Thomas Banchoff, Shiing-Shen Chern, and
William Pohl

Differential Geometry of Curves and Surfaces, 1st Edition

SPIN C346 Differential Geometry Banchoff/Chern/Pohl

Monograph – Mathematics –

January 29, 2003

Springer
Berlin Heidelberg New York
Barcelona Hong Kong
London Milan Paris
Tokyo

Part I

Introduction

1 Review of Euclidean Geometry

1.1 Motions

Three-dimensional Euclidean space E consists of points which have as coordinates ordered triples of real numbers x_1, x_2, x_3. In vector notation, we write $\mathbf{x} = (x_1, x_2, x_3)$. The *distance* between two points is given by the formula

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{3} (x_i - y_i)^2} \ . \tag{1.1}$$

Note that $d(\mathbf{x}, \mathbf{y}) \geq 0$ and $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$.

An *affine transformation* T from R^3 to R^3 is defined by $T(\mathbf{x}) = \mathsf{A}\mathbf{x} + \mathbf{b}$, where A is a 3×3 matrix and $\mathbf{b}$ is a vector in R^3. An affine transformation that preserves distance between points, such that $d(T(\mathbf{x}), T(\mathbf{y})) = d(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}$ and $\mathbf{y}$ is called a *motion* of R^3.

Proposition 1. *An affine transformation is a motion if and only if* A *is an orthogonal matrix, i.e. a matrix with columns that are mutually perpendicular unit vectors.*

Proof. Let the points (x_1, x_2, x_3) and (y_1, y_2, y_3) be mapped by an affine transformation T onto the points (x'_1, x'_2, x'_3) and (y'_1, y'_2, y'_3) respectively, so that

$$x'_i = \sum_{j=1}^{3} a_{ij} x_j + b_i$$

$$y'_i = \sum_{j=1}^{3} a_{ij} y_j + b_i \ ,$$

where a_{ij} denotes the entries of the matrix A and the b_i denotes the components of $\mathbf{b}$. If we subtract these two equations, we get

$$x'_i - y'_i = \sum_{j=1}^{3} a_{ij} \left(x_j - y_j \right) \ .$$

Taking the sum of the squares gives us

$$\sum_i (x_i' - y_i')^2 = \sum_{i,j,k} a_{ij} a_{ik} (x_j - y_j)(x_k - y_k) \ ,$$

where all the indices run from 1 to 3. This equality will only hold true if

$$\sum_{i=1}^{3} a_{ij} a_{ik} = \delta_{jk} \ , \tag{1.2}$$

where $\delta_{jk} = 1$ if $j = k$ and 0 otherwise.

Given an affine transformation $x_i' = \sum_{j=1}^{3} a_{ij} x_j + b_i$, we can solve explicitly for the x_i in terms of the x_i'. We first set $x_i' - b_i = \sum_{j=1}^{3} a_{ij} x_j$ then multiply by a_{ji} and sum over i and j to get

$$\sum_{i=1}^{3} a_{ji}(x_i' - b_i) = x_j \ .$$

Remark 1. The quantities δ_{jk} defined in (1.2) are called *Kronecker deltas*. We have illustrated their usefulness in the above proof, and they will be used consistently.

It will be convenient to introduce the matrices

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \mathsf{A}^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \ .$$

The second matrix A^T is obtained from the first by interchanging the rows and columns and is called the *transpose* of A. Using this notation, (1.2) can be re-written as

$$\mathsf{A}\mathsf{A}^T = I \tag{1.3}$$

where I denotes the unit matrix (δ_{ij}). A matrix A with this property is called *orthogonal*.

We may rewrite the definition of a motion in terms of matrices as $\mathbf{x}' = \mathsf{A}\mathbf{x} + \mathbf{b}$ where $\mathbf{x}$ is the column matrix with entries x_i. We may then solve explicitly for $\mathbf{x}$ in terms of $\mathbf{x}'$ by writing $\mathbf{x}' - \mathbf{b} = \mathsf{A}\mathbf{x}$, so

$$\begin{aligned} \mathsf{A}^T(\mathbf{x}' - \mathbf{b}) &= \mathsf{A}^T(\mathsf{A}\mathbf{x}) \\ &= (\mathsf{A}^T\mathsf{A})\mathbf{x} \\ &= I\mathbf{x} = \mathbf{x}. \end{aligned}$$

A basic result in linear algebra states that, for square matrices,

$$(\mathsf{AC})^T = \mathsf{C}^T \mathsf{A}^T \ ,$$

where the order of the multiplication is important. From this it follows that if A and C are orthogonal matrices, then

$$(\mathsf{AC})^T \mathsf{AC} = \mathsf{C}^T \mathsf{A}^T \mathsf{AC} = \mathsf{C}^T \mathsf{C} = I$$

so the product AC is also an orthogonal matrix.

For an orthogonal matrix A, we have $\mathsf{A}^{-1} = \mathsf{A}^T$. Moreover, $\mathsf{A}^T(\mathsf{A}^{-1})^T = (\mathsf{A}^{-1}\mathsf{A})^T = I^T = I$ so $(\mathsf{A}^{-1})^T = (\mathsf{A}^T)^{-1}$. It follows that if A is orthogonal, then so is A^{-1}. A collection of matrices that is closed under multiplication such that the inverse of every element of the collection is also in the collection is called a *group*.

The succesive application of two motions, as in (1.3) above, is called their *product*. This multiplication is in general *not* commutative. It is easily seen that all of the motions in E form a group under this multiplication, called the *group of motions*. *Euclidean geometry* studies the properties of E that are invariant under the group of motions.

From (1.3) we find $(\det \mathsf{A})^2 = 1$ so that $\det \mathsf{A} = \pm 1$. The motion is called *proper* if the determinant is $+1$, and *improper* if it is -1. It is easily verified that the product of two proper motions is a proper motion, and it is a simple result that all proper motions form a *subgroup* of the group of motions.

Example 1. The mirror reflection, $(x_1, x_2, x_3) \to (-x_1, x_2, x_3)$, is an improper motion.

A motion of the form $x_i' = x_i + b_i$, for $i = 1, 2, 3$ is called a *translation*. A motion of the form

$$x_i' = \sum_{i=1}^{3} a_{ij} x_j$$

where $j = 1, 2, 3$ is called an *orthogonal transformation*. In matrix form, a translation can be written $\mathbf{x}' = \mathbf{x} + \mathbf{b}$ and an orthogonal transformation can be written $\mathbf{x}' = \mathsf{A}\mathbf{x}$. An orthogonal tranformation is called *proper* or *improper* according to the sign of $\det \mathsf{A}$. A proper orthogonal transformation can also be called a *reflection*. All of the translations form a group, as do all of the othogonal transformations and all the rotations. The group of all rotations can be characterized as the subgroup of all proper motions with the origin fixed.

Exercise 1. Prove that the quadratic polynomial

$$\sum_{i,j} \alpha_{ij} \xi_i \xi_j$$

where $\alpha_{ij} = \alpha_{ji}$ is zero for all ξ_i if and only if $\alpha_{ij} = 0$. Show that this is not true without the symmetry condition on the coefficients α_{ij}. This result is used in the proof of (1).

Exercise 2. Show that the inverse motion of $\mathbf{x}' = \mathsf{A}\mathbf{x} + \mathbf{b}$ is $\mathbf{x}' = \mathsf{A}^{-1}\mathbf{x} - \mathsf{A}^{-1}\mathbf{b}$. Let T_i be the motions $\mathbf{x}' = \mathsf{A}_i\mathbf{x} + \mathbf{b}_i$ where $i = 1, 2$ and show that T_1T_2 is the motion $\mathbf{x}' = \mathsf{A}_2\mathsf{A}_1\mathbf{x} + \mathsf{A}_2\mathbf{b}_1 + \mathbf{b}_2$.

Exercise 3. Let G be a group with the elements $\{e, a, b, \dots\}$, where e is the *unit element*. The *left* and *right inverses* of an element a are defined by

$$a_l^{-1}a = e, \text{ and } aa_r^{-1} = e,$$

respectively. Prove that $a_l^{-1} = a_r^{-1}$. Observe that the equivalence of the conditions $\mathsf{A}\mathsf{A}^T = I$ and $\mathsf{A}^T\mathsf{A} = I$ means group-theoretically that the matrix A has the same right and left inverse, which is A^T.

Exercise 4. Prove that the translations form a *normal subgroup* of the group of motions, while the rotations do not.

Exercise 5. Show that the helicoidal motions

$$\begin{aligned} x_1' &= \cos(t)x_1 + \sin(t)x_2 \\ x_2' &= -\sin(t)x_1 + \cos(t)x_2 \\ x_3' &= x_3 + bt, \end{aligned}$$

where b is a constant and t is a parameter, form a group. Draw the orbit of the point $(a, 0, 0)$, and distinguish the cases when $b < 0$ and $b > 0$.

Exercise 6. Prove that the rotation

$$\mathbf{x}_i' = \sum_j a_{ij}x_j$$

where $i, j = 1, 2, 3$ and $\det(a_{ij}) = 1$ has a line of fixed points through the origin, the axis of rotation. Hence prove that the group of rotations is *connected*. Prove also that the group of orthogonal transformations is not connected. (Note: A subgroup of motions is connected if any two points can be joined by a continuous arc.)

1.2 Vectors

Two ordered pairs of points, $p(x_1, x_2, x_3)$, $q(y_1, y_2, y_3)$ and $p'(x_1', x_2', x_3')$, $q'(y_1', y_2', y_3')$ are called equivalent if there is a translation T which maps p to p' and q to q'. The last property can be expressed by the conditions $x_i' = x_i + b_i$ and $y_i' = y_i + b_i$, where $i = 1, 2, 3$. It follows that a necessary and sufficient condition for the equivalence of the two ordered pairs of points is $y_i' - x_i' = y_i - x_i$. Such an equivalence class is called a *vector*. We denote the vector by $V = \overrightarrow{pq}$ and call

$$v_i = y_i - x_i, \tag{1.4}$$

where $i = 1, 2, 3$ denote its *components*. A vector is therefore completely determined by its components. Geometrically $\overrightarrow{pq} = \overrightarrow{p'q'}$ if and only if the segments $\overline{pq}$ and $\overline{p'q'}$ are of the same length and parallel in the same sense.

Using the origin $\mathbf{O}$ of our coordinate system, we can set up a one-to-one correspondence between the points p of R^3 and the vectors $\overrightarrow{\mathbf{O}p}$. The latter will be called the *position vector* of p. Notice that it is defined with reference to the origin $\mathbf{O}$.

Given two vectors $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{w} = (w_1, w_2, w_3)$ their *sum* is $\mathbf{v}+\mathbf{w} = (v_1+w_1, v_2+w_2, v_3+w_3)$, and multiplication of a vector by a number is defined by $\lambda\mathbf{v} = (\lambda v_1, \lambda v_2, \lambda v_3)$. Throughout this book real numbers are sometimes called *scalars*, in order to emphasize their difference from vectors.

By (1.4) we see that under a motion the vectors are transformed according to the equations

$$v'_i = \sum_j \mathsf{A}_{ij}\mathbf{v}_j \tag{1.5}$$

where $i, j = 1, 2, 3$, and $\mathbf{v}' = (v'_1, v'_2, v'_3)$ is the image vector. Using (1.5) and (1) we get

$$v_1'^2 + v_2'^2 v_3'^2 = v_1^2 + v_2^2 + v_3^2 \ .$$

This leads us to define

$$\mathbf{v}^2 = v_1^2 + v_2^2 + v_3^2 \ . \tag{1.6}$$

We call $+\sqrt{\mathbf{v}^2}$ the *length* of $\mathbf{v}$. Thus, the length of a vector is invariant under motions.

More generally, we find

$$\tfrac{1}{2}\{(\mathbf{v}+\mathbf{w})^2 - \mathbf{v}^2 - \mathbf{w}^2\} = \textstyle\sum_i v_i w_i \ ,$$

where $i = 1, 2, 3$. Since the left-hand side involves only lengths of vectors, which are invariant under motions, the same property holds for the expressions at the right-hand side. Generalizing (1.6), we introduce the notation

$$\mathbf{v}\cdot\mathbf{w} = \mathbf{w}\cdot\mathbf{v} = \sum_i v_i w_i \ ,$$

which is called the *scalar* or *dot product* of $\mathbf{v}$ and $\mathbf{w}$.

Relative to addition and scalar multiplication of vectors, the scalar product has the following properties: $(\mathbf{v}_1 + \mathbf{v}_2)\cdot\mathbf{w} = \mathbf{v}_1\cdot\mathbf{w} + \mathbf{v}_2\cdot\mathbf{w}$ and $(\lambda\mathbf{v})\cdot\mathbf{w} = \mathbf{v}\cdot(\lambda\mathbf{w}) = \lambda(\mathbf{v}\cdot\mathbf{w})$ where λ = scalar. The relation

$$(\mathbf{v}+\lambda\mathbf{w})^2 = \mathbf{v}^2 + 2\lambda\mathbf{v}\cdot\mathbf{w} + \lambda^2\mathbf{w}^2 \geq 0$$

is true for all λ. So by elementary algebra we get

$$\mathbf{v}^2\mathbf{w}^2 - (\mathbf{v}\cdot\mathbf{w})^2 \geq 0, \tag{1.7}$$

which is called the *Cauchy-Schwartz inequality*. The angle θ between $\mathbf{v}$ and $\mathbf{w}$ is defined by

$$\cos\theta = \frac{\mathbf{v}\cdot\mathbf{w}}{\sqrt{\mathbf{v}^2\mathbf{w}^2}} \ .$$

This is meaningful because by the Cauchy-Schwartz inequality the right-hand side has absolute value ≤ 1. The vectors $\mathbf{v}$ and $\mathbf{w}$ are *perpendicular* or *orthogonal* if $\mathbf{v}\cdot\mathbf{w} = 0$.

The *determinant* of three vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ with components u_i, v_i, w_i for $i = 1, 2, 3$ respectively, is defined by

$$\det(\mathbf{u},\mathbf{v},\mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \ .$$

Under the transformation in (1.5) the determinant $(\mathbf{u},\mathbf{v},\mathbf{w})$ is multiplied by $\det \mathsf{A} = \det(a_{ij})$. Hence it is invariant under proper motions and changes its sign under improper motions. The following properties follow immediately from the definition:

$$\begin{aligned} &(\mathbf{u}+\mathbf{u}_1,\mathbf{v},\mathbf{w}) = (\mathbf{u},\mathbf{v},\mathbf{w}) + (\mathbf{u}_1,\mathbf{v},\mathbf{w}), \\ &(\lambda\mathbf{u},\mathbf{v},\mathbf{w}) = \lambda(\mathbf{u},\mathbf{v},\mathbf{w}) \quad \text{where } \lambda = \text{scalar}, \\ &(\mathbf{u},\mathbf{v},\mathbf{w}) = -(\mathbf{v},\mathbf{u},\mathbf{w}) \;\; \text{etc.} \end{aligned}$$

The *vector product* of two vectors $\mathbf{v}, \mathbf{w}$ is the vector $\mathbf{z}$ such that the relation

$$(\mathbf{v},\mathbf{w},\mathbf{x}) = \mathbf{z}\cdot\mathbf{x} \tag{1.8}$$

holds for all vectors $\mathbf{x}$. It follows that $\mathbf{z}$ has the components

$$z_1 = v_2w_3 - v_3w_2,\; z_2 = v_3w_1 - v_1w_3,\; z_3 = v_1w_2 - v_2w_1 \ . \tag{1.9}$$

We write

$$\mathbf{z} = \mathbf{v}\times\mathbf{w} \ .$$

The vector product has the following properties:

$$\begin{aligned} &\mathbf{v}\times\mathbf{w} + \mathbf{w}\times\mathbf{v} = 0, \\ &(\mathbf{v}_1+\mathbf{v}_2)\times\mathbf{w} = \mathbf{v}_1\times\mathbf{w} + \mathbf{v}_2\times\mathbf{w}, \\ &(\lambda\mathbf{v})\times\mathbf{w} = \lambda(\mathbf{v}\times\mathbf{w}) \quad \text{where } \lambda = \text{scalar}. \end{aligned}$$

The vector $\mathbf{z}$ can be described geometrically as follows: We see from (1.9) that $\mathbf{v}\times\mathbf{w} = 0$ if and only if one of the vectors $\mathbf{v}, \mathbf{w}$ is a multiple of the other. In this case we say that $\mathbf{v}$ and $\mathbf{w}$ are *linearly dependent*. Suppose next that $\mathbf{v} \times \mathbf{w} \neq 0$, i.e., $\mathbf{v}$ and $\mathbf{w}$ are linearly independent. Putting $\mathbf{x} = \mathbf{v}, \mathbf{w}$ in (1.8), we get $\mathbf{z} \cdot \mathbf{v} = \mathbf{z} \cdot \mathbf{w} = 0$, so that $\mathbf{z}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$. We write $\mathbf{z} = \lambda \mathbf{u}$, where $\mathbf{u}$ is a unit vector orthogonal to $\mathbf{v}$ and $\mathbf{w}$. Thus from (1.8) we get

$$(\mathbf{v}, \mathbf{w}, \mathbf{u}) = \mathbf{z} \cdot \mathbf{u} = \lambda \mathbf{u}^2 = \lambda \; .$$

Hence we have

$$\mathbf{z} = \mathbf{v} \times \mathbf{w} = (\mathbf{v}, \mathbf{w}, \mathbf{u}) \quad \mathbf{u} \neq 0 \; .$$

The unit vector $\mathbf{u}$ is defined up to its sign and we can choose $\mathbf{u}$ such that $(\mathbf{v}, \mathbf{w}, \mathbf{u}) > 0$. $\mathbf{z}$ is therefore a multiple of $\mathbf{u}$ and of length $(\mathbf{v}, \mathbf{w}, \mathbf{u})$. This completely determines $\mathbf{z}$.

Three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are called *linearly dependent* if $(\mathbf{u}, \mathbf{v}, \mathbf{w}) = 0$; otherwise they are *linearly independent*. An ordered set of three linearly independent vectors is called a *right-handed* or *left-handed frame* according to the sign of its determinant. The property of right-handedness or left-handedness of a frame remains unchanged under proper motion, while they interchange under an improper motion. Also a right-handed (or left-handed) frame becomes left-handed (or right-handed) when any two of its vectors are interchanged.

Remark 2. The importance of vectors in analytic geometry is due to the algebraic structure. Two vectors can be added and a vector can be multiplied by a scalar. It is important to observe that corresponding operations are meaningless on the points of R^3, because they are not invariant under motions.

Exercise 7. a) The vector equation of a line is $\mathbf{x}(t) = at + b$, where a, b $= constant and \mathrm{a} \neq 0$. Find its angles with the coordinate axes.

b) The vector equation of a plane is $\mathbf{a} \cdot \mathbf{x} = b$ where $\mathbf{a} \neq 0$. Give the geometrical meaning of b when $\mathbf{a}$ is a unit vector.

c) The vector equation of a sphere is $(\mathbf{x} - \mathbf{a})^2 = r^2$. What are $\mathbf{a}$ and r?

In each case draw the relevant figure.

Exercise 8. Prove that $\mathbf{x}(t) = \mathsf{A} \cos t + \mathbf{b} \sin t$ for $\mathsf{A}, \mathbf{b} \neq 0$, represents an ellipse.

Exercise 9. Let $\mathbf{y}_i$, for $i = 1, 2, 3$, be three linearly independent vectors. Prove that any vector $\mathbf{x}$ can be written

$$\mathbf{x} = \sum_i \lambda_i \mathbf{y}_i \; .$$

Hence prove that $\mathbf{x} = 0$ if and only if $\mathbf{x} \cdot \mathbf{y}_i = 0$.

Exercise 10. Let $(\mathbf{u}, \mathbf{v}, \mathbf{w}) = 0$, $\mathbf{u} \times \mathbf{v} \neq 0$. Prove that $\mathbf{w}$ is a *linear combination* of $\mathbf{u}$ and $\mathbf{v}$, i.e., $\mathbf{w}$ can be written $\mathbf{w} = \lambda \mathbf{u} + \mu \mathbf{v}$, where λ, μ are scalars.

Exercise 11. Prove Lagrange's identity:

$$(\mathbf{v} \times \mathbf{w}) \cdot (\mathbf{x} \times \mathbf{y}) = (\mathbf{v} \cdot \mathbf{x})(\mathbf{w} \cdot \mathbf{y}) - (\mathbf{v} \cdot \mathbf{y})(\mathbf{w} \cdot \mathbf{x}) \,.$$

Hence prove that

$$(\mathbf{v} \times \mathbf{w}) \times \mathbf{x} = (\mathbf{v}\mathbf{x})\mathbf{w} - (\mathbf{w}\mathbf{x})\mathbf{v} \,.$$

Hint. To prove the first equation, write out both sides in components.

2 Curves

2.1 Arc Length

A *parametrized curve* in Euclidean three-space e^3 is given by a vector function

$$\mathbf{x}(t) = (x_1(t), x_2(t), x_3(t))$$

that assigns a vector to every value of a parameter t in a domain interval $[a, b]$. The *coordinate functions* of the curve are the functions $x_i(t)$. In order to apply the methods of calculus, we suppose the functions $x_i(t)$ to have as many continuous derivatives as needed in the following treatment.

For a curve $\mathbf{x}(t)$, we define the *first derivative* $\mathbf{x}'(t)$ to be the limit of the secant vector from $\mathbf{x}(t)$ to $\mathbf{x}(t+h)$ divided by h as h approaches 0, assuming that this limit exists. Thus,

$$\mathbf{x}'(t) = \lim_{h \to 0} \left(\frac{\mathbf{x}(t+h) - \mathbf{x}(t)}{h} \right) .$$

The first derivative vector $\mathbf{x}'(t)$ is tangent to the curve at $\mathbf{x}(t)$. If we think of the parameter t as representing time and we think of $\mathbf{x}(t)$ as representing the position of a moving particle at time t, then $\mathbf{x}'(t)$ represents the *velocity* of the particle at time t. It is straightforward to show that the coordinates of the first derivative vector are the derivatives of the coordinate functions, i.e.

$$\mathbf{x}'(t) = ({x_1}'(t), {x_2}'(t), {x_3}'(t)) .$$

For most of the curves we will be concerning ourselves with, we will make the "genericity assumption" that $\mathbf{x}'(t)$ is non-zero for all t. (MISSING SECTIONS) lengths of polygons inscribed in $\mathbf{x}$ as the lengths of the sides of these polygons tend to zero. By the fundamental theorem of calculus, this limit can be expressed as the integral of the *speed* $s'(t) = |\mathbf{x}'(t)|$ between the parameters of the end-points of the curve, a and b. That is,

$$s(b) - s(a) = \int_a^b |\mathbf{x}'(t)| \; dt = \int_a^b \sqrt{\sum_{i=1}^{3} x_i'(t)^2} \; dt .$$

For an arbitrary value $t \in (a, b)$, we may define the *distance function*

$$s(t) - s(a) = \int_a^t |\mathbf{x}'(u)|\ du\ ,$$

which gives us the distance from a to t along the curve.

Notice that this definition of arc length is independent of the parametrization of the curve. If we define a function $v(t)$ from the interval $[a, b]$ to itself such that $v(a) = a$, $v(b) = b$ and $v'(t) > 0$, then we may use the change of variables formula to express the arc length in terms of the new parameter v:

$$\int_a^b |\mathbf{x}'(t)|\ dt = \int_v (a) = a^v(b) = b\,|\mathbf{x}'(v(t))|\,v'(t)\ dt = \int_a^b |\mathbf{x}'(v)|\ dv\ .$$

We can also write this expression in the form of differentials:

$$ds = |\mathbf{x}'(t)|\,dt = |\mathbf{x}'(v)|\,dv.$$

This differential formalism becomes very significant, especially when we use it to study surfaces and higher dimensional objects, so we will reinterpret results that use integration or differentiation in differential notation as we go along. For example, the statement $s'(t) = \sqrt{\sum_{i=1}^3 x_i'(t)^2}$ can be rewritten as

$$\left(\frac{ds}{dt}\right)^2 = \sum_{i=1}^3 \left(\frac{dx_i}{dt}\right)^2\ ,$$

and this may be expressed in the form

$$ds^2 = \sum_{i=1}^3 dx_i^2\ ,$$

which has the advantage that it is independent of the parameter used to describe the curve. ds is called the *element of arc*. It can be visualized as the distance between two neighboring points.

One of the most useful ways to parametrize a curve is by the arc length s itself. If we let $s = s(t)$, then we have

$$s'(t) = |\mathbf{x}'(t)| = |\mathbf{x}'(s)|\,s'(t)\ ,$$

from which it follows that $|\mathbf{x}'(s)| = 1$ for all s. So the derivative of $\mathbf{x}$ with respect to arc length is always a unit vector.

This parameter s is defined up to the transformation $s \to \pm s + c$, where c is a constant. Geometrically, this means the freedom in the choice of initial point and direction in which to traverse the curve in measuring the arc length.

Exercise 12. One of the most important space curves is the *circular helix*

$$\mathbf{x}(t) = (a\cos t, a\sin t, bt)\ ,$$

where $a \neq 0$ and b are constants. Find the length of this curve over the interval $[0, 2\pi]$.

Exercise 13. Find a constant c such that the helix

$$\mathbf{x}(t) = (a\cos(ct), a\sin(ct), bt)$$

is parametrized by arclength, so that $|\mathbf{x}'(t)| = 1$ for all t.

Exercise 14. The *astroid* is the curve defined by

$$\mathbf{x}(t) = \left(a\cos^3 t, a\sin^3 t, 0\right) ,$$

on the domain $[0, 2\pi]$. Find the points at which $\mathbf{x}(t)$ does not define an immersion, i.e., the points for which $\mathbf{x}'(t) = 0$.

Exercise 15. The trefoil curve is defined by

$$\mathbf{x}(t) = ((a + b\cos(3t))\cos(2t), (a + b\cos(3t))\sin(2t), b\sin(3t)) ,$$

where a and b are constants with $a > b > 0$ and $0 \leq t \leq 2\pi$. Sketch this curve, and give an argument to show why it is knotted, i.e. why it cannot be deformed into a circle without intersecting itself in the process.

Exercise 16. (For the serious mathematician) Two parametrized curves $\mathbf{x}(t)$ and $\mathbf{y}(u)$ are said to be equivalent if there is a function $u(t)$ such that $u'(t) > 0$ for all $a < t < b$ and such that $\mathbf{y}(u(t)) = \mathbf{x}(t)$. Show that relation satisfies the following three properties:

1. Every curve $\mathbf{x}$ is equivalent to itself
2. If $\mathbf{x}$ is equivalent to $\mathbf{y}$, then $\mathbf{y}$ is equivalent to $\mathbf{x}$
3. If $\mathbf{x}$ is equivalent to $\mathbf{y}$ and if $\mathbf{y}$ is equivalent to $\mathbf{z}$, then $\mathbf{x}$ is equivalent to $\mathbf{z}$

A relation that satisfies these properties is called an equivalence relation. Precisely speaking, a curve is considered be an equivalence class of parametrized curves.

2.2 Curvature and Fenchel's Theorem

If $\mathbf{x}$ is an immersed curve, with $\mathbf{x}'(t) \neq 0$ for all t in the domain, then we may define the *unit tangent vector* $\mathbf{T}(t)$ to be $\frac{\mathbf{x}'(t)}{|\mathbf{x}'(t)|}$. If the parameter is arclength, then the unit tangent vector $\mathbf{T}(s)$ is given simply by $\mathbf{x}'(s)$. The line through $\mathbf{x}(t_0)$ in the direction of $\mathbf{T}(t_0)$ is called the *tangent line* at $\mathbf{x}(t_0)$. We can write this line as $\mathbf{y}(u) = \mathbf{x}(t_0) + u\mathbf{T}(t_0)$, where u is a parameter that can take on all real values.

Since $\mathbf{T}(t) \cdot \mathbf{T}(t) = 1$ for all t, we can differentiate both sides of this expression, and we obtain $2\mathbf{T}'(t) \cdot \mathbf{T}(t) = 0$. Therefore $\mathbf{T}'(t)$ is orthogonal to $\mathbf{T}(t)$. The *curvature* of the space curve $\mathbf{x}(t)$ is defined by the condition $\kappa(t) = \frac{|\mathbf{T}'(t)|}{|\mathbf{x}'(t)|}$, so $= \kappa(t)s'(t) = |\mathbf{T}'(t)|$. If the parameter is arclength, then $\mathbf{x}'(s) = \mathbf{T}(s)$ and $\kappa(s) = |\mathbf{T}'(s)| = |\mathbf{x}\text{"}(s)|$.

Proposition 2. *If $\kappa(t) = 0$ for all t, then the curve lies along a straight line.*

Proof. Since $\kappa(t) = 0$, we have $\mathbf{T}'(t) = 0$ and $\mathbf{T}(t) = \mathbf{a}$, a constant unit vector. Then $\mathbf{x}'(t) = s'(t)\mathbf{T}(t) = s'(t)\mathbf{a}$, so by integrating both sides of the equation, we obtain $\mathbf{x}(t) = s(t)\mathbf{a} + b$ for some constant b. Thus $\mathbf{x}(t)$ lies on the line through b in the direction of $\mathbf{a}$.

Curvature is one of the simplest and at the same time one of the most important properties of a curve. We may obtain insight into curvature by considering the *second derivative* vector $\mathbf{x}"(t)$, often called the *acceleration* vector when we think of $\mathbf{x}(t)$ as representing the path of a particle at time t. If the curve is parametrized by arclength, then $\mathbf{x}'(s)\cdot\mathbf{x}'(s) = 1$ so $\mathbf{x}"(s)\cdot\mathbf{x}'(t) = 0$ and $\kappa(s) = |\mathbf{x}"(s)|$. For a general parameter t, we have $\mathbf{x}'(t) = s'(t)\mathbf{T}(t)$ so $\mathbf{x}"(t) = s"(t)\mathbf{T}(t) + s'(t)\mathbf{T}'(t)$. If we take the cross product of both sides with $\mathbf{x}'(t)$ then the first term on the right is zero since $\mathbf{x}'(t)$ is parallel to $\mathbf{T}(t)$. Moreover $\mathbf{x}'(t)$ is perpendicular to $\mathbf{T}'(t)$ so

$$|\mathbf{T}'(t) \times \mathbf{x}'(t)| = |\mathbf{T}'(t)||\mathbf{x}'(t)| = s'(t)^2\kappa(t) \ .$$

Thus

$$\mathbf{x}"(t) \times \mathbf{x}'(t) = s'(t)\mathbf{T}'(t) \times \mathbf{x}'(t)$$

and

$$|\mathbf{x}"(t) \times \mathbf{x}'(t)| = s'(t)^3\kappa(t) \ .$$

This gives a convenient way of finding the curvature when the curve is defined with respect to an arbitrary parameter. We can write this simply as

$$\kappa(t) = \frac{|\mathbf{x}"(t) \times \mathbf{x}'(t)|}{|\mathbf{x}'(t)\mathbf{x}'(t)|^{3/2}} \ .$$

Note that the curvature $\kappa(t)$ of a space curve is non-negative for all t. The curvature can be zero, for example at every point of a curve lying along a straight line, or at an isolated point like $t = 0$ for the curve $\mathbf{x}(t) = (t, t^3, 0)$. A curve for which $\kappa(t) > 0$ for all t is called *non-inflectional*.

The unit tangent vectors emanating from the origin form a curve $\mathbf{T}(t)$ on the unit sphere called the *tangential indicatrix* of the curve $\mathbf{x}$. To calculate the length of the tangent indicatrix, we form the integral of $|\mathbf{T}'(t)| = \kappa(t)s'(t)$ with respect to t, so the length is $\kappa(t)s'(t)dt = \kappa(s)ds$. This significant integral is called the *total curvature* of the curve $\mathbf{x}$.

Up to this time, we have concentrated primarily on local properties of curves, determined at each point by the nature of the curve in an arbitrarily small neighborhood of the point. We are now in a position to prove our first result in *global differential geometry* or *differential geometry in the large.*

By a closed curve $\mathbf{x}(t)$, $a \leq t \leq b$, we mean a curve such that $\mathbf{x}(b) = \mathbf{x}(a)$. We will assume moreover that the derivative vectors match at the endpoints of the interval, so $\mathbf{x}'(b) = \mathbf{x}'(a)$.

Theorem 1 (Fenchel's Theorem). *The total curvature of a closed space curve* $\mathbf{x}$ *is greater than or equal to* 2π*.*

$$\kappa(s)ds \geq 2\pi$$

The first proof of this result was found independently by B. Segre in 1934 and later independently by H. Rutishauser and H. Samelson in 1948. The following proof depends on a lemma by R. Horn in 1971:

Lemma 1. *Let* $\mathbf{g}$ *be a closed curve on the unit sphere with length* $L < 2$*. Then there is a point* $\mathbf{m}$ *on the sphere that is the north pole of a hemisphere containing* $\mathbf{g}$*.*

To see this, consider two points $\mathbf{p}$ and $\mathbf{q}$ on the curve that break $\mathbf{g}$ up into two pieces $\mathbf{g}_1$ and $\mathbf{g}_2$ of equal length, therefore both less than π. Then the distance from $\mathbf{p}$ to $\mathbf{q}$ along the sphere is less than π so there is a unique minor arc from $\mathbf{p}$ to $\mathbf{q}$. Let $\mathbf{m}$ be the midpoint of this arc. We wish to show that no point of $\mathbf{g}$ hits the equatorial great circle with $\mathbf{m}$ as north pole. If a point on one of the curves, say $\mathbf{g}_1$, hits the equator at a point $\mathbf{r}$, then we may construct another curve $\mathbf{g}_1'$ by rotating $\mathbf{g}_1$ one-half turn about the axis through $\mathbf{m}$, so that $\mathbf{p}$ goes to $\mathbf{q}$ and $\mathbf{q}$ to $\mathbf{p}$ while $\mathbf{r}$ goes to the antipodal point $\mathbf{r}'$. The curve formed by $\mathbf{g}_1$ and $\mathbf{g}_1'$ has the same length as the original curve $\mathbf{g}$, but it contains a pair of antipodal points so it must have length at least 2π, contradicting the hypothesis that the length of $\mathbf{g}$ was less than 2π.

From this lemma, it follows that any curve on the sphere with length less than 2π is contained in a hemisphere centered at a point $\mathbf{m}$. However if $\mathbf{x}(t)$ is a closed curve, we may consider the differentiable function $f(t) = \mathbf{x}(t) \cdot \mathbf{m}$. At the maximum and minimum values of f on the closed curve $\mathbf{x}$, we have

$$0 = f'(t) = \mathbf{x}'(t) \cdot \mathbf{m} = s'(t)\mathbf{T}(t) \cdot \mathbf{m}$$

so there are at least two points on the curve such that the tangential image is perpendicular to $\mathbf{m}$. Therefore the tangential indicatrix of the closed curve $\mathbf{x}$ is not contained in a hemisphere, so by the lemma, the length of any such indicatrix is greater than 2π. Therefore the total curvature of the closed curve $\mathbf{x}$ is also greater than 2π.

Corollary 1. *If, for a closed curve* $\mathbf{x}$*, we have* $\kappa(t) \leq \frac{1}{R}$ *for all* t*, then the curve has length* $L \geq 2\pi R$*.*

Proof.

$$L = \int ds \geq \int R\kappa(s)ds = R \int \kappa(s)ds \geq 2\pi R$$

Fenchel also proved the stronger result that the total curvature of a closed curve equals 2π if and only if the curve is a convex plane curve.

I. Fáry and J. Milnor proved independently that the total curvature must be greater than 4π for any non-self-intersecting space curve that is knotted (not deformable to a circle without self-intersecting during the process.)

Exercise 17. Let $\mathbf{x}$ be a curve with $\mathbf{x}'(t_0) \neq 0$. Show that the tangent line at $\mathbf{x}(t_0)$ can be written as $\mathbf{y}(u) = \mathbf{x}(t_0) + u\mathbf{x}'(t_0)$ where u is a parameter that can take on all real values.

Exercise 18. The plane through a point $\mathbf{x}(t_0)$ perpendicular to the tangent line is called the *normal plane* at the point. Show that a point $\mathbf{y}$ is on the normal plane at $\mathbf{x}(t_0)$ if and only if

$$\mathbf{x}'(t_0) \cdot \mathbf{y} = \mathbf{x}'(t_0) \cdot \mathbf{x}(t_0)$$

Exercise 19. Show that the curvature κ of a circular helix

$$\mathbf{x}(t) = (r\cos(t), r\sin(t), pt)$$

is equal to the constant value $\kappa = \frac{|r|}{r^2+p^2}$. Are there any other curves with constant curvature? Give a plausible argument for your answer.

Exercise 20. Assuming that the level surfaces of two functions $f(x_1, x_2, x_3) = 0$ and $g(x_1, x_2, x_3) = 0$ meet in a curve, find an expression for the tangent vector to the curve at a point in terms of the gradient vectors of f and g (where we assume that these two gradient vectors are linearly independent at any intersection point.) Show that the two level surfaces $x_2 - x_1^2 = 0$ and $x_3x_1 - x_2^2 = 0$ consists of a line and a "twisted cubic" $x_1(t) = t$, $x_2(t) = t^2$, $x_3(t) = t^3$. What is the line?

Exercise 21. What is the geometric meaning of the function $f(t) = \mathbf{x}(t) \cdot \mathbf{m}$ used in the proof of Fenchel's theorem?

Exercise 22. Let $\mathbf{m}$ be a unit vector and let $\mathbf{x}$ be a space curve. Show that the projection of this curve into the plane perpendicular to $\mathbf{m}$ is given by

$$\mathbf{y}(t) = \mathbf{x}(t) - (\mathbf{x}(t) \cdot \mathbf{m})\mathbf{m} \,.$$

Under what conditions will there be a t_0 with $\mathbf{y}'(t_0) = 0$?

2.3 The Unit Normal Bundle and Total Twist

Consider a curve $\mathbf{x}(t)$ with $\mathbf{x}'(t) \neq 0$ for all t. A vector $\mathbf{z}$ perpendicular to the tangent vector $\mathbf{x}'(t_0)$ at $\mathbf{x}(t_0)$ is called a *normal vector* at $\mathbf{x}(t_0)$. Such a vector is characterized by the condition $\mathbf{z} \cdot \mathbf{x}(t_0) = 0$, and if $|\mathbf{z}| = 1$, then $\mathbf{z}$ is said to be a *unit normal vector* at $\mathbf{x}(t_0)$. The set of unit normal vectors at a point $\mathbf{x}(t_0)$ forms a great circle on the unit sphere. The *unit normal bundle* is the collection of all unit normal vectors at $\mathbf{x}(t)$ for all the points on a curve $\mathbf{x}$.

At every point of a parametrized curve $\mathbf{x}(t)$ at which $\mathbf{x}'(t) \neq 0$, we may consider a *frame* $E_2(t)$, $E_3(t)$, where $E_2(t)$ and $E_3(t)$ are mutually orthogonal unit normal vectors at $\mathbf{x}(t)$. If $\underline{E}_2(t)$, $\underline{E}_3(t)$ is another such frame, then there is an angular function $\phi(t)$ such that

$$E_2(t) = \cos(\phi(t))\underline{E}_2(t) - \sin(\phi(t))\underline{E}_3(t)$$

$$E_3(t) = \sin(\phi(t))\underline{E}_2(t) + \cos(\phi(t))\underline{E}_3(t)$$

or, equivalently,

$$\underline{E}_2(t) = \cos(\phi(t))E_2(t) + \sin(\phi(t))E_3(t)$$

$$\underline{E}_3(t) = \sin(\phi(t))E_2(t) + \cos(\phi(t))E_3(t) \ .$$

From these two representations, we may derive an important formula:

$$E_2'(t) \cdot E_3(t) = \underline{E}_2'(t) \cdot \underline{E}_3(t) - \phi'(t)$$

Expressed in the form of differentials, without specifying parameters, this formula becomes:

$$dE_2E_3 = d\underline{E}_2\underline{E}_3 - d\phi \ .$$

Since $E_3(t) = \mathbf{T}(t) \times E_2(t)$, we have:

$$E_2'(t) \cdot E_3(t) = -[E_2'(t), E_2(t), \mathbf{T}(t)]$$

or, in differentials:

$$dE_2E_3 = -[dE_2, E_2, \mathbf{T}] \ .$$

More generally, if $\mathbf{z}(t)$ is a unit vector in the normal space at $\mathbf{x}(t)$, then we may define a function $\mathbf{w}(t) = -[\mathbf{z}'(t), \mathbf{z}(t), \mathbf{T}(t)]$. This is called the *connection function* of the unit normal bundle. The corresponding differential form $w = -[d\mathbf{z}, \mathbf{z}, \mathbf{T}]$ is called the *connection form* of the unit normal bundle.

A vector function $\mathbf{z}(t)$ such that $|\mathbf{z}(t)| = 1$ for all t and $\mathbf{z}(t) \cdot \mathbf{x}'(t) = 0$ for all t is called a *unit normal vector field* along the curve $\mathbf{x}$. Such a vector field is said to be *parallel* along $\mathbf{x}$ if the connection function $w(t) = -[\mathbf{z}'(t), \mathbf{z}(t), \mathbf{T}(t)] = 0$ for all t. In the next section, we will encounter several unit normal vector fields naturally associated with a given space curve. For now, we prove some general theorems about such objects.

Proposition 3. *If $E_2(t)$ and $\underline{E}_2(t)$ are two unit normal vector fields that are both parallel along the curve $\mathbf{x}$, then the angle between $E_2(t)$ and $\underline{E}_2(t)$ is constant.*

Proof. From the computation above, then:

$$E_2'(t) \cdot (-E_2(t) \times \mathbf{T}(t)) = \underline{E}_2'(t) \cdot (-\underline{E}_2(t) \times \mathbf{T}(t)) - \phi'(t) \ .$$

But, by hypothesis,

$$E_2'(t) \cdot (-E_2(t) x \mathbf{T}(t)) = 0 = \underline{E}_2'(t)(-\underline{E}_2(t) \times \mathbf{T}(t))$$

so it follows that $\phi'(t) = 0$ for all t, i.e., the angle $\phi(t)$ between $E_2(t)$ and $\underline{E}_2(t)$ is constant.

Given a closed curve $\mathbf{x}$ and a unit normal vector field $\mathbf{z}$ with $\mathbf{z}(b) = \mathbf{z}(a)$, we define

$$\mu(\mathbf{x}, \mathbf{z}) = -\frac{1}{2\pi}\int [\mathbf{z}'(t), \mathbf{z}(t), \mathbf{T}(t)]dt = -\frac{1}{2\pi}[d\mathbf{z}, \mathbf{z}, \mathbf{T}] \ .$$

If $\underline{\mathbf{z}}$ is another such field, then

$$\begin{aligned}\mu(\mathbf{x}, \mathbf{z}) - \mu(\mathbf{x}, \underline{\mathbf{z}}) &= -\frac{1}{2\pi}\int [\mathbf{z}'(t), \mathbf{z}(t), \mathbf{T}(t)] - [\underline{\mathbf{z}}'(t), \underline{\mathbf{z}}(t), \mathbf{T}(t)]dt \\ &= -\frac{1}{2\pi}\int \phi'(t)dt = -\frac{1}{2\pi}[\phi(b) - \phi(a)] \ .\end{aligned}$$

Since the angle $\phi(b)$ at the end of the closed curve must coincide with the angle $\phi(a)$ at the beginning, up to an integer multiple of 2π, it follows that the real numbers $\mu(\mathbf{x}, \mathbf{z})$ and $\mu(\mathbf{x}, \underline{\mathbf{z}})$ differ by an integer. Therefore the fractional part of $\mu(\mathbf{x}, \mathbf{z})$ depends only on the curve $\mathbf{x}$ and not on the unit normal vector field used to define it. This common value $\mu(\mathbf{x})$ is called the *total twist* of the curve $\mathbf{x}$. It is a global invariant of the curve.

Proposition 4. *If a closed curve lies on a sphere, then its total twist is zero.*

Proof. If $\mathbf{x}$ lies on the surface of a sphere of radius r centered at the origin, then $|\mathbf{x}(t)|^2 = \mathbf{x}(t) \cdot \mathbf{x}(t) = r^2$ for all t. Thus $\mathbf{x}'(t) \cdot \mathbf{x}(t) = 0$ for all t, so $\mathbf{x}(t)$ is a normal vector at $\mathbf{x}(t)$. Therefore $\mathbf{z}(t) = \frac{\mathbf{x}(t)}{r}$ is a unit normal vector field defined along $\mathbf{x}$, and we may compute the total twist by evaluating

$$\mu(\mathbf{x}, \mathbf{z}) = -\frac{1}{2\pi}\int [\mathbf{z}'(t), \mathbf{z}(t), \mathbf{T}(t)]dt \ .$$

But

$$[\mathbf{z}'(t), \mathbf{z}(t), \mathbf{T}(t)] = [\frac{\mathbf{x}'(t)}{r}, \frac{\mathbf{x}(t)}{r}, \mathbf{T}(t)] = 0$$

for all t since $\mathbf{x}'(t)$ is a multiple of $\mathbf{T}(t)$. In differential form notation, we get the same result: $[d\mathbf{z}, \mathbf{z}, \mathbf{T}] = \frac{1}{r^2}[\mathbf{x}'(t), \mathbf{x}(t), \mathbf{T}(t)]dt = 0$. Therefore $\mu(\mathbf{x}, \mathbf{z}) = 0$, so the total twist of the curve $\mathbf{x}$ is zero.

Remark 3. W. Scherrer proved that this property characterized a sphere, i.e. if the total twist of every curve on a closed surface is zero, then the surface is a sphere.

Remark 4. T. Banchoff and J. White proved that the total twist of a closed curve is invariant under inversion with respect to a sphere with center not lying on the curve.

Remark 5. The total twist plays an important role in modern molecular biology, especially with respect to the structure of DNA.

Exercise 23. Let $\mathbf{x}$ be the circle $\mathbf{x}(t) = (r\cos(t), r\sin(t), 0)$, where r is a constant > 1. Describe the collection of points $\mathbf{x}(t) + \mathbf{z}(t)$ where $\mathbf{z}(t)$ is a unit normal vector at $\mathbf{x}(t)$.

Exercise 24. Let Σ be the sphere of radius $r > 0$ about the origin. The *inversion* through the sphere S maps a point $\mathbf{x}$ to the point $\underline{\mathbf{x}} = r^2 \frac{\mathbf{x}}{|\mathbf{x}|^2}$. Note that this mapping is not defined if $\mathbf{x} = 0$, the center of the sphere. Prove that the coordinates of the inversion of $\mathbf{x} = (x_1, x_2, x_3)$ through S are given by $\underline{\mathbf{x}}_i = \frac{r^2 x_i}{x_1^2 + x_2^2 + x_3^2}$. Prove also that inversion preserves point that lie on the sphere S itself, and that the image of a plane is a sphere through the origin, except for the origin itself.

Exercise 25. Prove that the total twist of a closed curve not passing through the origin is the same as the total twist of its image by inversion through the sphere S of radius r centered at the origin.

2.4 Moving Frames

In the previous section, we introduced the notion of a frame in the unit normal bundle of a space curve. We now consider a slightly more general notion. By a *frame*, or more precisely a *right-handed rectangular frame with origin*, we mean a point $\mathbf{x}$ and a triple of mutually orthogonal unit vectors E_1, E_2, E_3 forming a right-handed system. The point $\mathbf{x}$ is called the origin of the frame. Note that $E_i \cdot E_j = 1$ if $i = j$ and 0 if $i \neq j$.

Moreover,

$$E_1 \times E_2 = E_3, \quad E_2 = E_3 \times E_1, \text{and} E_3 = E_1 \times E_2.$$

In the remainder of this section, we will always assume that small Latin letters run from 1 to 3.

Note that given two different frames, $\mathbf{x}, E_1, E_2, E_3$ and $\underline{\mathbf{x}}, E_1, E_2, E_3$, there is exactly one affine motion of Euclidean space taking $\mathbf{x}$ to $\underline{\mathbf{x}}$ and taking E_i to $\underline{E}_i$. When $\mathbf{x}(t), E_1(t), E_2(t), E_3(t)$ is a family of frames depending on a parameter t, we say we have a *moving frame* along the curve.

Proposition 5. *A family of frames* $\mathbf{x}(t), E_1(t), E_2(t), E_3(t)$ *satisfies a system of differential equations:*

$$\mathbf{x}'(t) = \Sigma p_i(t) E_i(t)$$
$$E_i'(t) = \Sigma q_{ij}(t) E_j(t)$$

where $p_i(t) = \mathbf{x}'(t) \cdot E_i(t)$ *and* $q_i j(t) = E_i'(t) \cdot E_j(t)$.

Since $E_i(t) \cdot E_j(t) = 0$ *for* $i \neq j$, *it follows that*

$$q_{ij}(t) + q_{ji}(t) = E_i'(t) \cdot E_j(t) + E_i(t) \cdot E_j'(t) = 0$$

i.e. the coefficients $q_{ij}(t)$ *are* anti-symmetric *in* i *and* j. *This can be expressed by saying that the matrix* $((q_i j(t)))$ *is an* anti-symmetric matrix, *with* 0 *on the diagonal.*

In a very real sense, the function $p_i(t)$ and $q_{ij}(t)$ completely determine the family of moving frames.

Specifically we have:

Proposition 6. *If* $\mathbf{x}(t), E_1(t), E_2(t), E_3(t)$ *and* $\underline{\mathbf{x}}(t), \underline{E}_1(t), \underline{E}_2(t), \underline{E}_3(t)$ *are two families of moving frames such that* $p_i(t) = \underline{p}_i(t)$ *and* $q_{ij}(t) = \underline{q}_{ij}(t)$ *for all* t, *then there is a single affine motion that takes* $\mathbf{x}(t), E_1(t), E_2(t), E_3(t)$ *to* $\underline{\mathbf{x}}(t), \underline{E}_1(t), \underline{E}_2(t), \underline{E}_3(t))$ *for all* t.

Proof. Recall that for a specific value t_0, there is an affine motion taking $\mathbf{x}(t_0), E_1(t_0), E_2(t_0), E_3(t_0)$ to $\underline{\mathbf{x}}(t_0), \underline{E}_1(t0), \underline{E}_2(t_0), \underline{E}_3(t_0)$. We will show that this same motion takes $\mathbf{x}(t), E_1(t), E_2(t), E_3(t)$ to $\underline{\mathbf{x}}(t), \underline{E}_1(t), \underline{E}_2(t), \underline{E}_3(t)$ for all t. Assume that the motion has been carried out so that the frames $\mathbf{x}(t_0), E_1(t_0), E_2(t_0), E_3(t_0)$ and $\underline{\mathbf{x}}(t_0), \underline{E}_1(t0), \underline{E}_2(t_0), \underline{E}_3(t_0)$ coincide.

Now consider

$$\begin{aligned}
(\Sigma E_i(t) \cdot \underline{E}_i(t))' &= \Sigma E_i'(t) \cdot \underline{E}_i(t) + \Sigma E_i(t) \cdot \underline{E}_i'(t) \\
&= \Sigma\Sigma q_{ij}(t) E_j(t) \cdot \underline{E}_i(t) + \Sigma E_i(t) \cdot \Sigma q_{ij}(t) \underline{E}_j(t) \\
&= \Sigma\Sigma q_{ij}(t) E_j(t) \cdot \underline{E}_i(t) + \Sigma\Sigma q_{ij}(t) E_i(t) \cdot \underline{E}_j(t) \\
&= \Sigma\Sigma q_{ij}(t) E_j(t) \cdot \underline{E}_i(t) + \Sigma\Sigma q_{ji}(t) E_j(t) \cdot \underline{E}_i(t) \\
&= 0\ .
\end{aligned}$$

It follows that

$$\Sigma E_i(t) \cdot \underline{E}_i(t) = \Sigma E_i(t_0) \cdot \underline{E}_i(t_0) = \Sigma E_i(t_0) \cdot E_i(t_0) = 3$$

for all t. But since $|E_i(t) \cdot \underline{E}_i(t)| \leq 1$ for any pair of unit vectors, we must have $E_i(t) \cdot \underline{E}_i(t) = 1$ for all t. Therefore $E_i(t) = \underline{E}_i(t)$ for all t.

Next consider

$$(\mathbf{x}(t) - \underline{\mathbf{x}}(t))' = \Sigma p_i(t) E_i(t) - \Sigma p_i(t) \underline{E}_i(t) = \Sigma p_i(t) E_i(t) - \Sigma p_i(t) E_i(t) = 0\ .$$

Since the origins of the two frames coincide at the value t_0, we have

$$\mathbf{x}(t) - \underline{\mathbf{x}}(t) = \mathbf{x}(t_0) - \underline{\mathbf{x}}(t_0) = 0$$

for all t.

This completes the proof that two families of frames satisfying the same set of differential equations differ at most by a single affine motion.

Exercise 26. Prove that the equations $E_i'(t) = \Sigma q_{ij}(t)E_j(t)$ can be written $E_i'(t) = \mathbf{d}(t) \times E_i(t)$, where $\mathbf{d}(t) = q_{23}(t)E_1(t) + q_{31}(t)E_2(t) + q_{12}(t)E_3(t)$. This vector is called the *instantaneous axis of rotation.*

Exercise 27. Under a rotation about the x_3-axis, a point describes a circle $\mathbf{x}(t) = (a\cos(t), a\sin(t), b)$. Show that its velocity vector satisfies $\mathbf{x}'(t) = \mathbf{d} \times \mathbf{x}(t)$ where $\mathbf{d} = (0, 0, 1)$. (Compare with the previous exercise.).

Exercise 28. Prove that $(\mathbf{v}\cdot\mathbf{v})(\mathbf{w}\cdot\mathbf{w}) - (\mathbf{v}\cdot\mathbf{w})^2 = 0$ if and only if the vectors $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent.

2.5 Curves at a Non-inflexional Point and the Frenet Formulas

A curve $\mathbf{x}$ is called *non-inflectional* if the curvature $k(t)$ is never zero. By our earlier calculations, this condition is equivalent to the requirement that $\mathbf{x}'(t)$ and $\mathbf{x}''(t)$ are linearly independent at every point $\mathbf{x}(t)$, i.e. $\mathbf{x}'(t) \times \mathbf{x}''(t) \neq 0$ for all t. For such a non-inflectional curve $\mathbf{x}$, we may define a pair of natural unit normal vector fields along $\mathbf{x}$.

Let $\mathbf{b}(t) = \frac{\mathbf{x}'(t)\times\mathbf{x}''(t)}{|\mathbf{x}'(t)\times\mathbf{x}''(t)|}$, called the *binormal vector* to the curve $\mathbf{x}(t)$. Since $\mathbf{b}(t)$ is always perpendicular to $\mathbf{T}(t)$, this gives a unit normal vector field along $\mathbf{x}$.

We may then take the cross product of the vector fields $\mathbf{b}(t)$ and $\mathbf{T}(t)$ to obtain another unit normal vector field $\mathbf{N}(t) = \mathbf{b}(t)\times\mathbf{T}(t)$, called the *principal normal vector.* The vector $\mathbf{N}(t)$ is a unit vector perpendicular to $\mathbf{T}(t)$ and lying in the plane determined by $\mathbf{x}'(t)$ and $\mathbf{x}''(t)$. Moreover, $\mathbf{x}''(t) \cdot \mathbf{N}(t) = k(t)s'(t)^2$, a positive quantity.

Note that if the parameter is arclength, then $\mathbf{x}'(s) = \mathbf{T}(s)$ and $\mathbf{x}''(s)$ is already perpendicular to $\mathbf{T}(s)$. It follows that $\mathbf{x}''(s) = k(s)\mathbf{N}(s)$ so we may define $\mathbf{N}(s) = \frac{\mathbf{x}''(s)}{k(s)}$ and then define $\mathbf{b}(s) = \mathbf{T}(s) \times \mathbf{N}(s)$. This is the standard procedure when it happens that the parametrization is by arclength. The method above works for an arbitrary parametrization.

We then have defined an orthonormal frame $\mathbf{x}(t)\mathbf{T}(t)\mathbf{N}(t)\mathbf{b}(t)$ called the *Frenet frame* of the non-inflectional curve $\mathbf{x}$.

By the previous section, the derivatives of the vectors in the frame can be expressed in terms of the frame itself, with coefficients that form an anti-symmetric matrix. We already have $\mathbf{x}'(t) = s'(t)\mathbf{T}(t)$, so

$$p_1(t) = s'(t) \ , \ p_2(t) = 0 = p_3(t) \ .$$

Also $\mathbf{T}'(t) = k(t)s'(t)\mathbf{N}(t)$, so

$$q_{12}(t) = k(t)s'(t) \text{ and } q_{13}(t) = 0 \ .$$

We know that

$$\mathbf{b}'(t) = q_{31}(t)\mathbf{T}(t) + q_{32}(t)\mathbf{N}(t) \text{ , and } q_{31}(t) = -q_{13}(t) = 0 \text{ .}$$

Thus $\mathbf{b}'(t)$ is a multiple of $\mathbf{N}(t)$, and we define the *torsion* $w(t)$ of the curve by the condition

$$\mathbf{b}'(t) = -w(t)s'(t)\mathbf{N}(t) \text{ ,}$$

so $q_{32}(t) = -w(t)s'(t)$ for the Frenet frame. From the general computations about moving frames, it then follows that

$$\mathbf{N}'(t) = q_{21}(t)\mathbf{T}(t) + q_{23}(t)\mathbf{b}(t) = -k(t)s'(t)\mathbf{T}(t) + w(t)s'(t)\mathbf{b}(t) \text{ .}$$

The formulas for $\mathbf{T}'(t)$, $\mathbf{N}'(t)$, and $\mathbf{b}'(t)$ are called the *Frenet formulas* for the curve $\mathbf{x}$.

If the curve $\mathbf{x}$ is parametrized with respect to arclength, then the Frenet formulas take on a particularly simple form:

$$\begin{aligned} \mathbf{x}'(s) &= \mathbf{T}(s) \\ \mathbf{T}'(s) &= k(s)\mathbf{N}(s) \\ \mathbf{N}'(s) &= -k(s)\mathbf{T}(s) + w(s)\mathbf{b}(s) \\ \mathbf{b}'(s) &= -w(s)\mathbf{b}(s) \text{ .} \end{aligned}$$

The torsion function $w(t)$ that appears in the derivative of the binormal vector determines important properties of the curve. Just as the curvature measures deviation of the curve from lying along a straight line, the torsion measures deviation of the curve from lying in a plane. Analogous to the result for curvature, we have:

Proposition 7. *If $w(t) = 0$ for all points of a non-inflectional curve $\mathbf{x}$, then the curve is contained in a plane.*

Proof. We have $\mathbf{b}'(t) = -w(t)s'(t)\mathbf{N}(t) = 0$ for all t so $\mathbf{b}(t) = \mathbf{a}$, a constant unit vector. Then $\mathbf{T}(t)\mathbf{a} = 0$ for all t so $(\mathbf{x}(t) \cdot \mathbf{a})' = \mathbf{x}'(t) \cdot \mathbf{a} = 0$ and $\mathbf{x}(t) \cdot \mathbf{a} = \mathbf{x}(a) \cdot \mathbf{a}$, a constant. Therefore $(\mathbf{x}(t) - \mathbf{x}(a)) \cdot \mathbf{a} = 0$ and $\mathbf{x}$ lies in the plane through $\mathbf{x}(a)$ perpendicular to $\mathbf{a}$.

If $\mathbf{x}$ is a non-inflectional curve parametrized by arclength, then

$$w(s) = \mathbf{b}(s) \cdot \mathbf{N}'(s) = [\mathbf{T}(s), \mathbf{N}(s), \mathbf{N}'(s)] \text{ .}$$

Since $\mathbf{N}(s) = \frac{\mathbf{x}''(s)}{k(s)}$, we have

$$\mathbf{N}'(s) = \frac{\mathbf{x}'''(s)}{k(s)} + \mathbf{x}''(s)\frac{-k'(s)}{k(s)^2} \text{ ,}$$

so

$$w(s) = \left[\mathbf{x}'(s), \frac{\mathbf{x}''(s)}{k(s)}, \frac{\mathbf{x}'''(s)}{k(s)} + \mathbf{x}''(s)\frac{-k'(s)}{k(s)^2}\right] = \frac{[\mathbf{x}'(s), \mathbf{x}''(s), \mathbf{x}'''(s)]}{k(s)^2} \text{ .}$$

We can obtain a very similar formula for the torsion in terms of an arbitrary parametrization of the curve $\mathbf{x}$. Recall that

$$\mathbf{x}''(t) = s\text{"}(t)\mathbf{T}(t) + k(t)s'(t)\mathbf{T}'(t) = s''(t)\mathbf{T}(t) + k(t)s'(t)^2\mathbf{N}(t) ,$$

so

$$\mathbf{x}'''(t) = s'''(t)\mathbf{T}(t) + s''(t)s'(t)k(t)\mathbf{N}(t) + \left[k(t)s'(t)^2\right]'\mathbf{N}(t) + k(t)s'(t)^2\mathbf{N}'(t) .$$

Therefore

$$\mathbf{x}'''(t)\mathbf{b}(t) = k(t)s'(t)^2\mathbf{N}'(t)\mathbf{b}(t) = k(t)s'(t)^2 w(t)s'(t) ,$$

and

$$\mathbf{x}'''(t) \cdot \mathbf{x}'(t) \times \mathbf{x}''(t) = k^2(t)s'(t)^6 w(t) .$$

Thus we obtain the formula

$$w(t) = \frac{\mathbf{x}'''(t) \cdot \mathbf{x}'(t)x\mathbf{x}''(t)}{|\mathbf{x}'(t) \times \mathbf{x}''(t)|^2} ,$$

valid for any parametrization of $\mathbf{x}$.

Notice that although the curvature $k(t)$ is never negative, the torsion $w(t)$ can have either algebraic sign. For the circular helix $\mathbf{x}(t) = (r\cos(t), r\sin(t), pt)$ for example, we find $w(t) = \frac{p}{r^2+p^2}$, so the torsion has the same algebraic sign as p. In this way, the torsion can distinguish between a right-handed and a left-handed screw.

Changing the orientation of the curve from s to $-s$ changes $\mathbf{T}$ to $-\mathbf{T}$, and choosing the opposite sign for $k(s)$ changes $\mathbf{N}$ to $-\mathbf{N}$. With different choices, then, we can obtain four different right-handed orthonormal frames, $\mathbf{xTNb}$, $\mathbf{x}(-\mathbf{T})\mathbf{N}(-\mathbf{b})$, $\mathbf{xT}(-\mathbf{N})(-\mathbf{b})$, and $\mathbf{x}(-\mathbf{T})(-\mathbf{N})\mathbf{b}$. Under all these changes of the Frenet frame, the value of the torsion $w(t)$ remains unchanged.

A circular helix has the property that its curvature and its torsion are both constant. Furthermore the unit tangent vector $\mathbf{T}(t)$ makes a constant angle with the vertical axis. Although the circular helices are the only curves with constant curvature and torsion, there are other curves that have the second property. We characterize such curves, as an application of the Frenet frame.

Proposition 8. *The unit tangent vector* $\mathbf{T}(t)$ *of a non-inflectional space curve* $\mathbf{x}$ *makes a constant angle with a fixed unit vector* $\mathbf{a}$ *if and only if the ratio* $\frac{w(t)}{k(t)}$ *is constant.*

Proof. If $\mathbf{T}(t) \cdot \mathbf{a} =$ constant for all t, then differentiating both sides, we obtain

$$\mathbf{T}'(t) \cdot \mathbf{a} = 0 = k(t)s'(t)\mathbf{N}(t) \cdot \mathbf{a} ,$$

so $\mathbf{a}$ lies in the plane of $\mathbf{T}(t)$ and $\mathbf{b}(t)$. Thus we may write $\mathbf{a} = \cos(\phi)\mathbf{T}(t) + \sin(\phi)\mathbf{b}(t)$ for some angle ϕ. Differentiating this equation, we obtain

$$0 = \cos(\phi)\mathbf{T}'(t) + \sin(\phi)\mathbf{b}'(t) = \cos(\phi)k(t)s'(t)\mathbf{N}(t) - \sin(\phi)w(t)s'(t)\mathbf{N}(t) \ ,$$

so $\frac{w(t)}{k(t)} = \frac{\sin(\phi)}{\cos(\phi)} = \tan(\phi)$. This proves the first part of the proposition and identifies the constant ratio of the torsion and the curvature.

Conversely, if $\frac{w(t)}{k(t)} = \text{constant} = \tan(\phi)$ for some ϕ, then, by the same calculations, the expression $\cos(\phi)\mathbf{T}(t) + \sin(\phi)\mathbf{b}(t)$ has derivative 0 so it equals a constant unit vector. The angle between $\mathbf{T}(t)$ and this unit vector is the constant angle ϕ.

Curves with the property that the unit tangent vector makes a fixed angle with a particular unit vector are called *generalized helices.* Just as a circular helix lies on a circular cylinder, a generalized helix will lie on a general cylinder, consisting of a collection of lines through the curve parallel to a fixed unit vector. On this generalized cylinder, the unit tangent vectors make a fixed angle with these lines, and if we roll the cylinder out onto a plane, then the generalized helix is rolled out into a straight line on the plane.

We have shown in the previous section that a moving frame is completely determined up to an affine motion by the functions $p_i(t)$ and $q_{ij}(t)$. In the case of the Frenet frame, this means that if two curves $\mathbf{x}$ and $\underline{\mathbf{x}}$ have the same arclength $s(t)$, the same curvature $k(t)$, and the same torsion $w(t)$, then the curves are congruent, i.e. there is an affine motion of Euclidean three-space taking $\mathbf{x}(t)$ to $\underline{\mathbf{x}}(t)$ for all t. Another way of stating this result is:

Theorem 2. *The Fundamental Theorem of Space Curves. Two curves parametrized by arclength having the same curvature and torsion at corresponding points are congruent by an affine motion.*

Exercise 29. Compute the torsion of the circular helix. Show directly that the principal normals of the helix are perpendicular to the vertical axis, and show that the binormal vectors make a constant angle with this axis.

Exercise 30. Prove that if the curvature and torsion of a curve are both constant functions, then the curve is a circular helix (i.e. a helix on a circular cylinder).

Exercise 31. Prove that a necessary and sufficient condition for a curve $\mathbf{x}$ to be a generalized helix is that

$$\mathbf{x}''(t) \times \mathbf{x}'''(t) \cdot \mathbf{x}^{iv}(t) = 0 \ .$$

Exercise 32. Let $\mathbf{y}(t)$ be a curve on the unit sphere, so that $|\mathbf{y}(t)| = 1$ and $\mathbf{y}(t) \cdot \mathbf{y}'(t) \times \mathbf{y}''(t) \neq 0$ for all t. Show that the curve $\mathbf{x}(t) = c \int \mathbf{y}(u) \times \mathbf{y}''(u) du$ with $c \neq 0$ has constant torsion $\frac{1}{c}$.

Exercise 33. (For students familiar with complex variables) If the coordinate functions of the vectors in the Frenet frame are given by

$$\begin{aligned}\mathbf{T} &= (e_{11}, e_{12}, e_{13}) \ ,\\ \mathbf{N} &= (e_{21}, e_{22}, e_{23}) \ ,\\ \mathbf{b} &= (e_{31}, e_{32}, e_{33}) \ ,\end{aligned}$$

then we may form the three complex numbers

$$z_j = \frac{e_{1j} + ie_{2j}}{1 - e_{3j}} = \frac{1 + e_{3j}}{e_{1j} - ie_{2j}} \ .$$

Then the functions z_j satisfy the Riccati equation

$$z_j' = -ik(s)z_j + \frac{i}{2}w(s)(-1 + z_j^2) \ .$$

This result is due to S. Lie and G. Darboux.

2.6 Local Equations of a Curve

We can "see" the shape of a curve more clearly in the neighborhood of a point $\mathbf{x}(t_0)$ when we consider its parametric equations with respect to the Frenet frame at the point. For simplicity, we will assume that $t_0 = 0$, and we may then write the curve as

$$\mathbf{x}(t) = \mathbf{x}(0) + x_1(t)\mathbf{T}(0) + x_2(t)\mathbf{N}(0) + x_3(t)\mathbf{b}(0) \ .$$

On the other hand, using the Taylor series expansion of $\mathbf{x}(t)$ about the point $t = 0$, we obtain

$$\mathbf{x}(t) = \mathbf{x}(0) + t\mathbf{x}'(0) + \frac{t^2}{2}\mathbf{x}''(0) + \frac{t^3}{6}\mathbf{x}'''(0) + \text{higher order terms} \ .$$

From our earlier formulas, we have

$$\begin{aligned}\mathbf{x}'(0) &= s'(0)\mathbf{T}(0) \ ,\\ \mathbf{x}''(0) &= s''(0)\mathbf{T}(0) + k(0)s'(0)^2\mathbf{N}(0) \ ,\\ \mathbf{x}'''(0) &= s'''(0)\mathbf{T}(0) + s''(0)s'(0)k(0)\mathbf{N}(0) + (k(0)s'(0)^2)'\mathbf{N}(0)\\ &\quad + k(0)s'(0)^2(-k(0)s'(0)\mathbf{T}(0) + w(0)s'(0)\mathbf{b}(0)) \ .\end{aligned}$$

Substituting these equations in the Taylor series expression, we find:

$$\begin{aligned}\mathbf{x}(t) = \mathbf{x}(0) &+ \left(ts'(0) + \frac{t^2}{2}s''(0) + \frac{t^3}{6}\left[s'''(0) - k(0)^2s'(0)^3\right] + ...\right)\mathbf{T}(0)\\ &+ \left(\frac{t^2}{2}k(0)s'(0)^2 + \frac{t^3}{6}\left[s''(0)s'(0)k(0) + (k(0)s'(0)^2)'\right] + ...\right)\mathbf{N}(0)\\ &+ \left(\frac{t^3}{6}k(0)w(0)s'(0)^3 + ...\right)\mathbf{b}(0) \ .\end{aligned}$$

If the curve is parametrized by arclength, this representation is much simpler:

$$\begin{aligned}\mathbf{x}(s) = \mathbf{x}(0) &+ \left(s - \frac{k(0)^2}{6}s^3 + ...\right)\mathbf{T}(0) \\ &+ \left(\frac{k(0)}{2}s^2 + \frac{k'(0)}{6}s^3 + ...\right)\mathbf{N}(0) \\ &+ \left(\frac{k(0)w(0)}{6}s^3 + ...\right)\mathbf{b}(0) \ .\end{aligned}$$

Relative to the Frenet frame, the plane with equation $x_1 = 0$ is the *normal plane*; the plane with $x_2 = 0$ is the *rectifying plane*, and the plane with $x_3 = 0$ is the *osculating plane*. These planes are orthogonal respectively to the unit tangent vector, the principal normal vector, and the binormal vector of the curve.

2.7 Plane Curves and a Theorem on Turning Tangents

The general theory of curves developed above applies to plane curves. In the latter case there are, however, special features which will be important to bring out. We suppose our plane to be oriented. In the plane a vector has two components and a frame consists of an origin and an ordered set of two mutually perpendicular unit vectors forming a right-handed system. To an oriented curve C defined by $\mathbf{x}(s)$ the Frenet frame at s consists of the origin $\mathbf{x}(s)$, the unit tangent vector $\mathbf{T}(s)$ and the unit normal vector $\mathbf{N}(s)$. Unlike the case of space curves this Frenet frame is uniquely determined, under the assumption that both the plane and the curve are oriented.

The Frenet formulas are

$$\begin{aligned}\mathbf{x}' &= \mathbf{T} \ , \\ \mathbf{T}' &= k\mathbf{N} \ , \\ \mathbf{N}' &= -k\mathbf{T} \ .\end{aligned} \tag{2.1}$$

The *curvature* $k(s)$ is defined with sign. It changes its sign when the orientation of the plane or the curve is reversed.

The Frenet formulas in (2.1) can be written more explicitly. Let

$$\mathbf{x}(s) = (x_1(s) \ , \ x_2(s)) \tag{2.2}$$

Then

$$\begin{aligned}\mathbf{T}(s) &= (\ \ x_1'(s) \ , \ x_2'(s)) \ , \\ \mathbf{N}(s) &= (-x_2'(s) \ , \ x_1'(s)) \ .\end{aligned} \tag{2.3}$$

Expressing the last two equations of (2.1) in components, we have

$$x_1'' = -kx_2' \tag{2.4}$$

$$x_2'' = \quad kx_1' \ . \tag{2.5}$$

These equations are equivalent to (2.1).

Since $\mathbf{T}$ is a unit vector, we can put

$$\mathbf{T}(s) = (\cos \tau(s) \ , \ \sin \tau(s)) \ , \tag{2.6}$$

so that $\tau(s)$ is the angle of inclination of $\mathbf{T}$ with the x_1-axis. Then

$$\mathbf{N}(s) = (-\sin \tau(s) \ , \ \cos \tau(s)) \ , \tag{2.7}$$

and (2.1) gives

$$\frac{d\tau}{ds} = k(s) \tag{2.8}$$

This gives a geometrical interpretation of $k(s)$.

A curve C is called *simple* if it does not intersect itself. One of the most important theorems in global differential geometry is the theorem on turning tangents:

Theorem 3. *For a simple closed plane curve we have*

$$\frac{1}{2\pi} \oint k \, ds = \pm 1 \ .$$

To prove this theorem we give a geometrical interpretation of the integral at the left-hand side of (3). By (2.8)

$$\frac{1}{2\pi} \oint k \, ds = \frac{1}{2\pi} \oint d\tau \ .$$

But τ , as the angle of inclination of $\tau(s)$, is only defined up to an integral multiple of 2π, and this integral has to be studied with care.

Let O be a fixed point in the plane. Denote by Γ the unit circle about O; it is oriented by the orientation of the plane. The tangential mapping or Gauss mapping

$$g \colon C \mapsto \Gamma \tag{2.9}$$

is defined by sending the point $\mathbf{x}(s)$ of C to the point $\mathbf{T}(s)$ of Γ. In other words, $g(P)$, $P \in C$, is the end-point of the unit vector through O parallel to the unit tangent vector to C at P. Clearly g is a continuous mapping. If C is closed, it is intuitively clear that when a point goes along C once its image point under g goes along Γ a number of times. This integer is called the *rotation index* of C. It is to be defined rigorously as follows:

We consider O to be the origin of our coordinate system. As above we denote by $\tau(s)$ the angle of inclination of $\mathbf{T}(s)$ with the x_1-axis. In order to make the angle uniquely determined we suppose $O \leq \tau(s) < 2\pi$. But $\tau(s)$ is not necessarily continuous. For in every neighborhood of s_0 at which $\tau(s_c) = 0$ there may be values of $\tau(s)$ differing from 2π by arbitrarily small quantities. We have, however, the following lemma:

Lemma 2. *There exists a continuous function $\tilde{\tau}(s)$ such that $\tilde{\tau}(s) \equiv \tau(s)$ mod 2π .*

Proof. We suppose C to be a closed curve of total length L. The continuous mapping g is uniformly continuous. There exists therefore a number $\delta > 0$ such that for $|s_1 - s_2| < \delta$, $\mathbf{T}(s_1)$ and $\mathbf{T}(s_2)$ lie in the same open half-plane. Let $s_0(=O) < s_1 < \cdots < s_i(=L)$ satisfy $|s_i - s_{i-1}| < \delta$ for $i = 1, \ldots, m$. We put $\tilde{\tau}(s_0) = \tau(s_0)$. For $s_0 \leq s \leq s_1$, we define $\tilde{\tau}(s)$ to be $\tilde{\tau}(s_0)$ plus the angle of rotation from $g(s_0)$ to $g(s)$ remaining in the same half-plane. Carrying out this process in successive intervals, we define a continuous function $\tilde{\tau}(s)$ satisfying the condition in the lemma. The difference $\tilde{\tau}(L) - \tilde{\tau}(O)$ is an integral multiple of 2π. Thus, $\tilde{\tau}(L) - \tilde{\tau}(O) = \gamma 2\pi$. We assert that the integer γ is independent of the choice of the function $\tilde{\tau}$. In fact let $\tilde{\tau}'(s)$ be a function satisfying the same conditions. Then we have $\tilde{\tau}'(s) - \tilde{\tau}(s) = n(s)2\pi$ where $n(s)$ is an integer. Since $n(s)$ is continuous in s, it must be constant. It follows that $\tilde{\tau}'(L) - \tilde{\tau}'(O) = \tilde{\tau}(L) - \tilde{\tau}(O)$, which proves the independence of γ from the choice of $\tilde{\tau}$. We call γ the rotation index of C. In performing integration over C we should replace $\tau(s)$ by $\tilde{\tau}$ in (2.8). Then we have

$$\frac{1}{2\pi} \oint k \, ds = \frac{1}{2\pi} \oint d\,\tilde{\tau} = \gamma \,. \tag{2.10}$$

We consider the mapping h which sends an ordered pair of points $\mathbf{x}(s_1)$, $\mathbf{x}(s_2)$, $O \leq s_1 \leq s_2 \leq L$, of C into the end-point of the unit vector through O parallel to the secant joining $\mathbf{x}(s_1)$ to $\mathbf{x}(s_2)$. These ordered pairs of points can be represented as a triangle $\triangle$ in the (s_1, s_2)-plane defined by $O \leq s_1 \leq s_2 \leq L$. The mapping h of $\triangle$ into Γ is continuous. Moreover, its restriction to the side $s_1 = s_2$ is the tangential mapping g in (2.9).

To a point $p \in \triangle$ let $\tau(p)$ be the angle of inclination of $\overline{Oh(p)}$ to the $\mathbf{x}_1$-axis, satisfying $O \leq \tau(p) < 2\pi$. Again this function need not be continuous. We shall, however, prove that there exists a continuous function $\widetilde{\tau(p)}$, $p \in \triangle$, such that $\tilde{\tau}(p) \equiv \tau(p) \mod 2\pi$. In fact, let m be an interior point of $\triangle$. We cover $\triangle$ by the radii through m. By the argument used in the proof of the above lemma we can define a function $\tilde{\tau}(p)$, $p \in \triangle$, such that $\tilde{\tau}(p) \equiv \tau(p)$ mod 2π, and such that it is continuous on every radius through m. It remains to prove that it is continuous in $\triangle$.

For this purpose let $p_0 \in \triangle$. Since h is continuous, it follows from the compactness of the segment $\overline{mp_0}$ that there exists a number $\eta = \eta(p_0) > 0$, such that for $q_0 \in \overline{mp_0}$ and for any point $q \in \triangle$ for which the distance $d(q, q_0) < \eta$ the points $h(q)$ and $h(q_0)$ are never antipodal. The latter condition can be analytically expressed by

$$\tilde{\tau}(q) \not\equiv \tilde{\tau}(q_0) \mod \pi \,. \tag{2.11}$$

Now let $\epsilon > 0$, $\epsilon < \frac{\pi}{2}$ be given. We choose a neighborhood U of p_0 such that U is contained in the η-neighborhood of p_0 and such that, for $p \in U$, the angle between $\overline{Oh(p_0)}$ and $\overline{Oh(p)}$ is $< \epsilon$. This is possible, because the mapping h is continuous. The last condition can be expressed in the form

$$\tilde{\boldsymbol{\tau}}(p) - \tilde{\boldsymbol{\tau}}(p_0) = \epsilon' + 2k(p)\pi \,, \tag{2.12}$$

where $k(p)$ is an integer. Let q_0 be any point on the segment $\overline{mp_0}$. Draw the segment $\overline{qq_0}$ parallel to $\overline{pp_0}$, with q on $\overline{mp}$. The function $\tilde{\boldsymbol{\tau}}(q) - \tilde{\boldsymbol{\tau}}(q_0)$ is continuous in q along $\overline{mp}$ and equals O when q coincides with m. Since $d(q, q_0) < \eta$, it follows from (2.11) that $|\tilde{\boldsymbol{\tau}}(q) - \tilde{\boldsymbol{\tau}}(q_0)| < \pi$. In particular, for $q_o = p_0$ this gives $|\tilde{\boldsymbol{\tau}}(p) - \tilde{\boldsymbol{\tau}}(p_0)| < \pi$. Combining this with (2.12), we get $k(p) = 0$. Thus we have proved that $\tilde{\boldsymbol{\tau}}(p)$ is continuous in $\triangle$, as asserted above. Since $\tilde{\boldsymbol{\tau}}(p) \equiv \boldsymbol{\tau}(p) \mod 2\pi$, it is clear that $\tilde{\boldsymbol{\tau}}(p)$ is differentiable.

Now let $A(O, O)$, $B(O, L)$, $D(L, L)$ be the vertices of $\triangle$. The rotation index γ of C is, by (2.10), defined by the line integral

$$2\pi\gamma = \oint_{\overline{AD}} d\tilde{\boldsymbol{\tau}} \,.$$

Since $\tilde{\boldsymbol{\tau}}(p)$ is defined in $\triangle$, we have

$$\oint_{\overline{AD}} d\tilde{\boldsymbol{\tau}} = \oint_{\overline{AB}} d\tilde{\boldsymbol{\tau}} + \oint_{\overline{BD}} d\tilde{\boldsymbol{\tau}} \,.$$

To evaluate the line integrals at the right-hand side, we suppose the origin O to be the point $\mathbf{x}(O)$ and C to lie in the upper half-plane and to be tangent to the $\mathbf{x}_1$-axis at O. This is always possible for we only have to take $\mathbf{x}(O)$ to be the point on C at which the x_2-coordinate is a minimum. Then the $\mathbf{x}_1$-axis is either in the direction of the tangent vector to C at O or opposite to it. We can assume the former case, by reversing the orientation of C if necessary. The line integral along $\overline{AB}$ is then equal to the angle rotated by $\overline{OP}$ as P goes once along C. Since C lies in the upper half-plane, the vector $\overline{OP}$ never points downward. It follows that the integral along $\overline{AB}$ is equal to π. On the other hand, the line integral along $\overline{BD}$ is the angle rotated by $\overline{PO}$ as P goes once along C. Since the vector $\overline{PO}$ never points upward, this integral is also equal to π. Hence their sum is 2π and the rotation index γ is $+1$. Since we may have reversed the orientation of C, the rotation index is ± 1 in general.

Exercise 34. Consider the plane curve $\mathbf{x}(t) = (t, f(t))$. Use the Frenet formulas in (2.1) to prove that its curvature is given by

$$k(t) = \frac{\ddot{f}}{(1 + \dot{f}^2)^{3/2}} \,. \tag{2.13}$$

Exercise 35. Draw closed plane curves with rotation indices $0, -2, +3$ respectively.

Exercise 36. The theorem on turning tangents is also valid when the simple closed curve C has "corners." Give the theorem when C is a triangle consisting of three arcs. Observe that the theorem contains as a special case the theorem on the sum of angles of a rectilinear triangle.

Exercise 37. Give in detail the proof of the existence of $\eta = \eta(p_0)$ used in the proof of the theorem on turning tangents. $\eta = \eta(p_0)$.

2.8 Plane Convex Curves and the Four Vertex Theorem

A closed curve in the plane is called *convex*, if it lies at one side of every tangent line.

Proposition 9. *A simple closed curve is convex, if and only if it can be so oriented that its curvature k is ≥ 0.*

The definition of a convex curve makes use of the whole curve, while the curvature is a local property. The proposition therefore gives a relationship between a local property and a global property. The theorem is not true if the closed curve is not simple. Counter examples can be easily constructed.

Let $\tilde{\tau(s)}$ be the function constructed above, so that we have $k = \frac{d\tau}{ds}$. The condition $k \geq O$ is therefore equivalent to the assertion that $\tilde{\tau}(s))$ is a monotone non-decreasing function. We can assume that $\tilde{\tau(O)} = O$. By the theorem on turning tangents, we can suppose C so oriented that $\tilde{\tau(L)} = 2\pi$.

Suppose $\tilde{\tau}(s)$, $O \leq s \leq L$, be monotone non-decreasing and that C is not convex. There is a point $A = \mathbf{x}(s_0)$ on C such that there are points of C at both sides of the tangent λ to C at A. Choose a positive side of k and consider the oriented perpendicular distance from a point $\mathbf{x}(s)$ of C to λ. This is a continuous function in s and attains a maximum and a minimum at the points M and N respectively. Clearly M and N are not on λ and the tangents to C at M and N are parallel to $\mathbf{x}$. Among these two tangents and k itself there are two tangents parallel in the same sense. Call $s_1 < s_2$ the values of the parameters at the corresponding points of contact. Since $\tilde{\tau}(s)$ is monotone non-decreasing and $O \leq \tilde{\tau}(s) \leq 2\pi$, this happems only when $\tilde{\tau}(s) = \tilde{\tau(s_1)}$ for all s satisfying $s_1 \leq s \leq s_2$. It follows that the arc $s_1 \leq s \leq s_2$ is a line segment parallel to λ. But this is obviously impossible.

Next let C be convex. To prove that $\tilde{\tau}(s)$ is monotone non-decreasing, suppose $\tilde{\tau(s_1)} = \tilde{\tau(s_2)}$, $s_1 < s_2$. Then the tangents at $\mathbf{x}(s_1)$ and $\mathbf{x}(s_2)$ are parallel in the same sense. But there exists a tangent parallel to them in the opposite sense. From the convexity of C it follows that two of them coincide.

We are thus in the situation of a line λ tangent to C at two distinct points A and B. We claim that the segment $\overline{AB}$ must be a part of C. In fact, suppose this is not the case and let D be a point on $\overline{AB}$ not on C. Draw through D a perpendicular μ to λ in the half-plane which contains C. Then

μ intersects C in at least two points. Among these points of intersection let F be the farthest from λ and G the nearest one, so that $F \neq G$. Then G is an interior point of the triangle ABF. The tangent to C at G must have points of C in both sides which contradicts the convexity of C.

It follows that under our assumption, the segment $\overline{AB}$ is a part of C , so that the tangents at A and B are parallel in the same sense. This proves that the segment joining $\mathbf{x}(s_1)$ to $\mathbf{x}(s_2)$ belongs to C. Hence $\tilde{\boldsymbol{\tau}}(s)$ remains constant in the interval $s_1 \leq s \leq s_2$. We have therefore proved that $\tilde{\boldsymbol{\tau}}(s)$ is monotone and $K \geq O$.

A point on C at which $k' = 0$ is called a *vertex*. A closed curve has at least two vertices, e.g., the maximum and the minimum of k. Clearly a circle consists entirely of vertices. An ellipse with unequal axes has four vertices, which are its intersection with the axes.

Theorem 4 (Four-vertex Theorem.). *A simple closed convex cuxve has at least four vertices.*

Remark 6. This theorem was first given by Mukhopadhyaya (1909). The following proof was due to G.Herglotz. It is also true for non-convex curves, but the proof will be more difficult.

2.9 Isoperimetric Inequality in the Plane

Among all simple closed curves having a given length the circle bounds the largest area, and is the only curve with this property. We shall state the theorem as follows:

Theorem 5. *Let L be the length of a simple closed curve C and A be the area it bounds. Then*

$$L^2 - 4\pi A \geq 0 \, . \tag{2.14}$$

Moreover, the equality sign holds only when C is a circle.

The proof given below is due to E. Schmidt (1939).

We enclose C between two parallel lines g, g', such that C lies between g, g' and is tangent to them at the points P, Q respectively. Let $s = 0$, s_0 be the parameters of P, Q. Construct a circle C tangent to g, g' at P, Q respectively. Denote its radius by r and take its center to be the origin of a coordinate system. Let $\mathbf{x}(s) = (x_1(s), x_2(s))$ be the position vector of C, so that

$$(x_1(0), x_2(0)) = (x_1(L), x_2(L)) \, .$$

As the position vector of $\overline{C}$ we take $(\overline{x}_1(s), \overline{x}_2)$, such that

$$\begin{aligned}\overline{x_1}(s) &= x_1(s)\ ,\\ \overline{x_2}(s) &= -\sqrt{r^2 - x_1^2(s)}\ ,\ 0 \le s \le s_0\\ &= +\sqrt{r^2 - x_1^2(s)}\ ,\ s_0 \le s \le L\ .\end{aligned}$$

Denote by $\overline{A}$ the area bounded by $\overline{C}$. Now the area bounded by a closed curve can be expressed by the line integral

$$A = \int_0^L x_1 x_2' ds = -\int_0^L x_2 x_1' ds = \frac{1}{2}\int_0^L (x_1 x_2' - x_2 x_1') ds\ .$$

Applying this to our two curves C and $\overline{C}$, we get

$$\begin{aligned}A &= \int_0^L x_1 x_2' ds\ ,\\ \overline{A} = \pi r^2 &= -\int_0^L x_2 \overline{x}_1' ds = -\int_0^L \overline{x}_2 x_1' ds\ .\end{aligned}$$

Adding these two equations, we have

$$\begin{aligned}A + \pi r^2 &= \int_0^L (x_1 x_2' - \overline{x}_2 x_1') ds \le \int_0^L \sqrt{(x_1 x_2' - \overline{x}_2 x_1')^2} ds\\ &\le \int_0^L \sqrt{(x_1^2 + \overline{x}_2^2)(x_1'^2 + x_2'^2)} ds\\ &= \int_0^L \sqrt{x_1^2 + \overline{x}_2^2} ds = Lr\ .\end{aligned} \tag{2.15}$$

Since the geometric mean of two numbers is $\le$ their arithmetic mean, it follows that

$$\sqrt{A}\sqrt{\pi r^2} \le \frac{1}{2}(A + \pi r^2) \le \frac{1}{2} Lr\ .$$

This gives, after squaring and cancellation of r, the inequality (2.14).

Suppose now that the equality sign in (2.14) holds. A and πr^2 have then the same geometric and arithmetic mean, so that $A = \pi r^2$ and $L = 2\pi r$. The direction of the lines g, g' being arbitrary, this means that C has the same "width" in all directions. Moreover, we must have the equality sign everywhere in (2.15). It follows in particular that

$$(x_1 x_2' - \overline{x}_2 x_1')^2 = (x_1^2 + \overline{x}_2^2)(x_1'^2 + x_2'^2)\ ,$$

which gives

$$\frac{x_1}{x_2'} = \frac{-\overline{x}_2}{x_1'} = \frac{\sqrt{x_1^2 + \overline{x}_2^2}}{\sqrt{x_1'^2 + x_2'^2}} = \pm r\ .$$

From the first equality in (2.15) the factor of proportionality is seen to be r, i.e.,

$$x_1 = r x_2' \ , \ \overline{x}_2 = -r x_1' \ .$$

This remains true when we interchange x_1 and x_2, so that

$$x_2 = r x_1' \ .$$

Therefore we have

$$x_1^2 + x_2^2 = r^2 \ ,$$

which means that C is a circle.