DERIVATIVES AND CURVATURE IN A NUTSHELL

JIAYIN PAN

ABSTRACT. Additional lecture notes for Surface Theory.

CONTENTS

1. FIRST FUNDAMENTAL FORM AS MEASUREMENT WITHIN THE SURFACE

Let S be a surface in $\mathbb{R}^3$. Recall that for each $p \in S$, we have the tangent space T_pS and the normal space N_pS, both of which are linear subspaces in $\mathbb{R}^3$.

Definition 1.1. We define the *first fundamental form* $\mathbf{I} : T_pS \times T_pS \to \mathbb{R}$ at each point $p \in S$ by

$$\mathbf{I}(X,Y) = X \cdot Y,$$

where $\cdot$ is the dot product in $\mathbb{R}^3$.

For each $p \in S$, the first fundamental form is an inner product on T_pS, that is, a bilinear, symmetric, and positive definite map. With it, we can naturally talk about, for example, the length of tangent vectors and the angle between any pair. As the point p moves in the surface, the first fundamental form varies.

In a Nutshell. *The first fundamental form is an intrinsic measurement within the surface.*

Let $q(u^1,u^2)$ be a parametric equation for S, then $\{\partial_1 q, \partial_2 q\}$ forms a basis for T_pS for any $p \in S$. We can write any tangent vector $X \in T_pS$ under this basis as

$$X = X^i \partial_i q.$$

Then for all $X, Y \in T_pS$, we have formula

$$\mathbf{I}(X,Y) = \mathbf{I}(X^i\partial_i q, Y^j \partial_j q) = X^i Y^j g_{ij},$$

where $g_{ij} = \mathbf{I}(\partial_i q, \partial_j q)$ forms a 2×2 symmetric and positive definite matrix. We call this matrix (g_{ij}) the matrix representation of $\mathbf{I}$ under the parametric equation $q(u^1,u^2)$, usually written as $[\mathbf{I}]$. In matrix form, we obtain

$$\mathbf{I}(X,Y) = \begin{pmatrix} X^1 & X^2 \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} Y^1 \\ Y^2 \end{pmatrix}.$$

Remark 1.2. We remark that by definition $\mathbf{I}$ is clearly independent of the parametric equations, but the matrix representation $[\mathbf{I}]$ depends on $q(u^1, u^2)$.

Definition 1.3. Let S_1 and S_2 be two surfaces in $\mathbb{R}^3$ with first fundamental form $\mathbf{I}_1$ and $\mathbf{I}_2$, respectively. Let $F: S_1 \to S_2$ be a smooth map. We say that F is a (local) *isometry*, if F preserves the first fundamental form, that is,

$$\mathbf{I}_1(X, Y) = \mathbf{I}_2(DF_p(X), DF_p(Y))$$

for all $p \in S_1$ and all $X, Y \in T_p S_1$.

Proposition 1.4. *Let $q(u^1, u^2)$ be a parametric equation for S_1 and let $F: S_1 \to S_2$ be a smooth map. We write $[\mathbf{I}_1]$ as the first fundamental form of S_1 under $q(u^1, u^2)$, and $[\mathbf{I}_2]$ as the first fundamental form of S_2 under $F \circ q(u^1, u^2)$. Then F is an isometry if and only if $[\mathbf{I}_1] = [\mathbf{I}_2]$.*

Proof. Recall that we always have

$$DF(\partial_i q) = \partial_i (F \circ q).$$

Suppose that F is an isometry. Then

$$\mathbf{I}_1\left(\partial_i q, \partial_j q\right) = \mathbf{I}_2\left(DF(\partial_i q), DF(\partial_j q)\right) = \mathbf{I}_2(\partial_i (F \circ q), \partial_j (F \circ q)).$$

and thus $[\mathbf{I}_1] = [\mathbf{I}_2]$.

Conversely, suppose that $[\mathbf{I}_1] = [\mathbf{I}_2]$. For any vectors $X, Y \in T_p S$, we write

$$X = X^i \partial_i q, \quad Y = Y^i \partial_i q.$$

Then under DF, we obtain

$$DF(X) = X^i \partial_i (F \circ q), \quad DF(Y) = Y^i \partial_i (F \circ q).$$

Hence

$$\mathbf{I}(X, Y) = X^i Y^j \mathbf{I}_1(\partial_i q, \partial_j q) = X^i Y^j \mathbf{I}_2(\partial_i (F \circ q), \partial_j (F \circ q)) = \mathbf{I}_2(DF(X), DF(Y)).$$

□

To close this section, we prove a formula writing any vector Z under the basis $\{\partial_1 q, \partial_2 q, N\}$, where N is a unit normal vector field on S, by using the first fundamental form. This will be useful in later sections.

Lemma 1.5. *Let V be a linear space with a basis $\{e_1, ..., e_n\}$. Let $\mathbf{I}$ be an inner product on V. We denote $g_{ij} = \mathbf{I}(e_i, e_j)$. Then for any vector $X \in V$, we have*

$$X = g^{ij} \mathbf{I}(X, e_j) e_i,$$

where (g^{ij}) is the inverse matrix of (g_{ij}).

Proof. We write $X = X^i e_i$. By direct calculation,

$$g^{ij} \mathbf{I}(X, e_j) = g^{ij} \mathbf{I}(X^k e_k, e_j) = X^k g^{ij} g_{jk} = X^k \delta^i_k = X^i.$$

The result follows. □

Proposition 1.6. *Let S be a surface in $\mathbb{R}^3$ with a parametric equation $q(u^1,u^2)$. Then for all $p\in S$ and all $X\in T_pS$, it holds that*

$$X = g^{ij}\mathbf{I}(X,\partial_j q)\partial_i q$$

In general, for any vector $Z\in\mathbb{R}^3$, we have tangential and normal decomposition

$$Z = Z^{\mathrm{T}} + Z^{\perp} = g^{ij}(Z\cdot\partial_j q)\partial_i q + (Z\cdot N)N,$$

where N is the unit normal vector field over S.

Proof. Using Lemma 1.5 and $\{\partial_1 q,\partial_2 q\}$ as a basis of T_pS, the first part of the Proposition follows. For any vector Z in $\mathbb{R}^3$, we decompose it as $Z = Z^{\mathrm{T}}+Z^{\perp}$, where $Z^{\mathrm{T}}\in T_pS$ is the tangential part of Z and $Z^{\perp}\in N_pS$ is the normal part, respectively. Noting that

$$Z\cdot\partial_i q = Z^{\mathrm{T}}\cdot\partial_i q = \mathbf{I}(Z^{\mathrm{T}},\partial_i q),\quad Z^{\perp} = (Z\cdot N)N,$$

we conclude

$$Z = Z^{\mathrm{T}} + Z^{\perp} = g^{ij}(Z\cdot\partial_j q)\partial_i q + (Z\cdot N)N.$$

□

2. Second fundamental form as normal part of directional derivative

Let S be a surface in $\mathbb{R}^3$ and let $p\in S$. Let $X\in T_pS$ and Y be a vector field of S around p. The main goal of this section and Section 4 is to decompose the directional derivative D_XY into tangential and normal parts

$$D_XY = (D_XY)^{\mathrm{T}} + (D_XY)^{\perp}.$$

In this section, we address the normal part $(D_XY)^{\perp}$.

We first recall directional derivatives from vector calculus. Given a function $f: U\to\mathbb{R}$, where U is an open subset of $\mathbb{R}^3$, and a vector $X\in\mathbb{R}^3$, the directional derivative of f along X at a point $p\in\mathbb{R}^3$ is defined as

$$D_Xf = \frac{d(f\circ c)}{dt}\Big|_{t=0},$$

where c is a curve in $\mathbb{R}^3$ with $c(0)=p$ and $c'(0)=X$. (Usually, one may use $c(t)=p+tX$, a straight line, to define the directional derivative.) The definition D_Xf is independent of the choice of c. In fact, under a standard Cartesian coordinate of $\mathbb{R}^3$, we write

$$f = f(x^1,x^2,x^3),\quad X=(X^1,X^2,X^3),\quad c(t)=(x^1(t),x^2(t),x^3(t));$$

then $\frac{dx^i}{dt}(0)=X^i$ and by chain rule

$$\frac{d}{dt}\Big|_{t=0}(f\circ c) = \frac{\partial f}{\partial x^i}(p)\frac{dx^i}{dt}(0) = \frac{\partial f}{\partial x^i}(p)X^i.$$

Once the curve c is chosen, D_Xf only depends on the value of f on c. For a vector field Y in $U\subseteq\mathbb{R}^3$, we write

$$Y(x) = (Y^1(x),Y^2(x),Y^3(x)).$$

Then the directional derivative of Y along X at p is defined by taking directional derivative of each component Y^i, that is,

$$D_XY = (D_XY^1,D_XY^2,D_XY^3) = \frac{d(Y\circ c)}{dt}\Big|_{t=0},$$

where c is any curve in $\mathbb{R}^3$ with $c(0)=p$ and $c'(0)=X$.

Now we move on to surface theory. Let S be a surface and let $p \in S$. Let $X \in T_pS$ and Y be a vector field of S around p. Under a parametric equation $q(u^1, u^2)$, we write

$$X = X^i \partial_i q, \quad Y(x) = Y^i(x)\partial_i q.$$

We draw any curve $\gamma : I \to S$ such that $\gamma(0) = p$ and $\gamma'(0) = X$. We can write γ as $\gamma(t) = q(u^1(t), u^2(t))$. By construction, for all i we have

$$\frac{du^i}{dt}(0) = X^i.$$

Restricting the vector field Y to γ gives a vector field along γ, that is, $Y(\gamma(t))$. We will calculate the normal part of $D_X Y$.

$$\begin{aligned}\frac{d(Y\circ\gamma)}{dt} &= \frac{d}{dt}\left((Y^i\circ\gamma)\frac{\partial q}{\partial u^i}\right)\\ &= \left(\frac{d}{dt}(Y^i\circ\gamma)\right)\frac{\partial q}{\partial u^i} + (Y^i\circ\gamma)\frac{d}{dt}\left(\frac{\partial q}{\partial u^i}\right)\end{aligned}$$

The first term above is in the tangential direction so it won't contribute to the normal part. By chain rule, the second term is

$$(Y^i\circ\gamma)\frac{d}{dt}\left(\frac{\partial q}{\partial u^i}\right) = (Y^i\circ\gamma)\frac{\partial^2 q}{\partial u^j \partial u^i}\frac{du^j}{dt}.$$

To find $(D_X Y)^\perp$, it suffices to evaluate $N\cdot D_X Y$.

$$N\cdot D_X Y = N\cdot\frac{d(Y\circ\gamma)}{dt}(0) = Y^i(\gamma(0))\left(N\cdot\frac{\partial^2 q}{\partial u^j\partial u^i}\Big|_p\right)\frac{du^j}{dt}(0) = Y^i(p)L_{ij}X^j,$$

where (L_{ij}) is the 2×2 symmetric matrix with components $N\cdot\partial_{ij}q$ evaluated at p. Written in the matrix form, we have the formula

$$N\cdot D_X Y = \begin{pmatrix} Y^1(p) & Y^2(p)\end{pmatrix}\begin{pmatrix} L_{11} & L_{12}\\ L_{21} & L_{22}\end{pmatrix}\begin{pmatrix} X^1\\ X^2\end{pmatrix}.$$

Definition 2.1. We define the *second fundamental form* $\mathbf{II} : T_pS\times T_pS \to \mathbb{R}$ at each point $p\in S$ by

$$\mathbf{II}(X,Y) = N\cdot D_X\widetilde{Y},$$

where $\widetilde{Y}$ is a smooth vector field of S around p that extends Y (i.e., $\widetilde{Y}(p) = Y$).

Remarks 2.2. (1) Under a parametric equation $q(u^1,u^2)$, the above calculation yields

$$N\cdot D_X\widetilde{Y} = X^i\widetilde{Y}^j(p)L_{ij} = X^iY^jL_{ij}.$$

Hence $N\cdot D_X\widetilde{Y}$ does not depend on the choice of $\widetilde{Y}$. This allows us to define $\mathbf{II}$ with domain $T_pS\times T_pS$.
(2) It follows from $\mathbf{II}(X,Y) = X^iY^jL_{ij}$ and $L_{ij} = L_{ji}$ that $\mathbf{II}$ is bilinear and symmetric.
(3) It follows from $\mathbf{II}(X,Y) = N\cdot D_X\widetilde{Y}$ that $\mathbf{II}$ is independent of the parametric equations of S.
(4) To evaluate $\mathbf{II}(X,X)$, we can choose a curve γ in S with $\gamma(0) = p$, $\gamma'(0) = X$, and a vector field Y of S around p such that it restricts to the velocity vector field γ' on γ. Then $Y(p) = \gamma'(0) = X$ and

$$\mathbf{II}(X,X) = N\cdot\frac{d(Y\circ\gamma)}{dt}(0) = N\cdot\frac{d\gamma'}{dt}(0) = N\cdot\gamma''(0),$$

which is the normal component of γ''. Writing $X = \gamma'(0) = X^i \partial_i q$, we also have

$$N \cdot \gamma''(0) = \mathbf{II}(X, X) = X^i X^j L_{ij}.$$

This shows that the normal part of the acceleration $\gamma''(0)^\perp = (N \cdot \gamma''(0))N$ only depends on the velocity $\gamma'(0) = X$.

3. CURVATURE AS THE CHANGE OF NORMAL DIRECTION

Definition 3.1. Let S be a surface in $\mathbb{R}^3$ with a unit normal vector field N. We think of N as a map (called the *Gauss map*)

$$N : S \to \mathbb{S}^2,$$

where $\mathbb{S}^2$ is the unit sphere in $\mathbb{R}^3$ with center 0. The *shape operator* (or the *Weingarten map*) L at $p \in S$ is the negative of the differential of N, that is,

$$L = -DN_p : T_pS \to T_{N(p)}\mathbb{S}^2$$

(there is no standard agreement whether to use + or − in the definition of L.)

Because

$$T_{N(p)}\mathbb{S}^2 = \{v \in \mathbb{R}^3 | v \perp N(p)\} = T_pS,$$

we understand the shape operator at $p \in S$ as a linear map

$$L = -DN_p : T_pS \to T_pS$$

In a Nutshell. *Curvature at p measures the infinitesimal change of N at p. Mathematically, it is some quantity extracted from the linear map DN_p.*

The differential DN is closely related to directional derivative. In fact, we have

Lemma 3.2. *Let $X \in T_pS$, then $DN(X) = D_X N$.*

Proof. This follows from the definition of directional derivative and that of differential. In fact, let $\gamma : I \to S$ be a curve on S with $\gamma(0) = p$ and $\gamma'(0) = X(p)$, then

$$DN_p(X(p)) = \frac{d}{dt}(N \circ \gamma)(0) = D_{X(p)}N.$$

□

Proposition 3.3. $\mathbf{I}$, $\mathbf{II}$, *and $L = -DN$ are related by the formula*

$$\mathbf{I}(L(X), Y) = \mathbf{II}(X, Y).$$

Consequently, L is self-adjoint with respect to $\mathbf{I}$, *that is,*

$$\mathbf{I}(L(X), Y) = \mathbf{I}(X, L(Y)).$$

Proof. By direct calculation and Lemma 3.2,

$$\mathbf{I}(L(X), Y) = -DN(X) \cdot Y = -D_X N \cdot Y = -D_X(N \cdot Y) + N \cdot D_X Y = \mathbf{II}(X, Y).$$

Because $\mathbf{II}$ is symmetric, it follows that L is self-adjoint. □

Corollary 3.4. *Under a parametric equation $q(u^1, u^2)$, the matrix representations* $[\mathbf{I}]$, $[\mathbf{II}]$, *and* $[L]$ *are related by*

$$[L] = [\mathbf{I}]^{-1}[\mathbf{II}].$$

Proof. Here the matrix representation of L means the matrix $[L] = (L_i^j)$ such that

$$\begin{pmatrix} L(\partial_1 q) & L(\partial_2 q) \end{pmatrix} = \begin{pmatrix} \partial_1 q & \partial_2 q \end{pmatrix} \begin{pmatrix} L_1^1 & L_2^1 \\ L_1^2 & L_2^2 \end{pmatrix}.$$

Equivalently, we can write $L(\partial_i q) = L_i^j \partial_j q$. By Lemma 1.5 and Proposition 3.3,

$$L_i^j = g^{jk}\mathbf{I}(L(\partial_i q), \partial_k q) = g^{jk}\mathbf{II}(\partial_i q, \partial_k q) = g^{jk}L_{ki}.$$

In the matrix form, this reads $[L] = [\mathbf{I}]^{-1}[\mathbf{II}]$. □

Remark 3.5. Though L is self-adjoint, its matrix representation $[L]$ may not be symmetric because in general $\{\partial_i q\}$ is not an orthonormal basis.

One measurement of a linear map (that is invariant under change of basis) is the eigenvalues. For a self-adjoint one, eigenvalues actually describe all the invariance. We recall the result below from linear algebra.

Theorem 3.6. *Let V be an n-dimensional real linear space with an inner product $\mathbf{I}$. Let $L : V \to V$ be a self-adjoint linear map. Then*
(1) All eigenvalues of L are real;
(2) L has n many linearly independent eigenvectors $\{E_1, ..., E_n\}$ that form an orthonormal basis w.r.t. $\mathbf{I}$.

Applying Lemma 3.6 to the shape operator L, we derive

Proposition 3.7. *At every point $p \in S$, L has real eigenvalues κ_1, κ_2 and corresponding eigenvectors E_1, E_2 which form an orthonormal basis of T_pS w.r.t. $\mathbf{I}$.*

Definitions 3.8. We call the above κ_i and E_i the *principal curvature* and the *principal direction* of S at p, respectively.

We define the *Gaussian curvature* and the *mean curvature* at p, respectively, as

$$K = \det L = \kappa_1\kappa_2, \quad H = \frac{1}{2}\mathrm{tr} L = \frac{1}{2}(\kappa_1 + \kappa_2).$$

Remarks 3.9. (1) All curvatures above are independent of the parametric equations.
(2) For a different choice of the unit normal vector N, $L = -DN$ differs by a sign, then all κ_i and H change their signs, while K does not change.
(3) To calculate the curvature, one can always use $[L] = [\mathbf{I}]^{-1}[\mathbf{II}]$ to calculate the eigenvalues of $[L]$. Also, Gaussian curvature can be calculated by

$$K = \det[L] = \frac{\det[\mathbf{II}]}{\det[\mathbf{I}]}.$$

(4) For very special surfaces, there are other ways to calculate the curvature without going through parametric equations. For example, in Petersen's notes page 112, it was shown that $\mathbf{II} = \frac{1}{R}\mathbf{I}$ with an inward unit normal N on the sphere of radius R. Then for all $X, Y \in T_pS$ and all $p \in S$,

$$\mathbf{I}(L(X), Y) = \mathbf{II}(X, Y) = \frac{1}{R}\mathbf{I}(X, Y).$$

Hence $L(X) = \frac{1}{R}X$ for all $X \in T_pS$. That is, the sphere of radius R has principal curvature $\kappa_1 = \kappa_2 = 1/R$ and Gaussian curvature $1/R^2$.

Proposition 3.10. *The principal curvatures at $p \in S$ are the minimum and maximum of*

$$\{\mathbf{II}(X,X) | X \in T_pS,\ \mathbf{I}(X,X) = 1\}.$$

Proof. Let E_1, E_2 be the principal directions at p. We write any unit vector $E \in T_pS$ as

$$E = \cos\theta E_1 + \sin\theta E_2.$$

Then by linearity of $\mathbf{II}$,

$$\mathbf{II}(E,E) = \cos^2\theta \mathbf{II}(E_1,E_1) + 2\cos\theta\sin\theta \mathbf{II}(E_1,E_2) + \sin^2\theta \mathbf{II}(E_2,E_2).$$

We calculate $\mathbf{II}(E_i,E_j)$ by

$$\mathbf{II}(E_i,E_j) = \mathbf{I}(L(E_i),E_j) = \mathbf{I}(\kappa_i E_i, E_j) = \kappa_i \delta_{ij}.$$

Hence

$$\mathbf{II}(E,E) = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta =: f(\theta).$$

One can easily check that $f(\theta)$ has minimum and maximum as κ_1 and κ_2. □

4. Covariant derivative as tangential part of directional derivative

In Section 3, we used both fundamental forms $\mathbf{I}$ and $\mathbf{II}$ to calculate the curvature, which measures the change of N. Recall that the second fundamental form describes the accelerations in N direction while the first one has nothing to do with N. So naturally, one expects that $\mathbf{II}$ is required to understand the curvature. In *Gauss's Theorema Egregium*, he observed that the Gaussian curvature K can be computed by only knowing $\mathbf{I}$. In other words, Gaussian curvature is intrinsic, depending only on measurements within the surface, not related to the normal direction. To understand this surprising result by Gauss, we will first study the tangential part of D_XY in this section.

Recall that from Section 2, at every point $p \in S$ the normal part $(D_XY)^\perp = \mathbf{II}(X,Y)N$ only depends on the value of X,Y at p, while a derivative should depend on the local information of Y near p. In contrast, we will see that the tangential part $(D_XY)^{\mathrm{T}}$ behaves like derivatives (see Proposition 4.10); and we will call it the covariant derivative.

In a Nutshell. *The covariant derivative is the intrinsic vector calculus within the surface.*

Definition 4.1. Let $X \in T_pS$ and let Y be a smooth vector field on S. The *covariant derivative* of Y along X is defined as the tangential part of D_XY, that is,

$$\nabla_X Y = (D_XY)^{\mathrm{T}}.$$

We use the same setup as in Section 2, where we calculated the normal part of D_XY. Let $X \in T_pS$ and let $\gamma(t)$ be a curve on S with $\gamma(0) = p$ and $\gamma'(0) = X$. Let Y be a vector field of S around p. Under a parametric equation $q(u^1,u^2)$, we write

$$X = X^i \partial_i q, \quad Y = Y^i(x)\partial_i q.$$

We will calculate the tangential part of D_XY.

$$\begin{aligned}\frac{d(Y\circ\gamma)}{dt}(0) &= \frac{d(Y^i\circ\gamma)}{dt}(0)\frac{\partial q}{\partial u^i}(p) + (Y^i\circ\gamma(0))\frac{d}{dt}\Big|_{t=0}\left(\frac{\partial q}{\partial u^i}\right)\\ &= D_XY^i(p)\frac{\partial q}{\partial u^i}(p) + Y^i(p)\frac{\partial^2 q}{\partial u^j \partial u^i}(p)\frac{du^j}{dt}(0)\\ &= (D_XY^i)\partial_i q + Y^i X^j \partial_{ij} q.\end{aligned}$$

(We omitted p in the last line above for brevity, but everything should be evaluated at the point p). The first term is in the tangential direction and is already written under the basis $\{\partial_i q\}$. We apply Proposition 1.6 to rewrite the tangential part of the second term

$$(Y^i X^j \partial_{ij} q)^{\mathrm{T}} = Y^i X^j g^{kl} \mathbf{I}((\partial_{ij} q)^{\mathrm{T}}, \partial_l q) \partial_k q.$$

Introducing Christoffel symbols

$$\Gamma_{ij,l} = \partial_{ij} q \cdot \partial_l q = \mathbf{I}((\partial_{ij} q)^{\mathrm{T}}, \partial_l q), \quad \Gamma_{ij}^k = g^{kl} \Gamma_{ij,l},$$

we have expression

$$(Y^i X^j \partial_{ij} q)^{\mathrm{T}} = Y^i X^j g^{kl} \Gamma_{ij,l} \partial_k q = Y^i X^j \Gamma_{ij}^k \partial_k q.$$

We remark that $\Gamma_{ij,k}$ and Γ_{ij}^k are symmetric in i and j. To sum up,

$$\nabla_X Y = (D_X Y)^{\mathrm{T}} = (D_X Y^i) \partial_i q + (Y^i X^j \partial_{ij} q)^{\mathrm{T}} = \left(D_X Y^k + X^i Y^j \Gamma_{ij}^k \right) \partial_k q.$$

Together with the normal part in Section 2, we have

Proposition 4.2. *Let $X \in T_p S$ and Y be a smooth vector field of S around p. Then*

$$\begin{aligned} D_X Y &= \nabla_X Y + \mathbf{II}(X, Y) N \\ &= \left(D_X Y^k + X^i Y^j \Gamma_{ij}^k \right) \partial_k q + (X^i Y^j L_{ij}) N. \end{aligned}$$

Remarks 4.3. (1) Fixing i and j, we use

$$X = \partial_i q, \quad Y = \partial_j q.$$

Then $X^i = 1$ and $Y^j = 1$ with all other components functions of X and Y as 0. Then $D_{\partial_i q} = \frac{\partial}{\partial u^i}$ (see Remark 4.5 below) and Proposition 4.2 yield the *Gauss equations*

$$\partial_{ij} q = D_{\partial_i q} \partial_j q = \Gamma_{ij}^k \partial_k q + L_{ij} N.$$

(2) Sometimes, Y is written along the curve γ as

$$Y = Y^i(t) \partial_i q.$$

Then it follows from the same calculation above that

$$\nabla_X Y = \left(\frac{dY^k}{dt} + X^i Y^j \Gamma_{ij}^k \right) \partial_k q$$

(3) Let $\gamma(t) = q(u^1(t), u^2(t))$ be a curve in S. We use $X = Y$ as the velocity of γ:

$$\gamma'(t) = \frac{du^i}{dt}(t) \partial_i q.$$

Then γ'' has decomposition

$$\begin{aligned} \gamma'' &= \nabla_{\gamma'} \gamma' + \mathbf{II}(\gamma', \gamma') N \\ &= \left(\frac{d}{dt} \left(\frac{du^k}{dt} \right) + \frac{du^i}{dt} \frac{du^j}{dt} \Gamma_{ij}^k \right) \partial_k q + \mathbf{II}(\gamma', \gamma') N \\ &= \left(\frac{d^2 u^k}{dt^2} + \frac{du^i}{dt} \frac{du^j}{dt} \Gamma_{ij}^k \right) \partial_k q + \left(\frac{du^i}{dt} \frac{du^j}{dt} L_{ij} \right) N. \end{aligned}$$

Definition 4.4. A curve γ on S is called a *geodesic*, if $\nabla_{\gamma'} \gamma' \equiv 0$.

Under a parametric equation $q(u^1,u^2)$, a curve $\gamma = q(u^1(t),u^2(t))$ is a geodesic if and only if γ obeys the geodesic equations $(k = 1,2)$:

$$\frac{d^2u^k}{dt^2} + \frac{du^i}{dt}\frac{du^j}{dt}\Gamma^k_{ij} = 0.$$

Remark 4.5. We give a few more explanations on $D_{\partial_i q} = \frac{\partial}{\partial u^i}$. To be more precisely, for any vector field Z (not necessarily tangent to S), we have

$$(D_{\partial_i q}Z)|_{q(u^1,u^2)} = \frac{\partial(Z\circ q)}{\partial u^i}(u^1,u^2).$$

We prove the above formula for $i = 1$. Let $p = q(u^1_0,u^2_0) \in S$. The curve $\gamma(t) = q(u^1_0 + t, u^2_0)$ satisfies $\gamma(0) = p$ and $\gamma'(0) = \partial_1 q$. Thus at p,

$$D_{\partial_i q}Z = \frac{d(Z\circ\gamma)}{dt}\Big|_{t=0} = \frac{d}{dt}\Big|_{t=0}(Z\circ q(u^1_0 + t,u^2_0)) = \frac{\partial(Z\circ q)}{\partial u^1}(u^1_0,u^2_0).$$

For the coordinate vector field $\partial_j q$, we usually abuse the notation to treat it as a vector field on S though it has domain as the parameter space U, not S. With this understanding, we can write

$$D_{\partial_i q}\partial_j q = \partial_i(\partial_j q) = \partial_{ij} q.$$

Sometimes, we write the unit normal vector field N as $N(u^1,u^2)$ with domain being the parameter space U (for example, when we calculate N as the re-normalization of $\partial_1 q\times\partial_2 q$). Then with the same understanding, we write

$$D_{\partial_i q}N = \partial_i N.$$

Using Lemma 3.2 and notations in Section 3, the left hand side is

$$D_{\partial_i q}N = DN(\partial_i q) = -L(\partial_i q) = -L^j_i\partial_j q.$$

This leads to the *Weingarten equations*

$$\partial_i N = -L^j_i\partial_j q.$$

We summarize the Gauss and Weingarten equations as below.

Theorem 4.6. *Let S be a surface in $\mathbb{R}^3$ with a parametric equation $q(u^1,u^2)$ and let $N(u^1,u^2)$ be a unit normal vector field on S. Then the frame $\{\partial_1 q,\partial_2 q,N\}$ on S obeys the partial differential equations*

$$\partial_i(\partial_j q) = \Gamma^k_{ij}\partial_k q + L_{ij}N,$$
$$\partial_i N = -L^k_i\partial_k q.$$

Proposition 4.7. *Let $q(u^1,u^2)$ be a parametric equation of S, then*

$$\Gamma_{ij,k} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}).$$

Hence the Christoffel symbols $\Gamma_{ij,k}$ and Γ^k_{ij} can be computed from [**I**].

Proof. For each i, j, and k, we have

$$\partial_k g_{ij} = \partial_k(\partial_i q \cdot \partial_j q) = \partial_{ki} q \cdot \partial_j q + \partial_i q \cdot \partial_{kj} q = \Gamma_{ki,j} + \Gamma_{kj,i}.$$

Using the fact that $\Gamma_{ij,k}$ is symmetric in i and j, we obtain

$$\begin{aligned} &\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij} \\ =&\Gamma_{ij,k} + \Gamma_{ik,j} + \Gamma_{ji,k} + \Gamma_{jk,i} - \Gamma_{ki,j} - \Gamma_{kj,i} \\ =&2\Gamma_{ij,k}. \end{aligned}$$

□

Remark 4.8. For surfaces, the Christoffel symbol $\Gamma_{ij,k}$ always has a repeated pair in $\{i,j,k\} \subseteq \{1,2\}$. One can use this to derive easier computations; see Example 5.4.

We can take a step further to define ∇ as a map.

Definition 4.9. Let $\mathfrak{X}(S)$ be the set of all vector fields on S. We define the *Levi-Civita connection*

$$\nabla : \mathfrak{X}(S) \times \mathfrak{X}(S) \to \mathfrak{X}(S), \quad (X,Y) \mapsto \nabla_X Y$$

by $(\nabla_X Y)(p) = \nabla_{X(p)} Y$, where $p \in S$.

Proposition 4.10. *The Levi-Civita connection ∇ satisfies the following properties: (below X, Y, Z are smooth vector fields on S; f and g are smooth functions on S.)*
(1) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$.
(2) $\nabla_X(aY + bZ) = a\nabla_X Y + b\nabla_X Z$, *where* $a, b \in \mathbb{R}$.
(3) $\nabla_X(fY) = (D_X f)Y + f(\nabla_X Y)$.
(4) $D_X \mathbf{I}(Y,Z) = \mathbf{I}(\nabla_X Y, Z) + \mathbf{I}(X, \nabla_Y Z)$.
(5) $\nabla_X Y - \nabla_Y X = D_X Y - D_Y X$.

Proof. (1,2,3) are left to the readers.

(4) Note that

$$(D_X Y) \cdot Z = (\nabla_X Y) \cdot Z = \mathbf{I}(\nabla_X Y, Z)$$

because the normal part of $D_X Y$ won't contribute to the dot product of $D_X Y$ with a tangent vector. Then

$$D_X \mathbf{I}(Y,Z) = D_X(Y \cdot Z) = (D_X Y) \cdot Z + Y \cdot (D_X Z) = \mathbf{I}(\nabla_X Y, Z) + \mathbf{I}(X, \nabla_Y Z).$$

(5) We decompose $D_X Y - D_Y X$ into tangential and normal parts

$$D_X Y - D_Y X = ((\nabla_X Y) - \mathbf{II}(X,Y)N) - (\nabla_Y X - \mathbf{II}(Y,X)N).$$

The desired formula follows since **II** is symmetric. □

5. Gaussian curvature as a commutator of covariant derivatives

In this section, we prove Gauss's Theorema Egregium. The proof also gives an interpretation of the Gaussian curvature K as a commutator of covariant derivatives.

Recall that for the Euclidean plane $\mathbb{R}^2$ (or $\mathbb{R}^n$), the partial derivatives are always commutative $\partial_{xy} = \partial_{yx}$ as long as we apply them to smooth functions or vector fields. The same commutative feature also applies when we use the polar coordinate: $\partial_{r\theta} = \partial_{\theta r}$, or any other coordinate system of $\mathbb{R}^2$. This no longer holds in general when we consider the intrinsic derivatives on a surface; in other words $\nabla_{\partial_1 q}\nabla_{\partial_2 q}$ and $\nabla_{\partial_2 q}\nabla_{\partial_1 q}$ could be different.

In a Nutshell. *Gaussian curvature is a quantity that measures how commuting the covariant derivatives* $\nabla_{\partial_1 q}$ *and* $\nabla_{\partial_2 q}$ *are.*

Definition 5.1. Let $q(u^1, u^2)$ be a parametric equation for on S. We define

$$R(\partial_i q, \partial_j q)\partial_k q = \nabla_{\partial_i q}\nabla_{\partial_j q}\partial_k q - \nabla_{\partial_j q}\nabla_{\partial_i q}\partial_k q.$$
$$R_{ijkl} = \mathbf{I}(R(\partial_i q, \partial_j q)\partial_k q, \partial_l q).$$

The commutator $\nabla_{\partial_i q}\nabla_{\partial_j q} - \nabla_{\partial_j q}\nabla_{\partial_i q}$ tells how much the covariant derivatives $\nabla_{\partial_i q}$ and $\nabla_{\partial_j q}$ fail to be commutative. By Proposition 4.7, covariant derivatives, thus R_{ijkl}, can be computed from the first fundamental form $[\mathbf{I}]$.

Theorem 5.2. *(Gauss's Theorema Egregium) Under a parametric equation* $q(u^1, u^2)$*, the Gaussian curvature* K *has formula*

$$K = \frac{R_{1221}}{\det[\mathbf{I}]}.$$

In particular, K *can be computed by knowing only* $[\mathbf{I}]$.

Proof. By Remark 3.9(3), it suffices to show that $\det[\mathbf{II}] = R_{1221}$. According to Remark 4.3(1), we have tangential and normal decomposition of $\partial_{ij}q$ as

$$\partial_{ij}q = \nabla_{\partial_i q}\partial_j q + L_{ij}N.$$

Hence

$$\partial_{11}q \cdot \partial_{22}q = \mathbf{I}(\nabla_{\partial_1 q}\partial_1 q, \nabla_{\partial_2 q}\partial_2 q) + L_{11}L_{22}$$
$$\partial_{12}q \cdot \partial_{21}q = \mathbf{I}(\nabla_{\partial_1 q}\partial_2 q, \nabla_{\partial_2 q}\partial_1 q) + L_{12}L_{21}.$$

This allows us to calculate $\det[\mathbf{II}]$ by

$$\begin{aligned}\det[\mathbf{II}] =& L_{11}L_{22} - L_{12}L_{21} \\ =& \{\partial_{11}q \cdot \partial_{22}q - \partial_{12}q \cdot \partial_{21}q\} \\ &+ \{\mathbf{I}(\nabla_{\partial_1 q}\partial_2 q, \nabla_{\partial_2 q}\partial_1 q) - \mathbf{I}(\nabla_{\partial_1 q}\partial_1 q, \nabla_{\partial_2 q}\partial_2 q)\} \\ =:& \{A\} + \{B\}.\end{aligned}$$

We apply Proposition 4.10(4) to the second term $\{B\}$:

$$\begin{aligned}\{B\} =& \partial_2\mathbf{I}(\nabla_{\partial_1 q}\partial_2 q, \partial_1 q) - \mathbf{I}(\nabla_{\partial_2 q}\nabla_{\partial_1 q}\partial_2 q, \partial_1 q) \\ &- \partial_1\mathbf{I}(\partial_1 q, \nabla_{\partial_2 q}\partial_2 q) + \mathbf{I}(\partial_1 q, \nabla_{\partial_1 q}\nabla_{\partial_2 q}\partial_2 q) \\ =& \mathbf{I}(R(\partial_1 q, \partial_2 q)\partial_2 q, \partial_1 q) + \partial_2(\partial_{12}q \cdot \partial_1 q) - \partial_1(\partial_1 q \cdot \partial_{22}q) \\ =& R_{1221} + \partial_{212}q \cdot \partial_1 q + \partial_{12}q \cdot \partial_{21}q - \partial_{11}q \cdot \partial_{22}q - \partial_1 q \cdot \partial_{122}q \\ =& R_{1221} - \{A\}.\end{aligned}$$

This completes the proof of $\det[\mathbf{II}] = \{A\} + \{B\} = R_{1221}$. ☐

Corollary 5.3. *Gaussian curvature is invariant under isometries.*

Proof. This follows immediately from Theorem 5.2 and Proposition 1.4. ☐

Example 5.4. As an example, we calculate the Gaussian curvature of a surface S with first fundamental form

$$[\mathbf{I}] = \begin{pmatrix} 1 & 0 \\ 0 & f(r)^2 \end{pmatrix}$$

under a parametric equation $q(r,\theta)$ of S, where $r>0$. We first calculate the Christoffel symbols $\Gamma_{ij,k}$.

$$\begin{aligned}
\Gamma_{11,1} &= \partial_{11}q\cdot\partial_1 q = \frac{1}{2}\partial_1(\partial_1 q\cdot\partial_1 q) = 0,\\
\Gamma_{22,2} &= \partial_{22}q\cdot\partial_2 q = \frac{1}{2}\partial_2(\partial_2 q\cdot\partial_2 q) = 0,\\
\Gamma_{12,1} = \Gamma_{21,1} &= \partial_{21}q\cdot\partial_1 q = \frac{1}{2}\partial_2(\partial_1 q\cdot\partial_1 q) = 0,\\
\Gamma_{12,2} = \Gamma_{21,2} &= \partial_{12}q\cdot\partial_2 q = \frac{1}{2}\partial_1(\partial_2 q\cdot\partial_2 q) = ff',\\
\Gamma_{11,2} &= \partial_{11}q\cdot\partial_2 q = \partial_1(\partial_1 q\cdot\partial_2 q) - \partial_1 q\cdot\partial_{12}q = 0,\\
\Gamma_{22,1} &= \partial_{22}q\cdot\partial_1 q = \partial_2(\partial_2 q\cdot\partial_1 q) - \partial_2 q\cdot\partial_{21}q = -ff'.
\end{aligned}$$

Then we use $\Gamma_{ij}^k = \Gamma_{ij,l}g^{lk}$ to calculate Γ_{ij}^k.

$$\begin{aligned}
\Gamma_{11}^1 &= \Gamma_{11,l}g^{l1} = 0, \quad \Gamma_{22}^2 = \Gamma_{22,l}g^{l2} = 0,\\
\Gamma_{12}^1 = \Gamma_{21}^1 &= \Gamma_{12,l}g^{l1} = 0, \quad \Gamma_{12}^2 = \Gamma_{21}^2 = \Gamma_{12,l}g^{l2} = \frac{f'}{f},\\
\Gamma_{11}^2 &= \Gamma_{11,l}g^{l2} = 0, \quad \Gamma_{22}^1 = \Gamma_{22,l}g^{l1} = -ff'.
\end{aligned}$$

This leads to the covariant derivatives and their commutator.

$$\begin{aligned}
\nabla_{\partial_1 q}\partial_1 q &= \Gamma_{11}^k\partial_k q = 0,\\
\nabla_{\partial_2 q}\partial_2 q &= \Gamma_{22}^k\partial_k q = -ff'\partial_1 q,\\
\nabla_{\partial_1 q}\partial_2 q = \nabla_{\partial_2 q}\partial_1 q &= \Gamma_{12}^k\partial_k q = \frac{f'}{f}\partial_2 q,\\
\nabla_{\partial_1 q}\nabla_{\partial_2 q}\partial_2 q &= \nabla_{\partial_1 q}(-ff'\partial_1 q) = -(f''f+(f')^2)\partial_1 q,\\
\nabla_{\partial_2 q}\nabla_{\partial_1 q}\partial_2 q &= \nabla_{\partial_2 q}(\frac{f'}{f}\partial_2 q) = \frac{f'}{f}\cdot(-ff')\partial_1 q = -(f')^2\partial_1 q,\\
R(\partial_1 q,\partial_2 q)\partial_2 q &= -(f''f+(f')^2)\partial_1 q + (f')^2\partial_1 q = -f''f\partial_1 q.\\
R_{1221} &= \mathbf{I}(R(\partial_1 q,\partial_2 q)\partial_2 q,\partial_1 q) = -f''f.
\end{aligned}$$

Hence Gaussian curvature

$$K = \frac{R_{1221}}{\det[\mathbf{I}]} = -\frac{f''}{f}.$$

One can use this formula to derive $[\mathbf{I}] = \mathrm{diag}\{1, f^2(r)\}$ with constant curvature:

$$\begin{aligned}
f(r) &= \frac{1}{\sqrt{\kappa}}\sin(\sqrt{\kappa}r), \text{ where } \kappa>0, \quad K=\kappa>0.\\
f(r) &= r, \quad K=0.\\
f(r) &= \frac{1}{\sqrt{-\kappa}}\sinh(\sqrt{-\kappa}r), \text{ where } \kappa<0, \quad K=\kappa<0.
\end{aligned}$$

Remark 5.5. One may wonder besides R_{1221}, whether $\{R_{ijkl}\}$ gives other useful quantities. The answer is no. In fact, R_{ijkl} has skew-symmetry

$$R_{ijkl} = -R_{jikl}, \quad R_{ijkl} = -R_{ijlk}.$$

Hence for $\{i, j, k, l\} \subseteq \{1,2\}$, R_{1221} determines all the nonzero R_{ijkl}.

The skew-symmetry of R_{ijkl} in i and j follows directly from the definition. We prove the skew-symmetry in k and l below.

Proposition 5.6. *For any smooth vector field Z of S, we have*

$$\mathbf{I}(R(\partial_i q, \partial_j q)Z, Z) = 0,$$

where $R(\partial_i q, \partial_j q)Z = \nabla_{\partial_i q}\nabla_{\partial_j q}Z - \nabla_{\partial_j q}\nabla_{\partial_i q}Z$. Consequently, $R_{ijkl} = R_{ijlk}$.

Proof. Let us consider the function $f = \frac{1}{2}\mathbf{I}(Z, Z)$. By Proposition 4.10(4), we have

$$\begin{aligned}\partial_i f &= \frac{1}{2} D_{\partial_i q}\mathbf{I}(Z, Z) = \mathbf{I}(\nabla_{\partial_i q}Z, Z);\\ \partial_{ji} f &= D_{\partial_j q}\mathbf{I}(\nabla_{\partial_i q}Z, Z)\\ &= \mathbf{I}(\nabla_{\partial_j q}\nabla_{\partial_i q}Z, Z) + \mathbf{I}(\nabla_{\partial_i q}Z, \nabla_{\partial_j q}Z).\end{aligned}$$

Switching i and j yields

$$\partial_{ij} f = \mathbf{I}(\nabla_{\partial_i q}\nabla_{\partial_j q}Z, Z) + \mathbf{I}(\nabla_{\partial_j q}Z, \nabla_{\partial_i q}Z).$$

Subtracting one by the other, we have

$$0 = \mathbf{I}(\nabla_{\partial_i q}\nabla_{\partial_j q}Z, Z) - \mathbf{I}(\nabla_{\partial_j q}\nabla_{\partial_i q}Z, Z) = \mathbf{I}(R(\partial_i q, \partial_j q)Z, Z).$$

To see the skew-symmetry in k and l, we use $Z = \partial_k q + \partial_l q$. Then by Proposition 4.10(2),

$$\begin{aligned}0 =& \mathbf{I}(R(\partial_i q, \partial_j q)Z, Z)\\ =& \mathbf{I}(R(\partial_i q, \partial_j q)\partial_k q, \partial_k q) + \mathbf{I}(R(\partial_i q, \partial_j q)\partial_l q, \partial_l q)\\ &+ \mathbf{I}(R(\partial_i q, \partial_j q)\partial_k q, \partial_l q) + \mathbf{I}(R(\partial_i q, \partial_j q)\partial_l q, \partial_k q)\\ =& R_{ijkl} + R_{ijlk}.\end{aligned}$$

□