Closed Curves and Space Curves

(COMS 4770/5770 Notes)

Yan-Bin Jia

Nov 14, 2024

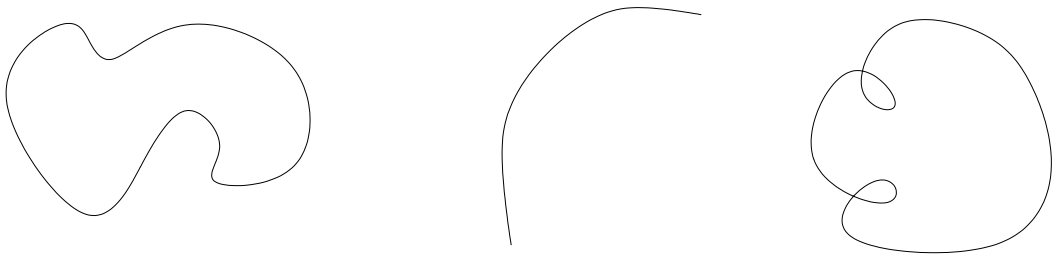
So far we have discussed only 'local' properties of (plane) curves. These properties depend only on the behavior of a curve near a given point, and not on the 'global' shape of the curve. Now let us look at some global results about curves. The most famous, and perhaps the oldest, of these is the 'isoperimetric inequality', which relates the length of a certain 'closed' curve to the area it contains. Later on, we will discuss space curves with an introduction to the celebrated Frenet formula.

1 Simple Closed Curves

Intuitively, simple closed curves are the curves that 'join up', but do not otherwise self-intersect. More precisely, a *simple closed curve* in $\mathbb{R}^2$ with period δ, where $\delta \in \mathbb{R}$, is a regular curve $\boldsymbol{\alpha} : \mathbb{R} \to \mathbb{R}^2$ such that

$$\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}(u) \qquad \text{if and only if} \qquad t - u = k\delta \text{ for some integer } k.$$

Thus, the point $\boldsymbol{\alpha}(t)$ returns to its starting point when t increases by δ but not before that. In the figure below, the first curve is simple closed, the second one is simple but not closed, and the third one is closed but not simple.



A simple closed curve is also referred to as a *Jordan curve*. A standard, but highly non-trivial, result of the topology of $\mathbb{R}^2$, called the *Jordan Curve Theorem*, states that any simple closed curve in the plane has an 'interior' and an 'exterior'. More precisely, according to the theorem, the set of points of $\mathbb{R}^2$ not on the curve is the disjoint union of two subsets of $\mathbb{R}^2$, denoted by $\text{int}(\boldsymbol{\alpha})$ and $\text{ext}(\boldsymbol{\alpha})$, such that

(i) $\text{int}(\boldsymbol{\alpha})$ is *bounded*;

(ii) $\text{ext}(\boldsymbol{\alpha})$ is *unbounded*;

(iii) both of the regions $\mathrm{int}(\boldsymbol{\alpha})$ and $\mathrm{ext}(\boldsymbol{\alpha})$ are *connected*, that is, they have the property that any two points in the same region can be joined by a curve lying entirely in the region.

EXAMPLE 1. The parametrized circle

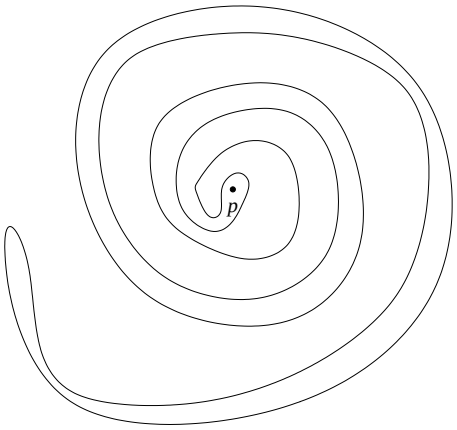
$$\boldsymbol{\alpha}(t) = \left(\cos\frac{2\pi t}{\delta}, \sin\frac{2\pi t}{\delta}\right)$$

is a simple closed curve with period δ. The interior and exterior of $\boldsymbol{\alpha}$ are given by

$$\{\,(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\,\} \qquad \text{and} \qquad \{\,(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 > 1\,\},$$

respectively.

Not all examples of simple closed curves have such an obvious interior and exterior, however. Look at the simple closed curve below. Is the point $\boldsymbol{p}$ in its interior or exterior?



Since every point on a simple closed curve $\boldsymbol{\alpha}$ of period δ is traced out as the parameter t of $\boldsymbol{\alpha}$ varies through any interval of length δ, for example, $[0, \delta]$, it is reasonable to define the *length* of $\boldsymbol{\alpha}$ to be

$$\ell(\boldsymbol{\alpha}) = \int_0^\delta \|\dot{\boldsymbol{\alpha}}(t)\|\, dt.$$

Since $\boldsymbol{\alpha}$ is regular, it has a unit-speed reparametrization $\tilde{\boldsymbol{\alpha}}$ with the arc length

$$s = \int_0^t \|\dot{\boldsymbol{\alpha}}(u)\|\, du$$

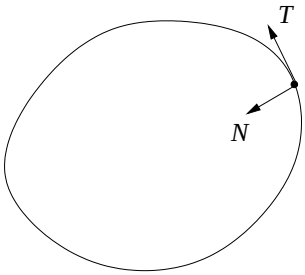
of $\boldsymbol{\alpha}$ as its parameter such that $\tilde{\boldsymbol{\alpha}}(s) = \boldsymbol{\alpha}(t)$. We have that

$$\begin{aligned} s(t+\delta) &= \int_0^{t+\delta} \|\dot{\boldsymbol{\alpha}}(u)\|\, du \\ &= \int_0^{\delta} \|\dot{\boldsymbol{\alpha}}(u)\|\, du + \int_{\delta}^{t+\delta} \|\dot{\boldsymbol{\alpha}}(u)\|\, du \\ &= \ell(\boldsymbol{\alpha}) + \int_0^{t} \|\dot{\boldsymbol{\alpha}}(v+\delta)\|\, dv \\ &= \ell(\boldsymbol{\alpha}) + \int_0^{t} \|\dot{\boldsymbol{\alpha}}(v)\|\, dv \\ &= \ell(\boldsymbol{\alpha}) + s(t). \end{aligned}$$

Hence we conclude that

$$\begin{aligned} \tilde{\boldsymbol{\alpha}}\Big(s(t)\Big) = \tilde{\boldsymbol{\alpha}}\Big(s(u)\Big) &\iff \boldsymbol{\alpha}(t) = \boldsymbol{\alpha}(u) \\ &\iff u - t = k\delta \qquad\qquad \text{some integer } k \\ &\iff s(u) - s(t) = k\ell(\boldsymbol{\alpha}). \end{aligned}$$

This shows that $\tilde{\boldsymbol{\alpha}}$ is a simple closed curve with period ℓ. In short, *we can always assume that a simple closed curve is unit-speed and that its period is equal to length.*



By convention, $\boldsymbol{\alpha}$ is *positively oriented*, that is, at every point t the unit normal $N(t)$ points into its interior. The *area* contained by $\boldsymbol{\alpha}$ is defined as

$$\mathcal{A}(\boldsymbol{\alpha}) = \int\!\!\int_{\mathrm{int}(\boldsymbol{\alpha})} dxdy.$$

This can be computed using *Green's Theorem*. According to the theorem, for all functions $f(x,y)$ and $g(x,y)$ continuously differentiable over the region $\mathrm{int}(\boldsymbol{\alpha})$,

$$\int\!\!\int_{\mathrm{int}(\boldsymbol{\alpha})} \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dxdy = \int_{\boldsymbol{\alpha}} f(x,y)\, dx + g(x,y)\, dy.$$

Proposition 1 *Let $\boldsymbol{\alpha}(t) = (x(t), y(t))$ be a positively-oriented simple closed curve in $\mathbb{R}^2$ with period δ. Then*

$$\mathcal{A}(\boldsymbol{\alpha}) = \frac{1}{2}\int_0^{\delta} (x\dot{y} - y\dot{x})\, dt.$$

Proof Let $f = -\frac{1}{2}y$, $g = \frac{1}{2}x$. Then by Green's theorem, we get

$$\begin{aligned} \mathcal{A}(\boldsymbol{\alpha}) &= \int\!\!\int_{\mathrm{int}(\boldsymbol{\alpha})} \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dxdy \\ &= \frac{1}{2}\int_{\boldsymbol{\alpha}} x\, dy - y\, dx \\ &= \frac{1}{2}\int_0^{\delta} (x\dot{y} - y\dot{x})\, dt. \end{aligned}$$

□

When applying Green's theorem, if we choose $f(x,y) = 0$ and $g(x,y) = x$, then we end up with a "simpler" form $\mathcal{A}(\boldsymbol{\alpha}) = \int_0^\delta x\dot{y}\,dt$. Or if we choose $f(x,y) = -y$ and $g(x,y) = 0$, we end up with $\mathcal{A}(\boldsymbol{\alpha}) = \int_0^\delta -y\dot{x}\,dt$.

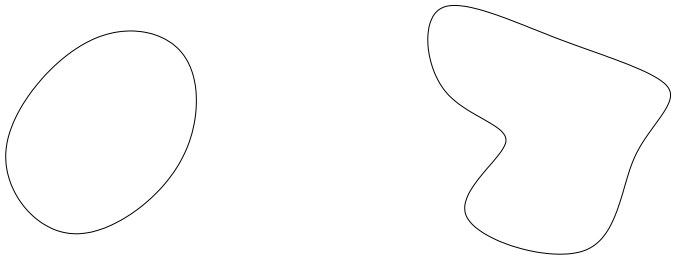
The most important global result about plane curves is given by the theorem below.

Theorem 2 (The Isoperimetric Inequality) *Let $\boldsymbol{\alpha}$ be a simple closed curve with length ℓ and area $\mathcal{A}$. Then*

$$A \leq \frac{1}{4\pi}\ell^2,$$

where equality holds if an only if $\boldsymbol{\alpha}$ is a circle.

We refer to [2, pp. 51–54] for a proof of the theorem. A simple closed curve $\boldsymbol{\alpha}$ is called *convex* if its interior $\text{int}(\boldsymbol{\alpha})$ is convex, in the sense that the straight line segment joining any two points in $\text{int}(\boldsymbol{\alpha})$ is contained entirely in $\text{int}(\boldsymbol{\alpha})$.



convex not convex

Theorem 3 (Four Vertex Theorem) *Every convex simple closed curve $\boldsymbol{\alpha}$ in $\mathbb{R}^2$ has at least four vertices.*

For a proof of this theorem, we refer to [2, pp. 56-57].

2 Space Curves

While a plane curve is determined by its curvature, this is no longer true for space curves. For example, a circle $(\cos t, \sin t)$ in the xy-plane and a circular helix $(\frac{1}{2}\cos t, \frac{1}{2}\sin t, \frac{1}{2}t)$ both have unit curvature everywhere, but it is impossible to change one curve into the other by any combination of rotations and translations. We shall define another type of curvature for space curves, called the *torsion*. As a matter of fact, the curvature and torsion of a curve together determine the curve up to a rigid motion.

Let $\boldsymbol{\gamma}(s)$ be a unit-speed curve in $\mathbb{R}^3$. Denote by T the unit tangent vector. Thus $T = \dot{\boldsymbol{\gamma}}(s)$. The real-valued function $\kappa(s)$ such that $\kappa(s) = \|\dot{T}(s)\| \geq 0$ is called the *curvature function* of $\boldsymbol{\gamma}$. Suppose the curvature $\kappa(s)$ is never zero. We define the *principal normal* of $\boldsymbol{\gamma}$ at the point $\boldsymbol{\gamma}(s)$ to be the vector

$$N(s) = \frac{1}{\kappa(s)}\dot{T}(s).$$

Since $\|\dot{T}(s)\| = \kappa$, N is a unit vector. Because $\dot{T} \cdot T = \frac{1}{2}\frac{d}{dt}(T \cdot T) = 0$, T and N are orthogonal to each other. It follows that

$$B = T \times N$$

is a unit vector orthogonal to both T and N. This vector $B(s)$ is called the *binormal vector* of $\boldsymbol{\gamma}$. Thus, $\{T(s), N(s), B(s)\}$ is an orthonormal basis of $\mathbb{R}^3$ such that

$$B = T \times N, \qquad N = B \times T, \qquad T = N \times B.$$

This basis is called the *Frenet frame*.

The key to the successful study of the geometry of $\boldsymbol{\gamma}$ is to use T, N, B, whenever possible, instead of the x-, y-, z-axes, which contain no information about the curve. The first and most important use of this idea is to express the derivatives $\dot{T}, \dot{N}, \dot{B}$ *in terms of* T, N, B. Since $T = \dot{\boldsymbol{\gamma}}$, we have $\dot{T} = \ddot{\boldsymbol{\gamma}} = \kappa N$. Next, consider $\dot{B}$. We differentiate the equation $B \cdot B = 1$ and obtain that $\dot{B} \cdot B = 0$. So $\dot{B}$ is orthogonal to B. Differentiating $B \cdot T = 0$ gives us $\dot{B} \cdot T + B \cdot \dot{T} = 0$. Thus

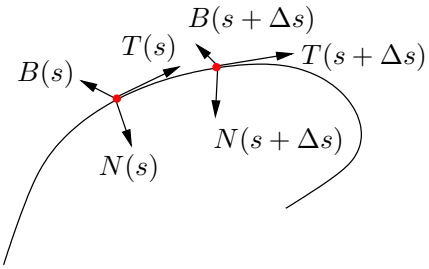
$$\dot{B} \cdot T = -B \cdot \dot{T} = -B \cdot \kappa N = 0.$$

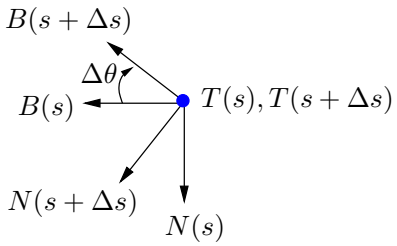
So $\dot{B}$ is orthogonal to T in addition to B. Therefore $\dot{B}$ must be a scalar multiple of N.

We can now define the *torsion* function τ of the curve $\boldsymbol{\gamma}$ to be the real-valued function on the curve domain I such that

$$\dot{B} = -\tau N. \tag{1}$$

The minus sign in the definition is traditional.





As curvature describes the rate of rotation of the tangent on a plane curve, torsion describes the rate of rotation of the binormal on a space curve. The above figure explains this. Consider traveling a distance Δs from the point $\boldsymbol{\gamma}(s)$ along the curve to reach a point $\boldsymbol{\gamma}(s + \Delta s)$. The two points have different Frenet frames attached to them. To isolate the effect due to binormal rotation from other geometric variations, imagine superposing the two points and aligning the two tangents $T(s)$ and $T(s + \Delta s)$ at them. Let these two vectors point inward, as shown on the right in the figure. Then the two principal normals $N(s)$ and $N(s + \Delta s)$, and the two binormals $B(s)$ and $B(s + \Delta s)$ are in a plane orthogonal to the two tangent vectors. Let $\Delta\theta$ be the angle of rotation from $B(s)$ to $B(s + \Delta s)$. Then we have

$$(B(s + \Delta s) - B(s)) \cdot N(s) \approx \Delta\theta \cdot \|B(s)\| = \Delta\theta,$$

and consequently,

$$\lim_{\Delta s \to 0} \frac{B(s + \Delta s) - B(s)}{\Delta s} \cdot N(s) = \lim_{\Delta s \to 0} \frac{\Delta\theta}{\Delta s}.$$

This just means that

$$\frac{d\theta}{ds} = \dot{B}(s) \cdot N(s) = -\tau.$$

Hence $-\tau$ measures the rate of binormal rotation.

EXAMPLE 2. Every plane curve $\boldsymbol{\gamma}(s) = (x(s), y(s))$ has zero torsion. This is because both T and N remains in the xy-plane, so $B = (0, 0, 1)$ and $\dot{B}(s) = \mathbf{0}$. By comparing this with (1) we obtain that $\tau = 0$.

We shall presently show that τ does measure the torsion, or twisting, of the curve $\boldsymbol{\gamma}$.

Theorem 4 (Frenet formulas) *If $\boldsymbol{\gamma} : I \to \mathbb{R}^3$ is a unit-speed curve with curvature $\kappa > 0$ and torsion τ, then*

$$\begin{aligned} \dot{T} &= \kappa N, \\ \dot{N} &= -\kappa T + \tau B, \\ \dot{B} &= -\tau N. \end{aligned}$$

Proof As we saw above, the first and third formulas are just the definitions of curvature and torsion. To prove the second, we express $\dot{N}$ in terms of T, N, B using orthonormal expansion:

$$\dot{N} = (\dot{N} \cdot T)T + (\dot{N} \cdot N)N + (\dot{N} \cdot B)B.$$

These coefficients are easily found. Differentiating $N \cdot T = 0$, we get $\dot{N} \cdot T + N \cdot \dot{T} = 0$; hence

$$\dot{N} \cdot T = -N \cdot \dot{T} = -N \cdot \kappa N = -\kappa.$$

That $N \cdot \dot{N} = 0$ follows from that N is a unit vector. Finally,

$$\dot{N} \cdot B = -N \cdot \dot{B} = -N(-\tau N) = \tau.$$

□

Notice that the matrix

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

which expresses $\dot{T}$, $\dot{N}$, and $\dot{B}$ in terms of T, N, and B is *skew symmetric* or *antisymmetric*, i.e., it is equal to the negative of its transpose.

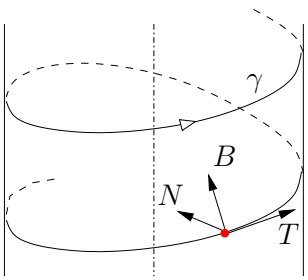
EXAMPLE 3. Let us compute the Frenet frame, the curvature and torsion of the unit-speed helix

$$\boldsymbol{\gamma}(s) = \left(a\cos\frac{s}{c}, a\sin\frac{s}{c}, \frac{bs}{c} \right),$$

where $c = \sqrt{a^2 + b^2}$ and $a > 0$.

Immediately, we obtain

$$T(s) = \dot{\boldsymbol{\gamma}}(s) = \left(-\frac{a}{c}\sin\frac{s}{c}, \frac{a}{c}\cos\frac{s}{c}, \frac{b}{c} \right).$$



Hence,

$$\dot{T}(s) = \left(-\frac{a}{c^2}\cos\frac{s}{c}, -\frac{a}{c^2}\sin\frac{s}{c}, 0\right).$$

Thus

$$\kappa(s) = \|\dot{T}(s)\| = \frac{a}{c^2} = \frac{a}{a^2+b^2} > 0.$$

From $\dot{T} = \kappa N$, we get

$$N(s) = \left(-\cos\frac{s}{c}, -\sin\frac{s}{c}, 0\right).$$

Note that regardless of what values a and b have, N always points straight toward the axis of the cylinder on which $\boldsymbol{\gamma}$ lies.

Next, we obtain

$$B(s) = T(s) \times N(s) = \left(\frac{b}{c}\sin\frac{s}{c}, -\frac{b}{c}\cos\frac{s}{c}, \frac{a}{c}\right).$$

To compute torsion, we calculate the derivative

$$\dot{B}(s) = \left(\frac{b}{c^2}\cos\frac{s}{c}, \frac{b}{c^2}\sin\frac{s}{c}, 0\right).$$

By definition, $\dot{B} = -\tau N$. Comparing the formulae for $\dot{B}$ and N, we conclude that

$$\tau(s) = \frac{b}{c^2} = \frac{b}{a^2+b^2}.$$

So the torsion of the helix is also constant.

Note that when the parameter b is zero, the helix reduces to a circle of radius a. The curvature of this circle is $\kappa = \frac{1}{a}$, and the torsion is identically zero.

Example 3 is a very special one. In general, neither the curvature nor the torsion functions of a curve need to be constant.

3 Approximation of Space Curves

We can now give an informative approximation of a given curve near an arbitrary point on the curve. The goal is to show how curvature and torsion influence the shape of the curve — and express this in terms of the Frenet frame at the selected point.

For simplicity, we shall consider the unit-speed curve $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)$ near the point $\boldsymbol{\beta}(0)$. For s small, each coordinate $\beta_i(s)$ is closely approximated by the initial terms of its Taylor series:

$$\beta_i(s) \approx \beta_i(0) + \frac{d\beta_i}{ds}(0)s + \frac{d^2\beta_i}{ds^2}(0)\frac{s^2}{2} + \frac{d^3\beta_i}{ds^3}(0)\frac{s^3}{6}.$$

Hence

$$\boldsymbol{\beta}(s) \approx \boldsymbol{\beta}(0) + s\dot{\boldsymbol{\beta}}(0) + \frac{s^2}{2}\ddot{\boldsymbol{\beta}}(0) + \frac{s^3}{6}\dddot{\boldsymbol{\beta}}(0).$$

But $\dot{\boldsymbol{\beta}}(0) = T_0$, and $\ddot{\boldsymbol{\beta}}(0) = \kappa_0 N_0$, where the subscript indicates evaluation at $s = 0$, and we assume $\kappa_0 \neq 0$. Now

$$\dddot{\boldsymbol{\beta}} = \frac{d}{ds}(\kappa N) = \frac{d\kappa}{ds}N + \kappa\dot{N}.$$

Thus by the Frenet formula for $\dot{N}$, we get

$$\dddot{\boldsymbol{\beta}}(0) = -\kappa_0^2 T_0 + \frac{d\kappa}{ds}(0) N_0 + \kappa_0 \tau_0 B_0.$$

Finally, substitute these derivatives into the approximation of $\boldsymbol{\beta}(s)$ given above, and keep only the dominant term in the coefficient of each of T_0, N_0, B_0 (that is, the one containing the smallest power of s), the result is

$$\boldsymbol{\beta}(s) \approx \boldsymbol{\beta}(0) + sT_0 + \kappa_0 \frac{s^2}{2} N_0 + \kappa_0 \tau_0 \frac{s^3}{6} B_0. \tag{2}$$

Denoting the right side by $\hat{\boldsymbol{\beta}}(s)$, we obtain a curve $\hat{\boldsymbol{\beta}}$ called the *Frenet approximation* of $\boldsymbol{\beta}$ near $s = 0$. We emphasize that $\boldsymbol{\beta}$ has a different Frenet approximation near each of its points; if 0 is replaced by an arbitrary number s_0, then s is replaced by $s - s_0$, as usual in Taylor expansions.

Let us now examine the Frenet approximation (2) given above. The first term in the expression for $\hat{\boldsymbol{\beta}}$ is just the point $\boldsymbol{\beta}(0)$. The first two terms give the tangent line $\boldsymbol{\beta}(0) + sT_0$ of $\boldsymbol{\beta}$ at $\boldsymbol{\beta}(0)$ — the best linear approximation of $\boldsymbol{\beta}$ near $\boldsymbol{\beta}(0)$. The first three terms give the parabola

$$\boldsymbol{\beta}(0) + sT_0 + \kappa_0 \frac{s^2}{2} N_0,$$

which is the best quadratic approximation of $\boldsymbol{\beta}$ near $\boldsymbol{\beta}(0)$. Note that this parabola lies in the plane through $\boldsymbol{\beta}(0)$ orthogonal to B_0, the *osculating plane* of $\boldsymbol{\beta}$ at $\boldsymbol{\beta}(0)$, as shown in Figure 1 (from [1, p. 60]). This parabola has the same shape as the parabola $y = \kappa_0 x^2/2$ in the xy plane, and is completely determined by the curvature κ_0 of $\boldsymbol{\beta}$ at $s = 0$.

Finally, the torsion τ_0, which appears in the last and smallest term of $\hat{\boldsymbol{\beta}}$, controls the motion of $\boldsymbol{\beta}$ orthogonal to the osculating plane at $\boldsymbol{\beta}(0)$.

We conclude with a result that a space curve is determined by it curvature and torsion functions up to a rigid motion (which consists of a rotation and a translation).

Theorem 5 (Fundamental Theorem of Space Curves) *Let $\boldsymbol{\beta}(s)$ and $\tilde{\boldsymbol{\beta}}(s)$ be two unit-speed curves in $\mathbb{R}^3$ with the same curvature $\kappa(s) > 0$ and the same torsion $\tau(s)$ for all s. Then there exists a rigid motion that transforms $\boldsymbol{\beta}$ into $\tilde{\boldsymbol{\beta}}$.*

Let $\kappa(s)$ and $\tau(s)$ be continuous functions for $s > 0$ such that $\kappa > 0$. Then there exists exactly one curve in $\mathbb{R}^3$, determined up to a Euclidean motion, which has arc length s, curvature κ, and torsion τ.

References

[1] B. O'Neill. *Elementary Differential Geometry.* Academic Press, Inc., 1966.

[2] A. Pressley. *Elementary Differential Geometry.* Springer-Verlag London, 2001.

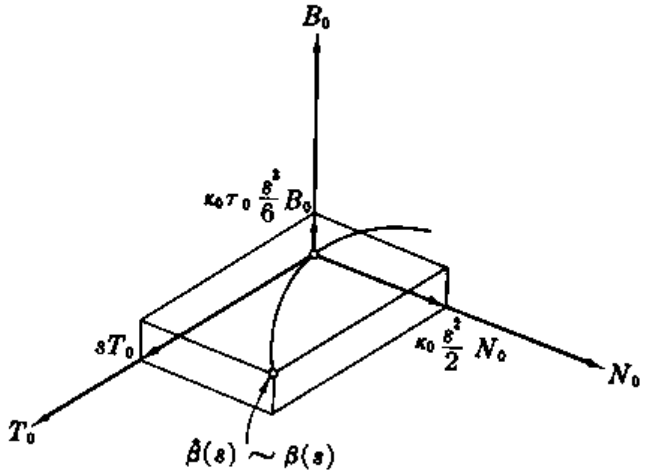


Figure 1: Space curve approximation, Fig. 2.10 on p. 60 of [1].