The Gauss-Bonnet Theorem

Bonnie Yang

June 10, 2023

1 Introduction

The Gauss-Bonnet theorem is a crowning result of surface theory that gives a fundamental connection between geometry and topology. Roughly speaking, geometry refers to the "local" properties—lengths, angles, curvature—of some fixed object, while topology seeks to identify the "global" properties that are unchanged by a continuous deformation, such as stretching or twisting. The theorem formalizes an intuitive idea: continuous changes to curvature on one region of a surface will be balanced out elsewhere, so the *total* curvature of the surface stays the same.

Explicitly, the Gauss-Bonnet theorem says that a surface's total curvature, defined using its local Gaussian curvature, is directly proportional to the number of holes in the surface, which comes from an invariant quantity called its Euler characteristic. The Euler characteristic is a way of classifying which surfaces can be continuously deformed into one another. For example, a sphere is topologically identical to a cube, but not the torus. Even though the sphere has uniform curvature everywhere and the cube has completely flat sides, according to the Gauss-Bonnet theorem, both objects will have total curvature 4π.

Our goal is to show

$$\int_{\mathcal{S}} K dA = 2\pi\chi(\mathcal{S}),$$

where $\mathcal{S}$ is a closed surface in $\mathbb{R}^3$, K is the Gaussian curvature, dA is the area element, and $\chi(\mathcal{S})$ is the Euler characteristic. The proof itself is delightfully systematic: we first find the total curvature of a curve on a plane, extend that result to curves on three-dimensional surfaces, extend *that* result to "polygons" on surfaces, and finally the entire surface.

In Section 2, we prove Hopf's Umlaufsatz for the total curvature of a simple closed curve in $\mathbb{R}^2$. Sections 3, 4, and 5 introduce concepts from differential geometry to define Gaussian curvature. In Section 6, we prove the local Gauss-Bonnet theorem for the total curvature of a surface polygon. At last, in Section 7, we prove the global Gauss-Bonnet theorem for compact surfaces by covering the surface with polygons and applying the local Gauss-Bonnet theorem to each one.

This paper assumes knowledge of multivariable calculus.

2 Plane curves and Hopf's Umlaufsatz

Hopf's Umlaufsatz [1] asserts that the total signed curvature of any simple closed curve in $\mathbb{R}^2$ is equal to $\pm 2\pi$, with sign depending on the curve's orientation. Although the theorem is about the curvature of a line and not a region with area, the Umlaufsatz does much of the heavy lifting for our later proof in $\mathbb{R}^3$. We begin with some preliminary theory of paths and curves.

Definition 2.1. A *(parametric) path* in $\mathbb{R}^n$ is a continuous function $\gamma : I \to \mathbb{R}^n$, where I is any interval of $\mathbb{R}$. The image of a path is called a *parametrized curve* in $\mathbb{R}^n$.

If γ is differentiable, the differential[2] $\dot{\gamma}(t)$ is called the *tangent vector* of γ at the point $\gamma(t)$. We say γ is *regular* if $\dot{\gamma}(t)$ is nonzero for all $t \in U$.

[1] From German *umlauf* (rotation) and *satz* (theorem). Sometimes translatd to "rotation angle theorem."

[2] The "overdot" notation is conventially used for a derivative taken with respect to time (i.e., $\dot{\gamma} = d\gamma/dt$ and $\ddot{\gamma} = d^2\gamma/dt$).

Remark. A particular curve can be the image of infinitely many paths. To see this, suppose γ_1 and γ_2 are two paths defined on the intervals I_1 and I_2, respectively. Since these are intervals of $\mathbb{R}$, we can define a bijection $\phi : I_1 \to I_2$ between their domains. Then if γ_1 and γ_2 are both injective with the same image curve, we can always *reparametrize* one path as the other by a composition $\gamma_2 = \gamma_1 \circ \phi$.

In practice, the terms *path* and *curve* are used interchangably to mean either a continuous function $\gamma : [a, b] \to \mathbb{R}^n$ or its image. The correct interpretation should be clear from context.

Unless otherwise specified, all curves discussed in this paper are assumed to be regular and *smooth*, meaning there exist continuous partial derivatives of all orders.

Definition 2.2. If $\gamma : [a, b] \to \mathbb{R}^n$ is a parametrized curve, then for any $a \leqslant t \leqslant b$, the *arc length* of γ from a to t is given by the function

$$s(t) = \int_a^t \|\dot{\gamma}_u\| du.$$

A regular curve γ is *unit-speed* if for all t, we have $\|\dot{\gamma}(t)\| = 1$. In this case, the arc length is $s(t) = t$, so γ is also said to be an *arc length parametrization.*

Remark. Every curve can be reparametrized to unit speed.

Definition 2.3. If γ is a unit-speed curve with unit tangent vector $\dot{\gamma}$, then we define the *unit normal vector* to γ at the point $\gamma(s)$ by

$$\mathbf{n} = \frac{\ddot{\gamma}(s)}{\|\ddot{\gamma}(s)\|},$$

where $\ddot{\gamma}$ represents the second derivative[3].

Remark. To see why $\mathbf{n}$ is indeed normal to $\dot{\gamma}(s)$, recall that $\langle \dot{\gamma}, \dot{\gamma} \rangle = 1$, so we can differentiate to obtain $\langle \ddot{\gamma}, \dot{\gamma} \rangle + \langle \dot{\gamma}, \ddot{\gamma} \rangle = 0$. Thus, the vectors $\dot{\gamma}$ and $\ddot{\gamma}$ are orthogonal.

For plane curves, which have two choices of unit normal for each tangent vector $\dot{\gamma}(s)$, we fix the *signed unit normal* $\mathbf{n}$ to be the one obtained by rotating $\dot{\gamma}$ counterclockwise by $\pi/2$. When γ is unit-speed, the *signed curvature* at any point is the scalar κ which satisfies

$$\ddot{\gamma} = \kappa \mathbf{n}.$$

This formulation of signed curvature is strictly local, since it arises from the behavior of a curve at a specific point: if $\gamma(s)$ is a point on a unit-speed curve, then we simply take $\kappa = \|\ddot{\gamma}(s)\|$. To see how Hopf's Umlaufsatz relates local curvature to a curve's topology, we must first get a sense of what global properties a curve has.

We start with a geometric interpretation of the tangent vector for plane curves. When $\gamma : [a, b] \to \mathbb{R}^2$ is unit-speed, the direction of each vector $\dot{\gamma}(s)$ is determined by the angle $\theta(s)$ for which $\dot{\gamma}(s) = e^{i\theta(s)}$. It is straightforward to show that our choice of $\theta(s)$ is smooth: briefly, if $\dot{\gamma}$ is indeed defined on the complex unit circle, then the chain rule implies

$$\ddot{\gamma}(s) = i\dot{\theta}(s) \cdot e^{i\theta(s)} = \dot{\theta}(s);$$

one can recover $\dot{\theta}$ as the scalar in this expression, and then the continuous map θ by taking an antiderivative.

Definition 2.4. Let $f : [a, b] \to S^1$ be any path in the unit circle, and let $p : \mathbb{R} \to S^1$ be defined by $p(t) = e^{it}$. An *angle function* for f is a smooth map $\theta : [a, b] \to \mathbb{R}$ which satisfies

$$f(s) = p \circ \theta = e^{i\theta(s)}.$$

If $f = \dot{\gamma}$ for some unit-speed plane curve γ, then θ is called a *tangent angle function* for γ.

Proposition 2.1. *Given a unit-speed curve $\gamma : [a, b] \to \mathbb{R}^2$ with a tangent angle function θ, the signed curvature of γ is defined by*

$$\kappa = \dot{\theta},$$

the rate at which the tangent vector $\dot{\gamma}$ rotates.

[3]When possible, we write s for the parameter to emphasize when γ is unit-speed.

Proof. Using the polar coordinate map for the circle, the tangent vector is given by $\dot{\gamma} = (\cos\theta, \sin\theta)$, and the fixed unit normal is $\mathbf{n} = (-\sin\theta, \cos\theta)$. Choosing $\kappa = \dot{\theta}$ and differentiating both sides, we see that

$$\ddot{\gamma} = \dot{\theta}(-\sin\theta, \cos\theta) = \kappa\mathbf{n}$$

which coincides with our original definition of signed curvature. $\square$

The upshot of this discussion is that we can express the tangent $\dot{\gamma}$ of any plane curve γ as a path in the unit circle! This is useful because every path in S^1 has a fixed *degree*, which counts how many times the curve "goes around" the circle counterclockwise. Defining a path $\dot{\gamma} : [a, b] \to S^1$ this way allows us to treat the degree of the tangent as a topological property of γ itself. Later, we will see that the proof of the Umlaufsatz is essentially an argument about the degree of $\dot{\gamma}$ when γ in a specific case, when γ is a simple closed curve.

A tangent angle function θ takes each point $\dot{\gamma}(s)$ on the circle to a number on the "unfolded" real line, which we call a *lift* of $\dot{\gamma}$ to $\mathbb{R}$. Notice that a tangent curve $\dot{\gamma}$ which winds around the circle n times will have tangent angle function increase by n. Using the fact that each corresponding angle $\theta(s)$ is unique up to an integer multiple of 2π, we can recover the degree of $\dot{\gamma}$ from this unfolding process.

$$\begin{tikzcd}
[a,b] \arrow[r, "\dot{\gamma}"] \arrow[d, "\theta"'] & S^1 \\
\mathbb{R} \arrow[ur, "p"'] &
\end{tikzcd}$$

Proposition 2.2. *Let $f : [a, b] \to S^1$ be a path in the circle and $\theta, \phi : [a, b] \to \mathbb{R}$ be any two angle functions for f. Then we have*

$$\theta(b) - \theta(a) = \phi(b) - \phi(a).$$

Equivalently, for any point $s_0 \in [a, b]$, there exists a unique *angle function such that $f(s_0) = e^{i\theta(s_0)}$.*

Proof. We will show that $e^{i\theta(s)}$ and $e^{i\phi(s)}$ to agree, the values $\theta(s)$ and $\phi(s)$ must differ by an integer multiple of 2π, and by continuity, the integer must be the same for all s.

First, since both expressions for $\dot{\gamma}(s)$ are points in S^1, the angles $\theta(s)$ and $\gamma(s)$ clearly differ by full rotations about unit circle. Formally, this means there exists some integer $n(s)$ such that for all $s \in [a, b]$, we have

$$\phi(s) - \theta(s) = 2\pi n(s).$$

Because θ and ϕ are continuous functions, n is continuous on the domain $[a, b]$ as well, and we apply the intermediate value theorem to conclude that n is a constant that does not depend on s. Thus, the integer term cancels, and we see

$$\phi(b) - \phi(a) = \theta(b) + 2\pi n(s) - \theta(a) - 2\pi n(s) = \theta(b) - \theta(a)$$

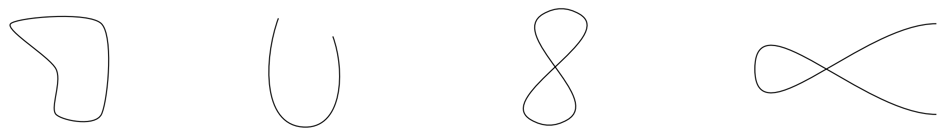
as desired. $\square$

Definition 2.5. Let $f : [a, b] \to S^1$ be a path in the circle and let $\theta : [a, b] \to \mathbb{R}$ be a tangent angle function of γ. The *degree* of f is defined as

$$\frac{\theta(b) - \theta(a)}{2\pi}.$$

If $\gamma : [a, b] \to \mathbb{R}^2$ is a unit-speed plane curve, then the degree of its tangent $\dot{\gamma}$ is called the *rotation index* of γ and denoted $I(\gamma)$.

Definition 2.6. Given a compact interval $[a, b] \subset \mathbb{R}$, we say $\gamma : [a, b] \to \mathbb{R}^n$ is a *closed curve* of period $b - a$ if $\gamma(a) = \gamma(b)$. If γ is injective on the open interval (a, b), then γ is called *simple*.



Simple and closed   Simple, not closed   Not simple, closed   Not simple and not closed

The Jordan curve theorem from topology tells us that any simple closed curve on a plane has an "interior" and an "exterior." Precisely, if γ is a simple closed curve in $\mathbb{R}^2$, then the *complement* of its image is the union of two subsets of $\mathbb{R}^2$, denoted $\text{int}(\gamma)$ and $\text{ext}(\gamma)$, which satisfy the following:

- $\text{int}(\gamma)$ and $\text{ext}(\gamma)$ are disjoint, so $\text{int}(\gamma) \cap \text{ext}(\gamma) = \varnothing$.
- $\text{int}(\gamma)$ is bounded and $\text{ext}(\gamma)$ is unbounded.
- Both $\text{int}(\gamma)$ and $\text{ext}(\gamma)$ are *connected*, so any two points in the same subset can be joined by a curve contained entirely in that subset.

This gives us a way to distinguish between two possible orientations of γ using geometry: we say γ is *positively-oriented* if the signed unit normal $\mathbf{n}$ points into $\text{int}(\gamma)$ at every point in the curve.

Now, when we claim that a property like the rotation index is global, we mean that it is invariant under a "continuous deformation." The following definition formalizes this notion for closed curves in $\mathbb{R}^2$.

Definition 2.7. An *isotopy* of closed plane curves of period ℓ is a family of curves $\gamma_t : \mathbb{R} \to \mathbb{R}^2$ such that

(i) For all $0 \leqslant t \leqslant 1$, the map $h : \mathbb{R} \times [0,1] \to \mathbb{R}^2$ defined by $h(s,t) = \gamma_t(s)$ is also a regular closed plane curve of period ℓ;

(ii) We have $h(s,0) = \gamma_0(s)$ and $h(s,1) = \gamma_1(s)$.

If such a family exists, we say that γ_0 is *isotopic* to γ_1.

Example. The reparametrizations discussed at the beginning of this section are given by isotopies of the form $h(s,t) = \gamma(s + s_0 t)$, where $s_0 \in \mathbb{R}$ is a constant. A *translation* of the curve is an isotopy of the form $h(s,t) = \gamma(s) + t\vec{x}$ for some point $\vec{x} \in \mathbb{R}^2$.

Lemma 2.3. *If γ_0 and γ_1 are closed plane curves connected by an isotopy, then $I(\gamma_0) = I(\gamma_1)$.*

Proof. Similar to the proof of Proposition 2.2, we show that the rotation index is an integer constant by continuity. First, notice that the rotation index for a closed curve is indeed an integer. Now let h be an isotopy from γ_0 to γ_1, and fix $\gamma_t(s) = h(s,t)$. Then the map from s to $I(\gamma_s)$ given by the equation in Definition 2.5 is a continuous function $[0,1] \to \mathbb{Z}$, so we apply the intermediate value theorem to conclude that $I(\gamma_s)$ is constant. $\square$

Theorem 2.4 (Hopf's Umlaufsatz)**.** *Let $\gamma : [a,b] \to \mathbb{R}$ be a unit-speed, simple closed curve on a plane. Then the total signed curvature is given by*

$$\int_\gamma \kappa ds = \pm 2\pi.$$

As promised, this reduces to a claim about the rotation index! Since $\kappa = \dot{\theta}$ for any curve by Proposition 2.1, the total signed curvature can be computed as

$$\int_\gamma \kappa ds = \int_a^b \dot{\theta}(s) ds = \theta(b) - \theta(a) = 2\pi I(\gamma).$$

Thus, the point of the Umlaufsatz is that for *simple* closed curves, we have $I(\gamma) = \pm 1$.

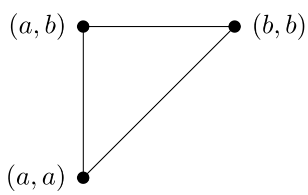
Proof of Theorem 2.4. Our strategy is to replace $\dot{\gamma} : [a,b] \to S^1$ with a map to the circle whose degree is easy to compute, the *secant line* between two points on the curve. Both the secant line and its angle function take two parameter inputs. When the two parameters are equal, the secant is the tangent line, and the secant angle function is continuously extended to the tangent angle function of γ at a single point. The domain of this secant map can be interpreted geometrically as a triangle formed by the points (a,a), (a,b), and (b,b), and the restriction of the secant map to the diagonal is exactly the tangent map $\dot{\gamma}$. A continuous deformation of the diagonal to the other two sides of the triangle preserves the endpoints (a,a) and (b,b), so the total change of the secant angle function is the same along this deformed path. Then to find $I(\gamma)$, it suffices to compute the degree of the secant map coming from the non-diagonal sides.

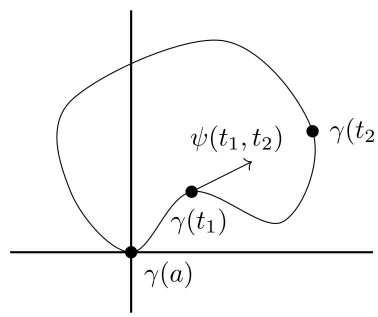
We begin by assuming, without loss of generality, that $\gamma(a)$ is the lowest point on the curve and located at the origin $(0,0)$. Since the projection of γ to its y-coordinate is continuous on $[a,b]$, we know there exists a $t_0 \in [a,b]$ such that the y-coordinate of $\gamma(t_0)$ is minimal. The assumptions follow because the rotation index is invariant under isotopy, including reparametrization and translation. Finally, because γ is unit-speed, we also have $\dot{\gamma}(a) = \pm e_1$, the first standard basis vector of $\mathbb{R}^2$.

Now we are ready to define the secant map. Let $\triangle = \{(t_1,t_2) \mid a \leqslant t_1 \leqslant t_2 \leqslant b\}$, and define the continuous function $\psi : \triangle \to S^1$ by

$$\psi(t_1,t_2) = \begin{cases} \dot{\gamma}(t_1) & t_1 = t_2 \\ -\dot{\gamma}(a) & (t_1,t_2) = (a,b) \\ \dfrac{\gamma(t_2) - \gamma(t_1)}{\|\gamma(t_2) - \gamma(t_1)\|} & \text{otherwise} \end{cases}.$$

This is a smooth function (see [1], pages 22-24), and the first two cases are straightforward to visualize. For parameters (t_1,t_2) which satisfy the third case, the vector $\psi(t_1,t_2)$ is precisely the unit vector with origin $\gamma(t_1)$ and pointing towards $\gamma(t_2)$. In particular, if (t_1,t_2) lies on a non-diagonal side of the triangle $\triangle$, then $\gamma(t_1)$ is fixed as $\gamma(t_2)$ travels along the curve (see [3] for nice animations).





By applying Proposition 2.2 in each coordinate, we see that there exists a smooth function $\tilde{\theta} : \mathbb{R}^2 \to S^1$ which gives the angle $\tilde{\theta}(t_1,t_2)$ between $\psi(t_1,t_2)$ and the horizontal. Because we defined $\psi = \dot{\gamma}$ along the diagonal, by Proposition 2.2, we know that

$$2\pi I(\gamma) = \theta(b) - \theta(a) = \tilde{\theta}(b,b) - \tilde{\theta}(a,a)$$

so $I(\gamma)$ is equal to the degree of ψ! Further, it is visually clear that we can compute the total change of $\tilde{\theta}$ the diagonal by computing the change from (a,a) to (a,b) and (a,b) to (b,b) separately, then taking a sum. Thus, we have

$$2\pi I(\gamma) = \tilde{\theta}(b,b) - \tilde{\theta}(a,a) = (\tilde{\theta}(a,b) - \tilde{\theta}(a,a)) + (\tilde{\theta}(b,b) - \tilde{\theta}(a,b)).$$

The last step is to compute the degree of ψ over the two non-diagonal segments. We will suppose γ is positively-oriented, so $\dot{\gamma}(a) = e_1$ and the secant angle is $\tilde{\theta}(a,a) = 0$ (an analogous argument holds for the opposite orientation, where $\tilde{\theta}(a,a) = \pi$). For the segment from (a,a) to (a,b), we know that the corresponding line $\psi(a,t)$ lies in the upper half-plane for all $t \in [a,b]$, so we must have $0 \leqslant \tilde{\theta}(a,t) \leqslant \pi$. Thus, we find $\tilde{\theta}(a,b) = \pi$. Meanwhile, on the segment from (a,b) to (b,b), we have the corresponding line $\psi(t,b) = -\psi(a,t)$, which implies $\tilde{\theta}(b,b) - \tilde{\theta}(a,b) = \pi$ as well. The degree of ψ is therefore

$$\frac{(\tilde{\theta}(a,b) - \tilde{\theta}(a,a)) + (\tilde{\theta}(b,b) - \tilde{\theta}(a,b))}{2\pi} = \frac{\pi + \pi}{2\pi} = 1$$

and -1 if the orientation of ψ is reversed. This shows $I(\gamma) = \pm 1$ as desired.

Altogether, we conclude that if θ is any tangent angle function for γ, then

$$\int_a^b \kappa ds = \theta(b) - \theta(a) = 2\pi I(\gamma) = \pm 2\pi,$$

which completes the proof. □

3 Regular surfaces and tangent planes

In the previous section, we showed that the two-dimensional circle can be locally unfolded to the one-dimensional real line using the function e^{it}, which gives a continuous deformation on sufficiently small intervals. Similarly, we interpret surfaces as three-dimensional objects which can be "flattened" to $\mathbb{R}^2$.

Definition 3.1. Given any subsets $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$, a invertible map $f : X \to Y$ is called a *homeomorphism* if both f and its inverse $f^{-1} : Y \to X$ are continuous. If such a map exists, we say X and Y are *homeomorphic*.

Remark. The paths defined in Section 2 are homeomorphisms from an interval of $\mathbb{R}$ to a curve in $\mathbb{R}^n$. In general, isotopies, which we only defined for simple closed plane curves, are continuous families of homeomorphisms.

Definition 3.2. A *regular surface* is a subset $\mathcal{S} \subset \mathbb{R}^3$ where for each point $p \in \mathcal{S}$, there exists an open neighborhood $V \subset \mathbb{R}^3$ containing p, an open subset $U \subset \mathbb{R}^2$, and a map $\sigma : U \to V \cap \mathcal{S}$ with the following properties:

(i) σ is a smooth function on U;

(ii) σ is a homeomorphism;

(iii) For all $q \in U$, the differential $d\sigma_q$ is injective.

In this case, the map σ is called a *surface patch* or *local parametrization* of the coordinate neighborhood $V \cap \mathcal{S}$. We will also only consider *connected* surfaces, meaning any two points in $\mathcal{S}$ can be joined by a curve lying entirely in $\mathcal{S}$.

Remark. Like paths, multiple surface patches may have the same image. Suppose the surface patches σ_1 and σ_2 are defined on the open subsets $U_1, U_2 \subset \mathbb{R}^2$ respectively. We say that two surface patches σ_1, σ_2 are *reparametrizations* of one another if there exists homeomorphism $\Phi : U_1 \to U_2$ that $\sigma_2 = \sigma_1 \circ \Phi$. In this case, the bijection Φ is called a reparametrization map. The upshot is that we can define any geometric property of a smooth surface by defining it up to reparametrization!

Condition (i) is basic for doing calculus on surfaces, like understanding what it means for a function on a surface to be differentiable. Condition (ii) ensures that the inverse $\sigma^{-1} : V \cap \sigma(U) \to U$ is continuous, so the surface has no self-intersections and the tangent to each point is unique. Condition (iii), sometimes called the regularity condition, allows us to apply the immersion theorem to conclude that σ is indeed "locally invertible" when the codomain is restricted to $V \cap \sigma(U)$.

Example. A surface is often the image of multiple surface patches. Given the unit sphere S^2, which has radius 1, we can define the smooth maps $\sigma_1, \sigma_2 : U \to S^2$ by

$$\sigma_1 \begin{pmatrix} \theta \\ \psi \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\psi \\ \cos\theta\sin\psi \\ \sin\theta \end{pmatrix} \qquad \sigma_2 \begin{pmatrix} \theta \\ \psi \end{pmatrix} = \begin{pmatrix} -\cos\theta\cos\psi \\ \sin\theta \\ -\cos\theta\sin\psi \end{pmatrix}$$

where θ and ψ are angles correspond to something like latitude and longitude, respectively. That is, if p is a point on the sphere, then we can draw a line through p which is parallel to the z-axis and intersects the xy-plane at a point q. Then θ is the angle between p and q, while ψ is the angle between q and the positive x-axis.

To ensure σ_1 and σ_2 are homeomorphisms, we take the domain to be the open set $U = (-\pi/2, \pi/2) \times (0, 2\pi) \subset \mathbb{R}^2$. Notice that σ_1 nor σ_2 cover all of S^2 when restricting the domain to U: the image of σ_1 misses points of the form $(x, 0, z)$ with $x \geqslant 0$, while the image of σ_2 misses points of the form $(x, y, 0)$ with $x \leqslant 0$. However, we have $S^2 = \sigma_1(U) \cup \sigma_2(U)$, so S^2 satisfies the definition of a surface.

Thus, the construction of a surface can be somewhat ad hoc. Our strategy also happens to be unnecessarily complicated for the sphere, which has a neat geometric origin we will introduce in the next example.

Example. A *surface of revolution* is obtained by rotating a simple plane curve, called the *profile curve*, around a straight line in the plane. Typically, the axis of revolution is the z-axis, and we define a path $\gamma : I \to \mathbb{R}^3$ on the xz-plane by $\gamma(u) = (f(u), 0, g(u))$. The surface obtained by rotating γ about the z-axis is parametrized with $\sigma : I \times [0, 2\pi) \to \mathbb{R}^3$ given by

$$\sigma(u, v) = (f(u) \cos v, f(u) \sin v,\ g(u)),$$

where v is the angle of rotation. To check for Definition 3.2 (iii), notice

$$\sigma_u \times \sigma_v = f(u)(-\dot{g}(u) \cos v,\ -\dot{g}(u) \sin v,\ \dot{f}(u)),$$

so $\sigma_u \times \sigma_v$ is nonzero if and only if $f(u) \neq 0$. Thus, the surface of revolution is indeed a surface when γ does not intesect the z-axis. In practice, we assume $f(u) > 0$ so that $f(u)$ is the distance between $\sigma(u, v)$ and the axis of rotation.

Example. The unit sphere S^2 in latitude-longitude coordinates, as in the first example, is a surface of revolution. Taking $u = \theta$ and $v = \psi$, the profile curve functions are $f(\theta) = \cos\theta$ and $g(\theta) = \sin\theta$.

Example. A torus is formed by rotating a circle in the xz-plane with center $(R, 0, 0)$ and radius r about the z-axis, with $R > r > 0$. This is a surface of revolution with profile curve

$$\gamma(\theta) = (R + r\cos\theta,\ 0,\ r\sin\theta),$$

and the parametrization is $\sigma : [0, 2\pi) \times [0, 2\pi) \to \mathbb{R}^3$ defined by

$$\sigma\begin{pmatrix}\theta \\ \psi\end{pmatrix} = \begin{pmatrix}(R + r\cos\theta)\cos\psi \\ (R + r\cos\theta)\sin\psi \\ r\sin\theta\end{pmatrix}$$

where θ is the angle in γ and ψ is the angle about the z-axis.

Now, condition (iii) of Definition 3.2 is also precisely what allows us to find the tangent *plane* to a point. It implies that the partials σ_u and σ_v are linearly independent, so their span must be a two-dimensional linear subspace. We begin defining the tangent by considering smooth curves on the surface.

Definition 3.3. Let p be any point on a surface $\mathcal{S} \subset \mathbb{R}^3$. If $\gamma : (-\epsilon, \epsilon) \to \mathcal{S}$ is a path with $\gamma(0) = p$, then *tangent vector* to $\mathcal{S}$ at p is precisely $\dot{\gamma}(0)$, the tangent vector to γ at p. The tangent space of $\mathcal{S}$ at p, denoted $T_p\mathcal{S}$, is the set of all vectors tangent to $\mathcal{S}$ at p.

Proposition 3.1. *Let p be a point on a surface $\mathcal{S} \subset \mathbb{R}^3$, and suppose $\sigma : U \to \mathbb{R}^3$ is a surface patch whose image contains p, say $p = \sigma(u_0, v_0)$. Then the tangent space of $\mathcal{S}$ at p is the vector subspace*

$$T_p\mathcal{S} = \operatorname{span}(\sigma_u, \sigma_v),$$

where σ_u, σ_v are the partial derivatives evaluated at p.

Proof. We will prove these two spaces are equal using double containment. First, if γ is a path in the image of a surface patch σ, then we have $\gamma(t) = \sigma(u(t), v(t))$ for some smooth functions $u(t)$ and $v(t)$. The existence of such smooth functions follows from properties (i)-(iii) of a surface, which imply σ^{-1} is smooth. Differentiating with the chain rule, we have

$$\dot{\gamma} = \sigma_u du + \sigma_v dv,$$

so every tangent vector of $\mathcal{S}$ can be written as a linear combination of the partials σ_u and σ_v.

On the other hand, we can write every vector $\vec{v} \in \text{span}(\sigma_u, \sigma_v)$ as a linear combination $\vec{v} = a_1\sigma_u + a_2\sigma_v$ for some with coefficients $a_1, a_2 \in \mathbb{R}$. Then we can define a curve

$$\gamma(t) = \sigma(u_0 + a_1 t, v_0 + a_2 t).$$

At the point $p = \gamma(0) \in \mathcal{S}$, we have

$$\dot{\gamma}(0) = a_1\sigma_u + a_2\sigma_v = \vec{v},$$

so every vector in the span is the tangent vector of $\mathcal{S}$ at some point p. □

4 The first fundamental form and surface area

To describe the local geometry of a surface, we need a way to make local measurements like lengths, angles, and areas. The first fundamental form allows us to compute the length of a curve on a surface using tangent vectors.

Definition 4.1. Let $p \in \mathcal{S}$ be any point of a surface. The *first fundamental form* of $\mathcal{S}$ at p is given by

$$\mathbf{I}_p(\vec{v}, \vec{w}) = \langle \vec{v}, \vec{w} \rangle,$$

where $\vec{v}, \vec{w} \in T_p\mathcal{S}$ are tangent vectors. That is, the first fundamental form I_p is the standard inner product on $\mathbb{R}^3$ restricted to the tangent space $T_p\mathcal{S}$.

In practice, this form is expressed in terms of surface patches. Suppose $p = \sigma(u_0, v_0)$ for some surface patch σ so that partial derivatives $\{\sigma_u, \sigma_v\}$ evaluated at p form a basis for the tangent plane $T_p\mathcal{S}$. Then any tangent vector $\vec{v} \in T_p\mathcal{S}$ is tangent to a curve γ in the image of σ given by $\gamma(t) = \sigma(u(t), v(t))$. As shown in the proof of Proposition 3.1, we can express the tangent vector as a linear combination $\vec{v} = \dot{\gamma}(0) = \sigma_u du + \sigma_v dv$.

We use the fact that the inner product is symmetric bilinear to expand $\mathbf{I}_p$ as the quadratic form

$$\begin{aligned}\mathbf{I}_p(\vec{v}, \vec{v}) &= \langle \sigma_u du + \sigma_v dv, \sigma_u du + \sigma_v dv \rangle \\ &= \langle \sigma_u, \sigma_u \rangle (du)^2 + 2\langle \sigma_u, \sigma_v \rangle du dv + \langle \sigma_v, \sigma_v \rangle (dv)^2.\end{aligned}$$

Traditionally, the inner product components of this form are denoted

$$E = \langle \sigma_u, \sigma_u \rangle \qquad F = \langle \sigma_u, \sigma_v \rangle \qquad G = \langle \sigma_v, \sigma_v \rangle,$$

and the expression $Edu^2 + 2Fdudv + Gdv^2$ is called the first fundamental form of the surface patch $\sigma(u, v)$. Note that the linear maps du, dv and metric coefficients E, F, G depend on choice of parametrization σ, but the form itself only depends on $\mathcal{S}$ and point p.

Finally, when γ is a curve in the image of a patch σ, we can substitute the first fundamental form of σ in the arc length formula to compute

$$\int \|\dot{\gamma}(t)\| \, dt = \int \sqrt{\langle \dot{\gamma}, \dot{\gamma} \rangle} \, dt = \int \sqrt{Edu^2 + 2Fdudv + Gdv^2} \, dt.$$

Example. For a surface of revolution with unit-speed profile curve $u \mapsto (f(u), 0, g(u))$, we have

$$\sigma_u = (\dot{f}\cos v,\ \dot{f}\sin v,\ \dot{g}) \qquad \sigma_v = (-f\sin v,\ f\cos v,\ 0).$$

Using the fact that $\dot{f}^2 + \dot{g}^2 = 1$ for the unit-speed curve, we compute the coefficients $E = 1$, $F = 0$, and $G = f^2$. Thus, the first fundamental form is

$$du^2 + f(u^2)dv^2.$$

Example. Taking $f(\theta) = \cos\theta$ and $g(\theta) = \sin\theta$, the first fundamental form of S^2 is $d\theta^2 + \cos^2\theta d\psi^2$.

Since the Gauss-Bonnet theorem involves integrating over a surface, we will briefly discuss areas of surface regions.

Definition 4.2. Given a surface patch $\sigma : U \to \mathbb{R}^3$ and a subset $R \subseteq U$, the *area* $A_\sigma(R)$ of the surface region $\sigma(R)$ is

$$A_\sigma(R) = \int_R \|\sigma_u \times \sigma_v\| du dv.$$

Using the first fundamental form to compute $\|\sigma_u \times \sigma_v\| = \sqrt{EG - F^2}$, we can further write

$$dA = \sqrt{EG - F^2} du dv.$$

Importantly, since the value $EG - F^2 = \det(\mathbf{I}_p)$ does not depend on choice of basis, the area of a surface region does not depend on choice of patch σ. This agrees with the remark about reparametrizations and geometric properties at the beginning of Section 3.

Example. The surface area of the unit sphere is

$$A(S^2) = \int_{S^2} 1 dA = \int_0^\pi \int_0^{2\pi} \sin\psi \, d\theta d\psi = 4\pi.$$

5 The second fundamental form and surface curvature

In the same way that a plane curve's signed curvature $\kappa = d\theta/ds$ is a ratio defined by associating an infinitesimal change $\dot{\gamma}$ with an infinitesimal angle $\dot{\theta}$ on the unit circle, the curvature of a surface in $\mathbb{R}^3$ is defined by associating an infinitesimal area element $dA = dudv$ with another infinitesimal area element $d\sigma$ on the unit sphere. The Gaussian curvature is precisely the ratio $K = dA/d\sigma$.

In practice, we can measure curvature by considering how the the unit normal $\mathbf{N}$ varies as we move around the surface. For the tangent plane $T_p\mathcal{S}$, Proposition 3.1 makes a choice of normal vector straightforward: if $\sigma : U \to \mathbb{R}^3$ is a surface patch which contains p, then the unit vector

$$\mathbf{N} = \frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|}$$

is perpendicular to every linear combination of σ_u and σ_v. We call $\mathbf{N}$ the *standard unit normal* of the patch σ at point p.

While $\pm\mathbf{N}$ does not depend on choice of surface patch σ, the parametrization determines the sign. In this paper, we will only consider surfaces which are *orientable*, meaning we have a smooth choice of normal $\mathbf{N}$ over the entire surface. Informally, an orientable surface has two sides; the typical example of a non-orientable surface is the Mobius strip (see [4], Example 4.5.3).

The values of $\mathbf{N}$ are given by the *Gauss map* $\mathbf{G} : \mathcal{S} \to S^2$, which sends each point $p \in \mathcal{S}$ to its standard unit normal $\mathbf{N}_p$ in the unit sphere. Since we are interested in the rate of change of $\mathbf{N}$, we need to define the derivative $d\mathbf{G}_p$ at each point. In general, given a map f between two surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$, the derivative of f is the linear map $df_p : T_p\mathcal{S}_1 \to T_{f(p)}\mathcal{S}_2$ which "pushes forward" the tangent vector to the curve $p = \gamma(0)$ in $\mathcal{S}_1$ to the tangent at $(f \circ \gamma)(0)$ in $\mathcal{S}_2$. Thus, the derivative of the Gauss map is a function

$$d\mathbf{G}_p : T_p\mathcal{S} \to T_{\mathbf{G}(p)}S^2.$$

Now by definition, $T_{\mathbf{N}_p}S^2$ is the plane through the origin perpendicular to the point $\mathbf{G}(p) = \mathbf{N}_p$, which is precisely $T_p\mathcal{S}$, so the derivative $d\mathbf{G}_p$ is actually a map from $T_p\mathcal{S}$ to itself. This brings us to a measurement of curvature known as the second fundamental form.

Definition 5.1. Let $\mathcal{S}$ be an orientable surface with Gauss map $\mathbf{G}$. For each $p \in \mathcal{S}$, the *Weingarten map* of $\mathcal{S}$ at p is the linear map $\mathbf{W} : T_p\mathcal{S} \to T_p\mathcal{S}$ is given by

$$\mathbf{W}_p = -d\mathbf{G}_p.$$

Definition 5.2. If $\mathbf{W}_p$ is the Weingarten map at a point $p \in \mathcal{S}$, the *Gaussian curvature* K of $\mathcal{S}$ at p is given by

$$K = \det(\mathbf{W}_p).$$

Remark. The Gaussian curvature does not depend on orientation of the tangent plane, as the determinant of the 2×2 matrix $\mathbf{W}_p$ is the same when every entry changes sign.

Example. The Gaussian curvature of S^2 is 1 everywhere, because the Gauss map at every point in S^2 is the precisely the identity map. Thus, the Weingarten map at every point is also the identity, and we have $K = \det(I) = 1$.

Unfortunately, most Weingarten maps are not so obvious. To get an explicit formula for K, we need to define a metric for curvature on a surface patch σ.

Definition 5.3. The *second fundamental form* of $\mathcal{S}$ at p is the bilinear map $\mathbf{II}_p : T_p\mathcal{S} \to \mathbb{R}$ defined by

$$\mathbf{II}_p = \langle \mathbf{W}_p(\vec{v}), \vec{w} \rangle$$

for some tangent vectors $\vec{v}, \vec{w} \in T_p\mathcal{S}$.

Unlike with the form $\mathbf{I}_p$, it is not immediately clear that $\mathbf{II}_p$ has a corresponding quadratic function.

Proposition 5.1. *The second fundamental form is* symmetric bilinear. *That is, for all tangent vectors $\vec{v}, \vec{w} \in T_p\mathcal{S}$, we have $\mathbf{II}_p(\vec{v}, \vec{w}) = \mathbf{II}_p(\vec{w}, \vec{v})$.*

Proof. First, let $p \in \mathcal{S}$ be a point in the image of a surface patch σ. Suppose $\gamma(t) = \sigma(u(t), v(t))$ is a curve in the patch with $\gamma(0) = p$, so $\dot{\gamma}(0) = \sigma_u du(0) + \sigma_v dv(0)$ is tangent to $\mathcal{S}$ at p. Then

$$\begin{aligned} \mathbf{W}_p(\dot{\gamma}(0)) &= -d\mathbf{G}_p(\sigma_u du(0) + \sigma_v dv(0)) \\ &= -\left.\frac{d}{dt}\mathbf{G}(u(t), v(t))\right|_{t=0} \\ &= -(\mathbf{G}_u du(0) + \mathbf{G}_v dv(0)). \end{aligned}$$

In particular, since

$$du(\sigma_u) = dv(\sigma_v) = 1 \qquad du(\sigma_v) = dv(\sigma_u) = 0,$$

we have $\mathbf{W}_p(\sigma_u) = -\mathbf{G}_u$ and $\mathbf{W}_p(\sigma_v) = -\mathbf{G}_v$.

Since $\{\sigma_u, \sigma_v\}$ is a basis for $T_p\mathcal{S}$, we can write our tangent vectors as linear combinations $\vec{v} = a_1\sigma_u + a_2\sigma_v$ and $\vec{w} = b_1\sigma_u + b_2\sigma_v$ for some $a_1, a_2, b_1, b_2 \in \mathbb{R}$. Using the fact that a bilinear form on $\mathbb{R}^n$ is linear in both inputs, we compute

$$\begin{aligned} \mathbf{II}_p(\vec{v}, \vec{w}) &= \langle \mathbf{W}_p(\vec{v}), \vec{w} \rangle = \langle -a_1\mathbf{G}_u - a_2\mathbf{G}_v, b_1\sigma_u + b_2\sigma_v \rangle \\ &= -a_1b_1\langle \mathbf{G}_u, \sigma_u \rangle - a_1b_2\langle \mathbf{G}_u, \sigma_v \rangle - a_2b_1\langle \mathbf{G}_v, \sigma_u \rangle - a_2b_2\langle \mathbf{G}_v, \sigma_v \rangle \\ &= \langle -b_1\mathbf{G}_u - b_2\mathbf{G}_v, a_1\sigma_u + a_2\sigma_v \rangle \\ &= \langle \mathbf{W}_p(\vec{w}), \vec{v} \rangle = \mathbf{II}_p(\vec{w}, \vec{v}), \end{aligned}$$

which shows the desired equality. ☐

We now obtain quadratic form: given a tangent vector $\vec{v} = \sigma_u du + \sigma_v dv$, we have

$$\mathbf{II}_p(\vec{v}, \vec{v}) = -\langle \mathbf{G}_u, \sigma_u \rangle (du)^2 - 2\langle \mathbf{G}_u, \sigma_v \rangle du dv - \langle \mathbf{G}_v, \sigma_v \rangle (dv)^2,$$

where the middle term uses the fact that $\langle \mathbf{G}_u, \sigma_v \rangle = \langle \mathbf{G}_v, \sigma_u \rangle$. The metric coefficients are traditionally denoted

$$L = -\langle \mathbf{G}_u, \sigma_u \rangle = \langle \sigma_{uu}, \mathbf{N} \rangle \qquad M = -\langle \mathbf{G}_u, \sigma_v \rangle = \langle \sigma_{uv}, \mathbf{N} \rangle \qquad N = -\langle \mathbf{G}_v, \sigma_v \rangle = \langle \sigma_{vv}, \mathbf{N} \rangle,$$

and we say $Ldu^2 + 2Mdudv + Ndv^2$ is the second fundamental form of the surface patch $\sigma(u, v)$.

Together with the first fundamental form, this gives us a very useful formula for Gaussian curvature. If we write $-\mathbf{G}_u$ and $-\mathbf{G}_v$ in terms of the basis $\{\sigma_u, \sigma_v\}$, then the explicit matrix for the Weingarten map with respect to this basis is

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} L & M \\ M & N \end{pmatrix}$$

(for the full derivation, see [4], Proposition 8.1.2). Thus, we have

$$K = \frac{LM - M^2}{EG - F^2}.$$

Example. A sphere of radius c has Gaussian curvature $1/c^2$ everywhere. This is because when a surface is scaled by some constant c, the coefficients E, F, G are multiplied by a factor of c^2 and the coefficients L, M, N are multiplied by a factor of c, so K changes by a factor of $1/c^2$.

Further, since the surface area changes by a factor of c^2, we find the *total curvature* of any sphere $\mathcal{S}$ is

$$\int_{\mathcal{S}} K dA = \frac{1}{a^2} \cdot 4\pi a^2 = 4\pi.$$

Example. Using the parametrization of the torus from Section 3, we compute the partials

$$\sigma_\theta = \begin{pmatrix} -r\sin\theta\cos\psi \\ -r\sin\theta\sin\psi \\ r\cos\psi \end{pmatrix} \qquad \sigma_\psi = \begin{pmatrix} -(R + r\cos\theta)\sin\psi \\ (R + r\cos\theta)\cos\psi \\ 0 \end{pmatrix}.$$

The coefficients for the first fundamental form are $E = r^2$, $F = 0$, and $G = (R + r\cos\theta)^2$, and the coefficients for the second are $L = r$, $M = 0$, and $N = (R + r\cos\theta)\cos\theta$. The Gaussian curvature is then

$$K = \frac{\cos\theta}{r(R + r\cos\theta)}.$$

Interestingly, the torus has both positive and negative curvature: we have $K \geqslant 0$ when $\pi/2 \leqslant \theta \leqslant \pi/2$, and $K \leqslant 0$ when $\pi/2 \leqslant \theta \leqslant 3\pi/2$.

6 The local Gauss-Bonnet theorem

The most basic version of the Gauss-Bonnet theorem applies to simple closed curves on a surface. In Section 2, we considered the particular case where the surface is a plane, where the Gaussian curvature is 0. Our next step is to extend the Umlaufsatz to curved surfaces.

Definition 6.1. Given an open subset $U \subset \mathbb{R}^2$ and a local parametrization $\sigma : U \to \mathbb{R}^3$, we say γ is a *simple closed curve* in $\sigma(U)$ if there exists a simple closed plane curve $\beta(t) = (u(t), v(t))$ such that $\gamma = \sigma \circ \beta$.

In this case, γ is *positively-oriented* if the signed unit normal $\mathbf{n}$ of β points into $\mathrm{int}(\beta) \subset \mathbb{R}^2$ at every point of β. Finally, $\mathrm{int}(\gamma) \subset \mathbb{R}^3$ is defined as the image of $\mathrm{int}(\beta)$ under the map σ.

Lemma 6.1. *In the situation above, we have*

$$\int_\gamma \dot{\theta}(s) ds = \pm 2\pi.$$

Proof. Briefly, we can find an isotopy between γ and any another simple closed curve $\tilde{\gamma}$ that is completely contained in $\mathrm{int}(\gamma)$. We choose $\tilde{\gamma}$ to be the image under surface patch σ of a *very* small circle in $\mathrm{int}(\beta)$, so the interior of $\tilde{\gamma}$ is essentially a subset of the plane in $\mathbb{R}^2$. Then using Lemma 2.3, we can replace γ with $\tilde{\gamma}$ in the above integral, and the equality follows from Hopf's Umlaufsatz.

For the first isotopy, let $p = \sigma(u_0, v_0)$ be a point in $\mathrm{int}(\gamma) = \sigma(\mathrm{int}(\beta))$. By Property (iii) of regular surfaces, we can scale the axes of $\mathbb{R}^3$ to obtain a patch $\tilde{\sigma}(V) \subset \sigma(U)$ containing p with

$$\tilde{\sigma}(x, y) = (x, y, f(x, y))$$

for some smooth map f. By the same property, we may translate the surface so that $p = \tilde{\sigma}(u_0, v_0)$. Then $\sigma^{-1}(\tilde{\sigma}(V))$ is an open subset of $U \subset \mathbb{R}^2$, so there exists an $\epsilon > 0$ such that $\sigma(B_\epsilon(p)) \subset \tilde{\sigma}(V)$.

Now, consider the isotopy of curves given by

$$h_1(s, t) = \sigma(t \cdot u(s), t \cdot v(s)).$$

By choosing sufficiently small t, such as $t = \epsilon/2$, we can find an isotopy between our original curve $\gamma = h_1(s, 1)$ and a curve in $\tilde{\sigma}(V)$. Note that such a curve has the form $\gamma_{\epsilon/2}(s) = (x(s), y(s), f(x(s), y(s)))$ for some smooth functions $x(s)$ and $y(s)$.

Using this, we define a second isotopy of curves in $\tilde{\sigma}(V)$ by

$$h_2(s,t) = (x(s), y(s), t \cdot f(x(s), y(s))).$$

This gives an isotopy between $\gamma_{\epsilon/2} = h_1(s, \epsilon/2) = h_2(s,1)$ and the simple plane curve $\tilde{\gamma} = h_2(s,0)$. Then by Lemma 2.3, we have

$$\int_\gamma \dot{\theta} ds = \int_{\gamma_{\epsilon/2}} \dot{\theta} ds = \int_{\tilde{\gamma}} \dot{\theta} ds,$$

and the final integral is equal to $\pm 2\pi$ by Theorem 2.4. $\square$

Remark. A more sophisticated version of this proof will define the *relative index* of a curve with respect to an orthonormal basis, then use the Gram-Schmidt process to produce a smooth family of bases for curves in $\tilde{\sigma}(V)$. After obtaining the plane curve $\tilde{\gamma}$, the final step is to show that the relative index of $\tilde{\gamma}$ coincides with the formula for $I(\tilde{\gamma})$ (see [6], Theorem 6.6).

Our definition of a curve's curvature also requires adjustment. Notice that given any curve γ on a surface $\mathcal{S}$, the set $\{\dot{\gamma}, \mathbf{N}, \mathbf{N} \times \dot{\gamma}\}$ is an orthonormal basis for $\mathbb{R}^3$. Recall that when γ is unit-speed, its signed curvature is given by $\kappa = \|\ddot{\gamma}\|$. There is a particular term for the projection of κ on the the tangent plane of $\mathcal{S}$.

Definition 6.2. If γ is a unit-speed curve on an surface $\mathcal{S}$, then the *geodesic curvature* of γ is defined by

$$\kappa_g = \ddot{\gamma} \cdot (\mathbf{N} \times \dot{\gamma}).$$

Remark. Informally, κ_g measures how far the curve is from being the shortest path between two points on a surface. When the surface is a plane, the shortest path is a straight line, so a plane curve in $\mathbb{R}^3$ has $\kappa_g = \kappa$ up to a sign. In general, the sign of the geodesic curvature κ_g of a curve depends on the orientation of both the surface and the curve itself.

We are now ready to prove the Gauss-Bonnet theorem for simple closed curves.

Theorem 6.2. *Let γ be a unit-speed simple closed curve on a surface patch σ, and suppose γ is positively-oriented. Then*

$$\int_\gamma \kappa_g ds = 2\pi - \int_{int(\gamma)} K dA$$

where κ_g is the geodesic curvature of γ, K is the Gaussian curvature of σ, and dA is the area element of σ. The integral over the area element is called the total curvature *of the region $int(\gamma)$.*

Proof. The argument is entirely computational. First, we will use a basis of the tangent plane to find an orthonormal basis for $\mathbb{R}^3$, then expand $\dot{\gamma}$ and $\ddot{\gamma}$ in terms of this basis. We then use this to compute κ_g, which allows us to rewrite the integral of κ_g over the curve γ as the difference of two integrals. Finally, we evaluate the integrals separately to obtain the expression on the right; the 2π term will come from a direct application of Hopf's Umlaufsatz for surface curves, while the area integral uses both fundamental forms of the surface patch σ.

Let $\{\mathbf{e}_1, \mathbf{e}_2\}$ be a smooth [4] orthonormal basis for the tangent plane at each point in the image of σ; one such choice is $\mathbf{e}_1 = \sigma_u / \|\sigma_u\|$ and $\mathbf{e}_2 = \mathbf{N} \times \mathbf{e}_1$. Then $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{N}\}$ is an orthonormal basis for $\mathbb{R}^3$. Note that since we can always swap values of $\mathbf{e}_1$ and $\mathbf{e}_2$ if necessary, we assume $\mathbf{N} = \mathbf{e}_1 \times \mathbf{e}_2$ without loss of generality.

Now, let $\theta(s)$ be the *oriented angle* between the tangent vector $\dot{\gamma}(s)$ and the basis vector $\mathbf{e}_1$. This is the angle by which $\mathbf{e}_1$ must be rotated to be parallel to $\dot{\gamma}$, when viewing the side of the surface which $\mathbf{N}$ points away from. That is, from this side, $\theta(s)$ is precisely the tangent angle from Definition 2.4 taken with respect to $\mathbf{e}_1$ instead of the standard basis. Thus, we have

$$\dot{\gamma} = \cos\theta \mathbf{e}_1 + \sin\theta \mathbf{e}_2$$

$$\ddot{\gamma} = \cos\theta \dot{\mathbf{e}}_1 + \sin\theta \dot{\mathbf{e}}_2 + \dot{\theta}(-\sin\theta \mathbf{e}_1 + \cos\theta \mathbf{e}_2),$$

[4] Here, "smooth" means that $\mathbf{e}_1, \mathbf{e}_2$ are smooth functions of the surface parameters (u, v).

where the expression for $\ddot{\gamma}$ uses the chain rule. Substituting these expressions and $\mathbf{N} = \mathbf{e}_1 \times \mathbf{e}_2$ into the formula for geodesic curvature, we find that $\kappa_g = \dot{\theta} - \mathbf{e}_1 \cdot \dot{\mathbf{e}}_2$ (for full computations, see [4], Theorem 13.1.2). We can therefore compute the left side of the claimed equality as

$$\int_\gamma \kappa_g ds = \int_\gamma \dot{\theta} ds - \int_\gamma \mathbf{e}_1 \cdot \dot{\mathbf{e}}_2 \, ds.$$

First, we know know from Lemma 6.1 that the integral of $\dot{\theta}$ around γ is equal to $\pm 2\pi$; since γ is positively-oriented, this is exactly 2π. It remains to show that

$$\int_\gamma \mathbf{e}_1 \cdot \dot{\mathbf{e}}_2 \, ds = \int_{\mathrm{int}(\gamma)} K dA.$$

Differentiating $\mathbf{e}_2$, we have

$$\begin{aligned}\int_\gamma \mathbf{e}_1 \cdot \dot{\mathbf{e}}_2 \, ds &= \int_\gamma \mathbf{e}_1 \cdot ((\mathbf{e}_2)_u \dot{u} + (\mathbf{e}_2)_v \dot{v}) \, ds = \int_\beta (\mathbf{e}_1 \cdot (\mathbf{e}_2)_u) du + (\mathbf{e}_1 \cdot (\mathbf{e}_2)_v) dv \\ &= \int_{\mathrm{int}(\beta)} \left[(\mathbf{e}_1 \cdot (\mathbf{e}_2)_v)_u - (\mathbf{e}_1 \cdot (\mathbf{e}_2)_u)_v \right] \, dudv,\end{aligned}$$

where the last equality uses Green's theorem (see [5], Appendix 2, Theorem 2.6). Now given the first and second fundamental forms of σ,

$$E du^2 + 2F dudv + G dv^2 \quad \text{and} \quad L du^2 + 2M dudv + N dv^2,$$

we can write the partial derivatives of $\mathbf{e}_1$ and $\mathbf{e}_2$ in terms of the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{N}\}$ to see that

$$(\mathbf{e}_1)_u \cdot (\mathbf{e}_2)_v - (\mathbf{e}_1)_v \cdot (\mathbf{e}_2)_u = \frac{LN - M^2}{(EG - F^2)^{1/2}}$$

(for full computations with coefficients, see [4], Lemma 13.1.3). Then applying the formulas for dA and K, this intergral becomes

$$\int_\gamma \mathbf{e}_1 \cdot \dot{\mathbf{e}}_2 \, ds = \int_{\mathrm{int}(\beta)} \frac{LN - M^2}{(EG - F^2)^{1/2}} \, dudv = \int_{\mathrm{int}(\gamma)} \frac{LN - M^2}{EG - F^2} \, dA = \int_{\mathrm{int}(\gamma)} K dA.$$

This completes the proof. □

For the remainder of this paper, our discussion will be in terms of regions on surfaces rather than curves. By region, we mean a compact, simply-connected subset $\triangle$ of a surface $\mathcal{S}$. We will only consider regions with *piecewise-smooth* boundaries, which means the boundary $\partial\triangle$ looks like a polygon with curved sides, or possibly a simple closed curve with no vertices. The boundary $\partial\triangle$ is *positively-oriented* if, for all t such that $\gamma_i(t)$ is not a vertex, the signed unit normal $\mathbf{n}$ obtained by rotating $\dot{\gamma}_i$ counterclockwise by $\pi/2$ points into $\triangle$.

The next version of the Gauss-Bonnet theorem accounts for boundary vertices, where a single oriented angle is undefined, by using exterior angles. Given a vertex v of the polygon, we have one curved edge γ_i traveling towards v and another edge γ_j traveling away. As in the beginning of the proof of Theorem 6.2, take $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{N}\}$ to be a smooth orthonormal basis of $\mathbb{R}^3$, and let θ_i and θ_j be the oriented angles of $\dot{\gamma}_i$ and $\dot{\gamma}_j$ at v, respectively. The *exterior angle* at v is given by $\delta = \theta_j - \theta_i$. Since this is only well-defined up to multiples of 2π, we assume $-\pi < \delta < \pi$.

Theorem 6.3 (Local Gauss-Bonnet)**.** *Let R be a simply-connected region with piecewise-smooth boundary in a surface path σ. If the boundary $\partial\triangle$ is positively-oriented, then we have*

$$\int_{\partial\triangle} \kappa_g ds = 2\pi - \sum_{i=1}^{n} \delta_i - \int_\triangle K dA,$$

where δ_i is the exterior angle for some vertex $i = 1, ..., n$.

Proof. This is essentially a generalization of Theorem 6.2 to curves with "corners." Applying the same argument as before, we find

$$\int_{\partial\triangle} \kappa_g ds = \int_{\partial\triangle} \dot{\theta} ds - \int_{\triangle} K dA.$$

It remains to show that

$$\int_{\partial\triangle} \dot{\theta} ds = 2\pi - \sum_{i=1}^{n} \delta_i.$$

The strategy is to approximate $\partial\triangle$ with a smooth curve γ which rounds off the corners. We know by Lemma 6.1 that the total turning angle going once around γ is exactly 2π. Now notice that since $\partial\triangle$ is piecewise-smooth, the integral on the left-hand side of the equality is really the sum of n integrals along the edges of the polygon, and the turning angle at each vertex is excluded from the total. We therefore take γ to be a close enough approximation such that the difference between 2π and $\int_{\partial\triangle} \dot{\theta}$ is only due to these vertex angles, and the equality follows (for a more rigorous argument, see [4], Theorem 13.2.2). ☐

Example. Consider an n-gon on the plane with straight edges. In this case, we have $K = 0$ and $\kappa_g = 0$ for each side of the polygon. An *internal angle* of the polygon is given by $\alpha_i = \pi - \delta_i$ for $i = 1, ..., n$ and $0 < \alpha_i < 2\pi$. Then Theorem 6.3 implies

$$\sum_{i=1}^{n} \alpha_i = (n-2)\pi,$$

a well-known formula from elementary geometry.

7 The global Gauss-Bonnet theorem

The most general version of the Gauss-Bonnet theorem applies to compact, oriented surfaces with piecewise-smooth boundary. Roughly speaking, any such surface may be covered with a specific arrangement of finitely many "polygons," and we can find the entire surface's curvature by applying the local Gauss-Bonnet theorem to each polygon and taking the sum.

Definition 7.1. A surface $\mathcal{S} \subset \mathbb{R}^3$ can be *triangulated* if it is possible to write $\mathcal{S} = \bigcup_{\lambda=1}^{F} \triangle_\lambda$, where

(i) Each $\triangle_\lambda$ is the image of a triangle under a local parametrization σ;

(ii) For all $\lambda \neq \mu$, the intersection $\triangle_\lambda \cap \triangle_\mu$ is either empty, a single vertex, or a single edge;

(iii) When $\triangle_\lambda \cap \triangle_\mu$ is a single edge, the orientations of the edge are opposite in $\triangle_\lambda$ and $\triangle_\mu$;

(iv) For all λ, at most one edge $\triangle_\lambda$ is contained in $\partial\mathcal{S}$.

In this case, each region $\triangle_\lambda$ is called a *face*, and a collection of such faces is called a *triangulation* $\mathbf{T}$ of $\mathcal{S}$.

Remark. The choice of compatible orientation in (iii) gives us an orientation on the boundary of $\mathcal{S}$, which comes from the normal $\mathbf{N}$ and orientation of $\mathcal{S}$ itself. However, we do not need to worry about boundary orientation in when integrating κ_g in the theorem. If we have instead $-\mathbf{N}$, then the orientation on $\partial\mathcal{S}$ swaps while $\mathbf{N} \times \dot{\gamma}$ is unchanged, so the sign of κ_g on $\partial\mathcal{S}$ does not depend on choice of orientation on $\mathcal{S}$.

Theorem 7.1. *Every compact surface has a triangulation with finitely many faces.*

This the proof of this theorem, which comes from algebraic topology, has a relatively simple idea. For every point $p \in S$, we can find a small disc containing p, and we know $\mathcal{S}$ can be covered by a finite collection of these discs because the surface is compact. We can triangulate the interior of each disc, then paste them together to make a surface homeomorphic to $\mathcal{S}$. The challenge with a formal proof is adjusting for how the discs may overlap (see [2]).

We now define the topological invariant of interest in the final Gauss-Bonnet theorem.

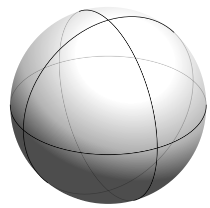
Definition 7.2. For any triangulation of a surface $\mathcal{S}$, the *Euler characteristic* of the triangulation is given by

$$\chi(\mathcal{S}) = V - E + F,$$

where V, E, F denote the total number of vertices, edges, and faces, respectively.

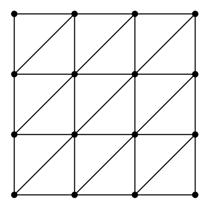
Theorem 7.2. *Let $\mathcal{S}$ be a surface equipped with a triangulation. If $\mathcal{S}$ is homeomorphic to another surface $\mathcal{S}'$, then $\chi(\mathcal{S}) = \chi(\mathcal{S}')$.*

Example. One triangulation of S^2 is found by intersecting the sphere with three coordinate planes.



This triangulation has eight faces, and its Euler characteristic is $6 - 12 + 8 = 2$.

Example. To triangulate the torus, we use the fact that the torus is homeomorphic to a square: roll the square into a tube, then stretch the tube so that the two ends meet as a donut. A triangulation of the square is shown below.



Taking into account that opposite sides of the squares will meet once rolled into the torus, we find the Euler characteristic of this triangulation to be $9 - 27 + 18 = 0$.

While different triangulations of a surface $\mathcal{S}$ may have different numbers of vertices, edges, and faces, the Euler characteristic $\chi(\mathcal{S})$ only depends on the surface itself. This important property is a consequence of the final Gauss-Bonnet theorem.

Theorem 7.3 (Global Gauss-Bonnet)**.** *Let $\mathcal{S} \subset \mathbb{R}^3$ be a compact, oriented surface with piecewise-smooth boundary. Then*

$$\int_{\partial\mathcal{S}} \kappa_g ds + \int_{\mathcal{S}} K dA + \sum_{i=1}^{n} \delta_i = 2\pi\chi(\mathcal{S}),$$

where δ_i with $i = 1, ..., n$ is an exterior angle of $\partial\mathcal{S}$.

Proof. Notice that because the left-hand side of the equality has nothing to do with a chosen triangulation, our proof will hold for any choice of triangulation for $\chi(\mathcal{S})$. As mentioned, the main idea is to apply the local Gauss-Bonnet theorem to each $\triangle_\lambda$ of the triangulation, then use the total to compute each term on the left-hand side. The integral values are easy to find, but we need some additional geometric reasoning to find the difference between the total exterior angles of $\partial\mathcal{S}$ and the total of exterior angles in the triangulation.

We begin by expressing the integrals over $\mathcal{S}$ in terms of the triangulation. For the integral of κ_g on the boundary, we know from Definition 7.1 (iii) that any edge of the triangulation which is not in $\partial\mathcal{S}$ will be paired with an edge of the opposite orientation. Because κ_g changes sign when the orientation of the curve is reversed, the integral of κ_g on non-boundary edges cancels out in pairs. As for the area integral, the area of $\mathcal{S}$ is the sum of each $\triangle_\lambda$ by definition. Thus, we have

$$\int_{\partial\mathcal{S}} \kappa_g ds = \sum_{\lambda=1}^{F} \int_{\partial\triangle_\lambda} \kappa_g ds \qquad \int_{\mathcal{S}} K dA = \sum_{\lambda=1}^{F} \int_{\partial\triangle_\lambda} K dA.$$

Now, we compute the total curvature of each region $\triangle_\lambda$. Let δ_{λ_j} for $j = 1, 2, 3$ denote an exterior angle of $\triangle_\lambda$. Applying Theorem 6.3, we have

$$\int_{\partial \triangle_\lambda} \kappa_g ds + \int_{\triangle_\lambda} K dA + \sum_{j=1}^{3} \delta_{\lambda_j} = 2\pi,$$

and the sum over all $\triangle_\lambda$ is

$$\int_{\partial \mathcal{S}} \kappa_g ds + \int_{\mathcal{S}} K dA + \sum_{\lambda=1}^{F} \sum_{j=1}^{3} \delta_{\lambda_j} = 2\pi F.$$

To complete the proof, we just need to show that the difference between the sum in the previous expression and the total exterior angle of $\partial\mathcal{S}$ is exactly

$$\sum_{\lambda,j} \delta_{\lambda_j} - \sum_{i=1}^{n} \delta_i = 2\pi(E - V).$$

This is merely a matter of counting. We first make a distinction between vertices of triangulation on the boundary and in the interior of $\mathcal{S}$, denoting the respective totals by V_B and V_I. We do the same for edges of the triangulation that are on the boundary, edges in the interior, and edges that join a boundary vertex to an interior vertex, denoting these totals by E_B, E_I, and E_{IB}.

Letting α_{λ_j} denote the interior angles of the region $\triangle_\lambda$, we have

$$\sum_{\substack{\text{interior} \\ \text{vertices}}} \delta_{\lambda_j} = \sum_{\substack{\text{interior} \\ \text{vertices}}} (\pi - \alpha_{\lambda_j}) = \pi(2E_I + E_{IB}) - 2\pi V_I.$$

This is because each interior edge contributes two interior vertices to the total, while each interior/boundary edge contributes one. Further, the interior angles at each interior vertex sum to 2π.

On the other hand, given a boundary vertex υ, we will denote the associated angle or number with a superscript (υ). Every boundary vertex υ is contained in $E_{IB}^{(\upsilon)} + 1$ faces. Moreover, the total interior angle at any boundary vertex is π if the vertex is on a smooth curve, and $\pi - \delta_i$ if the vertex is a "corner" of $\partial\mathcal{S}$ with exterior angle δ_i. Thus,

$$\begin{aligned} \sum_{\substack{\text{boundary} \\ \text{vertices } \upsilon}} \delta_{\lambda_j} &= \sum_{\substack{\text{boundary} \\ \text{vertices } \upsilon}} (\pi - \alpha_{\lambda_j}) = \sum_{\substack{\text{boundary} \\ \text{vertices } \upsilon}} \pi(E_{IB}^{(\upsilon)} + 1) - \left(\sum_{\text{smooth } \upsilon} \alpha_{\lambda_j} + \sum_{\text{corner } \upsilon} \alpha_{\lambda_j} \right) \\ &= \pi E_{IB} + \sum_{i=1}^{n} \delta_i. \end{aligned}$$

Using the fact that $V_B = E_B$ for the closed polygon $\partial\mathcal{S}$, we find

$$\begin{aligned} \sum_{\lambda,j} \delta_{\lambda_j} &= \sum_{\substack{\text{interior} \\ \text{vertices}}} \delta_{\lambda_j} + \sum_{\substack{\text{boundary} \\ \text{vertices}}} \delta_{\lambda_j} = 2\pi(E_I + E_{IB} - V_I) + \sum_{i=1}^{n} \delta_i \\ &= 2\pi(E_I + E_{IB} + E_B - V_I - V_B) + \sum_{i=1}^{n} \delta_i = 2\pi(E - V) + \sum_{i=1}^{n} \delta_i \end{aligned}$$

as desired. At last, we conclude

$$\int_{\partial \mathcal{S}} \kappa_g ds + \int_{\mathcal{S}} K dA + \sum_{i=1}^{n} \delta_i = 2\pi F - 2\pi(E - V) = 2\pi\chi(\mathcal{S}).$$

□

For surfaces without boundary, sometimes called *closed* surfaces, we have the following remarkable result.

Corollary 7.4. *When $\mathcal{S} \subset \mathbb{R}^3$ is a compact, oriented surface without boundary, the total curvature of $\mathcal{S}$ is*

$$\int_{\mathcal{S}} K dA = 2\pi\chi(\mathcal{S}).$$

Example. If $\mathcal{S}$ is any sphere, we know $\chi(\mathcal{S}) = 2$, so the Gauss-Bonnet theorem says

$$\int_{\mathcal{S}} K dA = 4\pi.$$

This agrees with our computation at the end of Section 5.

Example. If $\mathcal{S}$ is a torus, then $\chi(\mathcal{S}) = 0$ and the Gauss-Bonnet theorem says

$$\int_{\mathcal{S}} K dA = 0.$$

Earlier, we saw that the torus has both positively and negatively curved regions; we now know the positive and negative contributions cancel each other out.

References

[1] Rachad Bentbib. *The Gauss-Bonnet Theorem.* PhD thesis, Universite Cadi Ayyad, June 2017.

[2] P. H. Doyle and D. A. Moran. A short proof that compact 2-manifolds can be triangulated. *Inventiones Mathematicae*, 5(2):160–162, June 1968.

[3] Oliver Knill. Proof of the hopf umlaufsatz by deformation. `https://mathematik.com/Hopf/`, Jun 2006.

[4] Andrew Pressley. *Elementary Differential Geometry.* Springer Undergraduate Mathematics Series. Springer, London, New York, 2nd edition, 2010.

[5] Theodore Shifrin. Differential Geometry: A First Course in Curves and Surfaces.

[6] Andrew Swann. Supplementary Notes for MM08 Geometry I.