{uv} - 2\,\mathrm{I}\left(\partial_u \mathrm{q}, (\partial_v^2 \mathrm{q})^{\mathrm{I}}\right) \\
&= -2\,\mathrm{I}\left(\partial_u \mathrm{q}, (\partial_v^2 \mathrm{q})^{\mathrm{I}}\right),
\end{aligned}$$

which vanishes at $(0, v)$ since $v \mapsto \mathrm{q}(0, v)$ is a geodesic. In particular

$$\left(\sqrt{g_{vv}}\,\partial_u \sqrt{g_{vv}}\right)(0, v) = \partial_u \sqrt{g_{vv}}(0, v) = 0.$$

This shows that $\sqrt{g_{vv}(u, v)} = 1$ and hence that we have Cartesian coordinates in a neighborhood of a geodesic. $\square$

THEOREM 7.7.2. *(Minding, 1839) If two abstract surfaces have constant Gauss curvature K, then they are locally isometric to each other.*

PROOF. It suffices to show that if a surface has constant curvature K, then it has a parametrization around every point where the first fundamental form only depends on K.

As before we fix a geodesic coordinate system $\mathrm{q}(u,v)$ where all u-curves are unit speed geodesics and $\mathrm{q}(0,v)$ is a unit speed geodesic. The first fundamental form is

$$[\mathrm{I}] = \begin{bmatrix} 1 & 0 \\ 0 & g_{vv} \end{bmatrix},$$

where as in the proof above:

$$\begin{aligned} \sqrt{g_{vv}}(0,v) &= 1, \\ (\partial_u \sqrt{g_{vv}})(0,v) &= 0, \end{aligned}$$

and

$$K = -\frac{\partial_u^2 \sqrt{g_{vv}}}{\sqrt{g_{vv}}}.$$

The last equation dictates how $\sqrt{g_{vv}}$ changes along u curves and the two previous equations are the initial values. When $K=0$ we saw that $\sqrt{g_{vv}}=1$, otherwise

$$\sqrt{g_{vv}(u,v)} = \begin{cases} \cos\left(\sqrt{K}u\right), & K>0, \\ \cosh\left(\sqrt{-K}u\right), & K<0. \end{cases}$$

□

THEOREM 7.7.3. *Any complete simply connected surface M with constant curvature k is bijectively isometric to S_k^2.*

PROOF. We know from theorem 7.7.2 that given $x \in M$ sufficiently small balls $B(x,r) \subset M$ are isometric to balls $B(\bar{x},r) \subset S_k^2$. Furthermore, if $q \in B(x,r)$, $\bar{q} \in S_k^2$, and $L : T_qM \to T_{\bar{q}}S_k^2$ is a linear isometry, then there is a unique bijective isometry $F : B(x,r) \to B(F(x),r) \subset S_k^2$, where $F(q) = \bar{q}$ and $DF|_q = L$. Note that when $k \leq 0$, all metric balls in S_k^2 are convex, while when $k>0$ we need their radius to be $< \frac{\pi}{2\sqrt{k}}$ for this to be true. For the remainder of the proof assume that all metric balls are chosen to be isometric to convex balls in the space form. So for small radii the metrics balls are either disjoint or have connected intersection.

The construction of $F : M \to S_k^2$ proceeds basically in the same way one does analytic continuation on simply connected domains. Fix base points $p \in M$, $\bar{p} \in S_k^2$ and a linear isometry $L : T_pM \to T_{\bar{p}}S_k^2$. Next, let $x \in M$ be an arbitrary point. If $c \in \Omega_{p,x}$ is a curve from p to x in M, then we can cover c by a string of balls $B(p_i,r)$, $i=0,...,k$, where $p=p_0$, $x=p_k$, and $B(p_{i-1},r) \cap B(p_i,r) \neq \emptyset$. Define $F_0 : B(p_0,r) \to S_k^2$ so that $F(p)=\bar{p}$ and $DF_0|_{p_0} = L$. Then define $F_i : B(p_i,r) \to S_k^2$ successively to make it agree with F_{i-1} on $B(p_{i-1},r) \cap B(p_i,r)$ (this just requires their values and differentials agree at one point). Define a function $G : \Omega_{p,x} \to S_k^2$ by $G(c) = F_k(x)$. We have to check that it is well-defined in the sense that it doesn't depend on our specific way of covering the curve. This is easily done by selecting a different covering and then showing that the set of values in $[0,1]$ where the two choices agree is both open and closed.

If $\bar{c} \in \Omega_{p,x}$ is sufficiently close to c, then it lies inside a fixed covering of c, but then it is clear that $G(c) = G(\bar{c})$. This implies that G is locally constant. In particular, G has the same value on all curves in $\Omega_{p,x}$ that are homotopic to each other. Simple-connectivity simply means that all curves are homotopic to each other so G is constant on $\Omega_{p,x}$. This means that $F(x)$ becomes well-defined and a Riemannian isometry.

If M is geodesically complete at a point p, then any point $x \in M$ lies on a unit speed geodesic $\mathrm{q}(t) : [0, \infty) \to M$ so that $\mathrm{q}(0) = p$. The map F will take this to a unit speed geodesic from $\bar{p}$. Now any point in S^2_k lies on a unit speed geodesic that starts at $\bar{p}$, so this shows that F is onto.

If $F(x) = F(y)$, then we have two unit speed geodesics emanating from $\bar{p}$ that intersect at $F(x) = F(y)$. When $k \leq 0$ this is impossible unless the geodesics agree. Thus F is both onto and one-to-one when $k \leq 0$.

In case $k > 0$ two unit speed geodesics in S^2_k that start at $\bar{p}$ can only intersect at the antipodal point $-\bar{p}$. So if we have two different unit speed geodesics $\mathrm{q}_1, \mathrm{q}_2 : [0, \infty) \to M$ with $\mathrm{q}_i(0) = p$. Then $F \circ \mathrm{q}_i(t)$ are different unit speed geodesics emanating from $\bar{p}$ that intersect when $t = n\pi/\sqrt{k}$, $n = 1, 2, 3....$ In particular, $F : B(p, \pi/\sqrt{k}) \to S^2_k - \{-\bar{p}\}$ is one-to-one and $F(S(p, \pi/\sqrt{k})) = \{-\bar{p}\}$. Then $F^{-1} : S^2_k - \{-\bar{p}\} \to B(p, \pi/\sqrt{k})$ is a well-defined isometry that maps points close to $\bar{p}$ to points that are close to $S(p, \pi/\sqrt{k})$. Since points that are close to $\bar{p}$ are also close to each other it must follow that $S(p, \pi/\sqrt{k})$ consists of a single point q. This shows that all geodesics that start at p go through q. We can then conclude that $\bar{B}(p, \pi/\sqrt{k}) = M$ and that $F : M \to S^2_k$ is one-to-one. □

7.8. Comparison Results

In this section we prove several classical results for surfaces where the Gauss curvature is either bounded from below or above. Such results are often referred to as comparison results since they are obtained by a comparison with a corresponding constant curvature geometry.

We start by analyzing the second derivative of energy for some very specific variations.

LEMMA 7.8.1. *(Jacobi, 1842) Let* $\mathrm{q}(u, v)$ *be geodesic coordinates where all u-curves are geodesics along a unit speed geodesic* $\mathrm{q}(0, v)$*. Consider a variation:* $u = su(t)$ *and* $v = t$*, i.e.,* $\mathrm{q}(s, t) = \mathrm{q}(su(t), t)$*, then*

$$\frac{d^2E}{ds^2}|_{s=0} = \int_a^b \left(\dot{u}^2 - Ku^2\right) dt.$$

PROOF. We write the velocity out in coordinates

$$\frac{\partial \mathrm{q}}{\partial t} = s\dot{u}\partial_u \mathrm{q} + \partial_v \mathrm{q}$$

and obtain

$$\mathrm{I}\left(\frac{\partial \mathrm{q}}{\partial t}, \frac{\partial \mathrm{q}}{\partial t}\right) = s^2\dot{u}^2 + g_{vv}.$$

For fixed s the energy of $t \mapsto \mathrm{q}(su(t), t)$ is given by

$$E(s) = \frac{1}{2}\int_a^b \left(s^2\dot{u}^2 + g_{vv}\right) dt,$$

Keeping in mind that $g_{vv} = g_{vv}(su(t), t)$ the derivatives are easily calculated:

$$\frac{dE}{ds} = \int_a^b \left(s\dot{u}^2 + \frac{1}{2}u\partial_u g_{vv}\right) dt,$$

$$\frac{d^2E}{ds^2} = \int_a^b \left(\dot{u}^2 + \frac{1}{2} u^2 \partial_u^2 g_{vv} \right) dt.$$

From example 5.3.9 we have

$$K = -\frac{1}{2} \frac{\partial_u^2 g_{vv}}{g_{vv}} + \frac{1}{4} \left(\frac{\partial_u g_{vv}}{g_{vv}} \right)^2.$$

Since $\mathrm{q}(0,v)$ is a unit speed curve we have $g_{vv}(0,v) = 1$. The derivative is calculated as follows

$$\begin{aligned}
\partial_u g_{vv} &= 2\,\mathrm{I}\left((\partial_u \partial_v \mathrm{q})^{\mathrm{I}}, \partial_v \mathrm{q} \right) \\
&= 2\,\mathrm{I}\left((\partial_v \partial_u \mathrm{q})^{\mathrm{I}}, \partial_v \mathrm{q} \right) \\
&= 2\partial_v \mathrm{I}(\partial_u \mathrm{q}, \partial_v \mathrm{q}) - 2\,\mathrm{I}\left(\partial_u \mathrm{q}, \left(\partial_v^2 \mathrm{q}\right)^{\mathrm{I}} \right) \\
&= 2\partial_v g_{uv} - 2\,\mathrm{I}\left(\partial_u \mathrm{q}, \left(\partial_v^2 \mathrm{q}\right)^{\mathrm{I}} \right) \\
&= -2\,\mathrm{I}\left(\partial_u \mathrm{q}, \left(\partial_v^2 \mathrm{q}\right)^{\mathrm{I}} \right).
\end{aligned}$$

This vanishes when $u = 0$ since $\mathrm{q}(0,v)$ is a geodesic. The result now follows. □

COROLLARY 7.8.2. *(Bonnet, 1855) If $K \geq R^{-2} > 0$, then no geodesic of length $> \pi R$ is minimal.*

PROOF. We can assume that the geodesic doesn't intersect itself (if it does it is clearly not minimal) and construct geodesic coordinates where $\mathrm{q}(0,v)$ is the given geodesic parametrized by arclength on $[0,L]$. Then select a variation as in lemma 7.8.1 of the form $u(t) = \sin(t\pi/L)$. This will yield a proper variation with the second derivative of energy satisfying

$$\begin{aligned}
\frac{d^2E}{ds^2}\Big|_{s=0} &= \int_0^L \left(\dot{u}^2 - K u^2 \right) dt \\
&\leq \int_0^L \left(\left(\frac{\pi}{L}\right)^2 \cos^2(t\pi/L) - R^{-2} \sin^2(t\pi/L) \right) dt \\
&= \left(\frac{\pi}{L}\right)^2 \int_0^L \cos^2(t\pi/L)\, dt - R^{-2} \int_0^L \sin^2(t\pi/L)\, dt \\
&= \left(\left(\frac{\pi}{L}\right)^2 - R^{-2} \right) \frac{L}{2}.
\end{aligned}$$

This is strictly negative when $L > \pi R$ showing that the geodesic is a local maximum for the energy. Since the variation is fixed at the end points there will be nearby curves of strictly smaller energy with the same end points. Corollary 7.3.6 then shows that it can't be a minimum for the length functional. □

COROLLARY 7.8.3. *(Hopf-Rinow, 1931) If a complete surface satisfies $K \geq R^{-2} > 0$, then all distances are $\leq \pi R$ and must in particular be a closed surface.*

THEOREM 7.8.4. *If a closed surface has positive curvature, then any two closed geodesics intersect.*

PROOF. Assume otherwise and obtain a shortest geodesic between the two closed geodesics. This geodesic is perpendicular to both of the closed geodesics. In particular if we let it be the $\mathrm{q}(0,v)$ curve in a geodesic parametrization, then the curves $\mathrm{q}(u,0)$ and $\mathrm{q}(u,L)$ are our two closed geodesics. Now consider the variation where $s=u$ and $t=v$, then the second variation is given by

$$\frac{d^2E}{ds^2}|_{s=0} = \int_a^b \left(\dot{u}^2 - Ku^2\right) dt = \int_a^b -Ku^2 dt < 0.$$

This shows that the curves $v \mapsto \mathrm{q}(u,v)$ are shorter than L. As they are also curves between the two closed geodesics this contradicts that our original curve was the shortest such curve. □

THEOREM 7.8.5. *(Mangoldt, 1881, Hadamard, 1889?) A complete surface M with $K \leq 0$ admits a global parametrization $\mathrm{q}(u,v)$ where $(u,v) \in \mathbb{R}^2$. If on $\mathbb{R}^2$ we introduce the first fundamental form from M, then we obtain a complete metric on $\mathbb{R}^2$ with $K \leq 0$ where all geodesics are minimal.*

PROOF. The parametrization is given by the exponential map. Identify a fixed tangent space T_pM with $\mathbb{R}^2$ via a choice of orthonormal basis E_1, E_2 and introduce Cartesian (x,y) as well as polar coordinates (r,θ). We can use $\mathrm{q}(r,\theta) = \exp_p(r\cos\theta E_1 + r\sin\theta E_2)$ as a potential parametrization on M. Even when it isn't a parametrization as in lemma 7.4.4 we note that it is a geodesic variation with the radial lines as unit speed geodesics. We have the velocity fields $\partial_r \mathrm{q}, \partial_\theta \mathrm{q}$ for the r- and θ-curves which for each (r,θ) give us tangent vectors in $T_{\mathrm{q}(r,\theta)}M$. Since the r-curves are unit speed geodesics we have $|\partial_r \mathrm{q}| = 1$ everywhere. We can also show that $\mathrm{I}(\partial_r \mathrm{q}, \partial_\theta \mathrm{q}) = 0$. First note that it vanishes at $r=0$ since $\partial_\theta \mathrm{q}(0,\theta) = 0$. Next observe that $\mathrm{I}(\partial_r \mathrm{q}, \partial_\theta \mathrm{q}) = 0$ is constant since

$$\begin{aligned}
\partial_r \mathrm{I}(\partial_r \mathrm{q}, \partial_\theta \mathrm{q}) &= \mathrm{I}\left(\left(\partial_r^2 \mathrm{q}\right)^{\mathrm{I}}, \partial_\theta \mathrm{q}\right) + \mathrm{I}\left(\partial_r \mathrm{q}, \left(\partial_r\partial_\theta \mathrm{q}\right)^{\mathrm{I}}\right) \\
&= \mathrm{I}\left(\partial_r \mathrm{q}, \left(\partial_\theta\partial_r \mathrm{q}\right)^{\mathrm{I}}\right) \\
&= \frac{1}{2}\partial_\theta \mathrm{I}(\partial_r \mathrm{q}, \partial_r \mathrm{q}) \\
&= 0.
\end{aligned}$$

Thus $\mathrm{I}(\partial_r \mathrm{q}, \partial_\theta \mathrm{q}) = 0$ everywhere. It follows that $D\exp_p$ is nonsingular at a point (r,θ) precisely when $\mathrm{I}(\partial_\theta \mathrm{q}, \partial_\theta \mathrm{q}) > 0$ at (r,θ).

Define a first fundamental form on $\mathbb{R}^2$ by

$$\begin{bmatrix} g_{rr} & g_{r\theta} \\ g_{\theta r} & g_{\theta\theta} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & g_{\theta\theta} \end{bmatrix},$$

where

$$g_{\theta\theta} = \left|D\exp_p(-yE_1 + xE_2)\right|^2 = \mathrm{I}(\partial_\theta \mathrm{q}, \partial_\theta \mathrm{q}).$$

When $D\exp_p$ is nonsingular this corresponds precisely to the first fundamental form of M in this parametrization.

By continuity $\exp_p : T_pM \to M$ is nonsingular on some open set O that contains the origin. Let $B(0,R) \subset O$ be the largest ball inside O. We claim that $R = \infty$ and note that if $R < \infty$ then the closure $\bar{B}(0,R)$ cannot be contained in O. On $B(0,R)$ the (r,θ)-coordinates are geodesic polar coordinates with respect to

$$\begin{bmatrix} 1 & 0 \\ 0 & g_{\theta\theta} \end{bmatrix}.$$

Since they correspond to the first fundamental form on M the Gauss curvature satisfies

$$0 \geq K = -\frac{\partial_r^2 \sqrt{g_{\theta\theta}}}{\sqrt{g_{\theta\theta}}}.$$

In particular, $\partial_r^2 \sqrt{g_{\theta\theta}} \geq 0$. Consequently, $\sqrt{g_{\theta\theta}}$ is a nonnegative convex function in r. Moreover it vanishes at $r = 0$ and is positive for $r \in (0, R)$. Thus it is impossible for this function to vanish when $r = R$. This shows that $B(0, R) \subset O$ can't be maximal unless $R = \infty$.

This gives us the desired global parametrization on M with a first fundamental form on $\mathbb{R}^2$ that has $K \leq 0$. This will also help us establish the second part of the result. In fact, no metric on $\mathbb{R}^2$ with $K \leq 0$ can have geodesics that intersect at more than one point as that would violate Gauss-Bonnet. Consider two geodesics $\mathrm{q}_1(t)$ and $\mathrm{q}_2(t)$ with $\mathrm{q}_i(0) = p$. By lemma 7.4.4 they can't intersect near p. Therefore, if they intersect at some later point, then there will a point $q \neq p$ closest to p where they intersect. In this case we can after reparametrizing assume that $\mathrm{q}_i(1) = q$ and that when restricted to $t \in [0, 1]$ there are no other intersections between the geodesics. Now create a triangle by using p, q, and, say $\mathrm{q}_1(1/2)$, as vertices. This triangle has angle sum $> \pi$ as one angle is π. This however, violates the Gauss-Bonnet theorem as the whole triangle is a simple closed curve of rotation index 2π when oriented appropriately. Specifically, as the geodesic curvature vanishes the Gauss-Bonnet theorem 6.4.2 tells us

$$0 \geq \int_{\mathrm{q}(R)} K dA = 2\pi - \sum \theta_i,$$

where θ_i are the exterior angles at the three vertices. Since they are complementary to the interior angles α, β, γ we have

$$0 \geq \int_{\mathrm{q}(R)} K dA = 2\pi - \sum \theta_i = -\pi + \alpha + \beta + \gamma.$$

□

REMARK 7.8.6. There are different proofs of the latter part that do not appeal to the Gauss-Bonnet theorem.

CHAPTER 8

Riemannian Geometry

As with abstract surfaces we simply define what the dot products of the tangent fields should be:

$$[\mathrm{I}] = \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u^1} & \cdots & \frac{\partial \mathrm{q}}{\partial u^n} \end{bmatrix}^t \begin{bmatrix} \frac{\partial \mathrm{q}}{\partial u^1} & \cdots & \frac{\partial \mathrm{q}}{\partial u^n} \end{bmatrix} = \begin{bmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \cdots & g_{nn} \end{bmatrix}$$

The notation $\frac{\partial \mathrm{q}}{\partial u^i} = \partial_i \mathrm{q}$ for the tangent field that corresponds to the velocity of the u^i curves is borrowed from our view of what happens on a surface.

We have the very general formula for how vectors are expanded

$$\begin{aligned} V &= \begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix} \left(\begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix}^t \begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix} \right)^{-1} \begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix}^t V \\ &= \begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix} \begin{bmatrix} U_1 \cdot U_1 & \cdots & U_1 \cdot U_n \\ \vdots & \ddots & \vdots \\ U_n \cdot U_1 & \cdots & U_n \cdot U_n \end{bmatrix}^{-1} \begin{bmatrix} U_1 \cdot V \\ \vdots \\ U_n \cdot V \end{bmatrix} \end{aligned}$$

provided we know how to compute dot products of the basis vectors and dots products of V with the basis vectors. So we will now assume that were are given a symmetric matrix $[\mathrm{I}] = [g_{ij}]$ of functions on some domain $U \subset \mathbb{R}^n$ that uses u^i as parameters. We shall further assume that this first fundamental form has non-vanishing determinant so that we can calculate the inverse $[\mathrm{I}]^{-1} = \left[g^{ij}\right]$. We shall then think of $g_{ij} = \mathrm{I}(\partial_i \mathrm{q}, \partial_j \mathrm{q})$ as describing the inner product of the coordinate vector fields and q as a point on the space we are investigating. When dealing with surfaces we also used that this defined an inner product. For the moment we will not need this condition.

We can define the Christoffel symbols in relation to the tangent fields when we know the dot products of those tangent fields:

$$\begin{aligned} \Gamma_{ijk} &= \frac{1}{2}\left(\partial_j g_{ki} + \partial_i g_{kj} - \partial_k g_{ij}\right), \\ \Gamma_{ij}^k &= \sum_l g^{kl} \Gamma_{ijl}. \end{aligned}$$

PROPOSITION 8.0.1. *The metric and Christoffel symbols are also related by*

$$\begin{aligned} \partial_k g_{ij} &= \Gamma_{kij} + \Gamma_{kji} \\ \partial_k g^{ij} &= -\sum_l g^{il} \Gamma_{kl}^j + g^{jl} \Gamma_{kl}^i \end{aligned}$$

PROOF. The first formula follows directly from the definition

$$\begin{aligned}\Gamma_{kij}+\Gamma_{kji} &= \frac{1}{2}\left(\partial_k g_{ij}+\partial_i g_{kj}-\partial_j g_{ki}\right)\\ &\quad+\frac{1}{2}\left(\partial_k g_{ji}+\partial_j g_{ki}-\partial_i g_{kj}\right)\\ &= \partial_k g_{ji}.\end{aligned}$$

For the second we first have to calculate the derivative of the inverse of a matrix. Symbolically this is done as follows. If $I_n=\left[\delta^i_j\right]$ denotes the identity matrix then

$$\begin{aligned}I_n &= [\mathrm{I}]\,[\mathrm{I}]^{-1}\\ \delta_i^j &= g_{ik}g^{kj}\end{aligned}$$

so

$$\begin{aligned}0=\partial_s I_n &= \left(\partial_s[\mathrm{I}]\right)[\mathrm{I}]^{-1}+[\mathrm{I}]\,\partial_s[\mathrm{I}]^{-1}\\ 0=\partial_s\delta_i^j &= \sum_l \partial_s g_{il}g^{lj}+\sum_k g_{ik}\partial_s g^{kj}\end{aligned}$$

showing that

$$\begin{aligned}\partial_s[\mathrm{I}]^{-1} &= -[\mathrm{I}]^{-1}\left(\partial_s[\mathrm{I}]\right)[\mathrm{I}]^{-1}\\ \partial_s g^{kj} &= -\sum_{i,l} g^{ki}\partial_s g_{il}g^{lj}\end{aligned}$$

We can now use the first formula to prove the second

$$\begin{aligned}\partial_k g^{ij} &= -\sum_{s,t} g^{is}\partial_k g_{st}g^{tj}\\ &= -\sum_{s,t} g^{is}\left(\Gamma_{kst}+\Gamma_{kts}\right)g^{tj}\\ &= -\sum_{s,t} g^{is}\Gamma_{kst}g^{tj}-\sum_{s,t} g^{is}\Gamma_{kts}g^{tj}\\ &= -\sum_s g^{is}\Gamma^j_{ks}-\sum_t \Gamma^i_{kt}g^{tj}\\ &= -\sum_l g^{il}\Gamma^j_{kl}+g^{jl}\Gamma^i_{kl}\end{aligned}$$

□

While we have not yet specified where q is placed we can still attempt to define second partials intrinsically. This means that we imitate what happened for surfaces but assume that there is no normal vector.

To start with we should have

$$\partial^2_{ij}\,\mathrm{q}\cdot\partial_k\,\mathrm{q}=\Gamma_{ijk}$$

leading to

$$\begin{aligned}\partial^2_{ij}\,\mathrm{q} &= \begin{bmatrix}\partial_1\mathrm{q} & \cdots & \partial_n\mathrm{q}\end{bmatrix}[\mathrm{I}]^{-1}\begin{bmatrix}\Gamma_{ij1} & \cdots & \Gamma_{ijn}\end{bmatrix}^t\\ &= \begin{bmatrix}\partial_1\mathrm{q} & \cdots & \partial_n\mathrm{q}\end{bmatrix}\begin{bmatrix}\Gamma^1_{ij}\\ \vdots\\ \Gamma^n_{ij}\end{bmatrix}.\end{aligned}$$

Note that the symmetry of the metric and Christoffel symbols tell us that we still have

$$\partial_{ij}^2 \mathrm{q} = \partial_{ji}^2 \mathrm{q}\,.$$

This will allow us to define intrinsic acceleration and hence geodesics. It'll also allow us to show that the stationary curves for energy are geodesics. If in addition the metric is positive definite, i.e., $\mathrm{I}(V,V) > 0$ unless $V = 0$, then we can define the length of vectors and consider arc-length of curves. It will then also be true that short geodesics minimize arc-length.

To define curvature we collect the Gauss formulas

$$\begin{aligned}\partial_i \begin{bmatrix} \partial_1 \mathrm{q} & \cdots & \partial_n \mathrm{q} \end{bmatrix} &= \begin{bmatrix} \partial_1 \mathrm{q} & \cdots & \partial_n \mathrm{q} \end{bmatrix} \begin{bmatrix} \Gamma_{i1}^1 & \cdots & \Gamma_{in}^1 \\ \vdots & \ddots & \vdots \\ \Gamma_{i1}^n & \cdots & \Gamma_{in}^n \end{bmatrix} \\ &= \begin{bmatrix} \partial_1 \mathrm{q} & \cdots & \partial_n \mathrm{q} \end{bmatrix} [\Gamma_i]\end{aligned}$$

and form the expression

$$\partial_i [\Gamma_j] - \partial_j [\Gamma_i] + [\Gamma_i][\Gamma_j] - [\Gamma_j][\Gamma_i]$$

that we used to define the curvatures involved in the Gauss equations.

This time we don't have a Gauss curvature, but we can define the Riemann curvature as the k, l entry in this expression:

$$\begin{aligned}[R_{ij}] &= \partial_i [\Gamma_j] - \partial_j [\Gamma_i] + [\Gamma_i][\Gamma_j] - [\Gamma_j][\Gamma_i]\,, \\ R_{ijk}^l &= \partial_i \Gamma_{jk}^l - \partial_j \Gamma_{ik}^l + \begin{bmatrix} \Gamma_{i1}^l & \cdots & \Gamma_{in}^l \end{bmatrix} \begin{bmatrix} \Gamma_{jk}^1 \\ \vdots \\ \Gamma_{jk}^n \end{bmatrix} - \begin{bmatrix} \Gamma_{j1}^l & \cdots & \Gamma_{jn}^l \end{bmatrix} \begin{bmatrix} \Gamma_{ik}^1 \\ \vdots \\ \Gamma_{ik}^n \end{bmatrix}\end{aligned}$$

This expression shows how certain third order partials might not commute as this formula indicates that

$$\partial_{ijk}^3 \mathrm{q} - \partial_{jik}^3 \mathrm{q} = \begin{bmatrix} \partial_1 \mathrm{q} & \cdots & \partial_n \mathrm{q} \end{bmatrix} \begin{bmatrix} R_{ijk}^1 \\ \vdots \\ R_{ijk}^n \end{bmatrix}$$

But recall that since second order partials do commute we have

$$\partial_{ijk}^3 \mathrm{q} = \partial_{ikj}^3 \mathrm{q}$$

Thus the third order partials commute if and only if the Riemann curvature vanishes. This can be used to establish the difficult existence part of the next result.

THEOREM 8.0.2. [Riemann] *The Riemann curvature vanishes if and only if there are Cartesian coordinates around any point.*

PROOF. The easy direction is to assume that Cartesian coordinates exist. Certainly this shows that the curvatures vanish when we use Cartesian coordinates, but this does not guarantee that they also vanish in some arbitrary coordinate system. For that we need to figure out how the curvature terms change when we change coordinates. A long tedious calculation shows that if the new coordinates are called v^i and the curvature in these coordinates $\tilde{R}_{ijk}^l$, then

$$\tilde{R}_{ijk}^l = \frac{\partial u^\alpha}{\partial v^i} \frac{\partial u^\beta}{\partial v^j} \frac{\partial u^\gamma}{\partial v^k} \frac{\partial v^l}{\partial u^\delta} R_{\alpha\beta\gamma}^\delta.$$

Thus we see that if the all curvatures vanish in one coordinate system, then they vanish in all coordinate systems.

Conversely, to find Cartesian coordinates we set up a system of differential equations

$$\begin{aligned} \partial_i \mathrm{q} &= U_i \\ \partial_i \begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix} &= \begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix} [\Gamma_i] \end{aligned}$$

whose integrability conditions are a consequence of having vanishing curvature. We select a point $u_0 \in U$ in our given parametrization and assume that we are looking for a map $\mathrm{q} : U \to \mathbb{R}^n$ where $\mathrm{q}(u_0) = 0$ and $U_i(u_0) = u_i$ a suitable basis for $\mathbb{R}^n$.

The integrability conditions for the first set of equations are

$$\partial_i U_j = \partial_j U_i$$

which from the second set of equations imply that

$$\begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix} \begin{bmatrix} \Gamma^1_{ij} \\ \vdots \\ \Gamma^n_{ij} \end{bmatrix} = \begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix} \begin{bmatrix} \Gamma^1_{ji} \\ \vdots \\ \Gamma^n_{ji} \end{bmatrix}$$

These conditions holds since $\Gamma^k_{ij} = \Gamma^k_{ji}$.

For the second set of equations the integrability conditions are given by

$$\partial_i \left(\begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix} [\Gamma_j] \right) = \partial_j \left(\begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix} [\Gamma_i] \right)$$

which we know reduce to

$$\partial_i [\Gamma_j] + [\Gamma_i][\Gamma_j] = \partial_j [\Gamma_i] + [\Gamma_j][\Gamma_i]$$

These conditions hold because we assume that $[R_{ij}] = 0$.

This means that we can solve these equations on some neighborhood of $u_0 \in U$ with the specified initial conditions. We then have to show that the new parametrization is Cartesian. The new parameters are given by the coordinates for q, i.e.,

$$\left(x^1, ..., x^n\right) = \left(\mathrm{q}^1, ..., \mathrm{q}^n\right) = \mathrm{q}\left(u^1, ..., u^n\right)$$

Thus

$$\partial_j \mathrm{q}^k = \frac{\partial \mathrm{q}^k}{\partial u^j} = U^k_j$$

and

$$\partial_{ij} \mathrm{q}^k = \partial_i U^k_j = U^k_l \Gamma^l_{ij} = \partial_l \mathrm{q}^k \, \Gamma^l_{ij}$$

The new first fundamental form is then given by

$$\begin{aligned} \tilde{g}_{kl} &= \frac{\partial u^i}{\partial x^k} g_{ij} \frac{\partial u^j}{\partial x^l} \\ [\tilde{g}_{kl}] &= \left[\frac{\partial u^i}{\partial x^k}\right] [g_{ij}] \left[\frac{\partial u^j}{\partial x^l}\right] \end{aligned}$$

Unfortunately we don't know what the matrix $\left[\frac{\partial u^i}{\partial x^k}\right]$ is. It is given as the inverse of $\left[\frac{\partial x^k}{\partial u^i}\right]$ which in turn is the matrix $\begin{bmatrix} U_1 & \cdots & U_n \end{bmatrix}$ by our first equations. This means that we have

$$\tilde{g}^{kl} = \frac{\partial x^k}{\partial u^i} g^{ij} \frac{\partial u^l}{\partial x^j} = \partial_i \mathrm{q}^k \, g^{ij} \partial_j \mathrm{q}^l$$

We can now calculate the derivative of this as

$$\begin{aligned}
\partial_s \tilde{g}^{kl} &= \partial_{si}\,\mathrm{q}^k\, g^{ij}\partial_j\,\mathrm{q}^l + \partial_i\,\mathrm{q}^k\, g^{ij}\partial_{sj}\,\mathrm{q}^l \\
&\quad + \partial_i\,\mathrm{q}^k\,\partial_s g^{ij}\partial_j\,\mathrm{q}^l \\
&= \partial_t\,\mathrm{q}^k\,\Gamma^t_{si} g^{ij}\partial_j\,\mathrm{q}^l + \partial_i\,\mathrm{q}^k\, g^{ij}\partial_t\,\mathrm{q}^l\,\Gamma^t_{sj} \\
&\quad - \partial_i\,\mathrm{q}^k\left(g^{it}\Gamma^j_{st} + g^{jt}\Gamma^i_{st}\right)\partial_j\,\mathrm{q}^l \\
&= \partial_t\,\mathrm{q}^k\,\Gamma^t_{si} g^{ij}\partial_j\,\mathrm{q}^l - \partial_i\,\mathrm{q}^k\, g^{jt}\Gamma^i_{st}\partial_j\,\mathrm{q}^l \\
&\quad + \partial_i\,\mathrm{q}^k\, g^{ij}\partial_t\,\mathrm{q}^l\,\Gamma^t_{sj} - \partial_i\,\mathrm{q}^k\, g^{it}\Gamma^j_{st}\partial_j\,\mathrm{q}^l \\
&= 0 + 0
\end{aligned}$$

showing that the new metric coefficients are constant. We can then specify the basis u_i so that the new metric becomes Cartesian at u_0 and hence Cartesian everywhere since the metric coefficients are constant. $\square$

APPENDIX A

Vector Calculus

A.1. Vector and Matrix Notation

Given a basis e, f for a two-dimensional vector space we expand vectors using matrix multiplication

$$v = v^e e + v^f f = \begin{bmatrix} e & f \end{bmatrix} \begin{bmatrix} v^e \\ v^f \end{bmatrix}$$

The matrix representation $[L]$ for a linear map/transformation L can be found from

$$\begin{aligned} \begin{bmatrix} L(e) & L(f) \end{bmatrix} &= \begin{bmatrix} e & f \end{bmatrix} [L] \\ &= \begin{bmatrix} e & f \end{bmatrix} \begin{bmatrix} L_e^e & L_f^e \\ L_e^f & L_f^f \end{bmatrix} \end{aligned}$$

Next we relate matrix multiplication and the dot product in $\mathbb{R}^3$. We think of vectors as being columns or 3×1 matrices. Keeping that in mind and using transposition of matrices we immediately obtain:

$$\begin{aligned} X^t Y &= X \cdot Y, \\ X^t \begin{bmatrix} X_2 & Y_2 \end{bmatrix} &= \begin{bmatrix} X \cdot X_2 & X \cdot Y_2 \end{bmatrix} \\ \begin{bmatrix} X_1 & Y_1 \end{bmatrix}^t X &= \begin{bmatrix} X_1 \cdot X \\ Y_1 \cdot X \end{bmatrix} \\ \begin{bmatrix} X_1 & Y_1 \end{bmatrix}^t \begin{bmatrix} X_2 & Y_2 \end{bmatrix} &= \begin{bmatrix} X_1 \cdot X_2 & X_1 \cdot Y_2 \\ Y_1 \cdot X_2 & Y_1 \cdot Y_2 \end{bmatrix}, \\ \begin{bmatrix} X_1 & Y_1 & Z_1 \end{bmatrix}^t \begin{bmatrix} X_2 & Y_2 & Z_2 \end{bmatrix} &= \begin{bmatrix} X_1 \cdot X_2 & X_1 \cdot Y_2 & X_1 \cdot Z_2 \\ Y_1 \cdot X_2 & Y_1 \cdot Y_2 & Y_1 \cdot Z_2 \\ Z_1 \cdot X_2 & Z_1 \cdot Y_2 & Z_1 \cdot Z_2 \end{bmatrix} \end{aligned}$$

These formulas can be used to calculate the coefficients of a vector with respect to a general basis. Recall first that if E_1, E_2 is an orthonormal basis for $\mathbb{R}^2$, then

$$\begin{aligned} X &= (X \cdot E_1) E_1 + (X \cdot E_2) E_2 \\ &= \begin{bmatrix} E_1 & E_2 \end{bmatrix} \begin{bmatrix} E_1 & E_2 \end{bmatrix}^t X \end{aligned}$$

So the coefficients for X are simply the dot products with the basis vectors. More generally we have

Theorem A.1.1. *Let U, V be a basis for $\mathbb{R}^2$, then*

$$\begin{aligned} X &= \begin{bmatrix} U & V \end{bmatrix} \left(\begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} U & V \end{bmatrix} \right)^{-1} \begin{bmatrix} U & V \end{bmatrix}^t X \\ &= \begin{bmatrix} U & V \end{bmatrix} \left(\begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} U & V \end{bmatrix} \right)^{-1} \begin{bmatrix} U \cdot X \\ V \cdot X \end{bmatrix} \end{aligned}$$

PROOF. First write

$$X = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} X^u \\ X^v \end{bmatrix}$$

The goal is to find a formula for the coefficients X^u, X^v in terms of the dot products $X \cdot U, X \cdot V$. To that end we notice

$$\begin{aligned} \begin{bmatrix} U \cdot X \\ V \cdot X \end{bmatrix} &= \begin{bmatrix} U & V \end{bmatrix}^t X \\ &= \begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} X^u \\ X^v \end{bmatrix} \end{aligned}$$

Showing directly that

$$\begin{bmatrix} X^u \\ X^v \end{bmatrix} = \left(\begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} U & V \end{bmatrix} \right)^{-1} \begin{bmatrix} U \cdot X \\ V \cdot X \end{bmatrix}$$

and consequently

$$X = \begin{bmatrix} U & V \end{bmatrix} \left(\begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} U & V \end{bmatrix} \right)^{-1} \begin{bmatrix} U \cdot X \\ V \cdot X \end{bmatrix}$$

□

REMARK A.1.2. There is a similar formula in $\mathbb{R}^3$ which is a bit longer. In practice we shall only need it in the case where the third basis vector is perpendicular to the first two. Also note that if U, V are orthonormal then

$$\begin{bmatrix} U & V \end{bmatrix}^t \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and we recover the standard formula for the expansion of a vector in an orthonormal basis.

THEOREM A.1.3. *A real symmetric matrix, or symmetric linear operator on a finite dimensional Euclidean space, has an orthonormal basis of eigenvectors.*

PROOF. First observe that if we have two eigenvectors

$$Av = \lambda v, \ Aw = \mu w$$

where $\lambda \neq \mu$, then

$$\begin{aligned} (\lambda - \mu) \left(v^t w \right) &= (\lambda v)^t w - v^t (\mu w) \\ &= (Av)^t w - v^t (Aw) \\ &= v^t A^t w - v^t A w \\ &= v^t A w - v^t A w \\ &= 0 \end{aligned}$$

so it must follow that $v \perp w$.

This shows that the eigenspaces are all perpendicular to each other. Thus we are reduced to showing that such matrices only have real eigenvalues. There are many fascinating proofs of this. We give a fairly down to earth proof in the cases that are relevant to us.

For a 2×2 matrix

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

the characteristic polynomial is

$$\lambda^2 - (a+d)\,\lambda + ad - b^2$$

so the discriminant is

$$\Delta = (a+d)^2 - 4\left(ad - b^2\right) = (a-d)^2 + 4b^2 \geq 0$$

This shows that the roots must be real.

For a 3×3 matrix the characteristic polynomial is cubic. The intermediate value theorem then guarantees at least one real root. If we make a change of basis to another orthonormal basis where the first basis vector is an eigenvector then the new matrix will still be symmetric and look like

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & a & b \\ 0 & b & d \end{bmatrix}$$

The characteristic polynomial then looks like

$$(\lambda - \lambda_1)\left(\lambda^2 - (a+d)\,\lambda + ad - b^2\right)$$

where we see as before that $\lambda^2 - (a+d)\,\lambda + ad - b^2$ has two real roots. □

A.2. Geometry

Here are a few geometric formulas that use vector notation:

- The length or size of a vector X is denoted:
$$|X| = \sqrt{X^t \cdot X}$$
- The distance from X to a point P:
$$|X - P|$$
- The projection of a vector X onto another vector N:
$$\frac{X \cdot N}{|N|^2} N$$
- The signed distance from P to a plane that goes through X_0 and has normal N, i.e., given by $(X - X_0) \cdot N = 0$:
$$\frac{(P - X_0) \cdot N}{|N|}$$
the actual distance is the absolute value of the signed distance. This formula also works for the (signed) distance from a point to a line in $\mathbb{R}^2$.
- The distance from P to a line with direction N that passes through X_0:
$$\left|(P - X_0) - \frac{(P - X_0) \cdot N}{|N|^2} N\right| = \sqrt{|P - X_0|^2 - \frac{|(P - X_0) \cdot N|^2}{|N|^2}}$$
- The area of a parallelogram spanned by two vectors X, Y is
$$\sqrt{\det\left(\begin{bmatrix} X & Y \end{bmatrix}^t \begin{bmatrix} X & Y \end{bmatrix}\right)}$$
- If $X, Y \in \mathbb{R}^2$ there is also a signed area given by
$$\det \begin{bmatrix} X & Y \end{bmatrix}$$

- If $X, Y \in \mathbb{R}^3$ the area is also given by
$$|X \times Y|$$
- The volume of a parallelepiped spanned by X, Y, Z is
$$\sqrt{\det\left(\left[\begin{array}{ccc} X & Y & Z \end{array}\right]^t \left[\begin{array}{ccc} X & Y & Z \end{array}\right]\right)}$$
- If $X, Y, Z \in \mathbb{R}^3$ the signed volume is given by
$$\begin{aligned} \det\left[\begin{array}{ccc} X & Y & Z \end{array}\right] &= X \cdot (Y \times Z) \\ &= X^t (Y \times Z) \end{aligned}$$
- The

A.3. Geometry of Space-Time

We collect a few of the special features of space-time $\mathbb{R}^{2,1}$ where we use the inner product
$$X \cdot Y = X^x Y^x + X^y Y^y - X^z Y^z.$$

A.4. Differentiation and Integration

A.4.1. Directional Derivatives. If h is a function on $\mathbb{R}^3$ and $X = (P, Q, R)$ then
$$\begin{aligned} D_X h &= P\frac{\partial h}{\partial x} + Q\frac{\partial h}{\partial y} + R\frac{\partial h}{\partial z} \\ &= (\nabla h) \cdot X \\ &= [\nabla h]^t [X] \\ &= \left[\begin{array}{ccc} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{array}\right] [X] \end{aligned}$$
and for a vector field V we get
$$D_X V = \left[\begin{array}{ccc} \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} & \frac{\partial V}{\partial z} \end{array}\right] [X].$$
We can also calculate directional derivatives by selecting a curve such that $\dot{c}(0) = X$. Along the curve the chain rule says:
$$\frac{d(V \circ c)}{dt} = \left[\begin{array}{ccc} \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} & \frac{\partial V}{\partial z} \end{array}\right] \left[\frac{dc}{dt}\right] = D_{\dot{c}} V$$
Thus
$$D_X V = \frac{d(V \circ c)}{dt}(0)$$

A.4.2. Chain Rules. Consider a vector function $V : \mathbb{R}^3 \to \mathbb{R}^n$ and a curve $c : I \to \mathbb{R}^3$. That the curves goes in to space and the vector function is defined on the same space is important, but that it has dimension 3 is not. Note also that the vector function can have values in a higher or lower dimensional space.

The chain rule for calculating the derivative of the composition $V \circ c$ is:
$$\frac{d(V \circ c)}{dt} = \left[\begin{array}{ccc} \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} & \frac{\partial V}{\partial z} \end{array}\right] \left[\frac{dc}{dt}\right]$$

There is a very convenient short cut for writing such chain rules if we keep in mind that they simply involve matrix notation. Write

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and

$$c(t) = X(t)$$

Then this chain rule can be written as

$$\frac{d(V \circ c)}{dt} = \frac{\partial V}{\partial X}\frac{dX}{dt}$$

were we think of

$$\frac{\partial V}{\partial X} = \frac{\partial V}{\partial (x, y, z)} = \begin{bmatrix} \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} & \frac{\partial V}{\partial z} \end{bmatrix}$$

and

$$\frac{dX}{dt} = \frac{d}{dt}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

It is also sometimes convenient to have X be a function of several variables, say, $X(u, v)$. In that case we obtain

$$\frac{\partial V(X(u, v))}{\partial u} = \frac{\partial V}{\partial X}\frac{\partial X}{\partial u}$$

as partial derivatives are simple regular derivatives in one variable when all other variables are fixed.

A.4.3. Local Invertibility. Mention Inverse and Implicit Function Theorems. Lagrange multipliers.

A.4.4. Integration. Change of variables. Green's, divergence, and Stokes' thms. Use Green Thm to prove the change of variable formula, and similarly with Stokes.

A.5. Differential Equations

The basic existence and uniqueness theorem for systems of first order equations is contained in the following statement. The first part is standard and can be found in most text books. The second part about the assertion of smoothness in relation to the initial value is very important, but is somewhat trickier to establish.

THEOREM A.5.1. *Given a smooth function $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ the initial value problem*

$$\frac{d}{dt}x = F(t, x), \; x(0) = x_0$$

has a solution

$$x(t) = \begin{bmatrix} x^1(t) \\ \vdots \\ x^n(t) \end{bmatrix}$$

that is unique on some possibly small interval $(-\epsilon, \epsilon)$. When $|x_0| \leq R$, we can pick ϵ independently of x_0. Moreover this solution is smooth in both t and the initial *value x_0. In case $|F(t,x)| \leq M + C|x|$ for constants $M, C \geq 1$ we can choose $\epsilon = \infty$.*

Proof. The proof is quite long and consists of several different proof. The existence and uniqueness is relatively standard. The long term existence is less standard so we supply a proof below. The smoothness on initial values is also standard but not covered in all texts (see , however, MM for a good proof).

Long term existence:.................................... $\square$

The above result was strictly about ODEs (ordinary differential equations), but it can be used to prove certain results about PDEs (partial differential equations) as well.

We consider a system

$$\begin{aligned}\frac{\partial}{\partial u}x &= P(u,v,x)\\ \frac{\partial}{\partial v}x &= Q(u,v,x)\\ x(0,0) &= x_0\end{aligned}$$

where $x(u,v)$ is now a function of two variables with values in $\mathbb{R}^n$.

The standard situation from multivariable calculus is:

Theorem A.5.2. (Clairaut's Theorem) *When $P = P(u,v)$ and $Q = Q(u,v)$ do not depend on x a solution to*

$$\begin{aligned}\frac{\partial}{\partial u}x &= P(u,v)\\ \frac{\partial}{\partial v}x &= Q(u,v)\\ x(0,0) &= x_0\end{aligned}$$

can be found if and only if the system is exact*, i.e.,*

$$\frac{\partial}{\partial u}Q = \frac{\partial}{\partial v}P.$$

This solution will be defined on all of $\mathbb{R}^2$ provided $P, Q : \mathbb{R}^2 \to \mathbb{R}$.

Proof. If such a solution exists, then it follows that

$$\frac{\partial}{\partial u}Q = \frac{\partial^2 x}{\partial u \partial v} = \frac{\partial^2 x}{\partial v \partial u} = \frac{\partial}{\partial v}P.$$

Conversely start by defining $x_1(u)$ as

$$x_1(u) = x_0 + \int_0^u P(s,0)\,ds.$$

Next define the function $x(u,v)$ for a fixed u by

$$x(u,v) = x_1(u) + \int_0^v Q(u,t)\,dt.$$

This gives us

$$\frac{\partial x}{\partial v} = Q,\ x(0,0) = x_0.$$

Thus it remains to check that

$$\frac{\partial x}{\partial u} = P.$$

Note however that when $v=0$ we have

$$\frac{\partial x}{\partial u}(u,0)=\frac{dx_1}{du}(u)=P(u,0).$$

More generally the v-derivatives satisfy

$$\begin{aligned}\frac{\partial^2 x}{\partial v\partial u} &= \frac{\partial^2 x}{\partial u\partial v}\\ &= \frac{\partial Q}{\partial u}\\ &= \frac{\partial P}{\partial v}.\end{aligned}$$

So it follows that

$$\frac{\partial}{\partial v}\left(\frac{\partial x}{\partial u}-P\right)=0.$$

For fixed u this shows that

$$v\mapsto\frac{\partial x}{\partial u}-P$$

is constant. Since $\left(\frac{\partial x}{\partial u}-P\right)(u,0)=0$ this implies that $\frac{\partial x}{\partial u}=P$. $\square$

This result can be extended to the more general situation as follows. When computing the derivative of $P(u,v,x(u,v))$ with respect to v it is clearly necessary to use the chain rule

$$\frac{\partial}{\partial v}(P(u,v,x(u,v)))=\frac{\partial P}{\partial v}+\frac{\partial P}{\partial x}\frac{\partial x}{\partial v}=\frac{\partial P}{\partial v}+\frac{\partial P}{\partial x}Q$$

where $\frac{\partial P}{\partial v}$ is the partial derivative of P keeping v and x fixed. Similarly

$$\frac{\partial}{\partial u}(Q(u,v,x))=\frac{\partial Q}{\partial u}+\frac{\partial Q}{\partial x}\frac{\partial x}{\partial u}=\frac{\partial Q}{\partial u}+\frac{\partial Q}{\partial x}P$$

so if a solution exists the functions P and Q must satisfy the condition

$$\frac{\partial P}{\partial v}+\frac{\partial P}{\partial x}Q=\frac{\partial Q}{\partial u}+\frac{\partial Q}{\partial x}P.$$

This is called the *integrability condition* for the system. Conversely we have

THEOREM A.5.3. *Assume $P,Q:\mathbb{R}^2\times\mathbb{R}^n\to\mathbb{R}^n$ are two smooth functions that satisfy the integrability condition*

$$\frac{\partial P}{\partial v}+\frac{\partial P}{\partial x}Q=\frac{\partial Q}{\partial u}+\frac{\partial Q}{\partial x}P.$$

The solution

$$x(u,v)=\begin{bmatrix}x^1(u,v)\\ \vdots\\ x^n(u,v)\end{bmatrix}$$

for the initial value problem

$$\begin{aligned}\frac{\partial}{\partial u}x &= P(u,v,x),\\ \frac{\partial}{\partial v}x &= Q(u,v,x),\\ x(0,0) &= x_0\end{aligned}$$

exists and is unique on some possibly small domain $(-\epsilon,\epsilon)^2$. When $|P|,|Q|\leq M+C|x|$ for constants $M,C\geq 1$ the solution exists on all of $\mathbb{R}^2$.

PROOF. We invoke theorem A.5.3 to define x_1 as the unique solution to

$$\frac{d}{du}x_1(u) = P(u,0,x_1(u)),\ x_1(0) = x_0.$$

Next use theorem A.5.3 to define the function $x(u,v)$ for a fixed u as the solution to

$$\frac{d}{dv}x(u,v) = Q(u,v,x(u,v)),\ x(u,0) = x_1(u)$$

as well as to check that $x(u,v)$ is smooth in both variables. This gives us

$$\frac{\partial x}{\partial v} = Q,\ x(0,0) = x_0.$$

Thus it remains to check that

$$\frac{\partial x}{\partial u} = P.$$

Note however that when $v = 0$ we have

$$\frac{\partial x}{\partial u}(u,0) = \frac{dx_1}{du}(u) = P(u,0,x(u,0)).$$

More generally the v-derivatives satisfy

$$\begin{aligned}
\frac{\partial^2 x}{\partial v \partial u} &= \frac{\partial^2 x}{\partial u \partial v} \\
&= \frac{\partial}{\partial u}(Q(u,v,x)) \\
&= \frac{\partial Q}{\partial u} + \frac{\partial Q}{\partial x}\frac{\partial x}{\partial u} \\
&= \frac{\partial P}{\partial v} + \frac{\partial P}{\partial x}Q - \frac{\partial Q}{\partial x}P + \frac{\partial Q}{\partial x}\frac{\partial x}{\partial u}
\end{aligned}$$

and

$$\frac{\partial}{\partial v}P(u,v,x) = \frac{\partial P}{\partial v} + \frac{\partial P}{\partial x}\frac{\partial x}{\partial v} = \frac{\partial P}{\partial v} + \frac{\partial P}{\partial x}Q.$$

So it follows that

$$\frac{\partial}{\partial v}\left(\frac{\partial x}{\partial u} - P\right) = \frac{\partial Q}{\partial x}\left(\frac{\partial x}{\partial u} - P\right).$$

For fixed u this is a differential equation in $\frac{\partial x}{\partial u} - P$. Now $\left(\frac{\partial x}{\partial u} - P\right)(u,0) = 0$ and the zero function clearly solves this equation so it follows that

$$\frac{\partial x}{\partial u} - P = 0$$

for all v. As u was arbitrary this shows the claim.

In case $|P|,|Q| \leq M + C|x|$ we can invoke theorem A.5.1 to see that x is also defined for all $(u,v) \in \mathbb{R}^2$. $\square$

REMARK A.5.4. It is not difficult to expand this result to systems of m equations if x has m variables.

The most important case for us is when $x = X$ is a row matrix of vector functions

$$X = \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix},$$

where $U_i : \Omega \to V$ are defined on some domain $\Omega \subset \mathbb{R}^n$ and the vector space V is m-dimensional. We will generally assume that for each $p \in \Omega$ the vectors

$U_1(p), ..., U_m(p)$ form a basis for V. This implies that the derivatives of these vector functions are linear combinations of this basis. Thus we obtain a system

$$\frac{\partial}{\partial u^i}\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix} = \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}[D_i],$$

where $[D_i]$ is an $m \times m$ matrix whose columns represent the coefficients of the vectors on the left hand side

$$\frac{\partial U_j}{\partial u^i} = d_{ij}^1 U_1 + \cdots + d_{ij}^m U_m = \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}\begin{bmatrix} d_{ij}^1 \\ \vdots \\ d_{ij}^m \end{bmatrix}.$$

In this way each of the entries are functions on the domain $d_{ij}^k : \Omega \to \mathbb{R}$.

The necessary integrability conditions now become

$$\frac{\partial^2}{\partial u^i \partial u^j}\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix} = \frac{\partial^2}{\partial u^j \partial u^i}\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}.$$

As

$$\begin{aligned}
\frac{\partial^2}{\partial u^i \partial u^j}\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix} &= \frac{\partial}{\partial u^i}\left(\frac{\partial}{\partial u^j}\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}\right) \\
&= \frac{\partial}{\partial u^i}\left(\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}[D_j]\right) \\
&= \left(\frac{\partial}{\partial u^i}\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}\right)[D_j] + \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}\frac{\partial}{\partial u^i}[D_j] \\
&= \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}[D_i][D_j] + \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}\frac{\partial}{\partial u^i}[D_j] \\
&= \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}\left([D_i][D_j] + \frac{\partial}{\partial u^i}[D_j]\right)
\end{aligned}$$

and $U_1, ..., U_m$ form a basis the integrability conditions become

$$[D_i][D_j] + \frac{\partial}{\partial u^i}[D_j] = [D_j][D_i] + \frac{\partial}{\partial u^j}[D_i].$$

Depending on the specific context it might be possible to calculate $[D_i]$ without first finding the partial derivatives

$$\frac{\partial U_k}{\partial u^i}$$

but we can't expect this to always happen. Note, however, that if V comes with an inner product, then the product rule implies that

$$\frac{\partial (U_k \cdot U_l)}{\partial u^i} = \frac{\partial U_k}{\partial u^i} \cdot U_l + U_k \cdot \frac{\partial U_l}{\partial u^i}.$$

This means in matrix form that

$$\begin{aligned}
& \frac{\partial}{\partial u^i}\left(\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}^t \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}\right) \\
= & \left(\frac{\partial}{\partial u^i}\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}^t\right)\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix} + \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}^t \frac{\partial}{\partial u^i}\begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix} \\
= & [D_i]^t \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}^t \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix} + \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}^t \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}[D_i].
\end{aligned}$$

Or more condensed

$$\frac{\partial}{\partial u^i}\left(X^t X\right) = [D_i]^t X^t X + X^t X [D_i].$$

If we additionally assume that $d_{ij}^k = d_{ji}^k$, then we obtain the surprising formula:

$$d_{ij}^k = g^{kl}\left(\frac{\partial g_{li}}{\partial u^j} + \frac{\partial g_{lj}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^l}\right).$$

APPENDIX B

Special Coordinate Representations

The purpose of this appendix is to collect properties and formulas that are specific to the type of parametrization that is being used. These are used in several places in the text and also appear as exercises.

B.1. Cartesian and Oblique Coordinates

Cartesian coordinates on a surface is a parametrization where

$$[\mathrm{I}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Oblique coordinates more generally come from a parametrization where

$$[\mathrm{I}] = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

for constants a, b, d with $a, d > 0$ and $ad - b^2 > 0$.

Note that the Christoffel symbols all vanish if we have a parametrization where the metric coefficients are constant. In particular, the rather nasty formula we developed in the proof of Theorema Egregium shows that the Gauss curvature vanishes. This immediately tells us that Cartesian or oblique coordinates cannot exist if the Gauss curvature doesn't vanish. When we have defined geodesic coordinates below we'll also be able to show that even abstract surfaces with zero Gauss curvature admit Cartesian coordinates.

B.2. Surfaces of Revolution

Many features of surfaces show themselves for surfaces of revolution. While this is certainly a special class of surfaces it is broad enough to give a rich family examples.

We consider

$$\mathrm{q}(t,\mu) = (r(t)\cos\mu, r(t)\sin\mu, z(t)).$$

It is often convenient to select or reparametrize (r, z) so that it is a unit speed curve. In this case we use the parametrization

$$\begin{aligned} \mathrm{q}(s,\mu) &= (r(s)\cos\mu, r(s)\sin\mu, z(s)), \\ \dot{r}^2 + \dot{z}^2 &= 1 \end{aligned}$$

We get the unit sphere by using $r = \sin s$ and $z = \cos s$.

We get a cone, cylinder or plane, by considering $r = (\alpha t + \beta)$ and $h = \gamma t$. When $\gamma = 0$ these are simply polar coordinates in the q, y plane. When $\alpha = 0$ we get a cylinder, while if both α and γ are nontrivial we get a cone. When $\alpha^2 + \gamma^2 = 1$ we have a parametrization by arclength.

The basis is given by

$$\begin{aligned}
\frac{\partial \mathrm{q}}{\partial t} &= (\dot{r}\cos\mu, \dot{r}\sin\mu, \dot{z}), \\
\frac{\partial \mathrm{q}}{\partial \mu} &= (-r\sin\mu, r\cos\mu, 0), \\
\mathrm{n} &= \frac{(-\dot{z}\cos\mu, -\dot{z}\sin\mu, \dot{r})}{\sqrt{\dot{z}^2+\dot{r}^2}}
\end{aligned}$$

and first fundamental form by

$$\begin{aligned}
g_{tt} &= \dot{z}^2+\dot{r}^2, \\
g_{\mu\mu} &= r^2 \\
g_{t\mu} &= 0
\end{aligned}$$

Note that the cylinder has the same first fundamental form as the plane if we use Cartesian coordinates in the plane. The cone also allows for Cartesian coordinates, but they are less easy to construct directly. This is not so surprising as we just saw that it took different types of coordinates for the cylinder and the plane to recognize that they admitted Cartesian coordinates. Pictorially one can put Cartesian coordinates on the cone by slicing it open along a meridian and then unfolding it to be flat. Think of unfolding a lamp shade or the Cartesian grid on a waffle cone.

Taking a surface of revolution using the arclength parameter s, we see that

$$\begin{aligned}
\frac{\partial \mathrm{n}}{\partial s} &= \frac{\partial}{\partial s}(-\dot{z}\cos\mu, -\dot{z}\sin\mu, \dot{r}) \\
&= (-\ddot{z}\cos\mu, -\ddot{z}\sin\mu, \ddot{r}) \\
\frac{\partial \mathrm{n}}{\partial \mu} &= \frac{\partial}{\partial \mu}(-\dot{z}\cos\mu, -\dot{z}\sin\mu, \dot{r}) \\
&= (\dot{z}\sin\mu, -\dot{z}\cos\mu, 0)
\end{aligned}$$

The Weingarten map is now found by expanding these two vectors. For the last equation this is simply

$$\begin{aligned}
\frac{\partial \mathrm{n}}{\partial \mu} &= (\dot{z}\sin\mu, -\dot{z}\cos\mu, 0) \\
&= -\frac{\dot{z}}{r}(-r\sin\mu, r\cos\mu, 0) \\
&= -\frac{\dot{z}}{r}\frac{\partial \mathrm{q}}{\partial \mu}
\end{aligned}$$

Thus we have

$$\begin{aligned}
L_\mu^s &= L_s^\mu = 0, \\
L_\mu^\mu &= \frac{\dot{z}}{r}
\end{aligned}$$

This leaves us with finding L_s^s. Since $\frac{\partial \mathrm{q}}{\partial s}$ is a unit vector this is simply

$$\begin{aligned}
L_s^s &= -\frac{\partial \mathrm{n}}{\partial s}\cdot\frac{\partial \mathrm{q}}{\partial s} \\
&= (\ddot{z}\cos\mu, \ddot{z}\sin\mu, -\ddot{r})\cdot(\dot{r}\cos\mu, \dot{r}\sin\mu, \dot{z}) \\
&= \ddot{z}\dot{r}-\ddot{r}\dot{z}
\end{aligned}$$

Thus

$$
\begin{aligned}
K &= (\ddot{z}\dot{r} - \ddot{r}\dot{z})\frac{\dot{z}}{r} \\
H &= \frac{\dot{z}}{r} + \ddot{z}\dot{r} - \ddot{r}\dot{z}
\end{aligned}
$$

In the case of cylinder, plane, and cone we note that K vanishes, but H only vanishes when it is a plane. This means that we have a selection of surfaces all with Cartesian coordinates with different H.

We can in general simplify the Gauss curvature by using that

$$
\begin{aligned}
1 &= \dot{r}^2 + \dot{z}^2, \\
0 &= \frac{d\left(\dot{r}^2 + \dot{z}^2\right)}{dt} = 2\dot{r}\ddot{r} + 2\dot{z}\ddot{z}.
\end{aligned}
$$

This implies

$$
\begin{aligned}
K &= \left(\ddot{r}\frac{\dot{r}^2}{\dot{z}} - \ddot{r}\dot{z}\right)\frac{\dot{z}}{r} \\
&= \frac{\ddot{r}}{r}\left(-\dot{r}^2 - \dot{z}^2\right) \\
&= -\frac{\ddot{r}}{r} \\
&= -\frac{\frac{\partial^2}{\partial s^2}\left(\sqrt{g_{rr}}\right)}{\sqrt{g_{rr}}}.
\end{aligned}
$$

This makes it particularly easy to calculate the Gauss curvature and also to construct examples with a given curvature function. It also shows that the Gauss curvature can be computed directly from the first fundamental form! For instance if we want $K = -1$, then we can just use $r(s) = \exp(-s)$ for $s > 0$ and then adjust $z(s)$ for $s \in (0, \infty)$ such that

$$1 = \dot{r}^2 + \dot{z}^2$$

If we introduce a new parameter $t = \exp(s) > 1$, then we obtain a new parametrization of the same surface

$$
\begin{aligned}
\mathrm{q}(t, \mu) &= \mathrm{q}(\ln(t), \mu) \\
&= (\exp(-\ln t)\cos\mu, \exp(-\ln t)\sin\mu, z(\ln t)) \\
&= \left(\frac{1}{t}\cos\mu, \frac{1}{t}\sin\mu, z(\ln t)\right)
\end{aligned}
$$

To find the first fundamental form of this surface we have to calculate

$$
\begin{aligned}
\frac{d}{dt}z(\ln t) &= \frac{dz}{ds}\frac{1}{t} \\
&= \sqrt{1 - \dot{r}^2}\frac{1}{t} \\
&= \sqrt{1 - (-\exp(-s))^2}\frac{1}{t} \\
&= \sqrt{1 - \exp(-2\ln t)}\frac{1}{t} \\
&= \sqrt{1 - \frac{1}{t^2}}\frac{1}{t}
\end{aligned}
$$

Thus

$$\mathrm{I} = \begin{bmatrix} \frac{1}{t^4} + \left(1 - \frac{1}{t^2}\right)\frac{1}{t^2} & 0 \\ 0 & \frac{1}{t^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{t^2} & 0 \\ 0 & \frac{1}{t^2} \end{bmatrix}$$

This is exactly what the first fundamental form for the upper half plane looks like. But the domains for the two are quite different. What we have achieved is a local representation of part of the upper half plane.

Exercises.

(1) Show that geodesics on a surface of revolution satisfy Clairaut's condition: $r \sin\phi$ is constant, where ϕ is the angle the geodesic forms with the meridians.

B.3. Monge Patches

This is more complicated than the previous case, but that is only to be expected as all surfaces admit Monge patches. We consider $\mathrm{q}(u,v) = (u, v, f(u,v))$. Thus

$$\begin{aligned} \frac{\partial\,\mathrm{q}}{\partial u} &= \left(1, 0, \frac{\partial f}{\partial u}\right), \\ \frac{\partial\,\mathrm{q}}{\partial v} &= \left(0, 1, \frac{\partial f}{\partial v}\right) \end{aligned}$$

$$\mathrm{n} = -\frac{\left(\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, -1\right)}{\sqrt{1 + \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2}}$$

$$\begin{aligned} g_{uu} &= 1 + \left(\frac{\partial f}{\partial u}\right)^2, \\ g_{vv} &= 1 + \left(\frac{\partial f}{\partial v}\right)^2, \\ g_{uv} &= \frac{\partial f}{\partial u}\frac{\partial f}{\partial v}, \\ [\mathrm{I}] &= \begin{bmatrix} 1 + \left(\frac{\partial f}{\partial u}\right)^2 & \frac{\partial f}{\partial u}\frac{\partial f}{\partial v} \\ \frac{\partial f}{\partial u}\frac{\partial f}{\partial v} & 1 + \left(\frac{\partial f}{\partial v}\right)^2 \end{bmatrix} \\ \det[\mathrm{I}] &= 1 + \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2 \end{aligned}$$

$$\frac{\partial^2\,\mathrm{q}}{\partial w_1 \partial w_2} = \left(0, 0, \frac{\partial^2 f}{\partial w_1 \partial w_2}\right)$$

So we immediately get

$$\Gamma_{w_1 w_2 w_3} = \frac{\partial^2 f}{\partial w_1 \partial w_2}\frac{\partial f}{\partial w_3}$$

$$L_{w_1 w_2} = \frac{\frac{\partial^2 f}{\partial w_1 \partial w_2}}{\sqrt{1 + \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2}}$$

The Gauss curvature is then the determinant of

$$L=\begin{bmatrix} L_u^u & L_v^u \\ L_u^v & L_v^v \end{bmatrix}=\begin{bmatrix} g^{uu} & g^{uv} \\ g^{vu} & g^{vv} \end{bmatrix}\begin{bmatrix} L_{uu} & L_{uv} \\ L_{vu} & L_{vv} \end{bmatrix}$$

$$\begin{aligned} K &= \frac{1}{\det[\mathrm{I}]}\det\begin{bmatrix} L_{uu} & L_{uv} \\ L_{vu} & L_{vv} \end{bmatrix} \\ &= \frac{\frac{\partial^2 f}{\partial u^2}\frac{\partial^2 f}{\partial v^2}-\left(\frac{\partial^2 f}{\partial u\partial v}\right)^2}{\det[\mathrm{I}]^2}. \end{aligned}$$

We note that

$$\begin{aligned} [\mathrm{I}]^{-1} &= \frac{1}{\det[\mathrm{I}]}\begin{bmatrix} 1+\left(\frac{\partial f}{\partial v}\right)^2 & -\frac{\partial f}{\partial u}\frac{\partial f}{\partial v} \\ -\frac{\partial f}{\partial u}\frac{\partial f}{\partial v} & 1+\left(\frac{\partial f}{\partial u}\right)^2 \end{bmatrix}, \\ [\mathrm{II}] &= \frac{1}{\sqrt{\det[\mathrm{I}]}}\begin{bmatrix} \frac{\partial^2 f}{\partial u^2} & \frac{\partial^2 f}{\partial u\partial v} \\ \frac{\partial^2 f}{\partial u\partial v} & \frac{\partial^2 f}{\partial v^2} \end{bmatrix} \end{aligned}$$

and the Weingarten map

$$\begin{aligned} [L] &= [\mathrm{I}]^{-1}[\mathrm{II}] \\ &= \frac{1}{(\det[\mathrm{I}])^{\frac{3}{2}}}\begin{bmatrix} 1+\left(\frac{\partial f}{\partial v}\right)^2 & -\frac{\partial f}{\partial u}\frac{\partial f}{\partial v} \\ -\frac{\partial f}{\partial u}\frac{\partial f}{\partial v} & 1+\left(\frac{\partial f}{\partial u}\right)^2 \end{bmatrix}\begin{bmatrix} \frac{\partial^2 f}{\partial u^2} & \frac{\partial^2 f}{\partial u\partial v} \\ \frac{\partial^2 f}{\partial u\partial v} & \frac{\partial^2 f}{\partial v^2} \end{bmatrix} \\ &= \frac{1}{(\det[\mathrm{I}])^{\frac{3}{2}}}\begin{bmatrix} \left(1+\left(\frac{\partial f}{\partial v}\right)^2\right)\frac{\partial^2 f}{\partial u^2}-\frac{\partial f}{\partial u}\frac{\partial f}{\partial v}\frac{\partial^2 f}{\partial u\partial v} & \left(1+\left(\frac{\partial f}{\partial v}\right)^2\right)\frac{\partial^2 f}{\partial u\partial v}-\frac{\partial f}{\partial u}\frac{\partial f}{\partial v}\frac{\partial^2 f}{\partial v^2} \\ \left(1+\left(\frac{\partial f}{\partial u}\right)^2\right)\frac{\partial^2 f}{\partial u\partial v}-\frac{\partial f}{\partial u}\frac{\partial f}{\partial v}\frac{\partial^2 f}{\partial u^2} & \left(1+\left(\frac{\partial f}{\partial u}\right)^2\right)\frac{\partial^2 f}{\partial v^2}-\frac{\partial f}{\partial u}\frac{\partial f}{\partial v}\frac{\partial^2 f}{\partial u\partial v} \end{bmatrix} \end{aligned}$$

This gives us a general example where the Weingarten map might not be a symmetric matrix.

B.4. Surfaces Given by an Equation

This is again very general. Note that any Monge patch $(u,v,f(u,v))$ also yields a function $F(x,y,z)=z-f(x,y)$ such that the zero level of F is precisely the Monge patch. This case is also complicated by the fact that while the normal is easy to find, it is proportional to the gradient of F, we don't have a basis for the tangent space without resorting to a Monge patch. This is troublesome, but not insurmountable as we can solve for the derivatives of F. Assume that near some point p we know $\frac{\partial F}{\partial z}\neq 0$, then we can use x,y as coordinates. Our coordinates vector fields look like

$$\begin{aligned} \frac{\partial \mathrm{q}}{\partial u} &= \left(1,0,\frac{\partial f}{\partial u}\right), \\ \frac{\partial \mathrm{q}}{\partial v} &= \left(0,1,\frac{\partial f}{\partial v}\right) \end{aligned}$$

where

$$\frac{\partial f}{\partial w}=-\frac{\frac{\partial F}{\partial w}}{\frac{\partial F}{\partial z}}$$

Thus we actually get some explicit formulas

$$\begin{aligned}\frac{\partial \mathrm{q}}{\partial u} &= \left(1,0,-\frac{\frac{\partial F}{\partial u}}{\frac{\partial F}{\partial z}}\right),\\ \frac{\partial \mathrm{q}}{\partial v} &= \left(0,1,-\frac{\frac{\partial F}{\partial v}}{\frac{\partial F}{\partial z}}\right).\end{aligned}$$

We can however describe the second fundamental form without resorting to coordinates. We consider a surface given by an equation

$$F(x,y,z)=C.$$

The normal can be calculated directly as

$$\mathrm{n}=\frac{\nabla F}{|\nabla F|}.$$

This shows first of all that we have a simple equation defining the tangent space at each point p

$$T_pM=\left\{Y\in\mathbb{R}^3 \mid Y\cdot\nabla F(p)=0\right\}.$$

Next we make the claim that

$$\begin{aligned}\mathrm{II}(X,Y) &= -\frac{1}{|\nabla F|}\,\mathrm{I}(D_X\nabla F,Y)\\ &= -\frac{1}{|\nabla F|}Y\cdot D_X\nabla F,\end{aligned}$$

where D_X is the directional derivative. We can only evaluate II on tangent vectors, but $Y\cdot D_X\nabla F$ clearly makes sense for all vectors. This has the advantage that we can even use Cartesian coordinates in $\mathbb{R}^3$ for our tangent vectors. First we show that

$$L(X)=-D_X\,\mathrm{n}.$$

Select a parametrization $\mathrm{q}(u,v)$ such that

$$\frac{\frac{\partial \mathrm{q}}{\partial u}\times\frac{\partial \mathrm{q}}{\partial v}}{\left|\frac{\partial \mathrm{q}}{\partial u}\times\frac{\partial \mathrm{q}}{\partial v}\right|}=\frac{\nabla F}{|\nabla F|}.$$

The Weingarten equations then tell us that

$$L\left(\frac{\partial \mathrm{q}}{\partial w}\right)=-\frac{\partial \mathrm{n}}{\partial w}=-D_{\frac{\partial \mathrm{q}}{\partial w}}\,\mathrm{n}.$$

We can now return to the second fundamental form. Let Y be another tangent vector then, $Y\cdot\nabla F=0$ so

$$\begin{aligned}-\mathrm{II}(X,Y) &= -\mathrm{I}(L(X),Y)\\ &= Y\cdot D_X\,\mathrm{n}\\ &= Y\cdot\left(D_X\frac{1}{|\nabla F|}\right)\nabla F+Y\cdot\frac{1}{|\nabla F|}D_X\nabla F\\ &= Y\cdot\frac{1}{|\nabla F|}D_X\nabla F.\end{aligned}$$

Note that even when X is tangent it does not necessarily follow that $D_X\nabla F$ is also tangent to the surface.

In case $\frac{\partial F}{\partial z} \neq 0$ we get a relatively simple orthogonal basis for the tangent space. In case $\frac{\partial F}{\partial x} = \frac{\partial F}{\partial y} = 0$ we can simply use

$$X = (1,0,0)\,,\ Y = (0,1,0)$$

otherwise we obtain an orthogonal basis by using

$$\begin{aligned} X &= \left(-\frac{\partial F}{\partial y}, \frac{\partial F}{\partial x}, 0\right), \\ Y &= \left(\frac{\partial F}{\partial z}\frac{\partial F}{\partial x}, \frac{\partial F}{\partial z}\frac{\partial F}{\partial y}, -\left(\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2\right)\right). \end{aligned}$$

With that basis the Weingarten map can then be calculated as

$$\begin{aligned} [L] &= [\mathrm{I}]^{-1}[\mathrm{II}] \\ &= \begin{bmatrix} |X|^{-2} & 0 \\ 0 & |Y|^{-2} \end{bmatrix} \begin{bmatrix} \mathrm{II}(X,X) & \mathrm{II}(X,Y) \\ \mathrm{II}(X,Y) & \mathrm{II}(Y,Y) \end{bmatrix}. \end{aligned}$$

To calculate the second fundamental form we use that

$$\begin{bmatrix} \frac{\partial \nabla F}{\partial x} & \frac{\partial \nabla F}{\partial y} & \frac{\partial \nabla F}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial x \partial z} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial y \partial z} \\ \frac{\partial^2 F}{\partial z \partial x} & \frac{\partial^2 F}{\partial z \partial y} & \frac{\partial^2 F}{\partial z^2} \end{bmatrix}.$$

So

$$\mathrm{II}(X,X) = \frac{1}{|\nabla F|}\begin{bmatrix} -\frac{\partial F}{\partial y} & \frac{\partial F}{\partial x} & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial x \partial z} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial y \partial z} \\ \frac{\partial^2 F}{\partial z \partial x} & \frac{\partial^2 F}{\partial z \partial y} & \frac{\partial^2 F}{\partial z^2} \end{bmatrix} \begin{bmatrix} -\frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial x} \\ 0 \end{bmatrix},$$

$$\mathrm{II}(X,Y) = \frac{1}{|\nabla F|}\begin{bmatrix} -\frac{\partial F}{\partial y} & \frac{\partial F}{\partial x} & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial x \partial z} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial y \partial z} \\ \frac{\partial^2 F}{\partial z \partial x} & \frac{\partial^2 F}{\partial z \partial y} & \frac{\partial^2 F}{\partial z^2} \end{bmatrix} \begin{bmatrix} \frac{\partial F}{\partial z}\frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial z}\frac{\partial F}{\partial y} \\ -\left(\frac{\partial F}{\partial x}\right)^2 - \left(\frac{\partial F}{\partial y}\right)^2 \end{bmatrix},$$

$$\mathrm{II}(Y,Y) = \frac{1}{|\nabla F|}\begin{bmatrix} \frac{\partial F}{\partial z}\frac{\partial F}{\partial x} & \frac{\partial F}{\partial z}\frac{\partial F}{\partial y} & -\left(\frac{\partial F}{\partial x}\right)^2 - \left(\frac{\partial F}{\partial y}\right)^2 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial x \partial z} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial y \partial z} \\ \frac{\partial^2 F}{\partial z \partial x} & \frac{\partial^2 F}{\partial z \partial y} & \frac{\partial^2 F}{\partial z^2} \end{bmatrix} \begin{bmatrix} \frac{\partial F}{\partial z}\frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial z}\frac{\partial F}{\partial y} \\ -\left(\frac{\partial F}{\partial x}\right)^2 - \left(\frac{\partial F}{\partial y}\right)^2 \end{bmatrix}.$$

Exercises.

(1) If q is a curve, then it is a curve on $F = C$ if $\mathrm{q}(0)$ lies on the surface and $\dot{\mathrm{q}} \cdot \nabla F$ vanishes. If q is regular and a curve on $F = C$, then it can be reparametrized to be a geodesic if and only if the triple product $\det[\nabla F, \dot{\mathrm{q}}, \ddot{\mathrm{q}}] = 0$.

B.5. Geodesic Coordinates

This is a parametrization having a first fundamental form that looks like:

$$\mathrm{I} = \begin{bmatrix} 1 & 0 \\ 0 & g_{vv} \end{bmatrix}$$

This is as with surfaces of revolution, but now g_{vv} can depend on both u and v. Using a central v curve, we let the u curves be unit speed geodesics orthogonal to the fixed v curve. They are also often call Fermi coordinates after the famous physicist and seem to have been used in his thesis on general relativity. They were however also used by Gauss. These coordinates will be used time and again to simplify calculations in the proofs of several theorems. The v-curves are well defined as the curves that appear when u is constant. At $u=0$ the u and v curves are perpendicular by construction, so by continuity they can't be tangent as long as u is sufficiently small.

Exercises.

(1) Consider a parametrization $\mathrm{q}(s,t)$ where the s-curves are unit speed geodesics and $\frac{\partial \mathrm{q}}{\partial s}(s,0) \perp \frac{\partial \mathrm{q}}{\partial t}(s,0)$. Show that

$$\frac{\partial \mathrm{q}}{\partial s}(s,t) \perp \frac{\partial \mathrm{q}}{\partial t}(s,t)$$

and conclude that such a parametrization defines geodesic coordinates.

(2) Show that for geodesic coordinates:

$$\begin{aligned}
\Gamma_{uuu} &= 0, \\
\Gamma_{uvu} &= 0 = \Gamma_{vuu}, \\
\Gamma_{vvu} &= -\frac{1}{2}\frac{\partial g_{vv}}{\partial u}, \\
\Gamma_{vvv} &= \frac{1}{2}\frac{\partial g_{vv}}{\partial v}, \\
\Gamma_{uvv} &= \frac{1}{2}\frac{\partial g_{vv}}{\partial u} = \Gamma_{vuv}, \\
\Gamma_{uuv} &= 0, \\
\Gamma_{ij}^{u} &= \Gamma_{iju}, \\
\Gamma_{ij}^{v} &= \frac{1}{g_{vv}}\Gamma_{ijv},
\end{aligned}$$

and

$$K = -\frac{\partial_u^2\sqrt{g_{vv}}}{\sqrt{g_{vv}}} = -\frac{1}{2}\left(\frac{\partial_u^2 g_{vv}}{g_{vv}} - \left(\frac{\partial_u g_{vv}}{g_{vv}}\right)^2\right).$$

B.6. Chebyshev Nets

These correspond to a parametrization where the first fundamental form looks like:

$$\begin{aligned}
\mathrm{I} &= \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{bmatrix},
\end{aligned}$$

The idea is to have a material such as a fishnet where the fibers are not changed in length or stretched, but are allowed to change their mutual angles.

Note that such parametrizations are characterized as having unit speed parameter curves.

Exercises.

(1) ???Show that any surface locally admits Chebyshev nets. Hint: Fix a point $p = \mathrm{q}(u_0, v_0)$ for a given parametrization and define new parameters

$$\begin{aligned} s(u,v) &= \int_{u_0}^{u} \sqrt{g_{uu}(x,v)}dx \\ t(u,v) &= \int_{v_0}^{v} \sqrt{g_{vv}(u,y)}dy \end{aligned}$$

Show that $\frac{\partial s}{\partial v}(u_0, v_0) = 0 = \frac{\partial t}{\partial u}(u_0, v_0)$ and conclude that (s,t) defines a new parametrization that creates a Chebyshev net.

(2) Show that Chebyshev nets $\mathrm{q}(u,v)$ satisfy the following properties

$$\frac{\partial^2 \mathrm{q}}{\partial u \partial v} \perp T_p M,$$

$$\begin{aligned} \Gamma_{uvw} &= \Gamma_{uuu} = \Gamma_{vvv} = 0, \\ \Gamma_{uuv} &= -\frac{\partial \theta}{\partial u} \sin\theta, \\ \Gamma_{vvu} &= -\frac{\partial \theta}{\partial v} \sin\theta, \\ \frac{\partial^2 \theta}{\partial u \partial v} &= -K \sin\theta. \end{aligned}$$

(3) Show that the geodesic curvature κ_g of the u-coordinate curves in a Chebyshev net satisfy

$$\kappa_g = -\frac{\partial \theta}{\partial u}.$$

(4) (Hazzidakis) Show that $\sqrt{\det[\mathrm{I}]} = \sin\theta$, and that integrating the Gauss curvature over a coordinate rectangle yields:

$$-\int_{[a,b]\times[c,d]} K \sin\theta du dv = 2\pi - \alpha_1 - \alpha_2 - \alpha_3 - \alpha_4$$

where the angles α_i are the interior angles.

B.7. Isothermal Coordinates

These are also more generally known as *conformally flat coordinates* and have a first fundamental form that looks like:

$$\mathrm{I} = \begin{bmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{bmatrix}$$

The proof that these always exist is called the *local uniformization theorem*. It is not a simple result, but the importance of these types of coordinates in the development of both classical and modern surface theory cannot be understated. There is also a global result which we will mention at a later point. Gauss was the first to work with such coordinates, and Riemann also heavily depended on their use. They have the properties that

$$\begin{aligned}
\Gamma_{uuu} &= \frac{\partial \log \lambda}{\partial u} \\
\Gamma_{uvu} &= \frac{\partial \log \lambda}{\partial v} = \Gamma_{vuu} \\
\Gamma_{vvv} &= \frac{\partial \log \lambda}{\partial v} \\
\Gamma_{uvv} &= \frac{\partial \log \lambda}{\partial u} = \Gamma_{vuv} \\
\Gamma_{uuv} &= -\frac{\partial \log \lambda}{\partial v} \\
\Gamma_{vvu} &= -\frac{\partial \log \lambda}{\partial u}, \\
\Gamma_{w_1 w_2}^{w_3} &= \frac{1}{\lambda^2} \Gamma_{w_1 w_2 w_3},
\end{aligned}$$

$$K = -\frac{1}{\lambda^2}\left(\frac{\partial^2 \log \lambda}{\partial u^2} + \frac{\partial^2 \log \lambda}{\partial v^2}\right).$$

Using complex analysis one obtains isothermal parametrizations as follows. For a parametrized surface define $w = u + \mathrm{i}v$ and complex functions $\phi_i(w)$, $i = 1, 2, 3$

$$\phi = (\phi_1, \phi_2, \phi_3) = \frac{\partial \mathrm{q}}{\partial w} = \frac{\partial \mathrm{q}}{\partial u} - \mathrm{i}\frac{\partial \mathrm{q}}{\partial v}.$$

In other words

$$\mathrm{Re}\phi = \frac{\partial \mathrm{q}}{\partial u}, \ \mathrm{Im}\phi = -\frac{\partial \mathrm{q}}{\partial v}.$$

First observe that

$$\frac{\partial \phi}{\partial \bar{w}} = \frac{\partial \phi}{\partial u} + \mathrm{i}\frac{\partial \phi}{\partial v} = \frac{\partial^2 \mathrm{q}}{\partial \bar{w} \partial w} = \frac{\partial^2 \mathrm{q}}{\partial u^2} + \frac{\partial^2 \mathrm{q}}{\partial v^2}.$$

So ϕ is holomorphic if and only if q is harmonic.

Next note that

$$\phi_1^2 + \phi_2^2 + \phi_3^2 = \phi \cdot \phi = \left(\frac{\partial \mathrm{q}}{\partial u} - \mathrm{i}\frac{\partial \mathrm{q}}{\partial v}\right) \cdot \left(\frac{\partial \mathrm{q}}{\partial u} - \mathrm{i}\frac{\partial \mathrm{q}}{\partial v}\right) = g_{uu} - g_{vv} - 2\mathrm{i}g_{uv}.$$

So we obtain an isothermal parametrization when $\phi_1^2 + \phi_2^2 + \phi_3^2 = 0$. The conformal factor can then be calculated by noting that

$$\phi \cdot \bar{\phi} = |\phi_1|^2 + |\phi_2|^2 + |\phi_3|^2 = \left(\frac{\partial \mathrm{q}}{\partial u} - \mathrm{i}\frac{\partial \mathrm{q}}{\partial v}\right) \cdot \left(\frac{\partial \mathrm{q}}{\partial u} + \mathrm{i}\frac{\partial \mathrm{q}}{\partial v}\right) = g_{uu} + g_{vv}.$$

Conversely, starting with holomorphic functions $\phi_i(w)$, $i = 1, 2, 3$ such that $\phi_1^2 + \phi_2^2 + \phi_3^2 = 0$ we can define their antiderivatives and construct a surface by

$$\mathrm{q}(w) = \left(2\mathrm{Re}\left(\int \phi_1(w)\,dw\right), 2\mathrm{Re}\left(\int \phi_2(w)\,dw\right), 2\mathrm{Re}\left(\int \phi_3(w)\,dw\right)\right)$$

since $\frac{\partial \mathrm{Re} f}{\partial w} = \frac{1}{2}\left(\frac{\partial f}{\partial w} + \frac{\partial \bar{f}}{\partial w}\right) = \frac{1}{2}\frac{\partial f}{\partial w}$, when $f(w)$ is holomorphic and consequently $\bar{f}$ antiholomorphic.

A minimal surface evidently always has such a parametrization: First use that the Gauss map is conformal to conclude that it has an isothermal parametrization. This must be harmonic which in turn shows that ϕ is holomorphic since.

Examples:

Catenoid: $\phi_1 = \sinh w$, $\phi_2 = -\mathrm{i}\cosh w$, $\phi_3 = 1$.
Helicoid: $\phi_1 = \mathrm{i}\sinh w$, $\phi_2 = \cosh w$, $\phi_3 = \mathrm{i}$.
Enneper: $\phi_1 = 1 - w^2$, $\phi_2 = \mathrm{i}\left(1 + w^2\right)$, $\phi_3 = 2w$.
Scherk: $\phi_1 = \frac{2}{1+w^2}$, $\phi_2 = \frac{2\mathrm{i}}{1-w^2}$, $\phi_3 = \frac{4w}{1-w^4}$.
Catalan: $\phi_1 = 1 - \cosh(-\mathrm{i}w)$, $\phi_2 = \mathrm{i}\sinh(-\mathrm{i}w)$, $\phi_3 = 2\sinh\left(-\frac{\mathrm{i}w}{2}\right)$.
Can always use $\phi_1 = F\left(1 - G^2\right)$, $\phi_2 = \mathrm{i}F\left(1 + G^2\right)$, and $\phi_3 = 2FG$ with $2F = \phi_1 - \mathrm{i}\phi_2$.

Exercises.

(1) A particularly nice special case occurs when
$$\lambda^2(u, v) = U^2(u) + V^2(v)$$
These types of metrics are called *Liouville metrics*. Compute their Christoffel symbols, Gauss curvature, and show that when geodesics are written as $v(u)$ or $u(v)$ they they solve a separable differential equation. Show also that the geodesics have the property that
$$U^2 \sin^2\theta - V^2 \cos^2\theta$$
is constant, where θ is the angle the geodesic forms with the u curves.

(2) Show that when
$$\lambda = \frac{1}{a\left(u^2 + v^2\right) + b_u u + b_v v + c}$$
we obtain a metric with constant Gauss curvature
$$K = 4ac - b_u^2 - b_v^2$$
It can be shown that no other choices for λ will yield constant curvature.

Bibliography

[1] D. Hilbert and S. Cohn-Vossen, *Geometry and the Imagination*, New York: Chelsea, 1990.
[2] D.J. Struik, *Lectures on Classical Differential Geometry*, New York: Dover, 1988.