Notes for MATH 561 Differential Geometry

Zhan Cheng

Spring 2025

1 Lecture 1

1.1 Felix Klein's Erlangen-Nürnberg Theorem

Let (X, G) be a pair where:

- X is a set, and
- G is a group with operation $\cdot$.

A group G is defined as a set with:

1. A binary operation $G \times G \to G$, $\forall g_1, g_2, g_3 \in G$, $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$.
2. An identity element $e \in G$ such that $g \cdot e = g = e \cdot g$, $\forall g \in G$.
3. An inverse $g^{-1} \in G$ for every $g \in G$, satisfying $g \cdot g^{-1} = g^{-1} \cdot g = e$.

Lemma 1.1. *The group $(G, \cdot)$ satisfies:*

1. *The identity element is unique.*
2. *The inverse g^{-1} is unique for each $g \in G$.*

Proof. 1. Assume there are two identity elements e_1 and e_2. By the definition of the identity element:

$$e_1 = e_1 \cdot e_2 = e_2.$$

Hence, the identity element is unique.

2. Let $g \in G$ and suppose g has two inverses g_1 and g_2. By the definition of the inverse:

$$g_1 = g_1 \cdot e = g_1 \cdot (g \cdot g_2) = (g_1 \cdot g) \cdot g_2 = e \cdot g_2 = g_2.$$

Therefore, the inverse of g is unique.

□

1.2 Transformation Group

Definition 1.1 (Transformation Group). *A group G is called a transformation group on X if G is a set of bijections $\phi : X \to X$ and the group multiplication is given by the function composition.*

Example 1: Euclidean Isometries

Let $X = \mathbb{R}^n$ with the Euclidean distance $d_0(x, y) = \|x - y\|$. A Euclidean isometry is a mapping $\phi : \mathbb{R}^n \to \mathbb{R}^n$ such that:

$$d_0(\phi(x), \phi(y)) = d_0(x, y), \quad \forall x, y \in \mathbb{R}^n.$$

The group of isometries of $\mathbb{R}^n$ is:

$$\mathrm{Isom}(\mathbb{R}^n) = \{\phi \mid \phi(x) = Ax + b,\, b \in \mathbb{R}^n,\, A \in O(n)\}.$$

Here:

1. $O(n)$ is the set of all $n \times n$ orthogonal matrices, i.e., $A^T A = I$.
2. $b \in \mathbb{R}^n$ is a translation vector.

Question

Why do Euclidean isometries use orthogonal matrices?

Answer

Orthogonal matrices preserve the length of vectors and angles. For $\phi(x) = Ax + b$, $d_0(\phi(x), \phi(y)) = \|A(x-y)\| = \|x-y\|$, which holds only if $A^\top A = I$. Geometrically, orthogonal matrices represent rotations and reflections.

Example

1. Rotation Matrix:
$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$
This rotates vectors by angle θ around the origin.
2. Reflection Matrix:
$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$
This reflects vectors across the x-axis.

Example 2: Affine Transformations

Affine transformations extend Euclidean isometries by allowing general linear mappings and translations. An affine transformation ϕ is defined as:

$$\phi(x) = Ax + b, \quad A \in \mathrm{GL}(n, \mathbb{R}),\, b \in \mathbb{R}^n,$$

where:

1. $A \in \mathrm{GL}(n, \mathbb{R})$, the group of all invertible $n \times n$ matrices.
2. $b \in \mathbb{R}^n$ is a translation vector.

Question

Why must the matrix in an affine transformation be invertible?

Answer

Invertibility ensures:
1. No collapse to lower dimensions (e.g., projecting $\mathbb{R}^2$ to a line).
2. Bijectivity: every point has a unique preimage, enabling the inverse transformation:

$$\phi^{-1}(x) = A^{-1}(x - b).$$

Example

1. Non-invertible matrix:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

This projects all points onto the x-axis, losing information about the y-coordinate.
2. Invertible matrix:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

This preserves the full dimensionality of the space.

1.3 Metric Space

Definition 1.2 (Metric Space). *A metric space is a pair (X, d) where X is a set and $d : X \times X \to [0, \infty)$ is a distance function satisfying:*

1. *$d(p, q) \geq 0$, and $d(p, q) = 0 \iff p = q$, $\forall p, q \in X$.*
2. *$d(p, q) = d(q, p)$, $\forall p, q \in X$.*
3. *$d(p, q) \leq d(p, w) + d(w, q)$, $\forall p, q, w \in X$.*

Lemma 1.2. *For $(\mathbb{R}^n, d_0)$, $K \subset \mathbb{R}^n$, Two True, Force the Third:*

1. *K is compact.*
2. *K is sequentially compact.*
3. *K is bounded and closed.*

Question

What is compactness, sequential compactness, and the relationship with bounded and closed sets?

Answer

1. Compactness: Every open cover has a finite subcover. In $\mathbb{R}^n$, compactness $\iff$ bounded and closed.
2. Sequential Compactness: Every sequence in the set has a convergent subsequence within the set.
3. Bounded and Closed: A set is compact in $\mathbb{R}^n$ if it satisfies these properties.

Example

1. $[0,1]$: - Compact: Yes (bounded and closed). - Sequentially compact: Yes.
2. $(0,1)$: - Compact: No (not closed). - Sequentially compact: No (e.g., sequence $x_n = \frac{1}{n}$ converges to 0, which is not in $(0,1)$).
3. $\mathbb{R}$: - Compact: No (not bounded).

1.4 Arc Length

Definition (Arc Length): Let (X,d) be a metric space and $\gamma : [0,1] \to X$ be a curve. Then the arc length of γ is:

$$L(\gamma) = \sup_{k\in\mathbb{Z}_+} \sum_{j=1}^{k} d(\gamma(t_{j-1}), \gamma(t_j)),$$

where $0 = t_0 < t_1 < t_2 < \cdots < t_{k-1} < t_k = 1$ and $\{t_j\}_{j=0}^k$ is a partition of $[0,1]$. If $L(\gamma) < \infty$, the curve γ is called *rectifiable*.

Question

Why does the definition of arc length use sup instead of inf?

Answer

Arc length is defined as the supremum over all possible partitions of the sum of distances between consecutive points on a curve. Using sup ensures that finer partitions can better approximate the true curve length. If inf were used, coarse partitions might underestimate the length, failing to capture the curve's geometry.

Example

For the parabola $\gamma(t) = (t, t^2), t \in [0,1]$:
- Actual arc length: $L(\gamma) > 1.478$.
- Coarse partition (e.g., $t_0 = 0, t_1 = 1$): Straight-line length $\sqrt{2} \approx 1.414$, which underestimates the true length.

1.5 Length Metric

Definition (Length Metric): Let (X,d) be a metric space, the length metric $d^{\#}(p,q)$ is defined as:

$$d^{\#}(p,q) = \inf L(\gamma),$$

where $\gamma : [0, 1] \to X$ is a rectifiable curve connecting p and q.

Question

Why does the length metric use inf?

Answer

The length metric $d^{\#}(p, q)$ is defined as the infimum of arc lengths over all curves connecting p and q. This ensures that the metric captures the shortest possible distance, avoiding unnecessary detours or longer paths.

Example

1. In $\mathbb{R}^2$, the length metric $d^{\#}(p, q)$ equals the Euclidean distance $d(p, q)$ because the straight line minimizes the length.
2. On a sphere, $d^{\#}(p, q)$ corresponds to the length of the shortest great circle arc between p and q.

2 Lecture 2

2.1 Curve

Definition (Curve): A mapping $\gamma : I \to X$, where $I \subseteq \mathbb{R}$ is an interval, is said to be a **curve** if γ is continuous.

Lemma 2.1. *If (X, d) is a metric space, then $(X, d^{\#})$ is also a metric space.*

Theorem 2.1. *$(\mathbb{R}^n, d_0)$ is a connected and complete metric space (and is also path-connected).*

2.2 Connectedness

Definition (Connectedness): A metric space (X, d) is said to be **connected** if X does not admit any separation:

$$X = A \cup B, \quad A, B \neq \varnothing, \quad A \cap B = \varnothing, \quad \text{with } A, B \text{ both open.}$$

Definition (Path Connectedness): A metric space X is said to be **path connected** if $\forall p, q \in X$, there exists a continuous curve γ connecting p and q.

Lemma 2.2. *Path connected implies connected.*

2.3 Parameterization by Arc Length

Let $\gamma : [0, 1] \to \mathbb{R}^n$ be a C^1-curve. Denote:

- $\gamma'(t)$: tangent vector at t,
- $\|\gamma'(t)\|$: speed at t.

The length of γ is given by:

$$L(\gamma) = \int_0^1 \|\gamma'(t)\|dt.$$

Define the arc length function:

$$s(t) = \int_0^t \|\gamma'(\tau)\|d\tau, \quad s'(t) = \|\gamma'(t)\| \geq 0.$$

Definition (Parameterization by Arc Length): Let $\gamma : [0, L(\gamma)] \to \mathbb{R}^n$ and $L(\gamma\Big|_{[0,s]}) = s$.

Lemma 2.3. $\|\gamma'(s)\| = 1$.

Lemma 2.4. *Any C^1 curve meets a constant speed parameterization.*

2.4 Coincidence of Distance Structures

Proposition 2.1. *The two distance structures d_0 and $d_0^{\#}$ coincide on $(\mathbb{R}^n, d_0)$, where $d_0^{\#}$ is realized by $\|p - q\|$.*

Proof. Let $\gamma_0 : [0, 1] \to \mathbb{R}^n$ be a smooth curve connecting p and q, with constant speed v. We define the variation of a curve γ_0 as:

$$V(t, s) = \gamma_0(t) + s\psi(t),$$

where $-\epsilon \leq s \leq \epsilon$, ψ is a smooth function with $\psi(0) = \psi(1) = 0$.

Thus, $V(0, s) = p$, $V(1, s) = q$, and $L(s) = L(V(\cdot, s)) = \int_0^1 \|\partial_t V(t, s)\|dt$.

We compute $L'(0)$ as follows:

$$\begin{aligned}
L'(0) &= \frac{d}{ds}\Big|_{s=0} \int_0^1 \|\partial_t V(t, s)\|\, dt \\
&= \int_0^1 \frac{\partial}{\partial s}\Big|_{s=0} \|\partial_t V(t, s)\|\, dt \\
&= \int_0^1 \frac{\frac{\partial}{\partial s}\langle \partial_t V(t, s), \partial_t V(t, s)\rangle^{1/2}}{\|\partial_t V(t, s)\|}\Big|_{s=0} dt \\
&= \int_0^1 \frac{1}{2}\langle \partial_t V(t, s), \partial_t V(t, s)\rangle^{-1/2} \cdot 2\langle \frac{\partial^2 V}{\partial s \partial t}, \partial_t V(t, s)\rangle\Big|_{s=0} dt \\
&= \int_0^1 \frac{\langle \frac{\partial^2 V}{\partial s \partial t}, \partial_t V(t, s)\rangle}{\|\partial_t V(t, s)\|}\Big|_{s=0} dt \\
&= \int_0^1 v^{-1}\langle \psi'(t), \gamma_0'(t)\rangle\, dt \\
&= v^{-1}\langle \psi(t), \gamma_0'(t)\rangle\Big|_0^1 - v^{-1}\int_0^1 \langle \psi(t), \gamma_0''(t)\rangle\, dt \\
&= -v^{-1}\int_0^1 \langle \psi(t), \gamma_0''(t)\rangle\, dt,
\end{aligned}$$

where we used the fact that $\psi(0) = \psi(1) = 0$ to eliminate the boundary terms.

Thus, combining these results, we find that the variation $L'(0)$ vanishes when γ_0 is a straight line connecting p and q, since in this case $\gamma_0''(t) = 0$ for all t. This implies that the shortest path connecting p and q in the distance structure $d_0^{\#}$ is realized by the Euclidean norm $\|p - q\|$.

Finally, since d_0 also corresponds to the Euclidean distance, it follows that d_0 and $d_0^{\#}$ coincide on $(\mathbb{R}^n, d_0)$. □

3 Lecture 3

3.1 Length Space

Definition (Length Space): A metric space (X, d) is called a **length space** if $d^{\#} = d$.

Theorem 3.1. *$(\mathbb{R}^n, d_0)$ is a length space. In other words, for all $p, q \in \mathbb{R}^n$, we have $d_0(p, q) = d_0^{\#}(p, q)$. Moreover, the shortest curve connecting p and q is the line segment pq.*

Remark:

$$d_0(p, q) \leq d_0^{\#}(p, q) \leq \inf\{\ell(\gamma) \mid \gamma \text{ is a smooth curve connecting } p, q\} = d_0(p, q)$$

$$\Rightarrow d_0(p, q) = d(p, q).$$

3.2 Isometry

Definition (Isometry): A function $F : \mathbb{R}^n \to \mathbb{R}^n$ is called an **isometry** if F preserves distances:

$$d_0(F(x), F(y)) = d_0(x, y), \quad \forall x, y \in \mathbb{R}^n.$$

Lemma 3.1. *If $F(x) = Ax + b$, where $A \in O(n)$ and $b \in \mathbb{R}^n$, then F is an isometry.*

Proof.

$$\begin{aligned}
[d_0(F(x), F(y))]^2 &= ||Ax - Ay||^2 \\
&= \langle A(x - y), A(x - y)\rangle \\
&= \langle A^T A(x - y), x - y\rangle \\
&= ||x - y||^2 = [d_0(x, y)]^2.
\end{aligned}$$

Thus, $d_0(F(x), F(y)) = d_0(x, y)$. □

Lemma 3.2. *1. The set of all Euclidean isometries is a group with respect to composition, denoted by Isom$(\mathbb{R}^n)$*

2. The set of all affine mappings $x \to Ax + b$, where $A \in O(n)$ and $b \in \mathbb{R}^n$, is also a group with respect to composition.

Definition (Isometry): More generally, given two metric spaces (X, d_X) and (Y, d_Y), a function $F : X \to Y$ is called an *isometry* if:

1. It is an **isometric embedding**:

$$d_Y(F(p), F(q)) = d_X(p, q), \quad \forall p, q \in X.$$

2. F is subjective.

Lemma 3.3. *The following are equivalent:*

1. *$F : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry and $F(0) = 0$.*
2. *$F : \mathbb{R}^n \to \mathbb{R}^n$ preserves inner products $\langle \cdot, \cdot \rangle$.*

Proof. We prove (2) $\Rightarrow$ (1) here.

$$[d_0(F(x), F(y))]^2 = ||F(x) - F(y)||^2$$

$$= ||F(x)||^2 + ||F(y)||^2 - 2\langle F(x), F(y)\rangle.$$

Also, by

$$||F(0)||^2 = \langle F(0), F(0)\rangle = \langle 0, 0\rangle = 0,$$

we get

$$||F(x)||^2 = [d_0(F(x), F(0))]^2 = \langle F(x), F(0)\rangle = ||x||^2,$$
$$||F(y)||^2 = ||y||^2,$$

Since

$$\langle F(x), F(y)\rangle = \langle x, y\rangle,$$

we prove that

$$d_0(F(x), F(y)) = d_0(x, y).$$

□

Proposition 3.1. *If $\{e_j\}_{j=1}^n$ is an orthonormal basis of $\mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry satisfying:*

1. *$F(0) = 0$.*
2. *$F(e_j) = e_j, \quad \forall 1 \leq j \leq n$.*

Then F must be the identity.

Proof. For any $x = \sum_{j=1}^n x_j e_j$, we want to show:

$$\langle F(x), e_j\rangle = \langle x, e_j\rangle, \quad \forall 1 \leq j \leq n.$$

By Lemma 3.3, F preserves inner products:

$$\langle F(x), e_j\rangle = \langle F(x), F(e_j)\rangle = \langle x, e_j\rangle.$$

□

Theorem 3.2. *The following are equivalent:*

1. *$F : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry.*
2. *F is an affine mapping, where $F(x) = Ax + b$, where $A \in O(n)$ and $b \in \mathbb{R}^n$*

Proof. We prove (1) $\Rightarrow$ (2) here since (2) $\Rightarrow$ (1) is proved by Lemma 3.1.

Define $b = F(0) \in \mathbb{R}^n$ and set $\hat{F}(x) = F(x) - b$ so that $\hat{F}(0) = 0$. Let $\{e_j\}_{j=1}^n$ be an orthonormal basis. Define: $\hat{e}_j = \hat{F}(e_j), \quad \forall 1 \leq j \leq n$.

Since $\{\hat{e}_j\}$ is an orthonormal basis, there exists a unique $\hat{G} \in O(n)$ such that:

$$\hat{G}(\hat{e}_j) = e_j, \quad \forall 1 \leq j \leq n.$$

Thus, the mapping $\hat{G} \circ \hat{F} : \mathbb{R}^n \to \mathbb{R}^n$ preserves the origin and satisfies:

$$\hat{G} \circ \hat{F}(e_j) = e_j \quad \forall 1 \leq j \leq n,$$

which implies that

$$\hat{G} \circ \hat{F} = \mathrm{Id}$$

Therefore, $F(x) = \hat{G}^{-1}(x) + b$. □

3.3 Local Geometry of Euclidean Curves

Definition (Regular): A smooth curve $\gamma : [0,1] \to \mathbb{R}^n$ has a **regular** parameterization if:

$$\gamma'(t) \neq 0.$$

Let $\gamma : (0,l) \to \mathbb{R}^n$ be parameterized by arc length:

$$\|\gamma'(s)\| = 1, \quad \forall s \in (0,l).$$

We denote $T(s) = \gamma'(s)$, $N(s) \equiv \{\text{Unit vector in the direction of } \gamma''(s)\}$.

Definition (Curvature): The **curvature** of a curve is:

$$\kappa(s) = \|T'(s)\| = \|\gamma''(s)\|, \text{ where } \gamma''(s) = \kappa(s)N(s).$$

Example:

The curve $\gamma(s) = (\cos(s), \sin(s))$ has a constant curvature $\kappa = 1$.

4 Lecture 4

Lemma 4.1. *A regular curve $\gamma : [0,1] \to \mathbb{R}^n$ with $F : \mathbb{R}^n \to \mathbb{R}^n$, $F \in Isom(\mathbb{R}^n)$, we have*

$$\forall t_0 \in [0,1] \quad \|\frac{d}{dt}\Big|_{t=t_0}(F \circ \gamma)\| = \|\frac{d}{dt}\Big|_{t=t_0}\gamma\|$$

Proof. Let $\bar{\gamma} = F \circ \gamma$, we compute

$$\frac{d\bar{\gamma}}{dt}\Big|_{t=t_0} = \mathrm{DF}(\frac{d\gamma}{dt}\Big|_{t=t_0}) \quad \text{, where DF is the Jacobian matrix of } F,$$

given by

$$DF = \begin{bmatrix} \frac{\partial f_1}{\partial t_1} & \cdots & \frac{\partial f_1}{\partial t_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial t_1} & \cdots & \frac{\partial f_n}{\partial t_n} \end{bmatrix}$$

Then

$$LHS = \|\mathrm{DF}(\gamma')\| = \|A\gamma'\| = \|\gamma'\| = RHS,$$

where $F(x) = Ax + b$, $A \in O(n), b \in \mathbb{R}^n$. □

Note that Lemma 4.1 actually tells us that isometry doesn't change the curvature in $\mathbb{R}^n$.

Lemma 4.2. *Let $\gamma : [0, l] \to \mathbb{R}^n$ has an arc-length parameter s, with an isometry $F : \mathbb{R}^n \to \mathbb{R}^n$, $F \in Isom(\mathbb{R}^n)$, then*

$$\forall 0 \leq s \leq l \quad \kappa(\tilde{\gamma}(s)) = \kappa(\gamma(s)), \quad \textit{where } \tilde{\gamma} = F \circ \gamma$$

Proof.

$$\kappa(\tilde{\gamma}(s)) = \left\| \frac{d}{ds}\frac{d}{ds}(F \circ \gamma) \right\| = \left\| \frac{d}{ds}\bar{T}(s) \right\| = \left\| \frac{d}{ds}(\mathrm{DF}(\gamma'(s))) \right\| = \|A\gamma''(s)\| = \|\gamma''(s)\|$$

□

Let $\gamma : [0, 1] \to \mathbb{R}^n$ be a regular curve, then $\gamma(s)$ can be written as $\gamma(t(s))$. Thus,

$$\frac{d\gamma}{ds} = (\frac{ds}{dt})\frac{d\gamma}{dt} = \frac{1}{\|\gamma'(t)\|}\frac{d\gamma}{dt}.$$

We note that

$$s(t)\int_0^t = \|\gamma'(\tau)\|d\tau \Rightarrow \frac{ds}{dt} = \|\gamma'(t)\|.$$

Also,

$$\frac{d^2\gamma}{ds^2} = \frac{d}{ds}(\frac{\gamma'(t)}{\|\gamma'(t)\|}) = \frac{d}{dt}(\frac{\gamma'(t)}{\|\gamma'(t)\|})\frac{dt}{ds} = \frac{1}{\|\gamma'(t)\|}\frac{d}{dt}(\frac{\gamma'(t)}{\|\gamma'(t)\|}),$$

which implies that

$$\kappa(\gamma(s)) = \frac{\|\frac{d}{dt}(\frac{\gamma'(t)}{\|\gamma'(t)\|})\|}{\|\gamma'(t)\|}$$

4.1 Regular Curves in $\mathbb{R}^3$

Let $\gamma : [0, 1] \to \mathbb{R}^3$ be a regular curve, having an arc-length parameter s. As mentioned above, we denote $T(s) = \gamma'(s)$, $\frac{dT}{ds} = \|T'(s)\| = \|\gamma''(s)\|$, where $\gamma''(s) = \kappa(s)N(s)$.

Definition (Wedge product): The **Wedge product** between vector $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ is given by :

$$u \wedge v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

We denote that $\beta(s) = T(s) \wedge N(s)$. Since $\|\gamma'(s)\| = 1$ by definition, then

$$0 \equiv \frac{d}{ds}\langle\gamma', \gamma'\rangle = 2\langle\gamma'', \gamma'\rangle,$$

which implies that $\gamma'' \perp \gamma'$. Here $N \perp T$.

We can naturally give out that $\{T, N, \beta\}$ is an orthogonal basis on $\mathbb{R}^3$. Moreover,

$$T \wedge N = \beta, N \wedge \beta = T, \beta \wedge T = N.$$

This is known as **Frenet Frame**.

Theorem 4.1. *Let $\{T, N, \beta\}$ be the Frenet frame at $\gamma(s)$, then*

$$\begin{cases} \dfrac{dT}{ds} = \kappa N \\ \dfrac{dN}{ds} = -\tau\beta - \kappa T \\ \dfrac{d\beta}{ds} = \tau N \end{cases}$$

Here τ is called the torsion of γ, given by $\beta'(s) = \tau(s)N(s)$. Moreover, it can be written as

$$\begin{bmatrix} T \\ N \\ \beta \end{bmatrix}' = \Phi \begin{bmatrix} T \\ N \\ \beta \end{bmatrix},$$

where

$$\Phi = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & -\tau \\ 0 & \tau & 0 \end{bmatrix}.$$

Proof. Firstly, $\frac{dT}{ds} = \kappa N$ is proved by definition. Since $\langle N, N\rangle = 1$, we have

$$0 \equiv \frac{d}{ds}\langle N, N\rangle = 2\langle \frac{dN}{ds}, N\rangle \Rightarrow \frac{dN}{ds} \perp N \Rightarrow \frac{dN}{ds} \in \mathrm{Span}\{\beta, T\}.$$

Since $\beta = T \wedge N$, then

$$\frac{d\beta}{ds} = T'(s) \wedge N(s) + T(s) \wedge N'(s) = T(s) \wedge N'(s) \Rightarrow \frac{d\beta}{ds} \perp T \Rightarrow \frac{d\beta}{ds} \in \mathrm{Span}\{\beta, N\}.$$

Also,

$$0 \equiv \frac{d}{ds}\langle \beta, \beta\rangle = 2\langle \frac{d\beta}{ds}, \beta\rangle \Rightarrow \frac{d\beta}{ds} \perp \beta$$

Thus,

$$\frac{d\beta}{ds} \parallel N,$$

and we write $\beta'(s) = \tau(s)N(s)$.

Additionally,

$$\frac{dN}{ds} = \frac{d\beta}{ds} \wedge T + \beta \wedge \frac{dT}{ds} = \tau N \wedge T + \beta \wedge \kappa N = -\tau\beta - \kappa T.$$

□