Chapter 4

Tensors in Generalized Coordinates in Three Dimensions

1. Coordinate Curves and Coordinate Surfaces

Rectilinear coordinates in three dimensions have been defined in terms of three non-coplanar straight lines which intersect in a common point, the origin. The three lines are the coordinate axes. Taken by pairs, these axes also define the coordinate planes. Thus in Cartesian coordinates we have the mutually perpendicular $x-$, $y-$ and $z-$axes which define the $xy-$, $xz-$and $yz-$planes. Along each axis, one coordinate varies, the other two are zero. The axes may therefore be called **coordinate curves**. In the coordinate planes, one coordinate is fixed, the other two vary. The planes are therefore called **coordinates surfaces**.

More generally, we see that the lines defined by (1) $x^2 = b$, $x^3 = c$, (2) $x^1 = a$, $x^3 = c$ and (3) $x^1 = a$, $x^2 = b$ form a triply infinite family of coordinate curves which fill all of three-dimensional Euclidean space. Associated with them is the triply infinite family of planes (1) $x^1 = a$, (2) $x^2 = b$ and (3) $x^3 = c$. If the lines or planes are mutually orthogonal, the coordinate system is Cartesian; otherwise it is merely rectilinear.

Let us now waive the condition that the coordinate surfaces need be planes. Let them be simply any triply infinite families of surfaces such that (1) each family fills all of space or at least all of that region of space which is of interest and (2) the members of any one of the three families do not intersect others of the same family but do intersect members of the other families. In that event, any point of space may be located as being at the mutual intersection of three particular coordinate surfaces. The parameters of the three particular surfaces are the **generalized coordinates** of the point in question.

The intersections of the coordinate surfaces by pairs define the three coordinate curves. They thus consist of three families of curves such that one and only one member of each family passes through each point of space. As an example, consider the familiar spherical coordinate system. The coordinate surfaces are (1) all spheres about the origin, (2) all right circular cones with a common axis and a common vertex at the origin, and (3) all planes through the cones' common axis (see Fig. 71).

The coordinate curves are (1) all straight lines through the origin (radii), (2) all circles with centers on the axis and planes perpendicular to the axis (parallels of latitude), and (3) all circles with center at the origin and diameters along the axis (meridians of longitude).

As was the case in two dimensions, generalized or **curvilinear coordinates** are themselves not vector components, in general. Though every set of coordinates locates a point of space uniquely, those coordinates are in general not the components of a

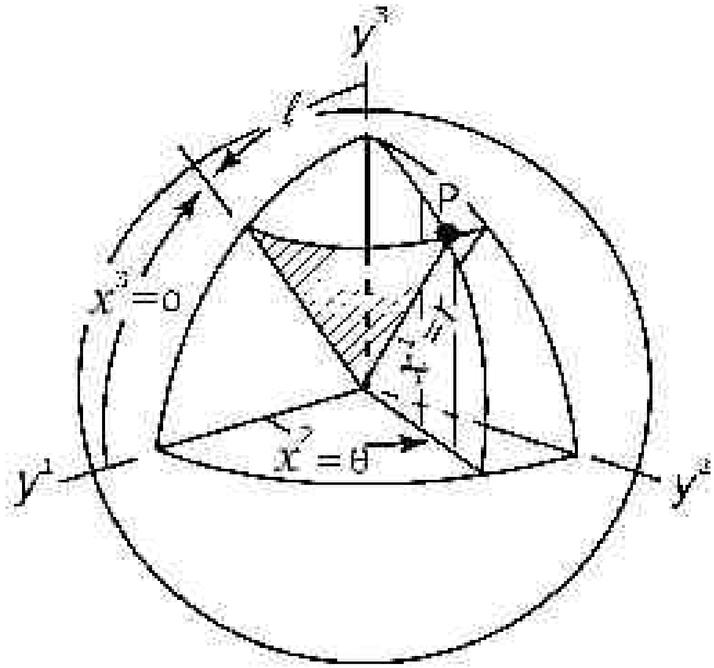

Figure 71

position vector to that point. This is one immediate and important distinction between rectilinear and curvilinear coordinates. It was on this account, for example, that the equations of motion of rigid bodies, which are extended objects, are more easily treated in rectilinear coordinates. On the other hand, various other types of problems are to be treated more conveniently in certain suitably chosen generalized coordinate systems.

Though the coordinates themselves are not position vector components in generalized coordinate systems, the differentials of the coordinates invariably are, for if $\bar{x}^r = \bar{x}^r(x^s)$ are new coordinates given as functions of old coordinates x^s, then

(1.1) $$d\bar{x}^r = \frac{\partial \bar{x}^r}{\partial x^s} dx^s$$

is not only the law of transformation of differentials but the contravariant vector transformation law as well, where the differential vector dx^s is used to find the components $d\bar{x}^r$ in the new coordinate system.

This connection between the transformation of differentials and the transformation law for differential vectors is no coincidence. It stems from the fact that at any point O the tangents to the coordinate curves through that point may be used to define a local rectilinear system x^i (see Fig. 72). The respective coordinate differentials at that point lie along these rectilinear axes, thus constitute the components of a differential contravariant vector. The tangent planes to the coordinate surfaces at O are the coordinate surfaces of this associated rectilinear system. We may thus define vectors and tensors of every sort in this rectilinear system and subject them to all the defined algebraic manipulations that were possible in rectilinear systems in general.

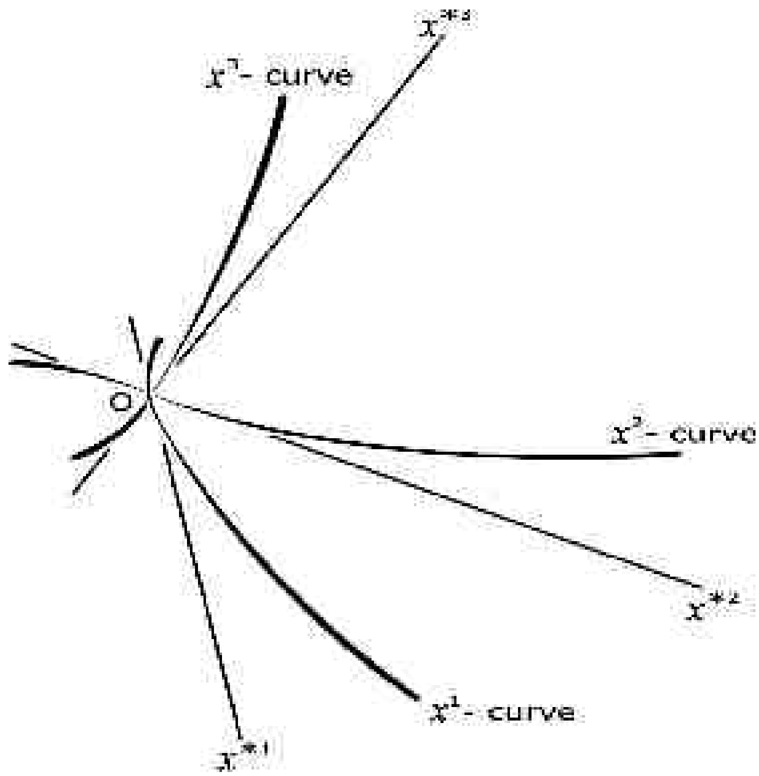

Figure 72

We now associate all vector and tensor quantities defined at O in the tangent rectilinear system with the curvilinear coordinate system itself. Any reversible transformation of coordinates will at most simply define a new tangent rectilinear system at O. Between this and the former system, the usual tensor transformation holds. It is necessary only that the transformation should be reversible at O, for which it is both necessary and sufficient that

$$\left| \frac{\partial \bar{x}^r}{\partial x^s} \right| \neq 0 . \tag{1.2}$$

The operations permitted by the previous argument all had to be performed strictly at the point O. These include addition, subtraction, multiplication, and the formation of inner and vector products. Not included was the operation of differentiation, which requires the formation of differences of vectors *not at the same point*. This same problem arose in considering curvilinear coordinates in two dimensions and may here be solved in an exactly analogous manner. The result is formally identical, and leads to the same definition of intrinsic and covariant differentiation. Hence these operations and associated formulae will be taken over intact. Only one minor difference need be noted, namely, that the indices of the Christoffel three-index symbols may now be all distinct, an impossibility in two dimensions.

Ex. (1.1) (a) Identify the coordinate curves and coordinate surfaces in a ***cylindrical coordinate system*** $\bar{x}^i$ *related to a Cartesian coordinate system* x^i *by the transformations*

$$0 \le \bar{x}^1 = r = \sqrt{(x^1)^2 + (x^2)^2} = \sqrt{(x)^2 + (y)^2} < \infty ,$$

$$0 \le \theta = \bar{x}^2 = \tan^{-1}\frac{x^2}{x^1} = \tan^{-1}\frac{y}{x} < 2\pi ,$$

$$-\infty < \bar{x}^3 = x^3 = z < \infty .$$

(b) Find the inverse transformations.

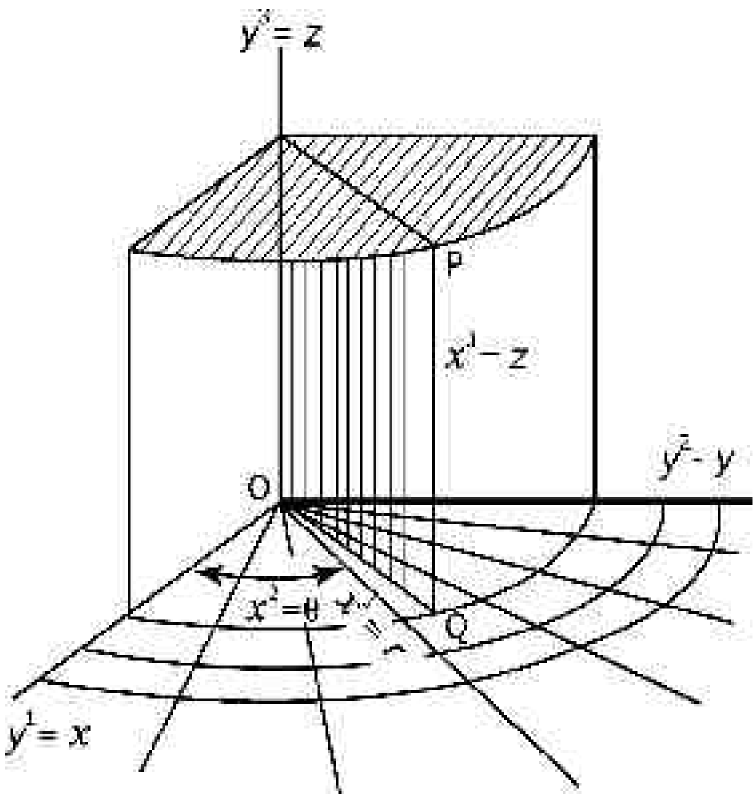

Figure 73

Ans. (a) The x^1*– curves (see Fig. 73) are any straight lines through the* x^3*– axis and perpendicular to it; the* x^2*– curves are circles about the* x^3*– axis and lying in a plane perpendicular to the* x^3*– axis; the* x^3*– curves are straight lines parallel to the* x^3*– axis. An* x^1*– surface is a right circular cylinder whose axis is the* x^3*– axis; an* x^2*– surface is any plane containing the* x^3*– axis; an* x^3*– surface is any plane perpendicular to the* x^3*– axis.*

(c) $x = x^1 = \bar{x}^1 \cos \bar{x}^2 = r \cos\theta$, $y = x^2 = \bar{x}^1 \sin \bar{x}^2 = r \sin\theta$,

$$z = x^3 = \bar{x}^3 .$$

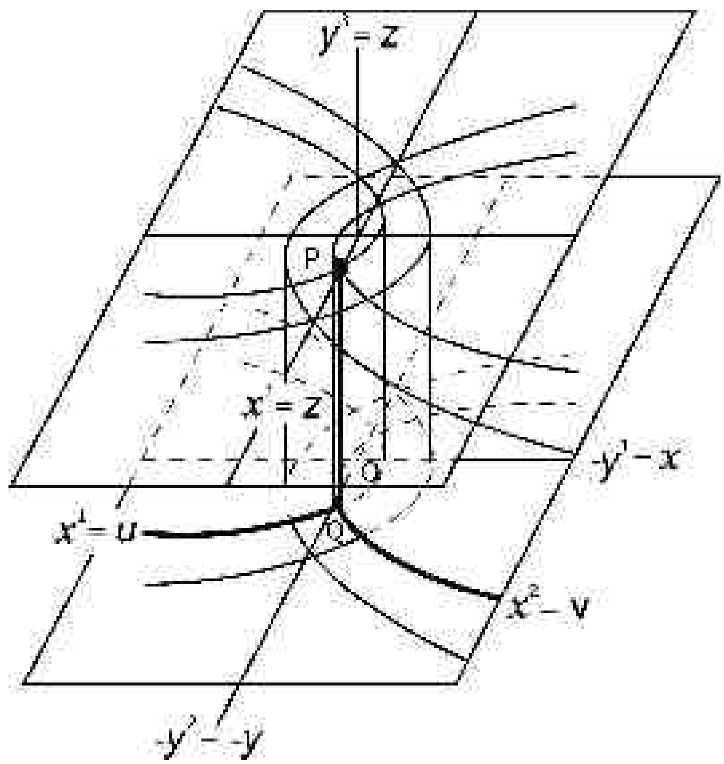

Figure 74

Ex. (1.2) (a) Identify the coordinates surfaces in a ***parabolic cylindrical coordinate system*** $\bar{x}^i$ *which is related to Cartesian coordinates* x^i *by the transformation*

$$\bar{x}^1 = u = \left[\sqrt{(x^1)^2 + (x^2)^2} + x^1\right]^{1/2} = \left[\sqrt{(x)^2 + (y)^2} + x\right]^{1/2}, \; -\infty < u < \infty,$$

$$\bar{x}^2 = \mathrm{v} = \left[\sqrt{(x^1)^2 + (x^2)^2} - x^1\right]^{1/2} = \left[\sqrt{(x)^2 + (y)^2} - x\right]^{1/2}, \; 0 \le \mathrm{v},$$

$$-\infty < \bar{x}^3 = z = x^3 < \infty .$$

(b) Find the inverse transformation.

Ans. (a) Setting $x^1 = u = \alpha = constant$ *(see part (b) below), we have*

$$x - \frac{\alpha^2}{2} = -(\mathrm{v})^2, \; y = \alpha \mathrm{v}, \; z = z,$$

which is the parametric representation of a family of parabolic cylinders opening to the negative $x-$*axis. Setting* $x^2 = \mathrm{v} = \beta = constant$, *we have*

$$x + \frac{\beta^2}{2} = u^2, \; y = \beta u, \; z = z,$$

which is the parametric representation of a family of parabolic cylinders opening to the positive $x-$axis. Setting $z = constant$ gives a plane perpendicular to the $z-$axis.

$$(b)\ x^1 = x = \frac{1}{2}(u^2 - v^2) = \frac{1}{2}\left[(\bar{x}^1)^2 - (\bar{x}^2)^2\right],$$
$$x^2 = y = u\,v = \bar{x}^1 \bar{x}^2,\ x^3 = z = \bar{z}.$$

Ex. (1.3) (a) What are the coordinate surfaces in a ***paraboloidal coordinate system*** *(see Fig. 74) $\bar{x}^i$ which is related to Cartesian coordinates by the transformations*

$$x^1 = x = u\,v\cos\theta = \bar{x}^1\bar{x}^2\cos\bar{x}^3,\ u \ge 0,\ v \ge 0,$$
$$x^2 = y = u\,v\sin\theta = \bar{x}^1\bar{x}^2\sin\bar{x}^3,\ 0 \le \theta \le \pi,$$
$$x^3 = z = u^2 - v^2 = (\bar{x}^1)^2 - (\bar{x}^2)^2.$$

(b) What is the inverse transformation?

Ans. (a) From the equation $z = u^2 - v^2$ it is clear that $v = constant$ is a paraboloid of revolution opening toward the positive z-axis, whereas $u = constant$ is a paraboloid of revolution opening toward the negative z-axis. From the first two equations of the transformation it is clear that

$$\bar{x}^3 = \theta = \tan^{-1}\frac{y}{x} = constant$$

is a plane containing the z-axis and making a dihedral angle θ with the xz-plane (see Fig. 75).

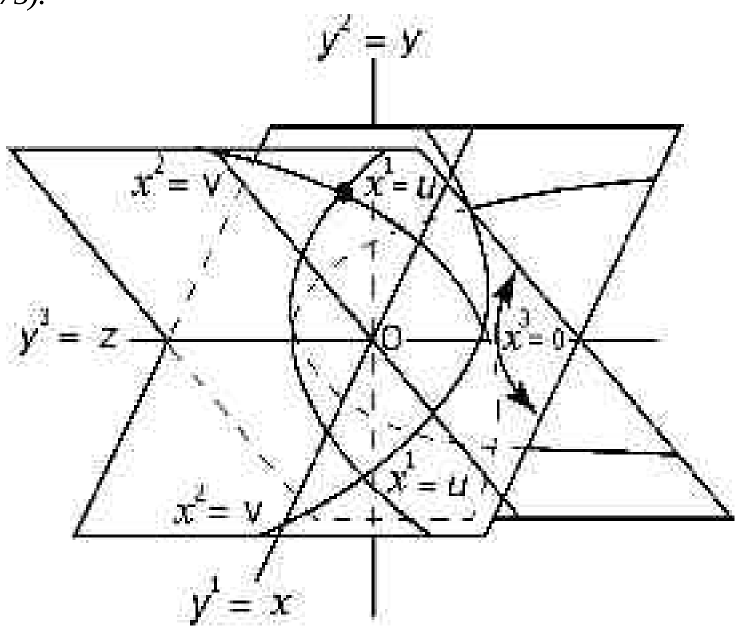

Figure 75

$$(b) \qquad \bar{x}^1 = u = \left[\frac{z}{2} \pm \sqrt{\left(\frac{z}{2}\right)^2 + \left(x^2 + y^2\right)} \right]^{1/2},$$

$$\bar{x}^2 = \mathrm{v} = \left[-\frac{z}{2} \pm \sqrt{\left(\frac{z}{2}\right)^2 + \left(x^2 + y^2\right)} \right]^{1/2}.$$

Ex. (1.4) (a) From the equations of transformation of plane bipolar coordinates, determine the transformation for ***bipolar cylindrical coordinates****. (b) Identify the coordinate surfaces. (c) What are the coordinate curves?*

Ans. (a)

$$\bar{x}^1 = \rho = \left[(x - a)^2 + (y)^2\right]^{1/2} \geq 0,$$

$$\bar{x}^2 = \sigma = \left[(x + a)^2 + (y)^2\right]^{1/2} \geq 0,$$

$$-\infty < \bar{x}^3 = z < +\infty.$$

(b) The $\bar{x}^1$*– surfaces are right circular cylinders with axes parallel to the* z*-axis and through the point* $x^i = (a, 0, 0)$*. The* $\bar{x}^2$*– surfaces are also right circular cylinders whose axes are the line* $x^1 = x = -a$, $x^2 = y = 0$*. The* $\bar{x}^3$*– surfaces are planes* $z = $ ***constant****. (See Fig. 76).*

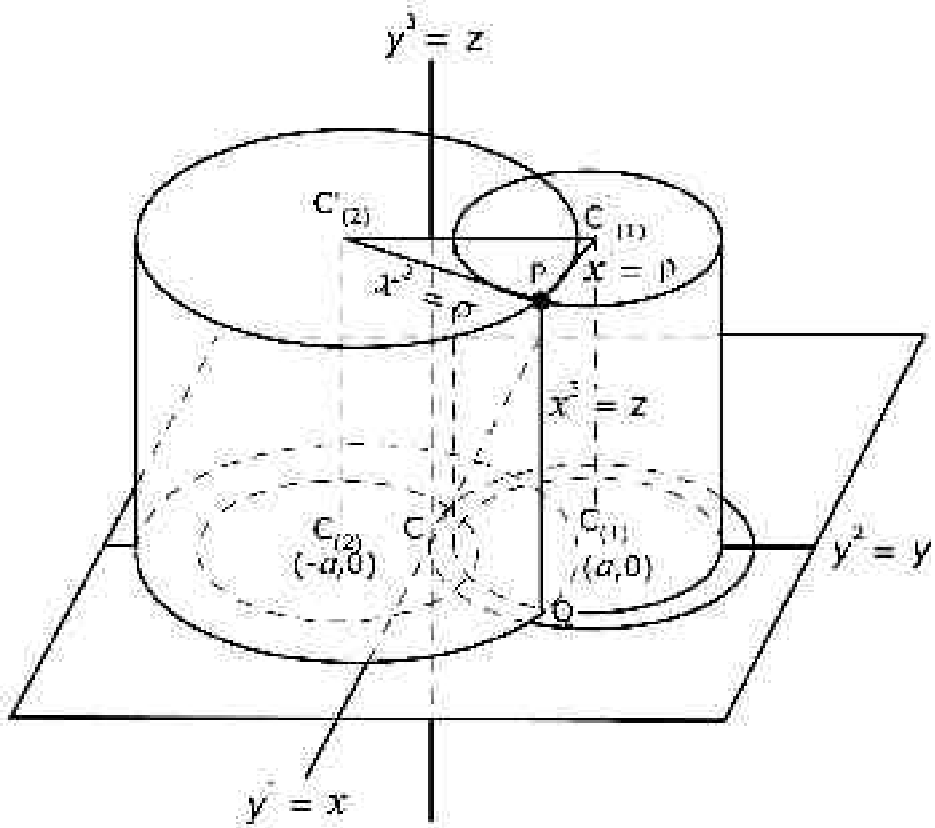

Figure 76

Ex. (1.5) Interpret the **bispherical coordinate system** *for which the transformations are*

$$\bar{x}^1 = \rho = \left[(x)^2 + (y)^2 + (z - a)^2\right]^{1/2} \geq 0 ,$$

$$\bar{x}^2 = \sigma = \left[(x)^2 + (y)^2 + (z + a\,)\right]^{1/2} \geq 0 ,$$

$$0 \leq \bar{x}^3 = \theta = \tan^{-1}\frac{y}{x} < 2\pi .$$

Ans. The $\bar{x}^1$*-surface (*$\rho = $ *constant) is a sphere of radius* ρ *and center at* $x^i = (0, 0, a)$. *The* $\bar{x}^2$*-surface* $\sigma = $ *constant is a sphere of radius* σ *and center at* $x^i = (0, 0, -a)$. *The* $\bar{x}^3$*-surface* $\theta = $ *constant is the plane containing the* z*-axis which makes an angle* θ *with the* xz*-plane.*

Ex. (1.6) From plane elliptical coordinates, determine the transformation from **elliptical cylindrical coordinates** *to Cartesian coordinates and identify the coordinate surfaces.*

Ans. $x^1 = x = a\cosh\xi\cos\eta = a\cosh\bar{x}^1\cos\bar{x}^2 ,$

$x^2 = y = a\sinh\xi\sin\eta = a\sinh\bar{x}^1\sin\bar{x}^2 ,$

$x^3 = z = \bar{x}^3 ,\ 0 \leq \xi ,\ 0 \leq \eta < 2\pi ,\ -\infty < z < +\infty .$

The $\bar{x}^1$*-surfaces are elliptical cylinders with axes along the* $\bar{x}^3$*-axis and foci at* $x^i = (\pm a, 0, 0)$. *The* $\bar{x}^2$*-surfaces are hyperbolic cylinders with axes along the* $\bar{x}^3$*-axis and foci at* $x^i = (\pm a, 0, 0)$. *The* $\bar{x}^3$*-surfaces are planes* $z = $ *constant.*

Ex. (1.7) Determine the transformation between bipolar cylindrical coordinates x^i *and elliptical coordinates* $\bar{x}^i$.

Ans.

$$a\cosh\bar{x}^1 = a\cosh\xi = \rho + \sigma = x^1 + x^2 ,$$

$$a\cos\bar{x}^2 = a\cos\eta = \rho - \sigma = x^1 - x^2 ,\ \bar{x}^3 = z = x^3 .$$

Ex. (1.8) Interpret the **spheroidal coordinate system**, *related to Cartesian coordinates by the transformations*

$$x^1 = x = a\cosh\xi\cos\eta\cos\theta = a\cosh\bar{x}^1\cos\bar{x}^2\cos\bar{x}^3 ,$$

$$x^2 = y\ a\cosh\xi\cos\eta\sin\theta = a\cosh x^1\cos x^2\sin x^3 ,$$

$$x^3 = z = a\sinh\xi\sin\eta = a\sinh x^1\sin x^2 .$$

Ans. From the first two equations we have

$$\bar{x}^3 = \theta = \tan^{-1}\frac{y}{x} = \tan^{-1}\frac{x^2}{x^1} .$$

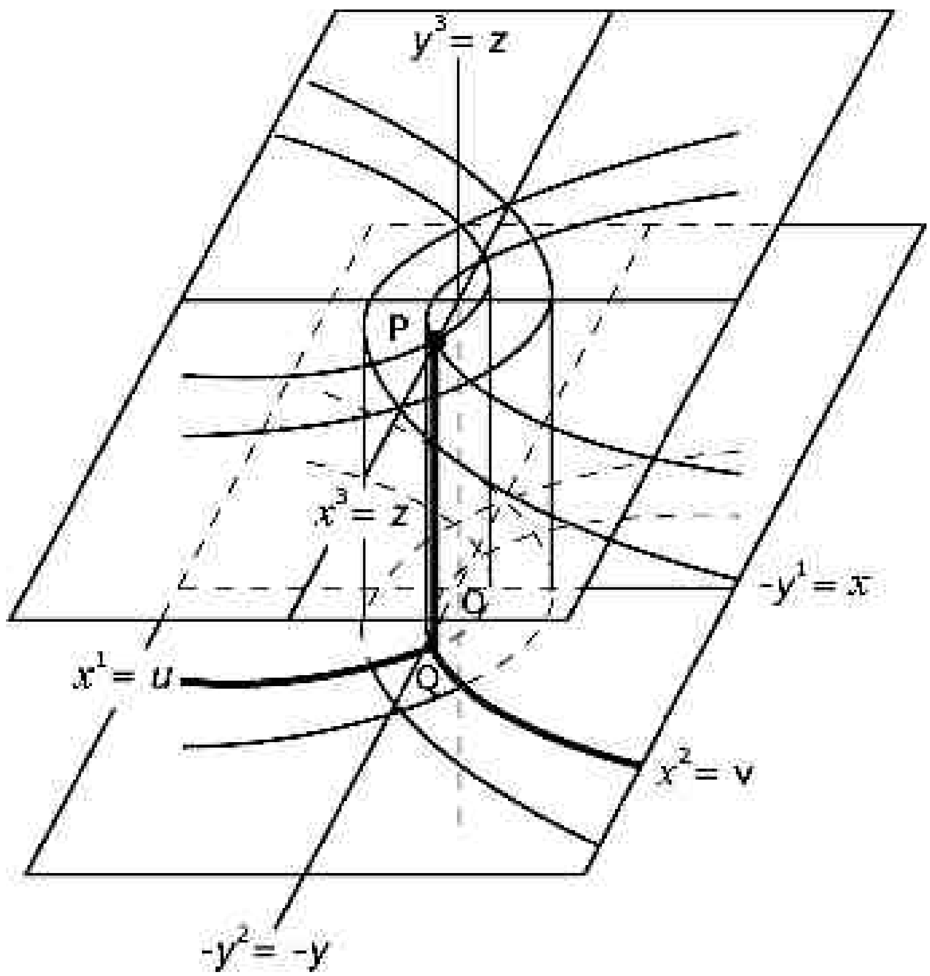

Figure 77

Hence the $\bar{x}^3$ – surfaces are planes through the z -axis making an angle θ with the xz-plane. By squaring all three equations and eliminating θ by adding the first two, we see that

$$\frac{x^2 + y^2}{a \cosh^2 \xi} + \frac{z^2}{a^2 \sinh^2 \xi} = 1 ,$$

$$\frac{x^2 + y^2}{a^2 \cos^2 \eta} - \frac{z^2}{a^2 \sin^2 \eta} = 1 .$$

*The $\bar{x}^1$ – surface (ξ = **constant**) is thus a spheroid symmetric about the z -axis and with major and minor axes $a \cosh \xi$, $a \sinh \xi$. The $\bar{x}^2$ – surface (η = **constant**) is clearly a hyperboloid symmetric about the z -axis with major and minor axes $a \cos \eta$ and $a \sin \eta$.*

Ex. (1.9) Interpret the ***ellipsoidal coordinates*** $\bar{x}^i$, *related to Cartesian coordinates* x^i *by the transformations*

$$x = \left[\frac{(a^2 - \bar{x}^1)(a^2 - \bar{x}^2)(a^2 - \bar{x}^3)}{(a^2 - b^2)(a^2 - c^2)} \right]^{1/2} = \left[\frac{(a^2 - \lambda)(a^2 - \mu)(a^2 - \nu)}{(a^2 - b^2)(a^2 - c^2)} \right]^{1/2},$$

$$y = \left[\frac{(b^2 - \bar{x}^1)(b^2 - \bar{x}^2)(b^2 - \bar{x}^3)}{(b^2 - a^2)(b^2 - c^2)} \right]^{1/2} = \left[\frac{(b^2 - \lambda)(b^2 - \mu)(b^2 - \nu)}{(b^2 - a^2)(b^2 - c^2)} \right]^{1/2},$$

$$z = \left[\frac{(c^2 - \bar{x}^1)(c^2 - \bar{x}^2)(c^2 - \bar{x}^3)}{(c^2 - a^2)(c^2 - b^2)} \right]^{1/2} = \left[\frac{(c^2 - \lambda)(c^2 - \mu)(c^2 - \nu)}{(c^2 - a^2)(c^2 - b^2)} \right]^{1/2},$$

$$\lambda = \bar{x}^1 \le a^2 < \mu = \bar{x}^2 \le b^2 < \nu = \bar{x}^3 \le c^2 .$$

Ans. From the first equation we have

$$\frac{(x)^2}{a^2 - \lambda} = \frac{(a^2 - \mu)(a^2 - \nu)}{(a^2 - b^2)(a^2 - c^2)} .$$

The second and third equations give

$$\frac{(y)^2}{b^2 - \lambda} = \frac{(b^2 - \lambda)(b^2 - \nu)}{(a^2 - c^2)(b^2 - c^2)} ,$$

$$\frac{(z)^2}{c^2 - \lambda} = \frac{(c^2 - \mu)(b^2 - \nu)}{(a^2 - c^2)(b^2 - c^2)} .$$

Adding the three equations gives

$$\frac{x^2}{a^2 - \lambda} = \frac{y^2}{b^2 - \lambda} + \frac{z^2}{c^2 - \lambda}$$

$$= \Big[(a^2 - \mu)(a^2 - \nu)(b^2 - c^2) - (b^2 - \mu)(b^2 - \nu)(a^2 - c^2)$$

$$+ (c^2 - \mu)(c^2 - \nu)(a^2 - b^2) \Big]$$

$$\div (a^2 - b^2)(b^2 - c^2)(a^2 - c^2) .$$

Expanding the numerator of the right hand side, we find that the terms in μ *and* ν *drop out and that the remaining terms are identically equal to the denominator. Hence*

$$\frac{x^2}{a^2 - \lambda} + \frac{y^2}{b^2 - \lambda} + \frac{z^2}{c^2 - \lambda} = 1 .$$

Therefore the $\bar{x}^1$*-surface* (λ = *constant*) *is an ellipsoid with axes* $\sqrt{a^2-\lambda}$, $\sqrt{b^2-\lambda}$ *and* $\sqrt{c^2-\lambda}$. *In similar fashion, we can show that the* $\bar{x}^2$*-surface* (μ = *constant*) *is*

$$\frac{x^2}{a^2-\mu}+\frac{y^2}{b^2-\mu}+\frac{z^2}{c^2-\mu}=1,$$

a hyperboloid of one sheet with axis the z*-axis. Finally, the* $\bar{x}^3$*-surface* (ν = *constant*) *has the equation*

$$\frac{x^2}{a^2-\nu}+\frac{y^2}{b^2-\nu}+\frac{z^2}{c^2-\nu}=1,$$

a hyperboloid of two sheets which open respectively to the positive and negative x*-axes. In general, the* x^i*-surface is the quadric whose equation is (see Fig. 78).*

$$\frac{(x^1)^2}{a^2-\bar{x}^i}+\frac{(x^2)^2}{b^2-\bar{x}^i}+\frac{(x^3)^2}{c^2-\bar{x}^i}=1.$$

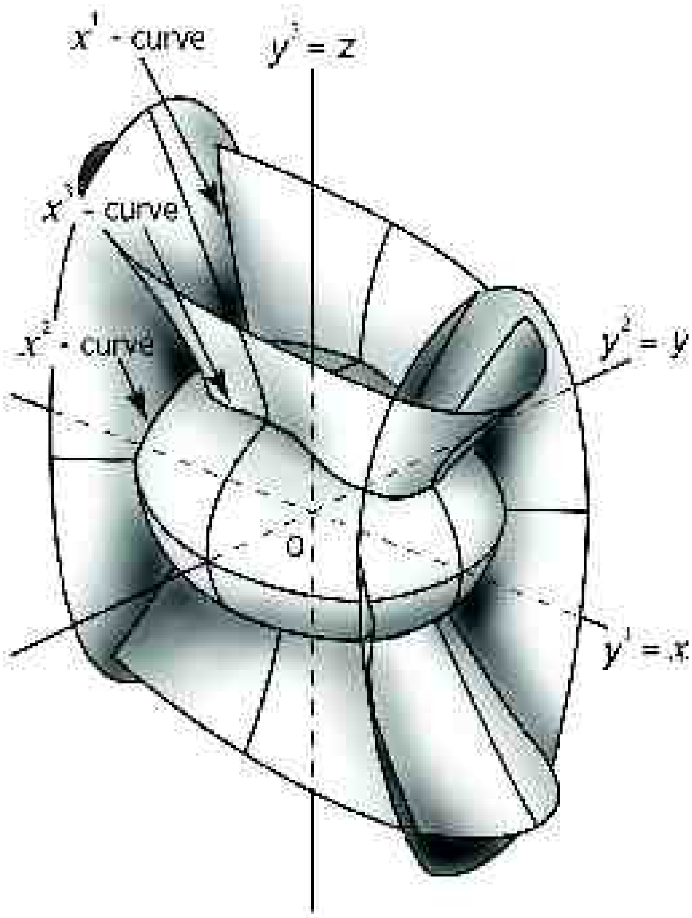

Figure 78

Ex. (1.10) Show that the line element in a cylindrical coordinate system is

$$(ds)^2 = (dr)^2 + (r)^2(d\theta)^2 + (dz)^2 .$$

Ex. (1.11) (a) Show that the line element for the spherical coordinate system of Fig. 71 is

$$(ds)^2 = (dr)^2 + r^2\left[\cos^2\varphi\,(d\theta)^2 + (d\varphi)^2\right].$$

(b) Determine the Christoffel symbols in this coordinate system.

Ans. Those which are not zero are

$$\left\{ {1 \atop 2\ 2} \right\} = -\,r\cos^2\varphi\,, \quad \left\{ {2 \atop 1\ 2} \right\} = \frac{1}{r} = \left\{ {2 \atop 2\ 1} \right\}, \quad \left\{ {2 \atop 2\ 3} \right\} = -\,\tan\varphi = \left\{ {2 \atop 3\ 2} \right\},$$

$$\left\{ {3 \atop 1\ 3} \right\} = \frac{1}{r} = \left\{ {3 \atop 3\ 1} \right\}, \quad \left\{ {3 \atop 2\ 2} \right\} = \sin\varphi\cos\varphi\,, \quad \left\{ {1 \atop 3\ 3} \right\} = -\,r\,.$$

Ex. (1.12) (a) Show that the line element for the hyperbolic cylindrical coordinates is

$$(ds)^2 = (u^2 + \mathrm{v}^2)\left[(du)^2 + (d\mathrm{v})^2\right] + (dz)^2$$

$$= \left[(x^1)^2 + (x^2)^2\right]\left[(dx^1)^2 + (dx^2)^2\right] + (dx^3)^2 .$$

(b) Determine the Christoffel symbols in this system.

Ans.

$$\left\{ {1 \atop 1\ 1} \right\} = \frac{u}{u^2 + \mathrm{v}^2}, \quad \left\{ {1 \atop 1\ 2} \right\} = \frac{\mathrm{v}}{u^2 + \mathrm{v}^2} = \left\{ {1 \atop 2\ 1} \right\},$$

$$\left\{ {1 \atop 2\ 2} \right\} = -\,\frac{u}{u^2 + \mathrm{v}^2}, \quad \left\{ {2 \atop 1\ 1} \right\} = -\,\frac{\mathrm{v}}{u^2 + \mathrm{v}^2},$$

$$\left\{ {2 \atop 1\ 2} \right\} = \frac{u}{u^2 + \mathrm{v}^2} = \left\{ {2 \atop 2\ 1} \right\}, \quad \left\{ {2 \atop 2\ 2} \right\} = \frac{\mathrm{v}}{u^2 + \mathrm{v}^2} .$$

Ex. (1.13) Show that the fundamental tensor in paraboloidal coordinates is

$$g_{ij} = \begin{Vmatrix} u^2 + \mathrm{v}^2 & 0 & 0 \\ 0 & u^2 + \mathrm{v}^2 & 0 \\ 0 & 0 & u^2 + \mathrm{v}^2 \end{Vmatrix}.$$

Ex. (1.14) Show that the fundamental tensor in elliptical cylindrical coordinates is

$$g_{ij} = \begin{Vmatrix} a^2[\sinh^2 u + \sin^2 v] & 0 & 0 \\ 0 & a^2[\sinh^2 u + \sin^2 v] & 0 \\ 0 & 0 & 1 \end{Vmatrix}.$$

Ex. (1.15) Show that the fundamental tensor in spheroidal coordinates is

$$g_{ij} = \begin{Vmatrix} a^2[\sinh^2\xi + \sin^2\eta] & 0 & 0 \\ 0 & a^2[\sinh^2\xi + \sin^2\eta] & 0 \\ 0 & 0 & a^2\cosh^2\xi\cos^2\eta \end{Vmatrix}.$$

Ex. (1.16) Show that in ellipsoidal coordinates the line element is

$$4(ds)^2 = \left[\frac{(\mu-\lambda)(\nu-\lambda)}{(a^2-\lambda)(b^2-\lambda)(c^2-\lambda)}\right](d\lambda)^2$$

$$+\left[\frac{(\nu-\lambda)(\lambda-\mu)}{(a^2-\mu)(b^2-\mu)(c^2-\mu)}\right](d\mu)^2$$

$$+\left[\frac{(\lambda-\nu)(\mu-\nu)}{(a^2-\nu)(b^2-\nu(c^2-\nu)}\right](d\nu)^2.$$

2. Space Curves

Consider any continuous curve C with a continuous tangent. Let x^m and $x^m + dx^m$ be the coordinates of nearby points on the curve. The distance along the curve between the points is the differential of arc length

$$ds = |dx^m| = \sqrt{g_{mn}\,dx^m\,dx^n}.$$

Therefore the vector

$$\lambda^m = \frac{dx^m}{|dx^n|} = \frac{dx^m}{ds} \tag{2.1}$$

is a unit vector along the curve at the point x^m. It is the **unit tangent** to the curve at that point.

Since λ^m is a unit vector, $\quad g_{mn}\lambda^m\lambda^n = 1$.

Taking the intrinsic derivative of both sides with respect to the arc length gives

$$g_{mn}\lambda^m\frac{\delta\lambda^n}{\delta s} + g_{mn}\frac{\delta\lambda^m}{\delta s}\lambda^n = 2g_{mn}\lambda^m\frac{\delta\lambda^n}{\delta s} = 0.$$

If we define κ to be the magnitude of $\delta\lambda^n / \delta s$ and μ^n to be a unit vector in the direction of $\delta\lambda^n / \delta s$, i.e. , if

$$\frac{\delta\lambda^n}{\delta s} = \kappa\,\mu^n\,,\ |\mu^n| = 1\,, \tag{2.2}$$

then it is clear from the preceding equation that

$$g_{mn}\,\lambda^m\,\mu^n = 0\,.$$

Evidently μ^m is orthogonal to λ^m. We therefore call μ^m the **principal normal** to the curve C. The invariant κ is the **curvature** of C. Thus far our results do not differ from those obtained in two dimensions.

Let us now form the intrinsic derivative of the preceding equation. It is

$$g_{mn}\frac{\delta\lambda^m}{\delta s}\mu^n + g_{mn}\,\lambda^m\frac{\delta\mu^n}{\delta s} = 0\,.$$

Using equation (2.2), this becomes

$$g_{mn}\,\lambda^m\frac{\delta\mu^n}{\delta s} = -\,\kappa = -\kappa\,g_{mn}\,\lambda^m\,\lambda^n\,.$$

Hence

$$g_{mn}\,\lambda^m\left[\frac{\delta\mu^n}{\delta s} + \kappa\,\lambda^n\right] = 0\,. \tag{2.3}$$

Therefore λ^m is orthogonal to the unit vector ν^m defined by

$$\frac{\delta\mu^n}{\delta s} + \kappa\,\lambda^n = \tau\,\nu^n\,,\ \left|\frac{\delta\mu^n}{\delta s} + \kappa\,\lambda^n\right| = \tau\,. \tag{2.4}$$

On the other hand, since

$$g_{mn}\,\mu^m\,\mu^n = 1\,,$$

it follows that

$$g_{mn}\,\mu^m\frac{\delta\mu^n}{\delta s} = 0\,.$$

Therefore μ^m is orthogonal to $\delta\mu^m / \delta s$ as well as to λ^m. Hence it is also orthogonal to ν^m. The vectors λ^n, μ^n, and ν^n thus form a mutually orthogonal triad. The sense of ν^n is so chosen that this triad is right-handed in the order λ^n, μ^n, ν^n; this requires that

$$\varepsilon_{ijk}\,\lambda^i\,\mu^j\,\nu^k = 1\,. \tag{2.5}$$

The sign of τ is chosen in equation (2.4) to make this so. The quantity τ is called the **torsion** of the curve, the vector ν^n is the curve's **binormal**.

Equations (2.2) and (2.4) express the intrinsic derivatives of two of the triad λ^n, μ^n, ν^n in terms of the others. This prompts the thought that perhaps it should be possible to express the third in a similar manner. It is easy to show that this is the case.

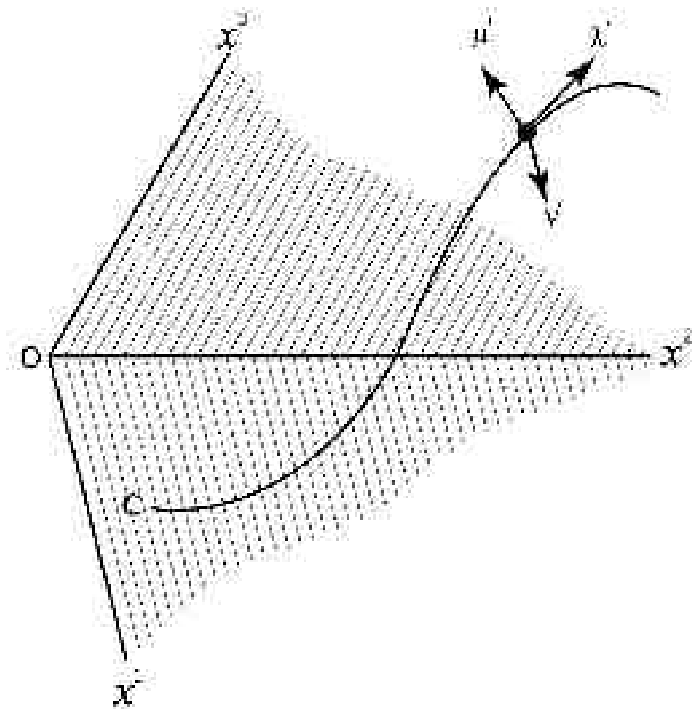

Figure 79

First, since λ^n, μ^n, ν^n forms an orthogonal triad of unit vectors, we may adopt it as a basis system. Then any other vector may be expressed as a linear combination of them; in particular, we may take

$$\frac{\delta\mu^n}{\delta s} = \alpha\,\lambda^n + \beta\,\mu^n + \gamma\,\nu^n,$$

where α, β and γ are the invariants

$$\alpha = g_{mn}\lambda^m\frac{\delta\nu^n}{\delta s},\ \beta = g_{mn}\mu^n\frac{\delta\nu^n}{\delta s},\ \gamma = \nu^m\frac{\delta\nu^n}{\delta s}.$$

The value $\gamma = 0$ follows at once from the fact that ν^n is a unit vector. To find the value of α, we use the fact that

$$g_{mn}\lambda^m\nu^n = 0,\ g_{mn}\left[\frac{\delta\lambda^m}{\delta s}\nu^n + \lambda^m\frac{\delta\nu^n}{\delta s}\right] = 0 = g_{mn}\left[\kappa\,\mu^m\nu^n + \lambda^m\frac{\delta\nu^n}{\delta s}\right],$$

implying that

$$\alpha = g_{mn}\lambda^m\frac{\delta\nu^n}{\delta s} = -\kappa\, g_{mn}\mu^m\nu^n = 0$$

because of the orthogonality of μ^n and ν^n.

In similar fashion, we find β by starting from the relation

$$g_{mn}\,\mu^m\,\nu^n \;=\; 0\,,$$

$$g_{mn}\left[\frac{\delta\mu^m}{\delta s}\,\nu^n + \mu^m\,\frac{\delta\nu^n}{\delta s}\right] = g_{mn}\left[(\tau\,\nu^m - \kappa\,\lambda^m)\,\nu^n + \mu^m\,\frac{\delta\nu^n}{\delta s}\right] = 0\,,$$

implying that

$$\beta = g_{mn}\,\mu^m\,\frac{\delta\nu^n}{\delta s} = -\,g_{mn}(\tau\,\nu^m - \kappa\,\lambda^m)\,\nu^n = -\,\tau$$

because ν^n is a unit vector and is orthogonal to λ^n.

If we use the values of α and β thus found and express the result along with equations (2.2) and (2.4) in symmetric fashion, we have the **Frenet formulae**

$$(2.6)\qquad \left\{\begin{aligned} \frac{\delta\lambda^n}{\delta s} &= \kappa\,\mu^n\,,\\ \frac{\delta\mu^n}{\delta s} &= \tau\,\nu^n - \kappa\,\lambda^n\,,\\ \frac{\delta\nu^n}{\delta s} &= -\,\tau\,\mu^n\,, \end{aligned}\right.$$

for a curve in space. In descriptive terms, the quantity κ measures "how curved" the curve is and the quantity τ measures "how twisted" the curve is.

Ex. (2.1) Find the tangent, principal normal, binormal, curvature and torsion to the regular circular helix whose parametric equations are

$$\rho = a\,,\;\; \theta = \left(\frac{\cos\alpha}{a}\right)s\,,\;\; a = constant > 0\,,\;\; \alpha \neq \frac{n\,\pi}{2}\,,$$

in cylindrical coordinates.

Ans. We have first that the unit tangent vector is

$$\lambda^i = \frac{dx^{\,i}}{ds} = \left(0\,,\frac{\cos\alpha}{a}\,,\sin\alpha\right).$$

Then, since the only Christoffel symbols which do not vanish are

$$\left\{ {1 \atop 2\;2} \right\} = -a\,,\;\; \left\{ {1 \atop 1\;2} \right\} = \frac{1}{a} = \left\{ {2 \atop 2\;1} \right\},$$

we have

$$\frac{\delta\lambda^i}{\delta s} = \left(-\,\frac{\cos^2\alpha}{a}\,,0\,,0\right) = \kappa\,\mu^i\,,$$

whence

$$\kappa = \frac{\cos^2\alpha}{a}, \ \mu^i = (-1, 0, 0).$$

Similarly,

$$\frac{\delta\mu^i}{\delta s} + \kappa\lambda^i = \frac{\sin\alpha\cos\alpha}{a}\left(0, -\frac{\sin\alpha}{a}, \cos\alpha\right) = \tau\nu^i,$$

whence

$$\tau = \frac{\sin\alpha\cos\alpha}{a}, \ \nu^i = \left(0, -\frac{\sin\alpha}{a}, \cos\alpha\right).$$

The angle α *may be seen to be the angle between the tangent vector and the plane perpendicular to the* z*-axis. (See Fig. 80).*

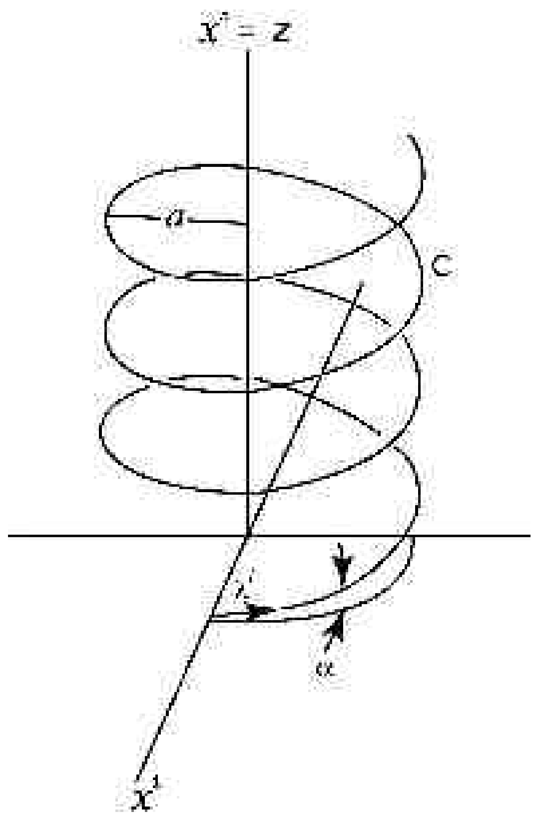

Figure 80

Ex. (2.2) (a) Find the unit tangent, curvature, and principal normal to the conical spiral (Fig. 81) whose equations are

$$\rho = a\theta, \ z = b\theta$$

in cylindrical coordinates. (b) Find the explicit expression for arc length s *in terms of angle* θ *.*

$$\textit{Ans.}\ (a)\quad \lambda^i = \frac{1}{\sqrt{c^2 + a^2\theta^2}}(a, 1, b),\ \textit{where}\ \ c^2 = a^2 + b^2;$$

$$\kappa = \frac{a^2}{(c^2 + a^2\theta^2)^2}[5c^2 - 2a^2c^2\theta + (3a^2 + 11c^2)a^2\theta^2 + 5a^4\theta^4];$$

$$\mu^i = \frac{(c^2 + a^2\theta^2)}{\kappa}(-a[a^2\theta + (c^2 + a^2\theta^2)], 2(c^2 + a^2\theta^2) - a^2\theta, -a^2b\theta).$$

$$(b)\quad s = \frac{\theta}{2}\sqrt{c^2 + a^2\theta^2} + \frac{c^2}{2a}\Big[-\log c + \log\Big(a\theta + \sqrt{c^2 + a^2\theta^2}\Big)\Big].$$

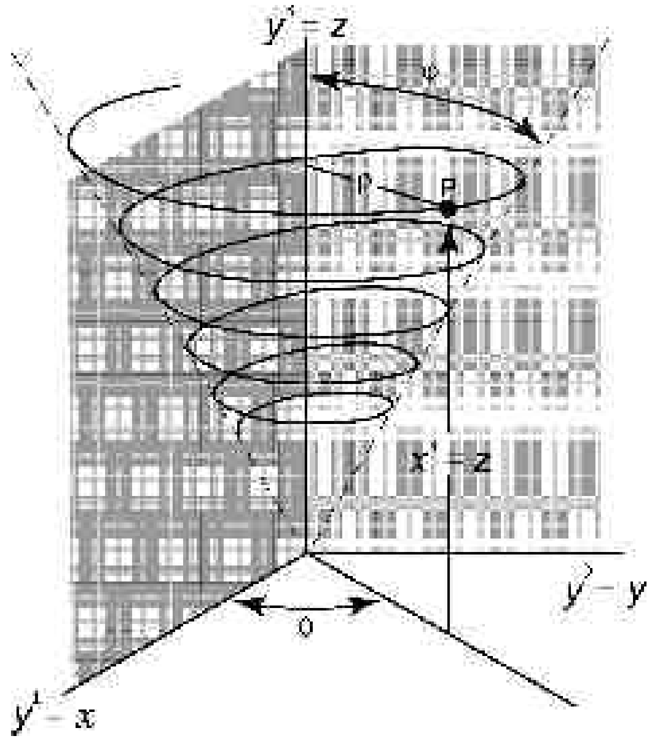

Figure 81

The foregoing algebraic derivation of the Frenet formulae and of the quantities associated with a curve — tangent, normal, binormal, curvature and torsion — may be buttressed by a parallel geometric discussion. To begin with, we define the unit tangent to a smooth curve, whether a plane curve or not, as the unit vector having the limiting position of the secant $\overline{\mathrm{PQ}}$ at a point P when Q approaches P. We may in this way define the tangent at all regular points of a curve.

Let us now consider the tangents at neighboring points P and Q. Let the tangent $\lambda^i(\mathrm{Q})$ define a parallel vector field, i. e., one satisfying

$$\frac{\delta\lambda^i(\mathrm{Q})}{\delta s} = 0\,.$$

The element of this parallel vector field at P may be used to determine the difference between $\lambda^i(\mathrm{P})$ and $\lambda^i(\mathrm{Q})$ at P. In the limit as Q approaches P, the direction and magnitude of $\delta\lambda^i/\delta s$ as $\mathrm{Q}\to\mathrm{P}$ is defined. Since $\lambda^i(\mathrm{Q})$ is a parallel vector field, and since $\delta\lambda^i/\delta s$ as $\mathrm{Q}\to\mathrm{P}$ is the limit of the ratio of the *difference* of $\lambda^i(\mathrm{P})$ and $\lambda^i(\mathrm{Q})$, it can have no part parallel to λ^i, must therefore lie in the plane normal to λ^i. This is the **normal plane**. The normal μ^i in the direction of $\delta\lambda^i/\delta s$ lies in the normal plane. Perpendicular to μ^i, through P, is the **rectifying plane**.

We may apply the procedure once more. Consider the normal $\mu^i(\mathrm{Q})$ at Q. Let it define a parallel vector field along the curve. The element of this vector field at P may be subtracted from $\mu^i(\mathrm{P})$ and the ratio

$$\frac{\mu^i(\mathrm{P}) - \mu^i(\mathrm{Q})}{s(\mathrm{P}) - s(\mathrm{Q})}$$

formed. Its limit as Q approaches P is $\delta\mu^i/\delta s$. It must be perpendicular to μ^i and, since μ^i is everywhere perpendicular to λ^i, it must be perpendicular to λ^i also. In this way, we generate the triad λ^i, μ^i, ν^i. Clearly, ν^i must lie in the rectifying plane. It is orthogonal to the plane of λ^i and μ^i, the **osculating plane**.

Ex. (2.3) Find (a) the normal plane, (b) the rectifying plane, and (c) the osculating plane at any point of the circular helix of Ex. (2.1).

Ans. (a) Since the normal plane is perpendicular to the tangent, it must be spanned by the unit vectors μ^i and ν^i. Hence the position vector of any point in the normal plane is given by

$$p^i = p^i_{(0)} + \gamma\,\mu^i + \beta\,\nu^i = \gamma(-1,0,0) + \beta\left(0, -\frac{\sin\alpha}{a}, \cos\alpha\right)$$
$$= \left(-\gamma, -\frac{\beta\sin\alpha}{a}, \beta\cos\alpha\right) + p^i_{(0)},$$

where γ and β are arbitrary scalar parameters and $p^i_{(0)}$ is a point on the curve. (b) In similar fashion, the rectifying plane is spanned by λ^i and μ^i so that the position vector q^i to any point in the rectifying plane is

$$q^i - q^i_{(0)} = \gamma\,\lambda^i + \beta\,\nu^i = \gamma\left(0, \frac{\cos\alpha}{a}, \sin\alpha\right) + \beta\left(0, -\frac{\sin\alpha}{a}, \cos\alpha\right)$$
$$= \left(0, \frac{\gamma\cos\alpha - \beta\sin\alpha}{a}, \gamma\sin\alpha + \beta\cos\alpha\right).$$

(c) The osculating plane, spanned by λ^i and μ^i, has at any point the position vector of the form

$$s - s^i_{(0)} = \gamma\,\lambda^i + \beta\,\mu^i = \gamma\left(0, \frac{\cos\alpha}{a}, \sin\alpha\right) + \beta(-1,0,0)$$
$$= \left(-\beta, \frac{\gamma\cos\alpha}{a}, \gamma\sin\alpha\right).$$

From the Frenet formulae we can deduce very simply certain results of considerable interest. Thus, consider a curve for which $\kappa = 0$, i.e., a curve of zero curvature. Along such a curve

$$\frac{\delta\lambda^n}{\delta s} = \frac{\delta}{\delta s}\left(\frac{dx^n}{ds}\right) = \frac{d^2x^n}{ds^2} + \left\{\begin{matrix} n \\ p\ q \end{matrix}\right\}\frac{dx^p}{ds}\frac{dx^q}{ds} = 0 .$$

This is identical with the equation for a geodesic in two dimensions. We may therefore *define* a geodesic as a curve whose curvature is zero. (As in two dimensions, we might also define a geodesic as a curve of minimum length between two given points. The development is identical and the results are the same.)

From the Frenet formulae, we see further that when the curve is a geodesic, i. e. , when $\kappa = 0$, the two remaining equations become

(2.7) $$\frac{\delta\mu^n}{\delta s} = \tau\nu^n , \quad \frac{\delta\nu^n}{\delta s} = -\tau\mu^n .$$

It appears that μ^i and ν^n are now wholly independent of λ^n. This is not true, however, for

$$\frac{\delta\mu^n}{\delta s} = \frac{d\mu^n}{ds} + \left\{\begin{matrix} n \\ p\ q \end{matrix}\right\}\mu^p\frac{dx^q}{ds} = \frac{d\mu^n}{ds} + \left\{\begin{matrix} n \\ p\ q \end{matrix}\right\}\mu^p\lambda^q ,$$

with a similar expression for $\delta\nu^n/\delta s$; the explicit dependence upon λ^q is here evident. However, the vectors μ^i and ν^n no longer have any necessary relation to λ^n other than orthogonality; in other words, at any point of the curve, we may choose μ^n to be any unit vector orthogonal to the unit tangent. At the same time, equations (2.7) will be satisfied for any τ if only μ^i and ν^n are mutually orthogonal. We may show this by multiplying the first by ν_n, the second by μ_n, and adding; the result is

$$\nu_n\frac{\delta\mu^n}{\delta s} + \mu^n\frac{\delta\nu^n}{\delta s} = \tau(\nu^n\nu_n) - \tau(\mu^n\mu_n) = 0 .$$

Hence

$$\nu_n\frac{\delta\mu^n}{\delta s} + \mu_n\frac{\delta\nu^n}{\delta s} = \frac{\delta}{\delta s}\left(g_{mn}\mu^m\nu^n\right) = \frac{d}{ds}\left(g_{mn}\mu^m\nu^n\right) = 0 ,$$

or $g_{mn}\mu^m\nu^n = \cos\theta = \text{constant}$ along the geodesic, where θ is the angle between μ^i and ν^n. In particular, if the angle θ is $90°$, the vectors μ^i and ν^n are orthogonal to each other as well as to λ^n. We take this to be the case in order to satisfy equation (2.5).

Let us now look upon equation (2.7) as defining τ. We choose μ^i and ν^n as any mutually orthogonal unit vectors which are both orthogonal to λ^n. The components of μ^i and ν^n may be any functions of s consistent with these conditions. Then from either of equations (2.7) we may define a τ which will satisfy the other equation. The simplest possible choice of τ is, of course, $\tau = 0$. In that case, the curve has no torsion and both normal and binormal are parallel vector fields along the geodesic. It is apparent, however, that a curve of zero curvature may have whatever torsion one chooses to give it; in this sense, the torsion of a geodesic is indefinite.

It is of further interest to note the special case of a curve for which $\kappa \neq 0$, $\tau = 0$. In this event, the binormal ν^n is a parallel vector field along the curve though the curve is in general not a geodesic. The Frenet formulae then reduce to*

$$\frac{\delta \lambda^n}{\delta s} = \kappa \mu^i, \quad \frac{\delta \mu^i}{\delta s} = -\kappa \lambda^n .$$

It is clear that we can apply the argument previously given for μ^i and ν^n to show that these equations are satisfied whenever μ^n and λ^n are orthogonal, as they must be by the definition of μ^n. The curve is then a plane curve.

Ex. (2.4) In three-dimensional Euclidean space, geodesics are straight lines. Without loss of generality, therefore, one may choose the z-axis of a Cartesian coordinate system to represent a geodesic. Along it, $ds = dz$, $s = z - z_0$, and the unit tangent is $\lambda^i = (0, 0, 1)$. Since $\kappa = 0$, we may choose the principal normal μ^i arbitrarily, then require only that the binormal ν^i be orthogonal to both λ^i and μ^i. (a) If μ^i is chosen to be

$$\mu^i = (\cos ks, \sin ks, 0)$$

(a corkscrew field), what is the binormal ν^i? (b) What is the torsion τ? (c) What form have μ^i and ν^i when we impose the condition $\tau = 0$?

Ans. (a) $\nu^i = (-\sin ks, \cos ks, 0)$. (b) $\tau = k$.

(c) $\mu^i = (1, 0, 0)$, $\nu^i = (0, 1, 0)$.

Ex. (2.5) In a spherical coordinate system (see Ex. 1.11)), the parametric representation of a curve is

$$\log \frac{r}{a} = \tan^{-1} u, \quad \theta = \tan^{-1}(\gamma u), \quad \varphi = \tan^{-1}\left(\left[\frac{1 - \gamma^2}{1 + \gamma^2 u^2}\right]^{1/2} u\right).$$

$$(a = constant > 0)$$

in terms of a parameter u and constant γ $(|\gamma| \leq 1)$. Determine (a) the unit tangent, (b) the principal normal, (c) the curvature, (d) the torsion, (e) the binormal. (f) Is the curve a plane curve? (g) Identify the curve.

Ans. (a) $\lambda^i = \dfrac{\sqrt{2}}{2}\left(1, \dfrac{\gamma(1 + u^2)}{r(1 + \gamma^2 u^2)}, \dfrac{1}{r}\left[\dfrac{1 - \gamma^2}{1 + \gamma^2 u^2}\right]^{1/2}\right)$.

(b) $\lambda^i = \dfrac{\sqrt{2}}{2}\left(-1, \dfrac{\gamma(1 + u^2)}{r(1 + \gamma^2 u^2)}, \dfrac{1}{r}\left[\dfrac{1 - \gamma^2}{1 + \gamma^2 u^2}\right]^{1/2}\right)$.

*It would appear at first glance that when $\tau = 0$ the Frenet formulae reduce to those for a *surface*, not necessarily a plane. It must be remembered, however, that the intrinsic derivatives have been formed by employing Christoffel symbols for *three* dimensions, not two. This distinction is essential.

$(c)\ \kappa = \dfrac{\sqrt{2}}{2}$. $\quad (d)\ \tau = 0$.

$(e)\ \nu^i = \varepsilon^{ijk}\lambda_j\mu_k = \left(0\,,\, \dfrac{\sqrt{(1-\gamma^2)(1-u^2)}}{r(1+\gamma^2 u^2)}\,,\, \dfrac{\gamma}{r}\left[\dfrac{1+u^2}{1+\gamma^2 u^2}\right]^{1/2}\right)$.

(f) Yes. (g) The curve is an exponential spiral lying in the plane tilted at an angle $\varepsilon = \cos^{-1}\gamma$ to the xy-plane and passing through the x-axis.

Let P be some point of space, located from the origin O by the position vector p^i. Through P let C be a curve which is continuous and whose equations

$$x^i = x^i(s)$$

are differentiable with respect to s to all required orders. Then the tangent to C is the vector

$$\lambda^i = \frac{\delta p^i}{\delta s}.$$

The principal normal is $\mu^i = \dfrac{1}{\kappa}\dfrac{\delta\lambda^i}{\delta s}$, where κ is the curvature of C. The unit vectors λ^i and μ^i define and span what is known as the **osculating plane** to C at P (see Fig. 82).

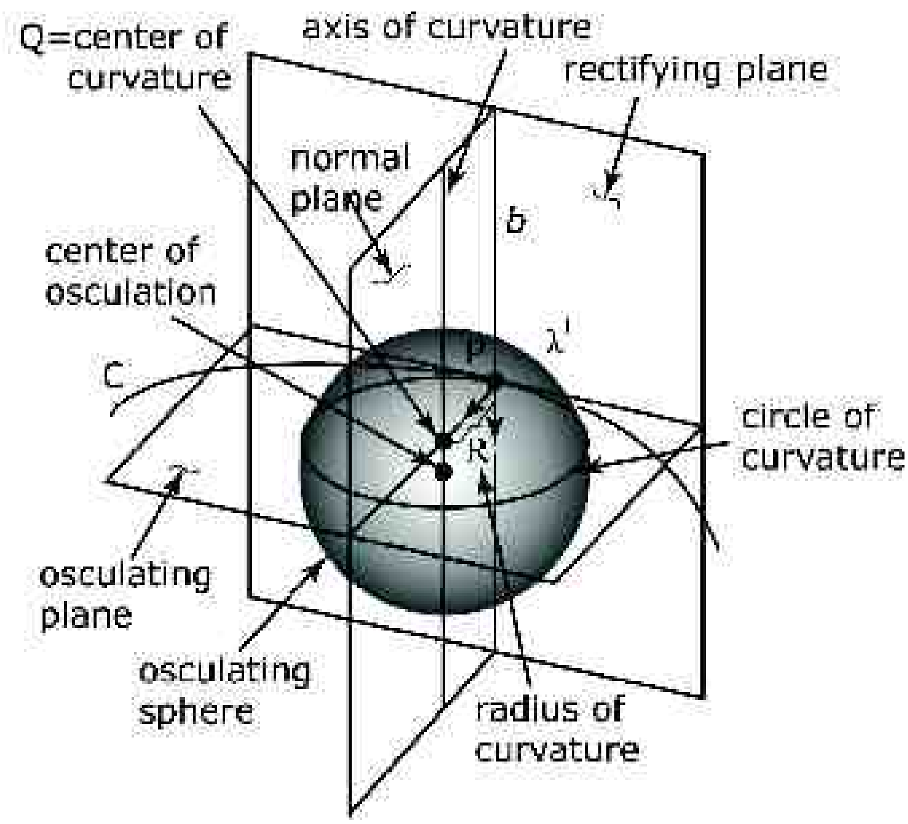

Figure 82

Consider the straight line L through P in the direction of μ^i. Its equation will be

$$c^i = p^i + b\,\mu^i,$$

where b is the parameter of the line. Let us pick a point such as Q on L, and draw in the osculating plane about Q a circle whose radius is b. This circle will obviously pass through P, as the manner of construction requires it should. Therefore the tangent to the circle is also a tangent to the curve C. In general, as P moves along the curve E, the point Q would also be displaced in space. We may ask, therefore, if there is some point Q which is stationary with respect to the displacement of P along C. In this case,

$$|c^i - p^i|^2 = (b)^2 = (c_i - p_i)(c^i - p^i)$$

would be stationary, having a contact of the third order. Differentiating the previous expression with respect to s gives first

$$-(c_i - p_i)\frac{\delta p^i}{\delta s} = 0 = (p_i - c_i)\lambda^i = -b\,\mu_i\lambda^i,$$

where we have set $\delta c_i / \delta s = 0$. This simply expresses the fact that λ^i and μ^i are orthogonal. In order that the circle with center at Q have a contact of the third order at P, we must differentiate once more the expression for b^2, giving

$$\frac{1}{2}\frac{d^2(b)^2}{ds^2} = \frac{\delta p_i}{\delta s}\lambda^i + (p_i - c_i)\frac{\delta\lambda^i}{\delta s} = 0,$$

$$\lambda_i\lambda^i - b\,\mu_i(\kappa\,\mu^i) = 0 = 1 - \kappa b,\ b = \frac{1}{\kappa}.$$

Thus the point Q at distance $\mathrm{R} = \dfrac{1}{\kappa}$ from P in the direction of μ^i is called the **center of curvature** of the curve C at point P. The quantity R is the **radius of curvature** of the curve at P.

Let us consider a further generalization. Suppose that c^i is

$$c^i = p^i + \frac{1}{\kappa}\mu^i + c\,\nu^i,\ |c^i - p^i|^2 = \frac{1}{\kappa^2} + (c)^2 = \mathrm{R}^2,$$

where c is an undetermined parameter. If taken as a curve parameter, c will evidently generate the straight line through Q $(c = 0)$ in the direction of ν^i — i. e. , perpendicular to the osculating plane. This straight line is called the **axis of curvature** of the curve C at P.

Now $\mathrm{R}^i = c^i - p^i = \dfrac{\mu^i}{\kappa} + c\,\nu^i$

is the radius of a circle through P with center on the axis of curvature. As P advances along the curve C,

$$\frac{\delta \mathrm{R}^i}{\delta s} = \lambda^i,\ \frac{\delta^2\mathrm{R}^i}{\delta s^2} = \kappa\,\mu^i,\ \frac{\delta^3\mathrm{R}^i}{\delta s^3} = \kappa'\mu^i + \kappa(\tau\,\nu^i - \kappa\,\lambda^i).$$

Hence

$$\frac{\delta (R)^2}{\delta s} = \frac{\delta}{\delta s}(R^i R_i) = 2 R_i \frac{\delta R^i}{\delta s} = \lambda_i (- \frac{\mu^i}{\kappa} + c \nu^i) = 0 ,$$

$$\frac{\delta^2 (R)^2}{\delta s^2} = 2\left(\frac{\delta R_i}{\delta s}\frac{\delta R^i}{\delta s} + R_i \frac{\delta^2 R^i}{\delta s^2} \right) = 2\left[\lambda_i \lambda^i + \left(- \frac{\mu}{\kappa} + c \nu^i \right) \kappa \mu^i \right] = 0 ,$$

$$\frac{\delta^3 (R)^2}{\delta s^3} = 2\left[3 \frac{\delta R_i}{\delta s}\frac{\delta^2 R^i}{\delta^2 s} + R_i \frac{\delta^3 R^i}{\delta s^3} \right]$$

$$2\left[3 \lambda_i \cdot \kappa \mu^i + \left(\frac{\mu^i}{\kappa} + c \nu^i \right) \left(\kappa' \mu_i + \kappa [\tau \nu_i - \kappa \lambda_i] \right) \right] = 2\left[\frac{1}{\kappa}\frac{d\kappa}{ds} + c \kappa \tau \right].$$

For a contact of the fourth order, the last expression must vanish. This requires that c have the value

$$c = - \frac{1}{\kappa^2 \tau}\frac{d\kappa}{ds} . \tag{2.8}$$

Hence the point

$$c^i = p^i + \frac{\mu^i}{\kappa} - \left(\frac{1}{\kappa^2 \tau}\frac{d\kappa}{ds} \right) \nu^i$$

is the center of the **osculating sphere**, having a contact of the fourth order with the curve C at P. It is the sphere which best fits the curve at P. Its intersection with the osculating plane is the **osculating circle**. Since $R = \dfrac{1}{\kappa}$,

$$\frac{dR}{ds} = - \frac{1}{\kappa^2}\frac{d\kappa}{ds} ,$$

$$c^i = p^i + R \mu^i + \frac{1}{\tau}\frac{dR}{ds} \nu^i , \tag{2.9}$$

$$| c^i - p^i | = \left[R^2 + \left(\frac{1}{\tau}\frac{dR}{ds} \right)^2 \right]^{1/2} . \tag{2.10}$$

In summary, we may say that the foregoing development shows how any curve C may be characterized at any point by its tangent λ^i, normal μ^i, binormal ν^i, and two invariants — the curvature κ and torsion τ. These quantities are intimately related by the Frenet formulae. It is reasonable to suppose, conversely, that *any curve may be characterized in its essentials by its curvature and torsion together with the orientation of the triad* λ^i, μ^i *and* ν^i *at some point*. This proposition may, in fact, be given a rigorous proof.

Ex. (2.6) A ***loxodrome*** *is a curve on the surface of a sphere which at every point makes a fixed angle* ε *with the meridian of longitude through the point. Thus its equation is*

$$r = a\,,\ \theta = (\tan\varepsilon)\ln(\sec\varphi + \tan\varphi)$$

in spherical coordinates. (Note that when $\varepsilon = \pm 90°$*, we require that* $\varphi = 0$*. Find (a) the unit tangent; (b) the normal; (c) the binormal; (d) the curvature; (e) the torsion. (Hint: use Ex. (1.11).)*

Ans. (*a*) $\lambda^i = \left(0\,,\dfrac{\sin\varepsilon\sec\varphi}{a}\,,\dfrac{\cos\varepsilon}{a}\right).$

$$(b)\quad \mu^i = \frac{\left(-1\,,-\dfrac{\sin\varepsilon\cos\varepsilon\tan\varphi\sec\varphi}{a}\,,\dfrac{\sin^2\varepsilon\tan\varphi}{a}\right)}{\left[1 + \sin^2\varepsilon\tan^2\varphi\right]^{1/2}}.$$

$$(c)\quad \nu^i = \frac{\left(\sin\varepsilon\tan\varphi\,,-\dfrac{\cos\varepsilon\sec\varphi}{a}\,,\dfrac{\sin\varepsilon}{a}\right)}{\left[1 + \sin^2\varepsilon\tan^2\right]^{1/2}}.$$

$$(d)\quad \kappa = \frac{1}{\mathrm{R}} = \frac{1}{a}\left[1 + \sin^2\varepsilon\tan^2\varphi\right]^{1/2}.$$

$$(e)\quad \tau = \frac{\sin\varepsilon\cos\varepsilon\sec^2\varphi}{a\left(1 + \sin^2\varepsilon\tan^2\varphi\right)} = \frac{\mathrm{R}^2}{a^3}\sin\varepsilon\cos\varepsilon\sec^2\varphi\,.$$

Ex. (2.7) For the helix of Ex. (2.1), (a) locate the center of curvature of any point on the curve; (b) determine the equation of the axis of curvature; (c) find the center of the osculating sphere; and (d) find the radius of the osculating sphere.

Ans. (a) Since the position vector is $p^i = (\rho\,,0\,,z)$ *at any point* $(\rho\,,\theta\,,z)$*, we have*

$$c^i = p^i + \mathrm{R}\,\mu^i = (-a\tan^2\alpha\,,0\,,z)\,;$$

the axis of curvature associated with any point of the helix is the line defined by the vector

$$c^i = p^i + \mathrm{R}\,\mu^i + c\,\nu^i = \left(-\tan^2\alpha\,,-\frac{c\sin\alpha}{a}\,,z + c\cos\alpha\right),$$

where c *is the curve parameter; (c) the center of the osculating sphere is at the point*

$$c^i = p^i + \mathrm{R}\,\mu^i + \frac{1}{\tau}\frac{d\mathrm{R}}{ds}\nu^i = \left(-a\tan^2\alpha\,,0\,,z\right);$$

(d) the radius of the osculating sphere is

$$|c^i - p^i| = \left[\mathrm{R}^2 + \left(\frac{1}{\tau}\frac{d\mathrm{R}}{ds}\right)^2\right]^{1/2},$$

but since $d\mathrm{R}/ds = 0$*, this is* $\mathrm{R} = a\sec^2\alpha$*.*

Ex. (2.8) For the loxodrome of Ex. (2.6), (a) locate the center of curvature, (b) find the center of the osculating sphere, and (c) determine the radius of the osculating sphere.

Ans. (a) Since the position vector of any point (r, θ, φ) *on the curve is* $p^i = (a, 0, 0)$, *we have*

$$c^i = p^i + \mathrm{R}\,\mu^i$$

$$= (a, 0, 0) + \frac{\mathrm{R}^2}{a}\left(-1\,,\, -\frac{\sin\varepsilon\cos\varepsilon\tan\varphi\sec\varphi}{a}\,,\, \frac{\sin^2\varepsilon\tan\varphi}{a}\right)$$

$$= \frac{\mathrm{R}^2}{a^2}\sin\varepsilon\tan\varphi\,(\,a\sin\varepsilon\tan\varphi\,,\, -\cos\varepsilon\sec\varphi\,,\,\sin\varepsilon\,)\,.$$

(b) $c^i = (0, 0, 0)$; *i. e., the center of the osculating sphere is at the center of the sphere on which the loxodrome lies.*

(c) $|c^i - p^i| = \left[\mathrm{R}^2 + \frac{1}{\tau}\left(\frac{d\mathrm{R}}{ds}\right)^2\right]^{1/2} = a\,.$

Ex. (2.9) (a) Show that for the loxodrome of Ex. (2.6),

$$\frac{\tau}{\kappa} = \left(\frac{\kappa'}{\tau\kappa^2}\right)',$$

where a prime denotes differentiation with respect to arc length. (b) Show that this relation is true for all curves which lie on a sphere. (Hint: as with the loxodrome, the radius of the osculating sphere is the radius of the sphere, hence a constant. Therefore, the derivative

$$\frac{d}{ds}|c^i - p^i| = 0\,.)$$

Ex. (2.10) Find the equation of the geodesic when the square of the line element is given in spherical coordinates as

$$(dl)^2 = (dr)^2 + r^2\left[(d\psi)^2 + (\sin^2\psi)^2(d\theta)^2\right],$$

ψ *being the co-latitude on the unit sphere about the origin and* θ *the longitude.*

Ans. The geodesic equations are

$$\ddot{r} - r\left[\psi^2 + \left(\sin^2\psi\right)\dot{\theta}^2\right] = 0\,, \qquad \frac{d}{dl}\left(r^2\psi\right) - r^2\sin\psi\cos\psi\left(\dot{\theta}\right)^2 = 0\,,$$

$$\frac{d}{dl}\left(\,[\,r^2\sin^2\psi\,]\,\dot{\theta}\right) = 0\,.$$

Here dots denote differentiation with respect to l . We may solve these equations most simply by noting that

$$(d\sigma)^2 = (d\psi)^2 + (\sin^2\psi)(d\theta)^2$$

is the square of the line element upon the unit sphere, whose geodesics have been determined in Ex. (2.7.10). The line element is now given in three dimensions as

$$(dl)^2 = (dr)^2 + r^2(d\sigma)^2 ,$$

for which the geodesic equations (see Ex. (2.7.9)) are

$$\ddot{r} - r\dot{\sigma}^2 = 0 , \quad \frac{d}{dl}\left(r^2\dot{\sigma}\right) = 0 .$$

A first integral of the last equation is

$$\dot{\sigma} = \frac{d\sigma}{dl} = \frac{\mathrm{H}}{r^2} , \quad \mathrm{H} = \mathit{constant} .$$

Therefore

$$r^2\frac{d\psi}{dl} = \left(r^2\frac{d\sigma}{dl}\right)\frac{d\psi}{d\sigma} .$$

Furthermore, the third of the geodesic equations has the integral

$$\left(r^2\sin^2\psi\right)\dot{\theta} = h = \mathit{constant} .$$

Using these integrals in the second geodesic equation gives

$$\frac{d^2\psi}{d\sigma^2} - \left(\frac{h}{\mathrm{H}}\right)^2\frac{\cos\psi}{\sin^3\psi} = 0 ,$$

and since

$$\frac{h}{\mathrm{H}} = \left(\sin^2\psi\right)\frac{d\theta}{d\sigma} ,$$

this may finally be given the form

$$\frac{d^2\psi}{d\sigma^2} - \sin\psi\cos\psi\left(\frac{d\theta}{d\sigma}\right)^2 = 0 ,$$

which is identical with the first of equations (2.7.33), and therefore has the same solution. At the same time, the preceding equation is of the same form as equation (2.7.34). Hence by combining the solutions of Exs. (2.7.9-10), we get as the equation of the geodesic

$$r\cos(\sigma - \sigma_0) = r\left[\cos\psi\cos\left(\frac{h}{\mathrm{H}}\right) + \sin\psi\sin\left(\frac{h}{\mathrm{H}}\right)\cos(\theta - \theta_0)\right] = \mathrm{H} .$$

Ex. (2.11) Find the equation of the geodesic when the square of the line element is given in geodesic polar coordinates as:

$(a)\ (dl)^2 = (d\rho)^2 + \sin^2\rho\left[(d\psi)^2 + \left(\sin^2\psi\right)(d\theta)^2\right];$

$(b)\ (dl)^2 = (d\rho)^2 + \sinh^2\rho\left[(d\psi)^2 + \left(\sin^2\psi\right)(d\theta)^2\right];$

$(c)\ (dl)^2 = (dr)^2 + r^2\left[(d\psi)^2 + \left(\sinh^2\psi\right)(d\theta)^2\right];$

$(d)\ (dl)^2 = (d\rho)^2 + \sin^2\rho\left[(d\psi)^2 + \left(\sinh^2\psi\right)(d\theta)^2\right];$

$(e)\ (dl)^2 = (d\rho)^2 + \sinh^2\rho\left[(d\psi)^2 + \left(\sinh^2\psi\right)(d\theta)^2\right];$

$(f)\ (dl)^2 = (d\rho)^2 + \sinh^2\rho\left[(dr)^2 + r^2(d\theta)^2\right];$

$(g)\ (dl)^2 = (dr)^2 + r^2\left[(d\psi)^2 + \psi^2(d\theta)^2\right];$

$(h)\ (dl)^2 = (d\rho)^2 + \sin^2\rho\left[(d\psi)^2 + \psi^2(d\theta)^2\right].$

3. The Dynamics of a Particle

A particle is conveniently defined as a mass point. This means that it has location, as specified by a set of coordinates x^i, and mass, as specified by an invariant m. A particle in motion will have coordinates which are a function of the invariant time t. Hence

$$\mathrm{v}^i = \frac{dx^i}{dt}$$

is the particle's **velocity vector**. Its **momentum vector** is

$$p^i = m\frac{dx^i}{dt} = m\mathrm{v}^i .$$

Newton's Second Law of Motion states that the force upon the particle is measured by the product of its mass and acceleration; that is,

$$\mathrm{F}^i = m\frac{\delta \mathrm{v}^i}{\delta t} = \frac{\delta p^i}{\delta t} \tag{3.1}$$

in an inertial system.*

*An **inertial system** is a frame of reference in which any particle continues in a state of rest or of uniform straight line motion unless acted upon by an external force. **Newton's First Law** thus defines inertial systems and it is only in such systems that Newton's Second and Third Laws are valid.

Ex. (3.1) What are the components of the force vector in spherical coordinates? (Hint: make use of the results of Ex. (1.11).)

Ans.

$$\mathrm{F}^1 = m\left[\ddot{r} - r\left(\cos^2\varphi\right)\dot{\theta}^2 - r\,\dot{\varphi}^2\right],$$

$$\mathrm{F}^2 = m\left[\ddot{\theta} + \frac{2\,\dot{r}\,\dot{\theta}}{r} - 2\left(\tan\varphi\right)\dot{\theta}\,\dot{\varphi}\right],$$

$$\mathrm{F}^3 = m\left[\ddot{\varphi} + \frac{2\,\dot{r}\,\dot{\theta}}{r} + \left(\sin\varphi\cos\varphi\right)\dot{\theta}^2\right],$$

where a dot denotes differentiation with respect to time.

An important invariant defined by the mass and velocity of a particle is its kinetic energy

$$\text{(3.2)} \qquad \mathrm{T} = \frac{1}{2}\,m\left(\frac{ds}{dt}\right)^2 = \frac{1}{2}\,m\,g_{ij}\,\frac{dx^i}{dt}\,\frac{dx^j}{dt} = \frac{1}{2}\,m\,g_{ij}\,\mathrm{v}^i\,\mathrm{v}^j\,.$$

One of the features of the kinetic energy which makes it particularly useful is the fact that from it the force vector may be calculated with comparative ease. Thus consider the equation

$$\mathrm{F}^i = m\,\frac{\delta v^i}{\delta t} = m\left[\frac{\delta}{\delta t}\left(\frac{dx^i}{dt}\right)\right] = m\left[\ddot{x} + \left\{ {i \atop p\;q} \right\}\dot{x}^p\,\dot{x}^q\right],$$

where dots denote differentiation with respect to t. Then

$$g_{ki}\,\mathrm{F}^i = \mathrm{F}_k = m\left(g_{ki}\,\ddot{x}^i + [pq\,,k]\,\dot{x}^p\,\dot{x}^q\right)$$

$$m\left[g_{ki}\,\ddot{x}^i + \frac{1}{2}\left(\frac{\partial g_{kp}}{\partial x^q} + \frac{\partial g_{kq}}{\partial x^p} - \frac{\partial g_{pq}}{\partial x^k}\right)\dot{x}^p\,\dot{x}^q\right]$$

$$= m\left[g_{ki}\,\ddot{x}^i + \frac{1}{2}\left(\frac{\partial g_{kp}}{\partial x^q} + \frac{\partial g_{kq}}{\partial x^p}\right)\dot{x}^p\,\dot{x}^q\right] - \frac{1}{2}\,\frac{\partial g_{pq}}{\partial x^k}\,\dot{x}^p\,\dot{x}^q$$

$$= m\left\{\left[g_{ki}\,\ddot{x}^i + \frac{\partial g_{ki}}{\partial x^q}\,\dot{x}^i\,\dot{x}^q\right] - \frac{1}{2}\,\frac{\partial_{pq}}{\partial x^k}\,\dot{x}^p\,\dot{x}^q\right\}$$

$$= m\left\{\frac{d}{dt}\left(g_{ki}\,\dot{x}^i\right) - \frac{\partial}{\partial x^k}\left(\frac{1}{2}\,g_{pq}\,\dot{x}^p\,\dot{x}^q\right)\right\}$$

$$= \frac{d}{dt}\left[\frac{\partial}{\partial \dot{x}^k}\left(\frac{1}{2}\,m\,g_{pq}\,\dot{x}^p\,\dot{x}^q\right)\right] - \frac{\partial}{\partial x^k}\left(\frac{1}{2}\,m\,g_{pq}\,\dot{x}^p\,\dot{x}^q\right).$$

From equation (3.2) it is clear that this may be written as

$$\text{(3.3)} \qquad \mathrm{F}_k = \frac{d}{dt}\left(\frac{\partial \mathrm{T}}{\partial \dot{x}^k}\right) - \frac{\partial \mathrm{T}}{\partial x^k}.$$

Thus the covariant force vector is obtainable directly in any coordinate system from the kinetic energy T. The expression on the right hand side of equation (3.3) is sometimes referred to as the **Lagrangian or Eulerian derivative** of T.

Ex. (3.2) Find (a) the kinetic energy of a particle and (b) the components of its Lagrangian derivatives in spherical coordinates. (c) From (b), determine the contravariant components of the Lagrangian derivative.

Ans. (a) $\quad \mathrm{T} = \frac{1}{2} m\left[\dot{r}^2 + \left(r^2 \cos^2 \varphi\right)\dot{\theta}^2 + r^2 \dot{\varphi}^2\right].$

(b) $\quad \mathrm{F}_1 = m\left[\ddot{r} - r\left(\cos^2 \varphi\right)\dot{\theta}^2 - r\dot{\varphi}^2\right], \quad \mathrm{F}_2 = m\frac{d}{dt}\left(r^2 [\cos^2 \varphi]\,\dot{\theta}\right),$

$$\mathrm{F}_3 = m\left[\frac{d}{dt}\left(r^2 \dot{\varphi}\right) + r^2 (\sin\varphi \cos\varphi)\,\dot{\theta}^2\right].$$

(c) *Same as Ex. (3.1).*

We can now show quite generally that the Lagrangian derivative of any invariant function $\Phi(x^k, \dot{x}^k)$ is a covariant vector. We do so by showing that it satisfies the covariant vector transformation. To this end, let us consider the two terms of the Lagrangian derivative separately. First, we note that $\Phi\left(\bar{x}^k, \dot{\bar{x}}^k\right) = \Phi\left(x^k, \dot{x}^k\right)$, and that $\bar{x}^k$ and $\dot{\bar{x}}^k$ in the left member can be expressed in terms of $\bar{x}^k$ and $\dot{\bar{x}}^k$ as

$$\bar{x}^i = \bar{x}^i(x^k), \; \dot{\bar{x}}^i = \frac{\partial x^i}{\partial x^k}\dot{x}^k,$$

whence it follows that

$$\frac{\partial \bar{x}^i}{\partial \dot{x}^k} = 0, \quad \frac{\partial \dot{\bar{x}}^k}{\partial \dot{x}^k} = \frac{\partial \bar{x}^i}{\partial x^k}.$$

Therefore the partial derivative of Φ with respect to $\dot{x}^k$ is

$$\text{(3.4)} \qquad \frac{\partial \Phi}{\partial \dot{x}^k} = \frac{\partial \Phi}{\partial \bar{x}^i}\frac{\partial \bar{x}^i}{\partial \dot{x}^k} + \frac{\partial \Phi}{\partial \dot{\bar{x}}^i}\frac{\partial \dot{\bar{x}}^i}{\partial \dot{x}^k},$$

which by the preceding relations becomes

$$\text{(3.5)} \qquad \frac{\partial \Phi}{\partial \dot{x}^k} = \frac{\partial \Phi}{\partial \dot{\bar{x}}^i}\frac{\partial \bar{x}^i}{\partial x^k},$$

showing that $\partial\Phi / \partial\dot{x}^k$ is a covariant vector. (When $\Phi = \mathrm{T}$, it is clear from any concrete example or from equation (3.2) that this vector is the covariant momentum vector.)

The first term of the Lagrangian derivative is now

$$\frac{d}{dt}\left(\frac{\partial\Phi}{\partial\dot{x}^k}\right) = \frac{\partial\bar{x}^i}{\partial x^k}\frac{d}{dt}\left(\frac{\partial\Phi}{\partial\dot{\bar{x}}^i}\right) + \frac{\partial\Phi}{\partial\dot{\bar{x}}^i}\frac{d}{dt}\left(\frac{\partial\bar{x}^i}{\partial x^k}\right)$$

$$= \frac{\partial\bar{x}^i}{\partial x^k}\frac{d}{dt}\left(\frac{\partial\Phi}{\partial\dot{\bar{x}}^i}\right) + \frac{\partial\Phi}{\partial\dot{\bar{x}}^i}\frac{\partial^2\bar{x}^i}{\partial x^m\partial x^k}\dot{x}^m .$$

This term is clearly not by itself a vector. At the same time, the second term of the Lagrangian derivative is

$$\frac{\partial\Phi}{\partial x^k} = \frac{\partial\Phi}{\partial\bar{x}^i}\frac{\partial\bar{x}^i}{\partial x^k} + \frac{\partial\Phi}{\partial\dot{\bar{x}}^i}\frac{\partial\dot{\bar{x}}^i}{\partial x^k} = \frac{\partial\bar{x}^i}{\partial x^k}\frac{\partial\Phi}{\partial\bar{x}^i} + \frac{\partial\Phi}{\partial\dot{\bar{x}}^i}\frac{\partial}{\partial x^k}\left[\frac{\partial\bar{x}^i}{\partial x^m}\dot{x}^m\right]$$

$$= \frac{\partial\bar{x}^i}{\partial x^k}\frac{\partial\Phi}{\partial\bar{x}^i} + \frac{\partial\Phi}{\partial\dot{\bar{x}}^i}\frac{\partial^2\bar{x}^i}{\partial x^m\partial x^k}\dot{x}^m .$$

By subtracting this last equation from the last of the previous equations, we see that

$$\left[\frac{d}{dt}\left(\frac{\partial\Phi}{\partial\dot{x}^k}\right) - \frac{\partial\Phi}{\partial x^k}\right] = \frac{\partial\bar{x}^i}{\partial x^k}\left[\frac{d}{dt}\left(\frac{\partial\Phi}{\partial\dot{\bar{x}}^i}\right) - \frac{\partial\Phi}{\partial\bar{x}^i}\right].$$

From this it is evident that the Lagrangian derivative of *any* invariant function of the coordinates and their derivatives is a covariant vector. In the special case when $\Phi = \mathrm{T}$, we see that the vector is the covariant force vector. The Lagrangian derivative of other invariant functions could be identified similarly when known in explicit coordinate systems.

The force defined in equation (3.1) is distinguished as the **kinematic form**. It is purely a definition, given solely in terms of the observed motion of the particle. One does not have an equation of motion until the kinematically defined force is equated to a force defined *dynamically* by the universe in which the particle is set. For example, the particle may be set in a universe with one other mass, to which it is attracted gravitationally ; the dynamical force is then the inverse square force of gravity. Other universes will clearly define other dynamical forces.

Of most interest in the dynamics of particles are those dynamical forces which may be expressed as the Lagrangian derivatives of an invariant function known as the **potential function**, V. For these, we have that

(3.6)
$$\mathrm{F}_k = \frac{d}{dt}\left(\frac{\partial\mathrm{V}}{\partial\dot{x}^k}\right) - \frac{\partial\mathrm{V}}{\partial x^k} .$$

Note that if V is independent of $\dot{x}^k$, the covariant force vector is given simply by the partial derivative of $-V$ with respect to x^k. This latter is in this case called the **gradient** of $-V$. It is important to realize that in general this is a vector only when V is independent of $\dot{x}^k$. Velocity dependent potential functions are the exception, however, so that it is often stated somewhat loosely that the force may be derived as the gradient of the potential.

Ex. (3.3) Given the gravitational potential function

$$V = -\frac{\mu M}{r}, \quad \mu = G(M + m),$$

in spherical coordinates, where G *is the constant of gravitation and* M *and* m *are the masses of the two attracting bodies, what are the components of the Lagrangian derivatives of* V*?*

Ans. $F_1 = -\dfrac{\mu M}{r^2}, \; F_2 = 0, \; F_3 = 0.$

Ex. (3.4) Consider the coordinate system related to spherical coordinates by the sub-tensor transformation

$$\bar{x}^1 = u = \frac{1}{r} = \frac{1}{x^1}, \; \bar{x}^2 = \theta = x^2, \; \bar{x}^3 = \varphi = x^3.$$

(a) Find the covariant components of the force vector in this coordinate system. (Hint: transform the equations obtained in Ex. (3.2)). (b) Express the gravitational potential (see Ex. (3.3)) in these coordinates and find its Lagrangian derivatives.

Ans. (a)

$$F_1 = \frac{G m M}{u^5}\left[u\ddot{u} - 2\dot{u}^2 + \left(u^2 \cos^2\varphi\right)\dot{\theta}^2 + u^2\dot{\varphi}^2\right],$$

$$F_2 = G m M \frac{d}{dt}\left(\frac{\cos^2\varphi}{u^2}\dot{\theta}\right), \; F_3 = G m M\left[\frac{d}{dt}\left(\frac{\dot{\varphi}}{u^2}\right) + \frac{\sin\varphi\cos\varphi}{u^2}\dot{\theta}^2\right].$$

(b) $F_1 = \mu, \; F_2 = 0 = F_3.$

Ex. (3.5) Find the Lagrangian derivatives of the velocity-dependent gravitational potential

$$V = -\mu u\left(1 + \frac{\dot{u}}{c u^2}\right)^{-2}, \; \mu = constant,$$

expressed in the coordinate system of Ex. (3.4), where c *is a universal constant. (This potential function may be seen to be an invariant function if we write it as*

$$V = -\frac{\mu}{|p^i|}\left[1 + \frac{\delta p^j}{\delta t} e_j^{(1)}\right]^{-2},$$

where p^i is the position vector, $\delta p^i/\delta t = \mathrm{v}^i$ the velocity vector, and $e_j^{(1)}$ the covariant basis vector along the x^1–curve in a spherical coordinate system. Under the transformation of Ex. (3.4) it takes the form given above.)

This classical potential function is of particular interest because it predicts correctly the advance of the perihelion of Mercury (see Appendix 4.1). It was proposed in 1898 by Paul Gerber (Zs. f. Math. u. Phys., v. xliii, pp. 93-104). He interprets the constant c as the velocity of gravitation, which is found to be the same as the velocity of light.

Ans.

$$\mathrm{F}_1 = \mu\left(1 + \frac{\dot{u}}{c\,u^2}\right)^{-4}\left\{1 + \frac{4\dot{u} - 6\,u\,\ddot{u} + 15\,\dot{u}^2}{c\,u^2}\right\},$$

$$\mathrm{F}_2 = 0 = \mathrm{F}_3 .$$

We may now equate the kinematic and dynamical forces defined respectively by equations (3.3) and (3.6). The resultant equation of motion is

$$\frac{d}{dt}\left(\frac{\partial \mathrm{T}}{\partial \dot{x}^k}\right) - \frac{\partial \mathrm{T}}{\partial x^k} = \frac{d}{dt}\left(\frac{\partial \mathrm{V}}{\partial \dot{x}^k}\right) - \frac{\partial \mathrm{V}}{\partial x^k} .$$

Let us define the **Lagrangian function** to be

(3.7) $$\mathrm{L} = \mathrm{T} - \mathrm{V} .$$

From the preceding equation it therefore follows that the equations of motion of a particle are

(3.8) $$\frac{d}{dt}\left(\frac{\partial \mathrm{L}}{\partial \dot{x}^k}\right) - \frac{\partial \mathrm{L}}{\partial x^k} = 0 .$$

They are called **Lagrange's equations**.

Ex. (3.6) Find Lagrange's equations, in a spherical coordinate system, for the harmonic oscillator.

Ans. Since $\mathrm{V} = -\,k\,r$, $k = \textit{constant}$, *the Lagrangian function is*

Hence

$$\mathrm{L} = \mathrm{T} - \mathrm{V} = \frac{1}{2}\,m\left\{\dot{r}^2 + r^2\left[\left(\cos^2\varphi\right)\dot{\theta}^2 + \dot{\varphi}^2\right]\right\} + k^2\,r .$$

$$m\,\ddot{r} - m\,r\left[\left(\cos^2\varphi\right)\dot{\theta}^2 + \dot{\varphi}^2\right] + k^2 = 0 ,$$

$$\frac{d}{dt}\left(\left[m\,r^2\cos^2\varphi\right]\dot{\theta}\right) = 0 ,$$

$$\frac{d}{dt}\left(m\,r^2\,\dot{\varphi}\right) + m\,r^2\left(\sin\varphi\cos\varphi\right)\dot{\theta}^2 = 0 .$$

Ex. (3.7) Determine Lagrange's equations for a simple pendulum. Take it to be a spherical bob of mass m at the end of a rigid weightless bar of length l. Use the displacement angle θ as variable (see Fig. 83).

Ans. Since $-V = mgl\cos\theta, \qquad L = \frac{1}{2}m\left(l\dot{\theta}\right)^2 + mgl\cos\theta.$

Hence we get

$$ml^2\ddot{\theta} + mgl\sin\theta = 0,\ \ddot{\theta} = -\frac{g}{l}\sin\theta.$$

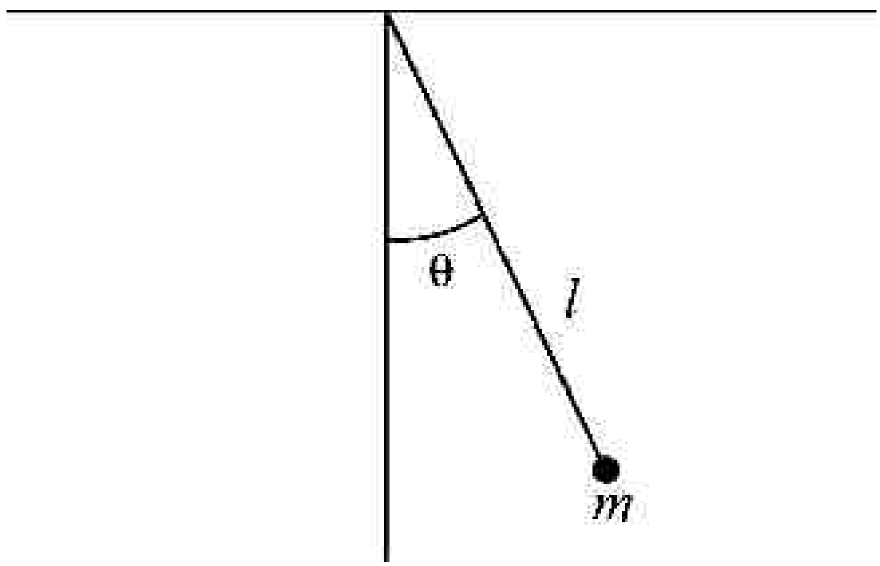

Figure 83

Ex. (3.8) Suppose that the pendulum of Ex. (3.7) is elastic rather than rigid. Then what do Lagrange's equations become?

Ans. Take r_0 to be the equilibrium length of the rod. Then extension or compression of the rod gives it a potential energy $V = \frac{1}{2}k^2\left(r - r_0\right)^2$, *where k^2 is the elastic constant. Therefore*

$$L = \frac{1}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - \frac{1}{2}k^2\left(r - r_0\right)^2 + mgr\cos\theta.$$

Hence

$$m\ddot{r} + k^2\left(r - r_0\right) + mg\cos\theta = 0,$$

$$\frac{d}{dt}\left(mr^2\dot{\theta}\right) + mgr\sin\theta = 0.$$

Ex. (3.9) Consider the compound pendulum (Fig. 84). It consists of masses m_1 and m_2 (subscripts here are not tensor indices) at the ends of rigid weightless rods of respective lengths $r_1 = l_1 = constant$ and $r_2 = l_2 = constant$ as shown. What is the Lagrangian function L of this system?

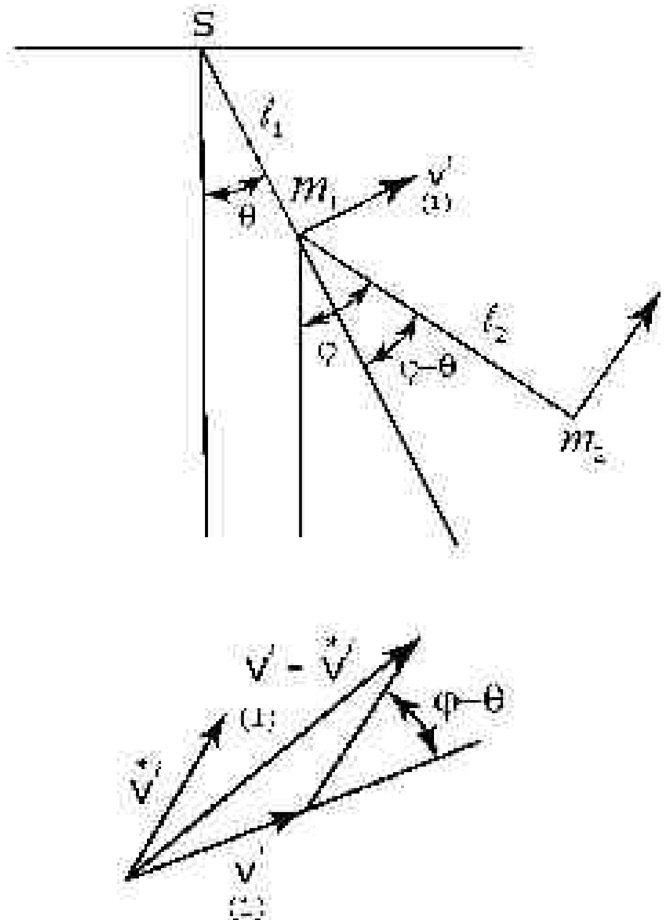

Figure 84

Ans. From the figure, we see that the velocity of m_1 *relative to the point of suspension* S *is* $v^i_{(1)}$, *where* $| v_{(1)} | = l_1 \dot{\theta}$. *The velocity of* m_2 *relative to* m_1 *is* v^{*i}, *where* $| v^{*i} | = l_2 \dot{\varphi}$. *Hence the velocity of* m_2 *relative to* S is $v^i_{(1)} + v^{*i} = v^i_{(2)}$ *and*

$$L = \frac{1}{2} m_1 v_i^{(1)} v^i_{(1)} + \frac{1}{2} m_2 v_i^{(2)} v^i_{(2)} - V$$

$$= \frac{1}{2} m_1 \left(l_1 \dot{\theta} \right)^2 + \frac{1}{2} m_2 \left[\left(l_1 \dot{\theta} \right)^2 + 2 l_1 l_2 \dot{\theta} \dot{\varphi} \cos (\varphi - \theta) + \left(l_2 \dot{\varphi} \right)^2 \right]$$

$$+ m_1 l_1 \cos \theta + m_2 \left(l_1 \cos \theta + l_2 \cos \varphi \right).$$

Ex. (3.10) Consider a mass μ $(1 > \mu > 0)$ *at* $(a, 0)$ *and a mass* $(1 - \mu)$ *at* $(-a, 0)$ *as in Fig. 85. A very small mass* m *is at a point* P *whose plane bipolar coordinates are* (ρ, σ). *If* μ *and* $(1 - \mu)$ *are fixed and if* m *is constrained to motion in the plane under the gravitational attraction, the problem becomes the celebrated* ***two-center problem****. It is the only three-body* *problem of celestial mechanics whose general solution can be expressed in terms of quadratures. Derive and solve Lagrange's equations for the small mass in elliptical coordinates.*

Ans. See Appendix (4.2).

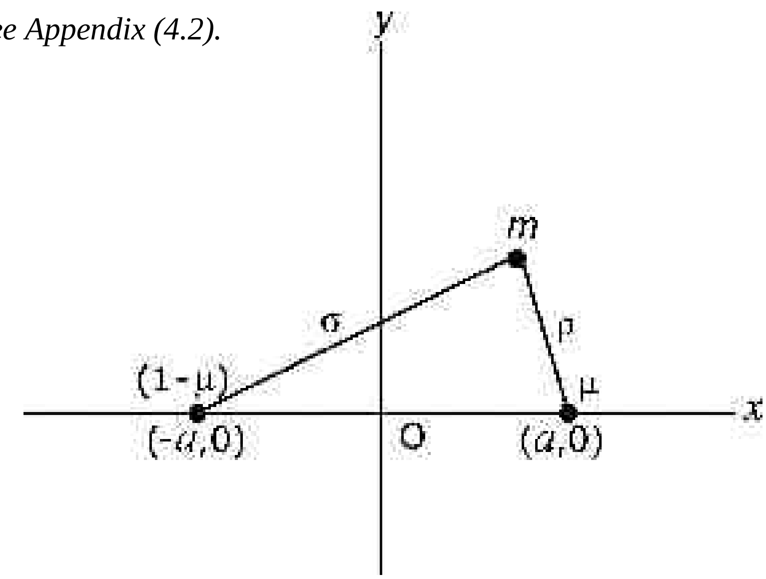

Figure 85

Ex. (3.11) Show that the kinetic energy of a rigid body is

$$T = \frac{1}{2} m \, g_{ij} \dot{\xi}^i \dot{\xi}^j + \frac{1}{2}\left[\left(g^{ij} I_{ij}\right) g_{pq} - I_{pq}\right] \omega^p \omega^q ,$$

where I_{ij} *is the moment of inertia tensor,* ω^i *the angular velocity,* m *the total mass, and* $\dot{\xi}^i$ *the velocity of the center of gravity.*

Ex. (3.12) A particle of mass m *is constrained by gravity to move frictionlessly over the outer surface of a hemisphere. Find the equations of motion.*

Ans. Take the center of the sphere of radius a *to be the origin of a spherical coordinate system. Then*

$$T = \frac{1}{2} m \, a^2 \left[\left(\cos^2 \varphi\right) \dot{\theta}^2 + \dot{\varphi}^2\right],$$

$$V = m \, g \, a \cos \varphi .$$

Then Lagrange's equations become

$$\frac{d}{dt}\left[\left(\cos^2 \varphi\right) \dot{\theta}\right] = 0 ,$$

$$\ddot{\varphi} - \left(\sin \varphi \cos \varphi\right) \dot{\theta}^2 + \frac{g \sin \varphi}{a} = 0 .$$

(Note that $\dot{\theta} = 0$ *leads to results similar to those of Ex. (3.7).)*

The importance and utility of Lagrange's equations can be grasped only very imperfectly from the few illustrations here presented. In Ex. (3.6) we see how they give Newton's equations 11 of motion in a very straightforward and economical way. In Ex. (3.7), we consider a standard problem which, however, has relatively simple yet interesting generalizations in Exs. (3.8) and (3.9). Each new variable required defines another **degree of freedom**. Clearly, the number of degrees of freedom may be increased indefinitely, independent of the number of dimensions of the space. In Ex. (3.10) we have an illustration of the value of certain specialized coordinate systems, without the use of which the solution of some problems would be very difficult or even impossible. Lagrange's equations apply also to rigid bodies, as may be seen in Ex. (3.11); in fact, they can be shown to hold for continuous media as well. And finally, problems which include constraints, as in Ex. (3.12), may be readily handled by Lagrange's equations.

Lagrange's equations are not only convenient as equations of motion but convenient also for the close relation they bear to other forms of the equations of motion.*

To consider these, let us first define the **generalized momentum**; it is

(3.9) $$p_k = \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\,.$$

If the potential V is independent of $\dot{x}^k$, this is the same as the ordinary momentum previously defined as

$$p_k' = \frac{\partial \mathrm{T}}{\partial \dot{x}^k}\,.$$

Otherwise it is more general, whence the designation as "generalized momentum". (It is clearly a covariant vector, obtained by setting $\Phi = \mathrm{L}$.) Now from Lagrange's equations we have that

(3.10) $$\dot{p}_k = \frac{\partial \mathrm{L}}{\partial x^k}\,.$$

We now look upon equation (3.9) as the definition of a new variable p_k. We write it as

$$p_k = p_k(x^s, \dot{x}^s)\,.$$

Assuming that this is a reversible transformation, we can solve it for $\dot{x}^s$ as

(3.11) $$\dot{x}^s = \dot{x}^s(p_k, x^k)\,.$$

---

* In particular, they are derived from a **Legendre transformation**, by which a function $f(x,y)$ becomes a function $g(x,z)$, where $z = \partial F/\partial y$ and such that $\partial g/\partial z = y$. In our application of the Legendre transformation, the function $f(x,y)$ is $\mathrm{L}(x^s,\dot{x}^s)$, where x stands for x^s and y for $\dot{x}^s$. Then $p_k = \partial \mathrm{L}/\partial \dot{x}^k$ becomes z, and, as is shown in what follows, $g(x,z)$ is the Hamiltonian function $\mathrm{H}(x^i,p_i) = \dot{x}^k p_k - \mathrm{L}(x^s,\dot{x}^s[x^i,p_i])$.

Note that this is not a simple point transformation of coordinates, such that the values of x^s determine the values of $\bar{x}^s$. It is a more general transformation which determines p_k from x^s and $\dot{x}^s$ or $\dot{x}^s$ from x^k and p_k.

Next, from the vectors x^k, p_k and the invariant $\mathrm{L}(x^s, \dot{x}^s[p_i, x^i])$ we form a new invariant

(3.12) $$\mathrm{H}(x^i, p_i) \equiv \dot{x}^k p_k - \mathrm{L}\,.$$

It is called the **Hamiltonian function**. Differentiating partially with respect to p_s, we get

(3.13) $$\frac{\partial \mathrm{H}}{\partial p_s} = \dot{x}^s + p_k \frac{\partial \dot{x}^k}{\partial p_s} - \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\frac{\partial \dot{x}^k}{\partial p_s} = \dot{x}^s + \left(p_k - \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\right)\frac{\partial \dot{x}^k}{\partial p_s} = \dot{x}^s$$

Differentiating partially with respect to x^s next, we get

(3.14) $$\frac{\partial \mathrm{H}}{\partial x^s} = p_k \frac{\partial \dot{x}^k}{\partial x^s} - \frac{\partial \mathrm{L}}{\partial x^s} - \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\frac{\partial \dot{x}^k}{\partial x^s} = \left(p_k - \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\right)\frac{\partial \dot{x}^k}{\partial x^s} - \dot{p}_s = -\dot{p}_s$$

according to equations (3.9) and (3.10). From equations (3.13) and (3.14), we thus have in final form **Hamilton's canonical equations**

(3.15) $$\dot{x}^s = \frac{\partial \mathrm{H}}{\partial p_s}, \quad \dot{p}_s + \frac{\partial \mathrm{H}}{\partial x^s} = 0\,.$$

It is interesting to note that $\dot{x}^s$ is a vector, since generalized velocities in any case are so. Therefore $\partial \mathrm{H}/\partial p_s$ is likewise a vector. On the other hand, though (as we have shown) p_s is a covariant vector, $\dot{p}_s$ is in general not a vector except in a rectilinear coordinate system, where it is identical to $\delta p_r/\delta t$. However, since

$$\dot{p}_r + \frac{\partial \mathrm{H}}{\partial x^r} = 0$$

in one coordinate system, it must be zero in all. Hence this quantity is a covariant vector even though its separate terms need not be. It is for this reason that the second of equations (3.15) has been put in this form. *In this form*, Hamilton's equations are vector equations.

Ex. (3.13) Find Hamilton's equations for the harmonic oscillator in three dimensions.

Ans. From Ex. (3.6), we have

$$p_1 = \frac{\partial \mathrm{L}}{\partial \dot{x}^1} = m\dot{r}\,,\ p_2 = \frac{\partial \mathrm{L}}{\partial \dot{x}^2} = m r^2(\cos^2\varphi)\,\dot{\theta}\,,\ p_3 = \frac{\partial \mathrm{L}}{\partial \dot{x}^3} = m r^2\,\dot{\varphi}\,.$$

Therefore

$$\mathrm{H} = \frac{1}{2m}\left[(p_1)^2 + \frac{(p_2)^2}{r^2\cos^2\varphi} + \frac{(p_3)^2}{r^2}\right] - k^2 r\,.$$

Hence

$$\dot{p}_1 = -\frac{\partial \mathrm{H}}{\partial x^1} = \frac{1}{m\,r^3}\left[\frac{(p_2)^2}{\cos^2\varphi} + (p_3)^2\right] + k^2 ,$$

$$\dot{p}_2 = -\frac{\partial \mathrm{H}}{\partial x^2} = 0 ,\ \dot{p}_3 = -\frac{\partial \mathrm{H}}{\partial x^3} = \frac{(p_2)^2 \sin\varphi}{m\,r^2\cos^3\varphi} ,$$

$$\dot{r} = \dot{x}^1 = \frac{\partial \mathrm{H}}{\partial p_1} = \frac{p_1}{m} ,\ \dot{\theta} = \dot{x}^2 = \frac{\partial \mathrm{H}}{\partial p_2} = \frac{p_2}{m\,r^2\cos^2\varphi} ,$$

$$\dot{\varphi} = \dot{x}^3 = \frac{\partial \mathrm{H}}{\partial p_3} = \frac{p_3}{m\,r^2} .$$

Ex. (3.14) Find Hamilton's equations for the simple pendulum.

Ans. From Ex. (3.7) we find that

$$\dot{p} = -\,m\,g\,l\sin\theta ,\ \dot{\theta} = \frac{p}{m\,l^2} .$$

Ex. (3.15) From Hamilton's equations, show that

$$\frac{\mathrm{dH}}{dt} = 0 ,\ \mathrm{H} = \mathrm{E} = \mathit{constant} .$$

Hence show that when the kinetic energy T *is a quadratic homogeneous function of* $\dot{x}^r$ *and the potential energy* V *is independent of* $\dot{x}^r$*, we have*

$$\mathrm{H} = \dot{x}^k p_k - \mathrm{L} = 2\,\mathrm{T} - \left(\mathrm{T} - \mathrm{V}\right) = \mathrm{T} + \mathrm{V} = \mathrm{E} ,$$

where E *is the total energy of the system. Such a system is said to be* ***conservative****.*

Ans. For any system whose Hamiltonian is $\mathrm{H}(x^r , p_r)$*, we have*

$$\frac{\mathrm{dH}}{dt} = \frac{\partial \mathrm{H}}{\partial x^r}\frac{dx^r}{dt} + \frac{\partial \mathrm{H}}{\partial p_r}\frac{dp_r}{dt} = \frac{\partial \mathrm{H}}{\partial x^r}\frac{\partial \mathrm{H}}{\partial p_r} - \frac{\partial \mathrm{H}}{\partial p_r}\frac{\partial \mathrm{H}}{\partial x^r} = 0 ,$$

whence the first result follows.

If this T *is a quadratic homogeneous function of* $\dot{x}^r$*, then by Gauss's theorem on homogeneous functions*

$$\dot{x}^r \frac{\partial \mathrm{T}}{\partial \dot{x}^r} = 2\,\mathrm{T} ,$$

whence the second result follows if we define the total energy to be the sum of the kinetic and potential energies.

4. Geometrical Dynamics

Lagrange's equations are also closely related to still another form of the equations of motion. We recognize the equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^k}\right) - \frac{\partial L}{\partial x^k} = 0$$

as identical with the equations to be satisfied by that curve along which the line integral

(4.1) $$A = \int_{P_0}^{P} L\, dt$$

is an extremum. The quantity A is called the **action** when L is the Lagrangian function. Lagrange's equations therefore require that the motion of a particle be such that its action integral is an extremum — in particular, a minimum. Therefore the motion of a particle is said to obey the **Principle of Least Action**.

As we have seen (Ex. (3.15)), for a conservative system it is true that $T + V = E = \text{constant}$ so that $L = T - V = 2T - E$. In this case, the action integral is

$$A = \int L\, dt = \int (2T - E)\, dt = \int 2T\, dt - E(t - t_0).$$

Therefore, if the times t and t_0 are fixed, and if we consider variations of the path under this condition, it will be true that

$$\delta A = 2\delta \int T\, dt = 2\delta \int (E - V)\, dt = 0$$

is the condition for an extremum. Now

$$2T = m\left(\frac{ds}{dt}\right)^2$$

so that

$$dt = \frac{dt}{ds}\, ds = \left[\frac{m}{2T}\right]^{1/2} = \left[\frac{m}{2(E - V)}\right]^{1/2} ds\, .$$

Therefore the principle of least action for a conservative system requires that

(4.2) $$\delta A = \sqrt{2m(E - V)}\, ds = 0\, .$$

Let us now define a conformal mapping (see Ch. 2, Sec. 8, Eq. (8.5))

(4.3) $$\bar{g}_{ij} = (E - V)\, g_{ij}\, .$$

In terms of the new line element, the action integral becomes

$$\bar{A} = \int \left(\bar{g}_{ij}\, dx^i\, dx^j\right)^{1/2} = \int dS\, .$$

Since

$$\delta \bar{A} = \delta \int d\mathrm{S} = 0 ,$$

it is clear that the action is stationary when the path length on the map is minimal - i. e., geodesic. The quantity $d\mathrm{S}$ is called the **action line element** to distinguish it from the **kinematic line element** ds. Thus *the geodesics of the action line element map into the natural trajectories of the kinematic line element, and conversely*. This process is sometimes referred to as the **geometrization of dynamics**. It is of considerable importance in the general theory of relativity.

Ex. (4.1) Show that

$$\left\{ \begin{matrix} r \\ i\ n \end{matrix} \right\} - \overline{\left\{ \begin{matrix} r \\ i\ n \end{matrix} \right\}} = \frac{1}{(\mathrm{E} - \mathrm{V})} \left(\delta_i^r \frac{\partial \mathrm{V}}{\partial x^n} + \delta_n^r \frac{\partial \mathrm{V}}{\partial x^i} - g_{in} g^{rs} \frac{\partial \mathrm{V}}{\partial x^s} \right) ,$$

where $\left\{ \begin{matrix} r \\ i\ n \end{matrix} \right\}$ *is the Christoffel symbol derived from the kinematic line element and* $\overline{\left\{ \begin{matrix} r \\ i\ n \end{matrix} \right\}}$ *is that derived from the action line element.*

Ex. (4.2) Show that Lagrange's equations in the kinematic metric become the geodesic equations in the action metric.

Ans. We write Lagrange's equations in the contravariant form

$$m \left[\frac{d^2 x^r}{dt^2} + \left\{ \begin{matrix} r \\ i\ n \end{matrix} \right\} \frac{dx^i}{dt} \frac{dx^n}{dt} \right] = - g^{rs} \frac{\partial \mathrm{V}}{\partial x^s} .$$

Substituting for $\left\{ \begin{matrix} r \\ i\ n \end{matrix} \right\}$ *from Ex. (4.1), we get*

$$m \left[\frac{d^2 x^r}{dt^2} + \overline{\left\{ \begin{matrix} r \\ i\ n \end{matrix} \right\}} \frac{dx^i}{dt} \frac{dx^n}{dt} + \frac{1}{(\mathrm{E} - \mathrm{V})} \left(\frac{\partial \mathrm{V}}{\partial x^k} \frac{dx^k}{dt} \right) \frac{dx^r}{dt} \right]$$

$$= - g^{rs} \frac{\partial \mathrm{V}}{\partial x^s} \left[1 - \frac{m g_{in}}{2(\mathrm{E} - \mathrm{V})} \frac{dx^i}{dt} \frac{dx^n}{dt} \right].$$

However, since

$$2(\mathrm{E} - \mathrm{V}) = m g_{in} \frac{dx^i}{dt} \frac{dx^n}{dt} ,$$

the right hand side vanishes. Hence

$$\frac{d^2 x^r}{dt^2} + \overline{\left\{ \begin{matrix} r \\ i\ n \end{matrix} \right\}} \frac{dx^i}{dt} \frac{dx^n}{dt} = - \left(\frac{1}{\mathrm{E} - \mathrm{V}} \frac{d\mathrm{V}}{dt} \right) \frac{dx^r}{dt}$$

$$= \left[\frac{d}{dt} \ln(\mathrm{E} - \mathrm{V}) \right] \frac{dx^r}{dt} .$$

By Ch. 2, §4, this is equivalent to

$$\frac{d^2x^r}{dS^2} + \overline{\begin{Bmatrix} r \\ i\ n \end{Bmatrix}} \frac{dx^i}{dS}\frac{dx^n}{dS} = 0$$

where

$$dS = \sqrt{2/m}\,(E - V)\,dt\,.$$

Ex. (4.3) Is the mapping defined by equation (4.3) geodesic? (Hint: see Ch. 2, Sec. 8, Eq. (8.11).)

Ans. No. (Caution: note that equation (8.11) must be modified for three-dimensional mappings. In this case

$$\varphi_k = \frac{1}{4} A^r_{kr}\,.$$

In n dimensions,

$$\varphi_k = \frac{1}{n+1} A^r_{kr}\,.)$$

5. Green's and Stokes' Theorems

We have seen that vectors of special interest may be generated by differentiating invariants; the Lagrangian derivative is an example. A case of particular interest arises when an invariant $V(x^s)$ is a function of the coordinates only. In that case, the vector

$$V_{,r} = \frac{\partial V}{\partial x^r}$$

to which, apart from sign, the Lagrangian reduces in this instance, is called the **gradient** of V. Examples of invariants whose gradients arise in physical applications are temperature, gravitational potential, chemical concentration, electric charge density, etc.

Gradients also figure frequently in a converse process of generation of invariants from derivatives. For example, consider a vector X^s. From its covariant derivatives $X^s_{,r}$ we can form the invariant $X^s_{,s}$. This quantity is called the **divergence** of the vector X^s. Clearly, the divergence of a covariant vector may also be formed simply by first putting it in a contravariant form. Thus

(5.1) $$X^s = g^{sr}X_r\,,\ X^s_{,s} = g^{sr}X_{r,s}\,,$$

so that the final expression is the divergence of X_r. When X_r is the gradient of V, this becomes

$$X_r = V_{,r}\,,\ g^{rs}V_{,rs} \equiv \nabla^2 V\,.$$

The operator here denoted by ∇^2 is called the **Laplacian**. It is an important invariant operator.

The divergence and Laplacian operators are of such frequent occurrence in mathematical physics that it is convenient to express them in a form most expedient for calculation in explicit coordinate systems. Consider the divergence of X^s. It is

$$X^s_{,s} = \frac{\partial X^s}{\partial x^s} + \left\{ {s \atop s\ p} \right\} X^p .$$

To find a useful expression for $\left\{ {s \atop s\ p} \right\}$, we first note that

$$\boldsymbol{\varepsilon}_{rst,p} = \frac{\partial \boldsymbol{\varepsilon}_{rst}}{\partial x^p} - \left\{ {m \atop r\ p} \right\} \boldsymbol{\varepsilon}_{mst} - \left\{ {m \atop s\ p} \right\} \boldsymbol{\varepsilon}_{rmt} - \left\{ {m \atop t\ p} \right\} \boldsymbol{\varepsilon}_{rsm} = 0 .$$

Substituting for $\boldsymbol{\varepsilon}_{rst}$ its equivalent, $\sqrt{g}\, e_{rst}$, this gives

$$e_{rst} \frac{\partial \sqrt{g}}{\partial x^p} = \sqrt{g} \left[\left\{ {m \atop r\ p} \right\} e_{mst} + \left\{ {m \atop s\ p} \right\} e_{rmt} + \left\{ {m \atop t\ p} \right\} e_{rsm} \right].$$

Consider the particular case when $r, s, t \equiv 1, 2, 3$. Then

$$\frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial x^p} = \left\{ {m \atop 1\ p} \right\} e_{m23} + \left\{ {m \atop 2\ p} \right\} e_{1m3} + \left\{ {m \atop 3\ p} \right\} e_{12m} .$$

Clearly, the only terms on the right hand side which survive the summation over m are the ones in which $m = 1$ in the first, $m = 2$ in the second, and $m = 3$ in the third. Hence (compare with Ex. (3.10) of Ch. 2)

$$\frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial x^p} = \left\{ {1 \atop 1\ p} \right\} + \left\{ {2 \atop 2\ p} \right\} + \left\{ {3 \atop 3\ p} \right\} = \left\{ {s \atop s\ p} \right\} . \tag{5.2}$$

We may now use this identity in the expression for the divergence of X^s. The result is

$$X^s_{,s} = \frac{\partial X^s}{\partial x^s} + \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial x^p} X^p = \frac{1}{\sqrt{g}} \left[\sqrt{g} \frac{\partial X^s}{\partial x^s} + X^s \frac{\partial \sqrt{g}}{\partial x^s} \right] \quad \text{or}$$

(5.3)

$$X^s_{,s} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^s} \left(\sqrt{g}\, X^s \right).$$

Because of its simplicity, this is the most expeditious formula for the divergence of a vector X^s.

If we now combine the results of equations (5.1) and (5.3), we find for the Laplacian of V the result

$$\nabla^2 V = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^s} \left(\sqrt{g}\, g^{sr} \frac{\partial V}{\partial x^r} \right) .$$

In Cartesian coordinates, the Laplacian takes the familiar form

$$\nabla^2 \mathrm{V} = \frac{\partial^2 \mathrm{V}}{\partial x^2} + \frac{\partial \mathrm{V}}{\partial y^2} + \frac{\partial \mathrm{V}}{\partial z^2}.$$

Ex. (5.1) Find the Laplacian of a function V *in (a) cylindrical coordinates, (b) spherical coordinates, and (c) parabolic cylindrical coordinates. (Hint: see Exs. (1.10, 1.11, 1.12).)*

Ans. (a) $$\nabla^2 \mathrm{V} = \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{\partial \mathrm{V}}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 \mathrm{V}}{\partial \theta^2} + \frac{\partial^2 \mathrm{V}}{\partial z^2}.$$

(b) $$\nabla^2 \mathrm{V} = \frac{1}{r^2}\left\{\frac{\partial}{\partial r}\left(r^2\frac{\partial \mathrm{V}}{\partial r}\right) + \frac{1}{\cos\varphi}\left[\frac{1}{\cos\varphi}\frac{\partial \mathrm{V}}{\partial \theta^2} + \frac{\partial}{\partial \varphi}\left(\cos\varphi\frac{\partial \mathrm{V}}{\partial \varphi}\right)\right]\right\}.$$

(c) $$\nabla^2 \mathrm{V} = \frac{1}{u^2 + \mathrm{v}^2}\left(\frac{\partial^2 \mathrm{V}}{\partial u^2} + \frac{\partial^2 \mathrm{V}}{\partial \mathrm{v}^2}\right) + \frac{\partial^2 \mathrm{V}}{\partial z^2}.$$

Ex. (5.2) Find the divergence of a vector f^s *in (a) cylindrical coordinates, (b) spherical coordinates, and (c) parabolic cylindrical coordinates.*

Ans. (a) $$f^s_{,s} = \frac{1}{r}\frac{\partial}{\partial r}\left(rf^1\right) + \frac{\partial f^2}{\partial \theta} + \frac{\partial f^3}{\partial z}.$$

(b) $$f^s_{,s} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 f^1\right) + \frac{\partial f^2}{\partial \theta} + \frac{1}{\cos\varphi}\frac{\partial}{\partial \varphi}\left(f^3\cos\varphi\right).$$

(c) $$f^s_{,s} = \frac{1}{u^2 + \mathrm{v}^2}\left\{\frac{\partial}{\partial u}\left[\left(u^2 + \mathrm{v}^2\right)f^1\right] + \frac{\partial}{\partial v}\left[\left(u^2 + \mathrm{v}^2\right)f^2\right]\right\} + \frac{\partial f^3}{\partial z}.$$

Perhaps the most important use of the divergence is in the statement of Green's theorem, which permits the transformation of surface integrals into volume integrals or vice versa. We state the theorem first in Cartesian coordinates, then translate it into generalized coordinates. The theorem says that *if* σ *is a closed surface bounding a volume* τ *(Fig. 86), and if* f^1, f^2 *and* f^3 *are three continuous functions having partial derivatives of the first order throughout* τ *and over* σ*, then*

$$\int_\sigma \left(n_1 f^1 + n_2 f^2 + n_3 f^3\right) d\sigma = \int_\tau \left(\frac{\partial f^1}{\partial x} + \frac{\partial f^2}{\partial y} + \frac{\partial f^3}{\partial z}\right) d\tau,$$

where n_1, n_2, and n_3 are the direction cosines of the outward normal of σ. This theorem, to be found in any text on advanced calculus, may be proven readily by integrating partially along lines parallel to the coordinate axes.

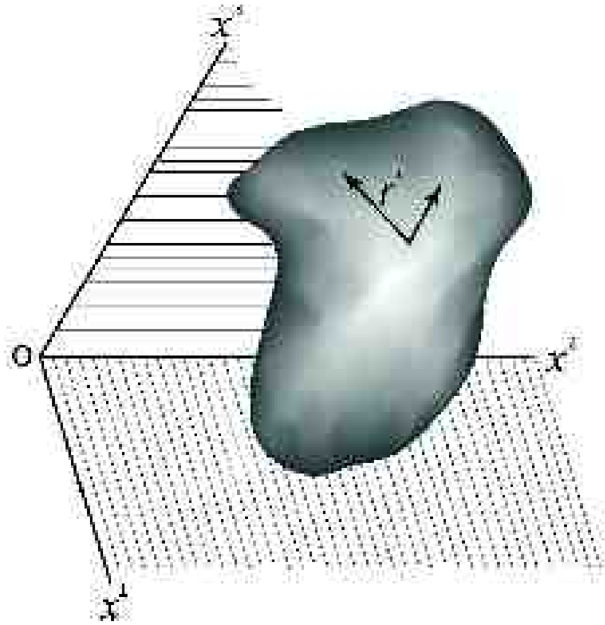

Figure 86

Let us now identify the integrand on the left hand side. The direction cosines are simply the covariant components of the unit normal n_i. The functions f^1, f^2 and f^3 may be taken as the components of a contravariant vector f^i. Hence the integrand on the left is nothing but $n_i f^i$, an invariant.

Now let us write the integrand in the right hand side as

$$\frac{\partial f^1}{\partial x} + \frac{\partial f^2}{\partial y} + \frac{\partial f^3}{\partial z} = \frac{\partial f^1}{\partial x^1} + \frac{\partial f^2}{\partial x^2} + \frac{\partial f^3}{\partial x^3} = \frac{\partial f^s}{\partial x^s}.$$

Remembering that in Cartesian coordinates

$$\frac{\partial f^s}{\partial x^s} = f^s_{,s},$$

we see that this integrand is the divergence of f^s, an invariant.

By identifying the integrands in Green's theorem as invariants, we are now able to express the theorem in any system of coordinates. It is contained simply in the equation

$$\int_\sigma n_i f^i \, d\sigma = \int_\tau f^i_{,i} \, d\tau . \tag{5.5}$$

The quantity $n_i f^i$ is called the **flux** of f^i across the surface σ. This terminology accords with the interpretation of $n_i f^i$ when f^i represents the rate of flow of matter, momentum, energy or similar quantity.

To interpret the divergence $f^i_{,i}$, let us consider a unit volume τ. If there is a positive net flux from τ, as measured by the left hand side of equation (5.5), this means that more of the quantity whose flow is represented by f^i has flowed out of τ than flowed into τ. There has thus been a creation of this quantity within τ. But the right hand side of equation (5.5) is the integral of $f^i_{,i}$ over the unit volume τ. Hence the divergence of f^i must represent the rate at which the quantity is created per unit volume. Therefore, if $f^i_{,i} > 0$, the quantity is said to have a **source** within τ; if $f^i_{,i} < 0$, the quantity is said to have as **sink** within τ. If $f^i_{,i} = 0$, the quantity is **conserved**.

Ex. (5.3) (a) Find the divergence of the vector whose Cartesian components are $f^i = (\sinh x \sin y, \cosh x \cos y, 0)$. *(b) Draw representative vectors of this field in the plane* $z = 0$. *(c) What is the integral of the flux of this vector over any closed surface* σ*?*

Ans. $(a)\quad f^i_{,i} = 0.\quad (c)\quad \int_\sigma n_i f^i \, d\sigma = 0.$

Ex. (5.4) Find the divergence of the vector whose components in a spherical coordinate system are

$$f^i = \left(\frac{1}{3} r \tan\varphi, k, 1 \right),$$

where k *is a constant.*

Ans. $f^i_{,i} = 0.$

Ex. (5.5) Find the divergence of the vectors:

(a) $f^i = (y, x, 0)$ *(Cartesian coordinates);*

(b) $f^i = (0, a, b)$, *where* a *and* b *are constants (cylindrical coordinates);*

(c) $f^i = (0, r, b)$, *where* b *is constant (cylindrical coordinates);*

(d) $f^i = (a, b - x, c)$, *(Cartesian coordinates);*

(e) $f^i = (r, 0, 0)$, *(spherical coordinates);*

(f) $f^i = (a, b/r, c)$, *where* a, b, c *are constants (cylindrical coordinates).*

Ans. (a) *0 ;* (b) *0 ;* (c) *0 ;* (d) *0 ;* (e) *3 ;* (f) a/r.

Ex. (5.6) The force of gravitation, f^i, *derives from matter. Hence the divergence of gravitational force must be a measure of the quantity of matter per unit volume, or mass density* ρ. *(a) Express this fact, therefore, in the form*

$$f^i_{,i} = k\rho,$$

which is satisfied by the gravitational force, where k is constant. (b) Show that the gravitational potential thus satisfies the equation

$$g^{ij}\,\mathrm{V}_{,ij} = \nabla^2 \mathrm{V} = -\,k\,\rho\,.$$

The customary form for the constant k is $4\,\pi\,\mathrm{G}$, where G is the ***constant of gravitation****. This equation is known as* ***Poisson's equation****.*

In addition to the examples thus far given, yet another vector of particular interest may be generated from the derivatives of a given vector. It is the **curl** of f_i, defined to be

(5.6) $$\mathrm{R}^i = -\,\varepsilon^{ijk} f_{j,k}\,.$$

We may express it quite simply just by writing out its components. Thus

$$\mathrm{R}^i = -\,\varepsilon^{ijk} f_{j,k} = \frac{1}{\sqrt{g}}\left(f_{2,3} - f_{3,2}\,, f_{3,1} - f_{1,3}\,, f_{1,2} - f_{2,1}\right).$$

Consider the first component. It is

$$\mathrm{R}^1 = \frac{1}{\sqrt{g}}\left(-f_{2,3} + f_{3,2}\right) = \frac{1}{\sqrt{g}}\left\{-\left[\frac{\partial f_2}{\partial x^3} - \left\{ {k \atop 2\;3} \right\} f_k\right] + \left[\frac{\partial f_3}{\partial x^2} - \left\{ {k \atop 3\;2} \right\} f_k\right]\right\}$$

$$= \frac{1}{\sqrt{g}}\left(\frac{\partial f_3}{\partial x^2} - \frac{\partial f_2}{\partial x^3}\right).$$

A similar computation for the remaining two components shows that

(5.7) $$\mathrm{R}^i = \frac{1}{\sqrt{g}}\left(\frac{\partial f_3}{\partial x^2} - \frac{\partial f_2}{\partial x^3}\,, \frac{\partial f_1}{\partial x^3} - \frac{\partial f_3}{\partial x^1}\,, \frac{\partial f_2}{\partial x^1} - \frac{\partial f_1}{\partial x^2}\right).$$

This is usually the most convenient form in which to calculate the curl of any covariant vector. The curl of any contravariant vector is to be found by lowering the index first, then proceeding as above.

The curl of a vector figures in the statement of another integral theorem, **Stokes' theorem**. Again, we state it first in Cartesian coordinates and from this derive a statement in any coordinate system. The theorem says that **if σ is a portion of a surface enclosed by a contour γ (see Fig. 87) and if f_1, f_2 and f_3 are three continuous functions having partial derivatives of the first order over σ, then**

$$\int_\gamma \left(f_1\,dx + f_2\,dy + f_3\,dz\right) = \int_\sigma \left\{n_1\left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}\right) + n_2\left(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}\right) + n_3\left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right)\right\} d\sigma\,,$$

where n_1, n_2 and n_3 are the direction cosines of the normal to σ and the line integral on the left is taken entirely around γ. Proof of this theorem, like that of Green's theorem, will be found in any text on advanced calculus.

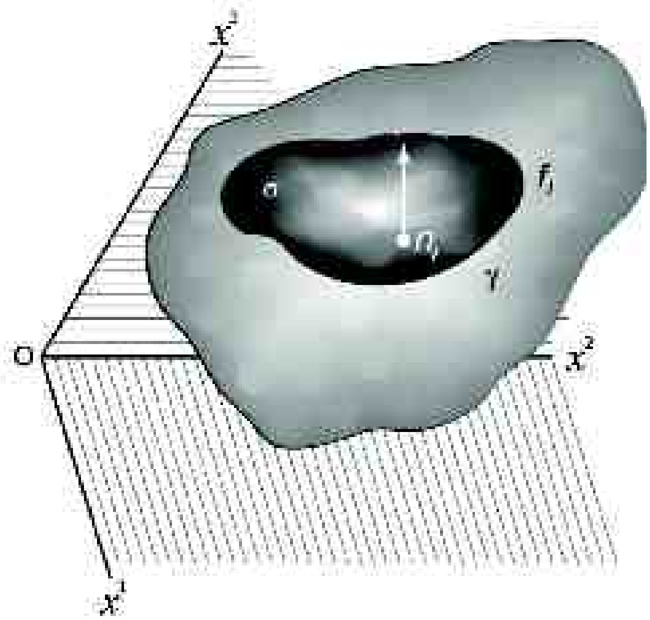

Figure 87

Again we proceed by identifying the integrands of the preceding equation as invariants. If $f_i = (f_1, f_2, f_3)$, then the integrand on the left hand side is the invariant

$$f_1\,dx + f_2\,dy + f_3\,dz = f_1\,dx^1 + f_2\,dx^2 + f_3\,dx^3 = f_i\,dx^i .$$

Reference to equation (5.7), together with the fact that in Cartesian coordinates $\sqrt{g} = 1$, allows us to write the integrand on the right side as

$$n_1\left(\frac{\partial f_3}{\partial x^2} - \frac{\partial f_2}{\partial x^3}\right) + n_2\left(\frac{\partial f_1}{\partial x^3} - \frac{\partial f_3}{\partial x^1}\right) + n_3\left(\frac{\partial f_2}{\partial x^1} - \frac{\partial f_1}{\partial x^2}\right)$$

$$= n_1 R^1 + n_2 R^2 + n_3 R^3 = n_i R^i .$$

Therefore the statement of Stokes' Theorem may be written in the generalized form

(5.8) $$\int_\gamma f_i\,dx^i = \int_\sigma \left(-\varepsilon^{ijk} f_{j,k}\right) n_i\,d\sigma .$$

The integral on the left is sometimes called the **circulation** of f^i about the curve γ.

Note that if f_j is the gradient of an invariant V, then the curl of f_j is zero, for

$$f_j = \frac{\partial V}{\partial x^j},$$

and by equation (5.7)

$$\frac{\partial f_j}{\partial x^k} - \frac{\partial f_k}{\partial x^j} = \frac{\partial^2 V}{\partial x^k \partial x^j} - \frac{\partial^2 V}{\partial x^j \partial x^k} = 0 .$$

Conversely, when the curl of a vector is everywhere zero, the vector must be expressible as the gradient of some invariant V. For then, by Stokes' Theorem,

$$\int_\gamma f_i \, dx^i = 0$$

for any curve γ. This means that the integrand is a perfect differential; that is,

$$f_i \, dx^i = dV ,$$

whence it follows by partial differentiation that

$$f_i = \frac{\partial V}{\partial x^i} .$$

Such a vector is called **irrotational**.

Ex. (5.7) Find the curl of the vectors of Ex. (5.5).

Ans. (a) $(0, 0, 0)$; (b) $(0, 0, 2a)$; (c) $(0, 0, 3r)$;
(d) $(0, 0, -1)$; (e) $(0, 0, 0)$; (f) $(0, 0, b/r)$.

Ex. (5.8) If $A^i = -\,\text{curl}\, W_j = \varepsilon^{ijk} w_{j,k}$, find $-\,\text{curl}\, A_l = -\,\text{curl}\, g_{li} A^i$.

Ans.

$$-\,\text{curl}\, A_l = \varepsilon^{mln}\left(g_{il} A^i\right)_{,n}$$

$$= \varepsilon^{mln}\left[g_{li}\left(\varepsilon^{ijk} W_{j,k}\right)_{,n}\right] = \varepsilon^{mln}\left[\varepsilon_{ljp}\, g^{pk} W^j_{,kn}\right]$$

$$= -\,\delta^{lmn}_{ljp}\, g^{pk} W^j_{,kn} = -\,\delta^{mn}_{jp}\, g^{pk} W^j_{,kn} = -\left(g^{kn} W^m_{,kn} - g^{mk} W^n_{,kn}\right) \quad \text{or}$$

$$\varepsilon^{mln} A_{l,n} = g^{mk}\left(W^n_{,n}\right)_{,n} - g^{kn} W^m_{,kn} .$$

Ex. (5.9) Show that the divergence of A^i *of Ex. (5.8) vanishes identically. Hence show that the divergence of the curl of any vector* W_j *is zero.*

We are now able to derive a result of great interest and usefulness. To do so, we need to make use of the fact* that the components of any continuous and differentiable vector field F^i may be expressed uniquely as the Laplacian of the components of a suitably constructed vector field W^i ; that is, we may take

$$F^i = g^{mn} W^i_{,mn} .$$

Now from Ex. (5.9),

$$g^{mn} W^i_{,mn} = g^{ij}\left(W^n_{,n}\right)_{,j} - \varepsilon^{iln} A_{l,n} .$$

Hence if we set $W^n_{,n} = - V$, we see that

(5.9) $$F^i = - g^{ij} V_{,j} - \varepsilon^{iln} A_{l,n} .$$

The scalar and vector potentials in terms of which F^i is expressed are not unique. We have thus shown that a suitably continuous and differentiable vector F^i may be expressed as the sum of the gradient of some invariant and the curl of some vector. The invariant V is the **scalar potential** and the vector A^i is the **vector potential**.

We see that an irrotational vector is one for which the second term on the right hand side of equation (5.9) vanishes. A non-zero vector F^i for which the first term on the right hand side of equation vanishes is called a **solenoidal vector**. By Ex. (5.9), *the divergence of a solenoidal vector is zero*. This result is the counterpart of the fact that *the curl of an irrotational vector is zero*.

Ex. (5.10) Show that a vector of the form $F^i = \varepsilon^{ijk} \varphi_{,j} \psi_{,k}$ *is solenoidal, where* φ *and* ψ *are invariants. (Hint: what is the divergence of* F^i*?)*

Ex. (5.11) Show that the vector potential for the vector F^i *of Ex. (5.10) may be taken to be* $A_i = \varphi \psi_{,i}$.

Ex. (5.12) Show that if a volume V *enclosed by a surface* S *contains a surface* σ *on which the vector* F^i *has a jump discontinuity, then*

$$\int_S n_r F^r dS + \int_\sigma \left\{\left(n_r F^r\right)_{(1)} + \left(n_r F^r\right)_{(2)}\right\} d\sigma = \int_V F^r_{,r} dV ,$$

where the parenthesized subscripts (1) and (2) denote the two sides of the surface.

Ex. (5.13) Show that the vector potential is indeterminate to the extent of an arbitrary irrotational vector.

---

*We will take this result of the theory of functions of real variables without proof. It is a familiar theorem from potential theory. A fuller discussion of it may be found in Phillips' "Vector Analysis" (Wiley, 1933), Ch. VIII, for example.

6. The Equations of Motion of a Continuous Medium

The laws of mechanics of continua are fundamentally the same as those of the mechanics of a particle. For continua, however, the particle becomes a small mass element, with the difference that the elements are contiguous and may interact at their common boundaries.

Consider an element of volume $d\tau$ in a continuous medium. If it contains an amount of matter dm, the matter of the medium is said to have a local density

$$\rho = \frac{dm}{d\tau} . \tag{6.1}$$

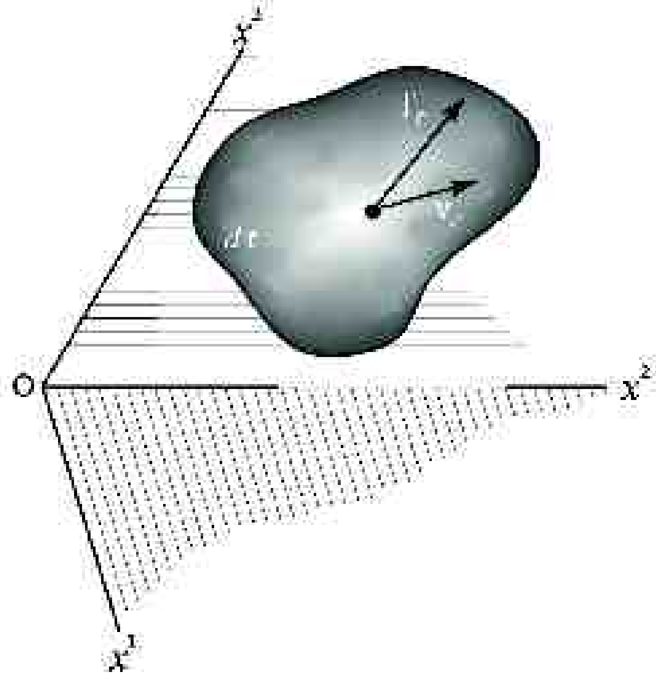

Figure 88

According to Newton's second law of motion, the acceleration a^s produced in such an element will be the result of the action of a force of amount

$$a^s \, dm = \rho \, a^s \, d\tau .$$

Such forces may originate either from a **force field** (gravitational, electrical, magnetic, etc.) or from the application of forces to the boundary surface of the medium which, by transmission throughout the medium, may be present as **internal forces**. Field force exists even in the absence of the medium; internal forces require the presence of the medium.

Suppose that the field produces an acceleration f^s, determinable from the nature of the field and the medium. Its contribution to the net force acting upon an element of the medium will be

$$f^s \, dm = \rho f^s \, d\tau .$$

The internal forces, which provide the remainder of the acceleration, must be considered in greater detail. To this end, consider a surface element $d\sigma$ bounding $d\tau$. Across this surface, the element $d\tau$ interacts with adjacent elements. In general, the force per unit area, which we denote by T^s, exerted by an adjacent element on $d\tau$, need not be in the direction of the normal ν^r to $d\sigma$; for example, the motion of one sticky surface across another calls into play such forces as drag, parallel to the surfaces. Geometrically, this means that the force at any point, represented by a directed line segment, need not lie along the unit normal to the surface. However, since both force and normal can be represented as straight line segments issuing from a common point, one may be transformed into the other by a linear transformation which will effect a rotation and a change of scale of the proper amount. In other words, we may be assured that there is a set of quantities E^s_l such that

$$\mathrm{T}^s = \mathrm{E}^s_l \nu^l . \tag{6.2}$$

The quantities E^s_l are the components of the **stress tensor**.

The stress tensor may be simply interpreted. Suppose, for example, that we choose a Cartesian coordinate system and that we select an orientation of the surface element $d\sigma$ such that the unit normal is parallel to the x^1– axis; the components of this unit normal will then be (1 , 0 , 0). In this case, therefore, we will have

$$\mathrm{T}^1 = \mathrm{E}^1_1 ,\ \mathrm{T}^2 = \mathrm{E}^2_1 ,\ \mathrm{T}^3 = \mathrm{E}^3_1 .$$

Evidently, the force per unit area has a component of amount E^1_1 along the unit normal as well as components E^2_1 and E^3_1 in the tangent plane of the surface element. Analogous interpretations may be given the remaining components of the stress tensor.

We must now somehow combine the effects of the volume and surface forces through the medium. To do this, let p_s be any parallel vector field, for which

$$p_{s,k} = 0 .$$

Then the projections on p_s of the volume and surface forces of the medium, summed over an arbitrary volume enclosed by an arbitrary surface σ will be

$$\int_\tau \rho\, a^s p_s\, d\tau = \int_\tau \rho f^s p_s\, d\tau + \int_\sigma \mathrm{T}^s p_s\, d\sigma .$$

Expressing T^s in terms of the stress tensor, this becomes

$$\int_\tau \left(\rho\, a^s - \rho f^s\right) p_s\, d\tau = \int_\sigma \left(\mathrm{E}^s_l \nu^l\right) p_s\, d\sigma .$$

We next transform the surface integral on the right hand side by applying Green's theorem,

$$\int_\sigma \left(\mathrm{E}^s_l \nu^l\right) p_s\, d\sigma = \int_\sigma \left(\mathrm{E}^{sl} p_s\right) \nu_l\, d\sigma = \int_\tau \left(\mathrm{E}^{sl} p_s\right)_{,l} d\tau$$

$$= \int_\tau \mathrm{E}^{sl}_{,l} p_s\, d\tau + \int_\tau \mathrm{E}^{sl} p_{s,l}\, d\tau = \int_\tau (\mathrm{E}^{lk} \mathrm{E}^s_{k,l} p_s\, d\tau ,$$

since by hypothesis $p_{s,l} = 0$. Using this transformation, the equation of motion requires that

$$\int_\tau \left(\rho\, a^s - \rho f^s - g^{lk} E^s_{k,l}\right) p_s \, d\tau = 0$$

for any parallel vector field p_s or volume τ. This can be true only if the quantity in parentheses vanishes, i. e., if

(6.3)
$$\rho\, a^s = \rho f^s + g^{lk} E^s_{k,l} .$$

This is the equation of motion in a continuous medium.

If, as is ordinarily the case, the material of the medium is conserved, this circumstance requires that, by Green's Theorem, the flux of matter $\rho\, v^i$ have a divergence which equals the rate of accumulation (positive or negative) of the matter. Thus, within a volume τ, the total amount of matter is changing at a rate of

$$\frac{dM}{dt} = \int_\tau \frac{\partial\rho}{\partial t} d\tau = -\int_\sigma \left(\rho\, v^i\right) n_i \, d\sigma = -\int_\tau \left(\rho\, v^i\right)_{,i} d\tau .$$

(The negative sign arises since n_i is the *outward* normal, whence a positive flux will reduce the amount of matter within τ.) Hence

$$\int \left\{ \frac{\partial\rho}{\partial t} + \left(\rho\, v^i\right)_{,i} \right\} d\tau = 0$$

for any volume τ. Therefore it must be true that

(6.4)
$$\frac{\partial\rho}{\partial t} + \left(\rho\, v^i\right)_{,i} = 0 .$$

This is the **equation of continuity** for the medium.

Ex. (6.1) Show that the equation of continuity may be written also in the form

$$\frac{d\rho}{dt} + \rho\, v^i_{,i} = 0 .$$

Thus show that an incompressible fluid ($v^i_{,i} = 0$ *) has constant density.*

Ex. (6.2) Write out the equation of continuity in (a) Cartesian coordinates, (b) spherical coordinates, (c) cylindrical coordinates.

Ans.

$$(a)\ \frac{d\rho}{dt} + \rho\left(\frac{\partial v^1}{\partial x} + \frac{\partial v^2}{\partial y} + \frac{\partial v^3}{\partial z}\right) = 0\,;$$

$$(b)\ \frac{\partial\rho}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\rho\, v^1\right) + \frac{\partial\left(\rho\, v^2\right)}{\partial\theta} + \frac{1}{\cos\varphi}\frac{\partial}{\partial\varphi}\left(\rho\, v^3\cos\varphi\right) = 0\,;$$

$$(c)\ \frac{\partial\rho}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}\left(r\,\rho\, v^1\right) + \frac{\partial\left(\rho\, v^2\right)}{\partial\theta} + \frac{\partial\left(\rho\, v^3\right)}{\partial z} = 0\,.$$

7. Isotropic and Viscous Media

Consider the case when the stress is isotropic. According to our previous definition of isotropy, this requires that

(7.1) $$E^s_l = -p\,\delta^s_l\,,$$

where p is an invariant. The meaning of p may be inferred from the fact that

$$T^s = E^s_l\, v^l = -p\,\delta^s_l\, v^l = -p\, v^s\,,$$

whence

$$|T^s| = |-p\, v^s| = |-p|\,.$$

In other words, the force per unit area, T^s, the same in all directions, has a magnitude whose value is p. Therefore p is called the **pressure** within the medium. We may also see that from equation (7.1) the value of p is given by

(7.2) $$-p\,\delta^s_s = -3p = E^s_s\,.$$

By this relation, pressure may be defined even in anisotropic media.

A fluid in which the stress is isotropic is called a **perfect fluid**. Combining equations (6.3) and (7.1), we obtain

(7.3) $$\rho\, a^s = \rho f^s - g^{lk}\left(p\,\delta^s_l\right)_{,k} = \rho f^s - g^{sk}\frac{\partial p}{\partial x^k}$$

as the equation of motion of a perfect fluid.

Now suppose that the fluid is not perfect; it is then said to be **viscous**. For a viscous fluid we may therefore write

(7.4) $$E^s_l + p\,\delta^s_l = E'^s_l \neq 0 \quad\text{or}\quad E'_{sl} = E_{sl} + p\, g_{sl}\,.$$

The tensor E'_{sl} is called the **viscosity tensor**. Clearly, the viscosity measures the extent to which the stress differs from a pure pressure. Alternatively, a fluid which is not perfect must support a strain. Now by equations (3.8.3) and (3.8.4) a strain is described by a rate of strain tensor

(7.5) $$e_{rs} = \frac{1}{2}\left(b_{rs} + b_{sr}\right)$$

where, from equation (3.8.2), $b_{rs} = \mathrm{v}_{r,s}$. Therefore

(7.6) $$e_{rs} = \frac{1}{2}\left(\mathrm{v}_{r,s} + \mathrm{v}_{s,r}\right)$$

is called the **velocity strain tensor**. It is clearly symmetric in r and s.

Any relation between viscosity and strain must be such that the viscosity vanishes with the strain. This condition can be secured most simply by requiring the viscosity tensor to be a linear function of the rate of strain e_{rs}. Thus, we get

(7.7) $$\mathrm{E}'_{rs} = \gamma_{rs}^{mn} e_{mn} = \gamma_{rs}^{mn} \mathrm{v}_{m,n},$$

where γ_{rs}^{mn} is a tensor of fourth order, symmetric in m and n. The components of γ_{rs}^{mn} are called the **viscosity coefficients**. With these, equation (6.3) takes the form

(7.8) $$\begin{aligned}\rho\, a_s &= \rho f_s - \frac{\partial p}{\partial x^s} + g^{lk}\left(\gamma_{sl}^{mn} \mathrm{v}_{m,n}\right)_{,k} \\ &= \rho f_s - \frac{\partial p}{\partial x^s} + g^{lk}\left[\gamma_{sl,k}^{mn} \mathrm{v}_{m,n} + \gamma_{sl}^{mn} \mathrm{v}_{m,nk}\right]\end{aligned}$$

as the equation of motion of a viscous fluid. If the field is homogeneous, then γ_{sl}^{mn} will constitute a parallel tensor field when both E'_{rs} and e_{rs} do so. The condition therefore is

$$\gamma_{rs,k}^{mn} = 0.$$

A circumstance of special interest is the case of isotropic viscosity. By our definition of isotropy, this means that the coefficient of viscosity tensor must be expressible in terms of the fundamental tensors g_{rs} and δ_r^s, together with whatever invariants may be needed. Remembering that γ_{rs}^{mn} is symmetric in m and n, we see that the most general isotropic coefficient of viscosity tensor symmetric in m and n will have the form

$$\gamma_{rs}^{mn} = \lambda g^{mn} + \mu\left(\delta_r^m \delta_s^n + \delta_r^n \delta_s^m\right),$$

where λ and μ are invariants.* From equation (7.7) we thus have for the isotropic viscosity tensor

(7.9) $$\begin{aligned}\mathrm{E}'_{rs} &= \left[\lambda g_{rs} g^{mn} + \mu\left(\delta_r^m \delta_s^n + \delta_r^n \delta_s^m\right)\right]\mathrm{v}_{m,n} \\ &= \lambda\left(g^{mn}\mathrm{v}_{m,n}\right)g_{rs} + \mu\left(\mathrm{v}_{r,s} + \mathrm{v}_{s,r}\right) = \mathrm{E}_{rs} + p\, g_{rs}.\end{aligned}$$

* Note that γ_{rs}^{mn} is now also symmetric in r and s.

The equation of motion of an isotropic viscous fluid is therefore

$$\rho\, a_s \;=\; \rho f_s \;+\; \frac{\partial}{\partial x^s}\Big[\lambda\, \mathrm{v}^m_{,m} \;-\; p\Big] \;+\; g^{rk}\Big[\mu\left(\mathrm{v}_{r,s} \;+\; \mathrm{v}_{s,r}\right)\Big]_{,k}. \tag{7.10}$$

This equation may be simplified somewhat. By our definition of pressure, it is true that

$$\mathrm{E}^s_l \;+\; p\,\delta^s_l \;=\; \mathrm{E}'^s_l,$$

whence

$$\mathrm{E}^s_s \;+\; 3p \;=\; \mathrm{E}'^s_s \;=\; g^{sl}\mathrm{E}'_{sl} \;=\; 0\,.$$

Therefore

$$g^{sl}\mathrm{E}'_{sl} \;=\; g^{sl}\Big[\lambda\, g_{sl}g^{mn} \;+\; \mu\left(\delta^m_s\,\delta^n_l \;+\; \delta^n_s\,\delta^m_l\right)\Big]\mathrm{v}_{m,n} \;=\; 0$$

for arbitrary v_m. This can be true only if

$$g^{sl}\Big[\lambda\, g_{sl}g^{mn} \;+\; \mu\left(\delta^m_s\,\delta^n_l \;+\; \delta^n_s\,\delta^m_l\right)\Big] \;=\; 3\lambda\, g^{mn} \;+\; \mu\left(g^{mn} \;+\; g^{nm}\right)$$

$$=\; (3\lambda \;+\; 2\mu)\,g^{mn} \;=\; 0\,.$$

Consequently, the invariants λ and μ must satisfy the condition that

$$3\lambda \;+\; 2\mu \;=\; 0\,.$$

The equation (7.10) then becomes

$$\rho\, a_s \;-\; \rho f_s \;=\; -\,\frac{\partial}{\partial x^s}\left[p \;+\; \frac{2}{3}\,\mu\,\mathrm{v}^m_{,m}\right] \;+\; g^{rk}\Big[\mu\left(\mathrm{v}_{r,s} \;+\; \mathrm{v}_{s,r}\right)\Big]_{,k}. \tag{7.11}$$

When μ is constant, and when $\mathrm{v}_{r,sk} \;=\; \mathrm{v}_{r,ks}$, as is true in any flat space, the last term on the right hand side may be written as

$$g^{rk}\Big[\mu\left(\mathrm{v}_{r,s} \;+\; \mathrm{v}_{s,r}\right)\Big]_{,k} \;=\; \mu\, g^{rk}\mathrm{v}_{r,sk} \;+\; \mu\, g^{rk}\mathrm{v}_{s,rk}$$

$$=\; \mu\left(g^{rk}\mathrm{v}_{r,k}\right)_{,s} \;+\; \mu g^{rk}\mathrm{v}_{s,rk} \;=\; \mu\left(\mathrm{v}^k_{,k}\right)_{,s} \;+\; \mu\, g^{rk}\mathrm{v}_{s,rk}\,.$$

Therefore equation (7.11) becomes

$$\rho\, a_s \;-\; \rho f_s \;=\; -\,\frac{\partial p}{\partial x^s} \;+\; \frac{1}{3}\,\mu\,\frac{\partial \mathrm{v}^m_{,m}}{\partial x^s} \;+\; \mu\, g^{rk}\mathrm{v}_{s,rk}\,. \tag{7.12}$$

Ex. (7.1) Consider a star in static equilibrium and in which the force field is derivable from a gravitational potential V*. Since it is static, the velocity strain tensor vanishes. Write down the equation of motion in spherical coordinates for the case when the potential* V is spherically symmetric.

Ans. $\rho\,\dfrac{d\mathrm{V}}{dr} \;=\; \dfrac{dp}{dr}\,.$

Ex. (7.2) Consider a rotating star of non-viscous fluid in which the force field is derivable from a combined gravitational and rotational potential

$$\mathrm{V} \; - \; \frac{1}{2} r^2 \left(\cos^2 \varphi \right) \omega^2 \, ,$$

where ω *is a constant angular velocity in a polar coordinate system in which* $\varphi = 90°$ *is the axis of rotation. What are the equations of motion?*

Ans.

$$\rho \frac{\partial \mathrm{V}}{\partial r} = \frac{\partial p}{\partial r} + r\,\omega^2 \cos^2 \varphi \, ,$$

$$\rho \frac{\partial \mathrm{V}}{\partial \varphi} = \frac{\partial p}{\partial \varphi} - r^2 \omega^2 \sin \varphi \cos \varphi \, .$$

Ex. (7.3) From equations (3.8.3) and (4.7.5) show that at any point in a fluid we may define a ***local angular velocity tensor***

$$\omega_{rs} = \frac{1}{2} \delta^{pq}_{rs} \, \mathrm{v}_{p,q} \, . \tag{7.13}$$

From this, show that the ***local angular velocity vector*** *is*

$$\omega^m = \frac{1}{2} \varepsilon^{mpq} \, \mathrm{v}_{p,q} \, . \tag{7.14}$$

This is sometimes also known as the ***vortex vector*** *or vorticity vector. The curves which satisfy the differential equation*

$$\lambda^r = \frac{dx^r}{ds} = \frac{\omega^r}{\omega} \, , \; \left(\omega = \left| \omega^i \right| \right)$$

are called ***vortex lines****.*

Ex. (7.4) From equation (2.14) and our previous definition of solenoidal vectors, we see that the angular velocity vector is solenoidal. Show, therefore, that $\omega^m = 0$ *characterizes an irrotational velocity field, hence derivable from a* ***velocity potential****.*

8. The Equations of Electromagnetism in Empty Space

It is well known that certain simple processes such as rubbing glass with silk can be made to alter the conditions of these materials in that they will attract small light objects such as bits of paper. Bodies whose condition has been so altered are said to have been **electrified**.

If a body has been electrified, its electrification may be transferred wholly or in part to other bodies by placing it in contact with them. By so doing, all bodies may thus be classified as **dielectrics** (**insulators**) or **conductors** according as the electrification of them in empty space may be arbitrarily localized or not.

The electrification of a body is attributed to the presence of **electric charge** upon its surface. Simple experiments such as placing charged bodies in one another's presence show that electric charges are of only two kinds, *positive* and *negative*. The same experiments establish that like charges repel and unlike charges attract.

If an electrically charged body is brought into the presence of a theretofore unelectrified body, the latter will be found to have acquired a charge distribution on or within it. These are called **induced charges**. It is by this process of induction that glass rubbed with silk is able to attract light objects nearby.

In principle, induction creates a difficulty for the investigation of charge distributions, for the presence of electric charges can be known only by the forces they produce upon other charges (**test charges**), and the latter modify the distribution of the former by the addition of induced charges. The perturbation of a charge distribution may be made arbitrarily small, however, by making the test charge arbitrarily small. All assertions about charge distributions imply the use of arbitrarily small test charges.

From the complete symmetry of a sphere and the fact that conductors are specifically characterized by the fact that electric charges cannot be localized on or in them, we conclude that in otherwise empty space the charge on the surface of a conducting sphere must be distributed uniformly. We may thus define a point charge as the limit approached by a conducting sphere as its volume approaches zero. Test charges are ideally point charges of infinitesimal strength. Charge distributions may be idealized as distributions of point charges over surfaces, through volumes, or at isolated points.

Since electric charges on any body give rise to forces on the body when in the presence of other charged bodies, the displacement of any charge can be effected only by the performance of work. The amount of work done in an infinitesimal displacement dx^r of a charge on which forces F^r act will by definition be

$$-\,d\mathrm{V} = \mathrm{F}_r\,dx^r = g_{rs}\,\mathrm{F}^r\,dx^s . \tag{8.1}$$

It is an experimental fact that *in the presence of a static charge distribution, the work done by a charged body in traversing any closed circuit in empty space is zero*. This implies that $d\mathrm{V}$ is the perfect differential of a single-valued function $\mathrm{V}(x^r)$. Hence

$$\mathrm{F}_r = -\frac{\partial \mathrm{V}}{\partial x^r} . \tag{8.2}$$

The quantity V is called the **electrostatic potential** of the given charge distribution. Like other forms of energy, electrostatic potentials of distinct charge distributions are additive.

Two problems now present themselves: (1) to find the form of the potential function V; and (2) to define quantity of charge by prescribing a method of measuring it. Now it can be shown experimentally that *within a closed conducting surface there are no electric forces due to charges on the surface,* for the work done by displacing a charge within the conducting surface is found to be independent of the surface charges. In particular, if no charges are present within the surface, no electrostatic forces will exist in it, no matter what distribution of charges exists on the enclosing surface.

This result suffices to fix partially the form of the potential function V. We note first that any charge on the conducting surface will distribute itself in such a way under the action of its own electrical forces that the net electrical force vanishes everywhere in the interior of the surface. The fact that the *net* force of surface charges is everywhere zero on a small test charge within the closed conducting surface does not mean that the force between it and each charged surface element is zero but only that the sum of the forces of all such elements together is zero. Since the force exerted by any surface element upon the test charge is equal and opposite to the force exerted by the test charge upon the surface element, we may find the total upon the test charge by summing the force over the conducting surface. This latter force is normal to the surface, else a redistribution of charge would take place. Hence the magnitude of the force at any point of the surface is $F_r n^r$, where n^r is the unit normal to the surface, and the net force over the surface is

$$\int_S F_r n^r dS = 0 .$$

By Green's Theorem, this implies that

$$\int_S F_r n^r dS = \int_V F^r_{,r} dV = 0$$

for any conducting surface S bounding the volume V. Therefore we must have

$$F^r_{,r} = g^{rs} F_{r,s} = 0 . \tag{8.3}$$

Expressing F^r in terms of the potential, this gives

$$g^{rs} V_{,rs} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^r}\left(\sqrt{g}\, g^{rs} \frac{\partial V}{\partial x^s} \right) = 0 .$$

That is, V satisfies **Laplace's equation** in empty space. In spherical coordinates, with radius r, longitude θ, latitude $\varphi = \sin^{-1}\mu$, this becomes

$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r} \right) + \frac{\partial}{\partial \mu}\left[\left(1 - \mu^2\right) \frac{\partial V}{\partial \mu} \right] + \frac{1}{1 - \mu^2} \frac{\partial^2 V}{\partial \theta^2} = 0 . \tag{8.4}$$

Inasmuch as (1) any charge distribution may be constructed from a distribution of point charges, (2) the potential functions of distinct charges are additive, and (3) Laplace's equation is linear in V, it will serve to determine the electrostatic potential of all charge distributions in otherwise empty space if we determine the potential of a single point charge. To this end, we continuously deform the surface S into a sphere with center at the origin, taken to be the position of the test charge. The purpose in so doing is to render all surface charges equivalent because of the sphere's symmetry about its center. Since this process in no way affects the preceding argument, the electrostatic potential must still satisfy Laplace's equation. Hence, by imposing spherical symmetry upon V — i.e., by making it independent of θ and μ — we find that the potential function of a point charge must satisfy the equation

(8.5)
$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r} \right) = 0\,, \quad - F_1 = \frac{\partial V}{\partial r} = \pm \frac{\text{constant}}{r^2}\,,$$
$$V - V_0 = \pm \frac{\text{constant}}{r}\,.$$

The arbitrary additive constant V_0 is taken to be zero so that V will vanish as r approaches infinity. The preceding equation expresses **Coulomb's law**.

Let us now turn to the problem of measuring the quantity of charge. We wish to define charge in terms of mechanically measurable quantities. For a point charge, this may evidently be done by computing the flux of force over any circumscribing surface not containing other charges, for since the very presence of the electric charge is known only by the force it exerts on other charges and since that force varies as the inverse square of the distance between point charges, the quantity

$$q = k' \int_S F_r n^r dS$$

is an invariant of the charge, k' being a proportionality factor dependent upon the choice of units.

Since the relation between two point charges is entirely reciprocal and since either may be used to measure the charge on the other, we must in fact require the flux of force to be proportional to the charge of each and therefore proportional to both. In other words, the *product* of the charges q and q' is given as

(8.6)
$$q\,q' = k \int_S F_r n^r dS\,,$$

In the **rational** system of units (or **Heaviside-Lorentz** units), $k = 1$. **Electrostatic units** are defined when unit charges $q = 1 = q'$ separated by unit distance ($r = 1$) produce unit force. Therefore

(8.7)
$$\frac{1}{k_{(s)}}\, q_{(s)}\, q'_{(s)} = \frac{1}{k_{(s)}} \times 1 \times 1 = \int_S F_r n^r dS = 4\pi = \frac{1}{k_{(s)}}\cdot$$

Electromagnetic units are defined by the relation $q_{(s)} = c \times q_{(m)}$, whence

(8.8)
$$k_{(m)} = \frac{q_{(s)}\, q'_{(s)}}{q_{(m)}\, q'_{(m)}} = \frac{1}{4\pi \times c^2}\,.$$

Having prescribed a purely mechanical method for measuring electric point charges, and having agreed upon a system of units, let us consider an arbitrary volume distribution of point charges. The quantity of charge within a surface S will be

$$q = \int_S F_r n^r dS$$

if we use as a test body a unit charge $(q' = 1)$. At the same time, if point charges are distributed with density σ through the volume V within S, then

$$q = \int_V \sigma \, dV .$$

By equating the two expressions for q and applying Green's Theorem, it follows that

$$\text{(8.9)} \qquad F^r_{,r} = g^{rs} F_{r,s} = \sigma .$$

We have tacitly assumed that the electric point charges are distributed through otherwise empty space. In this case, F_r represents the force per unit charge, a quantity defined to be the **electric field strength**, E_r. Equation (8.6) thus becomes

$$\text{(8.10)} \qquad E^r_{,r} = \sigma ,$$

one of the fundamental equations of electromagnetism. Note that essential modifications of the equations of electrostatics are necessary for charges distributed through material media.

Ex. (8.1) A ***line of force*** *is defined to be a curve whose unit tangent has the direction of the electric force vector, hence*

$$\frac{dx^r}{ds} = \frac{E^r}{|E^s|} .$$

Show that the lines of force are orthogonal to the ***equipotential surfaces*** $V = \mathit{constant}$.

Static distributions of electric charge are, as such, of little interest, for their static character guarantees that nothing happens to them. On the other hand, charges in motion display entirely new features in their modes of interaction. In particular, moving charges exert on other differently moving charges forces which are not electrostatic in character — i. e., not described by Coulomb's law. This is the substance of the discoveries made by Oersted in 1820 and Faraday in 1831. The new forces are of a type designated as **magnetic forces**. The simultaneous association of electric and magnetic forces with moving charges permits us to refer to them jointly as **electromagnetic forces**. It is with these that the laws of electromagnetism deal.

To introduce the magnetic forces, we compare the net force upon a moving charge of amount q with that upon an equal charge at rest at the same time and place. We find that the former, F^r, differs from the latter, $q\,E^r$; that is,

$$F^r - q\,E^r \neq 0 .$$

The difference is, in fact, proportional to q and to the magnitude of the velocity. Moreover, it is perpendicular to the velocity vector and, as variation of the velocity vector will show, is at the same time proportional to the sine of the angle between the velocity vector and a certain definite direction. Hence there must exist some vector H^r such that if v^s is the velocity of the charge, then

$$\mathrm{F}^r - q\,\mathrm{E}^r = \frac{q}{c}\,\varepsilon^{rst}\mathrm{v}_s\mathrm{H}_t$$

or, in a form known as the **Lorentz equation**

$$\mathrm{F}^r = q\left(\mathrm{E}^r + \frac{1}{c}\,\varepsilon^{rst}\mathrm{v}_s\mathrm{H}_t\right), \tag{8.11}$$

where c is the velocity of light in empty space (and hence a universal dimensional constant). The inclusion of the factor $1/c$ renders v^s/c dimensionless; in Cartesian coordinates, the vector H^t therefore has the same dimensions as E^t and equation (8.8) defines the **magnetic intensity vector** H^t *in electromagnetic units*. Henceforth, we regard E^t as defining the electric field and H^t the magnetic field in empty space. The term "field" indicates tacitly that the vectors E^t and H^t are determinate at any and every point which a test charge may occupy. Their components are to be found in the manner described and thus may be functions of position. This is the only sense in which the term "field" has a meaning. The theory of electromagnetism will therefore serve as a prototypical field theory.

Ex. (8.2) Show that in general coordinate systems, the vectors E^t *and* H^t *differ in dimensions by the factor* $1/\sqrt{g}$.

Ex. (8.3) Show that equation (8.9) implies that

$$g^{rs}\mathrm{V}_{,rs} = -\,\sigma\,. \tag{8.12}$$

Ex. (8.4) Show that the electrostatic potential represents the work done on a unit charge brought to any given point from infinity. Hence show that the work done upon charge of density σ *at a point where the electrostatic potential is* V *must be* $(1/2)\,\sigma\,\mathrm{V}$. *(Hint: bear in mind that if the charge is everywhere assembled by bringing increments of charge from infinity until they attain a density* σ *and create a potential* V*, the* ***average*** *charge during the process is only half the final value. This accounts for the factor* $1/2$.*)*

Ex. (8.5) Show that the energy of the entire electrostatic field is

$$\mathrm{W} = \frac{1}{2}\int_\tau \sigma\,\mathrm{V}\,d\tau$$

if there are no surface discontinuities. Here $d\tau$ *is a volume element and* τ *is all of space.*

Ex. (8.6) Show that the total energy of the electrostatic field is

$$\mathrm{W} = \frac{1}{2}\int_\tau g_{mn}\,\mathrm{E}^m\mathrm{E}^n\,d\tau\,.$$

Hence show that the energy density of the electrostatic field is

$$(1/2)\, g_{mn}\, \mathrm{E}^m \mathrm{E}^n .$$

Ans. From Ex. (8.5) and Ex. (8.3) we have

$$\mathrm{W} = -\int_\tau \mathrm{V}\, g^{rs} \mathrm{V}_{,rs}\, d\tau .$$

However, from Green's Theorem and Ex. (5.6) we have

$$-\int_\tau \mathrm{V}\, g^{rs} \mathrm{V}_{,rs}\, d\tau = \int_\tau g^{rs} \mathrm{V}_{,r} \mathrm{V}_{,s}\, d\tau - \int_\mathrm{S} \mathrm{V}\, \mathrm{V}_{,r}\, n^r\, d\mathrm{S} .$$

Since τ is to include the whole of space, the bounding surface S recedes to the sphere at infinity. The last integral therefore vanishes. Hence the required result follows.

The magnetic intensity vector, H^t, **though in some respects similar to the** electric intensity vector E^t, differs from it in a fundamental respect in that its existence cannot be traced to a "magnetic charge" which might have been expected to provide a counterpart to the electric charges q which give rise to the electric intensity vector. This fact may also be expressed by saying that the density of "magnetic charge" is identically zero everywhere. The magnetic intensity vector will therefore satisfy an equation analogous to equation (8.7) with the difference that we must set $\sigma = 0$. In other words,

(8.13) $$\mathrm{H}^r_{,r} = 0$$

is a second of the fundamental equations of electromagnetism.

To sum up, we have defined a measure of electric charge, stated Coulomb's law of electrostatic interaction, and, with the Lorentz equation, defined the magnetic intensity vector. The conservation of electric charge is expressed in equation (8.7) whereas the non-existence of "magnetic charge" is implied by equation (8.9). Let us now state the first law of electromagnetism, **Faraday's law**, which says that **the electromotive force induced in a circuit is proportional to the rate of decrease of the flux of the magnetic field through the circuit.** To see what this means, let C represent a closed curve and S that portion of surface bounded by C. By electromotive force is meant the integral

$$\int_\mathrm{C} \mathrm{E}_r\, dx^r .$$

It is clearly not a "force", but the work done on a unit charge by transporting it once around the curve C. The flux of the magnetic field is, according to the usual definition of any flux,

$$\int_\mathrm{S} \mathrm{H}_r\, n^r\, d\mathrm{S} ,$$

where n^r is the unit normal to S at any point. By Faraday's law, we then have between electromotive force and magnetic flux the relation

(8.14) $$\int_C \mathrm{E}_r\, dx^r = -\frac{1}{c}\frac{\partial}{\partial t}\int_S \mathrm{H}_r n^r\, d\mathrm{S}\,.$$

The constant c again occurs to preserve the relation of dimensions established in the Lorentz equation (8.8).

We now apply Stokes' theorem to equation (8.10), giving

$$\int\left\{-\,\varepsilon^{rst}\mathrm{E}_{s,t} + \frac{1}{c}\frac{\partial \mathrm{H}^r}{\partial t}\right\} n_r\, d\mathrm{S} = 0$$

for all surfaces S. Therefore it must be true that

(8.15) $$-\,\varepsilon^{rst}\mathrm{E}_{s,t} + \frac{1}{c}\frac{\partial \mathrm{H}^r}{\partial t} = 0\,.$$

This is **Faraday's law** in differential form.

A second fundamental law of electromagnetism is **Ampere's law**. It states that **the integral of the magnetic intensity vector around a closed circuit is proportional to the flux of electric current through the circuit**. In other words, in analogy to equation (8.10), we have

(8.16) $$\int_C \mathrm{H}_r\, dx^r = \frac{1}{c}\int_S \mathrm{C}^r n_r\, d\mathrm{S}\,,$$

where C^r is the **current vector**. It consists in part of a bodily transport of electric charge. This, the **convection current**, is $\sigma\,\mathrm{v}^r$, where v^r is the velocity of the charge of space density σ. More than this, the equation

$$q' = \int_S \mathrm{E}^r n_r\, d\mathrm{S}\,,$$

where S is a completely closed surface, suggests by differentiation with respect to time that a change in field strength through S should also be considered a current. In particular, since

$$\frac{\partial q'}{\partial t} = \int_S \frac{\partial \mathrm{E}^r}{\partial t} n_r\, d\mathrm{S}\,,$$

the current vector ought to be defined to be

$$\mathrm{C}^r = \sigma\,\mathrm{v}^r + \frac{\partial \mathrm{E}^r}{\partial t}\,.$$

Making use of this definition in equation (8.13) and applying Stokes' theorem to it gives

$$\int\left(-\,\varepsilon^{rst}\mathrm{H}_{s,t} - \frac{1}{c}\left[\frac{\partial \mathrm{E}^r}{\partial t} + \sigma\,\mathrm{v}^r\right]\right) n_r\, d\mathrm{S} = 0$$

for any surface S. Therefore it must be true that

(8.17) $$\varepsilon^{rst}\,\mathrm{H}_{s,t} + \frac{1}{c}\frac{\partial \mathrm{E}^{r}}{\partial t} = -\frac{\sigma}{c}\,\mathrm{v}^{r}\,.$$

This is an expression of **Ampere's law** in differential equation form. Together, equations (8.8), (8.10), (8.12) and (8.14) constitute a celebrated set of differential equations known as **Maxwell's equations** for free space.

Let us consider some alternate forms in which the equations of electromagnetism may be put. From equation (8.10), we see that $\mathrm{H}^{r}_{,r} = 0$. Therefore, by §5, there must be a vector A_s such that

$$\mathrm{H}^{m} = \varepsilon^{mns}\,\mathrm{A}_{s,n}\,.$$

Combining this with equation (8.12),

(8.18) $$-\,\varepsilon^{mns}\,\mathrm{E}_{s,n} + \frac{1}{c}\frac{\partial \mathrm{H}^{m}}{\partial t} = 0\,,$$

which is in all cases true; hence we get the result that

$$\varepsilon^{mns}\left(\mathrm{E}_s + \frac{1}{c}\frac{\partial \mathrm{A}_s}{\partial t}\right)_{,n} = 0\,.$$

As we have seen, it is therefore possible, with no loss of generality, to assume that

$$\mathrm{E}_s + \frac{1}{c}\frac{\partial \mathrm{A}_s}{\partial t} = -\,\varphi_{,s} = -\,\frac{\partial \varphi}{\partial x^{s}}$$

for some invariant function φ. Combining this result with equations (8.14) and (8.8), we have

(8.19) $$\varepsilon^{mns}\left(g_{sp}\,\mathrm{A}_{r,q}\right)_{,n} + \frac{1}{c^{2}}\frac{\partial^{2}\mathrm{A}^{m}}{\partial t^{2}} + \frac{1}{c}\frac{\partial}{\partial t}\left[g^{mn}\frac{\partial \varphi}{\partial x^{n}}\right] = \frac{\sigma}{c}\,\mathrm{v}^{m}$$

and

(8.20) $$\frac{1}{c}\frac{\partial \mathrm{A}^{m}_{,m}}{\partial t} + g^{mn}\,\varphi_{,mn} = \sigma\,.$$

However, we make use of the fact that

$$g_{sp}\,\varepsilon^{pqr} = g^{qu}\,g^{rv}\,\varepsilon_{suv}\,,$$

$$\varepsilon^{mns}\,g_{sp}\,\varepsilon^{pqr} = g^{qu}\,g^{rv}\left(\varepsilon^{mns}\,\varepsilon_{suv}\right)$$

$$= \delta^{mn}_{uv}\,g^{qu}\,g^{rv} = g^{qm}\,g^{rn} - g^{qn}\,g^{rm}\,.$$

Hence $$\varepsilon^{mns}\left(g_{sp}\,\varepsilon^{pqr}\,\mathrm{A}_{r,qn}\right) = g^{mq}\left(\mathrm{A}^r_{,r}\right)_{,q} - g^{nq}\,\mathrm{A}^m_{,nq}\,,$$

so that

$$-\,g^{nq}\,\mathrm{A}^m_{,nq} + \frac{1}{c^2}\frac{\partial^2\mathrm{A}^m}{\partial t^2} + g^{mq}\left(\mathrm{A}^r_{,r} + \frac{1}{c}\frac{\partial\varphi}{\partial t}\right) = \frac{\sigma}{c}\,\mathrm{v}^m\,.$$

Now, since A^m and φ are not wholly determined by the conditions thus far imposed upon them, we are at liberty to require also that

(8.21) $$\mathrm{A}^r_{,r} + \frac{1}{c}\frac{\partial\varphi}{\partial t} = 0\,,$$

a relation known as the **Lorentz condition**; convenience, not necessity, recommends this condition. The motive for imposing the Lorentz condition is clearly that it simplifies the preceding equation to the form

$$-\,g^{nq}\,\mathrm{A}^m_{,nq} + \frac{1}{c^2}\frac{\partial^2\mathrm{A}^m}{\partial t^2} = \frac{\sigma}{c}\,v^m\,.$$

At the same time, the equation (8.18) becomes

(8.22) $$-\,g^{nq}\,\varphi_{,nq} + \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = \sigma\,.$$

The last two equations are of a form known as a classical **wave equation**, for the solution of the homogeneous equation (i. e., when $\sigma = 0$), may be of the form of the superposition of plane waves.

(Note that equation (8.21) is actually a more stringent condition than is needed ; we require only that

$$g_{mq}\left(\mathrm{A}^r_{,r} + \frac{1}{c}\frac{\partial\varphi}{\partial t}\right)_{,q} = 0\,.$$

Hence equation (8.21) is sufficient, not necessary.)

The vector A^m and scalar φ are called the **general electrodynamic potentials**. From them, we can determine the electromagnetic field as

(8.23) $$\begin{cases} \mathrm{E}_s = -\,\varphi_{,s} - \dfrac{1}{c}\dfrac{\partial\mathrm{A}_s}{\partial t}\,, \\[2ex] \mathrm{H}^m = \varepsilon^{mns}\,\mathrm{A}_{s,n}\,. \end{cases}$$

We have seen that the electrodynamic potentials satisfy a wave equation. From this fact and their relations to E_s and H^m we can show also that the electric and magnetic fields must satisfy a wave equation as well. Thus, differentiating the wave equation of A_m with respect to t and taking the covariant derivative of both sides of the wave equation for φ , we have

(8.24)
$$\frac{1}{c}\left[g^{nq}\frac{\partial \mathrm{A}_{m,nq}}{\partial t} - \frac{1}{c^2}\frac{\partial}{\partial t}\left[\frac{\partial^2 \mathrm{A}_m}{\partial t^2}\right]\right] = -\frac{1}{c}\frac{\partial}{\partial t}\left(\sigma\, \mathrm{v}_m\right),$$
$$g^{nq}\varphi_{,nqm} - \frac{1}{c^2}\frac{\partial^2 \varphi_{,m}}{\partial t^2} = -\sigma_{,m}.$$

Adding and reversing the order of differentiation (the geometry is assumed Euclidean),

$$g^{nq}\left[\varphi_{,m} + \frac{1}{c}\frac{\partial \mathrm{A}_m}{\partial t}\right]_{,nq} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left[\varphi_{,m} + \frac{1}{c}\frac{\partial \mathrm{A}_m}{\partial t}\right] = -\sigma_{,m} - \frac{1}{c}\frac{\partial}{\partial t}\left(\sigma\, \mathrm{v}_m\right).$$

The quantities in square brackets are evidently $-\mathrm{E}_m$, whence

(8.25)
$$-g^{nq}\mathrm{E}_{m,nq} + \frac{1}{c^2}\frac{\partial^2 \mathrm{E}_m}{\partial t^2} = -\sigma_{,m} - \frac{1}{c}\frac{\partial}{\partial t}\left(\sigma\, \mathrm{v}_m\right).$$

This is the **wave equation** for E_m (or E^m). An analogous derivation yields a wave equation for H^m.

9. The Equations of Electromagnetism in Matter

We have thus far tacitly assumed that the electric charges under consider-ation have been isolated in otherwise empty space. We wish now to consider the form taken by the equations of electromagnetism when matter is present.

In a dielectric, electric charge is not freely mobile (detached) but is nonetheless partially mobile to the extent that an electric field will polarize the matter of the dielectric. Since the polarization vanishes with the electric field, we may conveniently take it to be associated with a field of the form

(9.1)
$$\mathrm{P}^r = \eta^r_s \mathrm{E}^s,$$

where η^r_s is the **electric susceptibility tensor**, having all components zero for empty space.

The “source” of the polarization vector is a polarization charge which, recalling the interpretation of the divergence, may be calculated as

$$\sigma^* = \mathrm{P}^r_{,r}.$$

The source of the electric field is still a charge of density

$$\sigma_{(0)} = \mathrm{E}^r_{,r}.$$

Therefore the total charge density, the sum of the two, must be

$$\sigma = \sigma_{(0)} + \sigma^* = \mathrm{E}^r_{,r} + \mathrm{P}^r_{,r} = \left(\mathrm{E}^r + \mathrm{P}^r\right)_{,r}.$$

This suggests defining a new vector

(9.2)
$$\mathrm{D}^r = \mathrm{E}^r + \mathrm{P}^r = \left(\delta^r_s + \eta^r_s\right)\mathrm{E}^s = \varepsilon^r_s \mathrm{E}^s,$$

the **displacement vector**. The equation

(9.3) $$D^r_{,r} = \sigma$$

then supersedes equation (8.7) within matter. The tensor ε^r_s is called the **dielectric tensor**. It remains true, however, that the electric field E^r determines the electric force upon a unit charge.

The case with the magnetic forces is very similar. In matter, the magnetic intensity vector induces a field

(9.4) $$B^r = \mu^r_s H^s .$$

The vector B^r is the **magnetic induction vector**, the tensor μ^r_s is the **magnetic permeability tensor**. The Lorentz force equation is now

(9.5) $$F^r = q\left(E^r + \varepsilon^{rst} v_s B_t\right).$$

This takes the form identical with that for empty space when $\mu^r_s = \delta^r_s$, as in the absence of matter.

It is now the case that

(9.6) $$B^r_{,r} = 0 ,$$

the magnetic analogue of equation (9.3). Setting

$$\mu^r_s = \delta^r_s + \lambda^r_s ,$$

where λ^r_s is the **magnetic susceptibility tensor** (the analogue of the electric susceptibility tensor), we define a vector I^r such that

(9.7) $$B^r = H^r + I^r , \; I^r = \lambda^r_s H^s .$$

The vector I^r is the **magnetization vector**, analogous to the electric polarization vector.

We note that equation (9.6) expresses the non-existence of independent “magnetic charge”. However, combined with equation (9.7) it implies only that

$$H^r_{,r} + I^r_{,r} = 0 ,$$

whence

$$H^r_{,r} = - I^r_{,r} = \kappa ,$$

not necessarily zero. There may also be some “permanent” magnetization, represented by a vector $I^r_{(0)}$.

Turning now to the remaining electromagnetic equations in matter, we find that Faraday’s law may be taken over merely by substituting the magnetic induction for the magnetic intensity. Thus, equation (8.11) becomes

(9.8) $$- \varepsilon^{rst} E_{s,t} + \frac{1}{c}\frac{\partial B^r}{\partial t} = 0 .$$

The expression of Ampere's law calls for two changes in the current vector. First, because the electric charge density is related to the electric induction D^r, we replace E^r by D^r. Second, in conducting matter there may exist a conduction current induced by the electric field, hence of the form

$$(9.9) \qquad i^r = \kappa^r_s E^s ,$$

where κ^r_s is the **conductivity tensor** of the medium. Equation (9.9) expresses **Ohm's law**. Therefore equation (8.13) becomes

$$(9.10) \qquad - \varepsilon^{rst} H_{s,t} - \frac{1}{c}\frac{\partial D^r}{\partial t} = \frac{1}{c}\left(\sigma v^r + \kappa^r_s E^s\right).$$

We now have the equations of electromagnetism in a form suitable for use in matter, reducible to those for empty space when matter is absent.

Ex. (9.1) Let $H^m_{,m} = - I^m_{,m} = \kappa$ *be the* ***density of magnetization****. It can then be shown in potential theory that* H_r *is the gradient of a potential function* Ω *which is given by*

$$\Omega = \frac{1}{4\pi}\int_V \frac{\kappa(x^i)}{r}\, dV ,$$

where r *is the distance from any point* P *to any point* x^i *at which the density of magnetization is* κ*, and* V *is the whole of space. In other words,*

$$H_r = \frac{\partial \Omega}{\partial x^r} .$$

Show that

$$\int_V B^r H_r \, dV = 0 .$$

(Hint: apply Green's theorem to the function $F_r = \Omega B_r$ *and use equation (9.6) and the fact that*

$$\int_S F_r v^r \, dS \rightarrow 0$$

as S *approaches the sphere at infinity.)*

Ex. (9.2) Since $dW = H_r dx^r = - \dfrac{\partial \Omega}{\partial x^r} dx^r = - d\Omega$ *is the work done on a unit magnetic pole through a displacement* dx^r*, we may regard* $-\Omega(P)$ *as the work done in bringing a unit magnetic pole from infinity to any point* P*. Hence show, as in Ex. (8.6), that the total energy in the magnetic field is*

$$W' = \frac{1}{2}\int_V \mu_{rs} H^r H^s \, dV .$$

Ex. (9.3) Show that the energy of the electric field is

$$W'' = \frac{1}{2}\int_V \varepsilon_{rs} E^r E^s \, dV .$$

Ex. (9.4) Show that the total energy of the electric and magnetic fields is

$$W = W' + W'' = \frac{1}{2}\int_V \left(\varepsilon_{rs} E^r E^s + \mu_{rs} H^r H^s\right) dV .$$

Hence show that the energy density of the field is

$$u = \frac{1}{2}\left(\varepsilon_{rs} E^r E^s + \mu_{rs} H^r H^s\right).$$

Ex. (9.5) If the dielectric tensor and the magnetic permeability tensor do not depend upon the time, show that

$$\frac{\partial}{\partial t}\int_V u \, dV = c\int_S \varepsilon^{rst} H_s E_t v_r \, dS - \int_V \left(\kappa^{rs} E_r E_s + \sigma v^r E_r\right) dV ,$$

where S *is a fixed surface enclosing the volume* V. *Hence show that the electromagnetic energy flux vector over* S *is*

$$S^r = c\,\varepsilon^{rst} H_s E_t . \tag{9.11}$$

This is known as the ***Poynting vector***.

Ans. We start with the fact that

$$\frac{\partial}{\partial t}\int_V u \, dV = \frac{\partial}{\partial t}\left[\frac{1}{2}\int_V \left(\varepsilon_{rs} E^r E^s + \mu_{rs} H^r H^s\right) dV\right]$$

$$= \int_V \left(E_s \frac{\partial D^s}{\partial t} + H^s \frac{\partial B^s}{\partial t} \right) dV .$$

From equations (9.10) and (9.8) we obtain by substitution

$$\frac{\partial}{\partial t}\int_V u \, dV = \int_V \left[E_s\left(- c\,\varepsilon^{spq} H_{p,q} - \sigma v^s - \kappa^{sp} E_p\right) + H_s\left(c\,\varepsilon^{spq} E_{p,q}\right)\right] dV .$$

Hence

$$\frac{\partial}{\partial t}\int_V u \, dV = \int_V c\,\varepsilon^{spq}\left(- E_s H_{p,q} + H_s E_{p,q}\right) dV$$
$$- \int_V \left(\sigma v^s + \kappa^{sp} E_p\right) E_s \, dV .$$

But

$$\boldsymbol{\varepsilon}^{spq}\left(-\mathrm{E}_s\mathrm{H}_{p,q}+\mathrm{H}_s\mathrm{E}_{p,q}\right)=\boldsymbol{\varepsilon}^{spq}\left(\mathrm{E}_p\mathrm{H}_{s,q}+\mathrm{E}_{p,q}\mathrm{H}_s\right)=\boldsymbol{\varepsilon}^{spq}\left(\mathrm{E}_p\mathrm{H}_s\right)_{,q}.$$

Therefore the first integral on the right can be transformed into a surface integral by Green's theorem, whence the required result follows.

10. Curves and Surfaces in Three Dimensions

We have until now considered surfaces to be two-dimensional spaces complete in themselves, quite independent of whether or not they are sub-spaces of manifolds of higher dimensions. It is of interest, however, to consider surfaces as sub-spaces of two dimensions embedded in three-dimensional space. We have, in fact, encountered such surfaces in the guise of coordinate surfaces — manifolds defined by assigning a constant value to any one of the three coordinates. We may assume, without loss of generality, that any surface of interest which does not intersect itself and which is continuous with a well defined tangent everywhere, may be taken to be a coordinate surface. Two other families of coordinate surfaces will then define coordinates over the given surface.

Let S be a continuous, smooth surface which does not intersect itself. Choose a coordinate system such that S is the surface $x^3 = c = \text{constant}$. Then in any other system of coordinates $\bar{x}^i$ we have

$$\bar{x}^i = \bar{x}^i(x^1, x^2, c) = \bar{x}^i(x^1, x^2)$$

and

$$x^1 = x^1(\bar{x}^i),\ x^2 = x^2(\bar{x}^i),\ x^3 = c.$$

Hence the surface is characterized by either of the conditions

$$x^3 = c \quad \text{or} \quad dx^3 = \left[\left(\frac{\partial x^3}{\partial \bar{x}^i}\right)_{x^3 = c}\right] d\bar{x}^i = 0.$$

Any displacement *in the surface* will therefore be one for which $dx^3 = 0$. Similarly, any contravariant vector tangent to the surface (or "lying in the surface") will be characterized by the condition $\mathrm{v}^3 = 0$, whence the condition

$$\mathrm{v}^3 = \frac{\partial x^3}{\partial \bar{x}^i}\bar{\mathrm{v}}^i = 0.$$

Such vectors are called **surface vectors** in S.

Consider a particular type of transformation of coordinates of the form

(10.1) $$\bar{x}^1 = \bar{x}^1(x^1, x^2),\ \bar{x}^2 = \bar{x}^2(x^1, x^2),\ \bar{x}^3 = x^3.$$

This transformation is a special kind known as a **sub-tensor transformation**. It is also a surface transformation, for the surface is characterized both by the condition $\bar{x}^3 = c$ and the condition $x^3 = c$; its equation is therefore not affected by the transformation, though coordinates on the surface are transformed. More generally, one could allow $\bar{x}^3 = \bar{x}^3(x^3)$; the surface is then given either by the condition $x^3 = c$ or the condition $\bar{x}^3 = \bar{x}^3(c) = \bar{c}$.

Consider what effect a surface transformation has upon a contravariant vector v^i. It becomes

$$\bar{\mathrm{v}}^i = \frac{\partial \bar{x}^i}{\partial x^j}\mathrm{v}^j = \frac{\partial \bar{x}^i}{\partial x^1}\mathrm{v}^1 + \frac{\partial \bar{x}^i}{\partial x^2}\mathrm{v}^2 + \frac{\partial \bar{x}^i}{\partial x^3}\mathrm{v}^3$$

or, written out in full,

$$\bar{\mathrm{v}}^1 = \frac{\partial \bar{x}^1}{\partial x^1}\mathrm{v}^1 + \frac{\partial \bar{x}^1}{\partial x^2}\mathrm{v}^2,\ \bar{\mathrm{v}}^2 = \frac{\partial \bar{x}^2}{\partial x^1}\mathrm{v}^1 + \frac{\partial \bar{x}^2}{\partial x^2},\ \bar{\mathrm{v}}^3 = \mathrm{v}^3 .$$

Thus, a surface transformation transforms an arbitrary contravariant vector as though (1) its components in the surface $(\mathrm{v}^1, \mathrm{v}^2)$ were a vector in two dimensions, and (2) its component out of the surface v^3 were an invariant.

In similar fashion, we see that the inverse of the surface transformation (assumed to be reversible) is likewise a surface transformation, namely

$$x^1 = x^1(\bar{x}^1, \bar{x}^2),\ x^2(\bar{x}^1, \bar{x}^2),\ x^3 = \bar{x}^3 .$$

The vector $\bar{\mathrm{v}}^i$ thus becomes v^i once again upon applying the inverse transformation. Moreover, we see that the transformation for a covariant vector v_i under a surface transformation yields an analogous result, namely

$$\bar{\mathrm{v}}_1 = \frac{\partial x^1}{\partial \bar{x}^1}\mathrm{v}_1 + \frac{\partial x^2}{\partial \bar{x}^1}\mathrm{v}_2,\ \bar{\mathrm{v}}_2 = \frac{\partial x^1}{\partial \bar{x}^2}\mathrm{v}_1 + \frac{\partial x^2}{\partial \bar{x}^2}\mathrm{v}_2,\ \bar{\mathrm{v}}_3 = \mathrm{v}_3 .$$

Thus, as before, a surface transformation transforms the first two components as though they formed a vector in the two-dimensional (x^1, x^2)– space and as though the third component were an invariant. Indeed, we can speak of v^3 or v_3 as a **surface invariant** because its value is not changed under a surface transformation. In particular, if $\mathrm{v}^3 = 0$, then $\bar{\mathrm{v}}^3 = 0$ so that any surface vector is transformed into another surface vector by a surface transformation.

With these results in mind, we can foresee the effect of a surface transformation upon a tensor of any order and type. Recalling that every tensor may be expressed as a sum of products of vectors of the appropriate type, we see that under a surface transformation the new components will have been transformed in a fashion determined by those indices which are *not* equal to 3. Thus, T^i_j transforms as $u^i\mathrm{v}_j$, which is to say as $u^i\mathrm{v}_j = \mathrm{T}^i_j\ (i, j \neq 3)$, $\mathrm{T}^i_3 = u^i\mathrm{v}_3\ (i \neq 3)$, $\mathrm{T}^3_j = u^3\mathrm{v}_j\ (j \neq 3)$ and the invariant $\varphi = \mathrm{T}^3_3 = u^3\mathrm{v}_3$.

We may represent this as

$$T^i_j = \left\| \begin{array}{cc|c} T^1_1 & T^1_2 & T^1_3 \\ \underline{T^2_1} & \underline{T^2_2} & \underline{T^2_3} \\ T^3_1 & T^3_2 & T^3_3 \end{array} \right\|.$$

The set of components

$$T^\alpha_\beta = \left\| \begin{array}{cc} T^1_1 & T^1_2 \\ T^2_1 & T^2_2 \end{array} \right\|, \ (\alpha, \beta = 1, 2)$$

transforms as a mixed tensor of the second order in two dimensions; the components $T^\alpha_3 = \left(T^1_3, T^2_3\right)$ transform as a contravariant vector in two dimensions; the components $T^3_\beta = \left(T^3_1, T^3_2\right)$ transform as a covariant vector in two dimensions; and the component T^3_3 transforms as an invariant under a surface transformation. In general, the several components of any tensor transform under a surface transformation as would a tensor of such order and type as results from suppressing all indices equal to 3. Henceforth we will distinguish indices not equal to 3 by Greek letters, as α and β.

Thus far, it has not been explicitly apparent what consequences flow from the existence of a third dimension — an out-of-surface dimension. One way of dramatizing the difference between a surface in two dimensions and a surface embedded in three dimensions is to note that the latter possesses a normal vector, something not defined for the former. Let n_i be the covariant form of the unit normal. Then, since it must be perpendicular to *every* surface vector, and since the latter must be of the form $v^i = \left(v^1, v^2, 0\right)$, it follows that n_i must have the form

$$n_i = (0, 0, n_3),$$

and since it is to be a unit vector, $n_3 = 1/\sqrt{g^{33}}$. The contravariant form of n_i will evidently be

$$n^i = \frac{1}{\sqrt{g^{33}}}\left(g^{13}, g^{23}, g^{33}\right).$$

Ex. (10.1) Consider the coordinate system which is related to spherical coordinates ($\theta = longitude$, $\varphi = latitude$*) by the transformations*

$$\bar{x}^1 = \theta = x^1, \ \bar{x}^2 = \varphi = x^2$$

$$\bar{x}^3 = \rho = r\left[1 - e^2\cos^2\varphi\right]^{1/2} = x^3\left[1 - e^2\cos^2(x^2)\right]^{1/2}$$

whose inverses are

$$x^1 = \theta = \bar{x}^1, \ x^2 = \varphi = \bar{x}^2, \ r = \frac{\rho}{\left[1 - e^2\cos^2\varphi\right]^{1/2}} = \frac{\bar{x}^3}{\left[1 - e^2\cos^2(x^2)\right]^{1/2}}.$$

The surfaces $\rho = a\sqrt{1 - e^2} = constant$ *are spheroids whose axial cross -sections are ellipses of major semi-axis a and eccentricity* e. *Find the covariant and contravariant forms of the unit normal to the* x^3 *– surface. (Hint: compare with Ch.2, Ex. (5.1).)*

Ans. Since

$$\bar{g}_{ij} = \left\| \begin{matrix} \dfrac{\rho^2 \cos^2 \varphi}{1 - e^2 \cos^2 \varphi} & 0 & 0 \\ 0 & \dfrac{\rho^2 [1 - (2e^2 - e^4)\cos^2 \varphi]}{(1 - e^2 \cos^2 \varphi)^3} & -\dfrac{e^2 \rho \cos \varphi \sin \varphi}{(1 - e^2 \cos^2 \varphi)^2} \\ 0 & -\dfrac{e^2 \rho \cos \varphi \sin \varphi}{(1 - e^2 \cos^2 \varphi)^2} & \dfrac{1}{1 - e^2 \cos^2 \varphi} \end{matrix} \right\|$$

and

$$\bar{g}^{ij} = \left\| \begin{matrix} \dfrac{1 - e^2 \cos^2 \varphi}{\rho^2 \cos^2 \varphi} & 0 & 0 \\ 0 & \dfrac{1 - e^2 \cos^2 \varphi}{\rho^2} & \dfrac{e^2 \sin \varphi \cos \varphi}{\rho} \\ 0 & \dfrac{e^2 \sin \varphi \cos \varphi}{\rho} & \dfrac{1 - (2e^2 - e^4)\cos^2 \varphi}{1 - e^2 \cos^2 \varphi} \end{matrix} \right\| ,$$

it follows that

$$n_i = \left(0, 0, \left[\frac{1 - e^2 \cos^2 \varphi}{1 - (2e^2 - e^4)\cos^2 \varphi} \right]^{1/2} \right)$$

and

$$\bar{n}^i = \left[\frac{1 - e^2 \cos^2 \varphi}{1 - \left(2e^2 - e^4\right)\cos^2 \varphi} \right]^{1/2} \left(0, \frac{e^2 \cos \varphi \sin\varphi}{\rho}, \frac{1 - \left(2e^2 - e^4\right)\cos^2 \varphi}{1 - e^2 \cos^2 \varphi} \right) .$$

With the normal vector, we may devise a projection operator which will separate the surface components of any vector or tensor from the rest. To construct this operator, let us note that the vector

$$p_i = \left(\mathrm{v}_j n^j\right) n_i = \left(\mathrm{v}^j n_j\right) n_i$$

is one whose magnitude is $|\mathrm{v}^j| \cos \theta$, where θ is the angle between the direction of v^j and that of n^j, and whose direction is the direction of n^i, normal to the x^3 – surface.

Therefore the tensor $N_i^j = n^j n_i$ is an operator which, when applied to any vector v_j gives the portion of v_j in the direction of the normal n_i. It has the property, characteristic of projection operators, that applying it two (or more) times gives exactly the same result as applying it once. That is,

$$N_i^j\left(N_j^k v_k\right) = \left(n^i n_i\right)\left(n^k n_j\right)v_k = \left(n^j n_j\right)\left(n^k n_i\right)v_k = n^k n_i v_k = N_i^k v_k .$$

The operator (tensor) N_i^k is therefore said to be **idempotent**. With the normal component of a vector v_i isolated in this way, the remainder of the vector must lie in the surface to which n^i is the normal. Thus

$$S_i = v_i - N_i^k v_k = \left(\delta_i^k - N_i^k\right)v_k = P_i^k v_k ,$$

(where $P_i^k = \left(\delta_i^k - n^k n_i\right)$) is the part of v_i which lies in the surface, i. e., the surface components of v_i . The surface and normal projection operators are clearly orthogonal, for

$$P_i^k N_k^j = \left(\delta_i^k - n^k n_i\right)n^j n_k = n^j n_i - n^j\left(n_k n^k\right)n_i = 0 .$$

Let us write out the components of P_i^k. They are

$$(10.2) \qquad P_i^k = \begin{Vmatrix} 1 & 0 & -n^1 n_3 \\ 0 & 1 & -n^2 n_3 \\ 0 & 0 & (1 - n^3 n_3) \end{Vmatrix} = \begin{Vmatrix} 1 & 0 & -n^1 n_3 \\ 0 & 1 & -n^2 n_3 \\ 0 & 0 & 0 \end{Vmatrix}$$

since $n^i n_i = n^3 n_3 = 1$. Clearly, P_i^k has no inverse (its determinant is zero); neither has N_i^k. This reflects the fact that a vector is not determined completely simply by its projection upon the x^3 – surface. This projection operator is, of necessity, also idempotent. It may be used to find the surface components of covariant or contravariant vectors. Thus

$$P_i^k v_k = \left(\delta_i^k - n^k n_i\right)v_k = v_i - \left(n^k v_k\right)n_i$$

$$= \left(v_1 , v_2 , \left[v_3 - \left(n_k v^k\right)\right]n_3\right) = \left(v_1 , v_2 , \left[v_3 - n_3 n_3 v^3\right]\right).$$

If the vector v^i is a surface vector, the component $v^3 = 0$ and the vector's surface components in covariant form are simply (v_1 , v_2 , v_3) . Otherwise, they are as given above.

Similarly, the surface components of any contravariant vector are

$$P_k^i v^k = \left(\delta_k^i - n^i n_k\right)v^k = v^i - \left(n_k v^k\right)n^i$$

$$= \left(v^1 - \left[n_3 v^3\right]n^1 , v^2 - \left[n_3 v^3\right]n^2 , v^3 - \left[n_3 v^3\right]n^3\right)$$

$$= \left(v^1 - \left[n_3 v^3\right]n^1 , v^2 - \left[n_3 v^3\right]n^2 , 0\right),$$

which is clearly a surface vector.

Ex. (10.2) Find the surface projection operator to the $\bar{x}^3$ – surface of Ex. (10.1).

Ans. From equation (10.2) and Ex. (10.1) we have

$$P_i^k = \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & -\dfrac{e^2 \cos\varphi \sin\varphi}{\rho}\left[\dfrac{\left(1 - e^2\cos^2\varphi\right)}{1 - \left(2e^2 - e^4\right)\cos^2\varphi}\right]^{1/2} \\ 0 & 0 & 0 \end{matrix} \right\| .$$

Since any tensor may be written as the sum of products of vectors of the appropriate type, we may obtain the surface components of any tensor by applying the appropriate surface projection operator to each index. For example, the fundamental tensor will have surface components

(10.3) $$a^{ij} \equiv P_k^i P_l^j g^{kl} = \left(\delta_k^i - n^i n_k\right)\left(\delta_l^j - n^j n_l\right) g^{kl}$$

$$\left(\delta_k^i - n^i n_k\right)\left(g^{kj} - n^j n^k\right) = g^{ij} - n^i n^j - n^i n^j + n^i n^j = g^{ij} - n^i n^j .$$

On the other hand, equation (10.2) shows that whenever i or j is equal to 3, the projection must be equal to zero. Hence

$$a^{ij} = \left\| \begin{matrix} a^{11} & a^{12} & 0 \\ a^{21} & a^{22} & 0 \\ 0 & 0 & 0 \end{matrix} \right\| .$$

Therefore

(10.4) $$g^{ij} = a^{ij} + n^i n^j = \left\| \begin{matrix} a^{11} + n^1 n^1 & a^{21} + n^1 n^2 & n^1 n^3 \\ a^{21} + n^2 n^1 & a^{22} + n^2 n^2 & n^2 n^3 \\ n^1 n^3 & n^2 n^3 & n^3 n^3 \end{matrix} \right\| .$$

The surface projection of the covariant fundamental tensor takes the somewhat surprising form

$$a_{ij} = P_i^k P_j^l g_{kl} = \left(\delta_l^k - n^k n_i\right)\left(\delta_j^l - n^l n_j\right) g_{kl} = \left(\delta_i^k - n^k n_i\right)\left(g_{kj} - n_k n_j\right)$$

$$= g_{ij} - n_i n_j - n_i n_j + n_i n_j = g_{ij} - n_i n_j$$

(10.5) $$= \left\| \begin{matrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} - n_3 n_3 \end{matrix} \right\| .$$

This could have been foreseen, however, simply by lowering the indices of equations (10.3) or (10.4) It is interesting to note also that

$$a_{ij}a^{jk} = \left(g_{ij} - n_i n_j\right)\left(g^{jk} - n^j n^k\right) = g_{ij}g^{jk} - n_i n^k - n_i n^k + n_i n^k$$

(10.6)
$$= \delta_i^k - n^k n_i = \begin{Vmatrix} 1 & 0 & -n^1 n_3 \\ 0 & 1 & -n^2 n_3 \\ 0 & 0 & 0 \end{Vmatrix}.$$

Thus $a_{\alpha\beta}$ and $a^{\alpha\beta}$ are mutual inverses, since $g_{\alpha\beta}g^{\beta\gamma} = \delta_\alpha^\gamma$.

Ex. (10.3) Derive the fundamental tensor for the surface of a spheroid by taking $d\rho = 0$. (Hint: see Ex. (10.1).) Then verify the components g^{ij} in Ex. (10.1) by calculating them from equation (10.4).)

Let us give application to these results by considering the Riemann-Christoffel curvature tensor. From equation (2.7.3) we have that the Riemann-Christoffel tensor is

$$\mathrm{R}_{prst} = \frac{1}{2}\left(\frac{\partial^2 g_{pt}}{\partial x^r \partial x^s} + \frac{\partial^2 g_{rs}}{\partial x^p \partial x^t} - \frac{\partial^2 g_{ps}}{\partial x^r \partial x^t} - \frac{\partial^2 g_{rt}}{\partial x^p \partial x^s}\right)$$
$$+ g^{mn}\left([rs, m][pt, n] - [rt, m][ps, n]\right).$$

Since there is nothing in the derivation of this result which limited it to two dimensions, it is true in three. In particular, when none of the free indices is equal to 3, we have

$$\mathrm{R}_{\alpha\beta\gamma\delta} = \frac{1}{2}\left(\frac{\partial^2 g_{\alpha\delta}}{\partial x^\beta \partial x^\gamma} + \frac{\partial^2 g_{\beta\gamma}}{\partial x^\alpha \partial x^\delta} - \frac{\partial^2 g_{\alpha\gamma}}{\partial x^\beta \partial x^\delta} - \frac{\partial^2 g_{\beta\delta}}{\partial x^\alpha \partial^\gamma}\right)$$
$$+ g^{mn}\left([\beta\gamma, m][\alpha\delta, n] - [\beta\delta, m][\alpha\gamma, n]\right).$$

(Note that whereas the Greek indices have the range (1, 2), the Latin indices range from 1 to 3.) Substituting from equations (10.4) and (10.5), we see that this may be written as

$$\mathrm{R}_{\alpha\beta\gamma\delta} = \frac{1}{2}\left(\frac{\partial^2 a_{\alpha\delta}}{\partial x^\beta \partial x^\gamma} + \frac{\partial^2 a_{\beta\gamma}}{\partial x^\alpha \partial x^\delta} - \frac{\partial^2 g_{\alpha\gamma}}{\partial x^\beta \partial x^\delta} - \frac{\partial^2 a_{\beta\delta}}{\partial x^\alpha \partial x^\gamma}\right)$$

$$+ a^{\mu\nu}\left([\beta\gamma, \mu][\alpha\delta, \nu] - [\beta\delta, \mu][\alpha\gamma, \nu]\right)$$

$$+ n^m n^n\left([\beta\gamma, m][\alpha\delta, n] - [\beta\delta, m][\alpha\gamma, n]\right).$$

Now the quantity in the first two lines is the Riemann-Christoffel tensor as calculated in the surface S, treated as a two-dimensional manifold independently of being the sub-space $x^3 = c$. Let us designate the surface Riemann-Christoffel tensor as $R^*_{\alpha\beta\gamma\delta}$. Then

$$R_{\alpha\beta\gamma\delta} = R^*_{\alpha\beta\gamma\delta} + \left(b_{\beta\gamma} b_{\alpha\delta} - b_{\beta\delta} b_{\alpha\gamma}\right), \tag{10.7}$$

where we have set

$$b_{\beta\gamma} = n^m [\beta\gamma, m] = n_m \left\{ \begin{matrix} m \\ \beta\ \gamma \end{matrix} \right\} = n_3 \left\{ \begin{matrix} 3 \\ \beta\ \gamma \end{matrix} \right\}. \tag{10.8}$$

As we have seen, n_3 is a surface invariant. Let us therefore examine the quantity $\left\{ \begin{matrix} 3 \\ \beta\ \gamma \end{matrix} \right\}$ to determine its nature under a surface transformation. Under any transformation we must have

$$\overline{\left\{ \begin{matrix} m \\ p\ q \end{matrix} \right\}} = \frac{\partial \bar{x}^m}{\partial x^j} \frac{\partial^2 x^j}{\partial \bar{x}^p \partial \bar{x}^q} + \left\{ \begin{matrix} j \\ k\ l \end{matrix} \right\} \frac{\partial \bar{x}^m}{\partial x^j} \frac{\partial x^k}{\partial \bar{x}^p} \frac{\partial x^l}{\partial \bar{x}^q}.$$

Therefore, under a surface transformation,

$$\overline{\left\{ \begin{matrix} 3 \\ p\ q \end{matrix} \right\}} = \delta^3_j \frac{\partial^2 x^j}{\partial \bar{x}^p \partial \bar{x}^q} + \left\{ \begin{matrix} 3 \\ k\ l \end{matrix} \right\} \frac{\partial x^k}{\partial \bar{x}^p} \frac{\partial x^l}{\partial \bar{x}^q}$$

$$= \frac{\partial^2 x^3}{\partial \bar{x}^p \partial \bar{x}^q} + \left\{ \begin{matrix} 3 \\ k\ l \end{matrix} \right\} \frac{\partial x^k}{\partial \bar{x}^p} \frac{\partial \bar{x}^l}{\partial \bar{x}^q} = \left\{ \begin{matrix} 3 \\ k\ l \end{matrix} \right\} \frac{\partial x^k}{\partial \bar{x}^p} \frac{\partial x^l}{\partial \bar{x}^q},$$

since the second partial of x^3 with respect to any pair of coordinates $\bar{x}^p$ and $\bar{x}^q$ vanishes in a surface transformation. Evidently, all such Christoffel symbols of the second kind transform as covariant tensors of the second rank under a surface transformation. Therefore $b_{\beta\gamma}$ is a doubly covariant symmetric surface tensor.

If, now, the three-dimensional space in which the surface is embedded is ordinary Euclidean space, then

$$\begin{aligned} R_{\alpha\beta\gamma\delta} &= 0 = R^*_{\alpha\beta\gamma\delta} + \left(b_{\beta\gamma} b_{\alpha\delta} - b_{\beta\delta} b_{\alpha\gamma}\right), \\ R^*_{\alpha\beta\gamma\delta} &= b_{\alpha\gamma} b_{\beta\delta} - b_{\beta\gamma} b_{\alpha\delta} = \delta^{\kappa\lambda}_{\gamma\delta} b_{\alpha\kappa} b_{\beta\lambda}. \end{aligned} \tag{10.9}$$

This relation is called the **Gauss equation of the surface**.

Ex. (10.4) Determine R_{1212} *for the surface of a spheroid. (Hint: use the equation (2.7.8) and the result of Ex. (7.2) of Chapter 2.)*

Ans. $$R^*_{1212} = \frac{\rho^2 (1 - e^2)^2 \cos^2\varphi}{\left(1 - e^2 \cos^2 \varphi\right)^2 \left[1 - \left(2e^2 - e^4\right)\cos^2 \varphi\right]}.$$

Ex. (10.5) Determine $b_{\beta\gamma}$ *for the surface of a spheroid. (Hint: use the first of equations (10.8) and the results of Ex. (10.1).)*

Ans.
$$\begin{cases} b_{11} = -\dfrac{\rho\,(1-e^2)\cos^2\varphi}{\left([1-e^2\cos^2\varphi][1-(2e^2-e^4)\cos^2\varphi]\right)^{1/2}}, \\ b_{12} = b_{21} = 0, \\ b_{22} = \dfrac{\rho\,(1-e^2)}{(1-e^2\cos^2\varphi)^{3/2}[1-(2e^2-e^4)\cos^2\varphi]^{1/2}}. \end{cases}$$

Ex. (10.6) Verify the answer to Ex. (10.4) by using Eq. (10.9).

Ex. (10.7) Show that if the three-dimensional space in which a surface is embedded is not Euclidean, then

$$R_{1212} = R^{*}_{1212} + (b_{12})^2 - b_{11}\,b_{22}\;.$$

The tensor $b_{\alpha\beta}$ as defined in equation (10.8), may be given a simple interpretation. To do so, consider the unit normal to S, the vector $n_i = (0, 0, n_3)$. Its covariant derivative with respect to the surface coordinates is

$$n_{i,\beta} = \frac{\partial n_i}{\partial x^\beta} - \left\{ {k \atop i\ \beta} \right\} n_k\,.$$

Therefore, when i is not 3,

$$n_{\alpha,\beta} = -\left\{ {3 \atop \alpha\ \beta} \right\} n_3 = -\,b_{\alpha\beta}\,.$$

Hence $b_{\alpha\beta}$ is the covariant derivative of the covariant surface components of the unit normal to the surface, taken with respect to the surface coordinates.

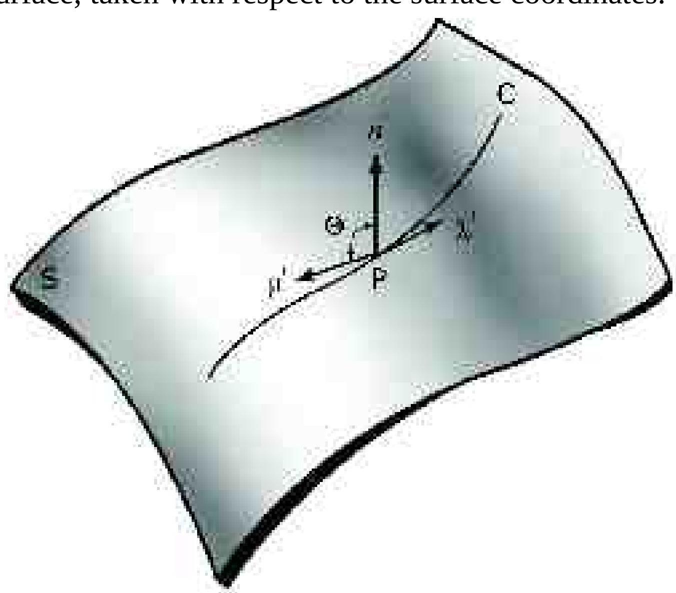

Figure 89

Now consider any curve on the surface S. Its tangent at any point is the surface vector $\lambda^i = \left(\lambda^1, \lambda^2, 0\right) = dx^i/ds$ For any curve whatsoever, though, we have found that

$$\frac{\delta \lambda^i}{\delta s} = \kappa \mu^i,$$

where κ is the principal curvature of the curve and μ^i the principal normal. Hence

$$n_i \frac{\delta \lambda^i}{\delta s} = \kappa \mu^i n_i .$$

But n^i and λ^i are by definition orthogonal, so that

$$n_i \frac{\delta\lambda^i}{\delta s} = \frac{\delta}{\delta s}\left(n_i \lambda^i\right) - \lambda^i \frac{\delta n_i}{\delta s} = 0 - \lambda^i \frac{\delta n_i}{\delta s} = - \lambda^i n_{i,j} \lambda^j .$$

However, since λ^i is a surface vector and therefore $\lambda^3 = 0$, we find that

(10.11)

$$\kappa\left(\mu^i n_i\right) = \kappa \cos\Theta = \kappa_{(n)} = n_i \frac{\delta\lambda^i}{\delta s} = - \lambda^i \frac{\delta n_i}{\delta s} = - n_{i,j} \lambda^i \lambda^j$$

$$= - n_{\alpha,\beta} \lambda^\alpha \lambda^\beta = - b_{\alpha\beta} \lambda^\alpha \lambda^\beta .$$

The quantity $\kappa_{(n)}$ is obviously the projection of $\kappa \mu^i$ upon the normal n^i; it is called the **normal curvature**. Clearly $\kappa_{(n)} \leq \kappa$. The angle Θ is the angle between n^i and μ^i. When $\Theta = 0$ or π, the curve at that point must lie in the plane of λ^i and n^i ($\equiv \pm \mu^i$); as we have seen, this is the osculating plane. Hence, all curves formed as plane sections of the surface and containing the normal, lie in the osculating planes; their torsions are therefore zero. The normal curvature to such a curve is equal to its principal curvature.

Ex. (10.7) Find the curvature of a parallel of latitude on a spheroid (such as the earth). (Hint: use the results of Ex. (10.1).)

Ans. Since $\lambda^i = \left(\dfrac{[1 - e^2 \cos^2\varphi]^{1/2}}{\rho\cos\varphi}, 0, 0 \right)$, *we find that*

$$\frac{\delta\lambda^i}{\delta s} = \left(0, \frac{\sin\varphi\,(1 - e^2\cos^2\varphi)}{\rho^2\cos\varphi}, - \frac{(1 - e^2)}{\rho} \right),$$

whence

$$\kappa = \left| \frac{\delta\lambda^i}{\delta s} \right| = \frac{[1 - e^2\cos^2\varphi]^{1/2}}{\rho\cos\varphi} .$$

Ex. (10.8) Find the angle Θ between the normal to the surface of the spheroid and the principal normal of a parallel of latitude.

Ans. From equation (10.11) and Ex. (10.4), we have

$$\kappa_{(n)} = b_{\alpha\beta}\,\lambda^\alpha\lambda^\beta = -\frac{(1-e^2)}{\rho}\left[\frac{1-e^2\cos^2\varphi}{1-(2e^2-e^4)\cos^2\varphi}\right]^{1/2}.$$

Hence from equation (10.11) and Ex. (10.7),

$$\cos\Theta = \frac{\kappa_{(n)}}{\kappa} = -\frac{(1-e^2)\cos\varphi}{\left[1-(2e^2-e^4)\cos^2\varphi\right]^{1/2}}.$$

Let us now form the orthonormal right-handed triad n^i, λ^i, and $m^i = \varepsilon^{ijk}n_j\lambda_k$. Clearly, since m^i is orthogonal to n^i, it therefore lies in the tangent plane and is a surface vector. Any vector may then be conveniently resolved into a linear combination of n^i, λ^i, and m^i. In particular,

$$\mu^i = a\,\lambda^i + b\,n^i + c\,m^i.$$

Since $\mu^i\lambda_i = 0$, it follows that $a = 0$, and since $b = \mu^i n_i = \cos\Theta$ whereas $\mu^i\mu_i = 1$,

$$\mu^i\mu_i = 1 = \cos^2\Theta + (c)^2,\ (c)^2 = 1 - \cos^2\Theta = \sin^2\Theta,\ c = \sin\Theta.$$

Hence

$$\mu^i = n^i\cos\Theta + m^i\sin\Theta,$$

(10.12)
$$\frac{\delta\lambda^i}{\delta s} = \kappa\,\mu^i = (\kappa\cos\Theta)\,n^i + (\kappa\sin\Theta)\,m^i = \kappa_{(n)}n^i + \kappa_{(g)}m^i,$$

where

(10.13)
$$\kappa_{(g)} = \kappa\sin\Theta.$$

We have previously found a geometric interpretation of the normal curvature $\kappa_{(n)}$. The quantity $\kappa_{(g)}$, the **geodesic curvature**, may also be given a geometric interpretation. Thus, when $\kappa_{(n)} = 0$, the curves which are normal sections of the surface have zero curvature and must therefore lie in a plane, the tangent plane. Then $\kappa_{(g)} = \kappa$, which is evidently the curvature of the curve in the tangent plane, or the **surface curvature**. More generally, curves on the surface which are not formed by normal sections of the surface, or which do not lie in the tangent plane, will nevertheless have curvature $\kappa_{(n)}$ normal to the surface, curvature $\kappa_{(g)}$ in the surface.

Ex. (10.9) Treating the surface of a spheroid as a space of two dimensions, find the curvature of a parallel of latitude.

Ans. From Ex. (10.1), we have

$$a_{ij} = \begin{Vmatrix} \dfrac{\rho^2 \cos^2 \varphi}{1 - e^2 \cos^2 \varphi} & 0 \\ 0 & \dfrac{\rho^2 [1 - (2e^2 - e^4) \cos^2 \varphi]}{(1 - e^2 \cos^2 \varphi)^3} \end{Vmatrix}$$

whence

$$\lambda^i = \left(\frac{[1 - e^2 \cos^2 \varphi]^{1/2}}{\rho \cos \varphi}, 0 \right),$$

$$\frac{\delta \lambda^i}{\delta s} = \left(0, \frac{\sin \varphi}{\rho^2 \cos \varphi}, \frac{(1 - e^2 \cos^2 \varphi)^2}{1 - (2e^2 - e^4) \cos^2 \varphi} \right).$$

Therefore

$$\kappa_{(g)} = \left| \frac{\delta \lambda^i}{\delta s} \right| = \frac{\sin \varphi}{\rho \cos \phi} \left[\frac{1 - e^2 \cos^2 \varphi}{1 - (2e^2 - e^4) \cos^2 \varphi} \right]^{1/2}.$$

Ex. (10.10) From Ex. (10.8) find $\sin \Theta$ *. Check your result by calculating* $\sin \Theta$ *independently from the relation*

$$\sin \Theta = \frac{\kappa_{(g)}}{\kappa}$$

(equation (10.13)).

Ans.
$$\sin \Theta = \frac{\sin \varphi}{\left[1 - (2e^2 - e^4) \cos^2 \varphi\right]^{1/2}}.$$

Ex. (10.11) Verify equation (10.12) for a parallel of latitude on the spheroid.

Since

$$\lambda^i = \left(\frac{[1 - e^2 \cos^2 \varphi]^{1/2}}{\rho \cos \varphi}, 0, 0 \right),$$

$$n^i = \left(0, \frac{e^2 \cos \varphi \sin \varphi}{\rho} \left[\frac{1 - e^2 \cos^2 \varphi}{1 - (2e^2 - e^4) \cos^2 \varphi} \right]^{1/2}, \left[\frac{1 - (2e^2 - e^4) \cos^2 \varphi}{1 - e^2 \cos^2 \varphi} \right]^{1/2} \right),$$

we find

$$m^i = \varepsilon^{ijk} n_j \lambda_k = \left(0 , \frac{(1 - e^2 \cos^2\varphi)^{3/2}}{\left[1 - (2e^2 - e^4)\cos^2\varphi\right]^{1/2}}, 0 \right) .$$

Substituting these vectors into equation (10.12), we get the same results as given by the direct calculation in Ex. (10.7).

Ex. (10.12) Show that when the curve on the surface is the x^i – curve, then

$$b_{ii} = \kappa^{(i)}_{(n)} a_{ii} .$$

(Hint: we have

$$b_{\alpha\beta} \lambda^{\alpha}_{(i)} \lambda^{\beta}_{(i)} = \kappa_{(n)} [a_{\alpha\beta} \lambda^{\alpha}_{(i)} \lambda^{\beta}_{(i)}] = \kappa_{(n)} .)$$

Still further light may be shed upon the properties of the embedded surface by considering yet another aspect of the tensor $b_{\alpha\beta}$. As we have seen, a symmetric tensor may have associated with it a pair of invariant eigenvalues and their associated eigenvectors. The eigenvalues are to be found as roots of the equation

$$(10.14) \qquad \frac{1}{a} \left| b_{\alpha\beta} - \kappa\, a_{\alpha\beta} \right| = \varepsilon^{\alpha\gamma} \varepsilon^{\beta\gamma} (b_{\alpha\beta} - \kappa\, a_{\alpha\beta}) (b_{\gamma\delta} - \kappa\, a_{\gamma\delta})$$

$$= \varepsilon^{\alpha\gamma} \varepsilon^{\beta\delta} (b_{\alpha\beta} b_{\gamma\delta} - \kappa [b_{\alpha\beta} a_{\gamma\delta} + b_{\gamma\delta} a_{\alpha\beta}] + \kappa^2 a_{\alpha\beta} a_{\gamma\delta}) = 0 .$$

Suppose them to be $\kappa^{(1)}$ and $\kappa^{(2)}$. Then the eigenvectors are solutions of the equations

$$(b_{\alpha\beta} - \kappa^{(\gamma)} a_{\alpha\beta}) \lambda^{\beta}_{(\gamma)} = 0 , \; (\gamma = 1, 2) ;$$

the eigenvectors will be assumed to have been normalized for convenience. If $\kappa^{(1)} \neq \kappa^{(2)}$, then $\lambda^{\beta}_{(1)}$ is orthogonal to $\lambda^{\beta}_{(2)}$; otherwise they may be chosen to be orthogonal.

Let us evaluate the coefficients of the powers of κ in equation (10.14) . They are, respectively,

$$\varepsilon^{\alpha\beta} \varepsilon^{\beta\delta} a_{\alpha\beta} a_{\gamma\delta} = \varepsilon_{\beta\delta} \varepsilon^{\beta\delta} = \delta^{\beta\delta}_{\beta\delta} = 2 ,$$

$$\varepsilon^{\alpha\gamma} \varepsilon^{\beta\delta} (b_{\alpha\beta} a_{\gamma\delta} + b_{\gamma\delta} a_{\alpha\beta}) = 2 \varepsilon^{\alpha\gamma} \varepsilon^{\beta\delta} b_{\alpha\beta} b_{\alpha\beta} a_{\gamma\delta}$$

$$= 2 a^{\alpha\lambda} \varepsilon_{\lambda\mu} \varepsilon^{\beta\delta} a_{\gamma\delta} = 2 (a^{\alpha\lambda} a^{\gamma\mu} \delta^{\beta\delta}_{\lambda\mu}) b_{\alpha\beta} a_{\gamma\delta}$$

$$= 2 (a^{\alpha\beta} a^{\gamma\delta} - a^{\alpha\delta} a^{\gamma\beta}) a_{\gamma\delta} b_{\alpha\beta} = 2 (a^{\alpha\beta} \delta^{\gamma}_{\gamma} - \delta^{\alpha}_{\gamma} a^{\gamma\beta}) b_{\alpha\beta} ,$$

$$= 2 (2 a^{\alpha\beta} - a^{\alpha\beta}) b_{\alpha\beta} = 2 a^{\alpha\beta} b_{\alpha\beta} ,$$

$$\varepsilon^{\alpha\gamma} \varepsilon^{\beta\delta} b_{\alpha\beta} b_{\gamma\delta} = a^{\alpha\mu} a^{\gamma\nu} \varepsilon_{\mu\nu} \varepsilon^{\beta\delta} b_{\alpha\beta} b_{\gamma\delta} = \left(\delta^{\beta\delta}_{\mu\nu} a^{\alpha\mu} a^{\gamma\nu} \right) b_{\alpha\beta} b_{\gamma\delta}$$

$$= \left(\delta^{\beta\delta}_{\mu\nu} b_{\alpha\beta} b_{\gamma\delta} \right) a^{\alpha\mu} a^{\gamma\nu} .$$

Now by equation (10.9) ,

$$\delta^{\beta\delta}_{\mu\nu} b_{\alpha\beta} b_{\gamma\delta} = R^*_{\alpha\gamma\mu\nu} \text{ so that } \varepsilon^{\alpha\gamma}\varepsilon^{\beta\delta} b_{\alpha\beta} b_{\gamma\delta} = R^*_{\alpha\gamma\mu\nu} a^{\alpha\mu} a^{\gamma\nu} .$$

On the other hand, by equation (2.7.8),

$$\begin{aligned} a^{\alpha\mu} a^{\gamma\nu} R^*_{\alpha\gamma\mu\nu} &= K(a_{\alpha\mu} a_{\gamma\nu} - a_{\alpha\nu} a_{\gamma\mu}) a^{\alpha\mu} a^{\gamma\nu} \\ &= K\left(\delta^\alpha_\alpha \delta^\gamma_\gamma - \delta^\mu_\nu \delta^\nu_\mu\right) = K\left(2^2 - 2\right) = 2K , \end{aligned}$$

where K is the Gaussian curvature. If, therefore, we set

$$a^{\alpha\beta} b_{\alpha\beta} = 2H , \tag{10.15}$$

the eigenvalue equation becomes

$$\kappa^2 - 2H\kappa + K = 0 . \tag{10.16}$$

Calling the roots of this quadratic equation $\kappa^{(1)}$ and $\kappa^{(2)}$, the eigenvalues which we seek, it is evident that

$$\kappa^{(1)} + \kappa^{(2)} = 2H , \quad \kappa^{(1)}\kappa^{(2)} = K . \tag{10.17}$$

In other words, the Gaussian curvature K is the product of the two **principal curvatures** of the surface, $\kappa^{(1)}$ and $\kappa^{(2)}$. On the other hand, the average of $\kappa^{(1)}$ and $\kappa^{(2)}$ is H, which is therefore called the **mean curvature** of the surface S.

Ex. (10.13) Find the principal curvatures on the surface of a spheroid. From them, find the mean curvature and the Gaussian curvature.

Ans.

$$\kappa^{(1)} = -\frac{(1 - e^2)}{\rho}\left[\frac{1 - e^2\cos^2\varphi}{1 - (2e^2 - e^4)\cos^2\varphi}\right]^{1/2} ,$$

$$\kappa^{(2)} = -\frac{(1 - e^2)}{\rho}\left[\frac{1 - e^2\cos^2\varphi}{1 - (2e^2 - e^4)\cos^2\varphi}\right]^{3/2}$$

$$H = -\frac{(1 - e^2)}{2\rho}\,\frac{(1 - e^2\cos^2\varphi)^{1/2}}{[1 - (2e^2 - e^4)\cos^2\varphi]^{3/2}}\left\{2 - (3e^2 - e^4)\cos^2\varphi\right\} ,$$

$$K = \frac{(1 - e^2)}{\rho^2}\left[\frac{1 - e^2\cos^2\varphi}{1 - (2e^2 - e^4)\cos^2\varphi}\right]^2 .$$

The principal curvatures have a relatively simple geometric meaning. With them are associated the two **principal directions**, defined by the two mutually orthogonal tangent vectors $\lambda^\alpha_{(1)}$ and $\lambda^\alpha_{(2)}$. Let $\lambda^\alpha_{(1)}$ and $\lambda^\alpha_{(2)}$ serve as an orthonormal basis for the resolution of all other surface vectors. Then any unit tangent vector λ^α which makes an angle ψ with $\lambda^\alpha_{(1)}$ and $90° - \psi$ with $\lambda_{(2)}$ will have components in this system given by

$$\lambda^\alpha = \lambda^\alpha_{(1)} \cos\psi + \lambda^\alpha_{(2)} \sin\psi .$$

We then have

$$\kappa_{(n)} = b_{\alpha\beta}\,\lambda^\alpha\lambda^\beta = \left[b_{\alpha\beta}\,\lambda^\alpha_{(1)}\lambda^\beta_{(1}\right]\cos^2\psi + \left[b_{\alpha\beta}\,\lambda^\alpha_{(2)}\lambda^\beta_{(2)}\right]\sin^2\psi$$

$$= \kappa^{(1)}\left[a_{\alpha\beta}\,\lambda^\alpha_{(1)}\lambda^\beta_{(1)}\right]\cos^2\psi + \kappa^{(2)}\left[a_{\alpha\beta}\,\lambda^\alpha_{(2)}\lambda^\beta_{(2)}\right]\sin^2\psi$$

$$= \kappa^{(1)}\cos^2\psi + \kappa^{(2)}\sin^2\psi .$$

This is **Euler's Theorem**.

If $\kappa^{(1)} = \kappa^{(2)}$, then $\kappa_{(n)}$ is their common value. In general, however, we may take $\kappa^{(1)} > \kappa^{(2)}$. Then, since both $\sin^2\psi$ and $\cos^2\psi$ are positive,

$$\kappa^{(1)} = \kappa^{(1)}\cos^2\psi + \kappa^{(1)}\sin^2\psi > \kappa^{(1)}\cos^2\psi + \kappa^{(2)}\sin^2\psi$$

$$= \kappa_{(n)} > \kappa^{(2)}\cos^2\psi + \kappa^{(2)}\sin^2\psi = \kappa^{(2)} .$$

In other words,

$$\kappa^{(1)} \geq \kappa_{(n)} \geq \kappa^{(2)} . \tag{10.19}$$

Therefore, $\kappa^{(1)}$ and $\kappa^{(2)}$ represent the maximum and minimum normal curvatures of all curves at a given point. The eigenvectors $\lambda^\alpha_{(1)}$ and $\lambda^\alpha_{(2)}$ are the tangents to these mutually orthogonal directions along which the surface curves have maximum and minimum normal curvature.

We may further note that when $\psi = 45°$, then

$$\kappa_{(n)} = \frac{1}{2}(\kappa^{(1)} + \kappa^{(2)}) = \mathrm{H} .$$

In other words, the normal curvature associated with tangent directions midway between the principal directions is equal to the mean curvature. We may also note that the mean curvature H vanishes only where $\kappa^{(1)} = -\kappa^{(2)}$; in general, therefore, zero mean curvature does not characterize a flat space. However, if either $\kappa^{(1)}$ or $\kappa^{(2)}$ is zero, the Gaussian curvature vanishes. Thus a plane is flat in the sense of having zero Gaussian curvature and also in having zero mean curvature. By contrast, a cylinder or cone is flat because for it $\mathrm{K} = 0$ though its mean curvature H is not zero.

The nature of the "geometry" depends upon the Gaussian curvature rather than the mean curvature. As we have seen, zero Gaussian curvature characterizes the plane, in which Euclidean geometry holds. However, if the Gaussian curvature is positive, then $\kappa^{(1)}$ and $\kappa^{(2)}$ are of the same sign; therefore the centers of curvature of all normal sections lie on the same side of the surface. The geometry of such a surface is called "elliptic" to distinguish it from plane geometry. If, however, $\kappa^{(1)}$ and $\kappa^{(2)}$ are of opposite sign, the centers of curvature of the normal sections containing the principal directions are on opposite sides of the surface. The geometry of such a surface is called "hyperbolic".

Ex. (10.14) Show that for the surface of a spheroid, the vectors defining the principal directions are

$$\lambda^{\alpha}_{(1)} = \left(\frac{[1 - e^2 \cos^2 \varphi]}{\rho \cos \varphi}, 0 \right)$$

and

$$\lambda^{\alpha}_{(2)} = \left(0, \frac{(1 - e^2 \cos^2 \varphi)^{3/2}}{\rho [1 - (2e^2 - e^4] \cos^2 \varphi]^{1/2}} \right).$$

Ex. (10.15) Find the Gaussian and mean curvatures of (a) a sphere, (b) a cylinder, and (c) a cone.

Ans. (a) $K = \frac{1}{a^2}$, $H = \frac{1}{a}$, *where a is the radius.*

(b) $K = 0$, $H = \frac{1}{2a}$, *where a is the radius.*

(c) $K = 0$, $H = \frac{|\tan \varphi_0|}{2r}$ *in a spherical coordinate system where* $\varphi = \varphi_0$ *defines the cone,* φ *being the latitude.*

Ex. (10.16) (a) Show that the unit tangent to the x^3*–curve in the* (θ, φ, ρ)*–coordinate system is*

$$u^i = \left(0, 0, \left[1 - e^2 \cos^2 \varphi\right]^{1/2}\right).$$

(b) Determine that this vector makes an angle υ *with the unit normal, where*

$$\tan \upsilon = \frac{e^2 \sin \varphi \cos \varphi}{1 - e^2 \cos^2 \varphi}.$$

(c) Show that the ***astronomical latitude*** $\psi = \upsilon + \phi$ *relates to the* ***geocentric latitude*** φ *according to the formula*

$$\tan \psi = \frac{\tan \varphi}{1 - e^2}.$$

(d) Show from Exs. (10.8) and (10.10) that

$$\tan \Theta = -\frac{\tan \varphi}{1 - e^2},$$

hence that $\Theta = 180 - \psi$, *where* Θ *is the angle between the normal to the spheroid and the principal normal of a parallel of latitude.*

Ex. (10.17) Let us define a ***geodesic spherical coordinate system*** *by considering all geodesics through a point* O *(the origin). Along each geodesic, let the arc length* s *be taken as the "radial" coordinate* x^1. *Then* $x^1 = $ ***constant*** *defines a particular coordinate surface* S. *At any point of* S *construct a geodesic polar coordinate system with coordinates* x^2 *and* x^3. *All* x^1- *curves containing any point of an* x^2- *curve span an* $(x^1, x^2)-$ *surface, etc.*

11. The Riemann-Christoffel Tensor in Three Dimensions

We have defined the Riemann-Christoffel tensor in a space of two dimensions and have related it to certain of the geometric properties of a surface. We may also define the Riemann-Christoffel tensor in three dimensions by the same identical equations, for there is nothing in the previous definition which implies any limitation to two dimensions. The previous symmetries and anti-symmetries in the indices remain valid. There are now six distinct components (aside from sign) rather than only one, however, and in general the tensor is not isotropic. We wish to explore further the properties of the Riemann-Christoffel tensor.

We have noted that there are six distinct non-zero components of the Riemann-Christoffel tensor in three dimensions: they are (1) three whose indices are of the form (*ijij*), and (2) three whose indices are of the form (*ikjk*). Let us now show that these six components may be expressed in terms of the six distinct components of a symmetric second rank tensor. We put

$$\mathrm{S}^{ij} = \frac{1}{4}\varepsilon^{ikl}\varepsilon^{jmn}\mathrm{R}_{klmn}. \tag{11.1}$$

Clearly S^{ij}, called the **Lamé tensor**, is a second rank contravariant tensor symmetric in i and j. From it we may recover the Riemann-Christoffel tensor, for

$$\begin{aligned}\varepsilon_{ipr}\varepsilon_{jst}\mathrm{S}^{ij} &= \frac{1}{4}\varepsilon_{ipr}\varepsilon^{ikl}\varepsilon_{jst}\varepsilon^{jmn}\mathrm{R}_{klmn},\\ \frac{1}{4}\delta^{ikl}_{ipr}\delta^{jmn}_{jst}\mathrm{R}_{klmn} &= \frac{1}{4}\delta^{kl}_{pr}\delta^{mn}_{st}\mathrm{R}_{klmn}\\ = \frac{1}{4}\delta^{mn}_{st}[\mathrm{R}_{prmn} - \mathrm{R}_{rpmn}] &= \frac{1}{2}\delta^{mn}_{st}\mathrm{R}_{prmn}\\ \frac{1}{2}\left[\mathrm{R}_{prst} - \mathrm{R}_{prts}\right] &= \mathrm{R}_{prst}.\end{aligned} \tag{11.2}$$

Now let us form directly from R_{qrsp} a symmetric covariant tensor, namely

$$\mathrm{R}_{rs} = g^{qp}\mathrm{R}_{qrsp}. \tag{11.3}$$

This is known as the **Ricci tensor**. From equation (11.1) we see that this is

$$R_{rs} = g^{qp} R_{qrsp} = g^{qp} \varepsilon_{iqr} \varepsilon_{jsp} S^{ij} = g_{im} g_{rn} \varepsilon^{mnp} \varepsilon_{jsp} S^{ij}$$

$$g_{im} g_{rn} \delta^{mnp}_{jsp} S^{ij} = - g_{im} g_{rn} \delta^{pmn}_{pjs} S^{ij}$$

$$- g_{im} g_{rn} \delta^{mn}_{js} S^{ij} = \left(- g_{ij} g_{rs} + g_{is} g_{rj}\right) S^{ij} = S_{rs} - S g_{rs}$$

where

$$S = g_{ij} S^{ij} .$$

Therefore the **curvature invariant** is

$$g^{rs} R_{rs} = g^{rs} g^{qp} R_{qrsp} \equiv R = g^{rs} S_{rs} - S g^{rs} g_{rs}$$

$$= g_{rs} S^{rs} - \delta^r_r S = S - 3 S = -2 S , \; S = - \frac{1}{2} R .$$

Hence

$$R_{rs} - \frac{1}{2} R g_{rs} = S_{rs} , \; S^{rs} = R^{rs} - \frac{1}{2} R g^{rs} ,$$

(11.4)

$$R_{prst} = \varepsilon_{ipr} \varepsilon_{jst} S^{ij} = \varepsilon_{ipr} \varepsilon_{jst} [R^{ij} - \frac{1}{2} R g^{ij}] .$$

However, we can easily show that

$$\varepsilon_{ipr} \varepsilon_{jst} g^{ij} = g_{ps} g_{rt} - g_{pt} g_{rs} .$$

Therefore

(11.5) $$R_{prst} = \varepsilon_{ipr} \varepsilon_{jst} R^{ij} + \frac{1}{2} R (g_{pt} g_{rs} - g_{ps} g_{rt}) .$$

The second term on the right hand side is evidently isotropic. The Riemann-Christoffel tensor as a whole, however, is isotropic only if R^{ij} is zero or isotropic, hence proportional to g^{ij}) .

We may therefore ask: under what condition is the Riemann-Christoffel tensor isotropic? In general the Ricci tensor R^{ij} will not be zero. We must therefore require that

(11.6) $$R^{ij} = \lambda g^{ij} , \; R_{ij} = \lambda g_{ij} ,$$

where λ is at most an invariant function of position. Its value may be found from the relation

$$g_{ij} R^{ij} = R = \lambda g_{ij} g^{ij} = 3 \lambda .$$

Consequently, $\lambda = R/3$. Therefore the Riemann-Christoffel tensor in three dimensions will be isotropic if and when

$$
\begin{aligned}
R_{prst} &= \varepsilon_{ipr}\,\varepsilon_{jst}\,[\,\tfrac{1}{3} - \tfrac{1}{2}\,]\,R\,g^{ij} \\
&= \frac{1}{6}R\left(g_{ps}g_{rt} - g_{pt}g_{rs}\right) = K\left(g_{ps}g_{rt} - g_{pt}g_{rs}\right), \qquad (11.7)\\
&\qquad - K \equiv \frac{1}{6}R\,.
\end{aligned}
$$

Ex. (11.1) Show that

$$R_{pijk} = \delta^{qr}_{pi}\,\delta^{st}_{jk}\left(\frac{\partial^2 g_{qt}}{\partial x^s \partial x^r} + g^{mn}[\,qt\,,m\,]\,[\,rs\,,n\,]\right).$$

(Hint: expand in straightforward fashion and compare the result with equation (2.7.3) of Ch.2).

Ex. (11.2) Show that

$$S^{ij} = \varepsilon^{iqr}\,\varepsilon^{jst}\left(\frac{\partial^2 g_{qt}}{\partial x^s \partial x^r} + g^{mn}[\,qt\,,m\,]\,[\,rs\,,n\,]\right).$$

Ex. (11.3)

$$R_{rs} = \frac{\partial}{\partial x^s}\left\{ {p \atop r\;p} \right\} - \frac{\partial}{\partial x^p}\left\{ {p \atop r\;s} \right\} + \left\{ {m \atop r\;p} \right\}\left\{ {p \atop m\;s} \right\} - \left\{ {m \atop r\;s} \right\}\left\{ {p \atop m\;p} \right\}.$$

(Hint: take $R_{rs} = R^p_{rsp}$ *and use equation (2.7.1).)*

Ex. (11.4) Show that

$$R = 2\left(g^{il}g^{jk} - g^{ik}\,g^{jl}\right)\left(\frac{\partial^2 g_{il}}{\partial x^j \partial x^k} + g^{mn}[\,il\,,m\,]\,[\,jk\,,n\,]\right).$$

Ex. (11.5) Given that

$$g_{ij} = \delta_{ij}\,[\,1 + k\,\delta_{mn}x^m x^n\,]^{-2},$$

(a) determine by direct calculation that

$$R_{1212} = R_{1313} = R_{2323} = 4k\left(1 + \delta_{mn}x^m x^n\right)^{-4},$$

$$R_{1213} = R_{1223} = R_{1323} = 0\,.$$

(Suggestion: calculate R^1_{212}*, then* $R_{1212} = g_{11}R^1_{212}$*, etc.).*

(b) Show that

$$\mathrm{R}_{ij} = -8k\delta_{ij}\Big[1 + \delta_{mn}x^m x^n\Big]^{-2} = -8kg_{ij}.$$

(c) Show that $\mathrm{R} = -24k$, *hence that* $\mathrm{K} = -\frac{1}{6}\mathrm{R} = 4k$, $k = \frac{1}{4}\mathrm{K}$.

The condition of isotropy is an even more stringent one than is apparent in the preceding result, however, for in three (or more) dimensions it requires also that R be a constant. This is **Schur's Theorem**. To prove Schur's theorem, we appeal to the two identities

$$\frac{1}{2}\delta^{ijk}_{abc}\mathrm{R}_{pijk} \equiv \mathrm{R}_{pabc} + \mathrm{R}_{pcab} + \mathrm{R}_{pbca} = 0 \tag{11.8}$$

and

$$\frac{1}{2}\delta^{ijk}_{abc}\mathrm{R}_{phij,k} \equiv \mathrm{R}_{phab,c} + \mathrm{R}_{phca,b} + \mathrm{R}_{phbc,a} = 0\,. \tag{11.9}$$

The latter is known as **Bianchi's identities**.

The former identity may be proven quite directly by starting from the expression for the Riemann-Christoffel tensor in the form (see Ex. (11.1))

$$\mathrm{R}_{pijk} = \delta^{qr}_{pi}\delta^{st}_{jk}\left(\frac{\partial^2 g_{qt}}{\partial x^s \partial x^r} + g^{mn}[qt, m][rs, n]\right).$$

Then

$$\frac{1}{2}\delta^{ijk}_{abc}\mathrm{R}_{pijk} = \frac{1}{2}\delta^{ijk}_{abc}\delta^{qr}_{pi}\delta^{st}_{jk}\left(\frac{\partial^2 g_{qt}}{\partial x^s \partial^r} + g^{mn}[qt, m][rs, n]\right).$$

Now

$$\frac{1}{2}\delta^{ijk}_{abc}\delta^{st}_{jk}\delta^{qr}_{pi} = \frac{1}{2}\left(\delta^{ist}_{abc} - \delta^{its}_{abc}\right)\delta^{qr}_{pi} = \delta^{ist}_{abc}\delta^{qr}_{pi}$$

since $\delta^{its}_{abc} = -\delta^{ist}_{abc}$. Further,

$$\delta^{ist}_{abc}\delta^{qr}_{pi} = \delta^{ist}_{abc}\left(\delta^q_p\delta^r_i - \delta^q_i\delta^r_p\right) = \delta^q_p\delta^{str}_{abc} - \delta^r_p\delta^{qst}_{abc}.$$

Therefore

$$\mathrm{R}_{pabc} + \mathrm{R}_{pcab} + \mathrm{R}_{pbca} = \frac{1}{2}\delta^{ijk}_{abc}\mathrm{R}_{pijk}$$

$$= \left(\delta^q_p\delta^{str}_{abc} - \delta^r_p\delta^{qst}_{abc}\right)\left(\frac{\partial^2 g_{qt}}{\partial x^s \partial x^r} + g^{mn}[qt,m][rs,n]\right)$$

$$= \delta^{str}_{abc}\left(\frac{\partial^2 g_{pt}}{\partial x^s \partial x^r} + g^{mn}[pt,m][rs,n]\right)$$

$$- \delta^{qst}_{abc}\left(\frac{\partial^2 g_{qt}}{\partial x^s \partial x^p} + g^{mn}[qt,m][ps,n]\right).$$

The first term on the right hand side vanishes because of the symmetry in the parenthesis in r and s as against the antisymmetry of δ^{str}_{abc} in r and s; the second term vanishes because of the symmetry of the parenthesis in q and t as against the antisymmetry of δ^{qst}_{abc} in q and t. This proves the first identity.

To prove the Bianchi identity, we start from the fact that

$$\delta^{lmn}_{ijk} = \left(\delta^l_i\delta^{mn}_{jk} + \delta^m_i\delta^{nl}_{jk} + \delta^n_i\delta^{lm}_{jk}\right) = \left(\delta^{lm}_{ij}\delta^n_k + \delta^{mn}_{ij}\delta^l_k + \delta^{nl}_{ij}\delta^m_k\right).$$

Then for any covariant vector w_h, using the first expansion,

$$\delta^{lmn}_{ijk}w_{h,lmn} = (\delta^{mn}_{jk}w_{h,i})_{,mn} + (\delta^{nl}_{jk}w_{h,l})_{,in} + (\delta^{lm}_{jk}w_{h,l})_{,mi}.$$

On the other hand, the second expansion gives

$$\delta^{lmn}_{ijk}w_{h,lmn} = (\delta^{lm}_{ij}w_{h,lm})_{,k} + (\delta^{mn}_{ij}w_{h,km})_{,n} + (\delta^{nl}_{ij}w_{h,lk})_{,n}.$$

Equating the two expansions and using the relation

$$\delta^{lm}_{ij}w_{h,lm} = \mathrm{R}^p_{kij}w_p,$$

we have

$$(\delta^{mn}_{jk}w_{h,i})_{,mn} + (w_{h,kij} - w_{h,jik}) + (w_{h,jki} - w_{h,kji})$$

$$= \left(\mathrm{R}^p_{hij}w_p\right)_{,k} + \left(w_{h,kij} - w_{h,kji}\right) + \left(w_{h,jki} - w_{h,ikj}\right).$$

After cancelling like terms on both sides, we are left with

$$\delta^{mn}_{jk}(w_{h,i})_{,mn} - w_{h,jik} + w_{h,ikj} = \mathrm{R}^p_{hij,k}w_p + \mathrm{R}^p_{hij}w_{p,k}.$$

The first term on the left is expressible in terms of the Riemann-Christoffel tensor, since $w_{h,i}$ is a covariant tensor of the second rank and therefore expressible as the sum of products of two covariant vectors. Let a typical term be $u_h \mathrm{v}_i$.

Then

$$\delta^{mn}_{ik}(u_h \mathrm{v}_i)_{,mn} = \delta^{mn}_{jk}(u_{h,m}\mathrm{v}_i + u_h \mathrm{v}_{i,m})_{,n}$$

$$= \delta^{mn}_{jk}[(u_{h,mn}\mathrm{v}_i + u_h \mathrm{v}_{i,mn}) + (u_{h,m}\mathrm{v}_{i,n} + u_{h,n}\mathrm{v}_{i,m})]$$

$$= \mathrm{R}^p_{hjk}(u_p \mathrm{v}_i) + \mathrm{R}^p_{ijk}(u_h \mathrm{v}_p) + \delta^{mn}_{jk}(u_{h,m}\mathrm{v}_{i,n} + u_{h,n}\mathrm{v}_{i,m}) .$$

Since the quantity in parenthesis in the last term is symmetric in *m* and *n*, this sum over *m* and *n* must vanish. Therefore, since a sum of terms of the form

$$\delta^{mn}_{jk}(u_h v_i)_{,mn} = \mathrm{R}^p_{hjk}(u_p v_i) + \mathrm{R}^p_{ijk}(u_h v_p)$$

is also of this form, we see that

$$\delta^{mn}_{jk}(w_{h,i})_{,mn} = \mathrm{R}^p_{hjk} w_{p,i} + \mathrm{R}^p_{ijk} w_{h,p} .$$

Therefore, making this substitution in our remaining terms, we have

$$\mathrm{R}^p_{hjk} w_{p,i} + \mathrm{R}^p_{ijk} w_{h,p} - \mathrm{R}^p_{hij} w_{p,k} + w_{h,ikj} - w_{h,jik} = \mathrm{R}^p_{hij,k} w_p .$$

We now multiply both sides of this equation by δ^{ijk}_{abc}. The second term on the left vanishes, by our previous identity, equation (11.8). From the first and third terms we get

$$\delta^{ijk}_{abc}\mathrm{R}^p_{hjk} w_{p,i} - \delta^{ijk}_{abc}\mathrm{R}^p_{hij} w_{p,k} = \delta^{ijk}_{abc}\mathrm{R}^p_{hjk} w_{p,i} - \delta^{kji}_{abc}\mathrm{R}^p_{hkj} w] w_{p\ i}$$

$$= [\delta^{ijk}_{abc}\mathrm{R}_{hkj} + \delta_{abc}\mathrm{R}^p_{hjk}] w_{p,i} , \quad \delta^{ijk}_{abc}[\mathrm{R}^p_{hjk} - \mathrm{R}^p_{hjk}] w_{p,i} = 0 .$$

The remaining two terms on the left give

$$\delta^{ijk}_{abc} w_{h,ihj} - \delta^{ijk}_{abc} w_{h,jk} = \delta^{ijk}_{abc} w_{h,ikj} - \delta^{kij}_{abc} w_{h,ikj}$$

$$= (\delta^{ijk}_{abc} - \delta^{kij}_{abc}) w_{h,ikj} = 0 .$$

The entire left hand side therefore vanishes and we have for an arbitrary vector w_p that

$$\delta^{ijk}_{abc}\mathrm{R}^p_{hij,k} w_p = 0 .$$

Hence

$$\frac{1}{2}\delta^{ijk}_{abc}\mathrm{R}^p_{hij,k} = \mathrm{R}^p_{hab,c} + \mathrm{R}^p_{hca,b} + \mathrm{R}^p_{hbc,a} = 0 ,$$

thus proving the Bianchi identities.

Let us now substitute $-\mathrm{K}$ for $\dfrac{\mathrm{R}}{6}$ and write equation (11.7) as

$$\mathrm{R}_{prst} = \mathrm{K}\,(g_{ps}\,g_{rt} - g_{pt}\,g_{rs}) = \mathrm{K}\,\delta^{vw}_{st}\,g_{pw}\,g_{rv}$$

as the form of the isotropic Riemann-Christoffel tensor. The Bianchi identity for this tensor is then

$$\frac{1}{2}\,\delta^{stu}_{abc}\,\mathrm{R}_{prst,u} = \frac{1}{2}\,\delta^{stu}_{abc}\,\delta^{vw}_{st}\,g_{pw}\,g_{rv}\,\mathrm{K}_{,u} = 0\,.$$

Multiplying through by $g^{pb}\,g^{ra}$, we find that

$$\frac{1}{2}\,\delta^{vw}_{st}\,\delta^{stu}_{abc}\Big[\,(g_{pw}\,g^{pb})\,(g_{rv}\,g^{ra})\,\Big]\mathrm{K}_{,u}$$

$$= \frac{1}{2}\,(\,\delta^{vwu}_{abc} - \delta^{wvu}_{abc}\,)\,\delta^{b}_{w}\,\delta^{a}_{v}\,\mathrm{K}_{,u} = (\,\delta^{vwu}_{abc}\,\delta^{a}_{v}\,\delta^{b}_{w}\,)\,\mathrm{K}_{,u}$$

$$= \delta^{abu}_{abc}\,\mathrm{K}_{,u} = 2\,\mathrm{K}_{,c} = 0\,.$$

Therefore

$$\frac{\partial \mathrm{K}}{\partial x^{c}} = 0\,,$$

and K is a constant, as Schur's theorem states.

Ex. (11.6) Given the metric tensor

$$g_{ij} = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & \mathrm{R}^2 \sin^2\!\left(\dfrac{r}{\mathrm{R}}\right)\cos^2\varphi & 0 \\ 0 & 0 & \mathrm{R}^2 \sin^2\!\left(\dfrac{r}{\mathrm{R}}\right) \end{Vmatrix},$$

with R *= constant, determine whether or not the Riemann-Christoffel tensor is isotropic. (Hint: establish from the preceding and following problems that the Riemannian curvature is constant, then appeal to Schur's Theorem.)*

Ans. Since $\mathrm{K} = \mathrm{R}^{-2} =$ ***constant***, *the Riemann-Christoffel tensor is isotropic.*

Ex. (11.7) Show that the space in which the line element in polar coordinates is

$$(ds)^2 = (dr)^2 + \mathrm{R}^2 \sinh^2\!\left(\frac{r}{\mathrm{R}}\right)[\,(\cos^2\varphi)\,(d\theta)^2 + (d\varphi)^2]$$

has a constant negative curvature $\mathrm{K} = -\dfrac{1}{\mathrm{R}^2}\,.$

Ex. (11.8) Show that the metric tensor of Ex. (11.6) can be gotten from that of Ex. (11.5) by the transformation

$$x^m = \left[2\,\mathrm{R} \tan\left(\frac{r}{2\,\mathrm{R}} \right) \right] u^m .$$

Here we understand x^m *to be the coordinates of some point* $\mathrm{P}(x^r)$ *relative to an origin* O. *Let* P *be connected to* O *by a geodesic. Then* u^m *is a unit tangent to the geodesic at* O, *hence*

$$u_m u^m = 1 \,, \; u^m du_m + u_m du^m = 0 \,.$$

The quantity r is the arc length along the geodesic from O *to* P. *Thus the* x^m *are both the coordinates of the point* P *and the components of the position vector of* P. *The quantity k is* $1/(4\,\mathrm{R}^2)$

Ans. First we have

$$1 + k\,\delta_{mn} x^m x^n = 1 + \frac{1}{4\,\mathrm{R}^2}\left[4\,\mathrm{R}^2 \tan^2\left(\frac{r}{\mathrm{R}} \right) \right] \delta_{mn} u^m u^n$$

$$= 1 + \left[\tan^2\left(\frac{r}{2\,\mathrm{R}} \right) \right] u_m u^m = 1 + \tan^2\left(\frac{r}{2\,\mathrm{R}} \right) = \sec^2\left(\frac{r}{2\,\mathrm{R}} \right) .$$

Then

$$dx^m = \left[2\,\mathrm{R} \tan\left(\frac{r}{2\,\mathrm{R}} \right) \right] du^m + u^m \left[2\,\mathrm{R} \sec^2\left(\frac{r}{2\,\mathrm{R}} \right) \right] dr \,,$$

whence

$$dx^m dx_m = 4\,\mathrm{R}^2 \tan^2\left(\frac{r}{2\,\mathrm{R}} \right) du^m du_m$$

$$+ \; 2\,\mathrm{R}\left(u^m du_m + u_m du^m \right) \tan\left(\frac{r}{2\,\mathrm{R}} \right) \sec^2\left(\frac{r}{2\,\mathrm{R}} \right) dr$$

$$+ \; (u^m u_m) \sec^4\left(\frac{r}{2\,\mathrm{R}} \right) (dr)^2 .$$

$$du^m du_m = \cos^2\varphi \, (d\theta)^2 + (d\varphi)^2 \;\; \textit{and}$$

$$u^m u_m = 1 \,, \; u^m du_m + u_m du^m = 0 \;\; \textit{so that}$$

$$dx^m dx_m = 4\,\mathrm{R}^2 \tan^2\left(\frac{r}{2\,\mathrm{R}} \right) [\cos^2\varphi \, (d\theta)^2 + (d\varphi)^2] + \sec^4\left(\frac{r}{2\,\mathrm{R}} \right) (dr)^2 .$$

Therefore

$$(ds)^2 = \frac{\sec^4\left(\dfrac{r}{2\,\mathrm{R}} \right) (dr)^2 + 4\,\mathrm{R}^2 \tan^2\left(\dfrac{r}{2\,\mathrm{R}} \right) [\cos^2 \varphi \, (d\theta)^2 + (d\phi)^2]}{\sec^4\left(\dfrac{r}{2\,\mathrm{R}} \right)}$$

$$= (dr)^2 + \left[2\,\mathrm{R} \sin\left(\frac{r}{2\,\mathrm{R}} \right) \cos\left(\frac{r}{2\,\mathrm{R}} \right) \right]^2 [\cos^2 \varphi \, (d\theta)^2 + (d\varphi)^2]$$

$$= (dr)^2 + \mathrm{R}^2 \sin^2\left(\frac{r}{\mathrm{R}} \right) [\cos^2 \varphi \, (d\theta)^2 + (d\varphi)^2] .$$

Ex. (11.9) Derive the line element of Ex. (11.7) from that of Ex. (11.5) by a transformation analogous to that of Ex. (11.8).

Ex. (11.10) In three dimensions, both the Ricci tensor and the Riemann-Christoffel tensor have six distinct components. Therefore solve equation (11.3) for the Riemann-Christoffel tensor.

Ans.

$$\mathrm{R}_{ijkl} = g_{il}\,\mathrm{R}_{jk} + g_{jk}\,\mathrm{R}_{il} - g_{ik}\,\mathrm{R}_{jl} - g_{jl}\,\mathrm{R}_{ik} + \frac{\mathrm{R}}{2}\left(g_{ik}\, g_{jl} - g_{il}\, g_{jk} \right),$$

where

$$\mathrm{R} = g^{ij}\,\mathrm{R}_{ij} .$$

Appendix 4.1 : The Advance of Perihelion

From Exs. (3.4) and (3.5) we have as the equation of motion

$$\frac{m}{u^5}\left[u\,\ddot{u} - 2\,\dot{u}^2 + u^2 \cos^2 \varphi \, \dot{\theta}^2 + u^2\,\dot{\varphi}^2 \right]$$

$$= \mathrm{G}\, m\, (\mathrm{M} + m) \left\{ 1 + \frac{1}{(c\, u^2)^2} [9\,\dot{u}^2 - 6\, u\, \ddot{u}] + \ldots \right\},$$

$$m \frac{d}{dt}\left(\frac{\cos^2\varphi}{u^2}\, \dot{\theta} \right) = 0 ,$$

$$m \left[\frac{d}{dt}\left(\frac{\dot{\varphi}}{u^2} \right) + \frac{\sin \varphi \cos \varphi}{u^2}\, \dot{\theta}^2 \right] = 0 ,$$

where we have made use of the fact that $|\dot{u}/u| << c$. Here M is the mass of the primary (sun) and *m* the mass of the much smaller secondary (planet). Guided by the well-known solution to the two-body problem, we recognize that there is no loss of generality in taking the coordinate system to be such that the orbital plane is the plane $\varphi = 0$. Then the third equation is automatically satisfied and the second becomes

$$\frac{d}{dt}\left(\frac{\dot{\theta}}{u^2}\right) = 0\,, \quad \dot{\theta} = h u^2\,, \quad h = \frac{2\pi a^2 \sqrt{1 - e^2}}{\mathrm{P}}\,,$$

where *a* is the major semi-axis of the orbital ellipse, *e* the eccentricity, and P the period of revolution. With this integral, we eliminate derivatives with respect to *t* in favor of derivatives with respect to θ according to the equation

$$\frac{d}{dt} \equiv h u^2 \frac{d}{d\theta}\,.$$

Then the first equation takes the form

$$u'' + u = \frac{\mu}{h^2}\left\{1 - 3\frac{h^2}{c^2}\left(2uu'' + u^2\right)\ldots\right\}.$$

Here primes denote derivatives with respect to θ. Since $\dfrac{h}{c} << 1$, we seek a solution of the form

$$u = u_0 + \left(\frac{\mu}{h^2}\right)^3 \frac{h^2}{c^2} u_1 + \ldots.$$

Then

$$u_0'' + u_0 = \frac{\mu}{h^2}\,, \quad u_1'' + u_1 = -3\left(\frac{h^2}{\mu}\right)^2\left(2u_0u_0'' + u_0'^2\right).$$

The first of these has the well-known Keplerian solution

$$u_0 = \frac{\mu}{h^2}[1 + e\cos\theta]\,,$$

where $0 \le e < 1$ for elliptical motion and $\theta = 0$ corresponds to the position of perihelion. Substituting this expression for u_0 into the second equation gives

$$u_1'' + u_1 = -3e^2 + 6e\cos\theta + 9e^2\cos^2\theta$$

$$= 6e^2 + 6e\cos\theta - 9e^2\sin^2\theta = f(\theta)\,.$$

Solving this equation by any standard method, such as variation of parameters, and imposing the initial conditions $u_1(0) = 0$, $u_1'(0) = 0$ for perihelion, we find that

$$u_1(\theta) = 3e\left(\theta\sin\theta + e\sin^2\theta\right).$$

Hence

$$u = \frac{\mu}{h^2}\left\{1 + e\left[\cos\theta + 3\left(\frac{\mu}{ch}\right)^2 \theta\sin\theta + 3e\left(\frac{\mu}{ch}\right)^2 \sin^2\theta\right] + \ldots\right\}$$

$$= \frac{\mu}{h^2}\left\{1 + e\left[\cos\theta + (\kappa\theta)\sin\theta\right] + e^2\left[\kappa\sin^2\theta\right] + \ldots\right\}$$

where
$$\kappa = 3\left(\frac{\mu}{ch}\right)^2 = \frac{12\pi^2 a^2}{(1 - e^2)P^2 c^2} << 1\,.$$

Since
$$\cos(\kappa\theta) \cong 1 \text{ and } \sin\kappa\theta \cong \kappa\theta\,,$$

we write the solution in the final form

$$u = \frac{\mu}{h^2}\left\{1 + e\left[\cos(\kappa\theta)\cos\theta + \sin(\kappa\theta)\sin\theta\right] + e^2\kappa\sin^2\theta + \ldots\right\}$$

$$= \frac{\mu}{h^2}\left\{1 + e\cos[(1 - \kappa)] + e^2\kappa\sin^2\theta + \ldots\right\}.$$

In the ordinary Keplerian solution, the secondary returns to perihelion after one sidereal revolution, i. e. when $\theta = 2\pi$. However, in this instance we see that

$$u(2\pi) = \frac{\mu}{h^2}\left\{1 + e\cos[2\pi(1 - \kappa)]\right\}.$$

In other words, the secondary falls short of its perihelion position by an angle $2\pi\kappa$; this is another way of saying that the perihelion has advanced by $2\pi\kappa$ radians. For the planet Mercury, $2\pi\kappa$ amounts to 43”/century.

Appendix 4.2 The Two-Center Problem

The gravitational potential function V, in the bipolar coordinates of Fig. 33 (or Fig. 76), is clearly

$$V = -\frac{\mu m}{\rho} - \frac{(1 - \mu)m}{\sigma}.$$

Since $\sigma = a[\cosh\xi - \cos\eta]$, $\rho = a[\cosh\xi + \cos\eta]$, the potential function therefore becomes

$$V = -\frac{\mu m}{a[\cosh\xi + \cos\eta]} - \frac{(1 - \mu)m}{a[\cosh\xi - \cos\eta]}.$$

On the other hand, the kinetic energy T is (see Ex. 2.1.8)

$$T = \frac{1}{2}ma^2(\cosh^2\xi - \cos^2\eta)(\dot{\xi}^2 + \dot{\eta}^2).$$

Hence

$$L = T - V = \frac{1}{2} m a^2 (\cosh^2 \xi - \cos^2 \eta)(\dot{\xi}^2 + \dot{\eta}^2)$$

$$+ \frac{\mu m}{a[\cosh \xi + \cos \eta]} + \frac{(1 - \mu) m}{a[\cosh \xi - \cos \eta]} .$$

Writing Lagrange's equations in the form

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}^i} \right) - \frac{\partial T}{\partial x^i} = \frac{d}{dt}\left(\frac{\partial V}{\partial \dot{x}^i} \right) - \frac{\partial V}{\partial x^i} ,$$

we have when $i = 1$ that

$$m a^2 \frac{d}{dt}[(\cosh^2 \xi - \cos^2 \eta)\dot{\xi}] - \frac{m}{2} a^2 \left(\dot{\xi}^2 + \dot{\eta}^2\right) \frac{\partial}{\partial \xi}(\cosh^2 \xi - \cos^2 \eta) = - \frac{\partial V}{\partial \xi} .$$

When $i = 2$ we get the corresponding equation

$$m a^2 \frac{d}{dt}[(\cosh^2 \xi - \cos^2 \eta)\dot{\eta}] - \frac{ma^2}{2}(\dot{\xi}^2 + \dot{\eta}^2) \frac{\partial}{\partial \eta}\left(\cosh^2 \xi - \cos^2 \eta\right) = - \frac{\partial V}{\partial \eta} .$$

To solve these equations, we multiply the first by $\dot{\xi}$, the second by $\dot{\eta}$ and add. This gives

$$m a^2 \left\{ \left(\cosh^2 \xi - \cos^2 \eta\right) \left(\dot{\xi} \frac{d\dot{\xi}}{dt} + \dot{\eta} \frac{d\dot{\eta}}{dt} \right) + \left(\dot{\xi}^2 + \dot{\eta}^2\right) \frac{d}{dt}\left(\cosh^2 \xi - \cos^2 \eta\right) \right.$$

$$\left. - \frac{1}{2}\left(\dot{\xi}^2 + \dot{\eta}^2\right) \left[\dot{\xi} \frac{\partial}{\partial \xi}\left(\cosh^2 \xi - \cos 2\eta\right) + \dot{\eta} \frac{\partial}{\partial \eta}\left(\cosh^2 \xi - \cos^2 \eta\right) \right] \right\}$$

$$= - \left[\frac{\partial V}{\partial \xi} \dot{\xi} + \frac{\partial V}{\partial \eta} \dot{\eta} \right].$$

After collecting terms and integrating, this becomes

(A4.2.1) $$\frac{1}{2} m a^2 (\cosh^2 \xi - \cos^2 \eta)(\dot{\xi}^2 + \dot{\eta}^2) \equiv T = E - V ,$$

where E is a constant, identifiable as the total energy.

Next, we multiply the first equation by $2\dot{\xi}(\cosh^2 \xi - \cos^2 \eta)$.

Then

$$m\,a^2 \frac{d}{dt}[\,(\cosh^2\xi \;-\; \cos^2\eta\,)\,\dot{\xi}\,]^2 \;=\; -\;2\,\dot{\xi}\,(\cosh^2 \;-\; \cos^2\eta\,)\,\frac{\partial V}{\partial \xi}$$

$$+\;m\,a^2\,(\cosh^2\xi \;-\; \cos^2\eta\,)\,(\,\dot{\xi}^2 \;+\; \dot{\eta}^2\,)\,\dot{\xi}\,\frac{\partial}{\partial \xi}(\cosh^2\;\xi \;-\; \cos^2\eta\,)$$

$$=\;-\;2\,\dot{\xi}\,(\cosh^2\xi \;-\; \cos^2\eta\,)\,\frac{\partial V}{\partial \xi} \;+\; (E \;-\; V)\,2\,\dot{\xi}\frac{\partial}{\partial \xi}(\cosh^2\xi \;-\; \cos^2\eta\,)$$

$$=\;2\,\dot{\xi}\,\frac{\partial}{\partial \xi}\left\{(E \;-\; V)\,(\cosh^2\xi \;-\; \cos^2\eta\,)\right\}$$

$$=\;2\,\dot{\xi}\,\frac{\partial}{\partial \xi}\left\{E\,(\cosh^2\xi \;-\; \cos^2\eta\,) \;+\; \frac{\mu\,m}{a}(\cosh\xi \;+\; \cos\eta\,)\right.$$

$$\left.+\;\frac{(1 \;-\; \mu\,)}{a}(\cosh\xi \;-\; \cos\eta\,)\right\}$$

$$=2\,\frac{d}{dt}\left\{E\cosh^2\xi \;+\; \frac{m}{a}\cosh\xi\right\}.$$

Integrating both sides, we get

$$\text{(A4.2.2)} \qquad \frac{m\,a^2}{2}(\cosh^2\xi \;-\; \cos^2\eta\,)^2\,\dot{\xi}^2 \;=\; E\cosh^2\xi \;+\; \frac{m}{a}\cosh\xi \;-\; \gamma\,,$$

where γ is a second constant of integration.

We now write equation (A4.2.1) in the form

$$(\cosh^2\xi - \cos^2\eta\,)\,T \;=\; \frac{m\,a^2}{2}(\cosh^2 - \cos^2)^2\,(\,\dot{\xi}^2 + \dot{\eta}^2$$

$$=\;(E - V)\,(\cosh^2\xi - \cos^2\theta\,)$$

$$=\;E\left(\cosh^2\xi \;-\; \cos^2\eta\right) \;+\; \frac{\mu m}{a}\left(\cosh\xi \;-\; \cos\eta\right) \;+\; \frac{(1 \;-\; \mu\,)}{a}\left(\cosh\xi \;+\; \cos\eta\right)$$

$$=\;E\left(\cosh^2\xi \;-\; \cos^2\eta\right) \;-\; \frac{2\,\mu\,m}{a}\cos\eta \;+\; \frac{m}{a}\cosh\xi\,.$$

Subtracting equation (A4.2.2) from this gives

$$\text{(A4.2.3)} \qquad \frac{m a^2}{2}(\cosh^2\xi \;-\; \cos^2\eta\,)\,\dot{\eta}^2 \;=\; -\;E\cos^2\eta \;-\; \frac{2\mu m}{a}\cos\eta \;+\; \gamma\,.$$

From equations (A4.2.2) and (A4.2.3) it is clear that

$$\frac{d\xi}{\left[\mathrm{E}\cosh^2\xi + \frac{m}{a}\cosh\xi - \gamma\right]^{1/2}} = \frac{dt}{a\left[\frac{m}{2}(\cosh^2\xi - \cos^2\eta)\right]^{1/2}}$$

$$= \frac{d\eta}{\left[-\mathrm{E}\cos^2\eta - \frac{2\mu m}{a}\cos\eta + \gamma\right]^{1/2}}.$$

The trajectory of the mass m is given implicitly by the integrals

$$\int\left[\mathrm{E}\cosh^2\xi + \frac{m}{a}\cosh\xi - \gamma\right]^{-1/2} d\xi = u = \left[\gamma - \frac{2\mu m}{a}\cos\eta - \mathrm{E}\cos^2\eta\right]^{-1/2} d\eta$$

while the motion is determined by the integral

$$t = a\int\left[\frac{m}{a}(\cosh^2\xi - \cos^2\eta)\right] du.$$

Notes - Chapter 4

§4.2 For a thorough treatment of the properties of space curves, see Laugwitz (8), Ch. I. Among the results of interest are:

(i) A curve is a plane curve if and only if its torsion vanishes.

(ii) In the neighborhood of a point P_0 on a curve whose parameter value is s_0, the position vector has the canonical expansion

$$p^i(s_0 + \Delta s) = p^i(s_0) + \Delta s\left[1 - \frac{\kappa^2}{6}(\Delta s)^2\right]\lambda^i$$
$$+ \frac{(\Delta s)^2}{2}\left[\kappa + \frac{\kappa'}{3}\Delta s\right]\mu^i + \left[\frac{(\Delta s)^3}{6}\kappa\tau\right]\nu^i + \ldots$$

through the third order in Δs. It is illuminating to interpret the contribution of each order of Δs.

(iii) To within a rigid body displacement, the functions $\kappa(s)$ and $\tau(s)$ determine uniquely the curve for which $\kappa(s)$, $\tau(s)$, and s are the curvature, torsion, and arc length, respectively. This is the content of what is known as the **Fundamental Theorem of the theory of curves** in Euclidean 3-space.

§4.3 (a) A detailed discussion of the derivation of the Euler-Lagrange equations under a wide variety of conditions and of their application to many problems of particle and continuum mechanics is given a model exposition in Lanczos (7).

(b) Many concrete, sophisticated applications of the Euler-Lagrange method are worked out fully in Wells (21).

(c) Whittaker (24) describes a formulation of the potential and kinetic energies by A. G. Walker which reduces to the traditional forms in the limit as the velocity of light $c \to \infty$ and yet which may be used with the classical Euler-Lagrange equations to derive the standard results of both special and general relativity, including the advance of perihelion and the gravitational refraction (see App. 4.1 and App. 6.1).

(d) The great utility and frequent economy of curvilinear coordinates is sometimes offset in part by the difficulty of translating the components of vector or tensor quantities in these coordinates into terms which may readily be apprehended physically. Even the dimensions of physical quantities may depend upon the choice of coordinate system; a familiar case in point is velocity, whose components are all linear in Cartesian coordinates but include angular components in spherical coordinates. For physical applications it is therefore frequently desirable to define the **physical components** of a vector or tensor. If λ^i is a unit vector in any direction, then $(g_{ij} v^i \lambda^j)$ is the physical component of the vector v^i in the direction of λ^i. By the use of the Product Theorem, one can readily define physical components for tensors of any order and type. For a more detailed discussion of the matter, see Synge and Schild (18), §5.1, pp. 142-149.

§4.4 (a) For other, more extensive expositions of geometrical dynamics, see Laugwitz (8), §14, pp. 167-176; McConnell (9), pp. 246-249; or Synge and Schild (18), §5.5, pp. 168-184.

(b) Closely related to geometrical dynamics is the content of **Helmholtz' Theorem** which states that **if there is free mobility of rigid bodies of any finite extension, then the space is Riemannian of constant curvature**. For a proof of this theorem, see Laugwitz (8), pp. 183-184.

§§4.3-4.4 Detailed consideration of many common problems in the mechanics of particles and rigid bodies will be found in Spiegel (16).

§4.5 Generalizations of the theorems of Green and Stokes as well as a discussion of relative tensors (as opposed to absolute tensors, to which we have confined our discussion) may be found in Synge and Schild (18), Ch. VII, pp. 240-281. Detailed discussion of divergence, curl, and Laplacian will be found in Davis (4), §§3.3-3.6, pp. 97-116.

The **jump discontinuity** referred to in Ex. (5.12) may be defined by saying that a vector F^r is said to have a jump discontinuity on a surface σ if there are two continuous vectors $F^r_{(1)}$ and $F^r_{(2)}$ defined over σ and such that the values of F^r at points near σ on one side approach the values of $F^r_{(1)}$ and on the other side approach the values of $F^r_{(2)}$.

§4.6 Strain, rigid body displacements, and infinitesimal deformations are discussed in McConnell (8), Ch. X, pp. 120-128 and in Synge and Schild (18), §§5.3-5.4, pp. 156-168.

§4.8 P. W. Bridgman, in his *The Logic of Modern Physics (Macmillan, New York, 1949, pp. 56-57)*, offers some very *a propos* comments on the subject of fields which may be applied more generally to all fields. They are worth quoting at length:

> "[A]n examination of the operations by which we determine the electric field at any point will show that it is a construct, in that it is not a direct datum of experience. To determine the electric field at any point, we place an exploring charge at the point, measure the force on it, and then calculate the ratio of the force to the charge. We then allow the exploring charge to become smaller and smaller, repeating our measurement of force on each smaller charge, and define the limit of the ratio of the force to the charge as the electric field intensity at the point in question, and the limiting direction of the force on a small charge as the direction of the field. We may extend this process to every point of space, and so obtain the concept of a field of force, by which every point of the space surrounding electric charges is tagged with the appropriate number and direction, the exploring charge having completely disappeared.
>
> "The field is, then, clearly a construct. Next, from the formal point of view of mathematics, it is a good construct, because there is a one to one correspondence between the electric field and the electric charges in terms of which it is defined, the field being uniquely determined by the charges, and conversely there being only one possible set of charges corresponding to a given field. Now nearly every physicist takes the next step, and ascribes physical reality to the electric field, in that he thinks that at every point of the field there is some real physical phenomenon taking place which is connected in a way not yet precisely determined with the number and direction which tag the point. At first this view most naturally involved as a corollary the existence of a medium, but lately it has become the fashion to say that the medium does not exist, and that only the field is real. ... I believe that a critical examination will show that the ascription of physical reality to the electric field is entirely without justification. I cannot find a single physical phenomenon or a single physical operation by which evidence of the existence of the field may be obtained independent of the operations which entered the definition. ...
>
> "It seems to me that any pragmatic justification in postulating reality for the electric field has now been exhausted, and that we have reached a stage where we should attempt to get closer to the actual facts by ridding the field concept of the implication of reality."

§4.9 For the equations of electromagnetism in matter, see McConnell (9), Ch. XIX, pp. 255-270.

§4.10 (a) For an alternative treatment of surfaces and curves on surfaces, see (i) Laugwitz (8), Ch. II, pp. 19-78; (ii) Kreyszig (5), Chs. III-VIII; and (iii) McConnell (9), Chs. XIV - XVI. These references also include discussions of many interesting matters not covered here, such as ruled and developable surfaces, minimal surfaces, and envelopes ; reference (ii) , in particular, includes many problems, with answers to odd-numbered exercises.

(b) Orientable surfaces are discussed in Kreyszig (5), p. 108, Synge and Schild (18), p. 261; and Davis (4), §4.5, pp. 151-160. All infinitesimal surface elements are orientable. Extended surfaces may not be orientable; the Moebius strip is a familiar example.

(c) See Note 1.15 (b) concerning projection operators.

(d) If a surface contains a straight line, that straight line is a geodesic curve on the surface. (Laugwitz (8), p. 45).

(e) A second very important relation for a surface, in addition to Gauss' equation, is the **equation of Codazzi**. We may easily derive it by noting that

$$\delta^{\mu\nu}_{\alpha\beta}\eta_{\gamma,\mu\nu} = R_{\gamma\alpha\beta\delta}\eta^{\delta}$$

$$= \eta_{\gamma,\alpha\beta} - \eta_{\gamma,\beta\alpha} = -b_{\gamma\alpha,\beta} + b_{\gamma\beta,\alpha}.$$

If, therefore, 3-dimensional space is flat, $R_{\gamma\alpha\beta\delta} = 0$ so that

$$b_{\gamma\beta,\alpha} - b_{\gamma\alpha,\beta} = 0.$$

But if $\gamma \neq 3$, then either $\gamma = \alpha$ or $\gamma = \beta$ Let $\gamma = \alpha$. Then

$$b_{\alpha\beta,\alpha} - b_{\alpha\alpha,\beta} = 0 \text{ (no summation over } \alpha\text{)}.$$

These are the **equations of Codazzi.**

(f) The fundamental theorem of the theory of surfaces states that the surface tensor $g_{\alpha\beta}$, assumed to be positive definite (i. e., $g_{\alpha\beta}dx^{\alpha}dx^{\beta} > 0$ if $dx^{\alpha} \neq 0$), and $b_{\alpha\beta}$ together determine a surface uniquely to within a rigid body displacement if they satisfy both the equations of Gauss and of Codazzi. For a more extended discussion, see Laugwitz (7), §11.07, or Sokolnikoff (14), pp. 184-186. In the former reference, it is noted (p. 131) that although the $g_{\alpha\beta}$ and $b_{\alpha\beta}$ determine a surface, these quantities are neither independent nor scalar. The problem of finding a complete set of invariants which fully characterize a surface is indeed a very difficult one, unlike the analogous problem for curves (see Note 4.2).

§4.11 (a) It is shown in Adler, Bazin, and Schiffer (1), pp. 81-83, that for any vector A_i it is identically true that

$$\delta^{ijk}_{pqr}A_{i,jk} = 0.$$

(b) It is shown in Sokolnikoff (14), pp. 92-96, that the necessary and sufficient condition that a geometry be Euclidean is that the Riemann-Christoffel tensor vanish. Schrödinger (13), pp. 43-52, relates the Riemann-Christoffel tensor to integrability about closed paths. This is the celebrated **Pfaffian problem**.

(c) Spaces of constant curvature are discussed in Synge and Schild (18), Ch. IV, and Laugwitz (8), §12.

(d) Laugwitz (8), p.127, interprets the curvature invariant in n dimensions as $n(n-1)$ times the arithmetic mean of all the Riemannian curvatures belonging to the planar directions which are obtainable from n mutually orthogonal vectors.