Differential Geometry *

Zhiyuan Bai

Compiled on June 5, 2022

This document serves as a set of revision materials for the Cambridge Mathematical Tripos Part II course *Differential Geometry* in Lent 2021. However, despite its primary focus, readers should note that it is NOT a verbatim recall of the lectures, since the author might have made further amendments in the content. Therefore, there should always be provisions for errors and typos while this material is being used.

Contents

*Based on the lectures under the same name taught by Dr. A. Keating in Lent 2021.

1 Smooth Manifolds and Smooth Maps

1.1 Definitions

Definition 1.1. For open $U \subset \mathbb{R}^n$, a map $f : U \to \mathbb{R}^m$ is smooth (or C^∞) if it has continuous partial derivatives of all orders.
For a general subset $X \subset \mathbb{R}^n$, a map $f : X \to \mathbb{R}^n$ is smooth (or C^∞) if for any $x \in X$, there is some open $W \subset \mathbb{R}^n$ containing x and smooth $F : W \to \mathbb{R}^m$ such that $F|_{W\cap X} = f|_{W\cap X}$.

Smoothness is what we call a "local property", in the sense that $f : X \to \mathbb{R}^m$ is smooth if and only if all $f|_{W_i \cap X}$ are smooth for some $\{W_i\}$ open in $\mathbb{R}^n$ such that $X \subset \bigcup_i W_i$.

Definition 1.2. A smooth map $f : X \to Y$ (with, say, $X \subset \mathbb{R}^n, Y \in \mathbb{R}^m$) is a diffeomorphism if it is bijective and has smooth inverse.
If such a diffeomorphism exists, then X, Y are said to be diffeomorphic.

Definition 1.3. $X \subset \mathbb{R}^N$ is a k-dimensional smooth manifold (k-manifold) if each point $x \in X$ has an open neighbourhood $V \subset X$ diffeomorphic to an open subset U of $\mathbb{R}^k$. Suppose the diffeomorphism is $\phi : U \to V$. We call ϕ a parameterisation of V and ϕ^{-1} a chart on V.

ϕ^{-1} is sometimes also called a coordinate system on V, since its components can be thought of the coordinates (or coordinate functions) on V.
Recall from algebraic topology that if X is both a k-dimensional smooth manifold and a l-dimensional one, then $k = l$. So if X is a k-manifold, we can confidently write $\dim X = k$.

Example 1.1. The unit sphere $S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ is a smooth manifold of dimension 2. Indeed, we can parameterise the upper hemisphere by $(x, y) \mapsto (x, y, \sqrt{1 - x^2 - y^2})$ (and other parts of the sphere similarly).
The k-sphere $S^k = \{x \in \mathbb{R}^{k+1} : \|x\| = 1\}$ is a smooth manifold of dimension k in a similar way.

You will show on example sheet that if $X \subset \mathbb{R}^m, Y \subset \mathbb{R}^n$ are smooth manifolds, then $X \times Y \subset \mathbb{R}^{m+n}$ is also a smooth manifold. Furthermore, it has dimension $\dim X + \dim Y$.

Definition 1.4. If X, Z are both manifolds in $\mathbb{R}^N$ and $Z \subset X$, then Z is called a submanifold of X. The codimension of Z in X is $\dim X - \dim Z$.

1.2 Tangent Space and Derivations

Recall from analysis that for an open $U \subset \mathbb{R}^n$, and a smooth $f : U \to \mathbb{R}^m$, the derivative of f at $x \in U$ is a linear map $Df_x : \mathbb{R}^n \to \mathbb{R}^m$ such that $f(x+h) = f(x) + Df_x(h) + \epsilon(h)$ where $\|\epsilon(h)\|/\|h\| \to 0$ as $\|h\| \to 0$. Furthermore, Df_x has the matrix

$$Df_x = \left(\frac{\partial f_i}{\partial x_j}\bigg|_x \right)_{i,j} = \begin{pmatrix} \partial f_1/\partial x_1 & \cdots & \partial f_1/\partial x_n \\ \vdots & \ddots & \vdots \\ \partial f_m/\partial x_1 & \cdots & \partial f_m/\partial x_n \end{pmatrix}\Bigg|_x$$

Definition 1.5. For an open $U \subset \mathbb{R}^n$ and $x \in U$, the tangent space T_xU to U at x is just $\mathbb{R}^n$.
For an arbitrary k-manifold $X \subset \mathbb{R}^N$, we take a parameterisation $\phi : U \to X$ for open $U \subset \mathbb{R}^k$ and define the tangent space T_xX of X at $x \in U \subset X$ to be the image of $D\phi_{\phi^{-1}(x)}$.

Lemma 1.1. *T_xX is independent of the choice of ϕ. Moreover, $D\phi_{\phi^{-1}(x)} : \mathbb{R}^k \to T_xX$ is an isomorphism. In particular, $\dim T_xX = \dim X$.*

Proof. Suppose $\phi : V \to X, \psi : U \to X$ are two distinct local parameterisations near x. WLOG $0 \in U \cap V, \phi(0) = \psi(0) = x$. We can also assume WLOG that $\phi(U) = \psi(V) = W \subset X$ by shrinking U, V. Now $h = \psi^{-1} \circ \phi : U \to V$ is a diffeomorphism, and by chain rule we have $D\phi_0 = D\psi_0 \circ Dh_0$. But Dh_0 is invertible since h is a diffeomorphism, so $D\phi_0$ and $D\psi_0$ have to have the same image.
As for the dimension of the tangent space, since $\phi^{-1} : \phi(U) \to U$ is smooth, we can choose $\tilde{W} \ni x$ open in $\mathbb{R}^N$ such that there is a smooth map $\bar{\phi} : \tilde{W} \to \mathbb{R}^k$ extending ϕ^{-1}. Then $\bar{\phi} \circ \phi$ is identity on a neighbourhood of $\phi^{-1}(x)$. By the chain rule, $D\phi_{\phi^{-1}(x)}$ and $D\bar{\phi}_x$ are inverses to each other which establishes the desired isomorphism. ☐

Remark. The cosets $x + T_xX \subset \mathbb{R}^N$ (the "affine tangent spaces") can be thought of as a linear approximation to X near x.

Definition 1.6. Suppose $f : X \to Y$ is a smooth map between manifolds $X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m$. Given $x \in X$, there is an open neighbourhood $x \in U \subset \mathbb{R}^n$ and smooth $F : U \to \mathbb{R}^m$ with $F|_{W \cap X} = f|_{W \cap X}$. We define the derivative $Df_x : T_xX \to T_{f(x)}Y$ of f at x to be the restriction of $DF_x : \mathbb{R}^n \to \mathbb{R}^m$ to T_xX.

Lemma 1.2. *Df_x is well-defined. That is, the restriction $DF_x|_{T_x(X)}$ does not depend on the choice of F and $DF_x(T_xX) \subset T_{f(x)}Y$.*

Proof. Let $y = f(x)$. Take parameterisations $\phi : U \to X$ and $\psi : V \to Y$ with $U \subset \mathbb{R}^k, V \subset \mathbb{R}^l$ open (so $\dim X = k, \dim Y = l$). WLOG $\phi(0) = x, \psi(0) = y, f(\phi(U)) \subset \psi(V)$. Say $\phi(U) = W \cap X$ with W open in $\mathbb{R}^n$. Again by shrinking we can assume the smooth function F extending f is defined on W.
There are a lot of notations so let's draw a commutative diagram

$$\begin{array}{ccc} \phi(U) = W \cap X & \xrightarrow{F} & \psi(V) \\ \phi \uparrow & & \uparrow \psi \\ U & \xrightarrow[h = \psi^{-1} \circ F \circ \phi]{} & V \end{array}$$

Note that $\psi^{-1} \circ f \circ \phi = \psi^{-1} \circ F \circ \phi$ on U, so h only depends on f and the parameterisations. Chain rule allows us to read off an accompanying diagram

$$\begin{array}{ccc} \mathbb{R}^m & \xrightarrow{DF_x} & \mathbb{R}^m \\ \uparrow & & \uparrow \\ T_xX & \xrightarrow{Df_x} & T_yY \\ {\scriptstyle D\phi_0}\uparrow & & \uparrow{\scriptstyle D\psi_0} \\ \mathbb{R}^k & \xrightarrow[Dh_0]{} & \mathbb{R}^l \end{array}$$

That is, $DF_x \circ D\phi_0 = D\psi_0 \circ Dh_0$ which implies $DF_x(T_xX) = DF_x(D\phi_0(\mathbb{R}^k)) = D\psi_0(Dh_0(\mathbb{R}^k)) \subset D\psi_0(\mathbb{R}^l) = T_yY$. Furthermore, since $D\psi_0$ and $D\phi_0$ are isomorphisms we can write $Dh_0 = (D\psi_0)^{-1} \circ DF_x \circ D\phi_0$, so $Df_x = DF_x|_{T_xX} = D\psi_0 \circ Dh_0 \circ (D\phi_0)^{-1}$ which is independent of the choice of F. □

Corollary 1.3 (Chain Rule on Smooth Manifolds)**.** *If $f : X \to Y, g : Y \to Z$ are smooth maps between manifolds, then $D(g \circ f)_x = Dg_{f(x)} \circ Df_x$.*

Proof. Immediate from the ordinary chain rule. □

Naturally, certain properties about smooth functions that we've seen in analysis can be translated nicely to the context of manifolds.

Definition 1.7. Suppose $f : X \to Y$ is a smooth map between manifolds. We say f is a local diffeomorphism at $x \in X$ if f restricts to a diffeomorphism from an open neighbourhood of x to an open neighbourhood of $f(x)$.

Theorem 1.4 (Inverse Function Theorem)**.** *If $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}^n$ is smooth such that $Df_x : \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism for $x \in U$, then f is a local diffeomorphism at x.*

Proof. Analysis. □

Theorem 1.5 (Inverse Function Theorem for Manifolds)**.** *Suppose $f : X \to Y$ is smooth and $Df_x : T_xX \to T_{f(x)}Y$ is an isomorphism, then f is a local diffeomorphism at x.*

Proof. Take parameterisations $\phi : U \to X, \psi : V \to Y$ with $x \in U, f(x) \in V$ and apply the ordinary inverse function theorem to $h = \psi^{-1} \circ f \circ \phi : U \to V$ (note that X, Y already have the same dimension due to the isomorphism). □

1.3 Regular Values and Critical Values

Definition 1.8. A smooth map $f : X \to Y$ is regular at $x \in X$ if $Df_x : T_xX \to T_{f(x)}Y$ is surjective.

If f is regular at x, we also say f is a submersion at x.

Definition 1.9. If f is not regular at x, we say x is a critical point (of f). For $y \in Y$, if $f^{-1}(y)$ contains some critical points, we call y a critical value; Otherwise, we call it a regular value.

Remark. If $\dim X < \dim Y$, then Df_x can never be surjective and every $x \in X$ is a critical point, so $y \in Y$ is regular iff $y \notin f(X)$.

Theorem 1.6 (Preimage Theorem)**.** *Let y be a regular value of a smooth $f : X \to Y$ with $\dim X \geq \dim Y$, then $f^{-1}(y)$ is a submanifold of X. If $y \in f(X)$, we have $\dim f^{-1}(y) = \dim X - \dim Y$.*

Proof. If $y \notin f(X)$ there is nothing to prove. Assuming henceforth that $f(x) = y$ for some $x \in X$. As y is a regular value, $Df_x : T_xX \to T_yY$ is surjective. Let $K = \ker Df_x$, then $p = \dim K = \dim X - \dim Y$.
Suppose $X \subset \mathbb{R}^n$. We can choose a linear $T : \mathbb{R}^n \to \mathbb{R}^p$ with $K \cap \ker T = \{0\}$. Consider $F : X \to Y \times \mathbb{R}^p$ given by $z \mapsto (f(z), T(z))$ which is smooth and $DF_x = (Df_x, T)$ is nonsingular by the choice of T. Inverse function theorem then says that F maps some open neighbourhood U of x diffeomorphically onto a neighbourhood V of $(y, T(x))$. By restriction, F maps $f^{-1}(y) \cap U$ diffeomorphically onto $(\{y\} \times \mathbb{R}^p) \cap V$ which induces a local chart around x. This completes the proof. □

Remark. The proof also shows that $T_xf^{-1}(y) = \ker Df_x$.

Example 1.2. 1. Take $f : \mathbb{R}^{n+1} \to \mathbb{R}$ and $f(x) = \sum_i x_i^2$, then $Df_a(h) = 2\sum_i a_ih_i$. The only critical value of f is 0, so for example $f^{-1}(1)$ (which is just S^n) is a smooth n-manifold.
2. The orthogonal group $O(n) = \{A \in M(n) : AA^\top = I\}$ (with the usual identification $M(n) \cong \mathbb{R}^{n\times n}$) is also a smooth manifold, of dimension $n(n-1)/2$: Let $S(n) \subset \mathbb{R}^{n\times n}$ be the submanifold of symmetric matrices which is naturally diffeomorphic to $\mathbb{R}^{n(n+1)/2}$. Note that we can then naturally identify $T_AS(n) = S(n)$ and $T_AM(n) = M(n)$.
The map $f : M(n) \to S(n)$ given by $f(A) = AA^\top$ has derivative $Df_A(H) = HA^\top + AH^\top$. If we can show that I is a regular value then we are done, since $O(n) = f^{-1}(I)$. Fix $A \in O(n)$, then given any $C \in T_{f(A)}S(n)$, there is some $H \in T_AM(n)$ such that $Df_A(H) = HA^\top + A^\top H = C$. Indeed, one can just take $H = CA/2$. Consequently, Df_A is always surjective for any $A \in f^{-1}(I)$, so I has to be a regular value.

Corollary 1.7. *Suppose $f : X \to Y$ is smooth and $\dim X = \dim Y$, then if X is compact and y is a regular value of f, then $f^{-1}(y)$ is a finite set of points.*

Proof. It is a 0-dimensional manifold by the preceding theorem. Compactness does the rest. □

Theorem 1.8 (Stack of Records Theorem)**.** *Suppose we are in the context of the preceding corollary and let $f^{-1}(y) = \{x_1, \dots, x_k\}$. Then y has an open neighbourhood U such that $f^{-1}(U)$ is a disjoint union $U_1 \sqcup \dots \sqcup U_k$ where U_i is an open neighbourhood of x_i and f maps each U_i diffeomorphically onto U.*

Proof. By inverse function theorem, we can find disjoint neighbourhoods W_i of x_i such that f maps W_i diffeomorphically onto $f(W_i) \ni y$. Then $X \setminus \bigcup_i W_i$ is closed hence compact by the compactness of X, and thus $f(X \setminus \bigcup_i W_i)$ is compact henced closed as Y is Hausdorff. Taking $U = (\bigcap_i f(W_i)) \setminus (f(X \setminus \bigcup_i W_i))$ finishes the proof. □

Remark. Consequently, the integer-valued function $y \mapsto |f^{-1}(y)|$ is locally constant as y ranges over regular values of f.

So regular values are really nice. But how about those critical values?

Definition 1.10. A set $A \subset \mathbb{R}^n$ has measure zero if it can be covered by a countable set of rectangular solids (i.e. product of intervals) with arbitrarily small total (n-dimensional) volume.
For an n-manifold X, $A \subset X$ has measure zero if, for every local parameterisation ϕ of X, $\phi^{-1}(A)$ has measure zero in $\mathbb{R}^n$.

Theorem 1.9 (Sard's Theorem)**.** *The set of critical values of a smooth map has measure zero.*

Proof. Omitted. □

Remark. A measure zero set on a manifold cannot contain a nonempty open subset, for any open subset of $\mathbb{R}^n$ has nonzero n-dimensional volume. Consequently, the set of regular values of a smooth function has to be dense.

1.4 Transversality

Suppose we have a smooth $f : X \to Y$ and a submanifold $Z \subset Y$. What can we say about $f^{-1}(Z)$?

Definition 1.11. A smooth map $f : X \to Y$ is transversal to a submanifold $Z \subset Y$ (written as $f \pitchfork Z$) if for any $x \in f^{-1}(Z)$ we have $Df_x(T_xX) + T_{f(x)}Z = T_{f(x)}Y$.

Example 1.3. A curve $f : \mathbb{R} \to \mathbb{R}^3$ is transversal to S^2 whenever the curve "pokes through" S^2; However, if it simply "glides through" it, then it is not transversal. Similarly, a surface $f : \mathbb{R}^2 \to \mathbb{R}^3$ is transversal to S^2 when the surface is somehow "inserted" into S^2.

Definition 1.12. A smooth $g : Z \to Y$ is an immersion at $y \in Z$ if $Dg_y : T_yZ \to T_{g(y)}Y$ is injective.

Theorem 1.10 (Local Immersion Theorem)**.** *If $g : Z \to Y$ is an immersion, then we can choose local coordinates on Z and Y such that g is locally given as $g(x_1, \ldots, x_k) = (x_1, \ldots, x_k, 0, \ldots, 0)$.*

Proof. Exercise. □

Theorem 1.11 (Generalised Preimage Theorem)**.** *If a smooth $f : X \to Y$ is transversal to a submanifold $Z \subset Y$, then $f^{-1}(Z)$ is a submanifold of X. Also, the codimension of $f^{-1}(Z)$ over X equals the codimension of Z over Y.*

Remark. 1. We recover Theorem 1.6 by taking Z to be the singleton set consisting of a regular value.
2. If $f : X \to Y$ is an inclusion, then it is transversal to another submanifold $Z \subset Y$ iff $T_xX + T_xZ = T_xY$ for all $x \in X \cap Z$.

Proof. By Theorem 1.10, Z can be written in a neighbourhood of a point $y = f(x) \in Z$ as the common zero set of smooth functions $h_1, \ldots, h_r$ where r is the codimension of Z in Y. Let $H = (h_1, \ldots, h_r)$, then $f^{-1}(Z)$ is the zero set of $H \circ f$. We are done by Theorem 1.6 as $0 \in \mathbb{R}^r$ is a regular value of $H \circ f$ from the transversality condition. □

Corollary 1.12. *The intersection of two transversal submanifolds X, Z of Y is again a submanifold whose codimension is the sum of codimensions of X and Z.*

Proof. Immediate. □

Remark. 1. Transversality is a "stable condition", i.e. a condition that survives under sufficiently small perturbations.
2. Transversality is also a "generic behaviour", in the sense that arbitrary smooth map can be deformed by an arbitrarily small amount to maps which are transversal to a given submanifold.

1.5 Manifolds with Boundary

Definition 1.13. The k-halfspace is defined as $\mathbb{H}^k = \{(x_1, \dots, x_k) \in \mathbb{R}^k : x_k \geq 0\}$. Its boundary is the hyperplane $\partial\mathbb{H}^k = \{(x_1, \dots, x_k) \in \mathbb{R}^k : x_k = 0\}$.

Definition 1.14. A subset $X \subset \mathbb{R}^n$ is called a (smooth) manifold with boundary if each $x \in X$ has an open neighbourhood in X diffeomorphic to an open set in $\mathbb{H}^k$. The diffeomorphism and its inverse are, as before, known as the chart and the parameterisation.
The boundary ∂X of X consists of points which belong to the image of $\partial\mathbb{H}^k$ under some local parameterisation. The interior $\operatorname{Int} X$ of X is defined to be $X \setminus \partial X$. For consistency, we use the convention $\operatorname{Int} X = X$ when X is a manifold without boundary.

In example sheet, you'll show that the boundary behaves nicely in the sense that if $x \in \partial X$, then any parameterisation ψ around x would have $\psi^{-1}(x) \in \partial\mathbb{H}^k$.
How would tangent spaces and derivatives at boundary points work? Say we have a parameterisation ϕ that maps $0 \in U \subset \mathbb{H}^k$ to a boundary point of X. We can continue ϕ on a neighbourhood $\tilde{U} \ni 0$ open in $\mathbb{R}^k$ via some smooth $\tilde{\phi}$. Then we have

$$D\tilde{\phi}_0 = \lim_{\operatorname{Int} \mathbb{H}^k \cap \tilde{U} \ni a \to 0} D\tilde{\phi}_a = \lim_{\operatorname{Int} \mathbb{H}^k \cap U \ni a \to 0} D\phi_a$$

As $\phi, \tilde{\phi}$ both have continuous partials. So we can define $D\phi_0 = D\tilde{\phi}_0$ which is independent of the choice of $\tilde{\phi}$.
For a boundary point $x \in X \subset \mathbb{R}^n$, let ϕ be a local parameterisation near x. ϕ^{-1} extends to a smooth $\bar{\phi}$ on an neighbourhood W of x that is open in $\mathbb{R}^n$. Then $\bar{\phi} \circ \phi$ is the identity on $\operatorname{Int} \mathbb{H}^k \cap U$, so $D\bar{\phi}_{\phi(a)} \circ D\phi_a = I$ for any $a \in \operatorname{Int} \mathbb{H}^k \cap U$. Sending $a \to \phi^{-1}(x)$ shows that $D\bar{\phi}_x \circ D\phi_{\phi^{-1}(x)} = I$. In particular, $D\phi_{\phi^{-1}(x)}$ is injective. We then define the tangent space $T_x X$ of X at x to be the image of $D\phi_{\phi^{-1}(x)}$, which has all the nice properties we want.

Lemma 1.13. *If X is a manifold without boundary, and $f : X \to \mathbb{R}$ is smooth with 0 a regular value, then $\{x \in X : f(x) \geq 0\}$ is a smooth manifold with boundary. Furthermore, the boundary is given by $f^{-1}(0)$.*

Example 1.4. The unit ball $B^k = \{x \in \mathbb{R}^k : |x| \geq 1\}$ can be constructed as $\{x \in \mathbb{R}^k : f(x) \geq 0\}$ where $f(x) = 1 - |x|^2$.

Proof. The set $\{x \in X : f(x) > 0\}$ is open in X, hence is a manifold (without boundary) of the same dimension k, so it naturally inherits the charts of X. As for points $v \in f^{-1}(0)$, we reiterate the idea in the proof of Theorem 1.6 to conclude that $\{x : f(x) \geq 0\} \cap U$ is diffeomorphic to $\{\mathbb{R}_{\geq 0} \times \mathbb{R}^{k-1}\} \cap V$ for some neighbourhood V of $\{0\} \times T(v)$ where T is chosen such that $(\ker T) \cap (\ker Df_0) = \{0\}$. This gives the desired chart. ☐

Theorem 1.14 (Preimage Theorem for Manifolds with Boundary)**.** *Let $f : X \to Y$ be a smooth map where X is an m-manifold with boundary, Y an n-manifold without boundary and $m > n$. Suppose $y \in Y$ is a regular value for both f and $f|_{\partial X}$. Then $f^{-1}(y)$ is a smooth $(m-n)$-manifold with boundary $f^{-1}(y) \cap \partial X$.*

Proof. Since we are just gonna construct charts, we can just work locally. In the local context, f can be realised as a smooth map $V \to \mathbb{R}^n$ where $V \subset \mathbb{H}^m$ is open. For $z \in f^{-1}(y)$, if $z \in \operatorname{Int} \mathbb{H}^m$ then we are done by Theorem 1.6. If however $z \in \partial \mathbb{H}^m$, we continue f to $F : U \to \mathbb{R}^n$ where U is a neighbourhood of z open in $\mathbb{R}^m$. WLOG F has no critical points in U, then $F^{-1}(y) \subset U$ is a smooth manifold of dimension $m - n$.
Consider $\pi : F^{-1}(y) \to \mathbb{R}$ given by $(x_1, \ldots, x_m) \mapsto x_m$. We shall show that 0 is a regular value of π. Indeed, if we let $g = f|_{\partial V} = f \circ i$ where i is the inclusion $\partial V \hookrightarrow V$. If 0 is a critical value of π, then $D\pi_x = 0$ for some $x \in \pi^{-1}(0)$. We have $T_x F^{-1}(y) = \ker DF_x = \ker Df_x$. This implies that $Dg_x = Df_x|_{\mathbb{R}^{m-1}}$ has rank $n - 1 < m - 1$, so x is a critical point of $g = f|_{\partial X}$. This makes y a critical value of $f|_{\partial X}$, which is a contradiction.
Back to our proof. We now know that 0 is a regular value of π, then $F^{-1}(y) \cap \mathbb{H}^m = f^{-1}(y) \cap U = \{x \in F^{-1}(y) : \pi(x) \geq 0\}$ is a smooth manifold with boundary $\pi^{-1}(0)$ by the preceding lemma. This completes the proof. ☐

Remark. The very same proof works in the case where Y is a manifold with boundary as well.

One can generalise the preceding theorem and Theorem 1.11 to the following statement:

Theorem 1.15. *Let $f : X \to Y$ be a smooth map with X a manifold with boundary and Y a manifold without boundary. Suppose both f and $f|_{\partial X}$ are transversal to a submanifold (without boundary) Z of Y. Then $f^{-1}(Z)$ is a manifold with boundary $f^{-1}(Z) \cap \partial X$ and its codimension over X equals the codimension of Z over Y.*

Proof. Omitted but worth a try. ☐

1.6 Degree modulo 2

Definition 1.15. Two smooth maps $f, g : X \to Y$ are smoothly homotopic (written $f \sim g$) if there is a smooth $F : X \times [0,1] \to Y$ with $F(-,0) = f, F(-,1) = g$. The map F is called a smooth homotopy between f and g.

We often use the notation $f_t = F(-,t)$ to denote the one-parameter family of "intermediate" smooth maps when f is "deforming into" g via F. One can check that $\sim$ is an equivalence relation. An equivalence class of this relation is called a homotopy class.

Definition 1.16. Two diffeomorphisms $f, g : X \to Y$ are smoothly isotopic if there is a smooth homotopy $X \times [0,1] \to Y$ from f to g such that $f_t = F(-,t)$ is a diffeomorphism for all t.

Theorem 1.16 (Classification of Compact 1-Manifolds)**.** *Every compact* 1*-manifold is diffeomorphic to either* S^1 *or* $[0,1]$.

Proof. Look it up I guess. □

Corollary 1.17. *The boundary of any compact* 1*-manifold consists of an even number of points.*

Lemma 1.18 (Homotopy Lemma)**.** *Suppose* $f, g : X \to Y$ *are smoothly isotopic,* X *is boundaryless and compact, and* $\dim X = \dim Y$*. If* $y \in Y$ *is a regular value for both* f *and* g*, then* $|f^{-1}(y)| \equiv |g^{-1}(y)| \pmod 2$

Proof. We already know that both quantity are finite (from Corollary 1.7). Let $F : X \times [0,1] \to Y$ be the smooth homotopy taking f to g. If y is also a regular value of F, then $F^{-1}(y)$ is a compact 1-manifold with boundary $F^{-1}(y) \cap (X \times \{0,1\})$ by Theorem 1.14. By the preceding corollary we know that $|f^{-1}(y)| + |g^{-1}(y)| = |\partial F^{-1}(y)|$ is even, which implies the result.
What if y is a critical value of F? We make use of the Theorem 1.8 which shows that $w \mapsto |f^{-1}(w)|, w \mapsto |g^{-1}(w)|$ are locally constant if w is a regular value for the respective functions. Choose open neighbourhoods V, W of y such that V consists of regular values of f and W consists of regular values of g, and let U be the connected component of $V \cap W$ containing y. By Theorem 1.9, U cannot consist exclusively of critical values of F, so there is some $z \in U$ which is a regular value for F. Then $|f^{-1}(y)| \equiv |f^{-1}(z)| \equiv |g^{-1}(z)| \equiv |g^{-1}(y)| \pmod 2$ by the first part of the proof. □

Lemma 1.19 (Homogeneity Lemma)**.** *Let* X *be a smooth connected manifold with or without boundary. Let* $y, z \in \operatorname{Int} X$*, then there is a diffeomorphism* $h : X \to X$ *smoothly isotopic to* id_X *such that* $h(y) = z$.

Proof. Let $B \subset \mathbb{R}^n$ be the closed unit ball and $z \in \operatorname{Int} B$. It suffices to show that for any $z \in B$, there is a diffeomorphism $\mathbb{R}^n \to \mathbb{R}^n$ taking 0 to z that is smoothly isotopic to $\mathrm{id}_{\mathbb{R}^n}$ and fixes that all points outside B. If $z = 0$ then the statement is trivial. Otherwise, let $\psi : \mathbb{R}^n \to \mathbb{R}$ be smooth such that $\psi(x) > 0$ when $\|x\| < 1$ and $\psi(x) = 0$ when $\|x\| \geq 1$ (obtained, say, via the classical bump function construction). Set $c = z/\|z\|$ and consider the ODE $(\mathrm{d}/\mathrm{d}t)F(t,x) = \psi(F(t,x))c$ with the condition $F(0,x) = x$ for all x. By the celebrated Picard-Lindelöf theorem, the ODE has a unique smooth solution.
Write $F_t(x) = F(t,x)$. We have $F_{s+t}(x) = F_s \circ F_t(x)$ by the uniqueness of solution. Each F_t is then a diffeomorphism since it admits F_{-t} as its inverse. Consequently, each of them is smoothly isotopic to the identity via F itself (with some scaling of time). Sure enough $F_l(0) = z$ for some suitable l. This completes the proof. □

Remark. The family $\{F_t\}_{t \in \mathbb{R}}$ above is an example of a flow.

Theorem 1.20 (Degree modulo 2)**.** *Let* X *be a compact boundaryless manifold,* Y *a connected manifold (with ot without boundary) of the same dimension. Suppose* $f : X \to Y$ *is smooth and has regular values* y, z*. Then* $|f^{-1}(y)| = |f^{-1}(z)| \pmod 2$.

Proof. Take a diffeomorphism h of Y isotopic to id_Y such that $h(y) = z$. Then z is regular for $h \circ f$ and $h \circ f$ is smoothly homotopic to f. By Lemma 1.18, we have $|f^{-1}(y)| \equiv |(h \circ f)^{-1}(z)| \equiv |f^{-1}(z)| \pmod 2$. □

Definition 1.17. The value $|f^{-1}(y)| \bmod 2$ for any regular value y of $f : X \to Y$ is known as the degree modulo 2 of f, written $\deg_2 f$.

Lemma 1.21. $\deg_2 f$ *is a homotopy invariant.*

Proof. If g is smoothly homotopic to f, then Theorem 1.9 shows that there is some $y \in Y$ that is regular for both f and g. The result then follows from Lemma 1.18. □

Example 1.5. If X is a compact manifold without boundary, then $\deg_2 \mathrm{id}_X = 1$ and $\deg_2 c = 0$ for any constant c. Consequently, the identity cannot be homotopic to a constant map.

Corollary 1.22. *There is no smooth* $f : B^{k+1} \to S^k = \partial B^{k+1}$ *that fixes* S^k.

Proof. Suppose for the sake of contradiction that f exists. Consider $F : S^k \times [0, 1] \to S^k$ given by $(x, t) \mapsto f(tx)$ which is a homotopy between the constant map and the identity, whose nonexistence is already known. □

Theorem 1.23 (Smooth Brouwer's Fixed Point Theorem)**.** *Any smooth* $f : B^k \to B^k$ *has a fixed point.*

Proof. Suppose we have some $f : B_k \to B_k$ with no fixed point, then we shall construct from it some smooth $g : B^k \to S^{k-1}$ that fixes S^{k-1} which will provide the desired contradiction. Indeed, for $x \in B^k$, we can just let $g(x) \in S^{k-1}$ be the intersection of the ray from $f(x)$ to x and S^{k-1}. □

Remark. The theorem is still true if we replace "smooth" by just "continuous". This can either be done using algebraic topology or from the smooth version via Stone-Weierstrass.

We now want to generalise this notion of degree modulo 2 from smooth maps to submanifolds. Suppose X is compact and boundaryless, Z a boundaryless submanifold of a manifold Y, $f \pitchfork Z$ and $\dim X + \dim Z = \dim Y$. $f^{-1}(Z) \subset X$ is then a closed submanifold of dimension zero, hence (by compactness) a finite set of points.

Definition 1.18. The intersection number modulo 2 of f with respect to Z is $I_2(f, Z) = |f^{-1}(Z)| \bmod 2$.

Lemma 1.24 (Analogue to Lemma 1.18)**.** *If* f_0, f_1 *are transversal to* Z *and are homotopic, then* $I_2(f_0, Z) = I_2(f_1, Z)$*.*

Sketch of proof. Let $F : X \times [0, 1] \to Z$ be the homotopy taking f_0 to f_1. Write $F(x, t) = f_t(x)$. If $F \pitchfork Z$, then $F^{-1}(Z)$ is a 1-manifold with boundary $F^{-1}(Z) \cap (X \times \{0, 1\})$. Otherwise, we perturb F and make use of the genericity property of transversality. □

Remark. For general smooth f, we can define $I_2(f, Z)$ to be $I_2(g, Z)$ with g homotopic to F and $g \pitchfork Z$. It takes another genericity argument to see that such g exists, but once we know that, then the choice of g does not matter by the preceding lemma.

Suppose we have an inclusion $i : X \hookrightarrow Y$ with X compact such that $i \pitchfork Z$ (which we often write instead as $X \pitchfork Z$). Then we define the intersection number modulo 2 of X, Z to be $I_2(X, Z) = I_2(i, Z) = |X \cap Z| \bmod 2$. But of course we can drop the condition $X \pitchfork Z$ by the above remark.
If $I_2(X, Z) = 1$, then we can deduce that X cannot become a submanifold disjoint from Z via homotopy.

Example 1.6. Take $Y = S^1 \times S^1, X = S^2 \times \{*\}, Z = \{*\} \times S^1$, then $I_2(X, Z) = 1$ and one cannot stretch X to make it disjoint from Z.

When $2 \dim X = \dim Y$, one can consider $I_2(X, X)$, the self-intersection number modulo 2.

Example 1.7. Take Y to be the Möbius band and X to the central circle. Then by perturbing X a bit so that it becomes transversal with its original self, we conclude that $I_2(X, X) = 1$.

Another application of degree modulo 2 is the following:

Theorem 1.25 (Jordan-Brouwer Seperation Theorem)**.** *Suppose $X \subset \mathbb{R}^n$ is a compact connected hypersurface (i.e. manifold without boundary with codimension 1). Then $\mathbb{R}^n \setminus X$ has two connected components D_0, D_1 where $\bar{D}_1$ is compact and $\partial \bar{D}_1 = X$.*

Moreover, if X is orientable (which we will define later), we can make a consistent choice of normal vector that always points to D_0 ("the outside").

Brief sketch of proof. Take $z \notin X$ and consider $u : X \to S^{n-1}$ to be the map $u_z(x) = (x - z)/\|x - z\|$ and $w_2(z) = \deg_2 u_z$ which can be interpreted as some sort of a winding number modulo 2 of X around z. Then we just take $D_0 = \{z \in \mathbb{R}^n \setminus X : w_2(z) = 0\}$ and $D_1 = \{z \in \mathbb{R}^n \setminus X : w_2(z) = 1\}$ which (after some justifications) work. □

1.7 Abstract Manifolds and Whitney's Theorem

Often, differential geometry is done more abstractly (or indeed,, instrinstically) with a generalised notion of manifolds.

Definition 1.19. An abstract n-manifold X is a Hausdorff, second countable topological space equipped with an atlas of charts $\{(\phi_{U_i}, U_i) : i \in I\}$ where:
1. Each U_i is open in X and $\phi_{U_i} : U_i \to \mathbb{R}^n$ is a homeomorphism onto its image.
2. $\bigcup_{i \in I} U_i = X$.
3. $\phi_{U_i} \circ \phi_{U_j}^{-1}|_{\phi_{U_j}(U_i \cap U_j)}$ is a diffeomorphism (onto its image) for any $i, j \in I$.

However, it turns out that this notion does not give any more manifolds than our original definition does. Indeed, we can embed any n-manifold into $\mathbb{R}^N$ for very big N with bump functions. In fact, we can reduce the dimension of the ambient space we embed it in to a reasonably small number.

Theorem 1.26 (Whitney's Embedding Theorem)**.** *An abstract n-manifold can be embedded into $\mathbb{R}^{2n}$.*

This is very hard to prove. However, a weaker version is possible to attempt at this stage.

Theorem 1.27. *An abstract n-manifold can be embedded into $\mathbb{R}^{2n+1}$.*

Very brief sketch of proof. Embed the manifold into $\mathbb{R}^N$ for big N and project it generically $\mathbb{R}^{2n}$ so that it becomes an immersion. One can then embed it in $\mathbb{R}^{2n+1}$ by resolving the self-intersections via small perturbations in the extra dimension. $\square$

In fact, we can do something even stronger than Whitney's theorem.

Theorem 1.28. *Any compact orientable smooth n-manifold can be embedded in $\mathbb{R}^{2n-1}$ for $n > 1$.*

2 Length, Area and Curvature

2.1 Arc Length, Curvature and Torsion of Curves

Definition 2.1. A curve in a manifold X is a smooth function $\alpha : I \to X$ where I is an interval in $\mathbb{R}$. The velocity of α at $t \in I$ is $\dot{\alpha}(t) = D\alpha_t(1) \in T_{\alpha(t)}X$. α is regular if it is an immersion, i.e. has never-vanishing velocity.

Definition 2.2. Suppose $\alpha : I \to \mathbb{R}^N$ is a regular curve in $\mathbb{R}^N$ and $t_0, t \in I$, then the arc length of the curve from t_0 to t is

$$s(t) = s_{t_0}(t) = \int_{t_0}^{t} |\dot{\alpha}(t)| \, \mathrm{d}t$$

If $I = [a, b]$ is a closed interval, then the total length of α is

$$\ell(\alpha) = \int_{a}^{b} |\dot{\alpha}(t)| \, \mathrm{d}t$$

Definition 2.3. A curve α is parameterised by arc length if it has unit speed, i.e. $|\dot{\alpha}(t)| = 1$ for all t.

Any regular curve can be refined so that it is parameterised by arc length. Indeed, for such an α and fixed t_0, the function $t \mapsto s_{t_0}(t)$ is smooth and strictly increasing, hence has a smooth inverse. $\beta : s \mapsto \alpha(s_{t_0}^{-1}(s))$, which represents the same curve as α, is then parameterised by arc length.

Definition 2.4. Let $\alpha : I \to \mathbb{R}^n$ be a regular curve parameterised by arc length. The tangent vector to α at $s \in I$ is $t(s) = \dot{\alpha}(s)$. The curvature of it is $k(s) = |\dot{t}(s)| = |\ddot{\alpha}(s)|$. If $k(s) \neq 0$, the unit normal $n(s)$ of α at s is defined by $\ddot{\alpha}(s) = k(s)n(s)$.

Differentiating the both sides of $|\dot{\alpha}|^2 = 1$ gives $\langle \dot{\alpha}, \ddot{\alpha} \rangle = 0$ which means that $t(s)$ and $n(s)$ are orthogonal.

Definition 2.5. The auxiliary plane of α at $s \in I$ is the plane spanned by $t(s)$ and $n(s)$.

When the curve resides in $\mathbb{R}^3$, we can add in one more vector associated locally with the curve.

Definition 2.6. If $\alpha : I \to \mathbb{R}^3$ is a regular curve in $\mathbb{R}^3$ (again parameterised by arc length), we define the binormal vector to it at $s \in I$ to be $b(s) = t(s) \times n(s)$

Consequently, fixing any s at which α has nonzero curvature, $t(s), n(s), b(s)$ form a set of orthonormal basis for $\mathbb{R}^3$.

Definition 2.7. When $k(s) \neq 0$, the orthonormal triple $(t(s), n(s), b(s))$ is the Frenet trihedron of α at s.

We have $\dot{b} = \dot{t} \times n + t \times \dot{n} = t \times \dot{n}$ and $\langle b, \dot{b} \rangle = 0$ (from differentiating $\|b\|^2 = 1$), so $\dot{b}$ is orthogonal to both t and b, hence is a scalar multiple of n.

Definition 2.8. The torsion $\tau(s)$ of α at $s \in I$ is defined by the equation $\dot{b}(s) = \tau(s)n(s)$.

We have $n = b \times t$, so $\dot{n} = \dot{b} \times t + b \times \dot{t} = \tau n \times t + kb \times n = -\tau b - kt$. We conclude the Frenet formulae

$$\begin{pmatrix} \dot{t} \\ \dot{n} \\ \dot{b} \end{pmatrix} = \begin{pmatrix} & k & \\ -k & & -\tau \\ & \tau & \end{pmatrix} \begin{pmatrix} t \\ n \\ b \end{pmatrix}$$

Why do we care? Turns out, this allows the proof of the following classification theorem.

Theorem 2.1 (Fundamental Local Structure of Curves in $\mathbb{R}^3$)**.** *Let I be a finite interval. Given smooth functions $k(s), \tau(s)$ for $s \in I$ with $k(s) > 0$, there exists a regular curve $\alpha : I \to \mathbb{R}^3$ parameterised by arc length whose curvature is given by $k(s)$ and whose torsion is $\tau(s)$. Moreover, such α is unique up to an isometry of $\mathbb{R}^3$.*

Proof. The Frenet formulae are a ODE for $(t, n, b) \in \mathbb{R}^{3\times 3}$. Picard-Lindelöf then gives the existence and uniqueness of a solution given an initial condition (t_0, n_0, b_0). We still need to show that whenever (t_0, n_0, b_0) is an orthonormal set, so is $(t(s), n(s), b(s))$ for all s. But this follows directly from the fact that the matrix

$$\begin{pmatrix} & k & \\ -k & & -\tau \\ & \tau & \end{pmatrix}$$

is antisymmetric. Integrating $t(s)$ obtained this way gives the result. □

For plane curves $\alpha : I \to \mathbb{R}^2 \subset \mathbb{R}^3$, we can in fact give a sign to the curvature by modifying our definition a little bit: Instead of defining $n(s)$ after obtaining $k(s)$ as $|\dot{t}(s)|$, we define $n(s)$ by requiring $\{t(s), n(s)\}$ to have the same orientation as a chosen standard basis $\{e_1, e_2\}$ (usually via the chosen orientation of the plane). The signed curvature can then be defined with the (very same) equation $\dot{t}(s) = k(s)n(s)$.

2.2 The Isoperimetric Inequality in the Plane

Let $\Omega \subset \mathbb{R}^2$ be a domain (a connected open set with compact closure) whose boundary $\partial\Omega$ is assumed be a connected C^1 curve without self-intersection. Let $A(\Omega)$ be the area of Ω.

Theorem 2.2 (Isoperimetric Inequality)**.** $\ell(\partial\Omega)^2 \geq 4\pi A(\Omega)$ *with equality iff* Ω *is a disc.*

Lemma 2.3 (Wirtinger's Inequality)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a periodic C^1 function with period L. Suppose that*

$$\int_0^L f(t)\,\mathrm{d}t = 0$$

Then

$$\int_0^L |f'(t)|^2\,\mathrm{d}t \geq \frac{4\pi^2}{L^2}\int_0^L |f(t)|^2\,\mathrm{d}t$$

with equality iff there are constants $a_{\pm 1}$ with $f(t) = a_{-1}e^{-2\pi it/L} + a_1 e^{2\pi it/L}$.

Proof. Take the Fourier series

$$f(t) = \sum_{k=-\infty}^{\infty} a_k e^{2\pi ikt/L}, f'(t) = \sum_{k=-\infty}^{\infty} b_k e^{2\pi ikt/L}$$

which are absolutely and uniformly convergent due as f is C^1. $b_0 = 0$ as f is L-periodic; $a_0 = 0$ due to the integral assumption on f. We also have

$$b_k = \frac{1}{L}\int_0^L f'(t)e^{-2\pi ikt/L}\,\mathrm{d}t = \frac{2\pi ik}{L}\frac{1}{L}\int_0^L f(t)e^{-2\pi ikt/L} = \frac{2\pi ik}{L}a_k$$

via integration by part. Then, by Parseval's theorem,

$$\begin{aligned}\int_0^L |f'(t)|^2\,\mathrm{d}t &= L\sum_{k\neq 0}|b_k|^2 = \frac{4\pi^2}{L}\sum_{k\neq 0}k^2|a_k|^2 \\ &\geq \frac{4\pi^2}{L}\sum_{k\neq 0}|a_k|^2 = \frac{4\pi^2}{L^2}\int_0^L |f(t)|^2\,\mathrm{d}t\end{aligned}$$

Equality holds iff $a_k = 0$ for all $|k| > 1$, i.e. f has the said form. □

Proof of Theorem 2.2. Parameterise $\partial\Omega$ by arc length. Consider the vector field $X(x, y) = (x, y)$ in $\mathbb{R}^2$. By translation, we can assume WLOG that

$$\int_{\partial\Omega} X \cdot \mathrm{d}s = 0$$

Let $n : \partial\Omega \to \mathbb{R}^2$ be the unit normal pointing out from Ω. Divergence theorem gives

$$\begin{aligned}2A(\Omega) &= \int_\Omega \nabla\cdot X\,\mathrm{d}A = \int_{\partial\Omega}\langle X, n\rangle\,\mathrm{d}s \leq \int_{\partial\Omega}\|X\|\,\mathrm{d}s \\ &\leq \left(\int_{\partial\Omega}|X|^2\,\mathrm{d}s\right)^{1/2}\left(\int_{\partial\Omega}\mathrm{d}s\right)^{1/2} = \left(\int_{\partial\Omega}|X|^2\,\mathrm{d}s\right)^{1/2}\ell(\partial\Omega)^{1/2}\end{aligned}$$

by Cauchy-Schwartz inequality. The components of the vector field $X(s) = (x(s), y(s))$ on $\partial\Omega$ are C^1 by assumption, and are periodic of period $\ell(\partial\Omega)$. Wirtinger's inequality then gives

$$\begin{aligned}\left(\int_{\partial\Omega}|X|^2\,\mathrm{d}s\right)^{1/2} &= \left(\int_0^{\ell(\partial\Omega)}(x(s)^2 + y(s)^2)\,\mathrm{d}s\right)^{1/2} \\ &\leq \left(\frac{\ell(\partial\Omega)^2}{4\pi^2}\int_0^{\ell(\partial\Omega)}(x'(s)^2 + y'(s)^2)\,\mathrm{d}s\right)^{1/2} = \frac{\ell(\partial\Omega)^{3/2}}{2\pi}\end{aligned}$$

Combining this with the previous inequality gives the inequality. If the equality is achieved, then for example the equality condition for integral Cauchy-Schwartz commands $|X|$ to be constant on $\partial\Omega$, which means that Ω is a disc. $\square$

Remark. One can generalise this from the plane to any smooth surface. An example of such is the following statement: If S is orientable and has constant curvature K (which again will be defined later) and $\Omega \subset S$ is a domain on S, then $4\pi A(\Omega)\chi(\Omega) \leq \ell(\partial\Omega)^2 + KA(\Omega)^2$ where $A(\Omega)$ is the area of Ω on S and $\chi(\Omega)$ is the Euler characteristic of Ω.

2.3 First Fundamental Form and Area

In this section, we only care about smooth boundaryless 2-manifolds ("surfaces") embedded in $\mathbb{R}^3$.

Definition 2.9. Suppose $S \subset \mathbb{R}^3$ is a surface. The positive definite quadratic form I_p on $T_p(S)$ given by $\mathrm{I}_p(w) = \langle w, w\rangle = \|w\|^2$ is called the first fundamental form of S.

This is a special case of the notion of Riemannian metric for general manifolds.

Definition 2.10. If X is a smooth manifold, a Riemannian metric on X is a smooth family $\mathfrak{g} = \{\mathfrak{g}_p\}_{p\in X}$ where each $\mathfrak{g}_p$ is an inner product on T_pX.

If $f : X \to Y$ is an immersion and Y has a Riemannian metric $\mathfrak{g}$, then we have a natually induced Riemannian metric (the "pullback metric") $\mathfrak{h}$ on X given by $\mathfrak{h}_p(w) = \mathfrak{g}_{f(p)}(Df_p(w))$. So the first fundamental form on a surface $S \subset \mathbb{R}^3$ is simply the pullback metric of the standard inner product on $\mathbb{R}^3$ under the inclusion $S \hookrightarrow \mathbb{R}^3$.

Definition 2.11. Surfaces S_1, S_2 are isometric if there is a diffeomorphism $f : S_1 \to S_2$ such that Df_p is a linear isometry between T_pS_1 and $T_{f(p)}S_2$ (equipped with their respective first fundamental form as inner product) for all $p \in S_1$.

We've done everything so far in a coordinate-free way, which is wonderful conceptually, but doesn't quite help with calculation. Let's now see how the first fundamental form looks like if we had a local coordinate. Let $\phi : U \to S$ be a local parameterisation near $p \in \phi(U) \subset S$ where $S \subset \mathbb{R}^3$ is a surface. Let (u, v) be the coordiate on $U \subset \mathbb{R}^2$, then $\phi_u(u,v) = D\phi_{(u,v)}(1,0), \phi_v(u,v) = D\phi_{(u,v)}(0,1)$ would be a basis for $T_{\phi(u,v)}S$. Suppose $p = \phi(u,v)$, we set $E = \langle\phi_u(u,v), \phi_u(u,v)\rangle, F = \langle\phi_u(u,v), \phi_v(u,v)\rangle, G = \langle\phi_v(u,v), \phi_v(u,v)\rangle$. Given $w \in T_pS$, pick a curve α such that $\alpha(0) = p, \dot\alpha(0) = w$ with $\alpha(t) = \phi(u(t), v(t))$ for $t \in (-\epsilon, \epsilon)$ for some $\epsilon > 0$. Then $\dot\alpha = \phi_u\dot u + \phi_v\dot v$ and therefore $I_p(w) = I_p(\dot\alpha(0)) = E\dot u(0)^2 + 2F\dot u(0)\dot v(0) + G\dot v(0)^2$. We reflect this phenomenon by writing I_p as $E\,\mathrm{d}u^2 + 2F\,\mathrm{d}u\,\mathrm{d}v + G\,\mathrm{d}v^2$ given the coordinates (u, v) from a local parameterisation.

Example 2.1. Let S be the torus obtained as a surface of revolution. One of its local parameterisations can be given by $\phi(u,v) = ((a + r\cos u)\cos v, (a + r\cos u)\sin v, r\sin u)$ where r is the thickness of it. Then by calculation we have $E = r^2, F = 0, G = (a + r\cos u)^2$.

For a curve $\alpha : I = [0,1] \to S \subset \mathbb{R}^3$ given by $\alpha(t) = \phi(u(t), v(t))$, the formula $\dot\alpha = \phi_u \dot u + \phi_v \dot v$ then gives

$$\ell(\alpha) = \int_0^1 \langle \alpha, \dot\alpha \rangle^{1/2} \, \mathrm{d}t = \int_0^1 (E\dot u^2 + 2F\dot u \dot v + G\dot v^2)^{1/2} \, \mathrm{d}t$$

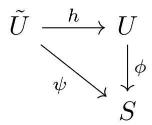
Also, $\|\phi_u \times \phi_v\| = \sqrt{EG - F^2}$ by the identity $\langle a, b\rangle^2 + \|a \times b\|^2 = \|a\|^2\|b\|^2$. Suppose a domain $\Omega \subset S$ is contained in the image of a local parameterisation $\phi : U \to S$. We want to define its area.

Lemma 2.4. *The integral*

$$\int_{\phi^{-1}(\Omega)} \|\phi_u \times \phi_v\| \, \mathrm{d}u \, \mathrm{d}v$$

does not depend on the choice of ϕ.

Proof. Say we have another parameterisation $\psi : \tilde U \to S$ whose image contains Ω. By restriction we can assume WLOG that $\psi(\tilde U) = \phi(U)$. Let $h = \phi^{-1} \circ \psi$ and draw a commutative diagram that isn't really useful per se:



h is actually just a change of variable: If we let $(u, v), (\tilde u, \tilde v)$ be the coordinates on $U, \tilde U$ respectively, then $h(\tilde u, \tilde v) = (u(\tilde u, \tilde v), v(\tilde u, \tilde v))$. There is no better thing to do than to bring out the Jacobian $J = \det(\partial(u, v)/\partial(\tilde u, \tilde v)) = \det Dh$. As $\psi = \phi \circ h$, we have $D\psi_a = D\phi_{h(a)} \circ Dh_a$, so $\|\psi_{\tilde u} \times \psi_{\tilde v}\|_a = \|\phi_u \times \phi_v\|_{h(a)} |J(a)|$. The change-of-variable formula from vector calculus gives the result. □

Definition 2.12. Let $\Omega \subset S$ be a bounded domain contained in the image of a local parameterisation $\phi : U \to S$. The area of Ω is

$$A(\Omega) = \int_{\phi^{-1}(\Omega)} \|\phi_u \times \phi_v\| \, \mathrm{d}u \, \mathrm{d}v = \int_{\phi^{-1}(\Omega)} \sqrt{EG - F^2} \, \mathrm{d}u \, \mathrm{d}v$$

Remark. 1. For bounded domains which are not contained in the image of a single parameterisation, we can still define its area by what's known as a partition of unity. We almost never need this. It turns out that we can always find a parameterisation whose image contains the domain except for possibly a finite number of curves (which don't contribute any area).
2. For a continuous $f : S \to \mathbb{R}$, we often write

$$\int_S f \, \mathrm{d}A = \int_U f(\phi(u, v)) \sqrt{EG - F^2} \, \mathrm{d}u \, \mathrm{d}v$$

where $\phi : U \to S$ is a parameterisation of S (after removing a finite number of curves). This is independent of the choice of ϕ by the exact same argument as above and the "area differential" $\mathrm{d}A = \sqrt{EG - F^2} \, \mathrm{d}u \, \mathrm{d}v$ is an example of a Riemannian measure on S.

Example 2.2. For the torus we considered earlier, we have $\sqrt{EG - F^2} = r(a + r\cos u)$, so it has total area $4\pi^2 ra$.

2.4 Gauss Map and Second Fundamental Form

Given a local parameterisation $\phi : U \to S$, we can choose a unit normal vector $\mathcal{N}(p)$ at any $p \in \phi(U)$ with the property that $\mathcal{N}(p) \perp T_pS$ via $\mathcal{N}(p) = \phi_u \times \phi_v / \|\phi_u \times \phi_v\|$ (evaluated at p). Clearly, any unit normal to S at p (i.e. unit vector perpendicular to T_pS) must be either $\mathcal{N}(p)$ or $-\mathcal{N}(p)$. This is a smooth function $\phi(U) \to S^2$, but there is no guarantee that we can extend this smoothly to the whole of S. Also, this does depend on the choice of ϕ since we can swap the coordinates on U and get a unit normal with reversed sign. These are no big deal – we just turn our attention to those surfaces where these problems are insignificant.

Definition 2.13. A smooth field of unit normal vectors on a surface S is a smooth map $\mathcal{N} : S \to S^2 \subset \mathbb{R}^3$ such that $\mathcal{N}(p) \perp T_pS$. A surface $S \subset \mathbb{R}^3$ is orientable if it admits such a choice $\mathcal{N}$ of smooth field of unit normal vectors (an "orientation"). $\mathcal{N}$ is called the Gauss map of S oriented as such.

Our previous definition of $\mathcal{N}$ shows that the image of any local parameterisation is orientable. Consequently, when defining and analysing local properties of a manifold, we can assume WLOG that it (locally) admits such a choice of $\mathcal{N}$. This also means that orientability is a global property – the surface can still be non-orientable even if there must be an orientable patch around each of its points.

Example 2.3. The Möbius band is non-orientable.

Remark. This definition suffers from the problem that it can only generalise to manifolds with codimension 1 over the ambient space we embed it in. We can instead define orientation as the consistent choice of ordered bases (u_p, v_p) on T_pS for each $p \in S$. This is the same as the unit normal definition in the codimension 1 case since we can then choose the unit normals by requiring $(u_p, v_p, \mathcal{N}(p))$ to be a right-handed (or left-handed) basis for $\mathbb{R}^3$.

Note that $T_pS = T_{\mathcal{N}(p)}S^2 = \mathcal{N}(p)^\perp \subset \mathbb{R}^3$, so $D\mathcal{N}_p : T_pS \to T_{\mathcal{N}(p)}S^2$ can be think of an endomorphism $T_pS \to T_pS$.

Lemma 2.5. *$D\mathcal{N}_p : T_pS \to T_pS$ is self-adjoint with respect to* I_p.

Proof. Let $\phi : U \to S$ be a parameterisation around p and suppose $\alpha(t) = \phi(u(t), v(t))$ is a curve in $\phi(U)$ with $\alpha(0) = p$. Then

$$\dot{u}(0)\mathcal{N}_u + \dot{v}(0)\mathcal{N}_v = \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{N}(u(t), v(t))\bigg|_{t=0} = D\mathcal{N}_p(\dot{\alpha}(0)) = D\mathcal{N}_p(\dot{u}(0)\phi_u + \dot{v}(0)\phi_v)$$

So $D\mathcal{N}_p(\phi_u) = \mathcal{N}_u, D\mathcal{N}_p(\phi_v) = \mathcal{N}_v$. $\{\phi_u, \phi_v\}$ is a basis for T_pS, so it remains to show that $\langle \mathcal{N}_u, \phi_v \rangle = \langle \phi_u, \mathcal{N}_v \rangle$. This follows from differentiating $\langle \mathcal{N}, \phi_u \rangle = \langle \mathcal{N}, \phi_v \rangle = 0$ (and the symmetry of mixed partials). □

Definition 2.14. The quadratic form on T_pS given by $\mathrm{II}_p(w) = -\langle D\mathcal{N}_p(w), w \rangle$ is known as the second fundamental form of S at p.

Let $\alpha : (-\epsilon, \epsilon) \to S$ be an arc length parameterised curve with $\alpha(0) = p \in S$. We have $\langle \mathcal{N}, \dot{\alpha} \rangle = 0$. We can differentiate this with respect to the arc length parameter to get $\langle \mathcal{N}(\alpha(s)), \ddot{\alpha}(s) \rangle = -\langle D\mathcal{N}_{\alpha(s)}(\dot{\alpha}(s)), \dot{\alpha}(s) \rangle$, which gives

$$\mathrm{II}_p(\dot{\alpha}(0)) = -\langle D\mathcal{N}_p(\dot{\alpha}(0)), \dot{\alpha}(0) \rangle = \langle \mathcal{N}(p), \ddot{\alpha}(0) \rangle = \langle \mathcal{N}(p), kn \rangle$$

where n is the normal to α at p and k is the curvature there.

Definition 2.15. $k_{\mathrm{n}} = \langle \mathcal{N}, kn \rangle$ is the normal curvature of α at p.

Note that the sign of this is dependent on the choice of orientation for S. Also, it depends on α only through the tangent vector $\dot{\alpha}(0) \in T_pS$.
As one might expect from any geometry course, there is another way to define normal curvature given $v \in T_pS$. Let $V \subset \mathbb{R}^3$ be a plane through p containing v and $\mathcal{N}_p$. $C = V \cap S$ is known as the normal section of S at p along V. Generalised preimage theorem shows that C is locally a smooth 1-manifold since V and S have to be transversal. Take a local arc length parameterisation $\alpha : (-\epsilon, \epsilon) \to S$ so that $\alpha(0) = p, \dot{\alpha}(0) = v/\|v\|$, then $\ddot{\alpha}(0) = \pm k\mathcal{N}(p)$, so $k_{\mathrm{n}}(p)$ of v is just (up to a sign) the curvature of α at p.
Back to $D\mathcal{N}$. Recall that any self-adjoint map has an orthonormal basis of eigenvectors.

Corollary 2.6. *There is an orthonormal basis* $\{e_1, e_2\}$ *of* T_pS *and real numbers* $k_1, k_2, k_1 \geq k_2$ *such that* $D\mathcal{N}_p(e_i) = -k_i e_i$. *Moreover,* k_1, k_2 *are the maximum and minimum possible values of* $\mathrm{II}_p|_{\{v \in T_pS : \mathrm{I}_p(v) = 1\}}$.

Definition 2.16. k_1, k_2 are called the principal curvature at p and e_1, e_2 the principal directions.

Remark. k_i are, consequently, extremal values of normal curvatures.

Definition 2.17. The Gaussian curvature $K(p)$ of S at p is $\det(D\mathcal{N}_p) = k_1k_2$. The mean curvature $H(p)$ of S at p is $-\operatorname{tr}(D\mathcal{N}_p)/2 = (k_1 + k_2)/2$.

Remark. 1. k_1, k_2 are then the roots of $k^2 - 2Hk + K = 0$.
2. Gaussian curvature does not depend on the specific orientation but mean curvature does (and it flips sign if one flips the orientation).

Definition 2.18. A point $p \in S$ is elliptic if $K(p) > 0$, hyperbolic if $K(p) < 0$, and parabolic if $K(p) = 0$.
p is called planar if $D\mathcal{N}_p = 0$.

Example 2.4. Every $p \in S^2$ is elliptic, so is every $p \in \{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2\}$. Every $p \in \{(x, y, z) \in \mathbb{R}^3 : z = x^2 - y^2\}$ is hyperbolic. Every $p \in \{(x, y, z) \in \mathbb{R}^3 : 1 = x^2 + y^2\}$ is parabolic.
The point $0 \in S = \{(x, y, z) \in \mathbb{R}^3 : z = x^2 - 3y^2x\}$ (the "monkey saddle") is actually planar. Indeed, we have $T_0S = \{z = 0\}$. Also, if we set $x = r\cos\theta, y = r\sin\theta$, then $z = r^3\cos\theta$, so we see that the every normal curvature vanishes at 0, so $0 \in S$ has to be planar.

Definition 2.19. A point $p \in S$ is umbilic if $k_1 = k_2$ at p.

You'll show in example sheet that if all points on a connected surface are umbilic, then it's contained in a sphere or a plane.

2.5 Second Fundamental Form in Local Coordinates

Definition 2.20. Let ϕ be a local parameterisation about $p \in S$. We define $e = \langle \mathcal{N}, \phi_{uu} \rangle = -\langle \mathcal{N}_u, \phi_u \rangle, f = \langle \mathcal{N}, \phi_{uv} \rangle = -\langle \mathcal{N}_u, \phi_v \rangle = -\langle \mathcal{N}_v, \phi_u \rangle, g = \langle \mathcal{N}, \phi_{vv} \rangle = -\langle \mathcal{N}_v, \phi_v \rangle$.

Lemma 2.7. *In local cooridnates* (u, v), II *is given by* $e\,\mathrm{d}u^2 + 2f\,\mathrm{d}u\,\mathrm{d}v + g\,\mathrm{d}v^2$.

Proof. If α is a curve on S with $\alpha(0) = p$, then

$$\begin{aligned}\mathrm{II}_p(\dot\alpha(0)) &= -\langle D\mathcal{N}_p(\dot\alpha(0)), \dot\alpha(0)\rangle = -\langle D\mathcal{N}_p(\phi_u\dot u + \phi_v\dot v), \phi_u\dot u + \phi_v\dot v\rangle \\ &= -\langle \mathcal{N}_u\dot u + \mathcal{N}_v\dot v, \phi_u\dot u + \phi_v\dot v\rangle = e\dot u^2 + 2f\dot u\dot v + g\dot v^2\end{aligned}$$

as desired. □

Lemma 2.8.

$$K = \frac{eg - f^2}{EG - F^2}, H = \frac{1}{2}\frac{eG - 2fF + gE}{EG - F^2}$$

Proof. Write

$$\begin{cases} D\mathcal{N}_p(\phi_u) = \mathcal{N}_u = a_{11}\phi_u + a_{21}\phi_v \\ D\mathcal{N}_p(\phi_v) = \mathcal{N}_v = a_{12}\phi_u + a_{22}\phi_v \end{cases}$$

where everything is evaluated at p. For both of these two equations, we take their inner product with ϕ_u and ϕ_v to get

$$\begin{pmatrix} -e & -f \\ -f & -g \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

But the matrix of $D\mathcal{N}_p$ in the basis $\{\phi_u, \phi_v\}$ is (a_{ij}), so $K = \det D\mathcal{N}_p = \det(a_{ij}) = (eg - f^2)/(EG - F^2)$ and $H = -\operatorname{tr}(D\mathcal{N}_p)/2 = -\operatorname{tr}(a_{ij})/2 = (eG - 2fF + gE)/(2(EG - F^2))$ □

Example 2.5. Take again the torus locally parameterised as $\phi(u, v) = ((a + r\cos u)\cos v, (a+r\cos u)\sin v, r\sin u)$. We have already calculated $E = r^2, F = 0, G = (a + r\cos u)^2$. We can also calculate $e = r, f = 0, g = (u + v\cos u)\cos u$. So $K = \cos u/(r(a+r\cos u))$. Consequently, points belonging to the "inside" of the torus ($\pi/2 < u < 3\pi/2$) are hyperbolic, points belonging to the "outside" of the torus ($0 < u < \pi/2$ or $3\pi/2 < u < 2\pi$) are elliptic, and the rest are parabolic.

There is a Taylor series interpretation of the second fundamental form as well. Paramaterise $\phi : U \to S$ around p (say $\phi(0) = p$). We expand the Taylor series of ϕ around 0 to get

$$\phi(h, k) - p = h\phi_u + k\phi_v + \frac{1}{2}(\phi_{uu}h^2 + 2\phi_{uv}hk + \phi_{vv}k^2) + \text{higher order terms}$$

If we take the inner product of $\mathcal{N}(p)$ with both sides of the equation, we get

$$\langle(\phi(h, k) - p), \mathcal{N}(p)\rangle = \frac{1}{2}(eh^2 + 2fhk + gk^2) = \frac{1}{2}\mathrm{II}_p(h, k)$$

up to second order terms. So II_p measures the normal displacement from the tangent plane in the direction (h, k). Similarly, for first fundamental form we have (by expanding $\phi(h, k)$ to first order)

$$\|\phi(h, k) - p\|^2 = Eh^2 + 2Fhk + Gk^2 = \mathrm{I}_p(h, k)$$

up to second order.
Consequently, if $K > 0$, then the two principal curvatures have the same sign, so the second fundamental form is either positive definite or negative definite and S stays (locally) on one side of T_pS. Equivalently, all the normal curvatures have the same sign.
If $K < 0$, the two principal curvatures have opposite signs and the second fundamental form is non-degenerate and indefinite.

2.6 Theorema Egregium

Theorem 2.9 (Theorema Egregium (Gauss, 1827))**.** *The Gaussian curvature of a surface is invariant under isometries.*

Proof. Isometries preserve the first fundamental form, so it suffices to show that K can be expressed purely in terms of E, F, G. Let $\phi : U \to S$ be a parameterisation. At each point of $\phi(U)$, we have a basis of $\mathbb{R}^3$ given by $\{\phi_u, \phi_v, \mathcal{N}\}$.
We first express the derivatives of ϕ_u and ϕ_v in this basis. Write

$$\begin{cases} \phi_{uu} = \Gamma_{11}^1 \phi_u + \Gamma_{11}^2 \phi_v + e\mathcal{N} \\ \phi_{uv} = \Gamma_{12}^1 \phi_u + \Gamma_{12}^2 \phi_v + f\mathcal{N} \\ \phi_{vu} = \Gamma_{21}^1 \phi_u + \Gamma_{21}^2 \phi_v + f\mathcal{N} \\ \phi_{vv} = \Gamma_{22}^1 \phi_u + \Gamma_{22}^2 \phi_v + g\mathcal{N} \end{cases}$$

Note that $\phi_{vu} = \phi_{uv}, \Gamma_{12}^1 = \Gamma_{21}^1, \Gamma_{12}^2 = \Gamma_{21}^2$. We put two separate equations there only for symmetry. Γ_{ij}^k here are called Christoffel symbols.
Next, we shall show that it's possible to express Γ_{ij}^k in terms of E, F, G and their first partials. Clearly we don't want any of e, f, g to hang around, so the obvious thing to do is to take inner product of the equations with ϕ_u and ϕ_v respectively. On the first equation, it gives

$$\begin{cases} \Gamma_{11}^1 E + \Gamma_{11}^2 F = \langle \phi_{uu}, \phi_u \rangle = E_u/2 \\ \Gamma_{11}^1 F + \Gamma_{11}^2 G = \langle \phi_{uu}, \phi_v \rangle = F_u - E_v/2 \end{cases}$$

We can solve this for Γ_{11}^k (noting that $EG - F^2 \neq 0$) by Cramer's rule and get an expression of them in terms of E, F, G and their first partials. Similar for other Christoffel symbols.
Lastly, we express K in terms of the Christoffel symbols, their derivatives, and E as well. As $\phi_{uuv} = \phi_{uvu}$, we have

$$\begin{aligned} &\Gamma_{11}^1 \phi_{uv} + \Gamma_{11}^2 \phi_{vv} + e\mathcal{N}_v + (\Gamma_{11}^1)_v \phi_u + (\Gamma_{11}^1)_v \phi_v + e_v \mathcal{N} \\ &= \Gamma_{12}^1 \phi_{uu} + \Gamma_{12}^2 \phi_{vu} + f\mathcal{N}_u + (\Gamma_{12}^1)_u \phi_u + (\Gamma_{12}^2)_u \phi_v + f_u \mathcal{N} \end{aligned}$$

But we can further expand the ϕ_{uu}, ϕ_{uv} and ϕ_{vv} using our original system. Using the notation in the proof of Lemma 2.8, we (after some tedious calculation) will arrive at what's known as the Gauss formula:

$$\begin{aligned} (\Gamma_{12}^2)_u - (\Gamma_{11}^2)_v + \Gamma_{12}^1 \Gamma_{11}^2 - \Gamma_{12}^2 \Gamma_{11}^1 + \Gamma_{12}^2 \Gamma_{21}^2 - \Gamma_{11}^2 \Gamma_{22}^2 &= -f a_{21} + e a_{22} \\ &= E \frac{eg - f^2}{EG - F^2} = -EK \end{aligned}$$

As $E = \langle \phi_u, \phi_u \rangle \neq 0$, this proves the theorem. □

Remark. If ϕ is an orthogonal parameterisation ($F = 0$), then the Gauss formula reduces to

$$K = -\frac{1}{2\sqrt{EG}} \left(\left(\frac{E_v}{\sqrt{EG}} \right)_v + \left(\frac{G_u}{\sqrt{EG}} \right)_u \right)$$

If ϕ is isothermal, i.e. orthogonal and $E = G = \lambda(u, v)^2$ for some λ, then Gauss formula further reduces to

$$\begin{aligned} K &= -\frac{1}{2\lambda^2}\left(\left(\frac{2\lambda_v}{\lambda}\right)_v + \left(\frac{2\lambda_U}{\lambda}\right)_u\right) = -\frac{1}{\lambda^2}((\log\lambda)_{vv} + (\log\lambda)_{uu}) \\ &= -\frac{1}{\lambda^2}\nabla^2(\log\lambda) \end{aligned}$$

In fact, every surface has an isothermal parameterisation (but it's slightly involved to prove).

Note however that the mean curvature H is not an isometry invariant. For example, $\mathbb{R}^2$ is isometric to the cylinder $\{(x, y, z) \in \mathbb{R}^3 : 1 = x^2 + y^2\}$, but the former has constantly vanishing mean curvature while the mean curvature of the latter never vanishes.

3 Stationary Points of Length and Area

The theme in this chapter is to analysis cases where the length and area functionals attain their stationary points. Those of the length functional are called geodesics. They can be obtained from the one-dimensional Euler-Lagrange equations, which are ODEs (i.e. easy to solve). However, the analysis of stationary points for the area functional (known as minimal surfaces) requires higher dimensional calculus of variations. The Euler-Lagrange equations in this case are PDEs which calls for much more delicate treatments.

3.1 Geodesics

Suppose $S \subset \mathbb{R}^3$ is a surface and $p, q \in S$. Let $\Omega(p, q)$ be the set of all curves $\alpha : [0, 1] \to S$ with $\alpha(0) = p, \alpha(1) = q$. As before, let ℓ be the length functional $\Omega(p, q) \to \mathbb{R}$ given by

$$\ell(\alpha) = \int_0^1 |\dot\alpha(t)| \,\mathrm{d}t$$

This is actually a bit hard to work with, since there's a square root involved in calculating $|\dot\alpha|$. We often turn to a closely related but more convenient functional known as the energy functional, given by

$$\mathrm{E}(\alpha) = \frac{1}{2}\int_0^1 |\dot\alpha|^2 \,\mathrm{d}t$$

A drawback of this is that E is now sensitive to reparameterisation, but worry you not:

Lemma 3.1. *For any $\alpha \in \Omega(p, q)$, $\ell(\alpha)^2 \le 2\mathrm{E}(\alpha)$ with equality if and only if α has constant speed.*

That is, minimising ℓ over constant speed curves (which is all we care here) is equivalent to minimising E over all curves.

Proof. Cauchy-Schwartz. □

Let's now find extremals for E. Consider a one-parameter smooth family of curves $\alpha_s \in \Omega(p,q)$ say for $s \in (-\epsilon, \epsilon)$ and set $\alpha = \alpha_0$, then

$$\left.\frac{\mathrm{dE}(\alpha_s)}{\mathrm{d}s}\right|_{s=0} = \int_0^1 \left\langle \frac{\partial^2 \alpha_s}{\partial s \partial t}, \frac{\partial \alpha_s}{\partial t} \right\rangle_{s=0} \mathrm{d}t = \int_0^1 \left\langle \frac{\mathrm{d}W(t)}{\mathrm{d}t}, \dot{\alpha} \right\rangle \mathrm{d}t$$

where $W(t) = \partial \alpha_s(t)/\partial s|_{s=0}$. Integration by part is a thing.

$$\begin{aligned} \left.\frac{\mathrm{dE}(\alpha_s)}{\mathrm{d}s}\right|_{s=0} &= \langle W(1), \dot{\alpha}(1)\rangle - \langle W(0), \dot{\alpha}(0)\rangle - \int_0^1 \langle W(t), \ddot{\alpha}\rangle \,\mathrm{d}t \\ &= -\int_0^1 \langle W(t), \ddot{\alpha}\rangle \,\mathrm{d}t \end{aligned}$$

So if $\ddot{\alpha} \perp T_{\alpha(t)}S$ for all t, then α is an extremum for E. This is in fact necessary and sufficient.

Definition 3.1. A curve $\alpha : I \to S$ is called a geodesic if $\forall t \in I, \ddot{\alpha} \perp T_{\alpha(t)}S$.

Remark. 1. One can easily generalise this to the definition of geodesics to any smooth manifolds $X \subset \mathbb{R}^N$.
2. As expected, any geodesics defined this way has constant speed as we have $(\mathrm{d}/\mathrm{d}t)\langle \dot{\alpha}(t), \dot{\alpha}(t)\rangle = 2\langle \ddot{\alpha}(t), \dot{\alpha}(t)\rangle = 0$.
3. Geodesics are stationary points of the length functional. It needs not to be global maximum or minimum of it.

Example 3.1. 1. For a curve in the plane $\mathbb{R}^2 \subset \mathbb{R}^3$, it is a geodesic iff its acceleration vanishes everywhere, i.e. it is a line segment parameterised to have constant speed.
2. The geodesics on S^2 are contained in great circles (which, if you take the major arcs, exhibits a case where a geodesic neither maximises nor minimises the length functional globally).

3.2 Covariant Derivative and Parallel Transport

Definition 3.2. Let $\alpha : I \to S$ be a curve. A vector field V along α is a smooth map $V : I \to \mathbb{R}^3$ such that $V(t) \in T_{\alpha(t)}S$ for all $t \in I$.

Definition 3.3. The projection of $\mathrm{d}V/\mathrm{d}t$ onto the tangent plane $T_{\alpha(t)}S$ is called the covariant derivative of V at t, denotes $\mathrm{D}V/\mathrm{d}t$.

Remark. Consequently, to say α is a geodesic is just to say $\mathrm{D}\dot{\alpha}/\mathrm{d}t = 0$ along α.

Definition 3.4. A vector field V along α is called parallel if $\mathrm{D}V/\mathrm{d}t = 0$

Example 3.2. For a plane curve $\alpha : I \to \mathbb{R}^2 \subset \mathbb{R}^3$, V is parallel iff it's constant.

Proposition 3.2. *Let V, W be parallel vector fields along $\alpha : I \to S$, then $\langle V(t), W(t)\rangle$ is constant.*

Proof. $(\mathrm{d}/\mathrm{d}t)\langle V(t), W(t)\rangle = \langle \mathrm{d}W/\mathrm{d}t, V\rangle + \langle W, \mathrm{d}V/\mathrm{d}t\rangle = 0$. □

Lemma 3.3. *The covariant derivative depends only on the first fundamental form (i.e. not on the embedding).*

Proof. Let $\phi : U \to S$ be a local parameterisation and $\alpha : I \to S$ a curve on $\phi(U)$. Write $\alpha(t) = \phi(u(t), v(t))$ and let V be a vector field along α. Let V be a vector field along α, then we have $V(t) = a(t)\phi_u + b(t)\phi_v$ for some smooth a, b. Then

$$\frac{\mathrm{d}V}{\mathrm{d}t} = a(\phi_{uu}\dot{u} + \phi_{uv}\dot{v}) + b(\phi_{vu}\dot{u} + \phi_{uv}\dot{v}) + \dot{a}\phi_u + \dot{b}\phi_v$$

But we know from the proof of Theorem 2.9 that the ϕ_u, ϕ_v coefficients in $\phi_{uu}, \phi_{uv}, \phi_{vv}$ can be expressed in terms of E, F, G (and their partials). This implies the result. □

In fact, we have

$$\begin{aligned}\frac{\mathrm{D}V}{\mathrm{d}t} &= (\dot{a} + \Gamma^1_{11}a\dot{u} + \Gamma^1_{12}a\dot{v} + \Gamma^1_{21}b\dot{u} + \Gamma^1_{22}b\dot{v})\phi_u \\ &\quad + (\dot{b} + \Gamma^2_{11}a\dot{u} + \Gamma^2_{12}a\dot{v} + \Gamma^2_{21}b\dot{u} + \Gamma^2_{22}b\dot{v})\phi_v\end{aligned}$$

If we regard the equation $\mathrm{D}V/\mathrm{d}t = 0$ as a linear ODE in a and b, then Picard-Lindelöf shows that for any $v_0 \in T_{\alpha(t_0)}S, t_0 \in I$, there exists a unique parallel vector field V along α with $V(t_0) = v_0$.

Definition 3.5. Let V be as above. For $t_1 \in I$, the vector $V(t_1)$ is the parallel transport of v_0 along α at t_1.

We write $\mathfrak{p} = \mathfrak{p}_{p,q} : T_pS \to T_qS$ to denote that map that assigns each $v \in T_pS$ its parallel transport along α to q. It's easy to see that $\mathfrak{p}$ is linear (as the ODE is linear) and is an isometry (Proposition 3.2).
The fact that α is a geodesic iff $\mathrm{D}\dot{\alpha}/\mathrm{d}t = 0$ along α brings us another consequence (noting that ϕ_u and ϕ_v are linearly independent):

Corollary 3.4. *In local coordinates, the equations for geodesics are*

$$\begin{cases}\dot{\alpha} = \dot{u}\phi_u + \dot{v}\phi_v \\ \ddot{u} + \Gamma^1_{11}\dot{u}^2 + 2\Gamma^1_{12}\dot{u}\dot{v} + \Gamma^1_{22}\dot{v}^2 = 0 \\ \ddot{v} + \Gamma^2_{11}\dot{u}^2 + 2\Gamma^2_{12}\dot{u}\dot{v} + \Gamma^2_{22}\dot{v}^2 = 0\end{cases}$$

Then, with Picard-Lindelöf,

Proposition 3.5. *Given a point $p \in S$ and a vector $v \in T_pS$, there is some $\epsilon > 0$ and a unique geodesic $\gamma : (-\epsilon, \epsilon) \to S$ with $\gamma(0) = p, \dot{\gamma}(0) = v$.*

3.3 Minimal Surfaces

Definition 3.6. A surface in $\mathbb{R}^3$ is minimal if its mean curvature vanishes everywhere.

Remark. You will show in example sheet that no minimal surface can be compact.

Let $\phi : U \to S$ be a parameterisation and let $D \subset U$ be a bounded domain with $\bar{D} \subset U$. Let $h : \bar{D} \to \mathbb{R}$ be a smooth function.

Definition 3.7. The normal variation of $\phi(\bar{D})$ determined by h is the map $\rho : \bar{D} \times (-\epsilon, \epsilon) \to \mathbb{R}^3$ given by $\rho(u, v, t) = \phi(u, v) + th(u, v)\mathcal{N}(\phi(u, v))$.

Remark. We often demand that $h|_{\partial \bar{D}} = 0$ especially when $\partial \bar{D}$ is nice, e.g. piecewise smooth.

For fixed $t \in (-\epsilon, \epsilon)$, we write $\rho^t = \rho(-,-,t) : D \to \mathbb{R}^3$. We don't really have anything better to do, so let's just differentiate it.

$$\begin{cases} \rho^t_u = \phi_u + th\mathcal{N}_u + th_u\mathcal{N} \\ \rho^t_v = \phi_v + th\mathcal{N}_v + th_v\mathcal{N} \end{cases}$$

which are linearly independent for sufficiently small t, so if we choose a small enough ϵ we can assume that $\rho^t(D)$ is a smooth surface parameterised by ρ^t. Let E^t, F^t, G^t be the coefficients of the first fundamental form of ρ^t, then

$$\begin{cases} E^t = E + 2th\langle \phi_u, \mathcal{N}_u\rangle + t^2h^2\langle \mathcal{N}_u, \mathcal{N}_u\rangle + t^2h_uh_u \\ F^t = F + th(\langle \phi_u, \mathcal{N}_v\rangle + \langle \phi_v, N_u\rangle) + t^2h^2\langle \mathcal{N}_u, \mathcal{N}_v\rangle + t^2h_uh_v \\ G^t = G + 2th\langle \phi_v, \mathcal{N}_v\rangle + t^2h^2\langle \mathcal{N}_v, \mathcal{N}_v\rangle + t^2h_vh_v \end{cases}$$

Of course $\langle \phi_u, \mathcal{N}_v\rangle = \langle \phi_v, N_u\rangle$, but it's probably less confusing if we put them separately. Anyhow, we have $E^tG^t - (F^t)^2 = EG - F^2 - 2th(eG - 2fF + gE) + O(t^2)$ after an unhealthy amount of calculations. If we now put in $H = (eG - 2fF + gE)/(2(EG - F^2))$, we obtain $E^tG^t - (F^t)^2 = (EG - F^2)(1 - 4thH) + r(t)$ where $r(t) = o(t)$. Along the 1-parameter family of these surfaces, the area function $t \mapsto A(\rho^t(D))$ is smooth and has expression

$$A(\rho^t(D)) = \int_D \sqrt{EG - F^2}\sqrt{1 - 4thH + \bar{r}}\, \mathrm{d}u\, \mathrm{d}v, \bar{r} = \frac{r}{EG - F^2} = o(t)$$

so its derivative at 0 would equal

$$\left.\frac{\mathrm{d}A(\rho^t(D))}{\mathrm{d}t}\right|_{t=0} = \lim_{t\to 0} \frac{A(\rho^t(D)) - A(0)}{t} = -\int_D 2hH\sqrt{EG - F^2}\, \mathrm{d}u\, \mathrm{d}v$$

Proposition 3.6. *$\phi(U)$ is minimal if and only if $\mathrm{d}A(\rho^t(D))/\mathrm{d}t|_{t=0}$ vanishes for all choices of D and h.*

Proof. If $H = 0$, then we already know $\mathrm{d}A(\rho^t(D))/\mathrm{d}t|_{t=0} = 0$ by the above expression. Conversely, if $H(q) \neq 0$ for some q, then we can just choose D such that $H \neq 0$ on $\bar{D}$ by continuity and set $h = H|_{\bar{D}}$, which gives

$$\left.\frac{\mathrm{d}A(\rho^t(D))}{\mathrm{d}t}\right|_{t=0} = -\int_D 2H^2\sqrt{EG - F^2}\, \mathrm{d}u\, \mathrm{d}v < 0$$

In particular, it does not vanish. □

Remark. As with geodesics, minimal surfaces are just stationary points of the area functional, and are not necessarily those surfaces that minimise area.

Definition 3.8. The mean value vector is defined as $\tilde{H} = H\mathcal{N}$, which is independent of the choice of orientation.

Example 3.3. Most naturally generated surfaces are minimal, e.g. (non self-intersecting) soap films.

3.4 Weierstrass Representation

Proposition 3.7. *Suppose that* $\phi : U \to \mathbb{R}^3$ *is an isothermal representation (i.e.* $E = G = \lambda^2, F = 0$*). Then* $\nabla^2\phi = \phi_{uu} + \phi_{vv} = 2\lambda^2\tilde{H}$*.*

Proof. We have $\langle \phi_u, \phi_u \rangle = \langle \phi_v, \phi_v \rangle$, so $\langle \phi_{uu}, \phi_u \rangle = \langle \phi_{uv}, \phi_v \rangle$. Also, $\langle \phi_u, \phi_v \rangle = 0$ gives $\langle \phi_{uv}, \phi_v \rangle = -\langle \phi_u, \phi_{vv} \rangle$. Combining them shows that $\phi_{uu} + \phi_{vv}$ is parallel to $\mathcal{N}$. We also have $H = (g+e)/(2\lambda^2)$, so $2\lambda^2 H = g + e = \langle \phi_{uu} + \phi_{vv}, \mathcal{N} \rangle$, i.e. $\phi_{uu} + \phi_{vv} = 2\lambda^2 H\mathcal{N} = 2\lambda^2\tilde{H}$. □

Corollary 3.8. *Suppose* $\phi : U \to \mathbb{R}^3$ *is an isothermal parameterisation, then* $\phi(U)$ *is minimal iff* $\nabla^2 u = 0$*.*

Remark. 1. Despite being quite hard in general, one can show that isothermal coordinates always exists for minimal surfaces fairly easily (example sheet).
2. The corollary means that the components of such a parameterisation are harmonic. It is a fact that the modulus of a nonconstant harmonic function has no local maxima nor minima, which implies that no minimal surface can be compact (although a proof in this way is quite an overkill).

Example 3.4. For the catenoid $\phi(u,v) = (a\cosh v\cos u, a\cosh v\sin u, av)$, we have $E = G = a^2\cosh^2 v, F = 0$, so ϕ is isothermal. Moreover, $\nabla^2\phi = 0$, so this is a minimal surface. In fact, it is the only kind of complete minimal surface of revolution. The catenoid is actually isometric to the helicoid $\phi(u,v) = (a\sinh v\cos u, a\sinh v\sin u, au)$ since the helicoid parameterised this way also has $E = G = a^2\cosh^2 v, F = 0$. One can actually deform one into the other whilst keeping the area (look up for videos!). The helicoid is also minimal (however, this is not a phenomenon in general: the plane and the cylinder are isometric but the plane is minimal but the cylinder is not).

It was Lagrange who first became curious about minimal surfaces in 1760s. Specifically, he asked whether there exists a minimal surface that is not part of a plane. Meusnier answered the question in 1776 by constructing the catenoid and the helicoid. Scherk constructed another minimal surface in the 1830s which we will see in a moment.
Back to our theory. We have seen that there are apparently some connections between minimal surfaces and harmonic functions on the plane. Of course, there is no better tool to study those functions other than introducing complex variables.

Proposition 3.9. *Let* $\phi(u,v) = (x(u,v), y(u,v), z(u,v))$ *be a local parameterisation of a surface* $S \subset \mathbb{R}^3$*. Let* $\theta_1 = x_u - ix_v, \theta_2 = y_u - iy_v, \theta_3 = z_u - iz_v$*.*
1. ϕ *is isothermal iff* $\theta_1^2 + \theta_2^2 + \theta_3^2 = 0$*.*
2. Assuming ϕ *is isothermal, then* S *is minimal iff* $\theta_1, \theta_2, \theta_3$ *are holomorphic.*

Proof. 1. Indeed $\theta_1^2 + \theta_2^2 + \theta_3^2 = E - G - 2iF$.
2. The Cauchy-Riemann equations for θ_1 are $x_{uv} = x_{vu}$ and $x_{uu} = -x_{vv}$. The first one is always true, the second one is saying that x is harmonic. Doing the same for θ_2, θ_3 and casting Corollary 3.8 finishes the proof. □

The equation $\theta_1^2 + \theta_2^2 + \theta_3^2 = 0$ looks interesting. Let's try to solve it.

Lemma 3.10. *Let $D \subset \mathbb{C}$ be a domain, g a meromorphic function on D and f a holomorphic function on D. Suppose that whenever g has a pole of order k at $z \in D$, f would have a zero of order at least $2k$. Then*

$$\theta_1 = \frac{1}{2} f(1-g^2), \theta_2 = \frac{i}{2} f(1+g^2), \theta_3 = fg$$

are holomorphic on D and $\theta_1^2 + \theta_2^2 + \theta_3^2 = 0$.
Conversely, every triple of holomorphic functions $\theta_1, \theta_2, \theta_3$ satisfying $\theta_1^2 + \theta_2^2 + \theta_3^2 = 0$ can be obtained in this way except for the case $\theta_1 = i\theta_2, \theta_3 = 0$.

The exceptional case isn't very interesting: Any minimal surface with this happening must be part of a plane.

Proof. The first part of the lemma is pretty much direct calculation. Suppose now that $\theta_1, \theta_2, \theta_3$ are holomorphic and satisfy $\theta_1^2 + \theta_2^2 + \theta_3^2 = 0$. Unless $\theta_1 = i\theta_2$ (which means $\theta_3 = 0$), we can set $f = \theta_1 - i\theta_2, g = \theta_3/f$ which satisfies $\theta_1 = \frac{1}{2} f(1-g^2), \theta_2 = \frac{i}{2} f(1+g^2), \theta_3 = fg$ by direct calculation. We also have $\theta_1 + i\theta_2 = -\theta_3^2/(\theta_1 - i\theta_2) = -fg^2$. Holomorphy of this implies the condition we set for the zeros of poles of f and g respectively. □

Expectly, we can recover a minimal surface (locally) from these data. Assume $D \subset \mathbb{C}$ is a simply connected domain and $\xi_0 \in D$. Suppose we have $\theta_1, \theta_2, \theta_3$ holomorphic on D with $\theta_1^2 + \theta_2^2 + \theta_3^2 = 0$ and let f, g be as above.

Proposition 3.11. *The functions*

$$x(u,v) = \operatorname{Re} \int_{\xi_0}^{u+iv} \theta_1 \, \mathrm{d}\xi, y(u,v) = \operatorname{Re} \int_{\xi_0}^{u+iv} \theta_2 \, \mathrm{d}\xi, z(u,v) = \operatorname{Re} \int_{\xi_0}^{u+iv} \theta_3 \, \mathrm{d}\xi$$

satisfy $x_u - ix_v = \theta_1, y_u - iy_v = \theta_2, z_u - iz_v = \theta_3$.

Proof. Write $\omega = u + iv$. The full integral

$$\chi(\omega) = \int_{\xi_0}^{u+iv} \theta_1(\xi) \, \mathrm{d}\xi = x(u,v) + i\beta(u,v)$$

is holomorphic. So $\theta_1(\omega) = \chi'(\omega) = x_u + i\beta_u = x_u - ix_v$ by its Cauchy-Riemann equations. Similar for y and z. □

Corollary 3.12. *Whenever $\phi(u,v) = (x(u,v), y(u,v), z(u,v))$ is an embedding, its image is a minimal surface.*

(f, g) is then known as the Weierstrass representation of the minimal surface.

Remark. 1. Checking that ϕ defined in this way is an embedding is not always easy. But to check it's an immersion is much easier. In example sheet, you'll show that ϕ is an immersion iff f vanishes only at the pole of g and the order of its zeros at such a point is exactly twice the order of the pole of g.
Of course, immersions are local embeddings, so we can almost always get away with restrictions. Global embeddings however are still tricky to play with in general.
2. Given the existence of isothermal coordinates, what we did means that any minimal surface has local Weierstrass representations with suitable choices of D, f, g and ξ_0.

Example 3.5. 1. The Enneper surface (1864) is given by the Weierstrass representation $D = \mathbb{C}, f(\xi) = 1, g(\xi) = \xi$. In terms of the actual parameterisation, it translates to

$$\phi(u,v) = \frac{1}{2}\left(u - \frac{u^3}{3} + uv^2, -v - \frac{v^3}{3} - u^2 v, u^2 - v^2\right)$$

which is an embedding on an neighbourhood of the origin. It has Gaussian curvature $K = -16/(1+u^2+v^2)^4$.
2. Scherk's doubly periodic surface (1835) is given by $D = D(0,1) = \{z \in \mathbb{C} : |z| < 1\}, \xi_0 = 0, f(\xi) = 4/(1-\xi^4), g(\xi) = \xi$. The explicit parameterisation, as one can verify, is given on D by

$$\phi(u,v) = \left(-\pi - \arg\frac{\xi+i}{\xi-i}, -\pi - \arg\frac{\xi+1}{\xi-1}, \log\left|\frac{\xi^2+1}{\xi^2-1}\right|\right)$$

where $\xi = u + iv$. The projection of $\phi(D)$ on the (x,y) plane is contained in the open square $(-\pi/2, \pi/2) \times (-\pi/2, \pi/2)$ and we also have $\cos y = e^z \cos x$, which justifies the phrase "doubly periodic". If one project the whole surface onto (x,y), then one would get a checkerboard, pretty cool if you ask me.

3.5 Interpreting Weierstrass Representation

It can't be a pure coincidence that both g and $\mathcal{N}$ has codomain in S^2, right? Let $\pi : S^2 \to \mathbb{C} \cup \{\infty\}$ be the usual stereographic projection. Then

$$\pi^{-1}(z) = \left(\frac{2\operatorname{Re} z}{1+|z|^2}, \frac{2\operatorname{Im} z}{1+|z|^2}, \frac{|z|^2-1}{|z|^2+1}\right)$$

Suppose ϕ is an immersion induced by a Weierstrass parameterisation, i.e. $\phi_u - i\phi_v = (\theta_1, \theta_2, \theta_3)$, then

$$\lambda^2 = E = G = \frac{1}{2}\langle \phi_u - i\phi_v, \phi_u + i\phi_v\rangle = \frac{1}{2}(|\theta_1|^2 + |\theta_2|^2 + |\theta_2|^2) = \left(\frac{|f|(1+|g|^2)}{2}\right)^2$$

One can check that

$$\phi_u \times \phi_v = (\theta_2\bar{\theta}_3, \theta_3\bar{\theta}_1, \theta_1\bar{\theta}_2) = \frac{|f|^2(1+|g|^2)}{4}(2\operatorname{Re} g, 2\operatorname{Im} g, |g|^2 - 1)$$

So

$$\mathcal{N} \circ \phi = \left(\frac{2\operatorname{Re} g}{1+|g|^2}, \frac{2\operatorname{Im} g}{1+|g|^2}, \frac{|g|^2-1}{|g|^2+1}\right) \implies g = \pi \circ \mathcal{N} \circ \phi$$

Remark. In example sheet, you will show that, as a function on D, the Gaussian curvature at (u,v) equals

$$K = -\left(\frac{4|g'|}{|f|(1+|g|^2)^2}\right)^2$$

A minimal surface doesn't really have to be globally orientable (so far we've been working on a coordinate patch which must be orientable), so we can't exactly give it a global complex structure (which isn't that big a shame since we don't really know that much about non-compact Riemann surfaces). That being said, using complex analysis locally can yield some nice results too. For instance, the principle of isolated zeros immediately gives

Corollary 3.13. *If K is not identically zero, then its zeros are isolated.*

The famous Picard's theorem also tells us that

Theorem 3.14. *If $D = \mathbb{C}$, then either the immersed minimal surface $\phi(D)$ lies in a plane, or the Gauss map takes all values in S^2 with at most 2 exceptions.*

The deduction is not hard but let's spell it out, you know, for fun.

Proof. $\phi(D)$ lies on a plane iff $\theta_1 = i\theta_2, \theta_3 = 0$ iff g is not defined. Assuming this is not the case, then we get a meromorphic $g : \mathbb{C} \to \mathbb{C} \cup \{\infty\}$. If g is constant then the surface also lies on a plane. Otherwise, Picard's theorem tells us that g takes all values in $\mathbb{C} \cup \{\infty\}$ except for at most two. We conclude the result by quoting $g = \pi \circ \mathcal{N} \circ \phi$. $\square$

Example 3.6. The Enneper surface misses exactly one value, and the catenoid misses exactly two values.

Albeit being very hard, we do have some results in the general case using more elaborate theories.

Theorem 3.15 (Xavier 1981, Fujimoto 1988)**.** *If S is a complete non-planar minimal surface in $\mathbb{R}^3$, then the image of its Gauss map misses at most 4 points.*

4 Global Riemannian Geometry

4.1 The Exponential Map and Geodesic Polar Coordinates

Let $S \subset \mathbb{R}^3$ be an embedded surface. Recall that for any $p \in S$ and $v \in T_pS$, there is some $\epsilon > 0$ and a unique geodesic $\gamma_v : (-\epsilon, \epsilon) \to S$ with $\gamma_v(0) = p, \dot{\gamma}_v(0) = v$. Suppose $\gamma_v(1)$ is defined, then we set $\exp_p(v) = \gamma_v(1)$ and $\exp_p(0) = p$.

Proposition 4.1. *$\exp_p$ is defined and is smooth for v close enough to 0.*

Proof. Smoothness follows from the fact that the solution to the geodesic ODE has smooth dependence on initial conditions. If γ_v is defined over $(-\epsilon, \epsilon)$, then $\gamma_{\lambda v}$ is defined over $(-\epsilon/|\lambda|, \epsilon/|\lambda|)$. This (along with some compactness arguments) shows the well-definedness. $\square$

Definition 4.1. $\exp_p$ is called the exponential map.

Proposition 4.2. *$\exp_p$ is a diffeomorphism from a neighbourhood of 0 onto its image.*

Proof. By the inverse function theorem, it suffices to show that $(D\exp_p)_0$ is nonsingular. For $v \in T_pS$, set $\alpha(t) = tv$ for small t, then $(\exp_p \circ \alpha)(t) = \exp_p(tv) = \gamma_{tv}(1) = \gamma_v(t)$ which has tangent vector v at $t = 0$. Then $(D\exp_p)_0(v) = (D\exp_p)_0(\dot{\alpha}(0)) = (D(\exp_p \circ \alpha))_0(1) = \dot{\gamma}_v(0) = v$. But then $(D\exp_p)_0$ is the identity which in particular is nonsingular. $\square$

Corollary 4.3. *When restricted to a sufficiently small domain near 0, $\exp_p$ is a parameterisation of S around p.*

Definition 4.2. The image of a parameterisation given by the exponential map is called a normal neighbourhood of p.

Remark. 1. If $S \subset \mathbb{R}^3$ is closed, then we can actually get $\exp_p$ to define over all of T_pS if you push the Picard-Lindelöf theory a little.
2. $\exp_p$ depends smoothly on p (again due to the smooth dependence of the solution to a Picard-Lindelöf ODE on the initial conditions).
3. Despite these, there is no guarantee that $\exp_p$ may be a global embedding since geodesics can intersect themselves.
4. We have seen that we can generalise the definition of geodesics to arbitrary manifolds $X \subset \mathbb{R}^N$. Similarly, we can generalise the notion of exponential map as well.

The reason it's called the exponential map comes from the study of matrix groups, specifically those that are closed submanifolds of $M(n) \cong \mathbb{R}^{n\times n}$ where the inner product takes the form $\langle L_1, L_2\rangle = \operatorname{tr}(L_1 L_2^\top)$. Suppose X is such a matrix group and $A \in T_I X$. Define $\alpha : \mathbb{R} \to M(n)$ by

$$\alpha(t) = \exp(At) = I + tA + \frac{t^2}{2}A^2 + \frac{t^3}{6}A^3 + \cdots$$

We have $\dot\alpha = A\alpha, \ddot\alpha = A^2\alpha$. One can also show that α is a curve on X and $\ddot\alpha(t) \perp T_{\alpha(t)}X$, so α is a geodesic on X. So $\exp_I(A) = \exp(A)$ which justifies the terminology.
Let (r, θ) be the polar coordinates on $\mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}$. Say $p \in S$ and take $\epsilon > 0$ such that $\exp_p : B_\epsilon(0) \to V \subset S$ is a diffeomorphism where $B_\epsilon(0)$ denotes the open disk of radius ϵ.

Definition 4.3. The geodesic polar coordinates about p are the images of the polar coordinates on $B_\epsilon(0) \subset \mathbb{R}^2$ under $\exp_p$.

Technically, the polar coordinate we defined only works for $r > 0$ and $\theta \in (0, 2\pi)$. Formally, the parameterisation we take is actually $\phi : (-\epsilon, \epsilon) \times \mathbb{R} \to V \ni p$ via

$$\phi(r, \theta) = \exp_p(r\cos\theta, r\sin\theta) = \exp_p(rv(\theta)) = \gamma_{v(\theta)}(r)$$

where $v(\theta)$ is the unit vector with inclination θ. One can check that this is perfectly good and does extend the θ coordinate. It is smooth at the origin but isn't quite an embedding there. Around any other points in V, however, it does give a local parameterisation.

Definition 4.4. The image of circles centred at the origin under $\exp_p$ are called geodesic circles. The image of lines through the origin under $\exp_p$ are called radical geodesics.

Proposition 4.4. *The coefficients for the first fundamental form in geodesic polar coordinates satisfies* $E \equiv 1, F \equiv 0, G(0, -) \equiv 0, (\sqrt{G})_r(0, -) = 1$. *Moreover, the Gaussian curvature has the expression* $K = -(\sqrt{G})_{rr}/\sqrt{G}$.

Proof. $\phi_r(v, \theta) = \dot\gamma_{v(\theta)}(r)$ and $v(\theta) = \dot\gamma_{v(\theta)}(0)$ is a unit vector, so $E(r, \theta) = 1$ as geodesics have constant speed.

As for F, G, let $w = \mathrm{d}v/\mathrm{d}\theta$. We have $\phi_\theta = (D\exp_p)_{rv}(rw) = r(D\exp_p)_{rv}(w)$, so

$$\begin{cases} F(r,\theta) = r\langle \dot\gamma_{v(\theta)}(r), (D\exp_p)_{rv}(w)\rangle \\ G(r,\theta) = r^2 |(D\exp_p)_{rv}(w)|^2 \end{cases}$$

We immediately have $(\sqrt{G})_r(0,\theta) = |(D\exp_p)_0(w)| = |w| = 1$ for all θ. Also, $F_r(r,\theta) = \langle \phi_{rr}, \phi_\theta\rangle + \langle \phi_r, \phi_{\theta r}\rangle$. $\langle \phi_{rr}, \phi_\theta\rangle = 0$ since $\phi(r,\theta) = \gamma_{v(\theta)}(r)$ is a geodesic. We also have $\langle \phi_r, \phi_{\theta r}\rangle = (1/2)\partial\langle \phi_r, \phi_r\rangle/\partial\theta = E_\theta/2 = 0$. We conclude that $F(r,\theta)$ is constant in r for fixed θ. But clearly $F(0,\theta) = 0$ for all θ, so $F \equiv 0$.
The expression for K follows pretty much immediately from the Gauss formula (for orthogonal parameterisations) appeared in the proof of Theorem 2.9. □

Remark. 1. The fact that the parameterisation is orthogonal means that the geodesic circles are everywhere orthogonal to the radical geodesics.
2. The geodesic polar coordinates on S makes it pretty easy to show that geodesics locally minimise length and energy. There is a question in example sheet that makes this precise.

4.2 Geodesic Curvature

Let w be a smooth field of unit vectors along a curve α on an oriented surface S. Since $\langle w, w\rangle \equiv 1$, $\mathrm{d}w/\mathrm{d}t$ is always orthogonal to w, so $\mathrm{D}w/\mathrm{d}t = \lambda\mathcal{N} \times w$ for some scalar $\lambda = \lambda(t)$.

Definition 4.5. $[\mathrm{D}w/\mathrm{d}t] = \lambda(t)$ is called the algebraic value of the covariant derivative of w at t.

Remark. The sign of $[\mathrm{D}w/\mathrm{d}t]$ clearly depends on the choice of orientation. Also, we have $[\mathrm{D}w/\mathrm{d}t] = \langle \mathrm{d}w/\mathrm{d}t, \mathcal{N} \times w\rangle$.

Definition 4.6. Let $\alpha : I \to S$ be a regular curve parameterised by arc length, then the algebraic value $k_\mathrm{g}(s)$ of the covariant derivative of $\dot\alpha$ at s is called the geodesic curvature of α at $\alpha(s) \in S$.

The sign of k_g also depends on the orientation. Explicitly, $k_\mathrm{g} = \langle \ddot\alpha, \mathcal{N} \times \dot\alpha\rangle$. The definition of geodesic then shows directly that a curve on the surface is a geodesic iff its geodesic curvature vanishes everywhere. We also have the formula $k^2 = |\ddot\alpha|^2 = k_\mathrm{g}^2 + k_\mathrm{n}^2$.

Example 4.1. In the case of a plane curve, k_g reduces to the signed curvature of α.

Let V, W be two smooth unit vector fields along a curve $\alpha : I \to S$. We write (terrible notation alert) iV to denote the unique unit vector field along α such that (V, iV) is always a positively oriented orthonormal basis for $T_{\alpha(s)}S$ (i.e. $(V, iV, \mathcal{N})$ is a right-handed basis for $\mathbb{R}^3$). Explicitly, $iV = \mathcal{N} \times V$. Write $W(t) = a(t)V(t) + b(t)iV(t)$ (necessarily $a^2 + b^2 = 1$) for a, b smooth.

Definition 4.7. Any smooth ψ with $a = \cos\psi, b = \sin\psi$ is called a smooth determination of the angle from V to W along α.

Clearly we can extend this definition to general smooth vector fields by, well, normalising them and repeating the same thing.
Such a smooth determination, if exists, is unique up to a constant since any such ψ must solve $\dot{\psi} = (a^2+b^2)\dot{\psi} = a\dot{\psi}\cos\psi + b\dot{\psi}\sin\psi = a\dot{b} - b\dot{a}$. But it must exist. Indeed, take any $t_0 \in I$ and ψ_0 such that $a(t_0) = \cos\psi_0, b(t_0) = \sin\psi_0$, then we can simply choose

$$\psi(t) = \psi_0 + \int_{t_0}^{t} (a(\tau)\dot{b}(\tau) - b(\tau)\dot{a}(\tau))\,\mathrm{d}\tau$$

Proposition 4.5. *Suppose V, W are smooth unit vector fields along a curve $\alpha : I \to S$, then $[\mathrm{D}W/\mathrm{d}t] - [\mathrm{D}V/\mathrm{d}t] = \mathrm{d}\psi/\mathrm{d}t$ where ψ is a smooth determination of the angle from V to W.*

Proof. By definition, $[\mathrm{D}W/\mathrm{d}t] = \langle \dot{W}, \mathcal{N} \times W\rangle$ and $[\mathrm{d}V/\mathrm{d}t] = \langle \dot{V}, \mathcal{N} \times V\rangle = \langle \dot{V}, iV\rangle = -\langle V, \mathrm{d}(iV)/\mathrm{d}t\rangle$. As $W = (\cos\psi)V + (\sin\psi)iV$, we have

$$\dot{W} = -(\dot{\psi}\sin\psi)V + (\dot{\psi}\cos\psi)iV + (\cos\psi)\dot{V} + (\sin\psi)\frac{\mathrm{d}(iV)}{\mathrm{d}t}$$

and $\mathcal{N} \times W = (\cos\psi)iV - (\sin\psi)V$, so

$$\begin{aligned}\left[\frac{\mathrm{D}W}{\mathrm{d}t}\right] &= \dot{\psi} + \left\langle -(\sin\psi)V + (\cos\psi)iV, (\cos\psi)\dot{V} + (\sin\psi)\frac{\mathrm{d}(iV)}{\mathrm{d}t}\right\rangle \\ &= \dot{\psi} + (\cos^2\psi)\langle iV, \dot{V}\rangle - (\sin^2\psi)\left\langle V, \frac{\mathrm{d}(iV)}{\mathrm{d}t}\right\rangle = \dot{\psi} + \left[\frac{\mathrm{D}V}{\mathrm{d}t}\right]\end{aligned}$$

as desired. □

Remark. Suppose $\alpha : I \to S$ is a regular curve parameterised by arc length and V is a parallel unit vector field along α. Let ψ be a smooth determination of angle from V to $\dot{\alpha}$, then the proposition shows that $\mathrm{d}\psi/\mathrm{d}t = [\mathrm{D}\dot{\alpha}/\mathrm{d}t] = k_{\mathrm{g}}(s)$. In particular, k_{g} reduces to the ordinary curvature k when S is a plane (by taking V to be constant), although showing this directly is much easier.

Definition 4.8. For an oriented surface S and a local parameterisation ϕ on S. We say that ϕ is compatible with orientation if (ϕ_u, ϕ_v) is positively oriented with respect to the orientation on S.

Proposition 4.6. *Let $\phi(u, v)$ be an orthogonal parameterisation on S that's compatible with its orientation. Suppose $W(t)$ is a smooth unit vector field along a local curve $\alpha(t) = \phi(u(t), v(t))$, then*

$$\left[\frac{\mathrm{D}W}{\mathrm{d}t}\right] = \frac{G_u\dot{v} - E_v\dot{u}}{2\sqrt{EG}} + \frac{\mathrm{d}\psi}{\mathrm{d}t}$$

where ψ is a smooth determination of angle from ϕ_u to W.
In particular, if α is parameterised by arc length, then the geodesic curvature of it has the form

$$k_{\mathrm{g}} = \frac{G_u\dot{v} - E_v\dot{u}}{2\sqrt{EG}} + \frac{\mathrm{d}\psi}{\mathrm{d}t}$$

where ψ is a smooth determination of angle from ϕ_u to $\dot{\alpha}$.

The first term in the expression for k_{g} we obtained here looks suspiciously like what you'd find in an application of Green's theorem when integrating (part of) the orthogonal Gauss formula – it's a surprise tool that will help us later.

Proof. Let $e_1 = \phi_u/\sqrt{E}, e_2 = \phi_v/\sqrt{G}$, then (e_1, e_2) is a positively oriented orthonormal basis (for the corresponding tangent space). By the preceding proposition, we have $[\mathrm{D}W/\mathrm{d}t] = [\mathrm{D}e_1/\mathrm{d}t] + \mathrm{d}\psi/\mathrm{d}t$. The rest are just calculations.

$$\left[\frac{\mathrm{D}e_1}{\mathrm{d}t}\right] = \langle \dot{e}_1, \mathcal{N} \times e_1 \rangle = \langle \dot{e}_1, e_2 \rangle = \dot{u}\langle (e_1)_u, e_2 \rangle + \dot{v}\langle (e_1)_v, e_2 \rangle$$

We've got

$$\begin{aligned}\langle (e_1)_u, e_2 \rangle = \left\langle \left(\frac{\phi_u}{\sqrt{E}}\right)_u, \frac{\phi_v}{\sqrt{G}} \right\rangle &= \frac{1}{\sqrt{EG}}\langle \phi_{uu}, \phi_v \rangle = \frac{1}{\sqrt{EG}}(F_u - \langle \phi_u, \phi_{uv} \rangle) \\ &= -\frac{1}{\sqrt{EG}}\langle \phi_u, \phi_{uv} \rangle = -\frac{1}{\sqrt{EG}}\frac{E_v}{2} = -\frac{E_v}{2\sqrt{EG}}\end{aligned}$$

Similarly $\langle (e_1)_v, e_2 \rangle = G_u/(2\sqrt{EG})$. Combining them gives the result. □

4.3 Local Gauss-Bonnet Theorem

Definition 4.9. Let S be an oriented surface. A region $\bar{R} \subset S$ is called simple if $\bar{R}$ is homeomorphic to a closed disc and $\partial\bar{R}$ is a simple closed curve $\alpha : [0, L] \to S$ that is smooth at all but finitely many points ("vertices") and is parameterised by arc length.
We say α is positively oriented if $\mathcal{N}(\alpha(s)) \times \dot{\alpha}$, whenever defined, points towards R (the interior of $\bar{R}$).

We only defined parameterisation by arc length for smooth curves, but since α has only finitely many vertices, we can generalise the definition immediately to accommodate this situation. It is also clear that we can reparameterise any simple closed piecewise-smooth curve by arc length.
The exterior angles at the vertices of α is the angle difference between the limiting tangent vectors from two sides of the singularity. Suppose $\bar{R} \subset \phi(U)$ for some local parameterisation $\phi : U \to S$ with U a disc. Then there exists a piecewise smooth $\psi : I \to \mathbb{R}$ which is a smooth determination of the angle from ϕ_u to $\dot{\alpha}$ in the smooth parts and jumps by the exterior angle at the vertices. Apparently $\psi(L) - \psi(0) = 2\pi$. It's a bit hectic to prove so we are just gonna assume it.

Theorem 4.7 (Local Gauss-Bonnet Theorem)**.** *Suppose $\bar{R}$ is a simple region contained in an orthogonal parameterisation $\phi : U \to S$ compatible with orientation, where U is a disk. Take an arc length parameterisation $\alpha : I = [0, L] \to \phi(U)$ of ∂R that's positively oriented. Then*

$$\int_I k_{\mathrm{g}}(s)\,\mathrm{d}s + \int_R K\,\mathrm{d}A = 2\pi - \sum_i \theta_i$$

where θ_i are the exterior angles at vertices of α.

Of course k_g is undefined at the vertices. When we write an integral of it over I, what we actually mean is the sum of the integrals of it over connected components of I after removing the singularities.
Gauss proved this for geodesic boundary (i.e. k_g is constantly zero) and Bonnet generalised this to the form as above.

Proof. Take a piecewise smooth $\psi : I \to \mathbb{R}$ such that it is a smooth determination of the angle from ϕ_u to $\dot{\alpha}$ and jumps by the respective exterior angles at vertices. Proposition 4.6 gives

$$k_\mathrm{g} = \frac{G_u\dot{v} - E_v\dot{u}}{2\sqrt{EG}} + \frac{\mathrm{d}\psi}{\mathrm{d}t}$$

sparing the vertices. We are staring at a formula with some derivatives. We are also not morons (citation needed), so let's integrate it.

$$\begin{aligned}
\int_I k_\mathrm{g}(s)\,\mathrm{d}s &= \int_I \frac{G_u\dot{v} - E_v\dot{u}}{2\sqrt{EG}}\,\mathrm{d}s + \int_I \frac{\mathrm{d}\psi}{\mathrm{d}s}\,\mathrm{d}s = \int_I \frac{G_u\dot{v} - E_v\dot{u}}{2\sqrt{EG}}\,\mathrm{d}s + 2\pi - \sum_i \theta_i \\
&= \int_{\phi^{-1}(R)} \left(\left(\frac{G_u}{2\sqrt{EG}}\right)_u + \left(\frac{E_v}{2\sqrt{EG}}\right)_v \right) \mathrm{d}u\,\mathrm{d}v + 2\pi - \sum_i \theta_i \\
&= -\int_{\phi^{-1}(R)} K\sqrt{EG}\,\mathrm{d}u\,\mathrm{d}v + 2\pi - \sum_i \theta_i \\
&= -\int_{\phi^{-1}(R)} K\sqrt{EG - F^2}\,\mathrm{d}u\,\mathrm{d}v + 2\pi - \sum_i \theta_i \\
&= -\int_R K\,\mathrm{d}A + 2\pi - \sum_i \theta_i
\end{aligned}$$

by Green's theorem (ta-da) and the orthogonal Gauss formula. □

Remark. If T is a geodesic triangle with interior angles $\alpha_1, \alpha_2, \alpha_3$, then the theorem shows that

$$\int_T K\,\mathrm{d}A = 2\pi - \sum_{i=1}^{3} (\pi - \alpha_i) = \alpha_1 + \alpha_2 + \alpha_3 - \pi$$

4.4 Interlude: Triangulation and Euler Characteristic

Definition 4.10. Let S be a compact surface, possibly with boundary. A triangulation of S consists of a finite number of topological triangles (i.e. topological embeddings of an ordinary triangle on S) $\{T_1, \ldots, T_F\}$ that covers S and any two distinct triangles either have a single vertex in common, have an entire edge in common, or have an entire edge in common.

Definition 4.11. The Euler characteristic of a triangulation is defined as $\chi = V + F - E$ where V is the number of vertices, F the number of faces and E the number of edges.

With a new definition this late into a course, you're gonna just accept some facts that don't get to be proved.

Proposition 4.8. *Every compact surface has a triangulation, even with the additional requirement that asks the edges to be smooth. Furthermore, any triangulation of can be approximated arbitrarily well by a smooth one.*

Note that this is not true for higher dimensional manifolds (where we replace triangles by simplices).
In fact, χ is independent of the triangulation of the surface. In this course, we will prove this using the global Gauss-Bonnet theorem we'll ultimately prove (in algebraic topology, you get to prove this for general manifolds with homology). So it makes sense to write $\chi = \chi(S)$ and intend it to mean the Euler characteristic of the surface S.

Theorem 4.9 (Classification of Compact Orientable Surfaces)**.** *Compact Orientable surfaces (without boundaries) are completely classified (up to diffeomorphism) according to their Euler characteristics. Specifically, any such surface is diffeomorphic to a sphere with g (the "genus") handles attached, which has Euler characteristic $2 - 2g$.*

4.5 Global Gauss-Bonnet Theorem

Proposition 4.10. *Given any $p \in S$, there exists a orthogonal parameterisation locally around p.*

Proof. We can of course use the geodesic polar coordinates, but one that centres at p won't quite work. There is a workaround using the smooth dependence of $\exp_p$ on p and local compactness, but we instead will introduce another set of coordinates called geodesic normal coordinates, y'know, just for fun.
Start with a unit speed geodesic $\gamma : (-\epsilon, \epsilon) \to S$ with $\gamma(0) = p$. For each point $\gamma(v_0)$, we take a unit speed geodesic $\gamma^{v_0}(u)$ with $\gamma^{v_0}(0) = \gamma(v_0)$ such that $\gamma_u^{v_0}(0), \gamma_v(v_0)$ are orthonormal and positively oriented. By the existence, uniqueness of geodesics and the fact that they depend smoothly on initial conditions, we conclude that $\phi(u, v) = \gamma^v(u)$ is smooth in a neighbourhood of 0. Moreover, $\phi_u(0, 0) = \gamma_u^0(0)$ and $\phi_v(0, 0) = \gamma_v(0)$ are orthonormal, in particular not parallel. Inverse function theorem then shows that, by further shrinking the neighbourhood around 0 where ϕ is defined, we can get it to be a diffeomorphism, i.e. a local parameterisation around $\phi(0) = \gamma^0(0) = p$. One can check that ϕ is orthogonal. □

For a compact surface S, we cover it (by compactness) by a finite number of open geodesic balls (regions inside geodesic circles) $B(x_j, 2\epsilon_j) = \{y \in S : \}$, each of whom admits an orthogonal parameterisation. Set $\epsilon = \min_j \epsilon_j$.
Take a smooth triangulation of S. We can subdivide it by dividing each triangule into four small triangles by connecting the midpoints of edges. This clearly does not change the Euler characteristic since (for each of these subdivisions) we get three more faces, six more edges and three more vertices. Also, if we repeat this process over and over (for finite number of times), we can arrive at a triangulation with the same Euler characteristic χ but each triangle has geodesic diameter at most ϵ. This means that each triangle must be contained in one of the geodesic balls.
What happens if we apply local Gauss-Bonnet to each triangle in this new triangulation?

Theorem 4.11 (Global Gauss-Bonnet Theorem). *Let S be a smooth compact orientable surface without boundary. Given any triangulation of S with Euler characteristic χ, we have*

$$\int_S K \, \mathrm{d}A = 2\pi\chi$$

Remark. What does this formula mean? The left-hand side is an analytic quantity, but the right-hand side is something that's entirely topological. If one perturb the surface smoothly, then K will change locally. But since χ must be invariant under such perturbations, what we might lose in K at one point must be compensated somewhere else. There are many other theorems in geometry that look for such connections between global analytic and topological behaviours in certain geometric objects.

Proof. The discussion we had before implies that we can assume WLOG that each triangle is contained in an open geodesic ball. Applying local Gauss-Bonnet to one of these triangles T_i, say with interior angles $\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3}$, gives

$$\int_{T_i} K \, \mathrm{d}A + \int_{\partial T_i} k_{\mathrm{g}} = \alpha_{i,1} + \alpha_{i,2} + \alpha_{i,3} - \pi$$

where ∂T_i is oriented positively. Note that we have

$$\sum_i \int_{\partial T_i} k_{\mathrm{g}} = 0$$

since the edges cancel out with each other when everything's consistently oriented positively, which is possible due to the orientability of S. So,

$$\int_S K \, \mathrm{d}A = \sum_i (\alpha_{i,1} + \alpha_{i,2} + \alpha_{i,3} - \pi) = 2\pi V - \pi F = 2\pi(V - E + F) = 2\pi\chi$$

since $3F = 2E$. □

Corollary 4.12. *The Euler characteristic for a smooth compact boundaryless surface does not depend on the choice of triangulation.*

Theorem 4.13 (Global Gauss-Bonnet Theorem for Surfaces with Boundary). *Let S be a smooth compact orientable surface and $R \subset S$ is a domain. Suppose R has compact closure and its boundary can be parameterised by n pieces of simple closed positively-oriented curves $\alpha_i : I_i \to S$ parameterised by arc length. Let $\{\theta_j\}$ be the exterior angles of the vertices (the endpoints of α_i, that is) and χ be the Euler characteristic of a triangulation of R, then*

$$\int_R K \, \mathrm{d}A + \sum_i \int_{\alpha_i} k_{\mathrm{g}} = 2\pi\chi - \sum_j \theta_j$$

Proof. Exactly the same, except you don't get to cancel out the exterior angles and the integrals of geodesic curvatures over ∂R. □

4.6 Applications of Gauss-Bonnet

Theorem 4.14. *A compact oriented surface S with positive Gaussian curvature everywhere is diffeomorphic to S^2. Moreover, any two simple closed geodesics on S must intersect.*

Proof. If $K > 0$ everywhere, then every point of S^2 is a regular value to the Gauss map $\mathcal{N} : S \to S^2$. By Theorem 1.8, $\mathcal{N}$ is a covering map (cf. algebraic topology). Since S^2 is simply connected, it does not have nontrivial covering spaces, so S has to be homeomorphic and hence diffeomorphic to S^2 via $\mathcal{N}$.
An alternative way to phrase it using languages we developed in this course is the following: Theorem 1.8 shows that a sufficiently fine triangulation on S^2 with F faces, E edges and V vertices lifts through $\mathcal{N}$ to one on S with $F' = kF$ faces, $E' = kE$ edges and $V' = kV$ vertices for some constant k. Then $\chi(S) = k\chi(S^2) = 2k$ which implies $k = 1, \chi(S) = 2$ and S is diffeomorphic to S^2 by the classification of surfaces.
As for the second part of the theorem, suppose we have simple closed geodesics γ_1, γ_2 on S with $\gamma_1 \cap \gamma_2 = \varnothing$, then they bound a domain R in between. Exploiting $S \cong S^2$ shows that R is homeomorphic to an annulus which has Euler characteristic 0. Consequently,

$$\int_R K \, \mathrm{d}A = 0$$

Contradiction. □

In example sheet, you'll prove another theorem that has a similar flavour.

Proposition 4.15. *Let S be a surface homeomorphic to a cylinder and has everywhere negative Gaussian curvature, then S has at most one simple closed geodesic.*

Example 4.2. Catenoid has precisely one simple closed geodesic, namely the one at the "waist" of it.

4.7 Fenchel's Theorem

Definition 4.12. Let $\alpha : [0, L] \to \mathbb{R}^3$ be a plane curve parameterised by arc length. The total curvature of α is

$$\int_0^L k(s) \, \mathrm{d}s$$

where $k = |k| = |\ddot{\alpha}|$ is the curvature of α.

Notice that if α is a plane curve, then it is always contained in a geodesic normal parameterisation, so local Gauss-Bonnet shows that

$$\int_0^L k_{\mathrm{g}}(s) \, \mathrm{d}s = 2\pi$$

So the total curvature of α is at least 2π as $k = |k_{\mathrm{g}}|$, with equality iff k_{g} never changes sign.
What happens in the general case?

Theorem 4.16 (Fenchel)**.** *Any simple regular closed curve with nowhere vanishing curvature has total curvature is at least 2π, with equality iff it's a plane curve.*

One can actually drop the condition of the curvature being nowhere vanishing with some more work.
To prove this, we are going to use the following fact: Any compact oriented surface S (possibly with boundary) in $\mathbb{R}^3$ has

$$\int_S |K| \,\mathrm{d}A = \int_{S^2} |\mathcal{N}^{-1}(y)| \,\mathrm{d}y$$

which can be derived from Theorem 1.8 and Theorem 1.9. The precise proof needs some amount of book-keeping so we'll leave out the details. Back to Fenchel's theorem.

Proof. Recall from example sheet that given such a curve α parameterised by arc length, we can form the tube T of it by taking the local parameterisations

$$\phi(s, v) = \alpha(s) + r(n(s)\cos v + b(s)\sin v)$$

for some small enough r so that there is no self-intersections. n, b are the normal and binormal to α respectively, the construction of which depends on α having nonvanishing curvature, but it's also clear that we can define the tube without reference to them (which is how one might prove the general case).
Make r even smaller so that $rk < 1$ everywhere, then ϕ is regular with $EG - F^2 = r^2(1 - rk\cos v)^2$ and $\mathcal{N} = -(n\cos v + b\sin v)$. The Frenet formulae shows that

$$\mathcal{N}_s \times \mathcal{N}_v = k\cos v(n\cos v + b\sin v) = -(k\cos v)\mathcal{N} = -\frac{k\cos v}{r(1 - rk\cos v)}\phi_s \times \phi_v$$

So $K = -(k\cos v)/(r(1 - rk\cos v))$, consequently the points of T where K vanishes are the points where b and $-b$ intersect T. Let $R = \{p \in T : K(p) \geq 0\}$ and $R_+ = \{p \in T : K(p) > 0\} \subset R$. Then

$$\begin{aligned}\int_R K \,\mathrm{d}A &= \int_0^L \int_{\pi/2}^{3\pi/2} K\sqrt{EG - F^2} \,\mathrm{d}s \,\mathrm{d}v = -\int_0^L \int_{\pi/2}^{3\pi/2} K\cos v \,\mathrm{d}s \,\mathrm{d}v \\ &= -\int_0^L k(s) \,\mathrm{d}s \int_{\pi/2}^{3\pi/2} \cos v \,\mathrm{d}v = 2\int_0^L k(s) \,\mathrm{d}s\end{aligned}$$

Clearly the Gauss map is surjective on R, so

$$2\int_0^L k(s) \,\mathrm{d}s = \int_R K \,\mathrm{d}A = \int_{S^2} |(\mathcal{N}|_R)^{-1}(y)| \,\mathrm{d}y \geq 4\pi$$

If we have equality, then $\mathcal{N}$ has to be injective on R_+. We claim that T must then lie on one side of the tangent plane at p for any $p \in R_+$.
Suppose not, then there exists $q \neq p$ on S with $\mathcal{N}(p) = \mathcal{N}(q)$ and $p \in R_+, q \in R$. Recall that the set of zeros of K only sit at the intersections of $b, -b$ with T, which is nowhere dense. Consequently, we can find points $q' \in T$ arbitrarily close to q with positive Gaussian curvature. Also, $\mathcal{N}$ is a local diffeomorphism near p, say between open sets $U \ni p$ and $V \ni \mathcal{N}(p) = \mathcal{N}(q)$. We can assume

$q \notin \bar{U}$ after possibly shrinking U. Choose our q' such that $q' \notin U, \mathcal{N}(q') \in V$ and choose $p' \in U$ such that $\mathcal{N}(p') = \mathcal{N}(q')$. Then both p', q' are in R_+ and $p' \neq q'$, contradiction.
By a limit argument, we further deduce that T must lie on one side of the tangent plane at p for any p having zero curvature. These are the intersections between $\pm b$ and T.
Fix a circle $s = s_0$ in T which is divided into two semicircles by the straight line parallel to the binormal to α. The two tangent planes P_1, P_2 (which are parallel since their normals are both parallel to the binormal to α) at the endpoints of this division then must squeeze T in between. α then must lie on the plane P halfway in between P_1, P_2. Indeed, suppose not, then we can find by compactness some s_1 such that $\alpha(s_1)$ is furthest away from P. $t(s_1)$ is parallel to P because it would otherwise produce a point on α even further from P. Then the circle $s = s_1$ (which has radius r) is contained on a plane that cuts P perpendicularly, so T cannot be squeezed betweem P_1 and P_2 unless $\alpha(s_1) \in P$. □

4.8 Fáry-Milnor Theorem

Definition 4.13. A regular closed curve is unknotted if its image is the boundary of an embedded disc.

Turns out this notion is only useful in the case of curves in $\mathbb{R}^3$, but it's interesting enough.

Example 4.3. The circle, of course, is unknotted. The trefoil is not unknotted ("knotted"), but it's not easy to prove it.

Theorem 4.17 (Fáry-Milnor)**.** *Any regular knotted simple closed curve in $\mathbb{R}^3$ with nowhere vanishing curvature has total curvature at least 4π.*

Milnor actually proved something slightly (or massively) stronger: The said total curvature is strictly more than 4π, and 4π is the best possible lower bound.

Proof. Let the curve be $\alpha : [0, L] \to \mathbb{R}^3$ with an arc length parameterisation. We contruct the tube T of α as usual, then just like we did before

$$\int_{S^2} |\mathcal{N}^{-1}(y)| \,\mathrm{d}y = \int_T |K| \,\mathrm{d}A = 4 \int_0^L k(s) \,\mathrm{d}s$$

Suppose that the total curvature of α is strictly less than 4π, then

$$\int_T |K| \,\mathrm{d}A < 16\pi$$

So some point $v \in S^2$ (WLOG a regular value of $\mathcal{N}$ by Theorem 1.9) is the image of at most 3 points of T. That is, $\dot{\alpha} \perp v$ for at most three values of s. The function $h : s \mapsto \langle v, \alpha(s) \rangle$ has derivative $\langle v, \alpha(s) \rangle$, hence has derivative 0 for at most 3 values of s. The number of local extrema of a smooth function on S^1 is even, and we have at least one maximum and one minimum. So these are all of them for h, i.e. α consist of two arcs joining at the maximum and the minimum of h. Then, for each s that is not one of the two, there is a unique $s' \neq s$ such that $h(s) = h(s')$. The union of all line segments joining such s and s' gives a disc with boundary α, so it is unknotted. □