Math 162A Lecture notes on Curves and Surfaces, Part I

by Chuu-Lian Terng, Winter quarter 2005
Department of Mathematics, University of California at Irvine

Contents

1. Curves in $\mathbb{R}^n$

1.1. **Parametrized curves.**

A map $\alpha : [a,b] \to \mathbb{R}^n$, $\alpha(t) = (x_1(t), \ldots, x_n(t))$, is *smooth* if all derivatives of x_j exist and are continuous for all $1 \le j \le n$. We use $\alpha'(t)$ to denote the derivative:

$$\alpha'(t) = \frac{d\alpha}{dt} = \left(\frac{dx_1}{dt}, \ldots, \frac{dx_n}{dt}\right).$$

A *parametrized curve in* $\mathbb{R}^n$ is a smooth map $\alpha : [a,b] \to \mathbb{R}^n$ such that $\alpha'(t) \neq 0$ for all $t \in [a,b]$, and $\alpha'(t)$, is called its *tangent vector* at the point $\alpha(t)$.

Example 1.1.1. Straight line
The straight line in $\mathbb{R}^2$ through $(1,2)$ and $(2,-3)$ can be parametrized by $\alpha(t) = (1,2) + t(1,-5) = (1+t, 2-5t)$.

Example 1.1.2. Circle
The circle C of radius 2 centered at $(1,-1)$ is given by the equation $(x-1)^2 + (y+1)^2 = 2^2$. To find a parametric equation for this circle we may take $x - 1 = 2\cos t$ and $y + 1 = 2\sin t$, so $x = 1 + 2\cos t$, $y = -1 + 2\sin t$. Define $\alpha : [0, 2\pi] \to \mathbb{R}^2$ by

$$\alpha(t) = (x(t), y(t)) = (1 + 2\cos t, -1 + 2\sin t).$$

Since $\alpha'(t) = (-2\sin t, 2\cos t)$ is never a zero vector for any $0 \le t \le 2\pi$, α is a parametrization of the circle C.

Example 1.1.3. Ellipse
To find a parametrization of the ellipse E given by $\frac{x^2}{4} + \frac{y^2}{9} = 1$, we can set $\frac{x}{2} = \cos t$ and $\frac{y}{3} = \sin t$. Then $x = 2\cos t$ and $y = 3\sin t$. It is easy to check that

$$\alpha(t) = (2\cos t, 3\sin t)$$

is a parametrization of E

Example 1.1.4. The graph of a function
If $f : [a,b] \to \mathbb{R}$ is a smooth function, then $\alpha(x) = (x, f(x))$ is a parametrization of the graph of f, $\{(x, f(x)) \mid x \in [a,b]\}$.

1.2. **arclength parameter.**

To define the arclength of a parametrized curve $\alpha : [a,b] \to \mathbb{R}^n$, we first subdivide the interval $[a,b]$ into N subintervals $[t_i, t_{i+1}]$ of length $\delta = \frac{b-a}{N}$, $i = 0, 1, \ldots, N-1$, by taking $t_i = a + i\delta$. We now get a polygonal approximation to the curve α, namely the polygon with the $N+1$ vertices $\alpha(t_i)$ and N edges $[\alpha(t_i), \alpha(t_{i+1})]$. The length of the i-th edge is clearly $||\alpha(t_{i+1} - \alpha(t_i)||$, so that intuitively, their sum, $\sum_{j=0}^{N-1} ||\alpha(t_{i+1}) - \alpha(t_i)||$ should approximate the length of the curve α. But of course, when N gets large (and hence δ gets small) $\alpha(t_{i+1}) - \alpha(t_i)$ is approximately $\alpha'(t_j)\delta$, so the length of the polygon is approximated by $\sum_{j=1}^{N} ||\alpha'(t_j)||\delta$, which we recognize as a Riemann sum for the integral $L = \int_a^b ||\alpha'(t)||\, dt$ corresponding to the partition $t_0, \ldots, t_N$ of the interval $[a, b]$, and so we define this integral L to be the arclength of the curve α from $t = a$ to $t = b$. (We can give a physical interpretation of the arclength. If $\alpha(t)$ represents the position at time t of a particle moving in $\mathbb{R}^3$, then $\alpha'(t)$ is its instantaneous velocity at time t, so that $||\alpha'(t)||$ is the instantaneous speed at time t, and the integral of $||\alpha'(t)||$ should represent the distance travelled by the particle, i.e., the arclength.)

Now let us define the arclength function s from $[a, b]$ to $[0, L]$ by

$$s(t) := \int_a^t ||\alpha'(u)||\, du,$$

i.e., the arclength of α restricted to the interval $[a, t]$. By the Fundamental Theorem of Calculus, $\frac{ds}{dt} = ||\alpha'(t)||$. Since α is a parametrized curve, by definition $||\alpha'(t)|| > 0$ for all $t \in [a, b]$. But a smooth function from $[a, b]$ to $[0, L]$ with positive derivative at every point must be strictly increasing. In other words, $s : [a, b] \to [0, L]$ is one to one and onto, so it has an inverse function $t = t(s)$, and moreover, from calculus we know that $\frac{dt}{ds} = 1/(\frac{ds}{dt})$

If $\alpha(t)$ has unit speed, i.e., $||\alpha'(t)|| = 1$ for all $t \in [a, b]$, then the arclength function for α is

$$s(t) = \int_a^t ||\alpha'(u)||\, du = \int_a^t 1\ du = t - a.$$

In other words, the parameter t and the arclength differ by a constant, and we call such parameter an arclength parameter, and we give a formal definition next.

Definition 1.2.1. A parametrized curve $\alpha : [c_0, c_1] \to \mathbb{R}^n$ is said to have *arclength parameter* if $\alpha'(t)$ is a unit vector for all $c_0 \leq t \leq c_1$, i.e., $||\alpha'(t)|| = 1$ for all $t \in [c_0, c_1]$.

Suppose $\alpha : [a, b] \to \mathbb{R}^n$ is a parametrized curve, $s : [a, b] \to [0, L]$ the arclength function, and $t = t(s)$ the inverse function of s. Then $\beta(s) = \alpha(t(s))$ is a parametrized curve with arclength parameter. Note that β and α give the same curve in $\mathbb{R}^n$ but with different parametrizations. So we have shown that we can always change parameter to make a curve parametrized by its arclength.

Example 1.2.2. Straight line

For Example 1.1.1, $\alpha'(t) = (1, -5)$. So $s(t) = \int_0^t ||\alpha'(x)||\, dx = \int_0^t \sqrt{1 + 25}\, dx = \sqrt{26}\ t$, and the inverse function is $t(s) = \frac{s}{\sqrt{26}}$. Hence

$$\beta(s) = \alpha(t(s)) = \alpha(\frac{s}{\sqrt{26}}) = (1 + \frac{t}{\sqrt{26}}, 2 - \frac{5t}{\sqrt{26}}t)$$

gives the same straight line, but now is parametrized by arclength.

Example 1.2.3. Circle

For Example 1.1.2, $\alpha : [0, 2\pi] \to \mathbb{R}^2$, $\alpha'(t) = (-2\sin t, 2\cos t)$, and $||\alpha'(t)||^2 = 4$. So $||\alpha'(t)|| = 2$ and $s(t) = \int_0^t ||\alpha'(x)||\, dx = \int_0^t 2\, dx = 2t$. Hence the inverse function $t = \frac{s}{2}$, and

$$\beta(s) = \alpha(t(s)) = \alpha(\frac{s}{2}) = (1 + 2\cos\frac{s}{2}, -1 + 2\sin\frac{s}{2})$$

gives the same circle but with arclength parameter.

Example 1.2.4. Ellipse

For Example 1.1.3, $\alpha'(t) = (-2\sin t, 3\cos t)$ and $||\alpha'(t)|| = \sqrt{4\cos^2 t + 9\sin^2 t}$. Although we can not give an explicit formula for $s(t) = \int_0^t \sqrt{4\cos^2 t + 9\sin^2 t}\, dt$, the inverse function $t = t(s)$ exists and $\beta(s) = \alpha(t(s))$ is parametrized by arclength.

1.3. **Curvature of a plane curve.**

Informally speaking, the curvature of a plane curve is the rate at which its direction is changing. We next turn this intuitive idea into a formal definition. Assume that $\alpha : [c_0, c_1] \to \mathbb{R}^2$ is a parametrized curve with arclength parameter, i.e., $||\alpha'(s)|| = 1$ for all $c_0 \le s \le c_1$. Since $\alpha'(s)$ is a unit vector, we can write

$$\alpha'(s) = (\cos\theta(s), \sin\theta(s)),$$

where $\theta(s)$ is the angle (measured counter-clockwise) between the positive x-axis and the tangent vector $\alpha'(s)$. Thus we can think of $\theta(s)$ as the direction of the curve α at the point $\alpha(s)$

The *curvature* of α is defined to be the instantaneous rate of change of θ with respect to the arclength, i.e.,

$$k(s) = \theta'(s) = \frac{d\theta}{ds}.$$

Exercise 1.3.1.

(1) Prove that the curvature of a straight line is identically zero.
(2) Prove that the curvature of a circle of radius r is the constant function $\frac{1}{r}$.

If $\alpha(t) = (x(t), y(t))$ is a parametrized curve and $s(t)$ its arclength function, then since $\frac{d\alpha}{ds}$ is in the same direction as $\alpha'(t)$, the curvature can be also viewed as the instantaneous rate of change of the angle $\theta(t)$ between $\frac{d\alpha}{dt}$ and $(1,0)$ with respect to the arclength parameter. Since $\alpha'(t) = (x'(t), y'(t))$, the angle $\theta(t) = \tan^{-1}(\frac{y'(t)}{x'(t)})$. Now the Chain rule implies that

$$\begin{aligned}\frac{d\theta}{ds} &= \frac{d\theta}{dt}\frac{dt}{ds} = \frac{\frac{y''x'-x''y'}{x'^2}}{1+(\frac{y'}{x'})^2}\frac{dt}{ds} \\ &= \frac{y''x'-x''y'}{x'^2+y'^2}\frac{1}{\sqrt{\alpha'(t)^2}} = \frac{y''x'-x''y'}{(x'^2+y'^2)^{\frac{3}{2}}},\end{aligned}$$

which is the curvature at $\alpha(t)$, i.e., we have proved

Proposition 1.3.1. *If $\alpha(t) = (x(t), y(t))$ is a parametrized curve, then its curvature function is*

$$k = \frac{y''x'-x''y'}{(x'^2+y'^2)^{\frac{3}{2}}}.$$

For example, the curvature of $\alpha(t) = (t, f(t))$ (the graph of f) is

$$k(t) = \frac{f''(t)}{(1+f'(t)^2)^{\frac{3}{2}}}.$$

1.4. **Some elementary facts about inner product.**

The dot product of two vectors $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ is

$$X \cdot Y = x_1 y_1 + \cdots + x_n y_n.$$

It can be checked easily that for any real numbers c_1, c_2 and $X, Y, Z \in \mathbb{R}^n$ the dot product has the following properties:

(1) $(c_1 X + c_2 Y) \cdot Z = c_1 X \cdot Z + c_2 Y \cdot Z$,
(2) $Z \cdot (c_1 X + c_2 Y) = c_1 Z \cdot X + c_2 Z \cdot Y$,
(3) $X \cdot Y = Y \cdot X$,
(4) $X \cdot X \geq 0$ for all $X \in \mathbb{R}^n$ and "=" holds if and only if $X = 0$.

The first two conditions mean that $\cdot$ is a *bilinear* map from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$, the third condition means that $\cdot$ is *symmetric*, and the fourth condition says that $\cdot$ is what is called *positive definite*.

More generally, we define an inner product on a real vector space V to be a positive definite, symmetric bilinear map $(\ ,\ ) : V \times V \to \mathbb{R}$, i.e.,

(1) $(c_1 v_1 + c_2 v_2, v) = c_1 (v_1, v) + c_2 (v_2, v)$ for all real numbers c_1, c_2 and $v_1, v_2, v \in V$,
(2) $(v_1, v_2) = (v_2, v_1)$ for all $v_1, v_2 \in V$,
(3) $(v, v) \geq 0$ for all $v \in V$ and $(v, v) = 0$ if and only if $v = 0$.

An inner product space is a vector space V together with a particular choice $(\ ,\ )$ of an inner product on V.

Example 1.4.1.

(1) The dot product is an inner product on $\mathbb{R}^n$.
(2) If V is a linear subspace of $\mathbb{R}^n$, then $(v_1, v_2) = v_1 \cdot v_2$ defines an inner product on V.

Exercise 1.4.1.

Let $\mathbb{R}^n$ denote the space of $n \times 1$ vectors, and $A = (a_{ij})$ a symmetric $n \times n$ matrices. For X, Y in $\mathbb{R}^n$ define

$$(X, Y) = X^t A Y.$$

(Note that X^t is $1 \times n$, A is $n \times n$, and Y is $n \times 1$, so $X^t A Y$ is a 1×1 matrix, which can be identified as a real number). Prove that $(\ ,\ )$ is symmetric and bilinear.

Let $(\ ,\ )$ be an inner product on V. The *length* (or the *norm*) of $v \in V$ is

$$||v|| = \sqrt{(v, v)},$$

and the angle θ between $v_1, v_2 \in V$ satisfies

$$\cos \theta = \frac{(v_1, v_2)}{||v_1|| \, ||v_2||}.$$

In particular, if $v_1 \cdot v_2 = 0$, then the angle between v_1 and v_2 is $\pi/2$, i.e., $v_1 \perp v_2$. Note that if $V = \mathbb{R}^2$ or $\mathbb{R}^3$ and the inner product is the dot product, then the above formulas are the usual length and angle formula.

Let V be a vector space. We review some definitions:

(1) $v_1, \ldots, v_k$ are *linearly independent* if $c_1 v_1 + \cdots + c_k v_k = 0$ for some real numbers $c_1, \ldots, c_k$ implies that all c_i's must be equal to zero.

(2) $v_1, \dots, v_k$ *spans* V if every element of V is a linear combination of $v_1, \dots, v_k$.
(3) $\{v_1, \dots, v_k\}$ is a *basis* of V if they are linearly independent and span V.
(4) If $(\ ,\ )$ is an inner product on V, then a basis $\{v_1, \dots, v_k\}$ is *orthonormal* if $(v_i, v_j) = \delta_{ij}$, where

$$\delta_{ij} = \begin{cases} 1, & \text{if } 1 \leq i = j \leq k, \\ 0, & \text{if } 1 \leq i \neq j \leq k. \end{cases}$$

In other words, all v_i's have unit length and are mutually perpendicular.

If $\{v_1, \dots, v_n\}$ is a basis of $\mathbb{R}^n$, then given any v we can solve a system of linear equations to find real numbers $c_1, \dots, c_n$ such that $v = c_1 v_1 + \cdots + c_n v_n$. But when $\{v_1, \dots, v_n\}$ is an orthonormal basis, then we can use the inner product to get these c_i's without solving a system of linear equations. In fact, we have

Proposition 1.4.2. *Suppose $(\ ,\ )$ is an inner product on a vector space V, and $\{v_1, \dots, v_n\}$ an orthonormal basis for V. If $v = \sum_{i=1}^n c_i v_i$, then $c_i = (v_i, v)$ for $1 \leq i \leq n$.*

Proof. Take the inner product of v and v_i to see that

$$\begin{aligned}(v, v_i) &= (c_1 v_1 + \cdots + c_n v_n, v_i) \\ &= c_1(v_1, v_i) + \ldots + c_i(v_i, v_i) + \ldots + c_n(v_n, v_i) = c_i.\end{aligned}$$

□

Proposition 1.4.3. *Suppose $e_i : [a, b] \to \mathbb{R}^n$ are smooth maps for $1 \leq i \leq n$ such that $\{e_1(t), \dots, e_n(t)\}$ is an orthonormal basis of $\mathbb{R}^n$ for all $t \in [a, b]$. Then*

(1) *$e_i'(t) = \sum_{j=1}^n a_{ji}(t) e_j(t)$, where $a_{ji}(t) = e_i'(t) \cdot e_j(t)$.*
(2) *$a_{ij}(t) = -a_{ji}(t)$.*

Proof. Since $\{e_1(t), \dots, e_n(t)\}$ is a basis of $\mathbb{R}^n$, $e_i'(t)$ is a linear combination of $e_1(t), \dots, e_n(t)$, say

$$e_i'(t) = a_{1i}(t) e_1(t) + \cdots + a_{ni}(t) e_n(t).$$

The above Proposition implies that $a_{ji}(t) = e_i'(t) \cdot e_j(t)$, which proves (1). Since $(e_i, e_j) = \delta_{ij}$ is a constant, the product rule of differentiation implies that $e_i' \cdot e_j + e_i \cdot e_j' = 0$. By (1), $a_{ji} = e_i' \cdot e_j = -e_j' \cdot e_i = -a_{ij}$. □

Let $A = (a_{ij})$ be an $n \times m$ matrix. The transpose of A denoted by A^t is a $m \times n$ matrix, whose i-th column is the An $n \times n$ matrix $A = (a_{ij})$ is *anti-symmetric* if $A^t = -A$, i.e., $a_{ij} = -a_{ji}$. So if $A = (a_{ij})$ is skew-symmetric, then $a_{ii} = 0$ for all $1 \leq i \leq n$ and $a_{ij} = -a_{ji}$.

Next we relate linear combinations of vectors in $\mathbb{R}^n$ and matrix product. Let A be an $n \times n$ matrix, v_i the i-th column of A for $1 \leq i \leq n$, and $b = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix}$ an $n \times 1$

matrix. Then Ab is an $n \times 1$ matrix, and

$$Ab = (v_1, \ldots, v_n) \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix} = b_1 v_1 + \cdots + b_n v_n.$$

Conversely, if $w \in \mathbb{R}^n$ is equal to $c_1 v_1 + \ldots + c_n v_n$, then $Ac = w$, where $c = (c_1, \ldots, c_n)^t$. If $B = (b_{ij})$ is an $n \times n$ matrix and u_i the i-th column of B for $1 \leq i \leq n$, then the i-th column of AB is Au_i, or

$$AB = A(u_1, \ldots, u_n) = (Au_1, \ldots, Au_n).$$

We will use these simple formulas over and over again in the study of curves and surfaces in $\mathbb{R}^3$.

From now on, we will view $\mathbb{R}^n$ as the space of all $n \times 1$ matrices, i.e., as column vectors. Suppose $e_i : [a, b] \to \mathbb{R}^n$ are smooth maps such that $\{e_1(t), \ldots, e_n(t)\}$ is an orthonormal basis for each $t \in [a, b]$. Let $A(t) = (a_{ij}(t))$, where $a_{ij} = e_j' \cdot e_i$. Then Proposition 1.1.4 implies that

$$(1.4.1) \qquad (e_1'(t), \ldots, e_n'(t)) = (e_1(t), \ldots, e_n(t))A(t),$$

where $A(t) = (a_{ij}(t))$ is skew-symmetric for each $t \in [a, b]$.

1.5. **Moving frames along a plane curve.**

If $\alpha(s) = \begin{pmatrix} x(s) \\ y(s) \end{pmatrix}$ is parametrized by arc length, then $e_1(s) = \alpha'(s) = \begin{pmatrix} x'(s) \\ y'(s) \end{pmatrix}$ is the unit tangent vector. Note that $\begin{pmatrix} -b \\ a \end{pmatrix} \perp \begin{pmatrix} a \\ b \end{pmatrix}$, which can be either checked by computing the dot product or geometrically seeing that $\begin{pmatrix} -b \\ a \end{pmatrix}$ is the vector obtained by rotating $\begin{pmatrix} a \\ b \end{pmatrix}$ by 90 degree counterclockwise. So $e_2(s) = \begin{pmatrix} -y'(s) \\ x'(s) \end{pmatrix}$ is perpendicular to $e_1(s)$ and also has length 1 ($e_2(s)$ is a normal vector to the curve at $\alpha(s)$). In other words, $\{e_1(s), e_2(s)\}$ is an orthonormal basis of $\mathbb{R}^2$ for each s, which is called an orthonormal moving frame along the plane curve α. Let $\theta(s)$ denote the angle from $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to $\alpha'(s)$. Then

$$\alpha'(s) = \begin{pmatrix} x'(s) \\ y'(s) \end{pmatrix} = \begin{pmatrix} \cos\theta(s) \\ \sin\theta(s) \end{pmatrix},$$

so

$$e_1' = \alpha'' = \begin{pmatrix} -\theta' \sin\theta \\ \theta' \cos\theta \end{pmatrix} = \theta' \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} = \theta' e_2 = k e_2.$$

By [],

$$(e_1', e_2') = (e_1, e_2)A$$

with A skew-symmetric. Note that the diagonal entries of a skew-symmetric matrix are zero. But $e_1' = ke_2$ implies that the 21-th entry of A is k, so $A = \begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix}$.

This shows that

$$(e_1', e_2') = (e_1, e_2)\begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix},$$

or equivalently,

$$\begin{cases} e_1' = ke_2, \\ e_2' = -ke_1. \end{cases}$$

Note also that $\alpha' = e_1$. So e_1, e_2, α satisfies the following system of ordinary differential equations (ODE):

$$\begin{cases} e_1' = ke_2, \\ e_2' = -ke_1, \\ \alpha' = e_1. \end{cases}$$

1.6. **Orthogonal matrices and rigid motions.**

In this section, we view $\mathbb{R}^n$ as the space of real $n \times 1$ matrices, i.e., view $X \in \mathbb{R}^n$ as column vectors.

An $n \times n$ matrix A is *orthogonal* if $A^tA = \mathrm{I}_n$, the $n \times n$ identity matrix. Let v_i denote the i-th column of the $n \times n$ matrix A, i.e., $A = (v_1, \ldots, v_n)$. Then by definition of matrix multiplication we see that $A^tA = \mathrm{I}$ if and only if $v_1, \ldots, v_n$ are orthonormal because

$$A^tA = \begin{pmatrix} v_1^t \\ \cdot \\ \cdot \\ v_n^t \end{pmatrix} (v_1, \ldots, v_n) = (v_i^tv_j) = \mathrm{I}$$

implies that $v_i^tv_j = \delta_{ij}$.

Exercise 1.6.1. 2×2 **rotation matrix**

Let $\mathbb{R}^2$ denote the space of 2×1 matrices, and ρ a constant angle. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the map defined by $f(v)$ = rotate v by ρ. We use polar coordinates to write down the map: Suppose $v = (x, y)^t$ has radius r and angle θ, so $x = r\cos\theta$ and $y = r\sin\theta$. Since $f(v)$ is obtained by rotating v by angle ρ, the radius of $f(v)$ is r and the angle is $\theta + \rho$. If we write $f(v) = (p, q)^t$, then

$$\begin{cases} p = r\cos(\theta+\rho) = r(\cos\theta\cos\rho - \sin\theta\sin\rho) = x\cos\rho - y\sin\rho, \\ q = r\sin(\theta+\rho) = r(\sin\theta\cos\rho + \cos\theta\sin\rho) = x\cos\rho + y\sin\rho. \end{cases}$$

This shows that

$$\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \cos\rho & -\sin\rho \\ \sin\rho & \cos\rho \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

In other words, $f(v) = R_\rho v$, where

$$R_\rho = \begin{pmatrix} \cos\rho & -\sin\rho \\ \sin\rho & \cos\rho \end{pmatrix}.$$

R_ρ is called a *rotation matrix*. Note that $\det(R_\rho) = 1$ and R_ρ is orthogonal.

Exercise 1.6.2. **2×2 reflection matrix**

Let $g:\mathbb{R}^2\to\mathbb{R}^2$ be the linear map defined by $g(v)=$ the reflection of v in the line $y=\tan(\rho/2)x$. Prove that

(1) If $v=(x,y)^t$ has polar coordinate (r,θ), then $g(v)=(p,q)^t$ has polar coordinate $(r,\rho-\theta)$.
(2) Prove that

$$\begin{pmatrix} p\\ q\end{pmatrix}=S_\rho\begin{pmatrix} x\\ y\end{pmatrix},\quad \text{where } S_\rho=\begin{pmatrix}\cos\rho & \sin\rho\\ \sin\rho & -\cos\rho\end{pmatrix}.$$

S_ρ is called a reflection matrix (reflection in the line $y=\tan(\rho/2)x$). Note that S_ρ is orthogonal and $\det(S_\theta)=-1$.

Next we will show that any 2×2 orthogonal matrix is either a rotation matrix or a reflection matrix.

Proposition 1.6.1. *If A is a 2×2 orthogonal matrix, then A is either a rotation matrix or a reflection matrix.*

Proof. If $A=\begin{pmatrix} a & b\\ c & d\end{pmatrix}$ is orthogonal, then $A^tA=\mathrm{I}_2$ implies that

$$\begin{cases} a^2+c^2=1,\\ b^2+d^2=1,\\ ab+cd=0.\end{cases}$$

So we may assume $a=\cos\rho$, $b=\sin\rho$ for some ρ. But the third equation implies that

$$\frac{a}{c}=-\frac{d}{b}=\cot\rho,$$

so $d=-b\cot\rho$. But

$$1=b^2+d^2=b^2+\cot^2\theta\, b^2=(1+\cot^2\rho)b^2=b^2\csc^2\rho,$$

hence $b^2=\sin^2\rho$. So either $b=\sin\rho$ and $d=-\cos\rho$ or $b=-\sin\rho$ and $d=\cos\rho$. In other words, we have shown that a 2×2 orthogonal matrix must be either is a rotation matrix R_ρ or a reflection matrix S_ρ. □

Note that the dot product of $X,Y\in\mathbb{R}^n$ can be written as $X\cdot Y=X^tY$.

Proposition 1.6.2. *Let $X,Y\in\mathbb{R}^n$, and A a $n\times n$ matrix. Then*

(1) $(AX)\cdot Y=X\cdot(A^tY)$,
(2) *A is orthogonal if and only if $(AX)\cdot(AY)=X\cdot Y$.*

Proof. (1) is true because $(AX)\cdot Y=(AX)^tY=X^tA^tY=X^t(A^tY)=X\cdot A^tY$.

But $(AX)\cdot(AY)=(AX)^t(AY)=X^tA^tAY$ is equal to X^tY if and only if $A^tA=\mathrm{I}$, this proves (2). □

Since the dot product on $\mathbb{R}^n$ gives the length and angle, A is orthogonal if and only if A preserves length and cosine of the angle.

Definition 1.6.3. A map $f:\mathbb{R}^n\to\mathbb{R}^n$ is called a *rigid motion* if there exists an orthogonal matrix A with $\det(A)=1$ and a constant vector $b\in\mathbb{R}^n$ such that $f(X)=AX+b$ for all $X\in\mathbb{R}^n$.

The reason we require that $\det(A) = 1$ in the definition of rigid motions is because we want rigid motions preserve the orientation of $\mathbb{R}^n$.

Given a constant angle ρ and a constant vector $b \in \mathbb{R}^2$, the map $f(X) = R_\rho X + b$ is a rigid motion of $\mathbb{R}^2$ and all rigid motions of $\mathbb{R}^2$ are of this form, i.e., a rigid motion is a rotation plus a translation.

Proposition 1.6.4. *Let $\alpha : [a, b] \to \mathbb{R}^2$ be a curve parametrized by its arc length, $f : \mathbb{R}^2 \to \mathbb{R}^2$ a rigid motion defined by $f(X) = R_\rho X + b$, and $\beta : [a, b] \to \mathbb{R}^2$ the curve defined by $\beta(s) = f(\alpha(s))$. Then α and β have the same curvature functions.*

Proof. Since $\beta(s) = f(\alpha(s)) = R_\rho \alpha(s) + b$, $\beta'(s) = R_\rho \alpha'(s)$. But R_ρ is orthogonal, so $||\beta'(s)|| = ||R_\rho \alpha'(s)|| = ||\alpha'(s)|| = 1$. This shows that β is parametrized by its arc length. Let $e_1(s) = \alpha'(s)$ and $e_2(s) =$ rotate $e_1(s)$ by $\pi/2$ counterclockwise, and $\tilde{e}_1(s) = \beta'(s)$ and $\tilde{e}_2(s) =$ rotate $\tilde{e}_1(s)$ by $\pi/2$ counterclockwise. Because R_ρ is a rotation and $R_\rho(e_1) = \tilde{e}_1$, $R_\rho(e_2(s)) = \tilde{e}_2(s)$. The curvature function k and $\tilde{k}$ of α and β are given by

$$k(s) = e_1'(s) \cdot e_2(s), \quad \tilde{k}(s) = \tilde{e}_1'(s) \cdot \tilde{e}_2(s)$$

respectively. But

$$\tilde{e}_1' \cdot \tilde{e}_2 = (R_\rho e_1)' \cdot (R_\rho e_2) = (R_\rho e_1') \cdot (R_\rho e_2),$$

which is equal to $e_1' \cdot e_2$ because R_ρ is orthogonal. So we have proved $k = \tilde{k}$. □

Exercise 1.6.3.

(1) Prove that if A, B are orthogonal matrices, then AB is also an orthogonal matrix.
(2) Prove that if A is an orthogonal matrix, then A^{-1} is orthogonal.
(3) Suppose A, B are 3×3 orthogonal matrices with $\det(A) = \det(B) = 1$, b, c constant 3×1 vectors, and f, g are maps from $\mathbb{R}^3$ to $\mathbb{R}^3$ defined by $f(X) = AX + b$ and $g(X) = Bx + c$, i.e., f and g are rigid motions of $\mathbb{R}^3$. Prove that the composition $f \circ g$ is also a rigid motion.
(4) Suppose A is a 3×3 orthogonal matrix, b a 3×1 vector, and $f(X) = AX + b$ is a rigid motion of $\mathbb{R}^3$. Find the formula for the inverse function of f and show that it is also a rigid motion.
(The above two exercise says that the set of all rigid motions of $\mathbb{R}^3$ is a group under composition).

1.7. **Fundamental Theorem of plane curves.**

First we recall some results from calculus:

Theorem 1.7.1. Fundamental Theorem of Calculus

Suppose $h : [a, b] \to \mathbb{R}$ is continuous, and $f : [a, b] \to \mathbb{R}$ is the function defined by $f(x) = \int_a^x h(t)\ dt$. Then f is differentiable and $f' = h$.

Theorem 1.7.2. *Given a continuous function $h : [a, b] \to \mathbb{R}$ and $c_0 \in [a, b]$, if $y : [a, b] \to \mathbb{R}$ satisfies $y' = h$ and $y(c_0) = y_0$, then*

$$y(s) = y_0 + \int_{c_0}^s h(t)\ dt.$$

Proof. From calculus, we know that $y(s) = C + \int_{c_0}^{s} h(t)\, dt$ for some constant C. But $y(c_0) = y_0 = C + \int_{c_0}^{c_0} h(t)\, dt = C + 0 = C$, so $C = y_0$. □

Lemma 1.7.3. *Given two unit vectors v_1, v_2 in $\mathbb{R}^2$ and $p_0, q_0 \in \mathbb{R}^2$, there exists a unique rigid motion f, i.e., $f(X) = R_\rho X + b$, such that*

$$f(p_0) = q_0, \quad R_\rho v_1 = v_2.$$

Proof. Since v_1, v_2 are unit vectors in $\mathbb{R}^2$, there is a unique angle $0 \leq \rho < 2\pi$ such that $v_2 = R_\rho v_1$. To get b, we need to solve $f(p_0) = R_\rho p_0 + b = q_0$, so $b = q_0 - R_\rho p_0$. □

Theorem 1.7.4. Fundamental Theorem of plane curves

(1) *Given a smooth function $k : [a, b] \to \mathbb{R}$, $p_0 = (x_0, y_0)^t \in \mathbb{R}^2$, $v_0 \in \mathbb{R}^2$ a unit vector, and $c_0 \in [a, b]$, then there exists a unique $\alpha : [a, b] \to \mathbb{R}^2$ parametrized by its arc length such that $\alpha(c_0) = p_0$, $\alpha'(c_0) = v_0$, and its curvature function is k.*

(2) *If $\alpha, \beta : [a, b] \to \mathbb{R}^2$ are parametrized by arc length and have the same curvature functions, then there is a unique rigid motion f of $\mathbb{R}^2$ such that $\beta = f \circ \alpha$.*

Proof. Since v_0 is a unit vector, there is a constant angle θ_0 such that $v_0 = (\cos\theta_0, \sin\theta_0)^t$. Set

$$\theta(s) = \theta_0 + \int_{c_0}^{s} k(t)dt,$$
$$x(s) = x_0 + \int_{c_0}^{s} \cos\theta(t)\, dt,$$
$$y(s) = y_0 + \int_{c_0}^{s} \sin\theta(t)\, dt,$$
$$\alpha(s) = (x(s), y(s))^t.$$

By the fundamental theorem of calculus, $\theta'(s) = k(s)$, $x'(s) = \cos\theta(s)$, and $y'(s) = \sin\theta(s)$. But $\alpha'(s) = (x'(s), y'(s)) = (\cos\theta(s), \sin\theta(s))^t$, so α is parametrized by arc length and the curvature for α is $\theta' = k$. This proves that k is the curvature function for the curve α, and (1) is proved.

The proof of (1) also implies that if $\gamma : [a, b] \to \mathbb{R}^2$ is a curve parametrized by arc length, $\gamma(c_0) = p_0$ and $\gamma'(c_0) = v_0$, then γ must be the one constructed in (1). Or equivalently, suppose two curves $\alpha_1, \alpha_2 : [a, b] \to \mathbb{R}^2$ parametrized by arc length have the same curvature function, $\alpha_1(c_0) = \alpha_2(c_0)$, and $\alpha_1'(c_0) = \alpha_2'(c_0)$. Then $\alpha_1 = \alpha_2$.

By Lemma 1.7.3, there is a unique rigid motion $f(X) = R_\rho X + b$ such that $f(\alpha(c_0)) = \beta(c_0)$ and $R_\rho \alpha'(c_0) = \beta'(c_0)$. Let $\gamma : [a, b] \to \mathbb{R}^2$ be defined by $\gamma(s) = f(\alpha(s))$. By Proposition 1.6.4, α and γ have the same curvature function. But $\gamma(0) = f(\alpha(0)) = \beta(0)$ and $\gamma'(0) = R_\rho \alpha'(0) = \beta'(0)$, so β and γ have the same curvature function, pass through the same point with same unit tangent when $s = c_0$. So $\beta = \gamma$, i.e., $\beta = f \circ \alpha$. □

Exercise 1.7.1. Use the method given in Theorem 1.7.4 to find the curve whose curvature is 2, passes through $(0, 1)^t$ whose tangent at $(1, 0)^t$ is $(1/2, \sqrt{3}/2)^t$.

1.8. **Parallel curves.**

Let $\alpha : [a, b] \to \mathbb{R}^2$ be a curve parametrized by arc length, and $\{e_1(s), e_2(s)\}$ the o.n. moving frame along α (here $e_1(s) = \alpha'(s)$ is the unit tangent and $e_2(s) = R_{\frac{\pi}{2}} e_1(s)$ the unit normal). Given a constant $r > 0$, the curve parallel to α having distance r to α is

$$\beta(s) = \alpha(s) + r\ e_2(s).$$

Note that

$$\beta'(s) = \alpha'(s) + r\ e_2'(s) = e_1(s) - r\ k(s)e_1(s) = (1 - r\ k(s))e_1(s),$$

so if $k(s) \neq 1/r$ for all $s \in [a, b]$ then β is a parametrized curve. Such a β is also called a wave front of α, and when $k(s_0) = 1/r$ then the curve fails to be a parametrized curve at s_0 (recall that β is a parametrized curve if $\beta'(s) \neq (0, 0)$ for all $s \in [a, b]$), and we say that β has a singular point at $s = s_0$.

Definition 1.8.1. Suppose $\alpha : [a, b] \to \mathbb{R}^2$ is a curve parametrized by arc length, and $k(s_0) \neq 0$. The *osculating circle* at $\alpha(s_0)$ is the circle centered at $\alpha(s_0) + \frac{1}{k(s_0)} e_2(s_0)$ with radius $\frac{1}{k(s_0)}$.

By definition, the osculating circle passes through $\alpha(s_0)$ and has the same tangent line as the curve α at $\alpha(s_0)$. Let e_1, e_2 denote the o. n. moving frame along α. Let β denote the arc-length parametrization of the osculating circle of α at $\alpha(s_0)$ such that $\beta(s_0) = \alpha(s_0)$ and $\beta'(s_0) = \alpha'(s_0)$. The curvature of β (a circle of radius $1/k(s_0)$) is the constant function $k(s_0)$, which is equal to the curvature of α at $\alpha(s_0)$. Let $\theta(s)$ and $\tau(s)$ denote the polar angle of α' and β' respectively, i.e.,

$$\alpha'(s) = (\cos\theta(s), \sin\theta(s)), \quad \beta'(s) = (\cos\tau(s), \sin\tau(s)).$$

Then θ' and τ' are the curvature functions for α and β respectively. But $\alpha'(s_0) = \beta'(s_0)$ implies that $\theta(s_0) = \tau(s_0)$, and the curvature of α and β at $s = s_0$ equal implies that $\theta'(s_0) = \tau'(s_0)$. Compute the second derivatives of α and β at $s = s_0$ by the chain rule to get

$$\alpha''(s_0) = \theta'(s_0)(-\sin\theta(s_0), -\cos\theta(s_0)), \quad \beta''(s_0) = \tau'(s_0)(-\sin\tau(s_0), \cos\tau(s_0)).$$

But we have shown that $\theta(s_0) = \tau(s_0)$ and $\theta'(s_0) = \tau'(s_0)$, so $\alpha''(s_0) = \beta''(s_0)$. Thus we have proved that

$$\alpha(s_0) = \beta(s_0),\ \alpha'(s_0) = \beta'(s_0),\ \alpha''(s_0) = \beta''(s_0).$$

By Taylor's Theorem, $\alpha(s)$ and $\beta(s)$ agree up to second order near s_0, or the osculating circle is the best circle approximation of the curve α at $\alpha(s_0)$.

1.9. **Space Curves and Frenet frame.**

Let $\alpha : [a, b] \to \mathbb{R}^3$ be an immersed curve parametrized by arc length. So $e_1(s) = \alpha'(s)$ is a unit tangent vector. First we construct a natural o. n. moving frame along α, then the coefficients of their derivatives as linear combinations of the frame should give us the local invariants for space curves as we did for plane curves. While there is a natural choice of the unit normal vector for plane curves, there are infinitely many unit vectors normal to a space curve. So we need to make a choice of the second vector field of the frame. But since $e_1 \cdot e_1 = 1$,

$$0 = (e_1 \cdot e_1)' = e_1' \cdot e_1 + e_1 \cdot e_1' = 2e_1' \cdot e_1.$$

Thus $e_1' \cdot e_1 = 0$. If $e_1'(s) = \alpha''(s)$ is never zero, then the direction of $e_1'(s)$ is a unit vector perpendicular to $e_1(s)$. Set $k(s) = ||e_1'(s)||$, $e_2(s) = e_1'(s)/||e_1'(s)||$, and $e_3(s) = e_1(s) \times e_2(s)$. Then $\{e_1, e_2, e_3\}$ is an o. n. moving frame along α. By definition of e_2 and k we have $e_1' = ke_2$. Write $e_i' = \sum_{j=1}^3 a_{ji} e_j$ for $1 \leq j \leq 3$, i.e.,

$$\begin{cases} e_1' = a_{11}e_1 + a_{21}e_2 + a_{31}e_3, \\ e_2' = a_{12}e_1 + a_{22}e_2 + a_{32}e_3, \\ e_3' = a_{13}e_1 + a_{23}e_2 + a_{33}e_3. \end{cases}$$

Since $e_1' = ke_2$, $a_{11} = 0$, $a_{21} = k$, and $a_{31} = 0$. By Proposition 1.4.3, $A = (a_{ij})$ is skew-symmetric, i.e., $a_{ji} = -a_{ij}$. So $a_{ii} = 0$ for all $1 \leq i \leq 3$, $a_{12} = -k$, and $a_{13} = 0$. Therefore the matrix A must be of the form

$$\begin{pmatrix} 0 & -k & 0 \\ k & 0 & -\tau \\ 0 & \tau & 0 \end{pmatrix}$$

for some function τ. In fact, $\tau = e_2' \cdot e_3$. So we have proved

Proposition 1.9.1. *Let $\alpha : [a, b] \to \mathbb{R}^3$ be parametrized by arc length. If $\alpha''(s)$ is never zero, then set $e_1(s) = \alpha'(s)$, $k(s) = ||\alpha''(s)||$, $e_2(s) = \alpha''(s)/||\alpha''(s)||$, $e_3 = e_1 \times e_2$, and $\tau = e_2' \cdot e_3$. Then*

$$(e_1', e_2', e_3') = (e_1, e_2, e_3) \begin{pmatrix} 0 & -k & 0 \\ k & 0 & -\tau \\ 0 & \tau & 0 \end{pmatrix},$$

or equivalently,

$$\begin{cases} e_1' = ke_2, \\ e_2' = -ke_1 + \tau e_3, \\ e_3' = -\tau e_2. \end{cases} \tag{1.9.1}$$

Equation (1.9.1) is called the *Frenet equation*, k is called the curvature, τ is the torsion of α.

If $k \equiv 0$, then $e_1' = \alpha'' = ke_2 \equiv 0$, so $\alpha'(s) = u_0$, and $\alpha(s) = u_0 s + u_1$ for some constant vector $u_0.u_1$ in $\mathbb{R}^3$, i.e., α is a straight line. In general, the curvature k measures the deviation of α from being a straight line.

If $\tau \equiv 0$, then $e_3' = -\tau e_2 = 0$. So $e_3 = v_0 = (a, b, c)^t$ is a constant vector. But $(\alpha \cdot v_0) = \alpha' \cdot v_0 = e_1 \cdot e_3 = 0$ because e_1, e_2, e_3 are orthonormal. Hence $\alpha \cdot v_0 = c_0$ a constant, which shows that $\alpha(s) = (x(s), y(s), z(s))$ lies in the plane $ax + by + cz = c_0$. In other words, if $\tau \equiv 0$, then α is a plane curve. In genreal, the torsion measures the deviation of α being a plane curve.

Example 1.9.2. helix

Let a, b be positive constants such that $a^2 + b^2 = 1$, $\alpha(s) = (a \cos s, a \sin s, bs)$. Then $\alpha'(s) = (-a \sin s, a \cos s, b)$ has length $\sqrt{a^2 + b^2} = 1$, i.e., α is parametrized by arc length. Note that α lies on the cyclinder $x^2 + y^2 = a^2$ and is a helix. But $e_1'(s) = \alpha''(s) = (-a \cos s, -a \sin s, 0)$ implies that $k = ||\alpha''(s)|| = a$ and $e_2 = (-\cos s, -\sin s, 0)$. So $e_3(s) = e_1(s) \times e_2(s) = (b \sin s, -b \cos s, a)$, $\{e_1, e_2, e_3\}$ is the Frenet frame for α, and the torsion $\tau = e_2' \cdot e_3 = b$.

Proposition 1.9.3. *Suppose α is a curve in $\mathbb{R}^3$ parametrized by arc length such that α'' is never zero, and $f(X) = AX + b$ is a rigid motion of $\mathbb{R}^3$. Then $\beta := f \circ \alpha$ has the same curvature and torsion. Moreover, if (e_1, e_2, e_3) is the Frenet frame along α, then (Ae_1, Ae_2, Ae_3) is the Frenet frame along β.*

Proof. Since $\beta(s) = A\alpha(s) + b$, $\beta'(s) = A\alpha'(s) = Ae_1(s)$. A is orthogonal implies that $||Ae_1|| = ||e_1|| = 1$. But $e_1' = ke_2$ and

$$(Ae_1)' = Ae_1' = A(ke_2) = kAe_2.$$

By definition of rigid motion, A is orthogonal and $\det A = 1$, so $A(e_1 \times e_2) = (Ae_1) \times (Ae_2)$. Hence (Ae_1, Ae_2, Ae_3) is the Frenet frame for β. But $(Ae_3)' = Ae_3' = A(-\tau e_2) = -\tau Ae_2$, so the torsion for β is also equal to τ. □

1.10. **The Initial Value Problem for an ODE.**

Suppose we know the wind velocity at every point of space and at every instant of time. A puff of smoke drifts by, and at a certain moment we observe the precise location of a particular smoke particle. Can we then predict where that particle will be at all future times? By making this metaphorical question precise we will be led to the concept of an initial value problem for an ordinary differential equation.

We will interpret "space" to mean $\mathbb{R}^n$, and an "instant of time" will be represented by a real number t. Thus, knowing the wind velocity at every point of space and at all instants of time means that we have a function $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ that associates to each (t, x) in $\mathbb{R} \times \mathbb{R}^n$ a vector $F(t, x)$ in $\mathbb{R}^n$ representing the wind velocity at x at time t. Such a mapping is called a *time-dependent vector field on* $\mathbb{R}^n$. We will always be working with such F that are at least continuous, and usually F will even be continuously differentiable. In case $F(t, x)$ does not actually depend on t then we call F a *time-independent vector field on* $\mathbb{R}^n$, or simply a vector field on $\mathbb{R}^n$. Note that this is the same as giving a map $F : \mathbb{R}^n \to \mathbb{R}^n$.

How should we model the path taken by the smoke particle? An ideal smoke particle is characterized by the fact that it "goes with the flow", i.e., it is carried along by the wind, meaning that if $x(t)$ is its location at a time t, then its velocity at time t will be the wind velocity at that point and time, namely $F(t, x(t))$. But the velocity of the particle at time t is $x'(t) = \frac{dx}{dt}$, so the path of a smoke particle will be a differentiable curve $x : (a, b) \to \mathbb{R}^n$ such that $x'(t) = F(t, x(t))$ for all $t \in (a, b)$. Such a curve is called a *solution curve* of the time-dependent vector field F and we also say that"x satisfies the ordinary differential equation $\frac{dx}{dt} = F(t, x)$".

Usually we will be interested in solution curves of a differential equation $\frac{dx}{dt} = F(t, x)$ that satisfy a particular *initial condition.* This means that we have singled out some special time t_0 (often, $t_0 = 0$), and some specific point $v_0 \in \mathbb{R}^n$, and we look for a solution of the ODE that satisfies $x(t_0) = v_0$. (In our smoke particle metaphor, this corresponds to observing the particle at v_0 as our clock reads t_0.) The pair of equations $\frac{dx}{dt} = F(t, x)$ and $x(t_0) = v_0$ is called an "initial value problem" (abbreviated as IVP) for the ODE $\frac{dx}{dt} = F(t, x)$. The reason that it is so important is that the so-called Existence and Uniqueness Theorem for ODE says (more or less) that, under reasonable assumptions on X, the initial value problem has a "unique" solution.

Remark 1.10.1. If the vector field F is time-independent, then the ODE $\frac{dx}{dt} = F(x)$ is often called *autonomous*.

Remark 1.10.2. Let $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, and $F(t,x) = (f_1(t,x), \ldots, f_n(t,x))$ so that written out in full, the ODE $\frac{dx}{dt} = F(t,x)$ looks like

$$\frac{dx_i}{dt} = f_i(t, x_1(t), \ldots, x_n(t)), \quad i = 1, \ldots n.$$

In this form it is usually referred to as a "system of ordinary differential equations".

Example 1.10.3. F a constant vector field, i.e., $F(t,x) = u$, where u is some fixed element of $\mathbb{R}^n$. The solution with initial condition $x(t_0) = v_0$ is clearly the straight line $x(t) := v_0 + (t - t_0)u$.

Example 1.10.4. F is the "identity" vector field, $F(t,x) = x$. The solution with initial condition $x(t_0) = v_0$ is clearly $x(t) := e^{(t-t_0)}v_0$. (Later we will see how to generalize this to an arbitrary linear vector field, i.e., one of the form $F(t,v) = Av$ where A is a constant $n \times n$ matrix.)

Example 1.10.5. A vector field that is "space-independent", i.e., $F(t,x) = \phi(t)$ where $\phi : \mathbb{R} \to \mathbb{R}^n$ is continuous. The solution with initial condition $x(t_0) = v_0$ is $x(t) = v_0 + \int_{t_o}^t \phi(s)\, ds$.

1.11. **The Local Existence and Uniqueness Theorem of ODE.**

In what follows, we formulate the discussion in the last section in precise mathematical terms.

Let $B_r(p_0)$ denote the ball in $\mathbb{R}^n$ with radius r centered at p_0, i.e.,

$$B_r(p_0) = \{p \in \mathbb{R}^n \mid ||p - p_0|| < r\}.$$

A subset U of $\mathbb{R}^n$ is *open* if for each $p \in U$ there exists some $r > 0$ (r depend on the point p) such that $B_r(p) \subset U$.

A map f from an open subset U of $\mathbb{R}^n$ to $\mathbb{R}^m$ is C^1 if $\frac{\partial f}{\partial x_i}$ is continuous for all $1 \le i \le n$.

Theorem 1.11.1. The Local Existence and Uniqueness Theorem of ODE
Let $U \subset \mathbb{R}^n$ be an open subset, $F : (a,b) \times U \to \mathbb{R}^n$ is C^1. Fix $c_0 \in (a,b)$. Then given any $p_0 \in U$, there exists $\epsilon > 0$ and a unique solution $\alpha : (c_0 - \epsilon, c_0 + \epsilon) \to U$ of the following initial value problem:

$$\begin{cases} \frac{d\alpha}{dt} = F(t, \alpha(t)), \\ \alpha(c_0) = p_0. \end{cases} \tag{1.11.1}$$

Remark 1.11.2. If F is only continuous, the above theorem may not hold. For example, let $F : (-1,1) \times (-1,1) \to \mathbb{R}$ be the map defined by $F(t,y) = y^{1/3}$. Consider the initial value problem:

$$\begin{cases} \frac{dy}{dt} = y^{1/3}, \\ y(0) = 0. \end{cases}$$

First note that the constant function $y = 0$ is a solution. But this is a separable ODE $y^{-1/3}dy = dt$. Integrate both sides to get $\frac{3}{2}y^{2/3} = t + c$. But the initial condition $y(0) = 0$ implies that $c = 0$. So $y = (2t/3)^{3/2}$. This shows that

(1) we have two solutions defined on $[0,1)$ for the initial value problem,
(2) the second solution is not defined on an open interval containing $t=0$.

Next we give an application of the uniqueness of solutions of ODE. Let $\mathcal{M}_{3\times 3}$ denote the space of all 3×3 matrices. Then we can identify $\mathcal{M}_{3\times 3}$ as $(\mathbb{R}^3)^3=\mathbb{R}^9$.

Proposition 1.11.3. *Suppose* $A:[a,b]\to\mathcal{M}_{3\times 3}$ *is smooth and* $A(t)$ *is skew-symmetric for all* $t\in[a,b]$*,* $c_0\in(a,b)$*, and* C *is a* 3×3 *orthogonal matrix. If* $g:[a,b]\to\mathcal{M}_{3\times 3}$ *a smooth solution to the following initial value problem*

$$\begin{cases}\frac{dg}{dt}=g(t)A(t),\\ g(c_0)=C.\end{cases}$$

Then $g(t)$ *is orthogonal for all* $t\in[a,b]$.

Proof. Set $y(t)=g(t)^Tg(t)$, where $g(t)^T$ is the transpose of $g(t)$. By the product rule, we have $y(c_0)=C^TC=\mathrm{I}$ and

$$\begin{aligned}y'&=(g^T)'g+g^Tg'=(g')^Tg+g^TgA=(gA)^Tg+g^TgA\\&=A^Tg^Tg+g^TgA=A^Ty+yA.\end{aligned}$$

But the constant function $z(t)=\mathrm{I}$ also satisfies the above equation because $z'=0$ and $A^Tz+zA=A^T+A=0$. So by the existence and uniqueness of solutions of ODE $y=z$. Hence $y(t)=g(t)^Tg(t)=\mathrm{I}$, and $g(t)$ is orthogonal. □

1.12. **Fundamental Theorem of space curves.**

Proposition 1.12.1. *Let* $p_0,q_0\in\mathbb{R}^3$*, and* $\{u_1,u_2,u_3\}$ *and* $\{v_1,v_2,v_3\}$ *o.n. bases of* $\mathbb{R}^3$ *such that* $\det(u_1,u_2,u_3)=1$ *and* $\det(v_1,v_2,v_3)=1$*. Then there exists a unique rigid motion* $f(x)=Ax+b$ *such that* $f(p_0)=q_0$ *and* $Au_i=v_i$ *for all* $1\le i\le 3$*, where* A *is an orthogonal* 3×3 *matrix with* $\det(A)=1$ *and* $b\in\mathbb{R}^3$.

Proof. Let $U=(u_1,u_2,u_3)$ and $V=(v_1,v_2,v_3)$ be the 3×3 matrices with u_i and v_i as the i-th column respectively for $i=1,2,3$. Since $\{u_1,u_2,u_3\}$ and $\{v_1,v_2,v_3\}$ are o.n., U and V are orthogonal matrices. We want to find A such that $Au_i=v_i$ for $1\le i\le 3$, i.e., $AU=V$. So $A=VU^{-1}=VU^T$. We need $b\in\mathbb{R}^3$ such that $q_0=Ap_0+b$. Therefore b should be equal to q_0-Ap_0. □

Theorem 1.12.2. (Fundamental Theorem of curves in $\mathbb{R}^3$)

(1) *Given smooth functions* $k,\tau:(a,b)\to\mathbb{R}$ *so that* $k(t)>0$*,* $t_0\in(a,b)$*,* $p_0\in\mathbb{R}^3$ *and* (u_1,u_2,u_3) *a fixed o.n. basis of* $\mathbb{R}^3$*, then there exists* $\delta>0$ *and a unique curve* $\alpha:(t_0-\delta,t_0+\delta)\to\mathbb{R}^3$ *parametrized by arc length such that* $\alpha(0)=p_0$*, and* (u_1,u_2,u_3) *is the Frenet frame of* α *at* $t=t_0$.
(2) *Suppose* $\alpha,\tilde{\alpha}:(a,b)\to\mathbb{R}^3$ *are curves parametrized by arc length, and* $\alpha,\tilde{\alpha}$ *have the same curvature function* k *and torsion function* τ*. Then there exists a rigid motion* f *so that* $\tilde{\alpha}=f(\alpha)$.

Proof. First note that $F:(a,b)\times\mathbb{R}^3\times\mathbb{R}^3\times\mathbb{R}^3\to\mathbb{R}^3\times\mathbb{R}^3\times\mathbb{R}^3$ defined by

$$F(t,e_1,e_2,e_3)=(k(t)e_2,-k(t)e_1+\tau(t)e_3,-\tau(t)e_3)$$

is C^1. So by the Existence and Uniqueness of solutions of ODE (Theorem 1.11.1), there exists $\delta > 0$ and a unique solution

$$g : (t_0 - \delta, t_0 + \delta) \to \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$$

of the initial value problem

$$\begin{cases} g'(t) = g(t)A(t), \\ g(t_0) = (u_1, u_2, u_3) \end{cases},$$

where

$$A(t) = \begin{pmatrix} 0 & -k(t) & 0 \\ k(t) & 0 & -\tau(t) \\ 0 & \tau(t) & 0 \end{pmatrix}.$$

Since the initial value $g(t_0) = (u_1, u_2, u_3)$ is orthogonal and A is smooth and is skew-symmetric, by Proposition 1.11.3 $g(t)$ must be orthogonal for all $t \in (t_0 - \delta, t_0 + \delta)$. Let $e_i(t)$ denote the i-th column of $g(t)$. Then $\{e_1(t), e_2(t), e_3(t)\}$ is an o.n. basis of $\mathbb{R}^3$ for all $t \in (t_0 - \delta, t_0 + \delta)$. Set

$$\alpha(t) = p_0 + \int_{t_0}^{t} e_1(s)\ ds.$$

Then $\alpha'(t) = e_1(t)$ has length 1, so α is a curve in $\mathbb{R}^3$ parametrized by arc length. Write $g(t) = (e_1(t), e_2(t), e_3(t))$. Then $g' = gA$ can be written as

$$\begin{cases} e_1' = ke_2, \\ e_2' = -ke_1 + \tau e_3, \\ e_3' = -\tau e_2, \\ (e_1, e_2, e_3)(t_0) = (u_1, u_2, u_3) \end{cases}$$

This imply that k and τ are the torsion and curvature for α and $\{e_1, e_2, e_3\}$ is its Frenet frame, which proves (1).

Fix $t_0 \in (a, b)$. Let $\{e_1, e_2, e_3\}$ be the Frenet frame of α, and $\{\tilde{e}_1, \tilde{e}_2, \tilde{e}_3\}$ the Frenet frame of $\tilde{\alpha}$. By Proposition 1.12.1, there is a unique rigid motion $f(x) = Ax + b$ of $\mathbb{R}^3$ such that $f(\alpha(t_0)) = \tilde{\alpha}(t_0)$ and $A(e_i(t_0)) = \tilde{e}_i(t_0)$ for $1 \leq i \leq 3$. Set $\beta = f \circ \alpha$. By Proposition 1.9.3, the curvature and torsion for β are also k, τ, and the Frenet frame of β is $\{Ae_1, Ae_2, Ae_3\}$. But both $(\tilde{\alpha}, \tilde{e}_1, \tilde{e}_2, \tilde{e}_3)$ and $(\beta, Ae_1, Ae_2, Ae_3)$ satisfy the following differential equation:

$$(x, y_1, y_2, y_3)' = (y_1, ky_2, -ky_1 + \tau y_3, -\tau y_2)$$

and have the same initial condition at t_0, i.e., $\beta(t_0) = f(\alpha(t_0)) = \tilde{\alpha}(t_0)$, $Ae_1(t_0) = \tilde{e}_1(t_0)$, $Ae_2(t_0) = \tilde{e}_2(t_0)$, and $Ae_3(t_0) = \tilde{e}_3(t_0)$. So the uniqueness of solutions of ODE implies that $\beta(t) = \tilde{\alpha}(t)$ for all $t \in (a, b)$. But $\beta = f \circ \alpha$, so $\tilde{\alpha} = f \circ \alpha$, which proves (2). □

Exercise 1.12.1.

(1) Let a, b be positive constant real numbers. We call $\alpha(t) = (a \cos t, a \sin t, bt)$ the helix defined by a, b.
   (a) Reparametrized the curve by arc length.
   (b) Find the Frenet frame, curvature, and torsion of the curve.

(c) Given positive constants c_1, c_2 prove that there exist a, b such that the helix defined by a, b has curvature c_1 and torsion c_2.
(d) Assume β is a curve in $\mathbb{R}^3$ whose curvature and torsion are positive constants. Prove that β must be a piece of a helix. (hint: use (c) and the Fundamental Theorem of space curves)

(2) Given a space curve α, the normal plane at $\alpha(s_0)$ is the plane through $\alpha(s_0)$ and perpendicular to the tangent $\alpha'(s_0)$. Prove that if all the normal planes of α pass through a fixed point p_0, then α lies in a sphere centered at p_0. (Hint: Take the derivative of the function $f(s) = (\alpha(s) - p_0) \cdot (\alpha(s) - p_0)$.)

(3) Let a, b be constants such that $a^2 + b^2 = 1$. Suppose the binormal of a curve α is $e_3(s) = (a\cos(s^2), a\sin(s^2), b)$ and the torsion of α is positive. Find the curvature and torsion of α. (hint: use the Frenet equation)

2. **Fundamental forms of parametrized surfaces**

2.1. **Parametrized surfaces in $\mathbb{R}^3$.**

Let $v_1 = (a_1, a_2, a_3)^T$ and $v_2 = (b_1, b_2, b_3)^T$ in $\mathbb{R}^3$. Then v_1, v_2 are *linearly independent* if $c_1v_1 + c_2v_2 = 0$ for some constants c_1, c_2 implies that $c_1 = c_2 = 0$. It is easy to see that $v_1, v_2 \in \mathbb{R}^3$ are linearly independent if and only if the cross product $v_1 \times v_2 \neq 0$.

Definition 2.1.1. Let $\mathcal{O}$ be an open subset of $\mathbb{R}^2$. A smooth map $f : \mathcal{O} \to \mathbb{R}^3$, $f(u, v) = (f_1(u, v), f_2(u, v), f_3(u, v))$, is an *immersion* if $\frac{\partial f}{\partial u}(u, v), \frac{\partial f}{\partial v}(u, v)$ are linearly independent for all (u, v) in $\mathcal{O}$. An immersion $f : \mathcal{O} \to \mathbb{R}^3$ will be called a *parametrized surface in* $\mathbb{R}^3$.

Example 2.1.2. The graph of a smooth function $h : \mathcal{O} \to \mathbb{R}$
Let $f : \mathcal{O} \to \mathbb{R}^3$ be defined by $f(u, v) = (u, v, h(u, v))^T$. Then

$$\frac{\partial f}{\partial u} = \left(1, 0, \frac{\partial h}{\partial u}\right)^T, \quad \frac{\partial f}{\partial v} = \left(0, 1, \frac{\partial h}{\partial v}\right)^T.$$

Set $h_u = \frac{\partial h}{\partial u}$ and $h_v = \frac{\partial h}{\partial v}$. Then

$$\frac{\partial f}{\partial u} \times \frac{\partial f}{\partial v} = \begin{vmatrix} i & j & k \\ 1 & 0 & h_u \\ 0 & 1 & h_v \end{vmatrix} = (-h_u, -h_v, 1)^T$$

is never zero for all $(u, v) \in \mathcal{O}$. So f is an immersion.

Example 2.1.3. Surface of revolution
Let $h : \mathbb{R} \to \mathbb{R}$ be a smooth function, and C the curve in the yz-plane given by $z = h(y)$, and M the surface obtained by rotating the curve C along the y-axis. Then M can be parametrized by

$$f(y, \theta) = (h(y)\cos\theta, y, h(y)\sin\theta)^T.$$

A direct computation implies that

$$f_y := \frac{\partial f}{\partial y} = (h'(y)\cos\theta, 1, h'(y)\sin\theta), \quad f_o := \frac{\partial f}{\partial \theta} = (-h(y)\sin\theta, 0, h(y)\cos\theta),$$

and

$$f_y \times f_\theta = (h\cos\theta, -hh', h\sin\theta)^T = h(\cos\theta, -h', \sin\theta)^T,$$

so f fails to be an immersion exactly at (y, θ) when $h(y) = 0$. However, if $h(y) > 0$ for all $y \in \mathbb{R}$, then f is an immersion.

2.2. **Tangent and normal vectors.**

Let $f : \mathcal{O} \to \mathbb{R}^3$ be a parametrized surface, and $(u_0, v_0) \in \mathcal{O}$. Suppose $c : (-\epsilon, \epsilon) \to \mathcal{O}$, $c(t) = (u(t), v(t))$, is smooth and $c(0) = (u_0, v_0)$. We call $(f \circ c)'(0)$ a *tangent vector* of f at $f(u_0, v_0)$. By the chain rule

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0} f(u(t), v(t)) &= \frac{d}{dt}\Big|_{t=0} \frac{\partial f}{\partial u}(u(t), v(t))\frac{du}{dt} + \frac{\partial f}{\partial v}(u(t), v(t))\frac{dv}{dt} \\ &= f_u(u_0, v_0)u'(0) + f_v(u_0, v_0)v'(0). \end{aligned}$$

In particular, if $c(t) = (u_0 + t, v_0)$, then $(f \circ c)'(0) = f_u(u_0, v_0)$, and if $c(t) = (u_0, v_0 + t)$ then $(f \circ c)'(0) = f_v(u_0, v_0)$. These computations imply that the space of all tangent vectors of f at $f(u_0, v_0)$ is a 2-dimensional linear subspace of $\mathbb{R}^3$ and $\{f_u(u_0, v_0), f_v(u_0, v_0)\}$ is a basis. We call the space Tf_p of all tangent vectors of f at $p = (u_0, v_0)$ the *tangent plane of f at p*.

Since $v_1 \times v_2$ is perpendicular to both v_1, v_2,

$$N(u_0, v_0) = \frac{f_u(u_0, v_0) \times f_v(u_0, v_0)}{||f_u(u_0, v_0) \times f_v(u_0, v_0)||}$$

is a unit vector perpendicular to the tangent plane of f at (u_0, v_0), i.e., $N(u_0, v_0)$ is a *unit normal* to f at (u_0, v_0).

Given a parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$, we have constructed a moving basis (frame) on the parametrized surface: f_u, f_v, N. If we imitate what we did for curves in $\mathbb{R}^3$, then we should take the derivative of the moving frame and write them as linear combinations of f_u, f_v, N. We expect the coefficients of these linear combinations will give us the geometric invariants for the surface. Write

$$\begin{aligned} (f_u, f_v, N)_u &:= (f_{uu}, f_{vu}, N_u) = (f_u, f_v, N)P, \\ (f_u, f_v, N)_v &:= (f_{uv}, f_{vv}, N_v) = (f_u, f_v, N)Q \end{aligned}$$

for some $P, Q : \mathcal{O} \to \mathcal{M}_{3\times 3}$, i.e.,

$$\begin{aligned} f_{uu} &= f_u p_{11} + f_v p_{21} + N p_{31}, \\ f_{vu} &= f_u p_{12} + f_v p_{22} + N p_{32}, \\ &\cdots \end{aligned}$$

Since $f_{uv} = f_{vu}$, the second column of P is equal to the first column of Q, so the entries of P, Q have $18 - 3 = 15$ functions. However, we will see later that entries of P, Q can be expressed in terms of two "fundamental forms" (depending on six functions) and they must satisfy a system of partial differential equations, which are the so called Gauss-Codazzi equations. The fundamental theorem of surfaces in $\mathbb{R}^3$ states that the two fundamental forms satisfying the Gauss-Codazzi equations determine the surfaces in $\mathbb{R}^3$ unique up to rigid motions. In order to explain the two fundamental forms, we need to review some more linear algebra of bilinear forms and linear operators.

2.3. **Quadratic forms.**

Recall that a bilinear form on a vector space V is a map $b : V \times V \to \mathbb{R}$ such that

$$\begin{aligned} b(c_1v_1 + c_2v_2, v) &= c_1b(v_1, v) + c_2b(v_2, v), \\ b(v, c_1v_1 + c_2v_2) &= c_1b(v, v_1) + c_2b(v, v_2) \end{aligned}$$

for all $v, v_1, v_2 \in V$ and $c_1, c_2 \in \mathbb{R}$. A bilinear form b is *symmetric* if $b(v, w) = b(w, v)$ for all $v, w \in V$.

Definition 2.3.1. A real-valued function Q on a vector space V is called a *quadratic form* if it can be written in the form $Q(v) = b(v, v)$ for some symmetric bilinear form b on V. (We say that Q is the *quadratic form* associated to b).

If b is a symmetric bilinear form, then

$$b(v+w, v+w) = b(v, v) + b(v, w) + b(w, v) + b(w, w) = b(v, v) + 2b(v, w) + b(w, w),$$

so $Q(v + w) = Q(v) + 2b(v, w) + Q(w)$. In other words, we can recover b from the corresponding quadratic form Q as

$$b(v, w) = \frac{1}{2}(Q(v + w) - Q(v) - Q(w)).$$

Example 2.3.2. Let A be an $n \times n$ matrix, $\mathbb{R}^n$ the space of all $n \times 1$ real matrices, and $b(X, Y) = X^T AY$ (here we identify 1×1 matrices as real numbers). Then b is a bilinear form. If $X = (x_1, \ldots, x_n)^T$, $Y = (y_1, \ldots, y_n)^T$, and $A = (a_{ij})$, then

$$\begin{aligned} b(X, Y) = X^T AY &= (x_1, \ldots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ y_n \end{pmatrix} \\ &= (\sum_{j=1}^n a_{j1}x_j, \sum_{j=1}^n a_{j2}x_j, \ldots, \sum_{j=1}^n a_{jn}x_j) \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ y_n \end{pmatrix} \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ji}x_jy_i = \sum_{i,j=1}^n a_{ij}x_iy_j. \end{aligned}$$

Let $v_1, \ldots, v_n$ be a basis of V, and b a bilinear form on V. Set $b_{ij} = b(v_i, v_j)$ for $1 \le i, j \le n$. The matrix $B = (b_{ij})$ is called the *coefficient matrix of* b *with respect to* $\{v_1, \ldots, v_n\}$. If $v = \sum_{i=1}^n x_iv_i$ and $w = \sum_{i=1}^n y_iv_i$, then $b(v, w) = \sum_{i,j=1}^n b_{ij}x_iy_j$. If b is symmetric, then $b_{ij} = b_{ji}$ and the corresponding quadratic form Q is $Q(v) = \sum_{i=1}^n b_{ij}x_ix_j$.

Exercise 2.3.1.

(1) Let $f : R^2 \to \mathbb{R}^3$ be the map defined by

$$f(z,t) = (e^z \cos t, e^z \sin t, z)^T.$$

Sketch the surface and prove that f is an immersion. (i.e., show tht $\partial f/\partial z, \partial f/\partial t$ are linearly independent at every point).

(2) Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be the map defined by $f(z,\theta) = (z\cos\theta, z\sin\theta, z)$. Sketch the surface and prove that f is not an immersion when $z = 0$.

(3) Let $\alpha(s) = (y(s), z(s))$ be a smooth curve parametrized by arc length in the yz-plane, and $f(s,\theta) = (z(s)\cos\theta, y(s), z(s)\sin\theta)$. Find all the points when f fails to be an immersion.

(4) Let $\alpha : (a,b) \to \mathbb{R}^3$ be a curve parametrized by arc length with positive curvature function, $e_1(s) = \alpha'(s)$, and $f : (a,b) \times \mathbb{R} \to \mathbb{R}^3$ the map defined by $f(s,t) = \alpha(s) + te_1(s)$. Prove that f is an immersion except at $(s,0)$.

2.4. **Linear operators.**

Let V be a vector space, $\{v_1, \ldots, v_n\}$ a basis of V, and $T : V \to V$ a linear map. Then for each $1 \le i \le n$ we can write $T(v_i)$ as a linear combination of $v_1, \ldots, v_n$,

$$T(v_j) = \sum_{i=1}^n a_{ij} v_i.$$

We call $A = (a_{ij})$ the *matrix associated to the linear map T with respect to the basis* $\{v_1, \ldots, v_n\}$. Note that if we write $v = \sum_{i=1}^n x_i v_i$ and $T(v) = \sum_{i=1}^n y_i v_i$, then $Y = AX$, where $X = (x_1, \ldots, x_n)^T$ and $Y = (b_1, \ldots, b_n)^T$. This is because

$$T(v) = T\left(\sum_{i=1}^n x_i v_i\right) = \sum_{i=1}^n x_i T(v_i) = \sum_{i,j=1}^n x_i a_{ji} v_j = \sum_{j=1}^n \left(\sum_{i=1}^n a_{ji} x_i\right) v_j,$$

so $y_j = \sum_{i=1}^n a_{ji} x_i$, i.e., $Y = AX$.

If we change basis, the corresponding matrix of T change by a conjugation:

Proposition 2.4.1. *Let $T : V \to V$ be a linear map, and A and B the matrices of T associated to bases $\{v_1, \cdots, v_n\}$ and $\{u_1, \cdots, u_n\}$ of V respectively. If $u_i = \sum_j c_{ji} v_j$ for $1 \le i \le n$, then $B = C^{-1}AC$, where $C = (c_{ij})$.*

Proof.

$$\begin{aligned} T(u_i) &= \sum_{k=1}^n b_{ki} u_k = \sum_{k=1}^n b_{ki} \sum_{m=1}^n c_{mk} v_m = \sum_{m=1}^n \left(\sum_{k=1}^n b_{ki} c_{mk}\right) v_m \\ &= T\left(\sum_{j=1}^n c_{ji} v_j\right) = \sum_{j=1}^n c_{ji} T(v_j) = \sum_{j=1}^n c_{ji} \sum_{m=1}^n a_{mj} v_m \\ &= \sum_{m=1}^n \left(\sum_{j=1}^n c_{ji} a_{mj}\right) v_m. \end{aligned}$$

But $\{v_1, \dots, v_m\}$ is a basis of V, so $T(u_i)$ can be written uniquely as a linear combination of $v_1, \dots, v_n$. Hence

$$\sum_{k=1}^{n} b_{ki} c_{mk} = \sum_{j=1}^{n} c_{ji} a_{mj},$$

which means that the mi-th entry of CB is equal to the mi-th entry of AC. So $CB = AC$, which implies that $B = C^{-1}AC$. □

Suppose $A = (a_{ij})$ and $B = (b_{ij})$ are the matrices associated to the linear maps $S : V \to V$ and $T : V \to V$ with respect to basis $\{v_1, \dots, v_n\}$ respectively. In other words,

$$S(v_j) = \sum_{i=1}^{n} a_{ij} v_i, \quad T(v_j) = \sum_{i=1}^{n} b_{ij} v_i, \quad 1 \le j \le n.$$

Let $S \circ T : V \to V$ denote the composition of S and T, i.e.,

$$(S \circ T)(v) = S(T(v)).$$

Proposition 2.4.2. *If $A = (a_{ij})$ and $B = (b_{ij})$ are the matrices associated to linear maps $S : V \to V$ and $T : V \to V$ with respect to the basis $\{v_1, \dots, v_n\}$, then the matrix associated to $S \circ T$ with respect to $\{v_1, \dots, v_n\}$ is AB.*

Proof. Suppose $C = (c_{ij})$ is the matrix associated to $S \circ T$ with respect to $v_1, \dots, v_n$, i.e.,

$$S(T(v_j)) = \sum_{k=1}^{n} c_{kj} v_k.$$

But

$$\begin{aligned} S(T(v_j)) &= S(\sum_{i=1}^{n} b_{ij} v_i) = \sum_{i=1}^{n} b_{ij} S(v_i) \\ &= \sum_{i=1}^{n} b_{ij} \sum_{k=1}^{n} a_{ki} v_k = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} a_{ki} b_{ij} \right) v_k, \end{aligned}$$

so $c_{kj} = \sum_{i=1}^{n} a_{ki} b_{ij} = (AB)_{kj}$. This proves $C = AB$. □

Definition 2.4.3. Let $(\ ,\ )$ be an inner product on V. A linear operator $T : V \to V$ is *self-adjoint* if $(Tv, w) = (v, Tw)$ for all $v, w \in V$.

Proposition 2.4.4. *Let $T : V \to V$ be a linear map, and $A = (a_{ij})$ the matrix associated to T with respect to an orthonormal basis $v_1, \dots, v_n$. Then $a_{ij} = (T(v_j), v_i)$ for $1 \le i, j \le n$. Moreover, if T is self-adjoint, then A is symmetric.*

Proof. Suppose $T(v_j) = \sum_{i=1}^{n} a_{ij} v_i$. Since $(v_i, v_j) = \delta_{ij}$, $(T(v_j), v_i) = a_{ij}$.

If T is self-adjoint, then $a_{ij} = (T(v_j), v_i) = (v_j, T(v_i)) = a_{ji}$. □

Suppose $\{v_1, \dots, v_n\}$ is a basis of V but not an orthonormal, then although $((T(v_i), v_j))$ is not the matrix associated to T with respect to this basis, we still can compute the matrix in terms of $(T(v_i), v_j)$ and (v_i, v_j) as follows:

Proposition 2.4.5. *Let $(V, (\ ,\ ))$ be an inner product space, $v_1, \dots, v_n$ a basis of V, $T : V \to V$ a self-adjoint operator, $A = (a_{ij})$ the matrices of T with respect to $v_1, \dots, v_n$, and $b_{ij} = (T(v_i), v_j)$. Let $g_{ij} = (v_i, v_j)$, $G = (g_{ij})$, $G^{-1} = (g^{ij})$ the inverse of G, and $B = (b_{ij})$. Then*

(1) $B = A^t G$,
(2) $A = G^{-1} B^t = G^{-1} B$,
(3) $\det(A) = \frac{\det(B)}{\det(G)}$, $\operatorname{tr}(A) = \sum_{i=1}^n a_{ii} = \sum_{i,j} b_{ij} g^{ij}$.

Proof. Compute directly to get

$$b_{ij} = (T(v_i), v_j) = (\sum_k a_{ki} v_k, v_j) = \sum_k a_{ki} g_{kj} = (A^t G)_{ij},$$

i.e., the ij-th entry of B is equal to the ij-th entry of $A^t G$, which implies (1). Then $A^t = BG^{-1}$. But B and G are symmetric matrices, so $A = (G^{-1})^t B^t = G^{-1} B$, which proves (2). Recall that $\det(A_1 A_2) = \det(A_1)\det(A_2)$ and $\det(G^{-1}) = \frac{1}{\det(G)}$. Thus $\det(A) = \frac{\det(B)}{\det(G)}$. □

Let $(\ ,\ )$ be an inner product on V, and $S : V \to V$ a self-adjoint linear operator. Define $b^S : V \times V \to \mathbb{R}$ by

$$b^S(v, w) = (S(v), w).$$

It is easy to check that

(i) b^S is a symmetric bilinear form on V,
(ii) the coefficient matrix of b^S is (s_{ij}), where $s_{ij} = (S(v_i), v_j)$.

The bilinear form b^S will be called the *bilinear associated to S*.

Proposition 2.4.6. *Suppose $(\ ,\ )$ is an inner product on V, and b a symmetric bilinear form on V. Then there exists a unique self-adjoint operator $S : V \to V$ such that $b = b^S$.*

Proof. Let $\{v_1, \ldots, v_n\}$ be an orthonormal basis of V, and $b_{ij} = b(v_i, v_j)$. Define $S : V \to V$ be the linear map such that $S(v_i) = \sum_{j=1}^n b_{ji} v_j$. It is easy to check that $b = b^S$. □

Exercise 2.4.1.

(1) Suppose the coefficient matrix of the bilinear form b on $\mathbb{R}^2$ with respect to basis $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ is $\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$. Find $b(v, v)$ for $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
(2) Let V be the linear subspace of $\mathbb{R}^3$ defined by $x + y + z = 0$, and $b(v, w) = v \cdot w$, the dot product of v and w. Prove that
   (a) $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ form a basis of V,
   (b) Find the coefficient matrix of b with respect to this basis.

2.5. **The first fundamental Form.**

Definition 2.5.1. Suppose $f : \mathcal{O} \to \mathbb{R}^3$ is a parametrized surface in $\mathbb{R}^3$. A *quadratic form Q on f*, is a function $p \mapsto Q_p$ that assigns to each p in $\mathcal{O}$ a quadratic form Q_p on the tangent plane Tf_p of f at p.

Remark 2.5.2. Making use of the bases $f_u(p), f_v(p)$ in the Tf_p, a quadratic form Q on f is described by the symmetric 2×2 matrix of real-valued functions $Q_{ij} : \mathcal{O} \to \mathbb{R}$ defined by

$$Q_{ij}(p) := Q(f_{x_i}(p), f_{x_j}(p)),$$

(where $x_1 = u$ and $x_2 = v$). We call the Q_{ij} the *coefficients* of the quadratic form Q, and we say that Q is of class C^k if its three coefficients are C^k. These three functions Q_{11}, $Q_{12} = Q_{21}$, and Q_{22} on $\mathcal{O}$ determine the quadratic form Q on f uniquely: if $w \in Tf_p$, then we can write $w = \xi f_u(p) + \eta f_v$, and $Q_p(w) = Q_{11}(p)\,\xi^2 + 2\,Q_{12}(p)\,\xi\,\eta + Q_{22}(p)\,\eta^2$.

Note that we can choose *any* three functions Q_{ij} and use the above formula for $Q_p(w)$ to define a unique quadratic form Q on f with these Q_{ij} as coefficients. This means that **we can identify quadratic forms on a surface with ordered triples of real-valued functions on its domain.**

Notation. Because of the preceding remark, it is convenient to have a simple way of referring to the quadratic form Q on a surface having the three coefficients A, B, C. There is a classical and standard notation for this, namely:

$$Q = A(u, v)\,du^2 + 2B(u, v)\,du\,dv + C(u, v)\,dv^2.$$

To see the reason for this notation—and better understand its meaning—consider a curve in $\mathcal{O}$ given parametrically by $t \mapsto (u(t), v(t))$, and the corresponding image curve $\alpha(t) := f(u(t), v(t))$ on f. Then $\alpha'(t) = f_u(u(t), v(t))u'(t) + f_v(u(t), v(t))v'(t)$ and

$$Q(\alpha'(t)) = A(u(t), v(t))(u'(t))^2 + B(u(t), v(t))(u'(t)v'(t)) + C(u(t), v(t))(v'(t))^2.$$

The point is that curves on f are nearly always given in the form $t \mapsto f(u(t), v(t))$, so a knowledge of the coefficients A, B, C as functions ot u, v is just what is needed in order to compute the values of the form on tangent vectors to such a curve from the parametric functions $u(t)$ and $v(t)$. As a first application we shall now develop a formula for the length of the curve α.

Definition of the First Fundamental Form of a Surface f

Since the tangent plane Tf_p is a linear subspace of $\mathbb{R}^3$ it becomes an inner-product space by using the restriction of the inner product on $\mathbb{R}^3$. Let $I_p : TM_p \times TM_p \to \mathbb{R}$ denote this inner product, i.e.,

$$\mathrm{I}_p(\xi, \eta) = \xi \cdot \eta$$

for all $\xi, \eta \in Tf_p$, which is the First Fundamental Form on f. The coefficient matrix for I with respect to the basis $\{f_u, f_v\}$ is (g_{ij}), where

$$\begin{aligned} g_{11} &= f_u \cdot f_u, \\ g_{12} &= g_{21} = f_u \cdot f_v, \\ g_{22} &= f_v \cdot f_v. \end{aligned}$$

Thus:

$$\mathrm{I} = g_{11}(u, v)\,du^2 + 2g_{12}(u, v)\,du\,dv + g_{22}(u, v)\,dv^2.$$

If ξ_1 and ξ_2 are tangent vectors of f at $p_0 = (u_0, v_0)$, then we can write $\xi_1 = a_1 f_u(p_0) + a_2 f_v(p_0)$ and $\xi_2 = b_1 f_u(p_0) + b_2 f_v(p_0)$ for some constants a_1, a_2, b_1, b_2.

So

$$\begin{aligned}\xi_1 \cdot \xi_2 &= (a_1 f_u(p_0) + a_2 f_v(p_0)) \cdot (b_1 f_u(p_0) + b_2 f_v(p_0)) \\ &= g_{11}(p_0)a_1b_1 + g_{12}(p_0)(a_1b_2 + a_2b_1) + g_{22}(p_0)a_2b_2.\end{aligned}$$

In other words, we can compute the inner product of two vectors of the tangent plane Tf_{p_0} (hence the length of a vector and the angle between two vectors in Tf_p) from $g_{ij}(p_0)$'s.

The Arc Length of a Curve on a Surface

Let $t \mapsto (u(t), v(t))$ be a parametric curve in $\mathcal{O}$ with domain $[a, b]$. The length, L, of the curve $\alpha : t \mapsto f(u(t), v(t))$ is:

$$\begin{aligned}L &= \int_a^b ||\alpha'(t)||\, dt \\ &= \int_a^b \sqrt{g_{11}(u(t), v(t))u'(t)^2 + 2g_{12}(u(t), v(t))u'(t)v'(t) + g_{22}(u(t), v(t))v'(t)^2)}\, dt.\end{aligned}$$

The Angle between two curves

Let $c_1(t) = (x_1(t), x_2(t))$ and $c_2(t) = (y_1(t), y_2(t))$ be two smooth curves in $\mathcal{O}$ such that $c_1(0) = c_2(0) = p_0 = (u_0, v_0)$. The angle θ between $\alpha_1 = f \circ c_1$ and $\alpha_2 = f \circ c_2$ is defined to be the angle between $\alpha_1'(0)$ and $\alpha_2'(0)$. So

$$\begin{aligned}\cos\theta &= \frac{\alpha_1'(0) \cdot \alpha_2'(0)}{||\alpha_1'(0)||||\alpha_2'(0)||} \\ &= \frac{\sum_{i,j=1}^2 g_{ij}(p_0)x_i'(0)y_j'(0)}{\sqrt{\left(\sum_{i,j=1}^2 g_{ij}(p_0)x_i'(0)x_j'(0)\right)\left(\sum_{i,j=1}^2 g_{ij}(p_0)y_i'(0)y_j'(0)\right)}}.\end{aligned}$$

The Area

First recall that the area of the parallelgram spanned by v_1, v_2 in $\mathbb{R}^3$ is equal to the base times the height, so it is $||v_1||||v_2|| \sin\theta$, where θ is the angle between v_1 and v_2. But $\cos\theta = \frac{v_1 \cdot v_2}{||v_1||||v_2||}$ implies that

$$\begin{aligned}||v_1||^2||v_2||^2 \sin^2\theta &= ||v_1||^2||v_2||^2(1 - \cos^2\theta) = ||v_1||^2||v_2||^2 \left(1 - \left(\frac{v_1 \cdot v_2}{||v_1||||v_2||}\right)^2\right) \\ &= ||v_1||^2||v_2||^2 - (v_1 \cdot v_2)^2 = \begin{vmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 \\ v_2 \cdot v_1 & v_2 \cdot v_2 \end{vmatrix} = ||v_1 \times v_2||^2.\end{aligned}$$

Next sketch the idea of how to compute the area of the surface f over a rectangular region $D = [a_1, b_1] \times [a_2, b_2]$ inside $\mathcal{O}$: we subdivide D into small rectangular pieces and use the parallelgram spanned by $f_u \triangle u$ and $f_v \triangle v$ to approximate the area of each small piece, and then add them up to get an approximation of the area of $f(D)$. When the sides of the small pieces tend to zero, the limit of the sum is the surface area of $f(D)$:

$$\int_{a_1}^{a_2} \int_{b_1}^{b_2} \sqrt{g_{11}g_{22} - g_{12}^2}\, dudv.$$

Alternative Notations for the First Fundamental Form.

The First Fundamental Form of a surface is so important that there are several other standard notational conventions for referring to it. One whose origin should be obvious is to denote to it by ds^2, and call $ds = \sqrt{ds^2}$ the "line element" of the surface.

2.6. **The shape operator and the second fundamental form.**

Suppose $f : \mathcal{O} \to \mathbb{R}^3$ is a parametrized surface, and $N : \mathcal{O} \to \mathbb{R}^3$ is the unit normal vector field defined by

$$N = \frac{f_{x_1} \times f_{x_2}}{||f_{x_1} \times f_{x_2}||}.$$

Since $N \cdot N = 1$, the product rule implies that

$$0 = \frac{\partial}{\partial x_i}(N \cdot N) = N_{x_i} \cdot N + N \cdot N_{x_i} = 2N_{x_i} \cdot N,$$

so both N_{x_1} and N_{x_2} are tangent to f.

Definition 2.6.1. *The shape operator* S_p at $p \in \mathcal{O}$ for the parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$ is the linear map from Tf_p to Tf_p such that $S(f_{x_i}(p)) = -N_{x_i}(p)$ for $i = 1, 2$.

Given any tangent vector ξ of f at $f(p)$, we write ξ as a linear combination of $f_{x_1}(p), f_{x_2}(p)$, $\xi = c_1 f_{x_1}(p) + c_2 f_{x_2}(p)$, then $S_p(\xi) = -c_1 N_{x_1}(p) - c_2 N_{x_2}(p)$.

Proposition 2.6.2. *The shape operator S_p is self-adjoint.*

Proof. By definition, $S_p(f_{x_1}) \cdot f_{x_2} = -N_{x_1} \cdot f_{x_2}$. But $N \cdot f_{x_2} = 0$ implies that

$$(N \cdot f_{x_2})_{x_1} = 0 = N_{x_1} \cdot f_{x_2} + N \cdot f_{x_2 x_1},$$

so $-N_{x_1} \cdot f_{x_2} = N \cdot f_{x_2 x_1}$. Similarly, $-N_{x_2} \cdot f_{x_1} = N \cdot f_{x_1 x_2}$. Because f is smooth, $f_{x_1 x_2} = f_{x_2 x_1}$. So

$$-N_{x_1} \cdot f_{x_2} = -N_{x_2} \cdot f_{x_1} = N \cdot f_{x_1 x_2},$$

which implies that

$$S_p(f_{x_1}(p)) \cdot f_{x_2}(p) = f_{x_1}(p) \cdot S_p(f_{x_2}(p)).$$

Set $v_1 = f_{x_1}(p)$, $v_2 = f_{x_2}(p)$. Next we want to prove S_p is self-adjoint. Given ξ, η in Tf_p, we write $\xi = c_1 v_1 + c_2 v_2$ and $\eta = d_1 v_1 + d_2 v_2$. Since S is linear,

$$S_p(\xi) \cdot \eta = c_1 d_1 S_p(v_1) \cdot v_1 + c_1 d_2 S_p(v_1) \cdot v_2 + c_2 d_1 S_p(v_2) \cdot v_1 + c_2 d_2 S_p(v_2, v_2),$$

$$S_p(\eta) \cdot \xi = c_1 d_1 S_p(v_1) \cdot v_1 + c_1 d_2 S_p(v_2) \cdot v_1 + c_2 d_1 S_p(v_1) \cdot v_2 + c_2 d_2 S_p(v_2) \cdot v_2.$$

But we have shown that $S_p(v_1) \cdot v_2 = v_1 \cdot S_p(v_2)$, so $S_p(\xi) \cdot \eta = \xi \cdot S_p(\eta)$. $\square$

Let II_p denote the symmetric bilinear form on Tf_p associated to the self-adjoint operator S_p, i.e.,

$$\mathrm{II}_p : Tf_p \times Tf_p \to \mathbb{R}$$

is defined by

$$\mathrm{II}_p(\xi, \eta) = S(\xi) \cdot \eta.$$

It is easy to check that II is bilinear, and is symmetric because S_p is self-adjoint. So the coefficient matrix (ℓ_{ij}) of II_p with respect to basis $\{f_{x_1}(p), f_{x_2}(p)\}$ of Tf_p are given by

$$\ell_{ij} = \mathrm{II}_p(f_{x_i}, f_{x_j}) = S_p(f_{x_i}) \cdot f_{x_j} = -N_{x_i} \cdot f_{x_j} = N \cdot f_{x_i x_j}.$$

Hence

$$\mathrm{II} = \ell_{11}dx_1^2 + 2\ell_{12}dx_1dx_2 + \ell_{22}dx_2^2,$$

which is called the *second fundamental form of* f.

Example 2.6.3. (**The graph of u**). Let $u : \mathcal{O} \to \mathbb{R}$ be a smooth function, and $f : \mathcal{O} \to \mathbb{R}^3$ the graph of u, i.e., $f(x_1, x_2) = (x_1, x_2, u(x_1, x_2))$. A direct computation implies that

$$f_{x_1} = (1, 0, u_{x_1}), \quad f_{x_2} = (0, 1, u_{x_2}), \quad N = \frac{f_{x_1} \times f_{x_2}}{||f_{x_1} \times f_{x_2}||} = \frac{(-u_{x_1}, -u_{x_2}, 1)}{\sqrt{1 + u_{x_1}^2 + u_{x_2}^2}}.$$

$$f_{x_i x_j} = (0, 0, u_{x_i x_j}).$$

So we have

$$g_{11} = f_{x_1} \cdot f_{x_1} = 1 + u_{x_1}^2, \quad g_{12} = f_{x_1} \cdot f_{x_2} = u_{x_1} u_{x_2}, \quad g_{22} = f_{x_2} \cdot f_{x_2} = 1 + u_{x_2}^2,$$

$$\ell_{ij} = N \cdot f_{x_i x_j} = \frac{u_{x_i x_j}}{\sqrt{1 + u_{x_1}^2 + u_{x_2}^2}}$$

and the two fundamental forms are

$$\mathrm{I} = \sum_{i,j=1}^{2} g_{ij} dx_i dx_j = (1 + u_{x_1}^2) dx_1^2 + 2u_{x_1} u_{x_2} dx_1 dx_2 + (1 + u_{x_2}^2) dx_2^2,$$

$$\mathrm{II} = \sum_{i,j=1}^{2} \ell_{ij} dx_i dx_j = \frac{1}{\sqrt{1 + u_{x_1}^2 + u_{x_2}^2}} (u_{x_1 x_1} dx_1^2 + 2u_{x_1 x_2} dx_1 dx_2 + u_{x_2 x_2} dx_2^2).$$

The area of the surface $f(D)$ is

$$\int\int_D \sqrt{g_{11} g_{22} - g_{12}^2}\; t dx_1 dx_2 = \int\int_D \sqrt{1 + u_{x_1}^2 + u_{x_2}^2}\; dx_1 dx_2.$$

The meaning of the second fundamental form II

Fix $p_0 \in \mathcal{O}$ and a unit tangent vector ξ of f at p_0, let σ denote the intersection of the surface $f(\mathcal{O})$ and the plane E spanned by ξ and the normal vector $N(p_0)$. Then σ is a curve lies on the plane E and will be called the *plane section* of f at p_0 defined by ξ.

Theorem 2.6.4. *(Meusnier's Theorem) The curvature of the plane section of a parametrized surface* $f : \mathcal{O} \to \mathbb{R}^3$ *at* p_0 *defined by a unit tangent vector* ξ *in* Tf_{p_o} *is equal to* $\mathrm{II}_{p_0}(\xi, \xi)$.

Proof. We may assume there is $c : (-\epsilon, \epsilon) \to \mathcal{O}$ such that $c(0) = p_0$, $\sigma(s) = f(c(s)) = f(c_1(s), c_2(s))$ parametrized by arc-length is the plane section defined by ξ. Then

$$\sigma'(s) = e_1(s) = f_{x_1}(c_1(s), c_2(s))c_1'(s) + f_{x_2}(c_1(s), c_2(s))c_2'(s),$$

$$\sigma''(s) = e_1'(s) = f_{x_1 x_1}(c(s))c_1'(s)^2 + 2f_{x_1 x_2}(c(s))c_1'(s)c_2'(s) + f_{x_2 x_2}(c(s))c_2'(s)^2.$$

The normal of σ at $s = 0$ is $N(p_0)$, i.e., $e_2(0) = N(p_0)$. But the curvature of σ at $s = 0$ is

$$\begin{aligned} sk(0) &= e_1'(0) \cdot e_2(0) = \sigma''(0) \cdot N(p_0) \\ &= c_1'(0)^2 f_{x_1 x_1}(p_0) \cdot N(p_0) + 2c_1'(0)c_2'(0) f_{x_1 x_2}(p_0) \cdot N(p_0) + c_2'(0)^2 f_{x_2 x_2}(p_) \cdot N(p_0) \\ &= \ell_{11}(p_0) c_1'(0)^2 + 2\ell_{12}(p_0) c_1'(0) c_2'(0) + \ell_{22} c_2'(0)^2 = \mathrm{II}_{p_0}(\xi, \xi). \end{aligned}$$

$\square$

Example 2.6.5.

(1) Let $f(x, y) = (x, y, 0)$ (a plane). Then $N = (0, 0, 1)$ and any plane section is a straight line. So $\mathrm{II}(\xi, \xi) = 0$ for all unit tangent vector ξ, thus II is zero.
(2) Let $f(x, y) = (\cos x, \sin x, y)$ (the cylinder with unit circle as the base). A simple computation implies that $N(x, y) = (-\sin x, \cos x, 0)$ is the unit normal. The plane section of f defined by $\xi = f_{x_2}(p_0) = (0, 0, 1)$ is a straight line, so the curvature is zero. Thus $\mathrm{II}_{p_0}(\xi, \xi) = 0$. The plane section of f defined by $\eta(p_0) = (-\sin x, \cos x, 0)$ is a circle of radius 1, so its curvature is 1, which implies that $\mathrm{II}_{p_0}(\eta, \eta) = 1$.
(3) Let $\mathcal{O} = \{(x, y) \mid x^2 + y^2 < 1\}$, and $f(x, y) = (x, y, \sqrt{1 - x^2 - y^2})$ (a hemi-sphere). Every plane section is a part of the great circle, hence the curvature is 1, which implies that $\mathrm{II}(\xi, \xi) = 1$ for all unit tangent vector ξ.

Example 2.6.6. Let $\mathcal{O} = \{(x_1, x_2) \mid x_1 \in (0, 2\pi), x_2 \in \mathbb{R}\}$, $f, h : \mathcal{O} \to \mathbb{R}^3$ defined by $f(x_1, x_2) = (x_1, x_2, 0)$ and $h(x_1, x_2) = (\cos x_1, \sin x_1, x_2)$. So f is the plane region $(0, 2\pi) \times \mathbb{R} \times \{0\}$ and h is the the cylinder minus the line $\{(1, 0, t) \mid t \in \mathbb{R}\}$. A simple computation implies that The first and second fundamental forms for f and h are

$$\begin{aligned} \mathrm{I} &= dx_1^2 + dx_2^2, \quad \mathrm{II} = 0, \\ \tilde{\mathrm{I}} &= dx_1^2 + dx_2^2, \ \tilde{}\ \mathrm{II} = dx_1^2, \end{aligned}$$

respectively. Note that f and h have the same first fundamental form, so the geometry involving arc length, angles and areas on the two surfaces are identical. The shortest curve jointing two points $(p, 0), (q, 0)$ in the surface f is the straight line $t \mapsto (p + t(q - p), 0)$, so the shortest curve joining p to q in the surface given by g is also $t \mapsto h(p + t(q - p))$. Thus the geometry of triangles on f and on h are the same. In fact, the map $f(x, y) \mapsto h(x, y)$ is an "isometry" (preserving the first fundamental form). Properties that depend only on the first fundamental form are called *intrinsic properties*.

2.7. **Eigenvalues and eigenvectors.**

Let V be a vector space, and $S : V \to V$ a linear map. A non-zero vector $v \in V$ is an *eigenvector* of S with *eigenvalue* λ_0 if $S(v) = \lambda_0 v$.

Recall that if A is an $n \times n$ real matrix, then for $\lambda_0 \in \mathbb{R}$, $\det(A - \lambda_0 \mathrm{I}) = 0$ if and only if there exists non-zero $u \in \mathbb{R}^n$ such that $Au = \lambda_0 u$. We call λ_0 an eigenvalue and u an eigenvector of the matrix A. If $S : \mathbb{R}^n \to \mathbb{R}^n$ is the linear map defined by $S(x) = Ax$, then eigenvalues and eigenvectors of A are the same as for the linear opearator S. For a linear map on a general vector space, we have

Proposition 2.7.1. *Let A denote the matrix associated to the linear operator $S : V \to V$ with respect to the basis $\{v_1, \ldots, v_n\}$. Then*

(1) *$\lambda_0 \in \mathbb{R}$ is an eigenvalues of S if and only if* $\det(A - \lambda_0 \mathrm{I}) = 0$,
(2) *if $Au = \lambda_0 u$, then $v = u_1 v_1 + \ldots + u_n v_n$ is an eigenvector of S, where $u = (u_1, \ldots, u_n)^t$.*

Proof. By definition, $S(v_i) = \sum_{j=1}^n a_{ji} v_j$. Since S is linear, we have

$$S(\sum_{i=1}^n u_i v_i) = \sum_{i=1}^n u_i S(v_i) = \sum_{i=1}^n u_i \sum_{j=1}^n a_{ji} v_j = \sum_{j=1}^n \sum_{i=1}^n a_{ji} u_i \ v_j.$$

But $\sum_{i=1}^n a_{ji} u_i$ is the j1-th entry of Au, which is $\lambda_0 u$, so it is equal to $\lambda_0 u_j$. Thus $S(\sum_{i=1}^n u_i v_i) = \lambda_0 \sum_{j=1}^n u_j v_j$, and $\sum_{i=1}^n u_i v_i$ is an eigenvector of S with eigenvalue λ_0. □

Proposition 2.7.2. *Let* $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ *(real, symmetric* 2×2 *matrix). Then*

(1) *A has two real eigenvalues λ_1, λ_2,*
(2) *$\lambda_1 \lambda_2 = \det(A)$, $\lambda_1 + \lambda_2 = \mathrm{tr}(A) = a + c$,*
(3) *there is an o.n. basis $\{v_1, v_2\}$ of $\mathbb{R}^2$ such that v_i is an eigenvector with eigenvalue λ_i for $i = 1, 2$ (we call $\{v_1, v_2\}$ an o.n. eigenbase of A).*

Proof. To find eigenvalues for A, we need to solve

$$\det(A - \lambda \mathrm{I}) = \begin{vmatrix} a - \lambda & b \\ b & c - \lambda \end{vmatrix} = (a - \lambda)(c - \lambda) - b^2 = \lambda^2 - (a + c)\lambda + ac - b^2 = 0.$$

Since the discriminant $\triangle$ is $(a+c)^2 - 4(ac - b^2) = (a-c)^2 + 4b^2 \geq 0$, this quadratic polynomial has two real roots:

$$\lambda_1 = \frac{1}{2}(a + c + \sqrt{(a-c)^2 + b^2}), \quad \lambda_1 = \frac{1}{2}(a + c - \sqrt{(a-c)^2 + b^2}).$$

This proves (1) and (2).

If the $\triangle$ is zero, then $a = c$ and $b = 0$, so $A = a\mathrm{I}$ and $\{(1,0)^t, (0,1)^t\}$ is an o.n. eigenbase of A. If $\triangle > 0$, then $(A - \lambda_i)\xi_i = 0$ can be solved easily:

$$\xi_1 = \left(-b, \frac{a - c - \sqrt{(a-c)^2 + b^2}}{2}\right)^t, \quad \xi_2 = \left(-b, \frac{a - c + \sqrt{(a-c)^2 + b^2}}{2}\right)^t$$

are eigenvectors of A with eigenvalues λ_1, λ_2 respectively. But $\xi_1 \cdot \xi_2 = 0$, so $\{v_1, v_2\}$ is an o.n. eigenbase for A, where $\{v_1 = \xi_1 / ||\xi_1||$ and $v_2 = \xi_2 / ||\xi_2||\}$. □

The above Proposition is a special case of the following general Theorem in linear algebra:

Theorem 2.7.3. *If A is a real, symmetric $n \times n$ matrix, then A has n real eigenvalues $\lambda_1, \ldots, \lambda_n$ and there is basis $\{u_1, \ldots, u_n\}$ of $\mathbb{R}^n$ such that $Au_i = \lambda_i u_i$ for all $1 \leq i \leq n$. (We call such basis an o.n. eigenbase for A).*

We have shown that the matrix associated to a self-adjoint operator with respect to an o.n. basis is symmetric. So we conclude:

Theorem 2.7.4. (Spectral Theorem)

Let $(\ ,)$ be an inner product on V, $\dim(V) = n$, and $S : V \to V$ a linear self-adjoint operator. Then

(1) *S has n real eigenvalues $\lambda_1, \ldots, \lambda_n$,*
(2) *there is an o.n. basis $\{v_1, \ldots, v_n\}$ of V such that $S(v_i) = \lambda_i v$ for $1 \leq i \leq n$.*

Proposition 2.7.5. *Let $(\ ,\ )$ be an inner product on V, $\dim(V) = n$, $S : V \to V$ a linear self-adjoint operator, $\{v_1, \ldots, v_n\}$ an o.n. eigenbase of S, and $S(v_i) = \lambda_i v_i$ for $1 \le i \le n$. Let $b^S : V \times V \to \mathbb{R}$ be the symmetric bilinear form associated to S, i.e., $b^S(\xi, \eta) = (S(\xi), \eta)$. Then*

$$\min_{||v||=1} b^S(v,v) = \lambda_1 = b^S(v_1), \qquad \max_{||v||=1} b^S(v,v) = \lambda_n = b^S(v_n).$$

Proof. Let $\{v_1, \ldots, v_n\}$ be an o.n. eigenbase of S with eigenvalues $\lambda_1, \ldots, \lambda_n$ respectively. Then

$$b^S\left(\sum_{i=1}^n x_i v_i, \sum_{i=1}^n x_i v_i\right) = \left(S\left(\sum_{i=1}^n x_i v_i\right), \sum_{j=1}^n x_j v_j\right) = \left(\sum_{i=1}^n x_i \lambda_i v_i, \sum_{j=1}^n x_j v_j\right)$$
$$= \sum_{i=1}^n \lambda_i x_i^2.$$

The norm $||v|| = 1$ if and only if $\sum_{i=1}^n x_i^2 = 1$. But the maximum and minimum value of $\lambda_1 x_1^2 + \ldots + \lambda_n x_n^2$ on the set $\sum_{i=1}^n x_i^2 = 1$ are λ_n and λ_1 respectively. □

2.8. **Principal, Gaussian, and mean curvatures.**

Since the shape operator S_p is a self-adjoint operator on the tangent plane Tf_p of a parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$, as a consequence of the Spectral Theorem we have

Proposition 2.8.1. *The shape operator of a parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$ at $f(p)$ has two real eigenvalues and an o.n. eigenbase.*

Definition 2.8.2.

(1) The eigenvalues k_1, k_2 of the shape operator S_p of the parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$ at p are called the *principal curvatures* and *principal directions*.
(2) The *Gaussian curvature* of f is $K = k_1 k_2$.
(3) The *mean curvature* of f is $H = k_1 + k_2$.
(4) The principal directions of f at p are the unit eigenvectors v_1, v_2 of the shape operator S_p.

We have shown before that the matrix of the shape operator S with respect to the basis $\{f_{x_1}, f_{x_2}\}$ is $A = G^{-1}L$, where $G = (g_{ij})$, $L = (\ell_{ij})$, g_{ij} and ℓ_{ij} are the coefficients for I and II. Note that if $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible, then it is easy to check that

$$B^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

By Proposition 2.7.1, $K = k_1 k_2 = \det(A)$ and $H = k_1 + k_2 = \operatorname{tr}(A)$, so

$$K = \frac{det(\ell_{ij})}{\det(g_{ij})} = \frac{\ell_{11}\ell_{22} - \ell_{12}^2}{g_{11}g_{22} - g_{12}^2}. \tag{2.8.1}$$

We compute the matrix A of the shape operator:

$$\begin{aligned} A = G^{-1}L &= \frac{1}{g_{11}g_{22} - g_{12}^2} \begin{pmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{pmatrix} \begin{pmatrix} \ell_{11} & \ell_{12} \\ \ell_{12} & \ell_{22} \end{pmatrix} \\ &= \frac{1}{g_{11}g_{22} - g_{12}^2} \begin{pmatrix} g_{22}\ell_{11} - g_{12}\ell_{12} & g_{22}\ell_{12} - g_{12}\ell_{22} \\ -g_{12}\ell_{11} + g_{11}\ell_{12} & -g_{12}\ell_{12} + g_{11}\ell_{22} \end{pmatrix}, \end{aligned}$$

so

$$H = \frac{g_{22}\ell_{11} - 2g_{12}\ell_{12} + g_{11}\ell_{22}}{g_{11}g_{22} - g_{12}^2}.$$

It follows from Proposition 2.7.1 that

(1) the principal curvatures k_1, k_2 are eigenvalues of $A = G^{-1}L$,

(2) if $\begin{pmatrix} r_1 \\ s_1 \end{pmatrix}$ and $\begin{pmatrix} r_2 \\ s_2 \end{pmatrix}$ are eigenvectors of $A = G^{-1}L$ with eigenvalues k_1, k_2, then $v_1 = r_1 f_{x_1} + s_1 f_{x_2}$ and $v_2 = r_2 f_{x_1} + s_2 f_{x_2}$ are in principal directions.

Example 2.8.3. For cylinder $f(x_1, x_2) = (\cos x_1, \sin x_1, x_2)$, we have

$$\begin{aligned} &f_{x_1} = (-\sin x_1, \cos x_1, 0), \quad f_{x_2} = (0, 0, 1), \\ &f_{x_1x_1} = (-\cos x_1, -\sin x_1, 0), \quad f_{x_1x_2} = (0, 0, 0), \quad f_{x_2x_2} = (0, 0, 0), \\ &N = \frac{f_{x_1} \times f_{x_2}}{||f_{x_1} \times f_{x_2}||} = (\cos x_1, \sin x_1, 0), \end{aligned}$$

so

$$\begin{aligned} &g_{11} = g_{22} = 1, \quad g_{12} = 0, \\ &\ell_{11} = -1, \quad \ell_{12} = \ell_{22} = 0, \end{aligned}$$

and the matrix for the shape operator is $A = G^{-1}L = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$. So the principal curvatures $k_1 = -1, k_2 = 0$ and f_{x_1}, f_{x_2} are the principal directions.

Example 2.8.4. If f is a parametrization of a piece of a sphere of radius r centered at the origin, then the unit normal vector at $f(p)$ is $N(p) = \frac{1}{r} f(p)$. So the shape operator $S_p(f_{x_i}) = N_{x_i} = \frac{1}{r} f_{x_i}$ for $i = 1, 2$. This implies that $S = -r\mathrm{Id}$ and $\mathrm{II}(\xi, \eta) = S(\xi) \cdot \eta = -r\xi \cdot \eta = -r\mathrm{I}$. The eigenvalues of S are $-\frac{1}{r}, -\frac{1}{r}$, hence $H = -2/r$ and $K = 1/r^2$.

Example 2.8.5. For $f(x, y) = (x, y, x^2 + y^2)$, we have

$$\begin{aligned} &f_x = (1, 0, 2x), \quad f_y = (0, 1, 2y), \quad N = \frac{(-2x, -2y, 1)}{\sqrt{1 + 4x^2 + 4y^2}}, \\ &f_{xx} = (0, 0, 2), \quad f_{xy} = 0, \quad f_{yy} = (0, 0, 2). \end{aligned}$$

So

$$\begin{aligned} &g_{11} = 1 + 4x^2, \quad g_{12} = 4xy, \quad g_{22} = 1 + 4y^2, \\ &\ell_{11} = \ell_{22} = \frac{2}{\sqrt{1 + 4x^2 + 4y^2}}, \quad \ell_{12} = 0, \\ &K = \frac{\det(\ell_{ij})}{\det(g_{ij})} = \frac{4}{(1 + 4x^2 + 4y^2)^2}, \quad H = \frac{4(1 + 2x^2 + 2y^2)}{(1 + 4x^2 + 4y^2)^{3/2}}. \end{aligned}$$

Exercise 2.8.1.

(1) Let $h : (0, \infty) \to \mathbb{R}$ be a smooth function, $\mathcal{O} = \{(x, y) \mid x > 0, y \in \mathbb{R}\}$, and $f : \mathcal{O} \to \mathbb{R}^3$ defined by

$$f(x, y) = \left(x, y, xh\left(\frac{y}{x}\right)\right).$$

Prove that all tangent planes of f pass through the origin $(0, 0, 0)$.

(2) Let $\mathcal{O} = \{(x, y) \mid x^2 + y^2 < 1\}$. Prove that if all normals to a parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$ pass through a fixed point q_0, then $f(\mathcal{O})$ must lie in a sphere centered at q_0.

Exercise 2.8.2.

Compute $\mathrm{I}, \mathrm{II}, H, K$ of the following parametrized surfaces:

(1) Let $\alpha : (a, b) \to \mathbb{R}^3$ be a smooth curve parametrized by arc-length with positive curvature, $\mathcal{O} = (a, b) \times (0, \infty)$, and $f : \mathcal{O} \to \mathbb{R}^3$ defined by

$$f(s, t) = \alpha(s) + t\alpha'(s).$$

(2) Let $\mathcal{O} = (0, 2\pi) \times (0, \pi)$, $f : \mathcal{O} \to \mathbb{R}^3$ defined by

$$f(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi).$$

(This is a sphere)

(3) $f(u, v) = (u - \frac{u^3}{3} + uv^2, v - \frac{v^3}{3} + vu^2, u^2 - v^2)$.

(4) Let $\alpha(s) = (x(s), y(s))$ be a plane curve parametrized by arc length such that $x(s) > 0$, and $f(s, t) = (x(s)\cos t, y(s), x(s)\sin t)$ (the surface obtained by rotating the curve α about the y-axis).

3. Fundamental Theorem of Surfaces in $\mathbb{R}^3$

3.1. Frobenius Theorem.

Let $\mathcal{O}$ be an open subset of $\mathbb{R}^2$, $f, g : \mathcal{O} \to \mathbb{R}$ smooth maps, $(x_0, y_0) \in \mathcal{O}$, and $c_0 \in \mathbb{R}$. Then the initial value problem for the following ODE system,

$$\begin{cases} \frac{\partial u}{\partial x} = f(x, y), \\ \frac{\partial u}{\partial y} = g(x, y), \\ u(x_0, y_0) = c_0, \end{cases}$$

has a smooth solution defined in some disk centered at (x_0, y_0) for any given $(x_0, y_0) \in \mathcal{O}$ if and only if f, g satisfy the *compatibility condition*

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x}$$

in $\mathcal{O}$. Moreover, we can use integration to find the solution as follows: Suppose $u(x, y)$ is a solution, then the Fundamental Theorem of Calculus implies that

$$u(x, y) = v(y) + \int_{x_0}^{x} f(t, y)\ dt$$

for some $v(y)$ such that $v(y_0) = c_0$. But

$$u_y = v'(y) + \int_{x_0}^{x} \frac{\partial f}{\partial y}\ dt = v'(y) + \int_{x_0}^{x} \frac{\partial g}{\partial x}\ dx = v'(y) + g(x,y) - g(x_0,y)$$

should be equal to $g(x,y)$, so $v'(y) = g(x_0,y)$. But $v(y_0) = c_0$, hence

$$v(y) = c_0 + \int_{y_0}^{y} g(x_0,s)ds.$$

In other words, the solution for the initial value problem is

$$u(x,y) = c_0 + \int_{y_0}^{y} g(x_0,s)\ ds + \int_{x_0}^{x} f(t,y)\ dt.$$

Given smooth maps $A, B; \mathcal{O} \times \mathbb{R} \to \mathbb{R}$, we now consider the following first order PDE system for $u : \mathcal{O} \to \mathbb{R}$:

$$\begin{cases} \frac{\partial u}{\partial x} = A(x,y,u(x,y)), \\ \frac{\partial u}{\partial y} = B(x,y,u(x,y)). \end{cases} \tag{3.1.1}$$

If we have a smooth solution u for (3.1.1), then $(u_x)_y = (u_y)_x$. But

$$\begin{aligned} &(u_x)_y = (A(x,y,u(x,y))_y = A_y + A_u u_y = A_y + A_u B \\ &= (u_y)_x = (B(x,y,u(x,y))_x = B_x + B_u u_x = B_x + B_u A. \end{aligned}$$

Thus A, B must satisfy the following condition:

$$A_y + A_u B = B_x + B_u A. \tag{3.1.2}$$

The Frobenius Theorem states that (3.1.2) is both a necessary and sufficient condition for the first order PDE system (3.1.1) to be solvable. We will see from the proof of this theorem that, although we are dealing with a PDE, the algorithm to construct solutions of this PDE is to first solve an ODE system in x variable (first equation of (3.1.1) on the line $y = y_0$), and then solve a family of ODE systems in y variables (the second equation of (3.1.1) for each x). The condition (3.1.2) guarantees that this process produces a solution of (3.1.1). We state the Theorem for $u : \mathcal{O} \to \mathbb{R}^n$:

Theorem 3.1.1. (Frobenius Theorem) *Let $U_1 \subset \mathbb{R}^2$ and $U_2 \subset \mathbb{R}^n$ be open subsets, $A = (A_1, \ldots, A_n), B = (B_1, \ldots, B_n) : U_1 \times U_2 \to \mathbb{R}^n$ smooth maps, $(x_0, y_0) \in U_1$, and $p_0 \in U_2$. Then the following first order system*

$$\begin{cases} \frac{\partial u}{\partial x} = A(x,y,u(x,y)), \\ \frac{\partial u}{\partial y} = B(x,y,u(x,y)), \\ u(x_0,y_0) = p_0, \end{cases} \tag{3.1.3}$$

has a smooth solution for u in a neighborhood of (x_0, y_0) for all possible $(x_0, y_0) \in U_1$ and $p_0 \in U_2$ if and only if

$$(A_i)_y + \sum_{j=1}^{n} \frac{\partial A_i}{\partial u_j} B_j = (B_i)_x + \sum_{j=1}^{n} \frac{\partial B_i}{\partial u_j} A_j, \qquad 1 \le i \le n, \tag{3.1.4}$$

hold identically on $U_1 \times U_2$.

Equation (3.1.3) written in coordinates gives the following system:

$$\begin{cases} \frac{\partial u_i}{\partial x} = A_i(x, y, u_1(x,y), \ldots, u_n(x,y)), & 1 \le i \le n, \\ \frac{\partial u_i}{\partial y} = B_i(x, y, u_1(x,y), \ldots, u_n(x,y)), & 1 \le i \le n, \\ u_i(x_0, y_0) = p_i^0, \end{cases}$$

where $p_0 = (p_1^0, \ldots, p_n^0)$.

We call (3.1.4) the *compatibility condition* for the first order PDE (3.1.3).

To prove the Frobenius Theorem we need to solve a family of ODEs depending smoothly on a parameter, and we need to know whether the solutions depend smoothly on the initial data and the parameter. This was answered by the following Theorem in ODE:

Theorem 3.1.2. *Let $\mathcal{O}$ be an open subset of $\mathbb{R}^n$, $t_0 \in (a_0, b_0)$, and $f : [a_0, b_0] \times \mathcal{O} \times [a_1, b_1] \to \mathbb{R}^n$ a smooth map. Given $p \in \mathcal{O}$ and $r \in [a_1, b_1]$, let $y^{p,r}$ denote the solution of*

$$\frac{dy}{dt} = f(t, y(t), r), \quad y(t_0) = p,$$

and $u(t, p, r) = y^{p,r}(t)$. Then u is smooth in t, p, r.

Proof. **Proof of Frobenius Theorem**

If $u = (u_1, \ldots, u_n)$ is a smooth solution of (3.1.3), then $\frac{\partial}{\partial y}\frac{\partial u_i}{\partial x} = \frac{\partial}{\partial x}\frac{\partial u}{\partial y}$. Use the chain rule to get

$$\begin{aligned} \frac{\partial}{\partial y}\frac{\partial u_i}{\partial x} &= \frac{\partial}{\partial y} A_i(x, y, u_1(x,y), \ldots, u_n(x,y)) = \frac{\partial A_i}{\partial y} + \sum_{j=1}^n \frac{\partial A_i}{\partial u_j}\frac{\partial u_j}{\partial y} \\ &= \frac{\partial A_i}{\partial y} + \sum_{j=1}^n \frac{\partial A_i}{\partial u_j} B_j, \\ \frac{\partial}{\partial x}\frac{\partial u_i}{\partial y} &= \frac{\partial}{\partial x} B_i(x, y, u(x,y)) = \frac{\partial B_i}{\partial x} + \sum_{j=1}^n \frac{\partial B_i}{\partial u_j}\frac{\partial u_j}{\partial x} \\ &= \frac{\partial B_i}{\partial x} + \sum_{j=1}^n \frac{\partial B_i}{\partial u_j} A_j, \end{aligned}$$

so the compatibility condition (3.1.4) must hold.

Conversely, assume A, B satisfy (3.1.4). To solve (3.1.3), we proceed as follows: The existence and uniqueness Theorem of solutions of ODE implies that there exist $\delta > 0$ and $\alpha : (x_0 - \delta, x_0 + \delta) \to U_2$ satisfying

$$\begin{cases} \frac{d\alpha}{dx} = A(x, y_0, \alpha(x)), \\ \alpha(x_0) = p_0. \end{cases} \tag{3.1.5}$$

For each fixed $x \in (x_0 - \delta, x_0 + \delta)$, let $\beta^x(y)$ denote the unique solution of the following ODE in y variable:

$$\begin{cases} \frac{d\beta^x}{dy} = B(x, y, \beta^x(y)), \\ \beta^x(y_0) = \alpha(x). \end{cases} \tag{3.1.6}$$

Set $u(x, y) = \beta^x(y)$. Note that (3.1.6) is a family of ODEs in y variable depending on the parameter x and B is smooth, so by Theorem 3.1.2, u is smooth in x, y.

Hence $(u_x)_y = (u_y)_x$. By construction, u satisfies the second equation of (3.1.3) and $u(x_0, y_0) = p_0$. It remains to prove u satisfies the first equation of (3.1.3). We will only prove this for the case $n = 1$, and the proof for general n is similar. First let

$$z(x, y) = u_x - A(x, y, u(x, y)).$$

But

$$\begin{aligned} z_y &= (u_x - A(x, y, u))_y = u_{xy} - A_y - A_u u_y = (u_y)_x - (A_y + A_u B) \\ &= (B(x, y, u))_x - (A_y + A_u B) = B_x + B_u u_x - (A_y + A_u B) \\ &= B_x + B_u u_x - (B_x + B_u A) = B_u(u_x - A) = B_u(x, y, u(x, y))z. \end{aligned}$$

This proves that for each x, $h^x(y) = z(x, y)$ is a solution of the following differential equation:

$$\frac{dh}{dy} = B_u(x, y, u(x, y))h. \tag{3.1.7}$$

Since α satisfies (3.1.5),

$$z(x, y_0) = u_x(x, y_0) - A(x, y_0, u(x, y_0)) = \alpha'(x) - A(x, y_0, \alpha(x)) = 0.$$

So h^x is the solution of (3.1.7) with initial data $h^x(y_0) = 0$. We observe that the zero function is also a solution of (3.1.7) with 0 initial data, so by the uniqueness of solutions of ODE we have $h^x = 0$, i.e., $z(x, y) = 0$, hence u satisfies the second equation of (3.1.3). □

Remark 3.1.3. The proof of Theorem 3.1.1 gives the following algorithm to construct numerical solution of (3.1.3):

1. Solve the ODE (3.1.5) on the horizontal line $y = y_0$ by a numerical method (for example Runge-Kutta) to get $u(x_k, y_0)$ for $x_k = x_0 + k\epsilon$ where ϵ is the step size in the numerical method.
2. Solve the ODE system (3.1.6) on the vertical line $x = x_k$ for each k to get the value $u(x_k, y_m)$.

If A and B satisfies the compatibility condition, then u solves (3.1.3).

Let $gl(n)$ denote the space of $n \times n$ real matrices. Note that $gl(n)$ can be identified as $\mathbb{R}^{n^2}$. For $P, Q \in gl(n)$, let $[P, Q]$ denote the *commutator* (also called the *bracket*) of P and Q defined by

$$[P, Q] = PQ - QP.$$

Corollary 3.1.4. *Let U be an open subset of $\mathbb{R}^2$, $(x_0, y_0) \in U$, $C \in gl(n)$, and $P, Q : U \to gl(n)$ smooth maps. Then the following initial value problem for $u : U \to gl(n)$*

$$\begin{cases} u_x = u(x, y)P(x, y), \\ u_y = u(x, y)Q(x, y), \\ u(x_0, y_0) = C \end{cases} \tag{3.1.8}$$

has a smooth solution u defined in some small disk centered at (x_0, y_0) for all possible (x_0, y_0) in U and $C \in gl(n)$ if and only if

$$P_y - Q_x = [P, Q]. \tag{3.1.9}$$

Proof. This Corollary follows from Theorem 3.1.1 with $A(x, y, u) = uP(x, y)$ and $B(x, y, u) = uQ(x, y)$. We can also compute the mixed derivatives directly as follows:

$$(u_x)_y = (uP)_y = u_y P + uP_y = (uQ)P + uP_y = u(QP + P_y),$$
$$(u_y)_x = (uQ)_x = u_x Q + uQ_x = (uP)Q + uQ_x = u(PQ + Q_x).$$

Thus $u(QP + P_y) = u(PQ + Q_x)$. So the compatibility condition is

$$QP + P_y = PQ + Q_x,$$

which is (3.1.9). □

Remark 3.1.5. Given $gl(3)$-valued smooth maps $P = (p_{ij})$ and $Q = (q_{ij})$, equation (3.1.8) is a system of 9 equations for 18 functions p_{ij}, q_{ij}. This is the type of equation we need for the Fundamental Theorem of surfaces in $\mathbb{R}^3$. However, if P and Q are skew symmetric, i.e., $P^T = -P$ and $Q^T = -Q$, then $[P, Q]$ is also skew-symmetric because

$$\begin{aligned}[P, Q]^T &= (PQ - QP)^T = Q^T P^T - P^T Q^T = (-Q)(-P) - (-P)(-Q) \\ &= QP - PQ = -[P, Q].\end{aligned}$$

In this case, equation (3.1.8) becomes a system of 3 first order PDE involving six functions $p_{12}, p_{13}, p_{23}, q_{12}, q_{13}, q_{23}$.

Proposition 3.1.6. *Let $\mathcal{O}$ be an open subset of $\mathbb{R}^2$, and $P, Q : \mathcal{O} \to gl(n)$ smooth maps such that $P^T = -P$ and $Q^T = -Q$. Suppose P, Q satisfy the compatibility condition* (3.1.9)*, and the initial data C is an orthogonal matrix. If $u : \mathcal{O}_0 \to gl(n)$ is the solution of* (3.1.8)*, then $u(x, y)$ is an orthogonal matrix for all $(x, y) \in \mathcal{O}_0$.*

Proof. Set

$$\xi(x, y) = u(x, y)^T u(x, y).$$

Then $\xi(x_0, y_0) = u(x_0, y_0)^T u(x_0, y_0) = \mathrm{I}$, the identity matrix. Compute directly to get

$$\xi_x = (u_x)^T u + u^T u_x = (uP)^T u + u^T (uP) = P^T u^T u + u^T uP = P^T \xi + \xi P,$$
$$\xi_y = (u_y)^T u + u^T u_y = (uQ)^T u + u^T (uQ) = Q^T u^T u + u^T uQ = Q^T \xi + \xi Q.$$

This shows that ξ satisfies

$$\begin{cases} \xi_x = P^T \xi + \xi P, \\ \xi_y = Q^T \xi + \xi Q, \\ \xi(x_0, y_0) = \mathrm{I}. \end{cases}$$

But we observe that the constant map $\eta(x, y) = \mathrm{I}$ is also a solution of the above initial value problem. By the uniqueness part of the Frobenius Theorem, $\xi = \eta$, so $u^T u = \mathrm{I}$, i.e., $u(x, y)$ is orthogonal for all $(x, y) \in \mathcal{O}_0$. □

Next we give some applications of the Frobenius Theorem 3.1.1:

Example 3.1.7. Given $c_0 > 0$, consider the following first order PDE

$$\begin{cases} u_x = 2 \sin u, \\ u_y = \frac{1}{2} \sin u, \\ u(0, 0) = c_0. \end{cases} \tag{3.1.10}$$

This is system (3.1.3) with $A(x,y,u) = 2\sin u$, $B(x,y,u) = \frac{1}{2}\sin u$. We check the compatibility condition next:

$$A_y + A_u B = 0 + (2\cos u)(\frac{1}{2}\sin u) = \cos u \sin u,$$
$$B_x + B_u A = 0 + (\frac{1}{2}\sin u)(2\cos u) = \sin u \cos u,$$

so $A_y + A_u B = B_x + B_u A$. Thus by Frobenius Theorem, (3.1.10) is solvable. Next we use the method outlined in the proof of Frobenius Theorem to solve (3.1.10).

(i) The ODE

$$\begin{cases} \frac{d\alpha}{dx} = 2\sin\alpha, \\ \alpha(0) = c_0 \end{cases}$$

is separable, i.e., $\frac{d\alpha}{\sin\alpha} = 2dx$, so $\int \frac{d\alpha}{\sin\alpha} = \int 2dx$. This integration can be solved explicitly:

$$\alpha(x) = 2\tan^{-1}\exp(2x + c).$$

But $\alpha(0) = c_0 = 2\tan^{-1} e^c$ implies that $c = \ln(\tan\frac{c_0}{2})$.

(ii) Solve

$$\begin{cases} \frac{du}{dy} = \frac{1}{2}\sin u, \\ u(x,0) = \alpha(x) = 2\tan^{-1}(2x + c). \end{cases}$$

We can solve this exactly the same way as in (i) to get

$$u(x,y) = 2\tan^{-1}(\exp(2x + \frac{y}{2} + c)),$$

where $c = \ln(\tan\frac{c_0}{2})$.

Moreover, if u is a solution for (3.1.10), then

$$(u_x)_y = (2\sin u)_y = 2\cos u\ u_y = (2\cos u)(\frac{1}{2}\sin u) = \cos u \sin u,$$

so u satisfies the following famous non-linear wave equation, the *sine-Gordon equation* (or SGE):

$$u_{xy} = \sin u \cos u.$$

The above example is a special case of the following Theorem of Bäcklund:

Theorem 3.1.8. *Given a smooth function $q : \mathbb{R}^2 \to \mathbb{R}$ and a non-zero real constant r, the following system of first order PDE is solvable for $u : \mathbb{R}^2 \to \mathbb{R}$:*

$$\begin{cases} u_s = -q_s + r\sin(u - q), \\ u_t = q_t + \frac{1}{r}\sin(u + q). \end{cases} \tag{3.1.11}$$

if and only if q satisfies the SGE:

$$q_{st} = \sin q \cos q, \qquad \text{SGE.}$$

Moreover, the solution u of (3.1.11) *is again a solution of the* SGE.

Proof. If (3.1.11) has a C^2 solution u, then the mixed derivatives must be equal. Compute directly to see that

$$\begin{aligned} (u_s)_t &= -q_{st} + r\cos(u - q)(u - q)_t \\ &= -q_{st} + r\cos(u - q)\left(\frac{1}{r}\ \sin(u + q)\right), \end{aligned}$$

so we get

$$(u_s)_t = -q_{st} + \cos(u-q)\sin(u+q). \tag{3.1.12}$$

A similar computation implies that

$$(u_t)_s = q_{ts} + \cos(u+q)\sin(u-q). \tag{3.1.13}$$

Since $u_{st} = u_{ts}$ and $\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B$, we get

$$-q_{st} + \cos(u-q)\sin(u+q) = q_{ts} + \cos(u+q)\sin(u-q),$$

so

$$2q_{st} = \sin(u+q)\cos(u-q) - \sin(u-q)\cos(u+q) = \sin(2q) = 2\sin q \cos q.$$

In other words, the first order PDE (3.1.11) is solvable if and only if q is a solution of the SGE.

Add (3.1.12) and (3.1.13) to get

$$2u_{st} = \sin(u+q)\cos(u-q) + \sin(u-q)\cos(u+q) = \sin(2u) = 2\sin u \cos u.$$

This shows that if u is a solution of (3.1.11) then $u_{st} = \sin u \cos u$. □

The above theorem says that if we know one solution q of the SGE then we can solve the first order system (3.1.11) to construct a family of solutions of the SGE (one for each real constant r). Note that $q = 0$ is a trivial solution of the SGE. Theorem 3.1.8 implies that system (3.1.11) can be solved for u with $q = 0$, i.e.,

$$\begin{cases} u_s = r \sin u, \\ u_t = \frac{1}{r} \sin u \end{cases} \tag{3.1.14}$$

is solvable. (3.1.14) can be solved exactly the same way as in Example 3.1.7, and we get

$$u(s,t) = 2 \arctan\left(e^{rs + \frac{1}{r}\ t}\right),$$

which are solutions of the SGE. The SGE is a so called "soliton" equation, and these solutions are called "1-solitons". A special feature of soliton equations is the existence of first order systems that can generate new solutions from an old one.

3.2. **Line of curvature coordinates.**

Definition 3.2.1. A parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$ is said to be parametrized by *line of curvature coordinates* if $g_{12} = \ell_{12} = 0$, or equivalently, both the first and second fundamental forms are in diagonal forms.

If $f : \mathcal{O} \to \mathbb{R}^3$ is a surface parametrized by line of curvature coordinates, then

$$\mathrm{I} = g_{11} dx_1^2 + g_{22} dx_2^2, \quad \mathrm{II} = \ell_{11} dx_1^2 + \ell_{22} dx_2^2.$$

The prinicipal, Gaussian and mean curvatures are given by the following formulas:

$$k_1 = \frac{\ell_{11}}{g_{11}}, \quad k_2 = \frac{\ell_{22}}{g_{22}}, \quad H = \frac{\ell_{11}}{g_{11}} + \frac{\ell_{22}}{g_{22}}, \quad K = \frac{\ell_{11}\ell_{22}}{g_{11}g_{22}}.$$

Example 3.2.2. Let $u : [a, b] \to \mathbb{R}$ be a smooth function, and

$$f(y, \theta) = (u(y)\sin\theta, y, u(y)\cos\theta),$$

so the image of f is the surface of revolution obtained by rotating the curve $z = u(y)$ in the yz-plane along the y-axis. Then

$$\begin{aligned} f_y &= (u'(y)\sin\theta, 1, u'(y)\cos\theta), \\ f_\theta &= (u(y)\cos\theta, 0, -u(y)\sin\theta), \\ N &= \frac{f_y \times f_\theta}{||f_y \times f_\theta||} = \frac{(-\sin\theta, u'(y), -\cos\theta)}{\sqrt{1 + (u'(y))^2}}, \\ f_{yy} &= (u''(y)\sin\theta, 0, u''(y)\cos\theta), \\ f_{y\theta} &= (u'(y)\cos\theta, 0, -u'(y)\sin\theta), \\ f_{\theta\theta} &= (-u(y)\sin\theta, 0, -u(y)\cos\theta). \end{aligned}$$

So

$$\begin{aligned} g_{11} &= 1 + (u'(y))^2, \quad g_{22} = u(y)^2, \quad g_{12} = 0, \\ \ell_{11} &= \frac{-u''(y)}{\sqrt{1 + (u'(y))^2}}, \quad \ell_{22} = \frac{u(y)}{\sqrt{1 + (u'(y))^2}}, \quad \ell_{12} = 0, \end{aligned}$$

i.e., (y, θ) is a line of curvature coordinate system.

Proposition 3.2.3. *If the principal curvatures $k_1(p_0) \neq k_2(p_0)$ for some $p_0 \in \mathcal{O}$, then there exists $\delta > 0$ such that the ball $B(p_0, \delta)$ of radius δ centered at p_0 is contained in $\mathcal{O}$ and $k_1(p) \neq k_2(p)$ for all $p \in B(p_0, r)$.*

Proof. The Gaussian and mean curvature K and H of the parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$ are smooth. The two principal curvatures are roots of $\lambda^2 - H\lambda + K = 0$, so we may assume

$$k_1 = \frac{H + \sqrt{H^2 - 4K}}{2}, \quad k_2 = \frac{H - \sqrt{H^2 - 4K}}{2}.$$

Note that the two real roots are distinct if and only if $u = H^2 - 4K > 0$. If $k_1(p_0) \neq k_2(p_0)$, then $u(p_0) > 0$. But u is continuous, so there exists $\delta > 0$ such that $u(p) > 0$ for all $p \in B(p_0, \delta)$, thus $k_1(p) \neq k_2(p)$ in this open disk. □

A smooth map $v : \mathcal{O} \to \mathbb{R}^3$ is called a *tangent vector field* of the parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$ if $v(p) \in Tf_p$ for all $p \in \mathcal{O}$.

Proposition 3.2.4. *Let $f : \mathcal{O} \to \mathbb{R}^3$ be a parametrized surface. If its principal curvatures $k_1(p) \neq k_2(p)$ for all $p \in \mathcal{O}$, then there exist smooth, o.n., tangent vector fields e_1, e_2 on f that are $e_1(p), e_2(p)$ are eigenvector for the shape operator S_p for all $p \in \mathcal{O}$.*

Proof. This Proposition follows from the facts that

1. self-adjoint operators have an o.n. basis consisting of eigenvectors,
2. the shape operator S_p is a self-adjoint linear opeartor from Tf_p to Tf_p,
3. the formula we gave for self-adjoint operator on a 2-dimensional inner product space implies that the eigenvectors of the shape operator are smooth maps.

□

We assume the following Proposition without a proof:

Proposition 3.2.5. *Let $f : \mathcal{O} \to \mathbb{R}^3$ be a parametrized surface. Suppose $\xi_1, \xi_2 : \mathcal{O} \to \mathbb{R}^3$ are tangent vector fields of f such that $\xi_1(p), \xi_2(p)$ are linearly independent for all $p \in \mathcal{O}$. Then given any $p_0 \in \mathcal{O}$ there exists $\delta > 0$, an open subset U of $\mathbb{R}^2$, and a diffeomorphism $\phi : U \to B(p_0, \delta)$ such that ξ_1 and ξ_2 are parallel to h_{x_1} and h_{x_2} respectively, where $h = f \circ \phi : U \to \mathbb{R}^3$.*

As a consequence of the above two Propositions, we see that

Corollary 3.2.6. *Let $f : \mathcal{O} \to \mathbb{R}^3$ be a parametrized surface, and $p_0 \in \mathcal{O}$. If $k_1(p_0) \neq k_2(p_0)$, then there exist an open subset $\mathcal{O}_0$ containing p_0, an open subset U of $\mathbb{R}^2$, and a diffeomorphism $\phi : U \to \mathcal{O}_0$ such that $f \circ \phi$ is parametrized by lines of curvature coordinates. In other words, we can change coordinates (or reparametrized the surface) near p_0 by lines of curvature coordinates.*

We call $f(p_0)$ an *umbilic point* of the parametrized surface $f : \mathcal{O} \to \mathbb{R}^3$ if $k_1(p_0) = k_2(p_0)$. The above Corollary implies that away from umbilic points, we can parametrized a surface by line of curvature coordinates locally.

3.3. **The Gauss-Codazzi equation in line of curvature coordinates.**

Suppose $f : \mathcal{O} \to \mathbb{R}^3$ is a surface parametrized by line of curvature coordinates, i.e.,

$$g_{12} = f_{x_1} \cdot f_{x_2} = 0, \quad \ell_{12} = f_{x_1 x_2} \cdot N = 0.$$

We define A_1, A_2, r_1, r_2 as follows:

$$\begin{aligned} g_{11} &= f_{x_1} \cdot f_{x_1} = A_1^2, \quad g_{22} = f_{x_2} \cdot f_{x_2} = A_2^2, \\ \ell_{11} &= f_{x_1 x_1} \cdot N = r_1 A_1, \quad \ell_{22} = f_{x_2 x_2} \cdot N = r_2 A_2. \end{aligned}$$

Or equivalently,

$$A_1 = \sqrt{g_{11}}, \quad A_2 = \sqrt{g_{22}}, \quad r_1 = \frac{\ell_{11}}{A_1}, \quad r_2 = \frac{\ell_{22}}{A_2}.$$

Set

$$e_1 = \frac{f_{x_1}}{A_1}, \quad e_2 = \frac{f_{x_2}}{A_2}, \quad e_3 = N.$$

Then (e_1, e_2, e_3) is an o.n. moving frame on the surface f. Recall that if $\{v_1, v_2, v_3\}$ is an orthonormal basis of $\mathbb{R}^3$, then given any $\xi \in \mathbb{R}^3$, $\xi = \sum_{i=1}^3 a_i e_i$, where $a_i = \xi \cdot e_i$. Since $(e_i)_{x_1}$ and $(e_i)_{x_2}$ are vectors in $\mathbb{R}^3$, we can write them as linear combinations of e_1, e_2 and e_3. We use p_{ij} to denote the coefficient of e_i for $(e_j)_{x_1}$ and use q_{ij} to denote the coefficient of e_i for $(e_j)_{x_2}$, i.e.,

$$(3.3.1) \qquad \begin{cases} (e_1)_{x_1} = p_{11}e_1 + p_{21}e_2 + p_{31}e_3, \\ (e_2)_{x_1} = p_{12}e_1 + p_{22}e_2 + p_{32}e_3, \\ (e_3)_{x_1} = p_{13}e_1 + p_{23}e_2 + p_{33}e_3, \\ (e_1)_{x_2} = q_{11}e_1 + q_{21}e_2 + q_{31}e_3, \\ (e_2)_{x_2} = q_{12}e_1 + q_{22}e_2 + q_{32}e_3, \\ (e_3)_{x_2} = q_{13}e_1 + q_{23}e_2 + q_{33}e_3, \end{cases}$$

where

$$p_{ij} = (e_j)_{x_1} \cdot e_i, \quad q_{ij} = (e_j)_{x_2} \cdot e_i.$$

Recall that the matrix $P = (p_{ij})$ and $Q = (q_{ij})$ must be skew symmetric because

$$(e_i \cdot e_j)_{x_1} = 0 = (e_i)_{x_1} \cdot e_j + e_i \cdot (e_j)_{x_1} = p_{ji} + p_{ij}.$$

Next we want to show that p_{ij} and q_{ij} can be written in terms of coefficients of the first and second fundamental forms. We proceed as follows:

$$\begin{aligned}
p_{12} &= (e_2)_{x_1} \cdot e_1 = \left(\frac{f_{x_2}}{A_2}\right)_{x_1} \cdot \frac{f_{x_1}}{A_1} = \left(\frac{f_{x_2x_1}}{A_2} - \frac{f_{x_2}(A_2)_{x_1}}{A_2^2}\right) \cdot \frac{f_{x_1}}{A_1} \\
&= \frac{f_{x_1x_2} \cdot f_{x_1}}{A_1A_2} - \frac{(A_2)_{x_1}}{A_1A_2^2} f_{x_2} \cdot f_{x_1} = \frac{\frac{1}{2}(f_{x_1} \cdot f_{x_1})_{x_2}}{A_1A_2} = \frac{\frac{1}{2}(A_1^2)_{x_2}}{A_1A_2} - 0 \\
&= \frac{A_1(A_1)_{x_2}}{A_1A_2} = \frac{(A_1)_{x_2}}{A_2}. \\
p_{31} &= (e_1)_{x_1} \cdot e_3 = \left(\frac{f_{x_1}}{A_1}\right)_{x_1} \cdot e_3 = \left(\frac{f_{x_1x_1}}{A_1} - \frac{f_{x_1}(A_1)_{x_1}}{A_1^2}\right) \cdot N \\
&= \frac{f_{x_1x_1} \cdot N}{A_1} - \frac{(A_1)_{x_1}}{A_1^2}\, f_{x_1} \cdot N = \frac{\ell_{11}}{A_1} - 0 = r_1. \\
p_{32} &= (e_2)_{x_1} \cdot e_3 = \left(\frac{f_{x_2}}{A_2}\right)_{x_1} \cdot N = \left(\frac{f_{x_2x_1}}{A_2} - \frac{f_{x_2}(A_2)_{x_2}}{A_2^2}\right) \cdot N = 0
\end{aligned}$$

So we have proved that

$$p_{12} = \frac{(A_1)_{x_2}}{A_2}, \quad p_{31} = r_1, \quad p_{32} = 0.$$

In the above computations we have used $f_{x_1} \cdot f_{x_2} = 0$, $f_{x_1x_2} \cdot N = 0$, $f_{x_1} \cdot N = f_{x_2} \cdot N = 0$. Similar computation gives

$$q_{12} = -\frac{(A_2)_{x_1}}{A_1}, \quad q_{31} = 0, \quad q_{32} = r_2.$$

Since P, Q are skew-symmetric, we have

$$(3.3.2) \qquad P = \begin{pmatrix} 0 & \frac{(A_1)_{x_2}}{A_2} & -r_1 \\ -\frac{(A_1)_{x_2}}{A_2} & 0 & 0 \\ r_1 & 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & -\frac{(A_2)_{x_1}}{A_1} & 0 \\ \frac{(A_2)_{x_1}}{A_1} & 0 & -r_2 \\ 0 & r_2 & 0 \end{pmatrix}.$$

So (3.3.1) becomes

$$\begin{cases} (e_1)_{x_1} = -\frac{(A_1)_{x_2}}{A_2} e_2 + r_1 e_3, \\ (e_2)_{x_1} = \frac{(A_1)_{x_2}}{A_2} e_1, \\ (e_3)_{x_1} = -r_1 e_1, \end{cases}$$

$$\begin{cases} (e_1)_{x_2} = \frac{(A_2)_{x_1}}{A_1} e_2, \\ (e_2)_{x_2} = -\frac{(A_2)_{x_1}}{A_1} e_1 + r_2 e_3, \\ (e_3)_{x_2} = -r_2 e_2. \end{cases}$$

We can also write (3.3.1) in matrix form:

$$(3.3.3) \qquad \begin{cases} (e_1, e_2, e_3)_{x_1} = (e_1, e_2, e_3)P, \\ (e_1, e_2, e_3)_{x_2} = (e_1, e_2, e_3)Q, \end{cases}$$

where P, Q are given by (3.3.2). It follows from Corollary 1.04 [] of the section on Frobenius Theorem [] that P, Q must satisfy the compatibility condition

$$P_{x_2} - Q_{x_1} = PQ - QP.$$

Use the formula of P, Q given by (3.3.2) to compute directly to get

$$\begin{aligned}
&PQ - QP \\
&= \begin{pmatrix} 0 & p & -r_1 \\ -p & 0 & 0 \\ r_1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & q & 0 \\ -q & 0 & -r_2 \\ 0 & r_2 & 0 \end{pmatrix} - \begin{pmatrix} 0 & q & 0 \\ -q & 0 & 0 \\ 0 & r_2 & 0 \end{pmatrix} \begin{pmatrix} 0 & p & -r_1 \\ -p & 0 & 0 \\ r_1 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -r_1 r_2 & -pr_2 \\ r_1 r_2 & 0 & -qr_1 \\ pr_2 & qr_1 & 0 \end{pmatrix}, \\
&P_{x_2} - Q_{x_1} = \begin{pmatrix} 0 & p_{x_2} - q_{x_1} & -(r_1)_{x_2} \\ -(p_{x_2} - q_{x_1}) & 0 & (r_2)_{x_1} \\ (r_1)_{x_2} & -(r_2)_{x_1} & 0 \end{pmatrix},
\end{aligned}$$

where

$$p = \frac{(A_1)_{x_2}}{A_2}, \quad q = -\frac{(A_2)_{x_1}}{A_1}.$$

Since

$$PQ - QP = P_{x_2} - Q_{x_1},$$

we get

$$\begin{cases} \left(\frac{(A_1)_{x_2}}{A_2}\right)_{x_2} + \left(\frac{(A_2)_{x_1}}{A_1}\right)_{x_1} = -r_1 r_2, \\ (r_1)_{x_2} = \frac{(A_1)_{x_2}}{A_2}\, r_2, \\ (r_2)_{x_1} = \frac{(A_2)_{x_1}}{A_1}\, r_1. \end{cases} \tag{3.3.4}$$

System (3.3.4) is called the *Gauss-Codazzi equation*. So we have proved:

Theorem 3.3.1. *Let $f : \mathcal{O} \to \mathbb{R}^3$ be a surface parametrized by line of curvature coordinates, and*

$$A_1 = \sqrt{g_{11}}, \quad A_2 = \sqrt{g_{22}}, \quad r_1 = \frac{\ell_{11}}{A_1}, \quad r_2 = \frac{\ell_{22}}{A_2}.$$

Set

$$e_1 = \frac{f_{x_1}}{A_1}, \quad e_2 = \frac{f_{x_2}}{A_2}, \quad e_3 = \frac{f_{x_1} \times f_{x_2}}{||f_{x_1} \times f_{x_2}||} = e_1 \times e_2.$$

Then A_1, A_2, r_1, r_2 satisfy the Gauss-Codazzi equation (3.3.4) *and*

$$\begin{cases} (f, e_1, e_2, e_3)_{x_1} = (e_1, e_2, e_3) \begin{pmatrix} A_1 & 0 & \frac{(A_1)_{x_2}}{A_2} & -r_1 \\ 0 & -\frac{(A_1)_{x_2}}{A_2} & 0 & 0 \\ 0 & r_1 & 0 & 0 \end{pmatrix} \\ (f, e_1, e_2, e_3)_{x_2} = (e_1, e_2, e_3) \begin{pmatrix} 0 & 0 & -\frac{(A_2)_{x_1}}{A_1} & 0 \\ A_2 & \frac{(A_2)_{x_1}}{A_1} & 0 & -r_2 \\ 0 & 0 & r_2 & 0 \end{pmatrix}. \end{cases} \tag{3.3.5}$$

3.4. **Fundamental Theorem of surfaces in line of curvature coordinates.**

The converse of Theorem 3.3.1 is also true, which is the Fundamental Theorem of surfaces in $\mathbb{R}^3$ with respect to line of curvature coordinates:

Theorem 3.4.1. *Supppose A_1, A_2, r_1, r_2 are smooth functions from $\mathcal{O}$ to $\mathbb{R}$, that satisfy the Gauss-Codazzi equation (3.3.4), and $A_1 > 0$, $A_2 > 0$. Given $p_0 \in \mathcal{O}$, $y_0 \in \mathbb{R}^3$, and an o.n. basis v_1, v_2, v_3 of $\mathbb{R}^3$, then there exist an open subset $\mathcal{O}_0$ of $\mathcal{O}$ containing p_0 and a unique solution $(f, e_1, e_2, e_3) : \mathcal{O}_0 \to (\mathbb{R}^3)^4$ of (3.3.5) that satisfies the initial condition*

$$(f, e_1, e_2, e_3)(p_0) = (y_0, v_1, v_2, v_3).$$

Moreover, f is a parametrized surface and its first and second fundmental forms are

$$\mathrm{I} = A_1^2\ dx_1^2 + A_2^2\ dx_2^2, \quad \mathrm{II} = r_1 A_1\ dx_1^2 + r_2 A_2\ dx_2^2.$$

Proof. We have proved that the compatibility condition for (3.3.3) is the Gauss-Codazzi equation (3.3.4), so by the Forbenius Theorem (3.3.3) is solvable. Let (e_1, e_2, e_3) be the solution with initial data

$$(e_1, e_2, e_3)(p_0) = (v_1, v_2, v_3).$$

Since P, Q are skew-symmetric and (v_1, v_2, v_3) is an orthogonal matrix, by Proposition 1.0.6 of the section on Frobenius Theorem that the solution $(e_1, e_2, e_3)(p)$ is an orthogonal matrix for all $p \in \mathcal{O}$. To construct the surface, we need to solve

$$\begin{cases} f_{x_1} = A_1 e_1, \\ f_{x_2} = A_2 e_2. \end{cases} \tag{3.4.1}$$

Note that the right hand side is known, so this system is solvable if and only if

$$(A_1 e_1)_{x_2} = (A_2 e_2)_{x_1}.$$

To see this, we compute

$$\begin{aligned} (A_1 e_1)_{x_2} &= (A_1)_{x_2} e_1 + A_1 (e_1)_{x_2} = (A_1)_{x_2} e_1 + A_1\ \frac{(A_2)_{x_1}}{A_1} e_2 \\ &= (A_1)_{x_2} e_1 + (A_2)_{x_1} e_2, \\ (A_2 e_2)_{x_1} &= (A_2)_{x_1} e_2 + A_2 (e_2)_{x_1} = (A_2)_{x_1} e_2 + A_2\ \frac{(A_1)_{x_2}}{A_2} e_1 \\ &= (A_2)_{x_1} e_2 + (A_1)_{x_2} e_1, \end{aligned}$$

and see that $(A_1 e_1)_{x_2} = (A_2 e_2)_{x_2}$, so (3.4.1) is solvable. Hence we can solve (3.4.1) by integration. It then follows that (f, e_1, e_2, e_3) is a solution of (3.3.5) with initial data (y_0, v_1, v_2, v_3). But (3.4.1) implies that f_{x_1}, f_{x_2} are linearly independent, so f is a parametrized surface, e_3 is normal to f, and $\mathrm{I} = A_1^2 dx_1^2 + A_2^2 dx_2^2$. Recall that

$$\ell_{ij} = f_{x_i x_j} \cdot N = f_{x_i x_j} \cdot e_3 = -(e_3)_{x_i} \cdot e_j.$$

So $\ell_{11} = -(e_3)_{x_1} \cdot f_{x_1} = -(-r_1 e_1) \cdot A_1 e_1 = r_1$, $\ell_{12} = -(e_3)_{x_2} \cdot A_1 e_1 = 0$, and $\ell_{22} = -(e_3)_{x_2} \cdot f_{x_2} = -(-r_2 e_2) \cdot A_2 e_2 = r_2 A_2$. Thus $\mathrm{II} = A_1 r_1 dx_1^2 + A_2 r_2 dx_2^2$. □

Corollary 3.4.2. *Suppose $f, g : \mathcal{O} \to \mathbb{R}^3$ are two surfaces parametrized by line of curvature coordinates, and f, g have the same first and second fundamental forms*

$$\mathrm{I} = A_1^2\ dx_1^2 + A_2^2\ dx_2^2, \quad \mathrm{II} = r_1 A_1\ dx_1^2 + r_2 A_2\ dx_2^2.$$

Then there exists a rigid motion ϕ of $\mathbb{R}^3$ such that $g = \phi \circ f$.

Proof. Let

$$e_1 = \frac{f_{x_1}}{A_1}, \quad e_2 = \frac{f_{x_2}}{A_2}, \quad e_3 = \frac{f_{x_1} \times f_{x_2}}{||f_{x_1} \times f_{x_2}||},$$
$$\xi_1 = \frac{g_{x_1}}{A_1}, \quad \xi_2 = \frac{g_{x_2}}{A_2}, \quad \xi_3 = \frac{g_{x_1} \times g_{x_2}}{||f_{x_1} \times g_{x_2}||}.$$

Fix $p_0 \in \mathcal{O}$, and let $\phi(x) = Tx + b$ be the rigid motion such that $\phi(f(p_0)) = g(p_0)$ and $T(e_i(p_0)) = \xi_i(p_0)$ for all $1 \leq i \leq 3$. Then

(1) $\phi \circ f$ have the same I, II as f, so $\phi \circ f$ and g have the same I, II,
(2) the o.n. moving frame for $\phi \circ f$ is (Te_1, Te_2, Te_3).

Thus both $(\phi \circ f, Te_1, Te_2, Te_3)$ and (g, ξ_1, ξ_2, ξ_3) are solutions of (3.3.5) with the same initial condition $(g(p_0), \xi_1(p_0), \xi_2(p_0), \xi_3(p_0))$. But Frobenius Theorem says that there is a unique solution for the initial value problem, hence

$$(\phi \circ f, T\xi_1, T\xi_2, T\xi_3) = (g, \xi_1, \xi_2, \xi_3).$$

In particular, this proves that $\phi \circ f = g$. □

3.5. **Gauss Theorem in line of curvature cooridnates.**

We know that the Gaussian curvature K depends on both I and II. The Gauss Theorem says that in fact K can be computed from I alone. We will first prove this when the surface is parametrized by line of curvatures.

Theorem 3.5.1. Gauss Theorem in line of curvature coordinates
Suppose $f : \mathcal{O} \to \mathbb{R}^3$ is a surface parametrized by line of curvatures, and

$$\mathrm{I} = A_1^2 dx_1^2 + A_2^2 dx_2^2, \mathrm{II} = r_1 A_1 dx_1^2 + r_2 A_2 dx_2^2.$$

Then

$$K = -\frac{\left(\frac{(A_1)_{x_2}}{A_2}\right)_{x_2} + \left(\frac{(A_2)_{x_1}}{A_1}\right)_{x_1}}{A_1 A_2},$$

so K can be computed from I *alone.*

Proof. Recall that

$$K = \frac{\det(\ell_{ij})}{\det(g_{ij})} = \frac{r_1 A_1 r_2 A_2}{A_1^2 A_2^2} = \frac{r_1 r_2}{A_1 A_2}.$$

But the first equation in the Gauss-Codazzi equation is

$$\left(\frac{(A_1)_{x_2}}{A_2}\right)_{x_2} + \left(\frac{(A_2)_{x_1}}{A_1}\right)_{x_1} = -r_1 r_2.$$

So

$$K = -\frac{\left(\frac{(A_1)_{x_2}}{A_2}\right)_{x_2} + \left(\frac{(A_2)_{x_1}}{A_1}\right)_{x_1}}{A_1 A_2}.$$

□

3.6. **Gauss-Codazzi equation in local coordinates.**

We will derive the Gauss-Codazzi equations for arbitrary parametrized surface $f : \mathcal{O} \to U \subset \mathbb{R}^3$.

Our experience in curve theory tells us that we should find a moving frame on the surface and then differentiate the moving frame to get relations among the invariants. We will use moving frames $F = (v_1, v_2, v_3)$ on the surface to derive the relations among local invariants, where $v_1 = f_{x_1}$, $v_2 = f_{x_2}$, and $v_3 = N$ the unit normal. Express the x and y derivatives of the local frame v_i in terms of v_1, v_2, v_3, then their coefficients can be written in terms of the two fundamental forms. Since $(v_i)_{xy} = (v_i)_{yx}$, we obtain a PDE relation for I and II. This is the Gauss-Codazzi equation of the surface. Conversely, given two symmetric bilinear forms g, b on an open subset $\mathcal{O}$ of $\mathbb{R}^2$ such that g is positive definite and g, b satisfies the Gauss-Codazzi equation, then by the Frobenius Theorem there exists a surface in $\mathbb{R}^3$ unique up to rigid motion having g, b as the first and second fundamental forms respectively.

We use the frame (f_{x_1}, f_{x_2}, N), where

$$N = \frac{f_{x_1} \times f_{x_2}}{||f_{x_1} \times f_{x_2}||}$$

is the unit normal vector field. Since f_{x_1}, f_{x_2}, N form a basis of $\mathbb{R}^3$, the partial derivatives of f_{x_i} and N can be written as linear combinations of f_{x_1}, f_{x_2} and N. So we have

$$\begin{cases} (f_{x_1}, f_{x_2}, N)_{x_1} = (f_{x_1}, f_{x_2}, N)P, \\ (f_{x_1}, f_{x_2}, N)_{x_2} = (f_{x_1}, f_{x_2}, N)Q, \end{cases} \tag{3.6.1}$$

where $P = (p_{ij}), Q = (q_{ij})$ are $gl(3)$-valued maps. This means that

$$\begin{cases} f_{x_1x_1} = p_{11}f_{x_1} + p_{21}f_{x_2} + p_{31}N, \\ f_{x_2x_1} = p_{12}f_{x_1} + p_{22}f_{x_2} + p_{32}N, \\ N_{x_1} = p_{13}f_{x_1} + p_{23}f_{x_3} + p_{33}N, \end{cases} \quad \begin{cases} f_{x_1x_2} = q_{11}f_{x_1} + q_{21}f_{x_2} + q_{31}N, \\ f_{x_2x_2} = q_{12}f_{x_1} + q_{22}f_{x_2} + q_{32}N, \\ N_{x_2} = q_{13}f_{x_1} + q_{23}f_{x_3} + q_{33}N. \end{cases}$$

Recall that the fundamental forms are given by

$$g_{ij} = f_{x_i} \cdot f_{x_j}, \quad \ell_{ij} = -f_{x_i} \cdot N_{x_j} = f_{x_ix_j} \cdot N.$$

We want to express P and Q in terms of g_{ij} and h_{ij}. To do this, we need the following Propositions.

Proposition 3.6.1. *Let V be a vector space with an inner product $(\ ,\ )$, $v_1, \cdots, v_n$ a basis of V, and $g_{ij} = (v_i, v_j)$. Let $\xi \in V$, $\xi_i = (\xi, v_i)$, and $\xi = \sum_{i=1}^n x_i v_i$. Then*

$$\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = G^{-1} \begin{pmatrix} \xi_1 \\ \cdot \\ \cdot \\ \xi_n \end{pmatrix},$$

where $G = (g_{ij})$.

Proof. Note that

$$\xi_i = (\xi, v_i) = (\sum_{j=1}^n x_j v_j, v_i) = \sum_{j=1}^n x_j (v_j, v_i) = \sum_{j=1}^n g_{ji} x_j.$$

So $(\xi_1, \cdots, \xi_n)^t = G(x_1, \cdots, x_n)^t$. □

Proposition 3.6.2. *The following statements are true:*

(1) *The $gl(3)$ valued functions $P = (p_{ij})$ and $Q = (q_{ij})$ in equation (3.6.1) can be written in terms of g_{ij}, ℓ_{ij}, and first partial derivatives of g_{ij}.*
(2) *The entries $\{p_{ij}, q_{ij} \mid 1 \le i, j \le 2\}$ can be computed from the first fundamental form.*

Proof. We claim that

$$f_{x_ix_j} \cdot f_{x_k}, \quad f_{x_ix_j} \cdot N, \quad N_{x_i} \cdot f_{x_j}, \quad N_{x_i} \cdot N,$$

can be expressed in terms of g_{ij}, ℓ_{ij} and first partial derivatives of g_{ij}. Then the Proposition follows from Proposition 3.6.1. To prove the claim, we proceed as follows:

$$\begin{cases} f_{x_ix_i} \cdot f_{x_i} = \frac{1}{2}(g_{ii})_{x_i}, & \\ f_{x_ix_j} \cdot f_{x_i} = \frac{1}{2}(g_{ii})_{x_j}, & \text{if } \ i \neq j, \\ f_{x_ix_i} \cdot f_{x_j} = (f_{x_i} \cdot f_{x_j})_{x_i} - f_{x_i} \cdot f_{x_jx_i} = (g_{ij})_{x_i} - \frac{1}{2}(g_{ii})_{x_j}, & \text{if } i \neq j \end{cases}$$

$$\begin{cases} f_{x_ix_j} \cdot N = \ell_{ij}, \\ N_{x_i} \cdot f_{x_j} = -\ell_{ij}, \\ N_{x_i} \cdot N = 0. \end{cases}$$

Let

$$G = \begin{pmatrix} g_{11} & g_{12} & 0 \\ g_{12} & g_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

By Proposition 3.6.1, we have

$$\text{(3.6.2)} \qquad \begin{cases} P = \begin{pmatrix} g^{11} & g^{12} & 0 \\ g^{12} & g^{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(g_{11})_{x_1} & \frac{1}{2}(g_{11})_{x_2} & -\ell_{11} \\ (g_{12})_{x_1} - \frac{1}{2}(g_{11})_{x_2} & \frac{1}{2}(g_{22})_{x_1} & -\ell_{12} \\ \ell_{11} & \ell_{12} & 0 \end{pmatrix} \\ \quad = G^{-1}A_1, \\ Q = \begin{pmatrix} g^{11} & g^{12} & 0 \\ g^{12} & g^{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(g_{11})_{x_2} & (g_{12})_{x_2} - \frac{1}{2}(g_{22})_{x_1} & -\ell_{12} \\ \frac{1}{2}(g_{22})_{x_1} & \frac{1}{2}(g_{22})_{x_2} & -\ell_{22} \\ \ell_{12} & \ell_{22} & 0 \end{pmatrix} \\ \quad = G^{-1}A_2. \end{cases}$$

This proves the Proposition. □

Formula (**??**) gives explicit formulas for entries of P and Q in terms of g_{ij} and ℓ_{ij}. Moreover, they are related to the Christofell symbols Γ^i_{jk} arise in the geodesic equation (**??**) in Theorem **??**. Recall that

$$\Gamma^k_{ij} = \frac{1}{2} g^{km}[ij, m],$$

where (g^{ij}) is the inverse matrix of (g_{ij}), $[ij, k] = g_{ki,j} + g_{jk,i} - g_{ij,k}$, and $g_{ij,k} = \frac{\partial g_{ij}}{\partial x_k}$.

Theorem 3.6.3. *For $1 \le i, j \le 2$, we have*

$$\text{(3.6.3)} \qquad p_{ji} = \Gamma^j_{i1}, \qquad q_{ji} = \Gamma^j_{i2}$$

Proof. Note that (**??**) implies

$$(3.6.4)\qquad \begin{cases} p_{11} &= \frac{1}{2}\ g^{11}g_{11,1} + g^{12}(g_{12,1} - \frac{1}{2}g_{11,2}) = \Gamma^1_{11}, \\ p_{12} &= \frac{1}{2}g^{11}g_{11,2} + \frac{1}{2}g^{12}g_{22,1} = \Gamma^1_{21}, \\ p_{21} &= \frac{1}{2}g^{12}g_{11,1} + g^{22}(g_{12,1} - \frac{1}{2}g_{11,2}) = \Gamma^2_{11}, \\ p_{22} &= \frac{1}{2}g^{12}g_{11,2} + \frac{1}{2}g^{22}g_{22,1} = \Gamma^2_{21}, \\ q_{11} &= \frac{1}{2}g^{11}g_{11,2} + \frac{1}{2}g^{12}g_{22,1} = \Gamma^1_{12}, \\ q_{12} &= g^{11}(g_{12,2} - \frac{1}{2}g_{22,1}) + \frac{1}{2}g^{12}g_{22,2} = \Gamma^1_{22}, \\ q_{21} &= \frac{1}{2}g^{12}g_{11,2} + \frac{1}{2}g^{22}g_{22,1} = \Gamma^2_{12}, \\ q_{22} &= g^{12}(g_{12,2} - \frac{1}{2}g_{22,1}) + \frac{1}{2}g^{22}g_{22,2} = \Gamma^2_{22}, \end{cases}$$

□

Note that

$$q_{11} = p_{12}, \quad q_{21} = p_{22}.$$

Theorem 3.6.4. The Fundamental Theorem of surfaces in $\mathbb{R}^3$.

Suppose $f : \mathcal{O} \to \mathbb{R}^3$ is a parametrized surface, and g_{ij}, ℓ_{ij} are the coefficients of I, II. *Let P, Q be the smooth $gl(3)$-valued maps defined in terms of g_{ij} and ℓ_{ij} by* (**??**). *Then P, Q satisfy*

$$(3.6.5)\qquad P_{x_2} - Q_{x_1} = [P, Q].$$

Conversely, let $\mathcal{O}$ be an open subset of $\mathbb{R}^2$, $(g_{ij}), (\ell_{ij}) : \mathcal{O} \to gl(2)$ smooth maps such that (g_{ij}) is positive definite and (ℓ_{ij}) is symmetric, and $P, Q : U \to gl(3)$ the maps defined by (3.6.2). *Suppose P, Q satisfies the compatibility equation* (3.6.5). *Let $(x_1^0, x_2^0) \in \mathcal{O}$, $p_0 \in \mathbb{R}^3$, and u_1, u_2, u_3 a basis of $\mathbb{R}^3$ so that $u_i \cdot u_j = g_{ij}(x_1^0, x_2^0)$ and $u_i \cdot u_3 = 0$ for $1 \leq i, j \leq 2$. Then there exists an open subset $\mathcal{O}_0 \subset \mathcal{O}$ of (x_1^0, x_2^0) and a unique immersion $f : \mathcal{O}_0 \to \mathbb{R}^3$ so that f maps $\mathcal{O}_0$ homeomorphically to $f(\mathcal{O}_0)$ such that*

1. *the first and second fundamental forms of the embedded surface $f(\mathcal{O}_0)$ are given by (g_{ij}) and (ℓ_{ij}) respectively,*
2. *$f(x_1^0, x_2^0) = p_0$, and $f_{x_i}(x_1^0, x_2^0) = u_i$ for $i = 1, 2$.*

Proof. We have proved the first half of the theorem, and it remains to prove the second half. We assume that P, Q satisfy the compatibility condition (3.6.5). So Frobenius Theorem **??** implies that the following system has a unique local solution

$$(3.6.6)\qquad \begin{cases} (v_1, v_2, v_3)_{x_1} = (v_1, v_2, v_3)P, \\ (v_1, v_2, v_3)_{x_2} = (v_1, v_2, v_3)Q, \\ (v_1, v_2, v_3)(x_1^0, x_2^0) = (u_1, u_2, u_3). \end{cases}$$

Next we want to solve

$$\begin{cases} f_{x_1} = v_1, \\ f_{x_2} = v_2, \\ f(x_1^0, x_2^0) = p_0. \end{cases}$$

The compatibility condition is $(v_1)_{x_2} = (v_2)_{x_1}$. But

$$(v_1)_{x_2} = \sum_{j=1}^{3} q_{j1} v_j, \quad (v_2)_{x_1} = \sum_{j=1}^{3} p_{j2} v_2.$$

It follows from (3.6.2) that the second column of P is equal to the first column of Q. So $(v_1)_{x_2} = (v_2)_{x_1}$, and hence there exists a unique f.

We will prove below that f is an immersion, v_3 is perpendicular to the surface f, $||v_3|| = 1$, $f_{x_i} \cdot f_{x_j} = g_{ij}$, and $(v_3)_{x_i} \cdot f_{x_j} = -\ell_{ij}$, i.e., f is a surface in $\mathbb{R}^3$ with

$$\mathrm{I} = \sum_{ij} g_{ij} dx_i dx_j, \quad \mathrm{II} = \sum_{ij} \ell_{ij} dx_i dx_j$$

as its first and second fundamental forms. The first step is to prove that the 3×3 matrix function $\Phi = (v_i \cdot v_j)$ is equal to the matrix G defined by (3.6.2). To do this, we compute the first derivative of Φ. Since v_1, v_2, v_3 satisfy (3.6.6), a direct computation gives

$$\begin{aligned}
(v_i \cdot v_j)_{x_1} &= (v_i)_{x_1} \cdot v_j + v_i \cdot (v_j)_{x_1} \\
&= \sum_k p_{ki} v_k \cdot v_j + p_{kj} v_k \cdot v_i = \sum_k p_{ki} g_{jk} + g_{ik} p_{kj} \\
&= (GP)_{ji} + (GP)_{ij} = (GP + (GP)^t)_{ij}.
\end{aligned}$$

Formula (3.6.2) implies that $GP = G(G^{-1}A_1) = A_1$ and $A_1 + A_1^t = G_{x_1}$. Hence $(GP)^t + GP = G_{x_1}$ and

$$\Phi_{x_1} = G_{x_1}.$$

A similar computation implies that

$$\Phi_{x_2} = G_{x_2}.$$

But the initial value $\Phi(x_1^0, x_2^0) = G(x_1^0, x_2^0)$. So $\Phi = G$. In other words, we have shown that

$$f_{x_i} \cdot f_{x_j} = g_{ij}, \quad f_{x_i} \cdot v_3 = 0.$$

Thus

(1) f_{x_1}, f_{x_2} are linearly independent, i.e., f is an immersion,
(2) v_3 is the unit normal field to the surface f,
(3) the first fundamental form of f is $\sum_{ij} g_{ij} dx_i dx_j$.

To compute the second fundamental form of f, we use (3.6.6) to compute

$$\begin{aligned}
-(v_3)_{x_1} \cdot v_j &= (g^{11}\ell_{11} + g^{12}\ell_{12}) v_1 \cdot v_j + (g^{12}\ell_{11} + g^{22}\ell_{12}) v_2 \cdot v_j \\
&= (g^{11}\ell_{11} + g^{12}\ell_{12}) g_{1j} + (g^{12}\ell_{11} + g^{22}\ell_{12}) g_{2j} \\
&= \ell_{11}(g^{11} g_{1j} + g^{12} g_{2j}) + \ell_{12}(g^{21} g_{1j} + g^{22} g_{2j}) \\
&= \ell_{11}\delta_{1j} + \ell_{12}\delta_{2j}.
\end{aligned}$$

So

$$-(v_3)_{x_1} \cdot v_1 = \ell_{11}, \quad -(v_3)_{x_1} \cdot v_2 = \ell_{12}.$$

Similar computations imply that

$$-(v_2)_{x_2} \cdot v_j = \ell_{2j}.$$

This proves that $\sum_{ij} \ell_{ij} dx_i dx_j$ is the second fundamental form of f. □

System (3.6.5) with P, Q defined by (3.6.2) is called the *Gauss-Codazzi equation for the surface* $f(\mathcal{O})$, which is a second order PDE with 9 equations for six functions g_{ij} and ℓ_{ij}. Equaqtion (3.6.5) is too complicated to memorize. It is more useful and simpler to just remember how to derive the Gauss-Codazzi equation.

It follows from (3.6.1), (3.6.2), and (3.6.3) that we have

$$f_{x_ix_1} = \sum_{j=1}^{2} p_{ji} f_{x_j} + \ell_{i1} N = \sum_{j=1}^{2} \Gamma^j_{i1} f_{x_j} + \ell_{i1} N,$$
$$f_{x_ix_2} = \sum_{j=1}^{2} q_{j2=i} f_{x_j} + \ell_{i2} N = \sum_{j=1}^{2} \Gamma^j_{i2} f_{x_j} + \ell_{i2} N,$$

where p_{ij} and q_{ij} are defined in (3.6.2).

So we have

$$f_{x_ix_j} = \Gamma^1_{ij} f_{x_1} + \Gamma^2_{ij} f_{x_2} + \ell_{ij} N, \tag{3.6.7}$$

Proposition 3.6.5. *Let $f : \mathcal{O} \to \mathbb{R}^3$ be a local coordinate system of an embedded surface M in $\mathbb{R}^3$, and $\alpha(t) = f(x_1(t), x_2(t))$. Then α satisfies the geodesic equation (**??**) if and only if $\alpha''(t)$ is normal to M at $\alpha(t)$ for all t.*

Proof. Differentiate α' to get $\alpha' = \sum_{i=1}^{2} f_{x_i} x_i'$. So

$$\begin{aligned} \alpha'' &= \sum_{i,j=1}^{2} f_{x_ix_j} x_i' x_j' + f_{x_i} x_i'' \\ &= \sum_{i,j,k=1}^{2} \Gamma^k_{ij} f_{x_k} x_i' x_j' + \ell_{ij} N + f_{x_i} x_i'' \\ &= \sum_{i,j=1}^{2} (\Gamma^k_{ij} x_i' x_j' + x_k'') f_{x_k} + \ell_{ij} N = 0 + \ell_{ij} N = \ell_{ij} N. \end{aligned}$$

□

3.7. The Gauss Theorem.

Equation (3.6.5) is the *Gauss-Codazzi equation* for M.

The Gaussian curvature K is defined to be the determinant of the shape operator $-dN$, which depends on both the first and second fundamental forms of the surface. In fact, by Proposition **??**

$$K = \frac{\ell_{11}\ell_{22} - \ell_{12}^2}{g_{11}g_{22} - g_{12}^2}.$$

We will show below that K can be computed in terms of g_{ij} alone. Equate the 12 entry of equation (3.6.5) to get

$$(p_{12})_{x_2} - (q_{12})_{x_1} = \sum_{j=1}^{3} p_{1j} q_{j2} - q_{1j} p_{j2}.$$

Recall that formula (3.6.4) gives $\{p_{ij}, q_{ij} \mid 1 \leq i, j \leq 2\}$ in terms of the first fundamental form I. We move terms involves p_{ij}, q_{ij} with $1 \leq i, j \leq 2$ to one side to get

$$(3.7.1) \qquad (p_{12})_{x_2} - (q_{12})_{x_1} - \sum_{j=1}^{2} p_{1j}q_{j2} - q_{1j}p_{j2} = p_{13}q_{32} - q_{13}p_{32}.$$

We claim that the right hand side of (3.7.1) is equal to

$$-g^{11}(\ell_{11}\ell_{22} - \ell_{12}^2) = -g^{11}(g_{11}g_{22} - g_{12}^2)K.$$

To prove this claim, use (3.6.2) to compute P, Q to get

$$p_{13} = -(g^{11}\ell_{11} + g^{12}\ell_{12}), \quad p_{32} = \ell_{12},$$
$$q_{13} = -(g^{11}\ell_{12} + g^{12}\ell_{22}), \quad q_{32} = \ell_{22}.$$

So we get

$$(3.7.2) \qquad (p_{12})_{x_2} - (q_{12})_{x_1} - \sum_{j=1}^{2} p_{1j}q_{j2} - q_{1j}p_{j3} = -g^{11}(g_{11}g_{22} - g_{12}^2)K.$$

Hence we have proved the claim and also obtained a formula of K purely in terms of g_{ij} and their derivatives:

$$K = -\frac{(p_{12})_{x_2} - (q_{12})_{x_1} - \sum_{j=1}^{2} p_{1j}q_{j2} - q_{1j}p_{j3}}{g^{11}(g_{11}g_{22} - g_{12}^2)}.$$

This proves

Theorem 3.7.1. Gauss Theorem. *The Gaussian curvature of a surface in* $\mathbb{R}^3$ *can be computed from the first fundamental form.*

The equation (3.7.2), obtained by equating the 12-entry of (3.6.5), is the *Gauss equation*.

A geometric quantity on an embedded surface M in $\mathbb{R}^n$ is called *intrinsic* if it only depends on the first fundamental form I. Otherwise, the property is called *extrinsic*, i.e., it depends on both I and II.

We have seen that the Gaussian curvature and geodesics are intrinsic quantities, and the mean curvature is extrinsic.

If $\phi : M_1 \to M_2$ is a diffeomorphism and $f(x_1, x_2)$ is a local coordinates on M_1, then $\phi \circ f(x_1, x_2)$ is a local coordinate system of M_2. The diffeomorphism ϕ is an *isometry* if the first fundamental forms for M_1, M_2 are the same written in terms of dx_1, dx_2. In particular,

(i) ϕ preserves anlges and arc length, i.e., the arc length of the curve $\phi(\alpha)$ is the same as the curve α and the angle between the curves $\phi(\alpha)$ and $\phi(\beta)$ is the same as the angle between α and β,

(ii) ϕ maps geodesics to geodesics.

Euclidean plane geometry studies the geometry of triangles. Note that triangles can be viewed as a triangle in the plane with each side being a geodesic. So a natural definition of a triangle on an embedded surface M is a piecewise smooth curve with three geodesic sides and any two sides meet at an angle lie in $(0, \pi)$. One important problem in geometric theory of M is to understand the geometry of triangles on M. For example, what is the sum of interior angles of a triangle on an embedded surface M? This will be answered by the Gauss-Bonnet Theorem.

Note that the first fundamental forms for the plane

$$f(x_1, x_2) = (x_1, x_2, 0)$$

and the cylinder

$$h(x_1, x_2) = (\cos x_1, \sin x_1, x_2)$$

have the same and is equal to $\mathrm{I} = dx_1^2 + dx_2^2$, and both surfaces have constant zero Gaussian curvature (cf. Examples **??** and **??**). We have also proved that geodesics are determined by I alone. So the geometry of triangles on the cylinder is the same as the geometry of triangles in the plane. For example, the sum of interior anlges of a triangle on the plane (and hence on the cylinder) must be π. In fact, let ϕ denote the map from $(0, 2\pi) \times \mathbb{R}$ to the cylinder minus the line $(1, 0, x_2)$ defined by

$$\phi(x_1, x_2, 0) = (\cos x_1, \sin x_1, x_2).$$

Then ϕ is an isometry.

3.8. **Gauss-Codazzi equation in orthogonal coordinates.**

If the local coordinates x_1, x_2 are orthogonal, i.e., $g_{12} = 0$, then the Gauss-Codazzi equation (3.6.5) becomes much simplier. Instead of putting $g_{12} = 0$ to (3.6.5), we derive the Gauss-Codazzi equation directly using an o.n. moving frame. We write

$$g_{11} = A_1^2, \quad g_{22} = A_2^2, \quad g_{12} = 0.$$

Let

$$e_1 = \frac{f_{x_1}}{A_1}, \quad e_2 = \frac{f_{x_2}}{A_2}, \quad e_3 = N.$$

Then (e_1, e_2, e_3) is an o.n. moving frame on M. Write

$$\begin{cases} (e_1, e_2, e_3)_{x_1} = (e_1, e_2, e_3)\tilde{P}, \\ (e_1, e_2, e_3)_{x_2} = (e_1, e_2, e_3)\tilde{Q}. \end{cases}$$

Since (e_1, e_2, e_3) is orthogonal, $\tilde{P}, \tilde{Q}$ are skew-symmetric. Moreover,

$$\tilde{p}_{ij} = (e_j)_{x_1} \cdot e_i, \quad \tilde{q}_{ij} = (e_j)_{x_2} \cdot e_i.$$

A direct computation gives

$$\begin{aligned} (e_1)_{x_1} \cdot e_2 &= \left(\frac{f_{x_1}}{A_1}\right)_{x_1} \cdot \frac{f_{x_2}}{A_2} = \frac{f_{x_1x_1} \cdot f_{x_2}}{A_1 A_2} \\ &= \frac{(f_{x_1} \cdot f_{x_2})_{x_1} - f_{x_1} \cdot f_{x_1x_2}}{A_1 A_2} \\ &= -\frac{(\frac{1}{2}A_1^2)_{x_2}}{A_1 A_2} = -\frac{(A_1)_{x_2}}{A_2}. \end{aligned}$$

Similar computation gives the coefficients $\tilde{p}_{ij}$ and $\tilde{q}_{ij}$:

$$\tilde{P} = \begin{pmatrix} 0 & \frac{(A_1)_{x_2}}{A_2} & -\frac{\ell_{11}}{A_1} \\ -\frac{(A_1)_{x_2}}{A_2} & 0 & -\frac{\ell_{12}}{A_2} \\ \frac{\ell_{11}}{A_1} & \frac{\ell_{12}}{A_2} & 0 \end{pmatrix}, \qquad \tilde{Q} = \begin{pmatrix} 0 & -\frac{(A_2)_{x_1}}{A_1} & -\frac{\ell_{12}}{A_1} \\ \frac{(A_2)_{x_1}}{A_1} & 0 & -\frac{\ell_{22}}{A_2} \\ \frac{\ell_{12}}{A_1} & \frac{\ell_{22}}{A_2} & 0 \end{pmatrix} \tag{3.8.1}$$

To get the Gauss-Codazzi equation of the surface parametrized by an orthogonal coordinates we only need to compute the 21-th, 31-th, and 32-the entry of the following equation

$$(\tilde{P})_{x_2} - (\tilde{Q})_{x_1} = [\tilde{P}, \tilde{Q}],$$

and we obtain

$$\begin{cases} -\left(\frac{(A_1)_{x_2}}{A_2}\right)_{x_2} - \left(\frac{(A_2)_{x_1}}{A_1}\right)_{x_1} = \frac{\ell_{11}\ell_{22}-\ell_{12}^2}{A_1A_2}, \\ \left(\frac{\ell_{11}}{A_1}\right)_{x_2} - \left(\frac{\ell_{12}}{A_1}\right)_{x_1} = \frac{\ell_{12}(A_2)_{x_1}}{A_1A_2} + \frac{\ell_{22}(A_1)_{x_2}}{A_2^2} \\ \left(\frac{\ell_{12}}{A_2}\right)_{x_2} - \left(\frac{\ell_{22}}{A_2}\right)_{x_1} = -\frac{\ell_{11}(A_2)_{x_1}}{A_1^2} - \frac{\ell_{12}(A_1)_{x_2}}{A_1A_2}. \end{cases} \tag{3.8.2}$$

The first equation of (3.8.2) is called the *Gauss equation*. Note that the Gaussian curvature is

$$K = \frac{\ell_{11}\ell_{22} - \ell_{12}^2}{(A_1A_2)^2}.$$

So we have

$$K = -\frac{\left(\frac{(A_1)_{x_2}}{A_2}\right)_{x_2} + \left(\frac{(A_2)_{x_1}}{A_1}\right)_{x_1}}{A_1A_2}. \tag{3.8.3}$$

We have seen that the Gauss-Codazzi equation becomes much simpler in orthogonal coordinates. Can we always find local orthogonal coordinates on a surface in $\mathbb{R}^3$? This question can be answered by the following theorem, which we state without a proof.

Theorem 3.8.1. *Suppose $f : \mathcal{O} \to \mathbb{R}^3$ be a surface, $x_0 \in \mathcal{O}$, and $Y_1, Y_2 : \mathcal{O} \to \mathbb{R}^3$ smooth maps so that $Y_1(x_0), Y_2(x_0)$ are linearly independent and tangent to $M = f(\mathcal{O})$ at $f(x_0)$. Then there exist open subset $\mathcal{O}_0$ of $\mathcal{O}$ containing x_0, open subset $\mathcal{O}_1$ of $\mathbb{R}^2$, and a diffeomorphism $h : \mathcal{O}_1 \to \mathcal{O}_0$ so that $(f \circ h)_{y_1}$ and $(f \circ h)_{y_2}$ are parallel to $Y_1 \circ h$ and $Y_2 \circ h$.*

The above theorem says that if we have two linearly independent vector fields Y_1, Y_2 on a surface, then we can find a local coordinate system $\phi(y_1, y_2)$ so that ϕ_{y_1}, ϕ_{y_2} are parallel to Y_1, Y_2 respectively.

Given an arbitrary local coordinate system $f(x_1, x_2)$ on M, we apply the Gram-Schmidt process to f_{x_1}, f_{x_2} to construct smooth o.n. vector fields e_1, e_2:

$$e_1 = \frac{f_{x_1}}{\sqrt{g_{11}}},$$
$$e_2 = \frac{\sqrt{g_{11}}(f_{x_2} - \frac{g_{12}}{g_{11}} f_{x_1})}{\sqrt{g_{11}g_{22} - g_{12}^2}},$$

By Theorem 3.8.1, there exists new local coordinate system $\tilde{f}(y_1, y_2)$ so that $\frac{\partial \tilde{f}}{\partial x_1}$ and $\frac{\partial \tilde{f}}{\partial x_2}$ are parallel to e_1 and e_2. So the first fundamental form written in this coordinate system has the form

$$\tilde{g}_{11} dy_1^2 + \tilde{g}_{22} dy_2^2.$$

However, in general we can not find coordinate system $\hat{f}(y_1, y_2)$ so that e_1 and e_2 are coordinate vector fields $\frac{\partial \hat{f}}{\partial y_1}$ and $\frac{\partial \hat{f}}{\partial y_2}$ because if we can then the first fundamental form of the surface is $\mathrm{I} = dy_1^2 + dy_2^2$, which implies that the Gaussian curvature of the surface must be zero.