Differential Geometry for GTD Applications

S. W. Lee

University of Illinois
Urbana, Illinois

Abstract

In applying GTD (ray techniques) to electromagnetic diffraction problems, some elementary knowledge of differential geometry is necessary. This report is written for those who are not familiar with this subject and wish to acquire a working knowledge in a rapid fashion. For more advanced readers, the report may provide a convenient collection of formulas in differential geometry relevant to GTD applications.

Acknowledgment This work was supported by the National Science Foundation under Grant No. NSF ENG 73-08218.

Preface

The material presented in this report was originally prepared for the Appendix of a monograph on GTD (which explains the letter "A" in section and figure numbers*. A small number of ditto copies of the Appendix were circulated in 1976. The monograph is far from completion. A surprisingly many requests have been received for the Appendix. Hence the Appendix is published herein as a technical report. The author appreciates comments and responses from readers.

*Note on Retyping: The letter "A" prefix used in the original document's section and figure numbering (e.g., A.1, Figure A-1) has been removed in this retyped manuscript for simpler referencing.

Contents

List of Figures

1 Introduction

For understanding and application of ray techniques in electromagnetics, it is necessary for the reader to have some elementary knowledge of differential geometry of curves and surfaces. This report is written for those readers who lack this knowledge. For easier comprehension of the subject, we will use ample examples and illustrations, while a few abstract concepts and proofs are omitted. In the first part (Sections 2–6) we discuss curves, and in the second part (Sections 7–12), surfaces. Key formulas and results are summarized in Section 13. All the materials, except for part of Sections 10 and 11, can be also found in standard textbooks of differential geometry. *

*We recommend two books: B. O'Neill, *Elementary Differential Geometry*, Academic Press, New York, 1966; and D. J. Struik, *Differential Geometry*, 2nd Edition, Addison-Wesley Publishing Co., Reading, Mass., 1961.

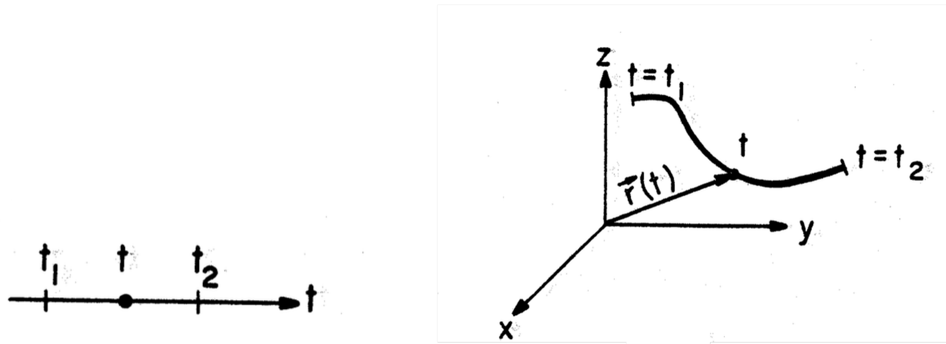
2 Representation of Curves

A curve may be pictured as a trip taken by a point in motion. Let us first concern ourselves with the description of the position of a point in three-dimensional space. In terms of the Cartesian coordinates (x, y, z) of the point, a position vector $\vec{r}$ is defined by

$$\vec{r} = x\hat{x} + y\hat{y} + z\hat{z} \tag{2.1}$$

where $(\hat{x}, \hat{y}, \hat{z})$ are unit vectors in the direction of the increasing (x, y, z), respectively. When the point is in motion, the locus traced out by the tip of the vector is a curve (Figure 1) and can be expressed as a vector $\vec{r}$ function of a parameter t in some open interval:

$$\vec{r}(t) = (x(t), y(t), z(t)); \quad t_1 < t < t_2 \tag{2.2}$$



(a) (b)

Figure 1: Curve defined by the tip of a vector in motion.

For engineers and physicists, it is convenient to think of t as the time, and we will use this association throughout this report. Let us now consider several examples of curves.

(i) **Straight lines.** The simplest curve in three-dimensional space is a straight line given by the equation

$$\vec{r}(t) = \vec{a} + \vec{b}t = (a_1 + b_1 t, a_2 + b_2 t, a_3 + b_3 t). \tag{2.3a}$$

where $\vec{a}$ and $\vec{b}$ are constant vectors, and $\vec{b}$ is not identically zero. Let the angles between the line and the three rectangular coordinate axes be $(\theta_1, \theta_2, \theta_3)$. Then the direction cosines of the line are given by

$$\cos\theta_n = \frac{b_n}{(b_1^2 + b_2^2 + b_3^2)^{1/2}}, \quad n = 1, 2, 3$$

Alternatively, (2.3a) may be written as

$$\begin{cases} y & = a_2 + (b_2/b_1)(x - a_1) \\ z & = a_3 + (b_3/b_1)(x - a_1) \end{cases} \tag{2.3b}$$

provided that $b_1 \neq 0$.

(ii) **Circular helix.** A point travels in the $x - y$ plane around a circle of radius a and rises along the z-direction at a constant speed b. Its trip is a circular helix:

$$\vec{r}(t) = (a\cos t, a\sin t, bt) \tag{2.4}$$

When $b > 0$, the helix is right-handed (Figure 2); when $b < 0$, the helix is left-handed.

(iii) **Conics.** Let us construct a plane curve C as follows: In the xz-plane, let L be a straight line (directrix) parallel to the x-axis, and F be a point (focus) on the z-axis (Figure 3). A typical point P on C satisfies the condition

$$\overline{PL} = e \cdot \overline{PF} \tag{2.5}$$

where $\overline{PL}$ is the distance from P to the straight line, and $\overline{PF}$ is the distance to F. The proportional constant e in (2.5), called **eccentricity**, is a positive real number. Curve C is a

- **parabola**, if $e = 1$
- **ellipse**, if $e < 1$
- **hyperbola**, if $e > 1$

The above three curves are known as conics, because they can be obtained as a section of a circular cone by a plane. We list below representations of these curves.

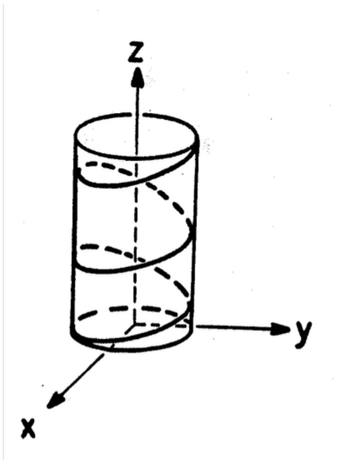


Figure 2: A right-handed circular helix.

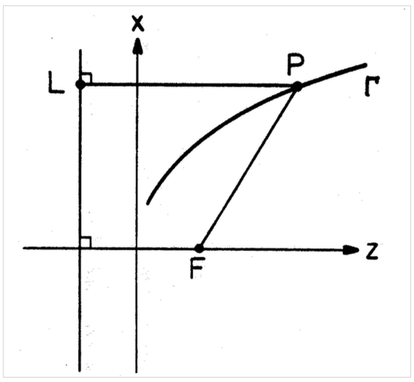


Figure 3: Construction of a conic Γ.

Parabola. (Figure 4).

$$\vec{r} = (2ft, 0, ft^2) \tag{2.6a}$$

$$z = \frac{x^2}{4f} \tag{2.6b}$$

$$R = 2f(1 - \cos\psi)^{-1}, \text{ for } R \geq 0 \tag{2.6c}$$

$$\text{eccentricity: } e = 1 \tag{2.7}$$

$$\text{foci: } (x = 0,\, z = f) \tag{2.8}$$

Ellipse (Figure 5).

$$\vec{r} = (a\cos t, 0, b\sin t) \tag{2.9a}$$

$$\left(\frac{x}{a}\right)^2 + \left(\frac{z}{b}\right)^2 = 1 \tag{2.9b}$$

$$R = a^2[b + \sqrt{b^2 - a^2}\cos\psi]^{-1}, \text{ assuming } b > a \tag{2.9c}$$

$$\text{eccentricity: } e = \sqrt{1 - (a/b)^2} \tag{2.10}$$

$$\text{foci: } (x = 0, z = \pm f), \text{ where } f = \sqrt{b^2 - a^2} \tag{2.11}$$

Hyperbola (Figure 6).

$$\vec{r} = (a\sinh t, 0, b\cosh t) \tag{2.12a}$$

$$\left(\frac{z}{b}\right)^2 - \left(\frac{x}{a}\right)^2 = 1 \tag{2.12b}$$

$$R = a^2[\pm b - \sqrt{a^2 + b^2}\cos\psi]^{-1} \tag{2.12c}$$

$$\text{eccentricity: } e = \sqrt{1 + (a/b)^2} \tag{2.13}$$

$$\text{foci: } (x = 0, z = \pm f), \text{ where } f = \sqrt{a^2 + b^2} \tag{2.14}$$

In (2.12c), the plus (minus) sign applies to the right- (left-) half hyperbola.

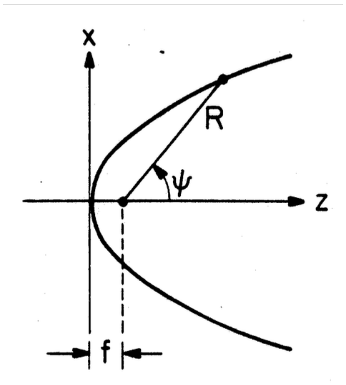


Figure 4: Parabola.

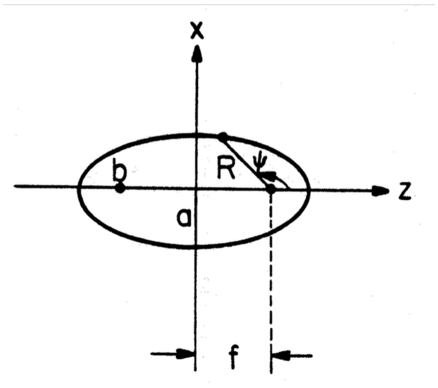


Figure 5: Ellipse.

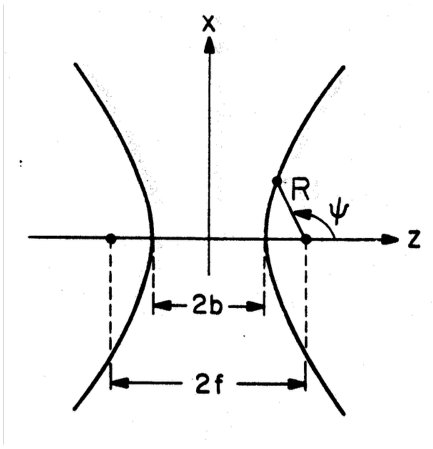


Figure 6: Hyperbola.

3 Tangent, Arc Length, and Reparameterization

Consider two points moving in space following an identical trajectory. We would say that they generate two identical curves. However, they may move at different speeds. To account for this difference, we will introduce a **tangent vector**. For a curve $\vec{r} = \vec{r}(t)$, its tangent vector at the point $\vec{r}(t)$ is the velocity vector

$$\frac{d\vec{r}}{dt} = \left(\frac{dx(t)}{dt}, \frac{dy(t)}{dt}, \frac{dz(t)}{dt}\right) \tag{3.1}$$

evaluated at the instant t. Its **speed** is the magnitude of the velocity vector:

$$v(t) = \left|\frac{d\vec{r}}{dt}\right| = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}. \tag{3.2}$$

A curve with $v(t) \equiv 1$ for all t is called a **unit-speed curve**; otherwise, it is an **arbitrary-speed curve**. Later, for a given curve with one speed, we may construct many "new" curves which have the same trajectory as the original one but travel at different speeds. For applications in electromagnetic diffraction problems, we often use unit-speed curves.

The distance travelled by a moving point is the **arc length** along a curve. In physics, the differential distance $d\sigma$ is equal to the product of the speed and the time interval: $d\sigma = v\, dt$. Thus, we define the arc length of a curve $\vec{r} = \vec{r}(t)$ from a reference instant t_0 to a variable instant t as

$$\sigma(t) = \int_{t_0}^{t} \left|\frac{d\vec{r}}{dt}\right| dt. \tag{3.3}$$

Clearly, $\sigma(t = t_0) = 0$, and $\sigma(t)$ can be positive or negative depending on whether $t > t_0$ or $t < t_0$.

Consider the circular helix in (2.4) as an example. Its velocity (or tangent) vector is

$$\frac{d\vec{r}}{dt} = (-a\sin t, a\cos t, b) \tag{3.4}$$

and its speed is

$$\left|\frac{d\vec{r}}{dt}\right| = \sqrt{\frac{d\vec{r}}{dt} \cdot \frac{d\vec{r}}{dt}} = \sqrt{a^2 + b^2} = c. \tag{3.5}$$

Assuming $c \neq 1$, then the helix, as given in (2.4), is an arbitrary-speed curve. Its arc length, measuring from $t = 0$, is

$$\sigma(t) = \int_0^t c\, dt = ct. \tag{3.6}$$

At this point, many readers may have already realized that a simple change of variable $t = \sigma/c$ in (2.4) gives rise to an equation $\vec{r} = \vec{r}(\sigma)$ which is a unit-speed curve. We have not changed the helix curve itself, we have changed its speed by passing from one parameter to another. Such a process is called **reparametrization**. Clearly, as many reparametrizations of a given curve exist as there are transformations for parameter t.

Of all reparametrizations, we are particularly interested in the one that results in a unit-speed curve. It can be shown that, for a curve $\vec{r} = \vec{r}(t)$, the reparametrization $\vec{r} = \vec{r}(\sigma)$ with

$$t = t(\sigma) \tag{3.7}$$

describes a unit-speed curve. In other words, a unit-speed curve is a curve whose arc length is its parameter.

4 Frenet Formula for Unit-Speed Curves

For a unit-speed curve $\vec{r} = \vec{r}(\sigma)$, the unit **tangent vector**

$$\hat{t} = \frac{d\vec{r}}{d\sigma} \tag{4.1}$$

indicates the direction of turning along the curve. Now, let us introduce two new parameters: (i) **curvature** κ, which measures the rate of turning; and (ii) **torsion** τ, which measures the rate of twisting. Consider the derivative of the unit tangent vector:

$$\frac{d\hat{t}}{d\sigma} = \frac{d^2\vec{r}}{d\sigma^2} = \kappa\hat{n}. \tag{4.2}$$

Here the unit vector $\hat{n}$ is in the direction of $d\hat{t}/d\sigma$ and is called **normal**. Since $\hat{t}\cdot\hat{t} = 1$, differentiation of this identity gives

$$\hat{t}\cdot\frac{d\hat{t}}{d\sigma} = 0, \quad \text{or} \quad \hat{t}\cdot\hat{n} = 0. \tag{4.3}$$

Hence $\hat{n}$ is orthogonal to $\hat{t}$.* The magnitude of $d\hat{t}/d\sigma$ is curvature κ, which by definition is nonnegative. As the curvature increases, the turning of the curve becomes sharper. From (4.2) we can deduce an alternative formula for curvature, namely,

$$\kappa(\sigma) = +\left(\frac{d^2\vec{r}}{d\sigma^2}\cdot\frac{d^2\vec{r}}{d\sigma^2}\right)^{1/2} \tag{4.4}$$

where the square root should take a nonnegative value.

At each point on a curve we have two orthonormal vectors $\hat{t}$ and $\hat{n}$. Define a third one, **binormal** $\hat{b}$, such that

$$\hat{b} = \hat{t}\times\hat{n}. \tag{4.5}$$

Then $(\hat{n}, \hat{b}, \hat{t})$ form a right-handed orthonormal basis for the three-dimensional space.† In general, they vary continuously along the curve, according to the turning and twisting of the curve. For this reason $(\hat{n}, \hat{b}, \hat{t})$ are known as the **moving trihedron**. In the study of the geometry of a curve, it is often more convenient to use $(\hat{n}, \hat{b}, \hat{t})$ as the base vectors instead of $(\hat{x}, \hat{y}, \hat{z})$, because the former contains information about the curve while the latter does not. The planes spanned by $(\hat{n}, \hat{b})$, $(\hat{b}, \hat{t})$, and $(\hat{t}, \hat{n})$ are called **normal plane**, **rectifying plane**, and **osculating plane**, respectively (Figure 7).

One of the most important applications of the moving trihedron of a curve concerns the expression of the derivatives of $(\hat{n}, \hat{b}, \hat{t})$. Consider first $d\hat{b}/d\sigma$, which measures the rate of change of the osculating plane. The differentiation of $\hat{b}\cdot\hat{t} = 0$ leads to

$$\frac{d\hat{b}}{d\sigma}\cdot\hat{t} = -\hat{b}\cdot\frac{d\hat{t}}{d\sigma} = -\kappa\hat{b}\cdot\hat{n} = 0 \tag{4.6}$$

where we have made use of (4.2). The differentiation of $\hat{b}\cdot\hat{b} = 1$ gives

$$\frac{d\hat{b}}{d\sigma}\cdot\hat{b} = 0. \tag{4.7}$$

*Since the curve has a constant speed (unit speed), the "acceleration" $d^2\vec{r}/d\sigma^2$ must be orthogonal to its velocity vector.

†Usually, the three vectors are written in the order of $(\hat{t}, \hat{n}, \hat{b})$. For our application, the order $(\hat{n}, \hat{b}, \hat{t})$ is preferred.

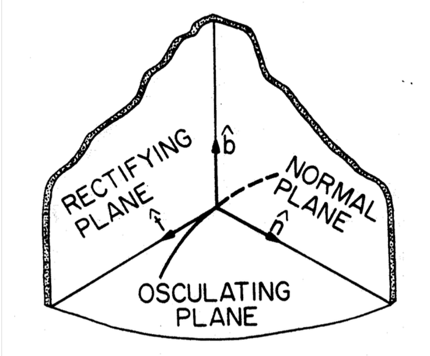


Figure 7: Moving trihedron of a curve..

From (4.6) and (4.7) we know that $d\hat{b}/d\sigma$ is orthogonal to both $\hat{b}$ and $\hat{t}$ and, therefore, in the direction of $\hat{n}$. Let us write it as

$$\frac{d\hat{b}}{d\sigma} = -\tau\hat{n} \tag{4.8}$$

where the **torsion** τ can be either positive or negative. For a planar curve, $\hat{b}(\sigma)$ points to a constant direction and, consequently, $\tau(\sigma) = 0$. Thus, τ measures the twisting of a curve from a planar one. Similar to the curvature formula in (4.4), we will derive a more explicit formula for τ. Starting with (4.8), we have

$$\begin{aligned}
\tau(\sigma) &= -\hat{n}\cdot\frac{d\hat{b}}{d\sigma} = -\hat{n}\cdot\frac{d}{d\sigma}(\hat{t}\times\hat{n}) \\
&= -\hat{n}\cdot\left(\hat{t}\times\frac{d\hat{n}}{d\sigma}\right) \\
&= -\frac{1}{\kappa}\frac{d^2\vec{r}}{d\sigma^2}\cdot\left(\frac{d\vec{r}}{d\sigma}\times\frac{d}{d\sigma}\frac{1}{\kappa}\frac{d^2\vec{r}}{d\sigma^2}\right) \\
&= \frac{1}{\kappa^2}\frac{d\vec{r}}{d\sigma}\cdot\frac{d^2\vec{r}}{d\sigma^2}\times\frac{d^3\vec{r}}{d\sigma^3}
\end{aligned}$$

or, when (4.4) is used,

$$\tau(\sigma) = \frac{d\vec{r}}{d\sigma}\cdot\left(\frac{d^2\vec{r}}{d\sigma^2}\times\frac{d^3\vec{r}}{d\sigma^3}\right)\left(\frac{d^2\vec{r}}{d\sigma^2}\cdot\frac{d^2\vec{r}}{d\sigma^2}\right)^{-1} \tag{4.9}$$

Next let us consider the representation of $d\hat{n}/d\sigma$ in terms of the moving trihedral. Its general form is

$$\frac{d\hat{n}}{d\sigma} = \left(\frac{d\hat{n}}{d\sigma}\cdot\hat{n}\right)\hat{n} + \left(\frac{d\hat{n}}{d\sigma}\cdot\hat{b}\right)\hat{b} + \left(\frac{d\hat{n}}{d\sigma}\cdot\hat{t}\right)\hat{t}. \tag{4.10}$$

Differentiation of $\hat{n}\cdot\hat{n} = 1$ gives

$$\frac{d\hat{n}}{d\sigma}\cdot\hat{n} = 0. \tag{4.11}$$

Differentiation of $\hat{n}\cdot\hat{b} = 0$ gives

$$\frac{d\hat{n}}{d\sigma}\cdot\hat{b} = -\hat{n}\cdot\frac{d\hat{b}}{d\sigma} = -\hat{n}\cdot(-\tau\hat{n}) = \tau. \tag{4.12}$$

Differentiation of $\hat{n}\cdot\hat{t} = 0$ gives

$$\frac{d\hat{n}}{d\sigma}\cdot\hat{t} = -\hat{n}\cdot\frac{d\hat{t}}{d\sigma} = -\hat{n}\cdot(\kappa\hat{n}) = -\kappa. \tag{4.13}$$

Substituting (4.11), (4.12), and (4.13) into (4.10) we have

$$\frac{d\hat{n}}{d\sigma} = \tau\hat{b} - \kappa\hat{t}. \tag{4.14}$$

Combining (4.2), (4.8), and (4.14), we have a set of equations describing the motion of the moving trihedron along a unit-speed curve:

$$\frac{d\hat{n}}{d\sigma} = \tau\hat{b} - \kappa\hat{t} \tag{4.15a}$$

$$\frac{d\hat{b}}{d\sigma} = -\tau\hat{n} \tag{4.15b}$$

$$\frac{d\hat{t}}{d\sigma} = \kappa\hat{n} \tag{4.15c}$$

Equation (4.15) is known as **Frenet formula** or **Serret-Frenet formula**, which was independently derived by F. Frenet and J. A. Serret around 1850.

In summary, at each point σ on a unit-speed curve $\vec{r} = \vec{r}(\sigma)$, there are five important fields: (i) the moving trihedron $(\hat{n}, \hat{b}, \hat{t})$, which may be computed from (4.2), (4.5), and (4.1); and, (ii) the curvature κ and the torsion τ, which may be computed either from (4.2) and (4.8), or more directly from (4.4) and (4.9). The variations of $(\hat{n}, \hat{b}, \hat{t})$ along the curve are described by the Frenet formula in (4.15).

Let us give an example to illustrate the computation of the Frenet apparatus. In terms of the arc length σ, a unit-speed circular helix has the representation

$$\vec{r}(\sigma) = \left(a\cos\frac{\sigma}{c}, a\sin\frac{\sigma}{c}, \frac{b}{c}\sigma\right) \tag{4.16}$$

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

$$\hat{t}(\sigma) = \frac{d\vec{r}}{d\sigma} = \left(-\frac{a}{c}\sin\frac{\sigma}{c}, \frac{a}{c}\cos\frac{\sigma}{c}, \frac{b}{c}\right) \tag{4.17}$$

and

$$\frac{d\hat{t}}{d\sigma} = \left(-\frac{a}{c^2}\cos\frac{\sigma}{c}, -\frac{a}{c^2}\sin\frac{\sigma}{c}, 0\right).$$

Recalling (4.15c) we have

$$\kappa(\sigma) = \left|\frac{d\hat{t}}{d\sigma}\right| = \frac{a}{c^2} \tag{4.18}$$

$$\hat{n}(\sigma) = \left(-\cos\frac{\sigma}{c}, -\sin\frac{\sigma}{c}, 0\right). \tag{4.19}$$

Note that $\hat{n}$ always points straight to the axis of the cylinder on which the helix lies. As sketched in Figure 8, the osculating plane determined by $\hat{t}$ and $\hat{n}$ is formed by wiggle lines. The binormal is

$$\hat{b}(\sigma) = \hat{t} \times \hat{n} = \left(\frac{b}{c}\sin\frac{\sigma}{c}, -\frac{b}{c}\cos\frac{\sigma}{c}, \frac{a}{c}\right) \tag{4.20}$$

and its derivative is

$$\frac{d\hat{b}}{d\sigma} = \left(\frac{b}{c^2}\cos\frac{\sigma}{c}, \frac{b}{c^2}\sin\frac{\sigma}{c}, 0\right). \tag{4.21}$$

From (4.15b), it follows immediately that

$$\tau(\sigma) = \left|\frac{d\hat{b}}{d\sigma}\right| = \frac{b}{c^2}$$

which is positive for a right-handed helix ($b > 0$), and negative for a left-handed helix ($b < 0$).

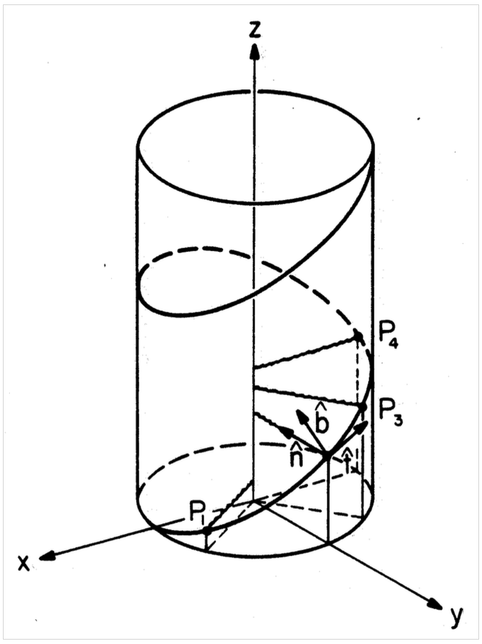


Figure 8: Moving trihedron of a right-handed circular helix.

5 Frenet Formula for Arbitrary-Speed Curves

The discussion in the previous section applies to unit-speed curves $\vec{r} = \vec{r}(\sigma)$ with σ being the arc length. For an arbitrary-speed curve $\vec{r} = \vec{r}(t)$, its speed

$$v = \left|\frac{d\vec{r}}{dt}\right| = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} \tag{5.1}$$

is not necessarily unit. In order to use the formulas in Section 4 and also those to be presented in Section 6, we may exercise the arc length reparametrization by letting

$$t = t(\sigma) \tag{5.2}$$

and obtain a new representation

$$\vec{r} = \vec{r}(t = t(\sigma)) = \vec{R}(\sigma) \tag{5.3}$$

For example, in the circular helix discussed in (2.4) and (4.16), we have

$$\vec{r}(t) = (a\cos t, a\sin t, bt)$$
$$\vec{R}(\sigma) = \left(a\cos\frac{\sigma}{c}, a\sin\frac{\sigma}{c}, b\frac{\sigma}{c}\right).$$

When there is no confusion, we will write $\vec{r}(\sigma)$ as $\vec{r}(\sigma)$

$$\vec{r}(\sigma) = \left(a\cos\frac{\sigma}{c}, a\sin\frac{\sigma}{c}, b\frac{\sigma}{c}\right).$$

When this convention is used, we should remember that $\vec{r}(\sigma)$ is obtained from $\vec{r}(t)$ via the transform in (5.2), but not by a simple substitution $t = \sigma$. After the arc length reparametrization, those formulas in Sections 4 and 6 can be applied to $\vec{r} = \vec{r}(\sigma)$.

Unfortunately, an explicit expression for (5.2) cannot always be found. Consider for example the curve

$$\vec{r}(t) = (1 + \cos t, \sin t, 2\sin(t/2)), \quad |t| < 2\pi \tag{5.4}$$

whose arc length, measured from $t = 0$, is

$$\sigma(t) = \int_0^t \sqrt{1 + \cos^2(t/2)}\, dt. \tag{5.5}$$

Thus, an explicit expression in the form of (5.2) is not available. For these cases, we will give an alternative set of formulas for computing the Frenet apparatus.

Note the basic relations of differentiation for $\vec{r}(t)$ and $\vec{r}(\sigma)$ (or more precisely $\vec{R}(\sigma)$):[‡]

$$\frac{d\vec{r}(t)}{dt} = \frac{d\sigma}{dt}\frac{d\vec{r}(\sigma)}{d\sigma} = v\frac{d\vec{r}(\sigma)}{d\sigma} = v\hat{t} \tag{5.6}$$

$$\begin{aligned}\frac{d^2\vec{r}(t)}{dt^2} &= \frac{d}{dt}\left(v\hat{t}\right) = \frac{dv}{dt}\hat{t} + v\frac{d\hat{t}}{dt} \\ &= \frac{dv}{dt}\hat{t} + v\frac{d\sigma}{dt}\frac{d\hat{t}}{d\sigma} \\ &= \frac{dv}{dt}\hat{t} + \kappa v^2\hat{n}\end{aligned} \tag{5.7}$$

[‡]In mechanics, $d^2\vec{r}/dt^2$ in (5.7) is the acceleration of a moving point. It has two orthogonal components: one along the tangential direction describing the change of speed, and the other along normal direction describing the change of direction of motion. For a unit-speed curve, $v \equiv 1$ and the tangential acceleration is identically zero as given in (4.3). Also, note that $\hat{t}$ is the tangent vector and is not a vector in the direction of increasing t.

Then straightforward manipulations lead to the following formulas for an arbitrary-speed curve $\vec{r} = \vec{r}(t)$:

$$\hat{n}(t) = \hat{b} \times \hat{t} \tag{5.8}$$

$$\hat{b}(t) = \frac{\vec{r}'(t) \times \vec{r}''(t)}{|\vec{r}'(t) \times \vec{r}''(t)|} \tag{5.9}$$

$$\hat{t}(t) = \frac{\vec{r}'(t)}{|\vec{r}'(t)|} \tag{5.10}$$

$$\kappa(t) = \frac{|\vec{r}'(t) \times \vec{r}''(t)|}{|\vec{r}'(t)|^3} \tag{5.11}$$

$$\tau(t) = \frac{(\vec{r}'(t) \times \vec{r}''(t)) \cdot \vec{r}'''(t)}{|\vec{r}'(t) \times \vec{r}''(t)|^2} \tag{5.12}$$

where

$$\vec{r}'(t) = \frac{d\vec{r}(t)}{dt}, \quad \vec{r}''(t) = \frac{d^2\vec{r}(t)}{dt^2}, \quad \vec{r}'''(t) = \frac{d^3\vec{r}(t)}{dt^3} \tag{5.13}$$

The symbol $|\cdot|$ indicates the magnitude of a vector.

Let us consider the computation of the Frenet apparatus for an arbitrary-speed representation of a circular helix. This representation is obtained from (2.4) by letting $t = t^2$,

$$\vec{r}(t) = (a\cos t^2, a\sin t^2, bt^2).$$

Its first three derivatives are

$$\vec{r}'(t) = (-2at\sin t^2, 2at\cos t^2, 2bt)$$

$$\vec{r}''(t) = (-2a(\sin t^2 + 2t^2\cos t^2), 2a(\cos t^2 - 2t^2\sin t^2), 2b)$$

$$\vec{r}'''(t) = (4at(-3\cos t^2 + 2t^2\sin t^2), 4at(-3\sin t^2 - 2t^2\cos t^2), 0).$$

Substitution of them into (5.8) through (5.12) leads to

$$\hat{n}(t) = (-\cos t^2, -\sin t^2, 0)$$

$$\hat{b}(t) = \left(\frac{b}{c}\sin t^2, -\frac{b}{c}\cos t^2, \frac{a}{c}\right)$$

$$\hat{t}(t) = \left(-\frac{a}{c}\sin t^2, \frac{a}{c}\cos t^2, \frac{b}{c}\right)$$

$$\kappa(t) = \frac{a}{c^2}, \quad \tau(t) = \frac{b}{c^2}.$$

Because $\sigma = ct^2$, the above results are exactly the same as those given in (4.17) through (4.21). Thus, the Frenet apparatus depends on the shape of the curve in space, not on its speed. In other words, reparametrization does **not** affect the Frenet apparatus.

6 Approximation of Unit-Speed Curves

A fundamental theorem of a unit-speed curve is stated below: Except for its position in space, a unit-speed curve is uniquely determined by its curvature $\kappa(\sigma)$ and torsion $\tau(\sigma)$. In other words, for two curves $\vec{r} = \vec{r}_1(\sigma)$ and $\vec{r} = \vec{r}_2(\sigma)$ with $\kappa_1(\sigma) = \kappa_2(\sigma)$ and $\tau_1(\sigma) = \pm\tau_2(\sigma)$ for all σ, these two curves are the same except possibly for their positions in space. The proof is simple. Concentrate on an arbitrary point $\sigma = \sigma_0$ on a unit-speed curve $\vec{r} = \vec{r}(\sigma)$. In its neighborhood, the curve can be represented by the Taylor series:

$$\vec{r}(\sigma) = \vec{r}(\sigma_0) + \frac{(\sigma-\sigma_0)}{1!}\left.\frac{d\vec{r}}{d\sigma}\right|_{\sigma_0} + \frac{(\sigma-\sigma_0)^2}{2!}\left.\frac{d^2\vec{r}}{d\sigma^2}\right|_{\sigma_0} + \frac{(\sigma-\sigma_0)^3}{3!}\left.\frac{d^3\vec{r}}{d\sigma^3}\right|_{\sigma_0} + \ldots \tag{6.1}$$

Note that, in terms of a moving trihedron, all the derivatives of $\vec{r}(\sigma)$ depend solely on $\kappa(\sigma)$ and $\tau(\sigma)$:

$$\frac{d\vec{r}}{d\sigma} = \hat{t} \tag{6.2}$$

$$\frac{d^2\vec{r}}{d\sigma^2} = \kappa\hat{n} \tag{6.3}$$

$$\frac{d^3\vec{r}}{d\sigma^3} = \frac{d\kappa}{d\sigma}\hat{n} + \kappa\tau\hat{b} - \kappa^2\hat{t} \tag{6.4}$$

$$\ldots\ldots$$

Thus, the Taylor series for the curve $\vec{r} = \vec{r}(\sigma)$, namely,

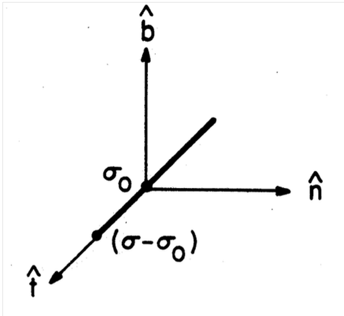
$$\vec{r}(\sigma) = \vec{r}(\sigma_0) + (\sigma-\sigma_0)\hat{t} + \frac{(\sigma-\sigma_0)^2}{2}\kappa\hat{n} + \frac{(\sigma-\sigma_0)^3}{6}\left(\frac{d\kappa}{d\sigma}\hat{n} + \kappa\tau\hat{b} - \kappa^2\hat{t}\right) + \ldots \tag{6.5}$$

is uniquely determined, except for its position in space, by $\kappa(\sigma)$ and $\tau(\sigma)$.

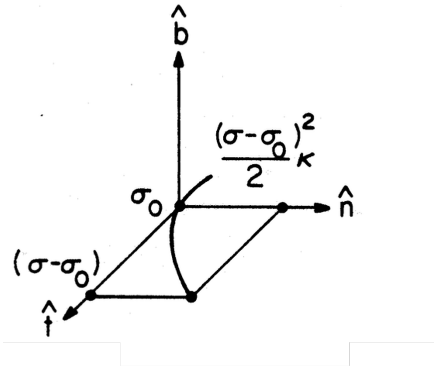
In many applications, a few terms in the Taylor series in (6.5) may be used to represent curve $\vec{r} = \vec{r}(\sigma)$ in a small neighborhood of $\sigma = \sigma_0$. If only the first two terms of (6.5) are used, we have a **linear approximation** (Figure 9a). If three terms are used, we have a **parabolic approximation** (Figure 9b), which lies in the osculating plane, and is determined by the curvature κ at $\sigma = \sigma_0$. If four terms are used, we have

$$\vec{r}(\sigma) \approx \vec{r}(\sigma_0) + (\sigma-\sigma_0)\hat{t} + \frac{(\sigma-\sigma_0)^2}{2}\kappa\hat{n} + \frac{(\sigma-\sigma_0)^3}{6}\kappa\tau\hat{b} \tag{6.6}$$

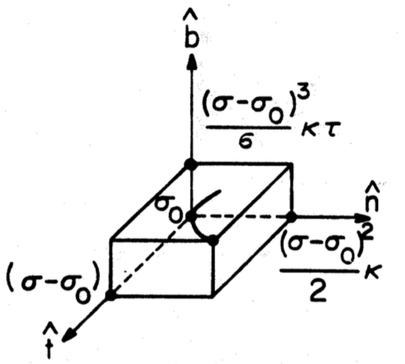
which is known as the **Frenet approximation** (Figure 9c). The torsion, which appeared in the last term of (6.6), controls the deviation from the osculating plane. The geometrical significance of the sign of τ can now be stated: If $\tau > 0$, the curve with increasing σ cuts through the osculating plane in the direction of the binormal $\hat{b}$, and if $\tau < 0$, in the opposite direction. If $\tau = 0$, no conclusion can be drawn. In that case, we have to study the higher-order terms in (6.5).



(a) linear approximation



(b) parabolic approximation



(c) Frenet approximation

Figure 9: Approximation of a curve near $\sigma = \sigma_0$, assuming $\tau > 0$.

7 Representation of Surfaces

A curve may be considered as a locus of points traced out by the tip of a vector $\vec{r}(t)$ when the parameter t takes values in an open interval $t_1 < t < t_2$ (Figure 1). Analogously, a **surface** is a locus of the tip of a vector $\vec{r}(u,v)$ depending now on two parameters, when (u,v) vary in a two-dimensional domain $u_1 < u < u_2$, $v_1 < v < v_2$, as graphically sketched in Figure 10. Corresponding to (2.2) for a curve we have a representation of a surface

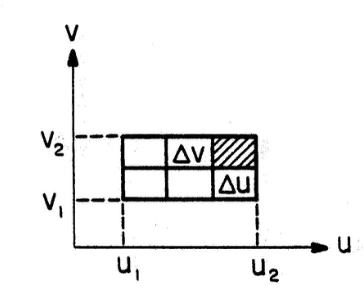
$$\vec{r}(u,v) = (x(u,v), y(u,v), z(u,v)); \quad u_1 < u < u_2 \text{ and } v_1 < v < v_2. \tag{7.1}$$

We call (u,v) the **curvilinear coordinates** of a point $\vec{r}$ on the surface. By holding v constant ($v = v_0$), $\vec{r}(u,v_0)$ defines a curve on the surface, called a u-parameter curve; by holding u constant ($u = u_0$), $\vec{r}(u_0,v)$ defines a v-parameter curve (Figure 11). At (u_0,v_0), the tangent (velocity) vectors along u- and v-parameter curves are denoted by $\vec{r}_u(u_0,v_0)$ and $\vec{r}_v(u_0,v_0)$, respectively, where

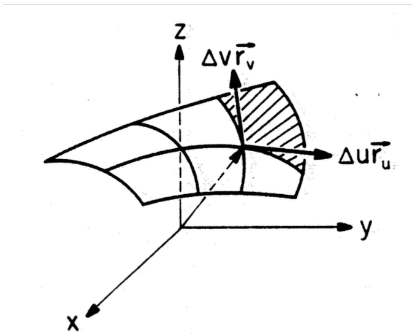
$$\vec{r}_u(u_0,v_0) = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right)_{u=u_0, v=v_0} \tag{7.2}$$

$$\vec{r}_v(u_0,v_0) = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right)_{u=u_0, v=v_0} \tag{7.3}$$

Thus, the subscript u of $\vec{r}_u$, for example, indicates the partial derivative of $\vec{r}$ with respect to u.



(a)



(b)

Figure 10: Surface defined by the tip o f a vector in motion.

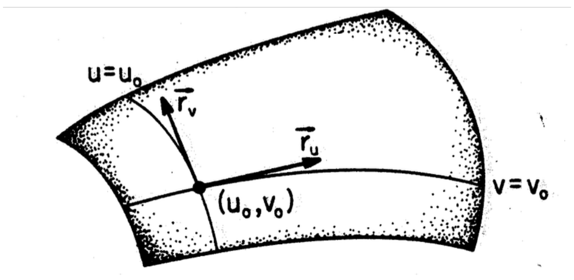


Figure 11: Veolcity vectors for $u-$ and $v-$ parameter curves.

Suppose that the inverse functions: $u = u(x,y)$, $v = v(x,y)$ can be found. Then we may use (x,y) instead of (u,v) as the parameters of a surface. Thus, an alternative form for a surface is

$$\vec{r} = \vec{r}(x,y) = (x, y, f(x,y)) \tag{7.4}$$

or simply

$$z = f(x,y), \quad \text{or} \quad g(x,y,z) = 0. \tag{7.5}$$

The two forms in (7.4) and (7.5) are often used in elementary calculus.

A familiar surface is a sphere which may be represented by

$$\vec{r}(u,v) = a(\sin u \cos v, \sin u \sin v, \cos u).$$

Clearly, the parameters (u,v) can be identified with (θ,ϕ) of the usual spherical coordinates. The tangent vectors are

$$\vec{r}_u(u,v) = a(\cos u \cos v, \cos u \sin v, -\sin u)$$

$$\vec{r}_v(u,v) = a(-\sin u \sin v, \sin u \cos v, 0)$$

with their cross product

$$\vec{r}_u \times \vec{r}_v = a^2 \sin u\,(\sin u \cos v, \sin u \sin v, \cos u) = (a \sin u)\,\vec{r}(u,v) \tag{7.6}$$

which is normal to the surface everywhere. If we express (u,v) in terms of (x,y), then an alternative representation of the sphere is

$$\vec{r} = (x, y, \sqrt{a^2 - (x^2+y^2)})$$

or

$$x^2 + y^2 + z^2 = a^2$$

which are in the forms of (7.4) and (7.5). Its tangent vectors are

$$\vec{r}_x(x,y) = \left(1, 0, \frac{-x}{\sqrt{a^2-(x^2+y^2)}}\right)$$

$$\vec{r}_y(x,y) = \left(0, 1, \frac{-y}{\sqrt{a^2-(x^2+y^2)}}\right)$$

with their cross product

$$\vec{r}_x \times \vec{r}_y = \left(\frac{x}{\sqrt{a^2-(x^2+y^2)}}, \frac{y}{\sqrt{a^2-(x^2+y^2)}}, 1\right) = \frac{1}{\sqrt{a^2-(x^2+y^2)}}\vec{r}(x,y) \tag{7.7}$$

which is again normal to the surface.

An elementary problem in calculus is to determine the area of a surface. Referring to Figure 10, we note that a differential rectangle $(\Delta u \times \Delta v)$ in (a) is mapped into a differential parallelogram with sides $\Delta u\, \vec{r}_u$ and $\Delta v\, \vec{r}_v$ in (b). The area of this differential parallelogram is

$$|\Delta u\, \vec{r}_u \times \Delta v\, \vec{r}_v| = |\vec{r}_u \times \vec{r}_v| \Delta u \Delta v.$$

Then the area over a region D is given by

$$\text{Area} = \iint_D |\vec{r}_u \times \vec{r}_v| du\, dv. \tag{7.8}$$

For the example of the sphere discussed above, we have from (7.6)

$$\text{Area} = \iint_D a^2 \sin u\; du\, dv \tag{7.9}$$

or, from (7.7),

$$\text{Area} = \iint_D \frac{a}{\sqrt{a^2-(x^2+y^2)}}\,dx\,dy. \tag{7.10}$$

Both formulas of course lead to the same results.

Before concluding this section, we list below several frequently encountered surfaces.

(i) **Elliptical cone** (Figure 12).

$$\vec{r} = (z\tan\theta_1\cos v, z\tan\theta_2\sin v, z) \tag{7.11a}$$

$$\left(\frac{x}{\tan\theta_1}\right)^2 + \left(\frac{y}{\tan\theta_2}\right)^2 = z^2 \tag{7.11b}$$

where $\theta_1(\theta_2)$ is the half-cone angle in the plane $y=0$ $(x=0)$.

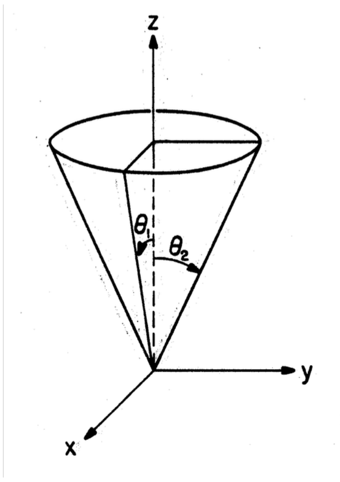


Figure 12: Elliptical cone.

(ii) **Elliptical paraboloid** (Figure 13).

$$\vec{r} = (au\cos v, bu\sin v, u^2) \tag{7.12a}$$

$$z = \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 \tag{7.12b}$$

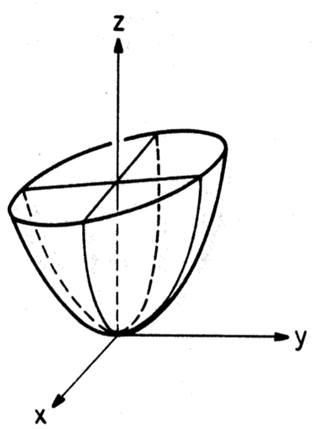


Figure 13: Elliptical paraboloid.

(iii) **Hyperbolic paraboloid** (Figure 14).

$$\vec{r} = (au\cosh v, bu\sinh v, u^2) \tag{7.13a}$$

$$z = \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 \tag{7.13b}$$

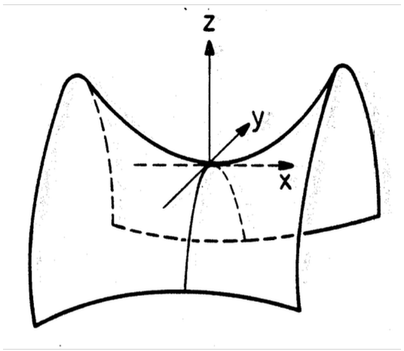


Figure 14: Hyperbolic paraboloid.

(iv) **Ellipsoid** (Figure 15).

$$\vec{r} = (a\cos u\cos v, b\cos u\sin v, c\sin u) \tag{7.14a}$$

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1 \tag{7.14b}$$

If $a = b$ and $a < c$, the ellipsoid has a rotational symmetry about its major axis, and it is said to be **prolate**. If $a = b$ and $a > c$, the ellipsoid has a rotational symmetry about its minor axis, and it is said to be **oblate**.

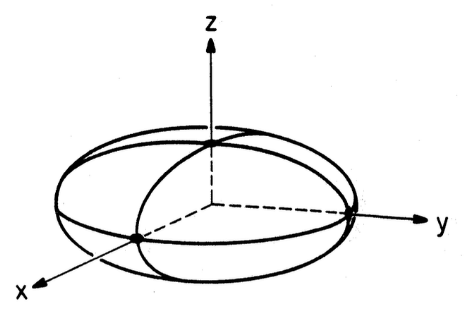


Figure 15: Ellipsoid.

(v) **Hyperboloid of one sheet** (Figure 16).

$$\vec{r} = (a\cosh u\cos v, b\cosh u\sin v, c\sinh u) \tag{7.15a}$$

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 - \left(\frac{z}{c}\right)^2 = 1 \tag{7.15b}$$

(vi) **Hyperboloid of two sheets** (Figure 17).

$$\vec{r} = (a\sinh u\cos v, b\sinh u\sin v, c\cosh u) \tag{7.16a}$$

$$\left(\frac{z}{c}\right)^2 - \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 1 \tag{7.16b}$$

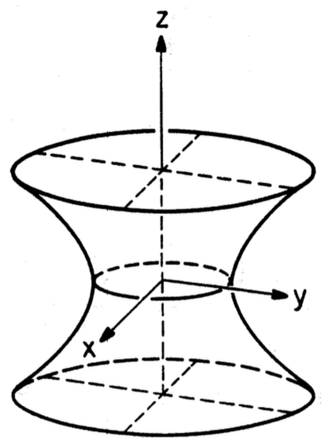


Figure 16: Hyperboloid of one sheet.

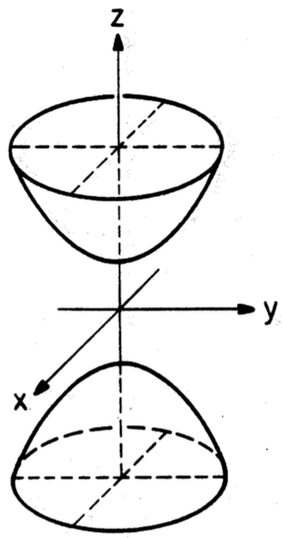


Figure 17: Hyperboloid of two sheets.

(vii) **Surface of revolution** (Figure 18). Consider a curve $z = f(\rho)$ lying in the plane $y = 0$. If this curve is rotated about the z-axis, it generates a surface of revolution, which may be represented by

$$x = \rho \cos\phi, \quad y = \rho \sin\phi, \quad z = f(\rho). \tag{7.17}$$

The curves $\rho =$ constant are the **parallels**, and the curves $\phi =$ constant are the **meridians** of the surface. A few examples are given below. For the special case $\theta_1 = \theta_2$, (7.17) describes a cone of revolution (circular cone). If $a = b$ in (7.12) through (7.16), all those surfaces become rotationally symmetrical. In fact, they may be generated by revolving conics (Section 2) about an axis.

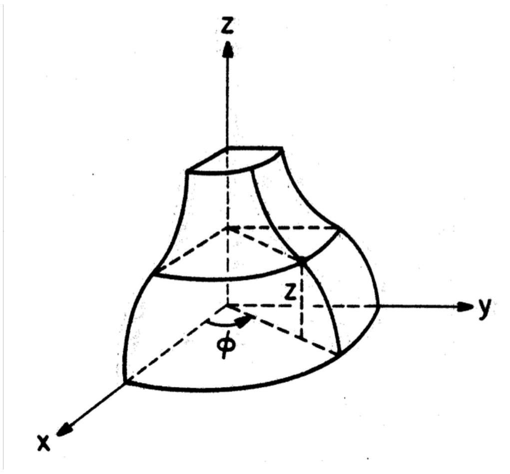


Figure 18: Surface of revolution.

8 Tangent Plane, Normal and Curvature

When the parameters (u, v) vary independently over a two-dimensional domain, the tip of the vector $\vec{r}(u, v)$ generates a surface (Figure 10). However, if (u, v) are not independent but vary according to a parameter t:

$$u = u(t), \quad v = v(t), \quad t_1 < t < t_2. \tag{8.1}$$

Then the tip of the vector $\vec{r}(u(t), v(t))$ or simply $\vec{r}(t)$ traces a curve on the surface. The tangent vector of the curve is given by

$$\frac{d\vec{r}(t)}{dt} = \vec{r}_u \frac{du}{dt} + \vec{r}_v \frac{dv}{dt} \tag{8.2}$$

where the vectors $\vec{r}_u$ and $\vec{r}_v$ were defined in (7.2) and (7.3). At any point P on the surface, the independent vectors $\vec{r}_u$ and $\vec{r}_v$ define a plane, called the **tangent plane** at P to the surface. The relation in (8.2) states that the tangents to all curves through P of the surface lie in the tangent plane (Figure 19). The **unit normal** at P of the surface is defined by

$$\hat{N} = \mu \frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|}, \quad \text{where } \mu = \pm 1 \tag{8.3}$$

(Note that we use capital $\hat{N}$ for the normal of a surface, and $\hat{n}$ for that of a curve.) The choice of the value of $\mu = \pm 1$ in (8.3) is arbitrary and can be made to suit the convenience of a particular problem. In application to EM diffraction problems, we always define the normal of a reflecting surface or a wavefront pointing toward the source.

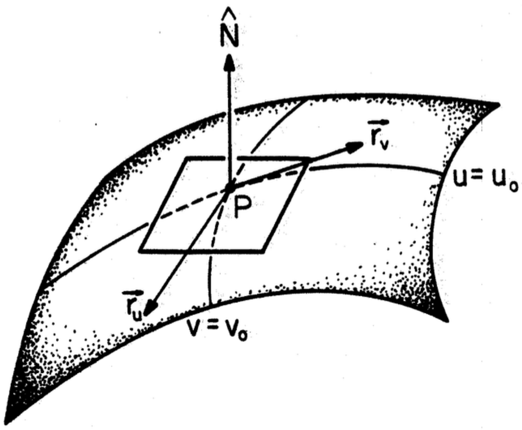


Figure 19: Tangent plane of a surface.

We will now study the bending of a surface by introducing a quantity of measurement called **normal curvature** or simply **curvature** κ[§]. At a point P on a surface Σ, let us consider a particular direction described by a unit tangent vector $\hat{t}$. The plane determined by $\hat{t}$ and the normal $\hat{N}$ cuts from Σ a planar curve C called the **normal section** of Σ in the direction of $\hat{t}$ (Figure 20). If we give a unit-speed reparametrization to planar curve C, denoted by $\vec{r} = \vec{r}(\sigma)$, then according to (4.2),

$$\frac{d^2\vec{r}}{d\sigma^2} = \kappa_c \hat{n}_c. \tag{8.4}$$

[§]We use the same symbol κ for the curvature of a surface and that of a curve. There is little chance of confusion.

The subscripts c of κ_c and $\hat{n}_c$ signify their association with planar curve C. By definition, κ_c is nonnegative. Since C lies in the plane spanned by $\hat{N}$ and $\hat{t}$, it is clear that $\hat{N} = \pm\hat{n}_c$. Now, we will define the curvature κ of surface Σ at point P in the direction $\hat{t}$ to be

$$\kappa(\hat{t}) = (\hat{n}_c \cdot \hat{N})\kappa_c. \tag{8.5}$$

Thus, $|\kappa| = |\kappa_c|$. The sign of $\kappa(\hat{t})$ has the following significance: (i) If $\kappa(\hat{t}) > 0$, then $\hat{N} = +\hat{n}_c$. The normal section C bends toward $\hat{N}$ at P (Figure 21a); (ii) If $\kappa(\hat{t}) < 0$, then $\hat{N} = -\hat{n}_c$. The normal section C bends away from $\hat{N}$ at P (Figure 21b).

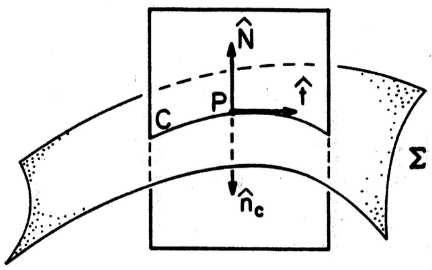
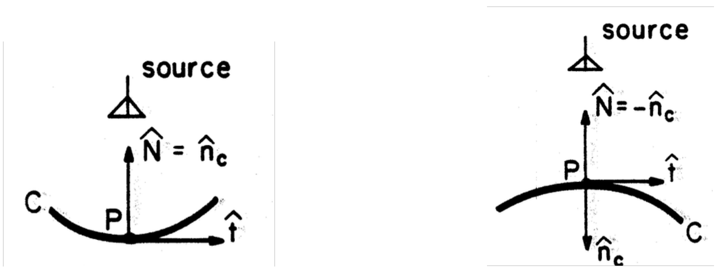


Figure 20: A normal section of surface Σ in the direction of $\hat{t}$.



(a) $\kappa > 0$ (convergent wavefront) **(b)** $\kappa < 0$ (divergent wavefront)

Figure 21: Sign convention of the curvature of a normal section.

In application to EM diffraction problems, two types of surfaces are frequently encountered: a wavefront and a reflecting surface of a scatterer. In either case, we choose $\hat{N}$ pointing toward the electromagnetic source. Because of this choice, the sign of κ has the following meaning:

(i) Let the surface be a wavefront. Since $\hat{N}$ points toward the source which excites the wave, $\hat{N}$ is in the opposite direction of the wave propagation. If $\kappa(\hat{t}) > 0$, the normal section C of the surface in the direction of $\hat{t}$ is divergent (Figure 21a), whereas if $\kappa(\hat{t}) < 0$, the normal section is convergent (Figure 21b).

(ii) Let the surface be a reflecting surface of a scatterer, and $\hat{N}$ point toward the source which illuminates the scatterer. If $\kappa(\hat{t}) > 0$, the normal section is concave (Figure 21a), whereas if $\kappa(\hat{t}) < 0$, the normal section is convex (Figure 21b).

From (8.4) and (8.5), it is clear that curvature $\kappa(\hat{t})$ has a dimension of (length)$^{-1}$. We define its reciprocal as the **radius of curvature** $R(\hat{t})$ of surface Σ at point P in the direction of $\hat{t}$:

$$R(\hat{t}) = \frac{1}{\kappa(\hat{t})}. \tag{8.6}$$

(In EM diffraction problems, we often use the radius of curvature for a surface, and curvature for a curve, in order to minimize confusion between κ and κ_c.) It is important to remember that at a given point P on the surface, $\kappa(\hat{t})$ and $R(\hat{t})$ are functions of tangent vector $\hat{t}$. Their variation with respect to $\hat{t}$ will be discussed in detail in the next section.

We will now develop a formula for computing $\kappa(\hat{t})$ for a given direction $\hat{t}$. Differentiation of $\hat{t}\cdot\hat{N}=0$ gives

$$\frac{d\hat{t}}{d\sigma}\cdot\hat{N}=-\hat{t}\cdot\frac{d\hat{N}}{d\sigma}=-\frac{d\vec{r}}{d\sigma}\cdot\frac{d\hat{N}}{d\sigma}=-\frac{d\vec{r}\cdot d\hat{N}}{d\vec{r}\cdot d\vec{r}}.$$

Making use of this relation and (8.4) in (8.5), we have

$$\kappa(\hat{t})=\frac{d^2\vec{r}}{d\sigma^2}\cdot\hat{N}=\frac{d\hat{t}}{d\sigma}\cdot\hat{N}=-\frac{d\vec{r}\cdot d\hat{N}}{d\vec{r}\cdot d\vec{r}}. \tag{8.7}$$

The numerator and denominator of the right-hand side of (8.7) represent two important quantities in the study of a surface. They are explained below:

The first fundamental form

$$I=d\vec{r}\cdot d\vec{r}=(d\sigma)^2=(dx)^2+(dy)^2+(dz)^2 \tag{8.8}$$

represents the square of the differential arc length along a curve on a surface. In terms of the parameters (u,v) we have

$$d\vec{r}=\vec{r}_u du+\vec{r}_v dv. \tag{8.9}$$

Then an alternative form for I is

$$I=d\vec{r}\cdot d\vec{r}=Edu^2+2Fdudv+Gdv^2 \tag{8.10}$$

where du^2 means $(du)^2$, and

$$E=\vec{r}_u\cdot\vec{r}_u,\quad F=\vec{r}_u\cdot\vec{r}_v,\quad G=\vec{r}_v\cdot\vec{r}_v. \tag{8.11}$$

Despite the fact that E, F, and G depend on (u,v), the first fundamental form is invariant with respect to change of parameters. This is obvious from the geometrical significance of arc length.

The second fundamental form

$$II=-d\vec{r}\cdot d\hat{N}, \tag{8.12}$$

which will be shown later, is twice the deviation of the surface at $(u+du,v+dv)$ from the tangent plane at (u,v) (Figure 22). For computational purposes we note

$$d\hat{N}=\hat{N}_u du+\hat{N}_v dv. \tag{8.13}$$

(Note that $\hat{N}$ is a unit vector but $\hat{N}_u$ is not). Then an alternative form of II is

$$II=-d\vec{r}\cdot d\hat{N}=edu^2+2fdudv+gdv^2 \tag{8.14}$$

where

$$e=-\vec{r}_u\cdot\hat{N}_u,\quad 2f=-(\vec{r}_u\cdot\hat{N}_v+\vec{r}_v\cdot\hat{N}_u),\quad g=-\vec{r}_v\cdot\hat{N}_v. \tag{8.15}$$

Since $\vec{r}_u\cdot\hat{N}=0$ and $\vec{r}_v\cdot\hat{N}=0$, more convenient formulas for e, f, and g are

$$e=\vec{r}_{uu}\cdot\hat{N}=\mu\frac{\vec{r}_{uu}\cdot(\vec{r}_u\times\vec{r}_v)}{\sqrt{EG-F^2}} \tag{8.16a}$$

$$f=\vec{r}_{uv}\cdot\hat{N}=\mu\frac{\vec{r}_{uv}\cdot(\vec{r}_u\times\vec{r}_v)}{\sqrt{EG-F^2}} \tag{8.16b}$$

$$g=\vec{r}_{vv}\cdot\hat{N}=\mu\frac{\vec{r}_{vv}\cdot(\vec{r}_u\times\vec{r}_v)}{\sqrt{EG-F^2}} \tag{8.16c}$$

where we used the definition of $\hat{N}$ in (8.3). The choice of μ ($\mu = +1$ or -1) should agree with that utilized in defining $\hat{N}$. Unlike the first fundamental form I, which is always positive, the second fundamental form $I\!I$ may be either positive or negative. It may be shown that the absolute value of $I\!I$ is also invariant with change of parameters (u, v). Its sign is preserved if the parameter transformation has a positive Jacobian, otherwise, its sign is reversed.

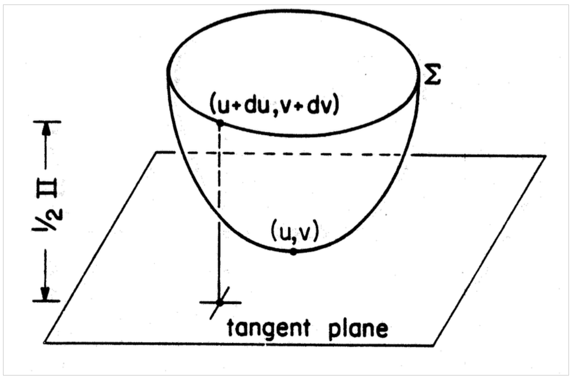


Figure 22: Second fundamental form $I\!I$ is twice the deviation of the surface from its tangent plane.

Return to the curvature in the direction $\hat{t}$ of a surface given in (8.7). Now, it can be written as

$$\kappa(\hat{t}) = \frac{I\!I}{I} = \frac{e du^2 + 2f du dv + g dv^2}{E du^2 + 2F du dv + G dv^2} \tag{8.17}$$

where (E, F, G) are defined in (8.11) and (e, f, g) in (8.16). The ratio dv/du determines the direction of $\hat{t}$. Alternatively we may write the tangent vector $\hat{t}$ as

$$\hat{t} = t_1 \vec{r}_u + t_2 \vec{r}_v. \tag{8.18}$$

Then the formula in (8.17) for the curvature becomes

$$\kappa(\hat{t}) = \frac{e t_1^2 + 2f t_1 t_2 + g t_2^2}{E t_1^2 + 2F t_1 t_2 + G t_2^2} \tag{8.19}$$

To illustrate the computation procedure of curvature, let us consider a cylinder with radius a as an example (Figure 23)

$$\vec{r}(u, v) = (a \cos u, a \sin u, v).$$

In terms of the familiar cylindrical coordinates, it is obvious that $u = \phi$ and $v = z$. Straightforward differentiation gives

$$\vec{r}_u = (-a \sin u, a \cos u, 0)$$
$$\vec{r}_v = (0, 0, 1)$$
$$\vec{r}_{uu} = (-a \cos u, -a \sin u, 0)$$
$$\vec{r}_{uv} = \vec{r}_{vv} = 0.$$

Substitution of the above results in (8.11) and (8.16) gives

$$E = a^2, \quad F = 0, \quad G = 1$$

$$e = -a, \quad f = 0, \quad g = 0.$$

Then the normal is found from (8.3), taking $\mu = +1$,

$$\hat{N} = (\cos u, \sin u, 0)$$

which points away from the axis of the cylinder (a convex cylinder). The two fundamental forms are

$$I = a^2 du^2 + dv^2$$

$$I\!I = -a du^2$$

and the curvature is given by (8.17), or

$$\kappa(\hat{t}) = -\frac{1}{a}\frac{1}{1+\frac{1}{a^2}\left(\frac{dv}{du}\right)^2}.$$

Introducing the angle α such that (Figure 23)

$$\tan\alpha = \frac{1}{a}\frac{dv}{du},$$

we have

$$\kappa(\hat{t}) = -\frac{1}{a}\cos^2\alpha.$$

Along the direction $\alpha = 0$, i.e., the direction of $\hat{t}$, $\kappa = -a^{-1}$, where the minus sign signifies the fact that the surface bends away from the outward normal $\hat{N}$. The magnitude of the curvature decreases continuously as α increases. At $\alpha = \pi/2$ (z direction), $\kappa = 0$ which is the direction of minimum bending of the surface.

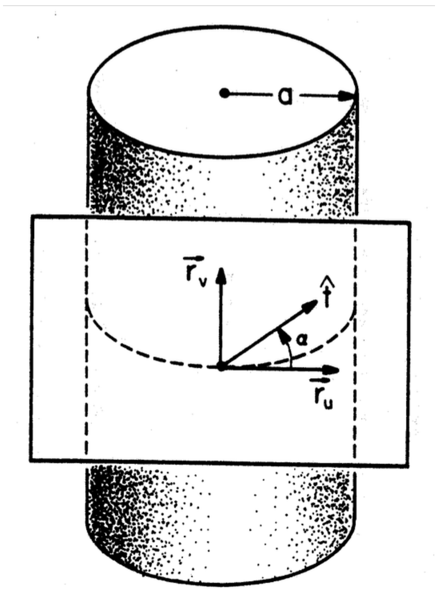


Figure 23: Curvature of a cylinder.

9 Principal Curvature

At a point P on a surface, curvature $\kappa(\hat{t})$ of the surface as given in (8.17) is a function of direction (or dv/du). A knowledge of $\kappa(\hat{t})$ for all $\hat{t}$ in the tangent plane determines the bending of the surface in the neighborhood of P. In this section, we will study the following important fact about $\kappa(\hat{t})$: As $\hat{t}$ revolves in the tangent plane, a pair of orthogonal directions exists for which $\kappa(\hat{t})$ assumes maximum and minimum values. These two directions are called **principal directions**, represented by two unit vectors $\hat{e}_1$ and $\hat{e}_2$; the two extreme values of κ are called **principal curvatures** denoted by κ_1 and κ_2. The curvature $\kappa(\hat{t})$ along an arbitrary direction can be actually expressed in terms of κ_1 and κ_2. Thus, there are only two degrees of freedom in $\kappa(\hat{t})$. From the simple example of the cylinder discussed at the end of the last section (Figure 23), it is obvious that $(\hat{e}_1, \hat{e}_2)$ are in the directions of $(\hat{\phi}, \hat{z})$, and $\kappa_1 = -a^{-1}, \kappa_2 = 0$.

Let us concentrate on the expression of $\kappa(\hat{t})$ given in (8.19). To determine the extreme values of $\kappa(\hat{t})$ when t_1 and t_2 vary, we require

$$\frac{\partial \kappa}{\partial t_1} = 0, \quad \frac{\partial \kappa}{\partial t_2} = 0 \tag{9.1}$$

which gives two homogeneous equations[¶]

$$(e - \kappa E)t_1 + (f - \kappa F)t_2 = 0 \tag{9.2}$$

$$(f - \kappa F)t_1 + (g - \kappa G)t_2 = 0 \tag{9.3}$$

For nontrivial solutions, the determinant of the coefficient matrix must be zero:

$$\det \begin{vmatrix} e - \kappa E & f - \kappa F \\ f - \kappa F & g - \kappa G \end{vmatrix} = 0 \tag{9.4}$$

Expanding the determinant, we have a quadratic equation for the extreme values of κ:

$$\kappa^2 - 2\kappa_M \kappa + \kappa_G = 0 \tag{9.5}$$

where the coefficients are:

$$\textbf{Mean curvature: } \kappa_M = \frac{\kappa_1 + \kappa_2}{2} = \frac{Eg - 2fF + eG}{2(EG - F^2)} \tag{9.6a}$$

$$\textbf{Gaussian curvature: } \kappa_G = \kappa_1 \kappa_2 = \frac{eg - f^2}{EG - F^2} \tag{9.6b}$$

It can be shown that the quadratic equation in (9.5) has two real roots, which are the principal curvatures κ_1 and κ_2:

$$\kappa_{1,2} = \kappa_M \pm \sqrt{\kappa_M^2 - \kappa_G} \tag{9.6c}$$

Substituting κ_1 and κ_2 into (9.2) or (9.3), we obtain solutions for (t_1, t_2), which determine the principal directions according to (8.18). More explicitly, the two (unit) principal directions

[¶]The formula of Rodrigues for lines of curvature reads

$$\kappa(\vec{r}_u du + \vec{r}_v dv) + (\hat{N}_u du + \hat{N}_v dv) = 0$$

which is the vector version of (9.2) and (9.3). It is a necessary and sufficient condition for dv/du to be the principal direction.

$(\hat{e}_1, \hat{e}_2)$ are given by

$$\hat{e}_1 = \frac{1}{\gamma_1}[1\vec{r}_u + \alpha\vec{r}_v] \tag{9.7a}$$

$$\hat{e}_2 = \frac{1}{\gamma_2}[\beta\vec{r}_u + 1\vec{r}_v] \tag{9.7b}$$

where

$$\alpha = \frac{e-\kappa_1 E}{\kappa_1 F - f} = \frac{f-\kappa_1 F}{\kappa_1 G - g} \tag{9.7c}$$

$$\beta = \frac{f-\kappa_2 F}{\kappa_2 E - e} = \frac{g-\kappa_2 G}{\kappa_2 F - f} \tag{9.7d}$$

$$\gamma_1 = (E + 2\alpha F + \alpha^2 G)^{1/2} \tag{9.7e}$$

$$\gamma_2 = (\beta^2 E + 2\beta F + G)^{1/2} \tag{9.7f}$$

If $\kappa_1 \neq \kappa_2$, $\hat{e}_1$ and $\hat{e}_2$ are orthogonal, which follows from the fact that

$$\hat{e}_1 \cdot \hat{e}_2 = \frac{1}{\gamma_1\gamma_2}[\beta E + (\alpha+\beta)F + \alpha G] = 0$$

If $\kappa_1 = \kappa_2$, the curvature at P is constant in all directions, and P is called **umbilic**.

Except for umbilics, at every point on a surface there are two mutually orthogonal principal directions $\hat{e}_1$ and $\hat{e}_2$. Curves on the surface that at all points are tangent to a principal direction are called **lines of curvature**. Many results can be greatly simplified if the lines of curvature are used for the u- and v-parameter curves.

Referring to Figure 24, when u- and v-parameter curves are themselves lines of curvature, we have

$$\frac{\vec{r}_u}{|\vec{r}_u|} = \hat{e}_1, \quad \frac{\vec{r}_v}{|\vec{r}_v|} = \hat{e}_2 \tag{9.8}$$

It can be shown from (9.7) that a necessary and sufficient condition for (9.8) is

$$F = 0, \quad f = 0 \tag{9.9}$$

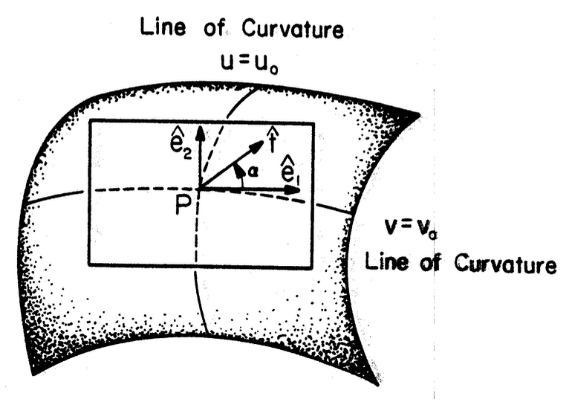


Figure 24: Lines of curvature are used as u- and v-parameter curves.

For an arbitrary tangent vector $\hat{t}$ at P, the representation in (8.18) now becomes

$$\hat{t} = \cos\alpha\,\hat{e}_1 + \sin\alpha\,\hat{e}_2 \tag{9.10}$$

where α is the angle of $\hat{t}$ measured from $\hat{e}_1$ in the tangent plane (Figure 24). Using (9.9) and (9.10) in (8.19), we have

$$\kappa(\hat{t}) = \frac{e}{E}\cos^2\alpha + \frac{g}{G}\sin^2\alpha \tag{9.11}$$

By simple differentiation of (9.11) with respect to α, we determine the two extreme values of κ:

$$\kappa_1 = \frac{e}{E}, \quad \text{if } \alpha = 0 \tag{9.12a}$$

$$\kappa_2 = \frac{g}{G}, \quad \text{if } \alpha = \pi/2 \tag{9.12b}$$

which by definition are the principal curvatures. Substituting (9.12) into (9.11) gives

$$\kappa(\hat{t}) = \kappa_1\cos^2\alpha + \kappa_2\sin^2\alpha \tag{9.13}$$

Thus, the curvature along an arbitrary direction defined in (9.10) is simply related to the two principal curvatures.

Ellipsoid. We will give an example for the calculation of curvatures. Consider an ellipsoid (Figure 15)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \tag{9.14a}$$

which may be represented by the following parametric equation:

$$\vec{r}(u,v) = (a\sin u\cos v, b\sin u\sin v, c\cos u) \tag{9.14b}$$

Straightforward differentiation gives

$$\vec{r}_u = (a\cos u\cos v, b\cos u\sin v, -c\sin u) \tag{9.15a}$$

$$\vec{r}_v = (-a\sin u\sin v, b\sin u\cos v, 0) \tag{9.15b}$$

$$\vec{r}_{uu} = (-a\sin u\cos v, -b\sin u\sin v, -c\cos u) \tag{9.15c}$$

$$\vec{r}_{uv} = (-a\cos u\sin v, b\cos u\cos v, 0) \tag{9.15d}$$

$$\vec{r}_{vv} = (-a\sin u\cos v, -b\sin u\sin v, 0) \tag{9.15e}$$

Substituting the above into (8.11) and (8.16) leads to

$$E = \cos^2 u(a^2\cos^2 v + b^2\sin^2 v) + c^2\sin^2 u \tag{9.16a}$$

$$F = (b^2 - a^2)\sin u\cos u\sin v\cos v \tag{9.16b}$$

$$G = \sin^2 u(a^2\sin^2 v + b^2\cos^2 v) \tag{9.16c}$$

$$e = \frac{-abc}{[c^2\sin^2 u(a^2\sin^2 v + b^2\cos^2 v) + a^2b^2\cos^2 u]^{1/2}} \tag{9.16d}$$

$$f = 0 \tag{9.16e}$$

$$g = e\sin^2 u \tag{9.16f}$$

The surface normal is found by (8.3), choosing $\mu = +1$,

$$\hat{N} = \frac{(\frac{\sin u\cos v}{a}, \frac{\sin u\sin v}{b}, \frac{\cos u}{c})}{[(\frac{\sin u\cos v}{a})^2 + (\frac{\sin u\sin v}{b})^2 + (\frac{\cos u}{c})^2]^{1/2}}$$

Alternatively it may be written in terms of (x, y, z)

$$\hat{N} = \frac{(\frac{x}{a^2}, \frac{y}{b^2}, \frac{z}{c^2})}{\sqrt{(\frac{x}{a^2})^2 + (\frac{y}{b^2})^2 + (\frac{z}{c^2})^2}}$$

From (9.6a) the mean curvature is found to be

$$\kappa_M = \frac{abc[a^2 \sin^2 u \cos^2 v + b^2 \sin^2 u \sin^2 v + c^2 \cos^2 u - (a^2 + b^2 + c^2)]}{2[c^2 \sin^2 u(a^2 \sin^2 v + b^2 \cos^2 v) + a^2 b^2 \cos^2 u]^{3/2}} \tag{9.17}$$

which becomes, in terms of (x, y, z),

$$\kappa_M = \frac{(x^2 + y^2 + z^2) - (a^2 + b^2 + c^2)}{2(abc)^2(\frac{x^2}{a^4} + \frac{y^2}{b^4} + \frac{z^2}{c^4})^{3/2}}$$

The Gaussian curvature is determined from (9.6b) with the result

$$\kappa_G = \frac{(abc)^2}{[c^2 \sin^2 u(a^2 \sin^2 v + b^2 \cos^2 v) + a^2 b^2 \cos^2 u]^2} = \frac{1}{(abc)^2(\frac{x^2}{a^4} + \frac{y^2}{b^4} + \frac{z^2}{c^4})^2} \tag{9.18}$$

The principal curvatures can be found from the quadratic equation in (9.5), or

$$\kappa_1, \kappa_2 = \kappa_M \pm \sqrt{\kappa_M^2 - \kappa_G} \tag{9.19}$$

At points where $\kappa_M^2 = \kappa_G$, we have identical curvature in all directions on the tangent plane. They are umbilics. For an ellipsoid with unequal axes, there are four umbilics. Assuming $a > b > c$, the coordinates of the umbilics are

$$x = \pm a\sqrt{\frac{a^2 - b^2}{a^2 - c^2}}, \quad y = 0, \quad z = \pm c\sqrt{\frac{b^2 - c^2}{a^2 - c^2}} \tag{9.20}$$

For the special case $a = b = c$, the ellipsoid reduces to a sphere of radius a. Then $\kappa_1 = \kappa_2 = (-1/a)$ and all points on a sphere are umbilics.

As a numerical example, let us set

$$a = 2, \quad b = \sqrt{2}, \quad c = 1 \tag{9.21}$$

The four umbilics are located at

$$x = \pm 2\sqrt{\frac{2}{3}}, \quad y = 0, \quad z = \pm\sqrt{\frac{2}{3}}$$

one of which is approximately indicated in Figure 25. At a point P where $u = v = \pi/4$ and has coordinates

$$x = 1, \quad y = \frac{1}{\sqrt{2}}, \quad z = \frac{1}{\sqrt{2}}$$

we have the following numerical values

$$\vec{r}_u = (1, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) \tag{9.22}$$

$$\vec{r}_v = (-1, \frac{1}{\sqrt{2}}, 0) \tag{9.23a}$$

$$E = 2, \quad F = -\frac{1}{2}, \quad G = \frac{3}{2} \tag{9.23b}$$

$$e = \frac{-4}{\sqrt{11}}, \quad f = 0, \quad g = -\frac{2}{\sqrt{11}} \tag{9.24a}$$

$$\kappa_M = \frac{-20}{11\sqrt{11}}, \quad \kappa_G = \frac{32}{121} \tag{9.25a}$$

$$\kappa_{1,2} = \frac{-4}{11\sqrt{11}}(5 \mp \sqrt{3}) \tag{9.25b}$$

The two principal directions calculated from (9.7) are

$$\hat{e}_1 = \frac{1}{\sqrt{9+4\sqrt{3}}}(2+\sqrt{3}, -\sqrt{\frac{3}{2}}, -\frac{1}{\sqrt{2}}) \tag{9.26a}$$

$$\hat{e}_2 = \frac{1}{\sqrt{16+6\sqrt{3}}}(-3-\sqrt{3}, \frac{1-\sqrt{3}}{\sqrt{2}}, \frac{1+\sqrt{3}}{\sqrt{2}}) \tag{9.26b}$$

which are obtained with the help of intermediate values

$$\alpha = -(1+\sqrt{3}), \quad \beta = \frac{1}{2}(1+\sqrt{3}) \tag{9.27a}$$

$$\gamma_1 = \sqrt{9+4\sqrt{3}}, \quad \gamma_2 = \frac{1}{2}\sqrt{16+6\sqrt{3}} \tag{9.27b}$$

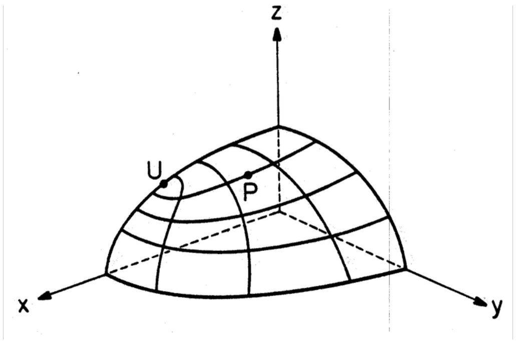


Figure 25: A sketch of the ellipsoid with $a = 2, b = \sqrt{2}, c = 1$. U is an umbilic. The lines are lines of curvature.

Caustic. As a last subject on principal curvatures, let us establish a fact that finds frequent use in EM diffraction problems. Consider a point P specified by parameters (u, v) on a surface W (in diffraction problems, W is recognized as a wavefront). Referring to an origin O, we may also specify P by a position vector $\vec{r}(u, v)$ (Figure 26). The surface normal at P is given by the equation

$$\vec{R}(u, v, \sigma) = \vec{r}(u, v) - \sigma\hat{N}(u, v) \tag{9.28}$$

Here $\vec{R}$ is the position vector of a typical point on the surface normal. $\hat{N}$ is the unit surface normal vector (pointing toward the source of the wavefront). σ is the arc length measured positively in the direction of $(-\hat{N})$, and $\sigma = 0$ at P. A question of interest is, under what condition does the normal at an adjacent point P' specified by parameters $(u + du, v + dv)$ on the surface intersect with that at P? The answer is that, **to the first-order approximation, normals drawn from adjacent points on a line of curvature intersect**. This will be shown below. The distance between the normal at P and that at P' is the magnitude of the vector

$$\begin{aligned}&\vec{R}(u + du, v + dv, \sigma + d\sigma) - \vec{R}(u, v, \sigma)\\&\quad = (\vec{R}_u du + \vec{R}_v dv + \vec{R}_\sigma d\sigma) + \text{higher-order terms}\end{aligned}$$

If the two normals intersect to the first order, we require

$$\vec{D} = \vec{R}_u du + \vec{R}_v dv + \vec{R}_\sigma d\sigma \tag{9.29}$$

to be zero. The substitution of (9.28) into (9.29) leads to

$$\vec{D} = (\vec{r}_u du + \vec{r}_v dv) - \hat{N} d\sigma - \sigma(\hat{N}_u du + \hat{N}_v dv) \tag{9.30}$$

If P and P' are both located on the same line of curvature, we may use the formula of Rodrigues (see footnotes in association with (9.2) and (9.3)):

$$\kappa(\vec{r}_u du + \vec{r}_v dv) + (\hat{N}_u du + \hat{N}_v dv) = 0 \tag{9.31}$$

Then (9.30) becomes

$$\vec{D} = -(1 + \sigma\kappa)(\vec{r}_u du + \vec{r}_v dv) - \hat{N} d\sigma$$

Clearly $\vec{D}$ is zero if

$$\sigma = -\frac{1}{\kappa} \quad \text{(a constant)}$$

and hence $d\sigma = 0$. In summary, to the first order, the normal at P and the normal at an adjacent point P' on the same line of curvature intersect at

$$\vec{R} = \vec{r}(u, v) + \frac{1}{\kappa}\hat{N} \tag{9.32}$$

Recall the sign convention of κ discussed in Section 8. If the normal section of W along PP' bends toward (away from) $\hat{N}$, κ is positive (negative). Therefore, regardless of the choice of $\hat{N}$, the intersection point $\vec{R}$ in (9.32) always lies on the concave side of the normal section of W.

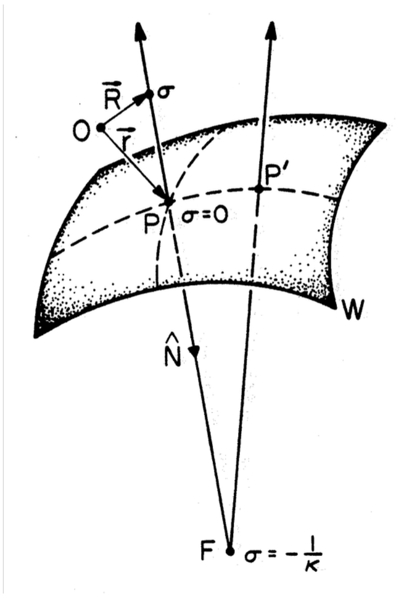


Figure 26: Intersection of normals drawn from adjacent points on a line of curvature.

In optics, point $\vec{R}$ in (9.32) is called a **focus**. Since, in general, there are two distinct principal curvatures $\kappa = \kappa_1$ and $\kappa = \kappa_2$, there are two foci on each normal (ray). As (u, v) varies on the given surface (wavefront), (9.32) describes two surfaces (one for $\kappa = \kappa_1$ and one for $\kappa = \kappa_2$), which are known as **caustics** (or **focal surfaces**). The condition that $\vec{D}$ defined in (9.29) is zero implies that $\vec{R}_u, \vec{R}_v, \vec{R}_\sigma$ are coplanar. Thus the normal $\hat{N}(u, v)$ is tangent to caustics at their foci. Two caustics F_1 and F_2 of a surface W are sketched in Figure 27. $(\hat{e}_1, \hat{e}_2)$ are principal directions, and $\hat{N}$, the normal. In this sketch, κ_1 is negative (the normal section bends away from $\hat{N}$), whereas κ_2 is positive.

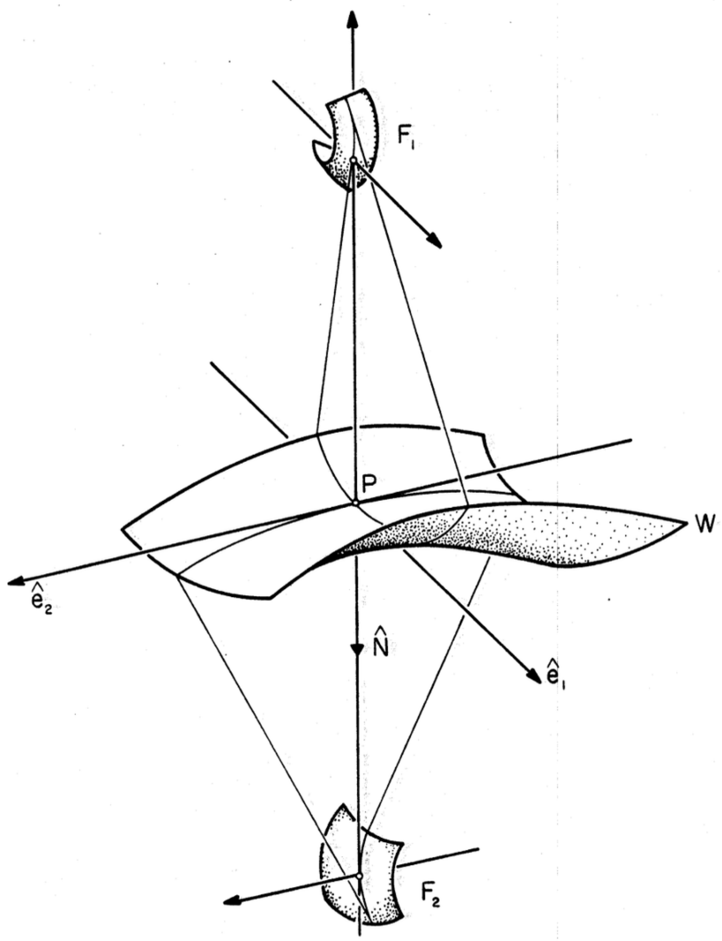


Figure 27: Two caustics F_1 and F_2 of a surface W. $(\hat{e}_1, \hat{e}_2)$ are principal directions. In this sketch, κ_1 is negative and κ_2 is positive.

10 Curvature Matrix

At a point P on a curve, the knowledge of curvature and torsion at P determines the local properties to the second degree of the curve in the neighborhood of P, as described in (6.6). At a point P on a surface, a similar role is played by a linear operator S, called **shape operator**, defined on the tangent plane at P. In our application to EM diffraction problems, it is convenient to represent S by a 2×2 matrix $\bar{\bar{Q}}$. We call $\bar{\bar{Q}}$ the **curvature matrix**. The explicit form of $\bar{\bar{Q}}$, of course, depends on the base vectors on the tangent plane. If the base vectors coincide with the principal directions of the surface, $\bar{\bar{Q}}$ is a diagonal matrix. Otherwise, $\bar{\bar{Q}}$ is not a diagonal matrix but can be related to the diagonalized one by a standard theory of linear transformation.

Instead of starting with the shape operator S, we will introduce $\bar{\bar{Q}}$ directly as below. On a surface Σ, there is a normal $\hat{N}$ at each point. As $\hat{N}$ moves away from a point in an arbitrary direction, its variation follows the bending of Σ in that direction. Take the cylindrical surface in Figure 23 as an example, along the direction of $\vec{r}_v$ (z-direction), $\hat{N}$ is a constant, indicating the fact that Σ does not bend in this direction. Along the direction of $\vec{r}_u$ (ϕ-direction), the differential variation of $\hat{N}$ is also in the direction of $\vec{r}_u$ and has a constant magnitude. This reflects the roundness of Σ in the direction of $\vec{r}_u$. Following this idea, we may introduce a quantity measuring the variation of $\hat{N}$ at each point on Σ. Since $\hat{N} \cdot \hat{N} = 1$, differentiations with respect to u and v give

$$\hat{N}_u \cdot \hat{N} = 0, \quad \hat{N}_v \cdot \hat{N} = 0 \tag{10.1}$$

which means that $(\hat{N}_u, \hat{N}_v)$ at a point P lie in the tangent plane of Σ at P. Hence we may express $(\hat{N}_u, \hat{N}_v)$ in terms of tangent vectors $(\vec{r}_u, \vec{r}_v)$ as follows:

$$-\hat{N}_u = Q_{11}\vec{r}_u + Q_{12}\vec{r}_v \tag{10.2a}$$

$$-\hat{N}_v = Q_{21}\vec{r}_u + Q_{22}\vec{r}_v \tag{10.2b}$$

The four parameters in (10.2) form a **curvature matrix** $\bar{\bar{Q}}$

$$\bar{\bar{Q}} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \tag{10.3}$$

In matrix notations, (10.2) may be rewritten as

$$[-\hat{N}_u \quad -\hat{N}_v]^T = \bar{\bar{Q}}[\vec{r}_u \quad \vec{r}_v]^T \tag{10.4}$$

Here T is the transpose operator. $[\vec{r}_u \quad \vec{r}_v]$ is a 3×2 matrix and is explicitly given by

$$[\vec{r}_u \quad \vec{r}_v] = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{bmatrix} \tag{10.5}$$

Similarly, $[\hat{N}_u \quad \hat{N}_v]$ is given by

$$[\hat{N}_u \quad \hat{N}_v] = \begin{bmatrix} \frac{\partial \hat{N}_x}{\partial u} & \frac{\partial \hat{N}_x}{\partial v} \\ \frac{\partial \hat{N}_y}{\partial u} & \frac{\partial \hat{N}_y}{\partial v} \\ \frac{\partial \hat{N}_z}{\partial u} & \frac{\partial \hat{N}_z}{\partial v} \end{bmatrix} \tag{10.6}$$

We regard (10.4) as the definition of $\bar{\bar{Q}}$.

For a given surface, it is desirable to develop a formula for calculating $\bar{\bar{Q}}$ directly. To this end, let us introduce the **first fundamental matrix** $\bar{\bar{I}}$ such that

$$\bar{\bar{I}} = \begin{bmatrix} E & F \\ F & G \end{bmatrix} = \begin{bmatrix} \vec{r}_u \cdot \vec{r}_u & \vec{r}_u \cdot \vec{r}_v \\ \vec{r}_v \cdot \vec{r}_u & \vec{r}_v \cdot \vec{r}_v \end{bmatrix} = [\vec{r}_u \quad \vec{r}_v]^T [\vec{r}_u \quad \vec{r}_v] \tag{10.7}$$

and the **second fundamental matrix** $\bar{\bar{II}}$ such that

$$\bar{\bar{II}} = \begin{bmatrix} e & f \\ f & g \end{bmatrix} = -\begin{bmatrix} \hat{N}_u \cdot \vec{r}_u & \hat{N}_u \cdot \vec{r}_v \\ \hat{N}_v \cdot \vec{r}_u & \hat{N}_v \cdot \vec{r}_v \end{bmatrix} = [-\hat{N}_u \quad -\hat{N}_v]^T [\vec{r}_u \quad \vec{r}_v] \tag{10.8}$$

Now multiply both sides of (10.4) by $[\vec{r}_u \quad \vec{r}_v]^T$, which yields

$$\bar{\bar{II}} = \bar{\bar{Q}}\bar{\bar{I}} \tag{10.9}$$

Since (u,v) are independent parameters, $\bar{\bar{I}}$ in (10.7) is nonsingular. Inversion of $\bar{\bar{I}}$ in (10.9) gives

$$\bar{\bar{Q}} = \bar{\bar{II}}(\bar{\bar{I}})^{-1} \tag{10.10a}$$

or more explicitly

$$Q_{11} = \frac{eG - fF}{EG - F^2} \qquad Q_{12} = \frac{fE - eF}{EG - F^2} \tag{10.10b}$$

$$Q_{21} = \frac{fG - gF}{EG - F^2} \qquad Q_{22} = \frac{gE - fF}{EG - F^2} \tag{10.10c}$$

The results in (10.2) and (10.10) are known as **Weingarten equations**. They enable us to calculate $\bar{\bar{Q}}$ directly for a given surface.

Once $\bar{\bar{Q}}$ is found at a point P on a surface, we may calculate from it the principal curvatures and directions at P. This will be demonstrated next. From (10.10) and the definitions in (9.6), it is readily shown that

$$\frac{1}{2} \cdot \text{trace } \bar{\bar{Q}} = \frac{1}{2}(\kappa_1 + \kappa_2) = \textbf{mean curvature } \kappa_M \tag{10.11a}$$

$$\det \bar{\bar{Q}} = \kappa_1 \kappa_2 = \textbf{Gaussian curvature } \kappa_G \tag{10.11b}$$

Therefore, κ_1 and κ_2 are the two **eigenvalues** of $\bar{\bar{Q}}$. Following a standard procedure, let us diagonalize $\bar{\bar{Q}}$. The two eigenvectors of $\bar{\bar{Q}}$ are denoted by

$$\vec{d}_1 = \begin{bmatrix} d_{11} \\ d_{21} \end{bmatrix}, \quad \vec{d}_2 = \begin{bmatrix} d_{12} \\ d_{22} \end{bmatrix} \tag{10.12}$$

which satisfy the relations

$$\bar{\bar{Q}}\vec{d}_n = \kappa_n \vec{d}_n, \quad n = 1, 2 \tag{10.13}$$

Explicitly the solutions of (10.13) are given by

$$\frac{d_{21}}{d_{11}} = \frac{\kappa_1 - Q_{11}}{Q_{12}} = \frac{Q_{21}}{\kappa_1 - Q_{22}} \tag{10.14a}$$

$$\frac{d_{12}}{d_{22}} = \frac{\kappa_2 - Q_{22}}{Q_{21}} = \frac{Q_{12}}{\kappa_2 - Q_{11}} \tag{10.14b}$$

which determine $\vec{d}_n$ within a normalization constant. Let us form a 2×2 matrix $\bar{\bar{D}}$ such that

$$\bar{\bar{D}} = [\vec{d}_1 \quad \vec{d}_2] = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \tag{10.15}$$

Then the matrix

$$\bar{\bar{D}}^{-1}\bar{\bar{Q}}\bar{\bar{D}} = \begin{bmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{bmatrix} \tag{10.16}$$

is the diagonalized curvature matrix. The first and second rows of the 2×3 matrix

$$\bar{\bar{D}}^{-1}[\vec{r}_u \quad \vec{r}_v]^T \tag{10.17}$$

give the principal directions. After normalization, the unit principal directions in (10.17) are explicitly given by

$$\hat{e}_1 = \frac{1}{\gamma_1}(1\vec{r}_u + \alpha\vec{r}_v) \tag{10.18a}$$

$$\hat{e}_2 = \frac{1}{\gamma_2}(\beta\vec{r}_u + 1\vec{r}_v) \tag{10.18b}$$

where

$$\alpha = \frac{d_{12}}{d_{22}} = \frac{Q_{22}-\kappa_2}{Q_{21}} = \frac{Q_{12}}{Q_{11}-\kappa_2} \tag{10.18c}$$

$$\beta = \frac{d_{21}}{d_{11}} = \frac{Q_{11}-\kappa_1}{Q_{12}} = \frac{Q_{21}}{Q_{22}-\kappa_1} \tag{10.18d}$$

$$\gamma_1 = (E + 2\alpha F + \alpha^2 G)^{1/2} \tag{10.18e}$$

$$\gamma_2 = (\beta^2 E + 2\beta F + G)^{1/2} \tag{10.18f}$$

By straightforward manipulation, it can be shown that (10.18) is identical to (9.7). In summary, curvature matrix $\bar{\bar{Q}}$ is defined in (10.2) and (10.3). It can be calculated from (10.10). Once $\bar{\bar{Q}}$ is known, the principal curvatures and direction are determined by (10.11) and (10.18), respectively.

The four vectors $\vec{r}_u, \vec{r}_v, \hat{e}_1$ and $\hat{e}_2$ lie in the tangent plane at P of a surface (Figure 19). In general, $\vec{r}_u$ and $\vec{r}_v$ are not normalized, nor mutually orthogonal, and not in principal directions, as shown in Figure 28a.

Now, let us consider the special case in Figure 28b, where

(i) $|\vec{r}_u| = |\vec{r}_v| = 1$ (10.19a)

(ii) $\vec{r}_u \cdot \vec{r}_v = 0$ (10.19b)

(iii) the angle measured counterclockwise from $\vec{r}_u$ to $\hat{e}_1$ is ψ (10.19c)

Then from the relations

$$E = 1, \quad F = 0, \quad G = 1, \tag{10.20}$$

the following facts may be established:

(a) Apart from a constant, $\bar{\bar{D}}$ in (10.15) is a unitary matrix given by

$$\bar{\bar{D}} = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} = (\bar{\bar{D}}^{-1})^T \tag{10.21}$$

(b) The curvature matrix $\bar{\bar{Q}}$ is symmetrical and is given by

$$\begin{aligned}\bar{\bar{Q}} = \begin{bmatrix} e & f \\ f & g \end{bmatrix} &= \bar{\bar{D}} \begin{bmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{bmatrix} \bar{\bar{D}}^T \\ &= \begin{bmatrix} \kappa_1\cos^2\psi + \kappa_2\sin^2\psi & (\kappa_1-\kappa_2)\sin\psi\cos\psi \\ (\kappa_1-\kappa_2)\sin\psi\cos\psi & \kappa_1\sin^2\psi + \kappa_2\cos^2\psi \end{bmatrix}\end{aligned} \tag{10.22}$$

If $\psi = 0$, $\bar{\bar{Q}}$ is further simplified to become a diagonal matrix

$$\bar{\bar{Q}} = \begin{bmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{bmatrix}, \quad \text{if } \psi = 0 \tag{10.23}$$

As discussed in (9.9), a necessary and sufficient condition for $\psi = 0$ is $F = 0$ and $f = 0$.

We conclude this section with an example. Let us calculate the curvature matrix at a point P with coordinates

$$x = 1, \quad y = \frac{1}{\sqrt{2}}, \quad z = \frac{1}{\sqrt{2}} \tag{10.24}$$

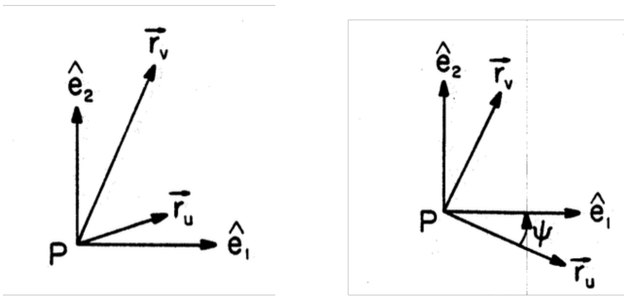
located on an ellipsoid (Figure 25)

$$\left(\frac{x}{2}\right)^2 + \left(\frac{y}{\sqrt{2}}\right)^2 + z^2 = 1 \tag{10.25}$$

This example was studied in Section 9. Substituting (9.24) into (10.10) gives immediately

$$\bar{\bar{Q}} = \frac{-4}{11\sqrt{11}} \begin{bmatrix} 6 & 2 \\ 1 & 4 \end{bmatrix} \tag{10.26}$$

We note that $\bar{\bar{Q}}$ is not diagonal because $(\vec{r}_u, \vec{r}_v)$ in (9.23) are not in the same directions as $(\hat{e}_1, \hat{e}_2)$ in (9.26); nor is it symmetrical because $\vec{r}_u$ and $\vec{r}_v$ are not orthogonal. From the given $\bar{\bar{Q}}$ in (10.26), we can calculate principal curvatures from (10.11) and principal directions from (10.18). These results are of course identical to those given in (9.25) and (9.26).



(a) **(b)**

Figure 28: Four vectors in the tangent plane at P of a surface. $(\hat{e}_1, \hat{e}_2)$ are principal directions.

11 Approximation of a Surface

In EM diffraction problems, there are typically two types of surface involved: (i) a perfectly conducting surface Σ where an electromagnetic field is reflected or diffracted, and (ii) a wavefront W of a ray pencil. We are often interested in the **local geometrical properties** of Σ or W at a point O. For this purpose instead of using the exact representation of the surface, a quadratic approximation based on the Taylor expansion is sufficient. This approximation is discussed in the present section.

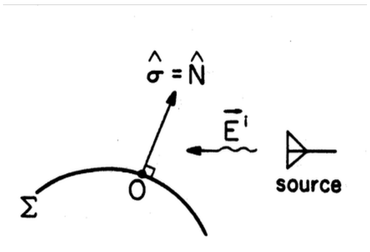
At point O, let us introduce a set of right-handed orthonormal base vectors $(\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma} = \hat{\sigma}_3)$. In the case of a conducting surface Σ, we choose

$$\hat{\sigma} = +\hat{N}, \quad \text{for } \Sigma \tag{11.1a}$$

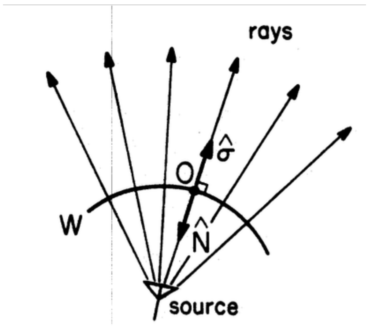
In the case of a wavefront W, we choose (Figure 29b)

$$\hat{\sigma} = -\hat{N}, \quad \text{for } W \tag{11.1b}$$

Thus, the normal $\hat{N}$ always points toward the source (of the incident field in the case of Σ, or of the ray pencil in the case of W), whereas $\hat{\sigma}$ points toward the incident field, or the direction of propagation of the ray pencil. The two orthogonal directions $(\hat{\sigma}_1, \hat{\sigma}_2)$ lie in the tangent plane of the surface. They may or may not coincide with the principal directions. In the remainder of this section, we will concentrate on the case in (11.1a). Results so obtained apply also to the case in (11.1b) after obvious modifications (See (13) in Section 13).



(a)



(b)

Figure 29: Choice of normal $\hat{N}$ in an electromagnetic diffraction problem. Note that $\hat{N}$ always points toward the electromagnetic source.

Consider a point P on Σ in the neighborhood of O (Figure 30). The position vector $\vec{r}$ of P in reference to O is

$$\vec{r} = \hat{\sigma}_1\sigma_1 + \hat{\sigma}_2\sigma_2 + \hat{\sigma}\sigma \tag{11.2}$$

where $(\sigma_1, \sigma_2, \sigma)$ are the rectangular coordinates of P. Since P is on Σ, there are only two degrees of freedom in $(\sigma_1, \sigma_2, \sigma)$. Let (σ_1, σ_2) be the independent parameters, and play the roles of (u, v). Then the relation

$$\sigma = \sigma(\sigma_1, \sigma_2) \tag{11.3}$$

describes Σ. To obtain an approximate version of (11.3) valid for small $|\sigma_1|$ and $|\sigma_2|$, let us replace (11.2) by its Taylor expansion around O, namely

$$\vec{r}(\sigma_1, \sigma_2) = \vec{r}_{\sigma_1}\sigma_1 + \vec{r}_{\sigma_2}\sigma_2 + \frac{1}{2}(\vec{r}_{\sigma_1\sigma_1}\sigma_1^2 + 2\vec{r}_{\sigma_1\sigma_2}\sigma_1\sigma_2 + \vec{r}_{\sigma_2\sigma_2}\sigma_2^2) + O(\sigma_{1,2}^3) \tag{11.4}$$

where $\vec{r}_{\sigma_1}$, for example, is the partial derivative of $\vec{r}$ with respect to σ_1 evaluated at $(\sigma_1 = 0, \sigma_2 = 0)$. In (11.4), $O(\sigma_1, \sigma_2^3)$ means terms of order $\sigma_1^n \sigma_2^m$ with $m+n=3$, which indicates that terms higher than quadratics of σ_1 and σ_2 have been neglected. From (11.4), σ is found to be

$$\sigma = \vec{r} \cdot \hat{\sigma} = \frac{1}{2}(e\sigma_1^2 + 2f\sigma_1\sigma_2 + g\sigma_2^2) + O(\sigma_{1,2}^3) \tag{11.5}$$

where e, f, and g were defined in (8.16) (recalling $u \to \sigma_1, v \to \sigma_2$ and $\hat{N} = \hat{\sigma}$). Note that at O we have

$$\vec{r}_{\sigma_1} = \frac{\partial}{\partial \sigma_1}(\hat{\sigma}_1\sigma_1 + \hat{\sigma}_2\sigma_2 + \hat{\sigma}\sigma)\Big|_{\sigma_1=0,\sigma_2=0} = \hat{\sigma}_1 + \hat{\sigma}(e\sigma_1 + f\sigma_2)\Big|_{\sigma_1=0,\sigma_2=0} = \hat{\sigma}_1$$

and similarly $\vec{r}_{\sigma_2} = \hat{\sigma}_2$. When these results are used in (8.10), we have

$$E = 1, \quad F = 0, \quad G = 1. \tag{11.6}$$

With the help of (10.10) and (11.6), we may rewrite (11.5) in terms of the curvature matrix $\bar{\bar{Q}}$, namely

$$\sigma = \frac{1}{2}\left[Q_{11}\sigma_1^2 + (Q_{12} + Q_{21})\sigma_1\sigma_2 + Q_{22}\sigma_2^2\right] + O(\sigma_{1,2}^3) \tag{11.7}$$

or, in matrix notation

$$\sigma = \frac{1}{2}\begin{bmatrix}\sigma_1 \\ \sigma_2\end{bmatrix} \cdot \bar{\bar{Q}} \begin{bmatrix}\sigma_1 \\ \sigma_2\end{bmatrix} + O(\sigma_{1,2}^3) \tag{11.8}$$

This is the desired quadratic approximation of a surface valid for small $|\sigma_1|$ and $|\sigma_2|$.

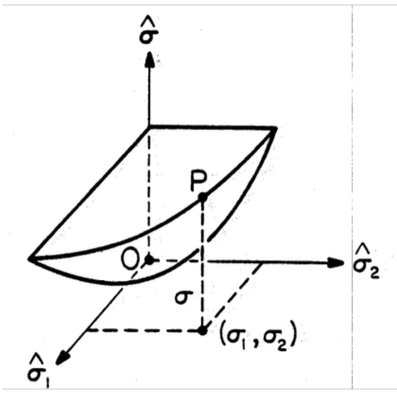


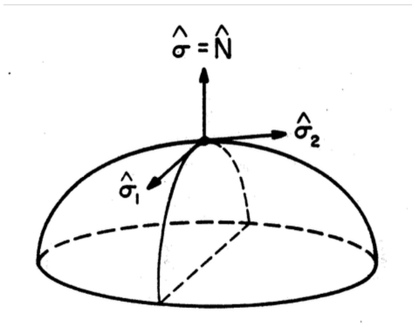
Figure 30: Quadratic approximation of a surface.

When $(\hat{\sigma}_1, \hat{\sigma}_2)$ coincide with the principal directions $(\hat{e}_1, \hat{e}_2)$, $\bar{\bar{Q}}$ is given in (10.23), and (11.8) becomes

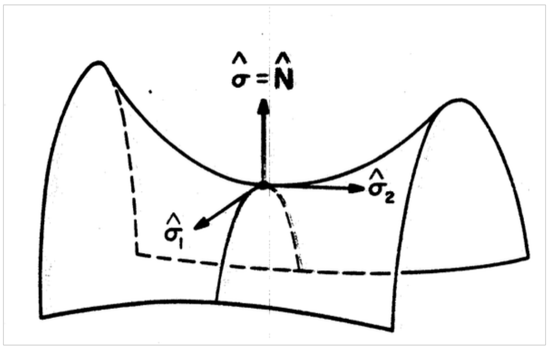
$$\sigma = \frac{1}{2}(\kappa_1\sigma_1^2 + \kappa_2\sigma_2^2) + O(\sigma_{1,2}^3) \tag{11.9}$$

This simple representation in (11.9) brings out clearly the geometrical significance of the sign of Gaussian curvature $\kappa_G = \kappa_1\kappa_2$, as discussed below : (i) If $\kappa_G > 0$, κ_1 and κ_2 have the same sign. The quadratic approximation in (11.9) describes a paraboloid (Figure 31a). (ii) If $\kappa_G < 0$, κ_1 and κ_2 have different signs. The quadratic approximation describes a hyperboloid (Figure 31b). (iii) If $\kappa_G = 0$ with $\kappa_1 \neq 0$ and $\kappa_2 = 0$, the quadratic approximation describes a cylinder (Figure 31c). (iv) If $\kappa_G = 0$ with $\kappa_1 = \kappa_2 = 0$, the quadratic approximation reduces to a plane. At most, the original surface can have small bending at $(\sigma_1 = 0, \sigma_2 = 0)$. To study this small bending, we have to examine the higher-order terms in the Taylor expansion. We emphasize that the Gaussian curvature, which equals $\det \bar{\bar{Q}}$, is an invariant geometrical quantity of a surface. Its value (and its sign) is independent of the choice of a particular coordinate system which describes the surface.

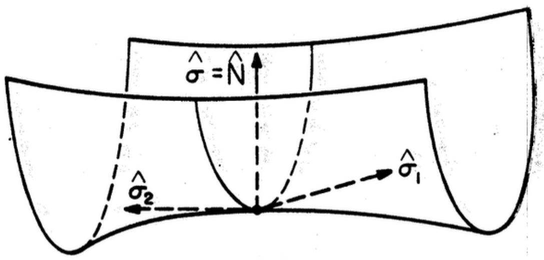
When $(\hat{\sigma}_1, \hat{\sigma}_2)$ make an angle ψ with the principal directions $(\hat{e}_1, \hat{e}_2)$ (cf. Figure 28b), $\bar{\bar{Q}}$ is given in (10.22). It should be added that although (11.8) was derived based on the assumption of orthonormal $(\hat{\sigma}_1, \hat{\sigma}_2)$, it is valid for any two independent vectors $(\hat{\sigma}_1, \hat{\sigma}_2)$ in the tangent plane of the surface. For example, coordinates (σ_1, σ_2) may refer to the base vectors $(\vec{r}_u, \vec{r}_v)$ in Figure 28a, and (11.8) remains valid. In the latter case, $\bar{\bar{Q}}$ is no longer symmetrical, and is given by the general form in (10.10).



(a) $\kappa_G > 0$ with $\kappa_1 < 0$ and $\kappa_2 < 0$



(b) $\kappa_G < 0$ with $\kappa_1 < 0$ and $\kappa_2 > 0$



(c) $\kappa_G = 0$ with $\kappa_1 > 0$ and $\kappa_2 = 0$

Figure 31: Sign of Gaussian curvature $\kappa_G = \kappa_1 \kappa_2$. In this sketch, $\hat{\sigma} = +\hat{N}$ is used.

12 Geodesics

For two given points P and Q on a surface Σ, what is the shortest arc joining them? If Σ is a plane, the answer is obvious: The shortest arc is the straight line segment joining P and Q. For a general surface, the shortest arc, if it exists, must be a **geodesic**. A straight line is characterized by the property that its curvature is zero, whereas geodesics are curves of zero geodesic curvature.

Let us first define geodesic curvature. Consider a curve C (not necessarily a geodesic) on Σ (Figure 32). The normal and tangent of C at a point P are denoted by $\hat{n}$ and $\hat{t}$, respectively. The curvature vector $\kappa\hat{n}$ of C lies in a plane perpendicular to $\hat{t}$, and can be resolved into two components

$$\kappa\hat{n} = \kappa_n \hat{N} + \kappa_g \hat{u} \tag{12.1}$$

where $\hat{N}$ is the normal to Σ at P (pointing to any one of the two possible directions), and $\hat{u} = \hat{N} \times \hat{t}$. We called κ_g in (12.1) the **geodesic curvature** of curve C at point P. From Figure 32, we note that, with a possible minus sign, κ_g is the curvature of curve C', which is the orthogonal projection of C on the tangent plane. As an example, consider the small circle C on a sphere of radius a (Figure 33). At any point on C, its geodesic curvature is

$$\kappa_g = \kappa\hat{n} \cdot \hat{u} = \frac{1}{a}\cos\theta \tag{12.2}$$

which varies between $+a^{-1}$ and $-a^{-1}$, and vanishes when C is a great circle ($\theta = \pi/2$).

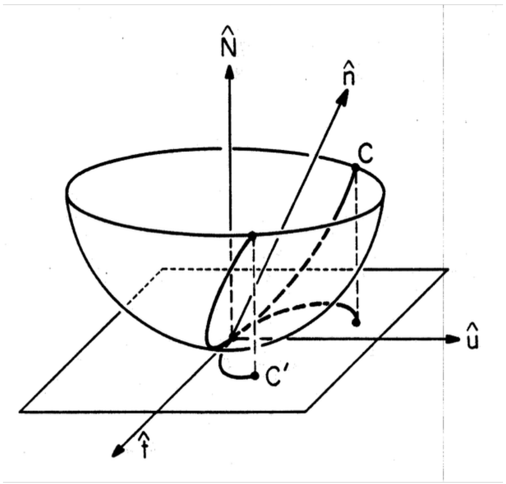


Figure 32: A curve C on a surface and its projection C' on the tangent plane.

In calculating κ_g, the following quantities, called **Christoffel symbols**, are needed:

$$\Gamma^1_{11} = \frac{GE_u - 2FF_u + FE_v}{2(EG - F^2)}, \qquad \Gamma^2_{11} = \frac{2EF_u - EE_v - FE_u}{2(EG - F^2)} \tag{12.3a}$$

$$\Gamma^1_{12} = \frac{GE_v - FG_u}{2(EG - F^2)}, \qquad \Gamma^2_{12} = \frac{EG_u - FE_v}{2(EG - F^2)} \tag{12.3b}$$

$$\Gamma^1_{22} = \frac{2GF_v - GG_u - FG_v}{2(EG - F^2)}, \qquad \Gamma^2_{22} = \frac{EG_v - 2FF_v + FG_u}{2(EG - F^2)} \tag{12.3c}$$

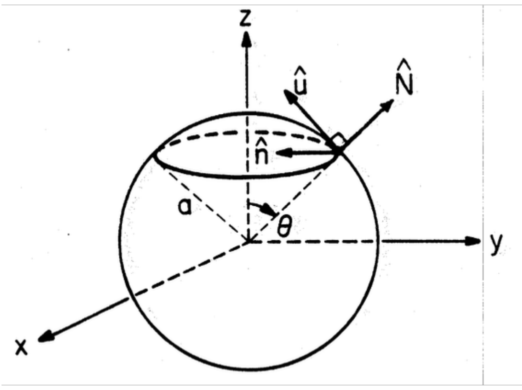


Figure 33: A small circle and its normal $\hat{n}$ on a sphere.

where the subscripts as usual indicate partial derivatives.

$$E_u = 2\vec{r}_u \cdot \vec{r}_{uu} \qquad E_v = 2\vec{r}_u \cdot \vec{r}_{uv} \tag{12.4a}$$

$$F_u = \vec{r}_u \cdot \vec{r}_{uv} + \vec{r}_v \cdot \vec{r}_{uu} \qquad F_v = \vec{r}_u \cdot \vec{r}_{vv} + \vec{r}_v \cdot \vec{r}_{uv} \tag{12.4b}$$

$$G_u = 2\vec{r}_v \cdot \vec{r}_{uv} \qquad G_v = 2\vec{r}_v \cdot \vec{r}_{vv} \tag{12.4c}$$

Note that Christoffel symbols depend only on the coefficients of the first fundamental form and their derivatives (not on e, f, g etc.). An explicit formula for κ_g in terms of Christoffel symbols can be found in standard differential geometry textbooks (See P. 128 of D. J. Struik).

The definition of geodesics is that they are curves of vanishing geodesic curvature

$$\kappa_g = 0 \tag{12.5}$$

From the condition in (12.5), two differential equations of the geodesic can be derived, namely

$$\frac{d^2u}{d\sigma^2} + \Gamma^1_{11}\left(\frac{du}{d\sigma}\right)^2 + 2\Gamma^1_{12}\frac{du}{d\sigma}\frac{dv}{d\sigma} + \Gamma^1_{22}\left(\frac{dv}{d\sigma}\right)^2 = 0 \tag{12.6a}$$

$$\frac{d^2v}{d\sigma^2} + \Gamma^2_{11}\left(\frac{du}{d\sigma}\right)^2 + 2\Gamma^2_{12}\frac{du}{d\sigma}\frac{dv}{d\sigma} + \Gamma^2_{22}\left(\frac{dv}{d\sigma}\right)^2 = 0 \tag{12.6b}$$

where σ is the arc length of a curve (geodesic). A solution of (12.6) is of the form

$$\begin{cases} u = f_1(\sigma) \\ v = f_2(\sigma) \end{cases}$$

which describes a geodesic. It should be remarked that (12.6a) and (12.6b) are not independent. They are related through the first fundamental form in (8.10) or

$$d\sigma^2 = Edu^2 + 2Fdudv + Gdv^2 \tag{12.7}$$

Eliminating $d\sigma$ from (12.6), we obtain a single differential equation of the geodesic, namely

$$\frac{d^2v}{du^2} = \Gamma^1_{22}\left(\frac{dv}{du}\right)^3 + 2(\Gamma^1_{12} - \Gamma^2_{22})\left(\frac{dv}{du}\right)^2 + (\Gamma^1_{11} - 2\Gamma^2_{12})\frac{dv}{du} - \Gamma^2_{11} \tag{12.8}$$

which has the following interpretation: At a given point (u, v) on a surface, once dv/du is given, d^2v/du^2 is determined, that is, the way in which the geodesic curve is continued. We list below several properties of geodesics:

(a) When $\kappa_g = 0$, we have from (12.1) that

$$\hat{n} = \pm\hat{N} \tag{12.9}$$

Thus, the normal of a geodesic is in the same (or opposite) direction of the surface normal.

(b) At a given point (initial point) on a surface, a geodesic is uniquely determined once a tangential direction is specified. This follows from (12.8) and an existence theorem of differential equations.

(c) For two given points P and Q on a surface, the minimum arc joining them, if it exists, must be a geodesic. The converse, however, is not true. For example, great circles of a sphere are geodesics. A great circle passing through P and Q has two arcs. In general, these two arcs are not equal and only one of them gives the shortest distance.

We conclude the section with several examples for determining geodesics.

(i) **Plane.** A point on the xy-plane may be described by

$$\vec{r}(x, y) = (x, y, z = \text{constant})$$

With $u = x$ and $v = y$, it may be shown that all Christoffel symbols in (12.3) vanish. Then (12.8) becomes

$$\frac{d^2y}{dx^2} = 0$$

Its solution is

$$y = ax + b$$

which is a straight line. Alternatively, the same problem may be solved by using cylindrical coordinates with

$$\vec{r}(\rho, \phi) = (\rho\cos\phi, \rho\sin\phi, z = \text{constant})$$

With $u = \rho$ and $v = \phi$, Christoffel symbols in (12.3) are all zeros except

$$\Gamma^1_{22} = -\rho, \quad \Gamma^2_{12} = \frac{1}{\rho}$$

Then (12.6) and (12.7) become

$$\frac{d^2\rho}{d\sigma^2} - \rho\left(\frac{d\phi}{d\sigma}\right)^2 = 0 \tag{12.10}$$

$$\frac{d^2\phi}{d\sigma^2} + \frac{2}{\rho}\frac{d\rho}{d\sigma}\frac{d\phi}{d\sigma} = 0 \tag{12.11}$$

$$d\sigma^2 = d\rho^2 + \rho^2 d\phi^2 \tag{12.12a}$$

As remarked earlier, not all the above three equations are independent. For example, let us write (12.12a) as

$$1 = \left(\frac{d\rho}{d\sigma}\right)^2 + \rho^2\left(\frac{d\phi}{d\sigma}\right)^2 \tag{12.12b}$$

Taking the derivative of both sides of (12.12b) with respect to σ gives

$$0 = \left(\frac{d\rho}{d\sigma}\right)\left[\frac{d^2\rho}{d\sigma^2} - \rho\left(\frac{d\phi}{d\sigma}\right)^2\right] + \rho\left(\frac{d\phi}{d\sigma}\right)\left[\rho\frac{d^2\phi}{d\sigma^2} + 2\frac{d\rho}{d\sigma}\frac{d\phi}{d\sigma}\right]$$

which is a proper combination of (12.10) and (12.11). Hence, we may concentrate on the solution of (12.11) and (12.12). Rewrite (12.11) as

$$\frac{(d^2\phi/d\sigma^2)}{(d\phi/d\sigma)}+\frac{2}{\rho}\frac{d\rho}{d\sigma}=0 \tag{12.13a}$$

or

$$\frac{d}{d\sigma}\left[\ln\left(\rho^2\frac{d\phi}{d\sigma}\right)\right]=0 \tag{12.13b}$$

whose solution is

$$d\sigma=\frac{1}{c}\rho^2 d\phi, \quad \text{where } c=\text{constant} \tag{12.14}$$

Substituting (12.14) into (12.12a), we have

$$\frac{c\,d\rho}{\rho\sqrt{\rho^2-c^2}}=d\phi \tag{12.15}$$

Integrating both sides of (12.15) leads to the solution

$$\rho=c\sec(\phi-\phi_0) \tag{12.16}$$

where ϕ_0 is another constant. As expected, (12.16) is also a straight line.

(ii) **Cylinder and developable surfaces.** Consider the helix curve on the surface of a cylinder shown in Figure 8. At every point the normal $\hat{n}$ of the helix is equal to $\pm\hat{N}$ of the cylinder. Therefore, by (12.9), a helix is a geodesic on a cylinder. For any two given points P and Q on a cylinder (Figure 34a), there are infinitely many helix curves (geodesics) joining them. This point may be best explained by using a "**developed**" cylinder. A **developable surface** is the one which ‖

(a) may be generated by a continuous motion of a straight line (the straight lines on the surface are called **generators**), and

(b) has the same tangent plane at all points on any given generator.

Examples of developable surfaces are cylinders and cones. If one cuts the cylinder along a generator ($\phi = 0$ in Figure 34a), it may be opened up to become a rectangle on a plane, without stretching or shrinking (Figure 34b). A helix on a cylinder now becomes a straight line on the developed cylinder. To account for the periodicity in ϕ, the rectangle may be repeated an infinite number of times in the manner shown in Figure 35. Then P has images $P_1, P_{-1}, P_2, P_{-2}, \ldots$. All possible geodesics joining P and Q on a cylinder may be constructed by drawing line segments between Q and $P, P_1, P_{-1}, \ldots$. For example, geodesic QP_1 goes from Q to P on the cylinder after revolving the cylinder once; and geodesic QP_{-2} goes from Q and P after revolving the cylinder twice in the opposite direction.

(iii) **Surface of revolution.** As discussed in Section 7 (Figure 18), a surface of revolution may be described by

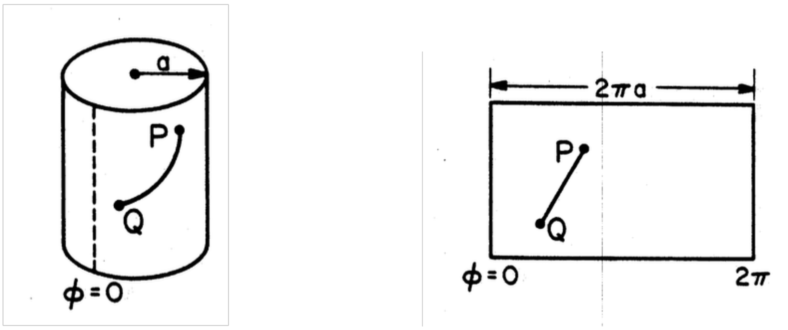
$$x=\rho\cos\phi, \quad y=\rho\sin\phi, \quad z=f(\rho) \tag{12.17}$$

Simple calculations lead to ($u=\rho$ and $v=\phi$)

$$E=1+(f')^2, \quad F=0, \quad G=\rho^2 \tag{12.18a}$$

$$e=\frac{f''}{\sqrt{1+(f')^2}}, \quad f=0, \quad g=\frac{\rho f'}{\sqrt{1+(f')^2}} \tag{12.18b}$$

‖ Surfaces with property (a) but not necessarily (b) are called **ruled surfaces**. A hyperbolic paraboloid (Figure 14), or a hyperboloid of one sheet (Figure 16) is a ruled surface, but not developable.



(a) **(b)**

Figure 34: A cylinder may be developed into a rectangle on a plane.

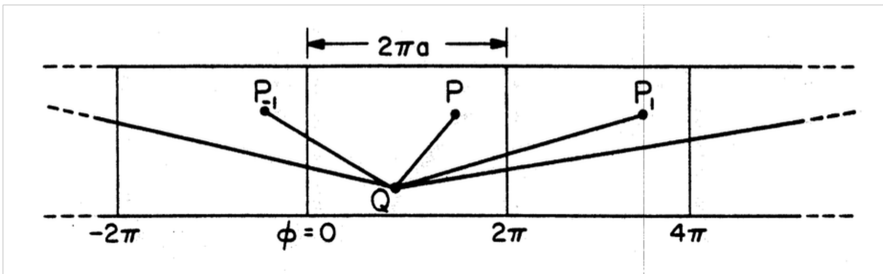


Figure 35: Developed cylinder and its "images."

where f' is the derivative of $f(\rho)$ with respect to ρ. Then Christoffel symbols in (12.3) are found to be

$$\Gamma^1_{11} = \frac{f'f''}{1+(f')^2}, \quad \Gamma^2_{11} = 0 \tag{12.19a}$$

$$\Gamma^1_{12} = 0, \quad \Gamma^2_{12} = \frac{1}{\rho} \tag{12.19b}$$

$$\Gamma^1_{22} = \frac{-\rho}{1+(f')^2}, \quad \Gamma^2_{22} = 0 \tag{12.19c}$$

Since the two equations in (12.6) are dependent, it is sufficient to take (12.6b), namely

$$\frac{d^2\phi}{d\sigma^2} + \frac{2}{\rho}\frac{d\rho}{d\sigma}\frac{d\phi}{d\sigma} = 0 \tag{12.20}$$

The above differential equation is the same as (12.13), whose solution is given in (12.14). Substituting (12.14) into the first fundamental form in (12.7) or

$$d\sigma^2 = [1+(f')^2]d\rho^2 + \rho^2 d\phi^2 \tag{12.21}$$

we have

$$d\phi = \frac{c}{\rho}\left[\frac{1+(f')^2}{\rho^2-c^2}\right]^{1/2} d\rho$$

or

$$\phi = \phi_0 + c\int \frac{1}{\rho}\left[\frac{1+(f')^2}{\rho^2-c^2}\right]^{1/2} d\rho \tag{12.22}$$

which is the desired equation of a geodesic defined by two constants ϕ_0 and c. For the special case $c=0$, (12.22) becomes

$$\phi = \phi_0 \tag{12.23}$$

Thus, all generating lines (meridians) of a surface of revolution are geodesics. As an example, let the surface be a circular cone with a half-cone angle θ_0. Then

$$f(\rho) = \rho \cot\theta_0 \tag{12.24}$$

Substituting (12.24) into (12.22) and carrying out the integral, we obtain the equation of a geodesic on a cone, namely

$$\rho = c \sec[(\phi - \phi_0)\sin\theta_0] \tag{12.25}$$

Using the spherical coordinate $r = \rho \csc\theta_0$, we may rewrite (12.25) as

$$r \cos[(\phi - \phi_0)\sin\theta_0] = r_0 \tag{12.26}$$

where r_0 is another constant. On a developed cone in Figure 36, (12.28) represents a straight line MP with $r_0 = \overline{OM}$ and $r = \overline{OP}$. The fact that the geodesics of a cone are straight lines on a developed cone is in agreement with our discussion in Example (ii).

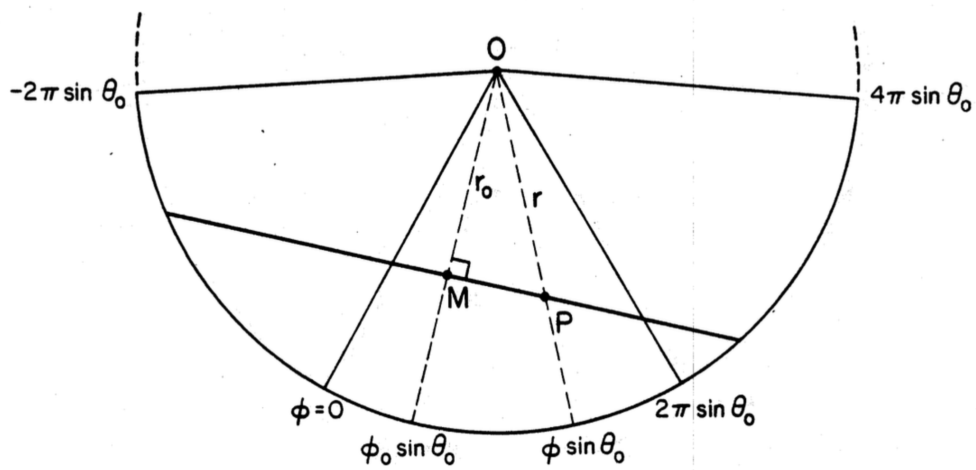


Figure 36: A geodesic MP on a developed cone and its "images."

13 Summary

Let us first summarize the results in Sections 2 to 6 for a curve:

(1) An arbitrary-speed curve can be represented by a parametric equation with parameter t:

$$\vec{r}(t) = (x(t), y(t), z(t)), \quad t_1 < t < t_2. \tag{13.1}$$

Its speed is defined by

$$v = \left|\frac{d\vec{r}}{dt}\right| = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}. \tag{13.2}$$

The same curve may have other representations with different speeds through a change of parameter (passing t to a new parameter).

(2) The arc length $\sigma(t)$ of a curve is defined by

$$\sigma(t) = \int_{t_0}^{t} \left|\frac{d\vec{r}}{dt}\right| dt \tag{13.3}$$

in which t_0 is a reference point.

(3) A unit-speed curve is a curve $\vec{r}(\sigma)$ whose parameter is arc length σ. The Frenet apparatus can be computed by the following formulas:

$$\hat{n}(\sigma) = \frac{d^2\vec{r}}{d\sigma^2} \Bigg/ \left|\frac{d^2\vec{r}}{d\sigma^2}\right| \tag{13.4a}$$

$$\hat{b}(\sigma) = \hat{t} \times \hat{n} \tag{13.4b}$$

$$\hat{t}(\sigma) = \frac{d\vec{r}}{d\sigma} \tag{13.4c}$$

$$\kappa(\sigma) = \left|\frac{d^2\vec{r}}{d\sigma^2}\right| \tag{13.5a}$$

$$\tau(\sigma) = \frac{d\vec{r}}{d\sigma} \cdot \left(\frac{d^2\vec{r}}{d\sigma^2} \times \frac{d^3\vec{r}}{d\sigma^3}\right) \Bigg/ \left|\frac{d^2\vec{r}}{d\sigma^2}\right|^2 \tag{13.5b}$$

The variation of $(\hat{n}, \hat{b}, \hat{t})$ is given by the Frenet formula:

$$\frac{d\hat{n}}{d\sigma} = \tau\hat{b} - \kappa\hat{t} \tag{13.6}$$

$$\frac{d\hat{b}}{d\sigma} = -\tau\hat{n} \tag{13.7}$$

$$\frac{d\hat{t}}{d\sigma} = \kappa\hat{n} \tag{13.8}$$

(4) For an arbitrary-speed curve $\vec{r}(t)$ the formulas for the Frenet apparatus are given by

$$\hat{n} = \hat{b} \times \hat{t} \tag{13.9a}$$

$$\hat{b}(t) = \frac{\vec{r}\,' \times \vec{r}\,''}{|\vec{r}\,' \times \vec{r}\,''|} \tag{13.9b}$$

$$\hat{t}(t) = \frac{\vec{r}\,'}{|\vec{r}\,'|} \tag{13.9c}$$

$$\kappa(t) = \frac{|\vec{r}' \times \vec{r}''|}{|\vec{r}'|^3} \tag{13.10a}$$

$$\tau(t) = \frac{(\vec{r}' \times \vec{r}'') \cdot \vec{r}'''}{|\vec{r}' \times \vec{r}''|^2} \tag{13.10b}$$

where

$$\vec{r}' = \frac{d\vec{r}(t)}{dt}, \vec{r}'' = \frac{d^2\vec{r}(t)}{dt^2}, \dots \tag{13.11}$$

(5) For a unit-speed curve $\vec{r}(\sigma)$, the Frenet approximation of the curve in the neighborhood of $\sigma = \sigma_0$ is (Figure 9)

$$\vec{r}(\sigma) \approx \vec{r}(\sigma_0) + (\sigma - \sigma_0)\hat{t} + \frac{(\sigma - \sigma_0)^2}{2}\kappa\hat{n} + \frac{(\sigma - \sigma_0)^3}{6}\kappa\tau\hat{b} \tag{13.12}$$

where $(\hat{n}, \hat{b}, \hat{t})$ are evaluated at $\sigma = \sigma_0$.

Next, let us summarize the results in Sections 7 through 12 for a surface:

(6) A surface can be represented by a parametric equation with parameters (u, v):

$$\vec{r}(u, v) = (x(u, v), y(u, v), z(u, v)), \quad u_1 < u < u_2 \text{ and } v_1 < v < v_2. \tag{13.13}$$

The unit surface normal at (u, v) is defined by

$$\hat{N}(u, v) = \mu\frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|} \tag{13.14}$$

where $\mu = \pm 1$. In EM diffraction problems, μ takes a value such that $\hat{N}$ always points toward the source.

(7) In calculating the properties of the surface in (13.13), the following parameters are often needed:

- **Coefficients of first and second fundamental forms**

$$E = \vec{r}_u \cdot \vec{r}_u, \quad F = \vec{r}_u \cdot \vec{r}_v, \quad G = \vec{r}_v \cdot \vec{r}_v \tag{13.15a}$$

$$e = \mu\frac{\vec{r}_{uu} \cdot (\vec{r}_u \times \vec{r}_v)}{\sqrt{EG - F^2}}, \quad f = \mu\frac{\vec{r}_{uv} \cdot (\vec{r}_u \times \vec{r}_v)}{\sqrt{EG - F^2}}, \quad g = \mu\frac{\vec{r}_{vv} \cdot (\vec{r}_u \times \vec{r}_v)}{\sqrt{EG - F^2}} \tag{13.15b}$$

- **Christoffel symbols**

$$\Gamma_{11}^1 = \frac{GE_u - 2FF_u + FE_v}{2(EG - F^2)}, \quad \Gamma_{11}^2 = \frac{2EF_u - EE_v - FE_u}{2(EG - F^2)} \tag{13.16a}$$

$$\Gamma_{12}^1 = \frac{GE_v - FG_u}{2(EG - F^2)}, \quad \Gamma_{12}^2 = \frac{EG_u - FE_v}{2(EG - F^2)} \tag{13.16b}$$

$$\Gamma_{22}^1 = \frac{2GF_v - GG_u - FG_v}{2(EG - F^2)}, \quad \Gamma_{22}^2 = \frac{EG_v - 2FF_v + FG_u}{2(EG - F^2)} \tag{13.16c}$$

When the surface is described by the special form

$$\vec{r}(x, y) = (x, y, f(x, y)) \tag{13.17}$$

the above parameters become

$$E = 1 + f_x^2, \quad F = f_x f_y, \quad G = 1 + f_y^2 \tag{13.18a}$$
$$e = \Delta f_{xx}, \quad f = \Delta f_{xy}, \quad g = \Delta f_{yy} \tag{13.18b}$$
$$\Delta = (1 + f_x^2 + f_y^2)^{-1/2} \tag{13.19}$$
$$\Gamma_{11}^1 = \Delta^2 f_x f_{xx}, \quad \Gamma_{11}^2 = \Delta^2 f_y f_{xx} \tag{13.20a}$$
$$\Gamma_{12}^1 = \Delta^2 f_x f_{xy}, \quad \Gamma_{12}^2 = \Delta^2 f_y f_{xy} \tag{13.20b}$$
$$\Gamma_{22}^1 = \Delta^2 f_x f_{yy}, \quad \Gamma_{22}^2 = \Delta^2 f_y f_{yy} \tag{13.20c}$$

When the surface is one of revolution, those parameters are given in (12.18) and (12.19).

(8) The curvature κ and the radius of curvature R in the direction dv/du are given by

$$\kappa = \frac{1}{R} = \frac{e(du)^2 + 2f\,du dv + g(dv)^2}{E(du)^2 + 2F\,du dv + G(dv)^2}. \tag{13.21}$$

The sign of κ (or R) computed above is positive if the normal section of the surface bends toward $\hat{N}$, and is negative if the normal section bends away from $\hat{N}$ (Figure 21).

(9) At any point on a surface (except at an umbilic) a pair of orthogonal directions exists for which κ assumes maximum and minimum values. These two directions are called **principal directions** $(\hat{e}_1, \hat{e}_2)$, and the two extreme values of κ are called **principal curvatures** (κ_1, κ_2). The principal curvatures are given by

$$\kappa_{1,2} = \kappa_M \pm \sqrt{\kappa_M^2 - \kappa_G} \tag{13.22}$$

where the mean curvature is

$$\kappa_M = \frac{\kappa_1 + \kappa_2}{2} = \frac{Eg - 2fF + eG}{2(EG - F^2)} \tag{13.23}$$

and the Gaussian curvature is

$$\kappa_G = \kappa_1 \kappa_2 = \frac{eg - f^2}{EG - F^2}. \tag{13.24}$$

The two principal directions are given by

$$\hat{e}_1 = \frac{1}{\gamma_1}[1\vec{r}_u + \alpha \vec{r}_v] \tag{13.25a}$$
$$\hat{e}_2 = \frac{1}{\gamma_2}[\beta \vec{r}_u + 1\vec{r}_v] \tag{13.25b}$$

where

$$\alpha = \frac{e - \kappa_1 E}{\kappa_1 F - f} = \frac{f - \kappa_1 F}{\kappa_1 G - g} \tag{13.25c}$$
$$\beta = \frac{f - \kappa_2 F}{\kappa_2 E - e} = \frac{g - \kappa_2 G}{\kappa_2 F - f} \tag{13.25d}$$
$$\gamma_1 = (E + 2\alpha F + \alpha^2 G)^{1/2} \tag{13.25e}$$
$$\gamma_2 = (\beta^2 E + 2\beta F + G)^{1/2} \tag{13.25f}$$

Referring to the principal direction $\hat{e}_1$, we find the curvature in the direction α (Figure 24) is given by

$$\kappa = \kappa_1 \cos^2 \alpha + \kappa_2 \sin^2 \alpha \tag{13.26}$$

(10) A necessary and sufficient condition for u- and v-parameter curves is also the lines of curvatures: $F = 0$ and $f = 0$. When these two conditions are met, then

$$\hat{e}_1 = \frac{\vec{r}_u}{|\vec{r}_u|}, \quad \hat{e}_2 = \frac{\vec{r}_v}{|\vec{r}_v|} \tag{13.27}$$

(11) The 2×2 **curvature matrix** $\bar{\bar{Q}}$ is defined on the tangent plane of a surface by the definition in (10.2). Its elements may be computed from the formula

$$Q_{11} = \frac{eG - fF}{EG - F^2}, \quad Q_{12} = \frac{fE - eF}{EG - F^2} \tag{13.28a}$$

$$Q_{21} = \frac{fG - gF}{EG - F^2}, \quad Q_{22} = \frac{gE - fF}{EG - F^2}. \tag{13.28b}$$

(12) Instead of using the formulas in (9), the principal curvatures and directions can be calculated from $\bar{\bar{Q}}$:

$$\kappa_M = \frac{1}{2}(Q_{11} + Q_{22}) \tag{13.29a}$$

$$\kappa_G = Q_{11}Q_{22} - Q_{12}Q_{21} \tag{13.29b}$$

$$\alpha = \frac{Q_{22} - \kappa_2}{Q_{21}} = \frac{Q_{12}}{Q_{11} - \kappa_2} \tag{13.29c}$$

$$\beta = \frac{Q_{11} - \kappa_1}{Q_{12}} = \frac{Q_{21}}{Q_{22} - \kappa_1} \tag{13.29d}$$

Substitution of (13.29) into (13.22) and (13.25) gives the principal curvatures and directions.

(13) To the first-order approximation, surface normals drawn from adjacent points on a line of curvature intersect. The intersecting point is located on one of the two **caustic surfaces** (Figure 27).

(14) Consider a **conducting surface** Σ in an EM diffraction problem. At a point O (point of reflection) on Σ, let us introduce three right-handed orthonormal base vectors $(\hat{\sigma}_1, \hat{\sigma}_2, \hat{s})$ such that $\hat{s} = +\hat{N}$, pointing toward the source of the incident field (Figure 29a). Then a quadratic approximation of Σ in the neighborhood of O is

$$\Sigma : \sigma = +\frac{1}{2}\begin{bmatrix}\sigma_1\\ \sigma_2\end{bmatrix}^T \cdot \bar{\bar{Q}} \begin{bmatrix}\sigma_1\\ \sigma_2\end{bmatrix} + o(\sigma_{1,2}^3) \tag{13.30}$$

If $(\hat{\sigma}_1, \hat{\sigma}_2)$ coincide with the principal directions, $\bar{\bar{Q}}$ is a **diagonal matrix** given by

$$\bar{\bar{Q}} = \begin{bmatrix}\kappa_1 & 0\\ 0 & \kappa_2\end{bmatrix} \tag{13.31}$$

where (κ_1, κ_2) are **positive (negative)** if their normal sections bend toward (away from) $\hat{s} = \hat{N}$. If $(\hat{\sigma}_1, \hat{\sigma}_2)$ make an angle ψ with principal directions, then

$$\bar{\bar{Q}} = \begin{bmatrix}\cos\psi & -\sin\psi\\ \sin\psi & \cos\psi\end{bmatrix}\begin{bmatrix}\kappa_1 & 0\\ 0 & \kappa_2\end{bmatrix}\begin{bmatrix}\cos\psi & -\sin\psi\\ \sin\psi & \cos\psi\end{bmatrix}^T \tag{13.32}$$

(15) Consider a **wavefront** W of a ray pencil in an EM diffraction problem. At a point O (a reference point on the axial ray) on W, let us introduce three right-handed orthonormal

base vectors $(\hat{\sigma}_1, \hat{\sigma}_2, \hat{s})$ such that $\hat{s} = -\hat{N}$, pointing toward the direction of wave propagation (Figure 29b). Then a quadratic approximation of W in the neighborhood of O is

$$W : \sigma = -\frac{1}{2}\begin{bmatrix}\sigma_1\\ \sigma_2\end{bmatrix}^T \cdot \bar{\bar{Q}} \begin{bmatrix}\sigma_1\\ \sigma_2\end{bmatrix} + o(\sigma_{1,2}^3) \tag{13.33}$$

where $\bar{\bar{Q}}$ is given again by (13.31) or (13.32). The principal curvatures (κ_1, κ_2) are **positive (negative)** if their normal sections of the wavefront are **divergent (convergent)**.

(16) A **geodesic** on a surface is a curve of vanishing geodesic curvature. It may be described by any one of the following three differential equations:

$$\frac{d^2u}{d\sigma^2} + \Gamma_{11}^1\left(\frac{du}{d\sigma}\right)^2 + 2\Gamma_{12}^1\frac{du}{d\sigma}\frac{dv}{d\sigma} + \Gamma_{22}^1\left(\frac{dv}{d\sigma}\right)^2 = 0 \tag{13.34a}$$

$$\frac{d^2v}{d\sigma^2} + \Gamma_{11}^2\left(\frac{du}{d\sigma}\right)^2 + 2\Gamma_{12}^2\frac{du}{d\sigma}\frac{dv}{d\sigma} + \Gamma_{22}^2\left(\frac{dv}{d\sigma}\right)^2 = 0 \tag{13.34b}$$

$$\frac{d^2v}{du^2} = \Gamma_{22}^1\left(\frac{dv}{du}\right)^3 + (2\Gamma_{12}^1 - \Gamma_{22}^2)\left(\frac{dv}{du}\right)^2 + (\Gamma_{11}^1 - 2\Gamma_{12}^2)\frac{dv}{du} - \Gamma_{11}^2 \tag{13.35}$$

where σ is the arc length of the geodesic. With the aid of the first fundamental form

$$d\sigma^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2 \tag{13.36}$$

we may eliminate σ in (13.34) and recover (13.35).

(17) For two points on a surface, the minimum arc joining them, if it exists, must be a geodesic.

(18) For a surface of revolution, the differential equation in (13.31) has been solved, and the solution for the geodesic is given in (12.22).

Subject Index