Chapter 2

Tensors in Generalized Coordinates in Two Dimensions

1. Non-Rectilinear Plane Coordinate Systems

We have thus far confined our consideration to coordinate systems defined by two families of parallel straight lines. One member of each family was selected as a reference line or axis. Points of the plane were located by identifying the member of each of the family of curves which passed through the point in question.

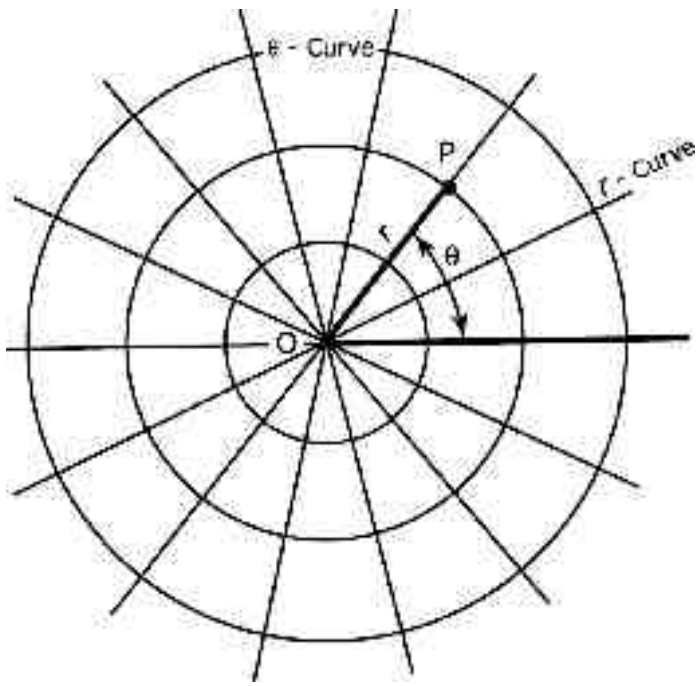

Figure 31

It is clear that a coordinate mesh may be defined equally well by any two families of curves whose members do not intersect other curves of the same family; they need not be families of parallel straight lines. Thus in polar coordinates, the two families of curves are, as in Fig. 31, (1) the family of straight lines through a point O (the origin) and (2) the family of circles about the point. Along each straight line through the origin, the angle θ is constant and the radius r alone varies; for this reason, such a curve is called an $r-$curve. Along each circle about the origin, the radius r is constant and the angle θ alone varies; for this reason, such a curve is called a $\theta-$curve. Elliptical, parabolic, bipolar and infinitely many other curvilinear or **generalized coordinates** are evidently possible.

Ex. (1.1) Given the transformation of coordinates, determine what the curves $\bar{x}^i$ = ***constant*** *are in the following curvilinear systems (*x *and* y *are ordinary Cartesian coordinates):*

$$(a) \quad \bar{x}^1 = r = \sqrt{(x)^2 + (y)^2} = [(x^1)^2 + (x^2)^2]^{1/2},$$

$$\bar{x}^2 = \theta = \cos^{-1}\left(\frac{x}{r}\right) = \cos^{-1}\frac{x^1}{\sqrt{(x^1)^2 + (x^2)^2}}$$

$$= \sin^{-1}\left(\frac{y}{r}\right) = \sin^{-1}\frac{x^2}{\sqrt{(x^1)^2 + (x^2)^2}} = \tan^{-1}\frac{y}{x} = \tan^{-1}\frac{x^2}{x^1}.$$

Ans. (a) Polar coordinates, with $r-$*curves straight lines through the origin,* $\theta-$*curves circles about the origin (see Fig. 32).*

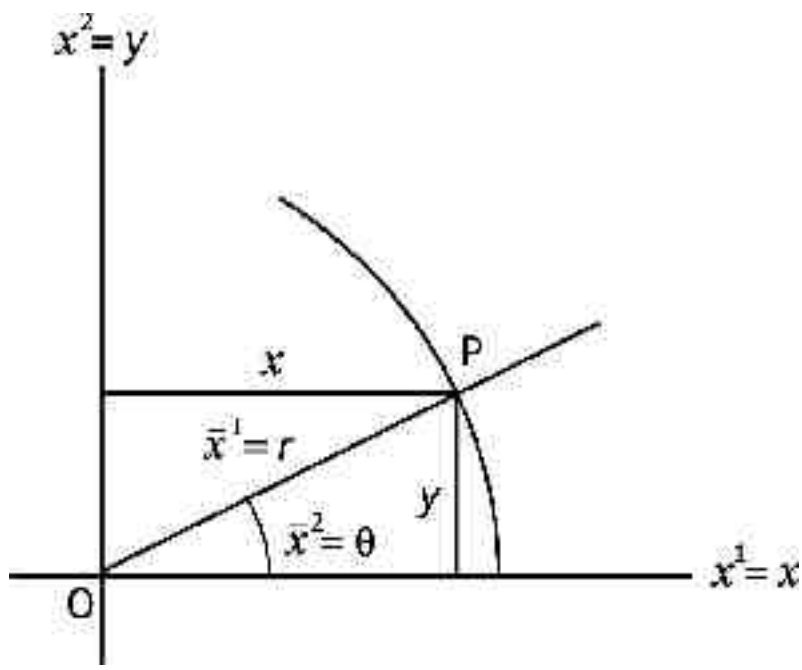

Figure 32

$$(b) \quad \bar{x}^1 = \rho = \sqrt{(x - \frac{a}{2})^2 + y^2} = \left[(x^1 - \frac{a}{2})^2 + (x^2)^2\right]^{1/2},$$

$$\bar{x}^2 = \sigma = \sqrt{(x + \frac{a}{2})^2 + (y)^2} = \left[(x^1 + \frac{a}{2})^2 + (x^2)^2\right]^{1/2}.$$

Ans. (b) Bipolar coordinates, with $\rho-$*curves circles about the point with Cartesian coordinates* $(\frac{a}{2}, 0)$ *and* $\sigma-$*curves circles about the point with Cartesian coordinates* $(-\frac{a}{2}, 0)$ *(See Fig. 33).*

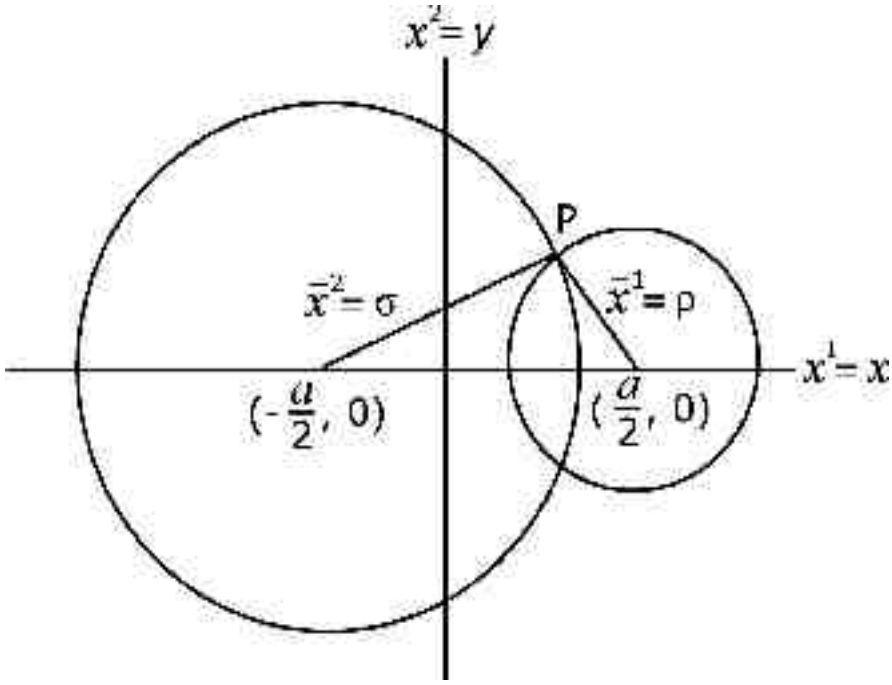

Figure 33

$$(c) \quad \bar{x}^1 = \xi = \cosh^{-1}\left(\frac{\rho + \sigma}{a}\right), \ \bar{x}^2 = \eta = \cos^{-1}\left(\frac{\sigma - \rho}{a}\right).$$

Ans. (c) Elliptical coordinates, with x^1*– curves ellipses with foci at the points with Cartesian coordinates* $(\pm \frac{a}{2}, 0)$, *and* η*– curves confocal hyperbolae (see Fig. 34).*

$$(d) \quad \bar{x}^1 = \alpha = \frac{y}{(x)^2} = x^2 (x^1)^{-2}, \ \bar{x}^2 = \beta = \frac{x}{(y^2)} = x^1 (x^2)^{-2}.$$

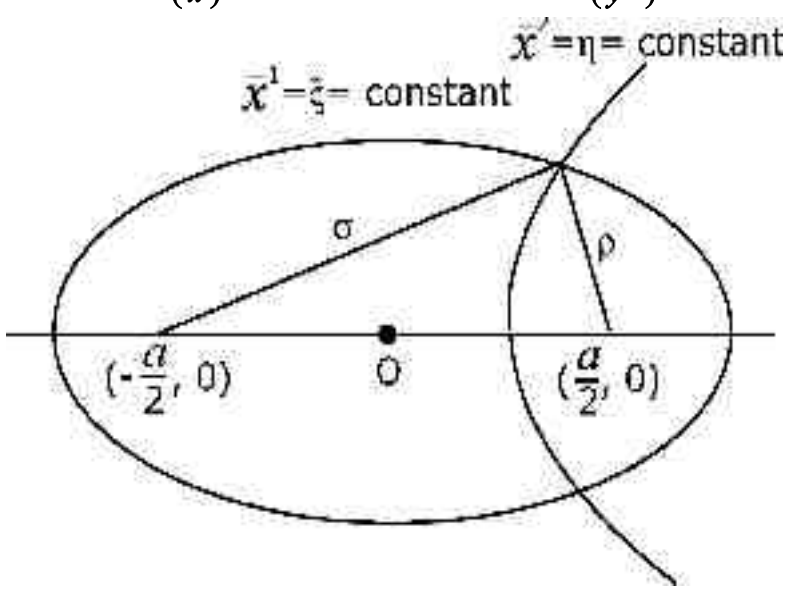

Figure 34

Ans. (d) Biparabolic coordinates, with $\alpha-$*curves parabolae tangent at their vertices to the* $x-$*axis at the origin and axes along the* $\pm y-$*axis,* $\beta-$*curves parabolae tangent to the* $y-$*axis and axes along the* $\pm x-$*axis. (Note: the positive* $x-$*axis is defined by* $\alpha = 0$, $\beta = +\infty$, $\lim\limits_{\alpha\to 0,\,\beta\to\infty} (\alpha^2\beta) = (x)^3$, *where* x *is the coordinate of a particular point on the* $x-$*axis, etc. See Fig. 35).*

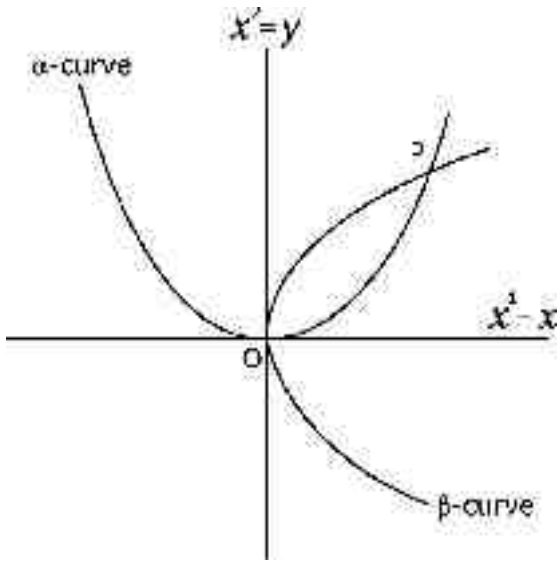

Figure 35

Ex. (1.2) Determine the inverses of the transformations in Ex. (1.1).

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

$$(b)\quad x^1 = \frac{\sigma^2 - \rho^2}{2a} = \frac{(\bar{x}^2)^2 - (\bar{x}^1)^2}{2a},$$

$$x^2 = y = \pm\frac{2}{a}\sqrt{s(s-\rho)(s-\sigma)(s-a)}, \quad 2s = \rho + \sigma + a$$

$$\text{or}\quad x^2 = \pm\frac{2}{a}\sqrt{s(s-\bar{x}^1)(s-\bar{x}^2)(s-a)} \quad (2s = \bar{x}^1 + \bar{x}^2 + a)$$

$$x^2 = \frac{1}{2a}[2a^2\rho^2 + 2a^2\sigma^2 + 2\rho^2\sigma^2 - (a^4 + \rho^4 + \sigma^4)]^{1/2}$$

$$= \frac{1}{2a}\left[2a^2\left((\bar{x}^1)^2 + (\bar{x}^2)^2\right) + 2(\bar{x}^1)^2(\bar{x}^2)^2 - (a^4 + (\bar{x}^1)^4 + (\bar{x}^2)^4)\right]^{1/2}.$$

$$(c)\quad x^1 = x = \frac{a}{2}\sinh\xi\cos\eta = \frac{a}{2}\cosh\bar{x}^1\cos\bar{x}^2,$$

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

(Hint: Use the facts that

$$\sigma + \rho = a\cosh\xi, \quad a\sinh\xi = 2\sqrt{s(s-a)},$$
$$\sigma - \rho = a\cos\eta, \quad a\sin\eta = 2\sqrt{(s-\rho)(s-\sigma)}$$

together with the results of part (b) above.)

$$(d) \quad x^1 = x = (\alpha^2\beta)^{-1/3}, \; x^2 = y = (\alpha\beta^2)^{-1/3} = [\bar{x}^1(\bar{x}^2)^2]^{-1/3}.$$

It is clear at once that there is no evident sense in which we may refer to generalized coordinates defined in such manner as being either contravariant or covariant. It might therefore seem that generalized coordinates might have to be excluded from consideration in discussing vectors and tensors. To see that such is not the case, however, consider Fig. 36. The point $\mathbf{P}$ has both rectilinear coordinates $x^i(\mathbf{P})$ and curvilinear (generalized) coordinates $\bar{x}^i(\mathbf{P})$. At $\mathbf{P}$, let the tangents to the $\bar{x}^1-$ and $\bar{x}^2-$ curves be $\mathbf{T}_{(1)}$ and $\mathbf{T}_{(2)}$, respectively. These tangents may be used to define a local rectilinear coordinate system $\bar{y}^i$ at $\mathbf{P}$. For this purpose, transfer the origin $\mathbf{O}$ to $\mathbf{P}$ and take

$$\bar{y}^j = \left(\frac{\partial \bar{x}^j}{\partial x^i}\right)_{\mathbf{P}} \left(x^i - x^i(\mathbf{P})\right), \tag{1.1}$$

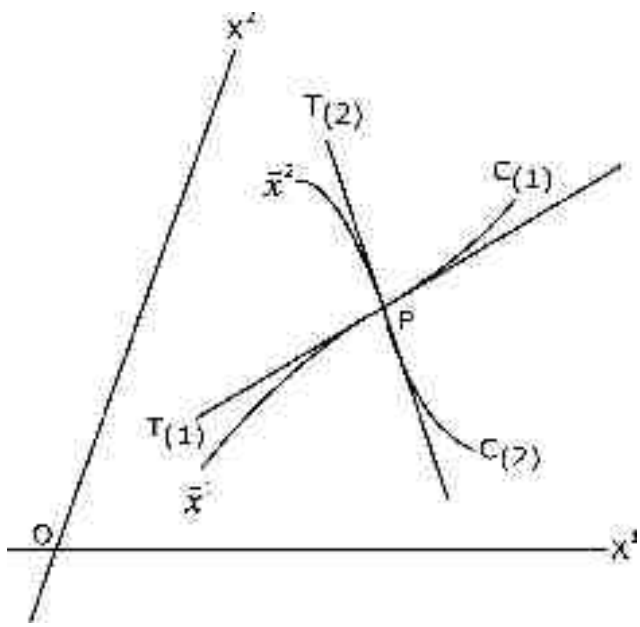

Figure 36

where the partials are evaluated at the point $\mathbf{P}$ and are therefore constants so far as the definition of the $\bar{y}^j$ is concerned. At any other point $\mathbf{P}'$, these coefficients would in general have a different set of values.

The choice of equation (1.1) as the relation for defining the new rectilinear system at $\mathbf{P}$ is suggested by the fact that

$$d\bar{y}^j = \left(\frac{\partial \bar{x}^j}{\partial x^i} \right)_{\mathrm{P}} dx^i .$$

But the expression on the right hand side is seen at once to be $d\bar{x}^j$. Therefore $d\bar{y}^j = d\bar{x}^j$. We have thus made a choice of transformations (1.1) which makes the contravariant vector $d\bar{x}^j$ identically equal to the contravariant rectilinear differential vector $d\bar{y}^j$. This suggests, particularly in view of the Product Theorem, that the components of any tensor $\mathrm{T}^{i\ldots}_{j\ldots}$ in any $\bar{x}^j$– coordinate system be given the values at P which they would have under the **rectilinear** transformation (1.1). At any point P, therefore, the vector and tensor transformation laws remain valid. Consequently, **the Quotient Theorem, the Product Theorem, and all algebraic results thus far derived remain true in generalized coordinates.**

Ex. (1.3) Given a vector $\mathrm{v}^i = (\mathrm{v}^1, \mathrm{v}^2)$ *in Cartesian coordinates, what are its components in plane polar coordinates?*

Ans. Since at every point of the plane we must have

$$\bar{\mathrm{v}}^i = \frac{\partial \bar{x}^i}{\partial x^j} \mathrm{v}^j = \frac{\partial \bar{x}^i}{\partial x^1} \mathrm{v}^1 + \frac{\partial \bar{x}^i}{\partial x^2} \mathrm{v}^2 ,$$

we have only to substitute the values of the partial derivatives. These are

$$\frac{\partial \bar{x}^1}{\partial x^1} = \frac{\partial r}{\partial x} = \frac{x}{r} = \cos\theta , \quad \frac{\partial \bar{x}^1}{\partial x^2} = \frac{\partial r}{\partial y} = \frac{y}{r} = \sin\theta ,$$

$$\frac{\partial \bar{x}^2}{\partial x^1} = \frac{\partial \theta}{\partial x} = -\frac{\sin\theta}{r} , \quad \frac{\partial \bar{x}^2}{\partial x^2} = \frac{\partial \theta}{\partial y} = \frac{\cos\theta}{r} .$$

Therefore

$$\bar{\mathrm{v}}^1 = (\cos\theta)\mathrm{v}^1 + (\sin\theta)\mathrm{v}^2 , \quad \bar{\mathrm{v}}^2 = -\left(\frac{\sin\theta}{r}\right)\mathrm{v}^1 + \left(\frac{\cos\theta}{r}\right)\mathrm{v}^2 .$$

Ex. (1.4) (a) What are the components in polar coordinates of the vector whose Cartesian components are (24, 7)? (b) What are the components in polar coordinates of the vector whose Cartesian components are $p^i = (x, y)$*? (c) What are these when* $p^i = (7, 24)$*? Check the results of (a) against those of (b). (d) Interpret these components geometrically.*

Ans. $(a)\ \bar{\mathrm{v}}^1 = 24\cos\theta + 7\sin\theta , \quad \bar{\mathrm{v}}^2 = -24\frac{\sin\theta}{r} + 7\frac{\cos\theta}{r} ;$

$(b)\ \bar{p}^1 = (\cos\theta)x + (\sin\theta)y = r\cos^2\theta + r\sin^2\theta = r ;$

$(c)\ \bar{p}^2 = -\left(\frac{\sin\theta}{r}\right)x + \left(\frac{\cos\theta}{r}\right)y = -\sin\theta\cos\theta + \sin\theta\cos\theta = 0 .$

Note here a somewhat subtle point. In (a) we did not find it necessary to state at what point of the plane the Cartesian components of $\mathbf{v}^i$ *were* (24,7), *because in rectilinear coordinates the same components represent the same vector at* ***all*** *points. Not so in polar coordinates, where the components depend upon the coordinates* $\bar{x}^i = (r, \theta)$.

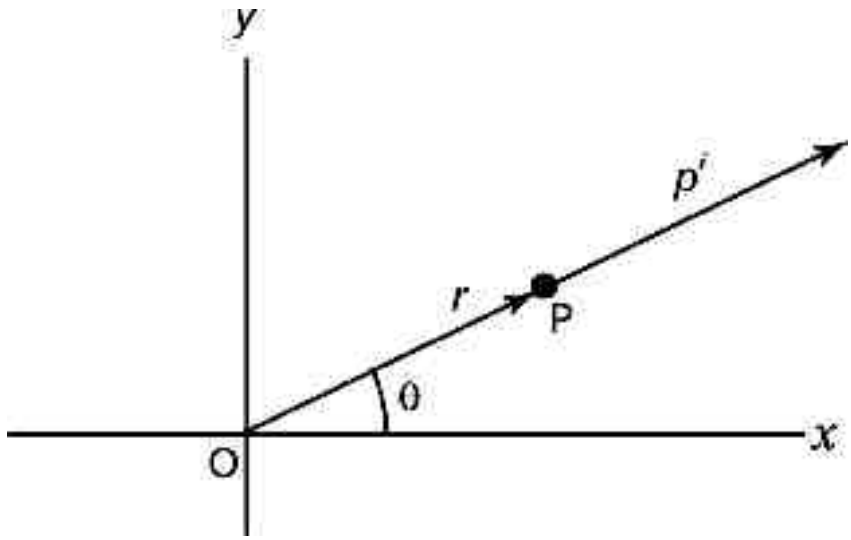

Figure 37

This is brought out more clearly in (b), where we consider a vector $p^i = (x, y)$ ***at the point*** $x^i = (x, y)$. *The same vector* p^i, *being independent of location in Cartesian coordinates, also has the same components* (x, y) *even at the point* $x^i = (0, 0)$, *the origin. It is important to recognize (cf. Ex. (1.3)) that this is not true in polar coordinates where, in fact, at the origin we have no tangent curves and therefore no projections defined.*

This has a bearing on the common practice of regarding the position vector of the point P *to be the directed line segment* $\overline{\mathrm{OP}}$. *This is permissible in Cartesian (or rectilinear) coordinates. In polar coordinates* $\bar{x}^i$, *however, the position vector is more properly* $p^i = (x, y)$ *or* $\bar{p}^i = (r, 0)$ *associated with* ***and located at*** *the point* P *with coordinates* $x^i = (x, y)$ *or* $\bar{x}^i = (r, \theta)$. *No harm comes from failure to make this distinction but much puzzlement may result if it is overlooked.*

(c) $\bar{p}^i = (25, 0)$.

(d) The vector $\bar{p}^i$ *has a projection of length* 25 *parallel to the* r*-curve (radial line) and zero projection parallel to the* θ*-curve (i. e., the tangent to the circle through* P *about* O*) hence perpendicular to the radius.*

Ex. (1.5) Given a vector $\mathbf{v}_i = (\mathbf{v}_1, \mathbf{v}_2)$ *in Cartesian coordinates, what are its components in elliptical coordinates?*

Ans. Since $\bar{\mathrm{v}}_i = \dfrac{\partial x^j}{\partial \bar{x}^i}\mathrm{v}_j$, *and since (see Ex. (1.2))*

$$\frac{\partial x^1}{\partial \bar{x}^1} = \frac{\partial x}{\partial \xi} = \frac{a}{2}\sinh\xi\cos\eta\,, \quad \frac{\partial x^2}{\partial \bar{x}^1} = \frac{\partial y}{\partial \xi} = \frac{a}{2}\cosh\xi\sin\eta\,,$$

$$\frac{\partial x^1}{\partial \bar{x}^2} = \frac{\partial x}{\partial \eta} = -\frac{a}{2}\cosh\xi\cos\eta\,, \quad \frac{\partial x^2}{\partial \bar{x}^2} = \frac{\partial y}{\partial \eta} = \frac{a}{2}\sinh\xi\cos\eta\,,$$

we have that

$$\bar{\mathrm{v}}_1 = \frac{a}{2}(\sinh\xi\cos\eta)\mathrm{v}_1 + \frac{a}{2}(\cosh\xi\sin\eta)\mathrm{v}_2\,,$$

$$\bar{\mathrm{v}}_2 = -\frac{a}{2}(\cosh\xi\sin\eta)\mathrm{v}_1 + \frac{a}{2}(\sinh\xi\cos\eta)\mathrm{v}_2\,.$$

Ex. (1.6) (a) What are the components in elliptical coordinates of the vector whose Cartesian components are $\mathrm{v}_i = \left(\dfrac{225}{574}a, \dfrac{960}{574}a\right)$ *? (b) Of the position vector whose Cartesian components are* $p_i = (x, y)$ *? (c) What are these when* $p_i = \left(\dfrac{225}{574}a, \dfrac{960}{574}a\right)$ *? Use the results of (a) and (b) to check (c).*

Ans. (a) $\quad \bar{\mathrm{v}}_1 = \dfrac{225}{1148}a^2\sinh\xi\cos\eta + \dfrac{960}{1148}a^2\cosh\xi\sin\eta\,,$

$$\bar{\mathrm{v}}_2 = -\frac{225}{1148}a^2\cosh\xi\sin\eta + \frac{960}{1148}a^2\sinh\xi\cos\eta\,.$$

(b) $\quad \bar{\mathrm{v}}_1 = \dfrac{a}{2}\sinh\xi\cosh\eta\,, \quad \bar{\mathrm{v}}_2 = -\dfrac{a}{2}\sin\xi\cos\eta\,.$

(c) $\quad \bar{\mathrm{v}}_i = \left(\dfrac{50\,a^2}{49}, \dfrac{90\,a^2}{1681}\right).$

Ex. (1.7) (a) Find the components of the fundamental tensor in polar coordinates. (b) determine from the result the angle between the coordinate curves and the scale factor along each curve. (c) Determine the components of the contravariant fundamental tensor.

Ans. (a) We have

$$\bar{g}_{ij} = \frac{\partial x^k}{\partial \bar{x}^i}\frac{\partial x^l}{\partial \bar{x}^j}\delta_{kl} = \frac{\partial x^1}{\partial \bar{x}^i}\frac{\partial x^1}{\partial \bar{x}^j} + \frac{\partial x^2}{\partial \bar{x}^i}\frac{\partial x^2}{\partial \bar{x}^j}\,.$$

Therefore, making use of the results of Ex. (1.2), we find that

$$\frac{\partial x^1}{\partial \bar{x}^1} = \frac{\partial x}{\partial r} = \cos\theta\,, \quad \frac{\partial x^1}{\partial \bar{x}^2} = \frac{\partial x}{\partial \theta} = -\,r\sin\theta\,,$$

$$\frac{\partial x^2}{\partial \bar{x}^1} = \frac{\partial y}{\partial r} = \sin\theta\,, \quad \frac{\partial x^2}{\partial \bar{x}^2} = \frac{\partial y}{\partial \theta} = r\cos\theta\,.$$

Hence

$$\bar{g}_{11} = (\cos\theta)^2 + (\sin\theta)^2 = 1\,,$$

$$\bar{g}_{12} = (\cos\theta)(-r\sin\theta) + (\sin\theta)(r\cos\theta) = 0 = \bar{g}_{21}\,,$$

$$\bar{g}_{22} = (-r\sin\theta)^2 + (r\cos\theta)^2 = r^2\,.$$

(b) The coordinate curves are orthogonal and $s_{(1)} = 1\,,\ s_{(2)} = r\,.$

(c) $\bar{g}^{11} = 1\,,\ \ \bar{g}^{12} = 0 = \bar{g}^{21},\ \ \bar{g}^{22} = \dfrac{1}{r^2}\,.$

Ex. (1.8) (a) Find the components of the fundamental tensor in bipolar coordinates. (b) Find from these the components of the contravariant fundamental tensor. (c) What is the expression for the line element in this coordinate system?

Ans. (a) Calling the angle opposite the base α *and using the relations*

$$\cos\alpha = \frac{\rho^2 + \sigma^2 - a^2}{2\rho\sigma}\,,$$

$$\sin\alpha = \frac{1}{2\rho\sigma}\left[2(a^2\rho^2 + a^2\sigma^2 + \rho^2\sigma^2) - (a^4 + \rho^4 + \sigma^4)\right]^{1/2},$$

we find for the fundamental tensor the components

$$\bar{g}_{ij} = \frac{1}{\sin^2\alpha}\begin{Vmatrix} 1 & -\cos\alpha \\ -\cos\alpha & 1 \end{Vmatrix}.$$

$$(b) \qquad \bar{g}^{ij} = \begin{Vmatrix} 1 & -\cos\alpha \\ -\cos\alpha & 1 \end{Vmatrix}.$$

$$(c) \qquad (ds)^2 = \frac{1}{\sin^2\alpha}\left[(d\rho)^2 - 2(\cos\alpha)\,d\rho\,d\sigma + (d\sigma)^2\right].$$

Ex. (1.9) Given that the line element in a particular coordinate system is

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

what are the components of the fundamental tensor?

Ans. By inspection of the coefficients of $d\bar{x}^i\, d\bar{x}^j$ *we see at once that they are*

$$\bar{g}_{ij} = \begin{Vmatrix} (u)^2 + (\mathrm{v})^2 & 0 \\ 0 & (u)^2 + (\mathrm{v})^2 \end{Vmatrix}.$$

Note: these coordinates are commonly called ***parabolic coordinates****, related to Cartesian coordinates by the transformation*

$$u = \pm\sqrt{x + \sqrt{x^2 + y^2}}, \quad \mathrm{v} = \sqrt{-x + \sqrt{x^2 + y^2}},$$

$$-\infty < u < +\infty, \quad \mathrm{v} \geq 0,$$

$$x = \frac{1}{2}(u^2 - \mathrm{v}^2), \quad y = u\mathrm{v}.$$

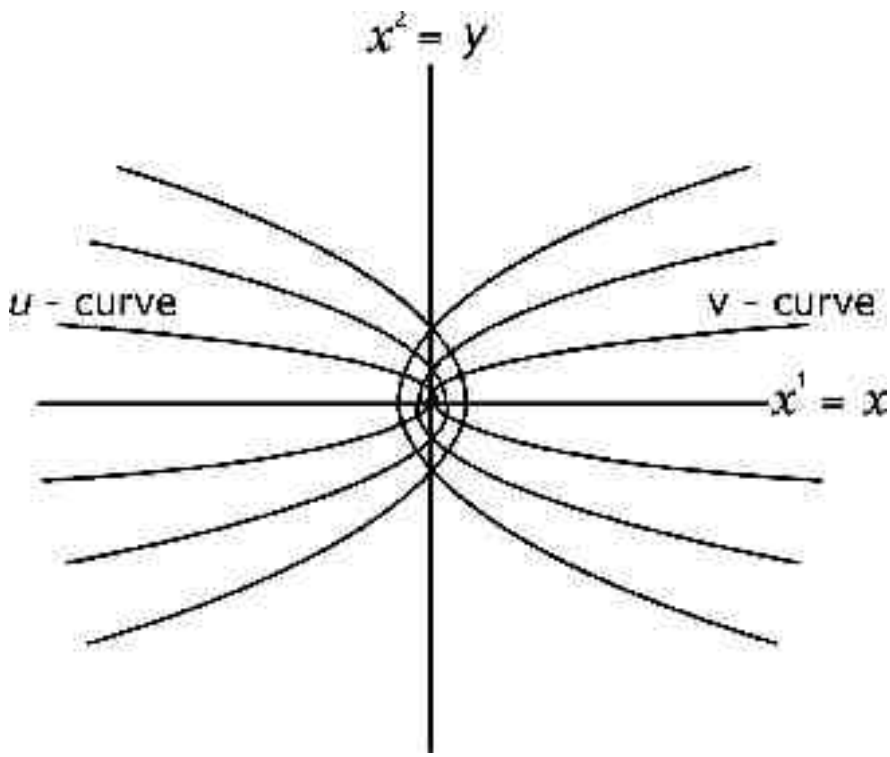

Figure 38

The curves $\mathrm{v} = \textbf{\textit{constant}}$ *are parabolae opening to the positive* $x-$*axis, the curves* $u = \textbf{\textit{constant}}$ *are parabolae opening to the negative* $x-$*axis, with* u *having the sign of* y*. The origin is the common focus. (See Fig. 38). The coordinate grid in the variables* $\bar{x}^i = (u, \mathrm{v})$ *has the interesting property that the set of points* $(2m, 2n)$ *are all at integral distance from the origin if m and n are arbitrary integers.*

Ex. (1.10) Show that if $\varphi(x^i)$ *is an invariant function of the coordinates* x^i, *then* $\dfrac{\partial\varphi}{\partial x^i}$ *is a covariant vector.*

Ans. Since $\varphi(\bar{x}^i) = \varphi(x^j(\bar{x}^i))$, *by the rule for the differentiation of a function of a function we have*

$$\frac{\partial\varphi}{\partial\bar{x}^i} = \frac{\partial\varphi}{\partial x^j}\frac{\partial x^j}{\partial\bar{x}^i},$$

which is the covariant transformation law applied to the vector $g_j = \dfrac{\partial\varphi}{\partial x^j}$.

Ex. (1.11) Given the function $\varphi(r,\theta) = r^2[1 - e^2\cos^2\theta]$ *in a polar coordinate system* $(0 \le e < 1)$. *(a) Find the components of* $\dfrac{\partial\varphi}{\partial x^i} = \varphi_{,i}$.

(b) Find the unit vector in the direction of $\varphi_{,i}$. *(c) What are the contravariant components of this unit vector?*

Ans. (a) $\dfrac{\partial\varphi}{\partial x^i} = (2r[1 - e^2\cos^2\theta],\ -2e^2r^2\sin\theta\cos\theta)$.

(b) $n_i = \dfrac{\varphi_{,i}}{|\varphi_{,i}|} = \dfrac{1}{\mathrm{R}}(1 - e^2\cos^2\theta,\ -re^2\sin\theta\cos\theta)$, *where*

$$\mathrm{R} = [1 - (2e^2 - e^4)\cos^2\theta]^{1/2}.$$

(c) $n^i = \dfrac{1}{\mathrm{R}}\left(1 - e^2\cos^2\theta,\ -\dfrac{e^2\sin\theta\cos\theta}{r}\right)$.

Ex. (1.12) Show that n^i *is the unit normal to the surface* φ = *constant. (Hint:* $d\varphi = \varphi_{,i}dx^i = 0$ *when* dx^i *lies in the surface.)*

We must always remember that the coefficients $\dfrac{\partial\bar{x}^i}{\partial x^j}$ **and** $\dfrac{\partial x^i}{\partial\bar{x}^j}$ **of the** contravariant and covariant transformations, respectively, will in general vary from point to point. This means that in general or curvilinear coordinates, vectors and tensors are transformed to axes whose orientations vary and along which scales of length may change as a function of position. We may therefore say, in a sense, that vectors are always defined in rectilinear coordinates and only in such, but in other coordinate systems it is possible to relate such rectilinear systems to curvilinear systems in a unique and intimate way. Indeed, this is done in such a subtle but straightforward way that we speak somewhat loosely of vectors and tensors in curvilinear coordinates. It is obvious that being able to express vectors and tensors in any desired coordinate system vastly augments their convenience and utility.

2. Intrinsic and Covariant Differentiation of a Contravariant Vector

Let us consider a vector field $\mathbf{v}^i$ defined along a curve $x^i(t)$; the vector might represent a velocity field, for example, and t the time. We may ask whether the derivative of $\mathbf{v}^i$ with respect to t is also a vector. The necessary and sufficient condition, as we have seen, is that the law of vector transformation be satisfied. To determine whether or not it is, we start from the fact that $\mathbf{v}^i$ itself is a vector, i.e., that it satisfies for all values of t the relation

$$\bar{\mathbf{v}}^j = \frac{\partial \bar{x}^j}{\partial x^i}\mathbf{v}^i .$$

Differentiating both sides, we see that

$$\text{(2.1)} \qquad \frac{d\bar{\mathbf{v}}^j}{dt} = \frac{\partial \bar{x}^j}{\partial x^i}\frac{d\bar{x}^i}{dt} + \mathbf{v}^i \frac{d}{dt}\left(\frac{\partial \bar{x}^j}{\partial x^i}\right) = \frac{\partial \bar{x}^j}{\partial x^i}\frac{d\mathbf{v}^i}{dt} + \frac{\partial^2 \bar{x}^j}{\partial x^k \partial x^i}\mathbf{v}^i \frac{dx^k}{dt} .$$

Clearly, this equation is disqualified as a vector transformation for $d\mathbf{v}^i/dt$ unless the second term on the right is zero no matter what $\mathbf{v}^i$ or dx^i/dt may be. This means that we must have

$$\frac{\partial^2 \bar{x}^j}{\partial x^k \partial x^i} = 0 \text{ or } \frac{\partial \bar{x}^j}{\partial x^i} = \text{constant},$$

a condition which is satisfied only in rectilinear coordinates. Thus, as we have seen, the derivatives of vectors in rectilinear coordinates are vectors whereas the derivatives of vectors in curvilinear coordinates are not vectors.

Though equation (2.1) is not a vector equation in curvilinear coordinates we may nevertheless devise an artifice by which from equation (2.1) a vector equation may be constructed. To this end, let us define a set of non-tensor quantities $\bar{\Gamma}^i_{kl}$ which obey a non-tensor transformation law of the form

$$\text{(2.2)} \qquad \bar{\Gamma}^j_{kl}\bar{\mathbf{v}}^k \frac{d\bar{x}^l}{dt} = \frac{\partial \bar{x}^j}{\partial x^i}\Gamma^i_{pq}\mathbf{v}^p \frac{dx^q}{dt} - \frac{\partial^2 \bar{x}^j}{\partial x^k \partial x^i}\mathbf{v}^i \frac{dx^k}{dt} .$$

Our purpose in proposing such a set of quantities is transparently that of eliminating the second term on the right in equation (2.1) and at the same time putting the remaining terms into the form required by the law of vector transformation. That we have achieved such a goal is evident simply by adding equations (2.1) and (2.2). The result is

$$\frac{d\bar{\mathbf{v}}^j}{dt} + \bar{\Gamma}^j_{kl}\bar{\mathbf{v}}^k \frac{d\bar{x}^l}{dt} = \frac{\partial \bar{x}^j}{\partial x^i}\left[\frac{d\mathbf{v}^i}{dt} + \Gamma^i_{pq}\mathbf{v}^p \frac{dx^q}{dt}\right].$$

Clearly, the bracketed quantity

$$\text{(2.3)} \qquad \frac{\delta \mathbf{v}^i}{\delta t} \equiv \frac{d\mathbf{v}^i}{dt} + \Gamma^i_{pq}\mathbf{v}^p \frac{dx^q}{dt}$$

satisfies the law of vector transformation and is therefore a vector. It is called the **intrinsic derivative of $\mathbf{v}^i$ with respect to t**; it is distinguished notationally from the ordinary derivative $d\mathbf{v}^i/dt$ by writing it as $\delta\mathbf{v}^i/\delta t$. Note that the intrinsic derivative of each component of a vector is a function of *all* the components of the vector, not alone of the given component.

If $\mathbf{v}^i$ does not depend explicitly on t, then

$$\frac{d\mathbf{v}^i}{dt} = \frac{\partial \mathbf{v}^i}{\partial x^q}\frac{dx^q}{dt},$$

whence equation (2.3) may be written as

$$\frac{\delta \mathbf{v}^i}{\delta t} = \left[\frac{\partial \mathbf{v}^i}{\partial x^q} + \Gamma^i_{pq}\mathbf{v}^p\right]\frac{dx^q}{dt} = \mathbf{v}^i_{,q}\frac{dx^q}{dt}.$$

Since the left hand side, the intrinsic derivative of $\mathbf{v}^i$ with respect to t, is a vector, and since dx^q/dt is an arbitrary vector, the quantity in square brackets is, by the Quotient Theorem, a mixed tensor. Thus the quantity

$$\mathbf{v}^i_{,q} \equiv \frac{\partial \mathbf{v}^i}{\partial x^q} + \Gamma^i_{pq}\mathbf{v}^p \tag{2.4}$$

is a mixed tensor formed from the vector $\mathbf{v}^i$ and its partial derivatives; it is called **the covariant derivative of $\mathbf{v}^i$ with respect to x^q**. A comma before a subscript indicates covariant differentiation with respect to the subscript(s) which follow(s) it.

Before turning to the problem of how the quantities Γ^i_{pq} ought to be chosen, let us note that equation (2.2) which relates the values of these quantities in different coordinate systems was written in the form best suited for the immediate purpose at hand. It is, however, a form which may be simplified by eliminating from it the vectors $\mathbf{v}^i$ and dx^k/dt. Thus, using the vector transformation law on $\bar{\mathbf{v}}^k$ and $d\bar{x}^k/dt$ and replacing summation indices i and k in the last term on the right by p and q, we have

$$\left[\bar{\Gamma}^j_{kl}\frac{\partial \bar{x}^k}{\partial x^p}\frac{\partial \bar{x}^l}{\partial x^q} - \frac{\partial \bar{x}^j}{\partial x^i}\Gamma^i_{pq} + \frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p}\right]\mathbf{v}^p\frac{dx^q}{dt} = 0$$

for arbitrary $\mathbf{v}^p$ and dx^q/dt. These equations can be true, then, only if each quantity within the square brackets vanishes. Therefore

$$\bar{\Gamma}^j_{kl}\frac{\partial \bar{x}^k}{\partial x^p}\frac{\partial \bar{x}^l}{\partial x^q} + \frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} = \frac{\partial \bar{x}^j}{\partial x^i}\Gamma^i_{pq}.$$

Multiplying both sides by $\partial x^m/\partial \bar{x}^j$,

$$\frac{\partial x^m}{\partial \bar{x}^j}\frac{\partial \bar{x}^j}{\partial x^i}\Gamma^i_{pq} = \delta^m_i\Gamma^i_{pq} = \Gamma^m_{pq} = \frac{\partial x^m}{\partial \bar{x}^j}\frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} + \bar{\Gamma}^j_{kl}\frac{\partial x^m}{\partial \bar{x}^j}\frac{\partial \bar{x}^k}{\partial x^p}\frac{\partial \bar{x}^i}{\partial x^{qx}}. \tag{2.5}$$

It is apparent that if the $\bar{\Gamma}^j_{kl}$ are known in the $\bar{x}^i$– coordinate system, their counterparts in a new x^i– coordinate system may be found from equation (2.5).

It is now clear that the first two terms in equation (2.2) had to include, of necessity, factors $\bar{\mathrm{v}}^k\, d\bar{x}^l/dt$ and $\mathrm{v}^p\, dx^q/dt$, respectively. Only then would the transformation of the Γ^i_{jk} not depend upon the particular vector being differentiated; i. e., it would not otherwise have been possible to derive a result such as equation (2.5) free of v^p and dx^q/dt.

Let us now turn to the problem of how the Γ^i_{jk} ought to be chosen. In principle, their choice in some particular coordinate system is completely arbitrary; in any other coordinate system they are determined by equation (2.5). In practice, however, it will clearly serve the ends of simplicity and convenience if we choose the Γ^i_{jk} so that the intrinsic derivative of a contravariant vector in rectilinear coordinates is identical with its ordinary derivative, which we have previously found to be a vector. It should be expressly understood that such a choice is not necessary, only convenient. In that case, in a system of rectilinear coordinates $\bar{x}^i$,

$$\frac{\delta \bar{\mathrm{v}}^i}{\delta t} = \frac{d\bar{\mathrm{v}}^i}{dt},$$

from which it is apparent that

$$\bar{\Gamma}^j_{kl} = 0, \text{ all } j, k, l. \tag{2.6}$$

Hence, by equation (2.5), in a system of curvilinear coordinates we must have

$$\Gamma^m_{pq} = \frac{\partial x^m}{\partial \bar{x}^j} \frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} \equiv \left\{ \begin{matrix} m \\ p\ q \end{matrix} \right\}. \tag{2.7}$$

With this particular choice of $\bar{\Gamma}^m_{pq}$ in rectilinear coordinate systems (equation (2.6)),we are led to form (2.7) in generalized coordinates, in which the quantities $\left\{ \begin{matrix} m \\ p\ q \end{matrix} \right\}$ are called the **Christoffel symbols of the second kind.**

The Christoffel symbols bear a close relation to the fundamental tensor. To demonstrate this, consider first the first factor on the right hand side of equation (2.7). Let us write it as

$$\frac{\partial x^m}{\partial \bar{x}^j} = \delta^m_{\mathrm{v}} \frac{\partial x^{\mathrm{v}}}{\partial \bar{x}^j} = g^{mn} g_{n\mathrm{v}} \frac{\partial x^{\mathrm{v}}}{\partial \bar{x}^j} = g^{mn} \left(\bar{\delta}_{kl} \frac{\partial \bar{x}^k}{\partial x^n} \frac{\partial \bar{x}^l}{\partial x^{\mathrm{v}}} \right) \frac{\partial x^{\mathrm{v}}}{\partial \bar{x}^j} \tag{2.8}$$

$$= g^{mn} \bar{\delta}_{kl} \frac{\partial \bar{x}^k}{\partial x^n} \delta^l_j = g^{mn} \bar{\delta}_{kj} \frac{\partial \bar{x}^k}{\partial x^n}.$$

Substituting this into equation (2.7) gives

$$\left\{ \begin{matrix} m \\ p\ q \end{matrix} \right\} = g^{mn} \left(\bar{\delta}_{kj} \frac{\partial \bar{x}^k}{\partial x^n} \frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} \right). \tag{2.9}$$

Since $g_{np} = \bar{\delta}_{kj} \dfrac{\partial \bar{x}^k}{\partial x^n} \dfrac{\partial \bar{x}^j}{\partial x^p}$, the quantity in parenthesis is

$$\bar{\delta}_{kj} \frac{\partial \bar{x}^k}{\partial x^n} \frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} = \frac{\partial g_{np}}{\partial x^q} - \bar{\delta}_{kj} \frac{\partial^2 \bar{x}^k}{\partial x^n \partial x^q} \frac{\partial \bar{x}^j}{\partial x^p}$$

$$= \frac{\partial g_{np}}{\partial x^q} - \left(\frac{\partial g_{pq}}{\partial x^n} - \bar{\delta}_{kj} \frac{\partial^2 \bar{x}^j}{\partial x^p \partial x^n} \frac{\partial \bar{x}^k}{\partial x^q} \right)$$

$$= \frac{\partial g_{np}}{\partial x^q} - \frac{\partial g_{pq}}{\partial x^n} + \left(\frac{\partial g_{qn}}{\partial x^p} - \bar{\delta}_{kj} \frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} \frac{\partial \bar{x}^k}{\partial x^n} \right).$$

Taking the last term to the left hand side and dividing by 2,

$$\bar{\delta}_{kj} \frac{\partial \bar{x}^k}{\partial x^n} \frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} = \frac{1}{2} \left(\frac{\partial g_{np}}{\partial x^q} + \frac{\partial g_{qn}}{\partial x^p} - \frac{\partial g_{pq}}{\partial x^n} \right) \tag{2.10}$$

$$\equiv [pq, n] = [qp, n].$$

The quantity $[pq, n]$ is called the **Christoffel symbol of the first kind**. From equations (2.9) and (2.10) we infer between the Christoffel symbols of the first and second kinds the relation

$$\left\{ \begin{matrix} m \\ p\ q \end{matrix} \right\} = \left\{ \begin{matrix} m \\ q\ p \end{matrix} \right\} = g^{mn} [pq, n]. \tag{2.11}$$

Since the $\left\{ \begin{matrix} m \\ p\ q \end{matrix} \right\}$ are *not* tensors whereas the g^{mn} are, it may be concluded that the $[pq, n]$ are also *not* tensors. Equations (2.10) and (2.11) provide a relatively simple and direct way of determining the Christoffel symbols in any coordinate system.

Ex. (2.1) Show quite generally that

$$\left\{ \begin{matrix} m \\ p\ q \end{matrix} \right\} = \frac{\partial x^m}{\partial \bar{x}^j} \frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} + \overline{\left\{ \begin{matrix} j \\ k\ l \end{matrix} \right\}} \frac{\partial x^m}{\partial \bar{x}^j} \frac{\partial \bar{x}^k}{\partial x^p} \frac{\partial \bar{x}^l}{\partial x^q},$$

and hence we may choose $\Gamma^m_{pq} = \left\{ \begin{matrix} m \\ p\ q \end{matrix} \right\}$, *regardless of whether or not we start from a rectilinear coordinate system.*

Ans. Let x^i *and* $\bar{x}^i$ *be any two coordinate systems connected by a reversible transformation; neither need be rectilinear. Then, replacing* $\bar{\delta}_{kl}$ *by* $\bar{g}_{kl}$ *in equation (2.8), we have that*

$$\frac{\partial x^m}{\partial \bar{x}^j} = g^{mn} \bar{g}_{kj} \frac{\partial \bar{x}^k}{\partial x^n},$$

hence by definition

$$\frac{\partial x^m}{\partial \bar{x}^j}\frac{\partial^2 \bar{x}^j}{\partial x^p \partial x^q} = g^{mn}\bar{g}_{kj}\frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p}.$$

Now

$$g_{np} = \bar{g}_{kj}\frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial \bar{x}^j}{\partial x^p},$$

so that

$$\frac{\partial g_{np}}{\partial x^q} = \bar{g}_{kj}\frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} + \bar{g}_{kj}\frac{\partial \bar{x}^j}{\partial x^p}\frac{\partial^2 \bar{x}^k}{\partial x^n \partial x^q} + \frac{\partial \bar{g}_{kj}}{\partial x^q}\frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial \bar{x}^j}{\partial x^p}.$$

If we now permute indices, substituting n for p, p for q, and q for n, we get a second equation

$$\frac{\partial g_{qn}}{\partial x^p} = \bar{g}_{kj}\frac{\partial \bar{x}^k}{\partial x^q}\frac{\partial^2 \bar{x}^j}{\partial x^p \partial x^n} + \bar{g}_{kj}\frac{\partial \bar{x}^j}{\partial x^n}\frac{\partial^2 \bar{x}^k}{\partial x^q \partial x^p} + \frac{\partial \bar{g}_{kj}}{\partial x^p}\frac{\partial \bar{x}^k}{\partial x^q}\frac{\partial \bar{x}^j}{\partial x^n}$$

Permuting once more, substituting in this equation n for p, q for n, and p for q, we have a third equation

$$\frac{\partial g_{pq}}{\partial x^n} = \bar{g}_{kj}\frac{\partial \bar{x}^k}{\partial x^p}\frac{\partial^2 \bar{x}^j}{\partial x^n \partial x^q} + \bar{g}_{kj}\frac{\partial \bar{x}^j}{\partial x^q}\frac{\partial \bar{x}^k}{\partial x^p \partial x^n} + \frac{\partial \bar{g}_{kj}}{\partial x^n}\frac{\partial \bar{x}^k}{\partial x^p}\frac{\partial \bar{x}^j}{\partial x^q}.$$

Adding half the first two and subtracting half the third, we get

$$\frac{1}{2}\left[\frac{\partial g_{np}}{\partial x^q} + \frac{\partial g_{qn}}{\partial x^p} - \frac{\partial g_{pq}}{\partial x^n}\right] \equiv [pq, n]$$

$$= \frac{1}{2}\bar{g}_{kj}\left[\frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} + \frac{\partial \bar{x}^j}{\partial x^n}\frac{\partial^2 \bar{x}^k}{\partial x^q \partial x^p}\right]$$

$$+ \frac{1}{2}\frac{\partial \bar{g}_{kj}}{\partial \bar{x}^r}\left[\frac{\partial \bar{x}^r}{\partial x^q}\frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial \bar{x}^j}{\partial x^p} + \frac{\partial \bar{x}^r}{\partial x^p}\frac{\partial \bar{x}^k}{\partial x^q}\frac{\partial \bar{x}^j}{\partial x^n} - \frac{\partial \bar{x}^r}{\partial x^n}\frac{\partial \bar{x}^k}{\partial x^p}\frac{\partial \bar{x}^j}{\partial x^q}\right].$$

Because the first term on the right hand side is symmetric in the summation indices k and j , the two terms in the square bracket may be coalesced. Within the second square bracket on the right hand side, we replace in the second term k by r, j by k, and r by j, whereas in the third term we substitute k for r, j for k, and r for j. Then the preceding equation becomes

$$[pq, n] = \bar{g}_{kj}\frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} + \overline{[jr, k]}\frac{\partial \bar{x}^j}{\partial x^p}\frac{\partial \bar{x}^r}{\partial x^q}\frac{\partial \bar{x}^k}{\partial x^n}.$$

Multiplying this equation by $g^{mn} = \dfrac{\partial x^m}{\partial \bar{x}^a}\dfrac{\partial x^n}{\partial \bar{x}^b}\bar{g}^{ab}$ *gives*

$$g^{mn}[pq,n] \equiv \begin{Bmatrix} m \\ pq \end{Bmatrix} = \frac{\partial x^m}{\partial \bar{x}^a}\frac{\partial x^n}{\partial \bar{x}^b}\bar{g}^{ab}\left(\bar{g}_{kj}\frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} + \overline{[jr,k]}\frac{\partial \bar{x}^r}{\partial x^q}\frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial \bar{x}^j}{\partial x^p}\right)$$

$$= \frac{\partial x^m}{\partial \bar{x}^a}\left\{\left[\frac{\partial x^n}{\partial \bar{x}^b}\frac{\partial \bar{x}^k}{\partial x^n}\bar{g}_{kj}\right]\bar{g}^{ab}\right\}\frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p}$$

$$+ \overline{[jr,k]}\left\{\left(\frac{\partial x^n}{\partial \bar{x}^b}\frac{\partial \bar{x}^k}{\partial x^n}\right)\bar{g}^{ab}\right\}\frac{\partial x^m}{\partial \bar{x}^a}\frac{\partial \bar{x}^r}{\partial x^q}\frac{\partial \bar{x}^j}{\partial x^p}$$

$$= \frac{\partial x^m}{\partial \bar{x}^j}\frac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} + \overline{\begin{Bmatrix} a \\ j\ r \end{Bmatrix}}\frac{\partial x^m}{\partial \bar{x}^a}\frac{\partial \bar{x}^j}{\partial x^p}\frac{\partial \bar{x}^r}{\partial x^r},$$

as was to be shown.

Ex. (2.2) Derive the equation of transformation between Christoffel symbols of the first kind. (Hint: See Ex. (2.1)).

Ans. $[pq,n] = \bar{g}_{kj}\dfrac{\partial \bar{x}^k}{\partial x^n}\dfrac{\partial^2 \bar{x}^j}{\partial x^q \partial x^p} + \overline{[jr,k]}\dfrac{\partial \bar{x}^j}{\partial x^p}\dfrac{\partial \bar{x}^r}{\partial x^q}\dfrac{\partial \bar{x}^k}{\partial x^n}.$

Ex. (2.3) (a) Determine the Christoffel symbols of the first kind in polar coordinates. (b) Determine the Christoffel symbols of the second kind in polar coordinates.

Ans. (a) Since the only one of the g_{ij} *which is not constant is* $g_{22} = r^2$, *only* $\partial g_{22}/\partial r = \partial g_{22}/\partial x^1$ *does not vanish. Hence*

$$[21,2] = [12,2] = -[22,1] = \frac{1}{2}\frac{\partial g_{22}}{\partial x^1} = r.$$

(b) Since $g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & r^{-2} \end{pmatrix}$, *from (a) we find that*

$$\begin{Bmatrix} 2 \\ 21 \end{Bmatrix} = g^{2m}[21,m] = g^{22}[21,2] = \frac{1}{r} = \begin{Bmatrix} 2 \\ 12 \end{Bmatrix},$$

$$\begin{Bmatrix} 1 \\ 22 \end{Bmatrix} = g^{1m}[22,m] = g^{11}[22,1] = -r.$$

All others are zero.

Ex. (2.4) Determine the Christoffel symbols in bipolar coordinates. (Hint: see Ex. (1.8).)

Ans. $[11,1] = [12,2] = [21,2] = \dfrac{\cos\alpha}{\rho\,\sigma\sin^4\alpha}(\rho - \sigma\cos\alpha)$,

$$[22,2] = [12,1] = [21,1] = \frac{\cos\alpha}{\sigma\,\rho\sin^4\alpha}(\sigma - \rho\cos\alpha),$$

$$[11,2] = \frac{\sigma\cos^3\alpha - \rho}{\sigma\,\rho\sin^4\alpha}, \quad [22,1] = \frac{\rho\cos^3\alpha - \sigma}{\sigma\,\rho\sin^4\alpha},$$

$$\begin{Bmatrix} 1 \\ 1\,1 \end{Bmatrix} = -\frac{\cos^2\alpha}{\rho\sin^2\alpha}, \quad \begin{Bmatrix} 1 \\ 1\,2 \end{Bmatrix} = \frac{\cos\alpha}{\rho\sin^2\alpha} = \begin{Bmatrix} 1 \\ 2\,1 \end{Bmatrix},$$

$$\begin{Bmatrix} 1 \\ 2\,2 \end{Bmatrix} = -\frac{1}{\rho\sin^2\alpha}, \quad \begin{Bmatrix} 2 \\ 1\,1 \end{Bmatrix} = -\frac{1}{\sigma\sin^2\alpha},$$

$$\begin{Bmatrix} 2 \\ 2\,2 \end{Bmatrix} = -\frac{\cos^2\alpha}{\sigma\sin^2\alpha}, \quad \begin{Bmatrix} 2 \\ 1\,2 \end{Bmatrix} = \frac{\cos\alpha}{\sigma\sin^2\alpha} = \begin{Bmatrix} 2 \\ 2\,1 \end{Bmatrix}.$$

Ex. (2.5) Determine the Christoffel symbols in parabolic coordinates. (Hint: see Ex. (1.9).)

Ans. $[11,1] = u$, $[11,2] = -\mathrm{v}$, $[22,1] = -u$, $[22,2] = \mathrm{v}$,

$$[12,1] = \mathrm{v} = [21,1], \; [12,2] = u = [21,2].$$

$$\begin{Bmatrix} 1 \\ 1\,1 \end{Bmatrix} = \frac{u}{u^2 + \mathrm{v}^2}, \begin{Bmatrix} 1 \\ 2\,2 \end{Bmatrix} = -\frac{u}{u^2 + \mathrm{v}^2}, \begin{Bmatrix} 2 \\ 1\,1 \end{Bmatrix} = -\frac{\mathrm{v}}{u^2 + \mathrm{v}^2},$$

$$\begin{Bmatrix} 2 \\ 2\,2 \end{Bmatrix} = \frac{\mathrm{v}}{u^2 + \mathrm{v}^2}, \begin{Bmatrix} 1 \\ 1\,2 \end{Bmatrix} = \frac{\mathrm{v}}{u^2 + \mathrm{v}^2} = \begin{Bmatrix} 1 \\ 2\,1 \end{Bmatrix}, \begin{Bmatrix} 2 \\ 1\,2 \end{Bmatrix} = -\frac{u}{u^2 + \mathrm{v}^2} = \begin{Bmatrix} 2 \\ 2\,1 \end{Bmatrix}.$$

Ex. (2.6) Show that $[pq,l] = g_{lm}\begin{Bmatrix} m \\ pq \end{Bmatrix}$. *(Hint: use equation (2.11).)*

To see what the Christoffel symbols mean, let us recall that they were so chosen that the intrinsic derivative and ordinary derivative of any contravariant vector would be the same in a rectilinear coordinates system. By equation (2.4), it is then also true that the covariant derivative becomes a partial derivative in a rectilinear coordinate system. Thus the covariant derivative is the counterpart in generalized coordinates of the partial derivative in rectilinear coordinates.

Let us seek a geometrical insight into the Christoffel symbols. To this end, consider the vector fields defined by the contravariant basis vectors in a curvilinear coordinate system. As a concrete example, consider the vector field $e^i_{(1)} = (1,0)$ in a plane polar coordinate system. At each point of the plane, it is a vector of unit length and directed radially outward from the origin, as in Fig. 39.

In what manner do the vectors change as one goes from point to point? To answer this, consider the covariant derivatives of $e^i_{(1)}$. They are

$$e^1_{(1),1} = \begin{Bmatrix} 1 \\ 11 \end{Bmatrix},\ e^1_{(1),2} = \begin{Bmatrix} 1 \\ 21 \end{Bmatrix},\ e^2_{(1),1} = \begin{Bmatrix} 2 \\ 11 \end{Bmatrix},\ e^2_{(1),2} = \begin{Bmatrix} 2 \\ 21 \end{Bmatrix}.$$

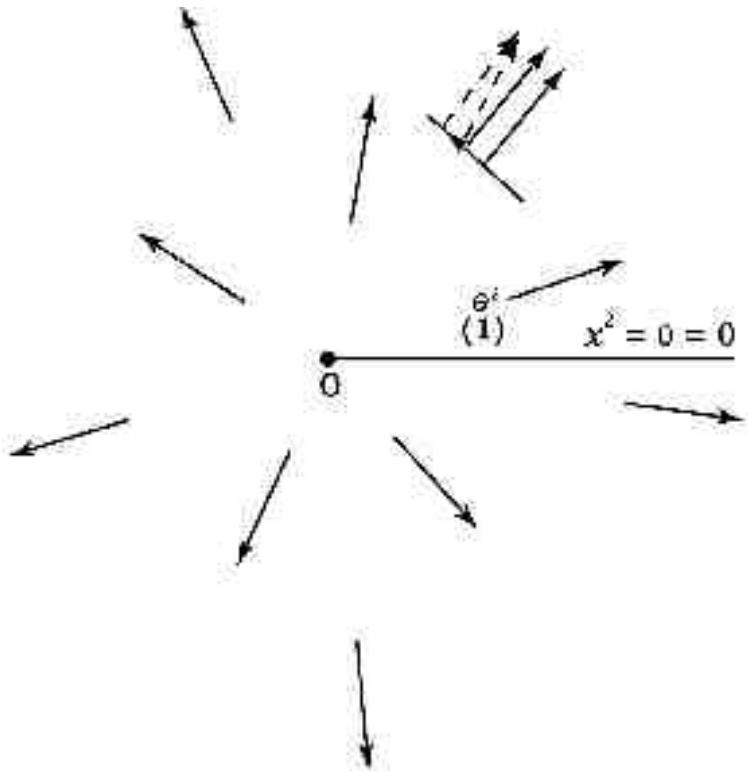

Figure 39

From Ex. (2.3) we see that the values of these in a polar coordinate system are

$$e^1_{(1),1} = 0,\ e^1_{(1),2} = 0,\ e^2_{(1),1} = 0,\ e^2_{(1),2} = \frac{1}{r}.$$

This tells us that over the field $e^j_{(1)}$ only its second component changes, and that only in the θ – direction, and at a rate $\frac{1}{r}$. In similar fashion, we see that the field $e^j_{(2)}$ (see Fig. 40) consists of vectors of length r, at each point perpendicular to the radius to the point. Furthermore,

$$e^1_{(2),1} = \begin{Bmatrix} 1 \\ 12 \end{Bmatrix} = 0,\ e^1_{(2),2} = \begin{Bmatrix} 1 \\ 22 \end{Bmatrix} = -r,$$

$$e^2_{(2),1} = \begin{Bmatrix} 2 \\ 21 \end{Bmatrix} = \frac{1}{r},\ e^2_{(2),2} = \begin{Bmatrix} 2 \\ 22 \end{Bmatrix} = 0.$$

From this we see that over the field $e^j_{(2)}$ the first component changes at a rate $-r$ in the $\theta-$direction, the second component at a rate $\frac{1}{r}$ in the $r-$direction.

Hence we see that the Christoffel symbols are equal to the covariant derivatives of the basis vectors and that these are the counterparts in curvilinear coordinates of the partial derivatives of the basis vectors in a tangent rectilinear system at each point.

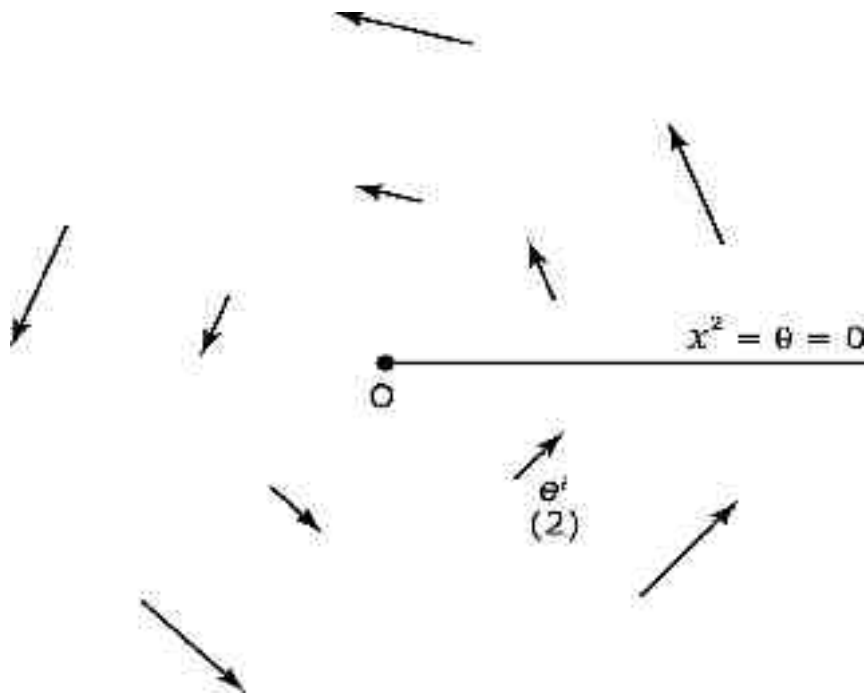

Figure 40

We may also look upon the Christoffel symbols in another light. Referring to equation (2.1), we see that the change in $\bar{v}^j$ is due to two effects: (1) the change in the vector itself, either in direction or magnitude, as reckoned by the first term on the right hand side; and (2) the change in the reference system from each point to the next, as allowed for in the second term on the right hand side. In a rectilinear coordinate system this second term vanishes, hence so do the Christoffel symbols in such coordinates systems. In curvilinear coordinates, the Christoffel symbols automatically compensate for the variation of the local frame of reference.

Let us illustrate the latter point of view by considering the position vector $p^i = (r, 0)$. Its intrinsic derivative with respect to t (time, for example), is

$$v^1 = \frac{\delta p^1}{\delta t} = \dot{p}^1 = \dot{p}^1 + \begin{Bmatrix} 1 \\ ij \end{Bmatrix} p^i \frac{dx^j}{dt} = \dot{r} + \begin{Bmatrix} 1 \\ 22 \end{Bmatrix} p^2 \dot{\theta} = \dot{r},$$

$$v^2 = \frac{\delta p^2}{\delta t} = \dot{p}^2 + \begin{Bmatrix} 2 \\ ij \end{Bmatrix} p^i \frac{dx^j}{dt} = 0 + \begin{Bmatrix} 2 \\ 12 \end{Bmatrix} p^1 \dot{\theta} + \begin{Bmatrix} 2 \\ 21 \end{Bmatrix} p^2 \dot{r} = \frac{1}{r} \cdot r \cdot \dot{\theta} = \dot{\theta}.$$

It is clear that $\delta p^1/\delta t = \dot{r}$ derives from the rate of change of $p^1 = r$ whereas $\delta p^2/\delta t = \dot{\theta}$ derives from the term $\begin{Bmatrix} 2 \\ 12 \end{Bmatrix} p^1 \dot{\theta}$. The former is the rate of change of the vector due to the rate of change of its component, the latter is the rate of change of the vector due to the rate of change of the reference system.

Carrying the process one step further will provide additional illumination. Thus, the intrinsic derivative of the generalized velocity vector $\mathrm{v}^i = (\dot{r}, \dot{\theta})$ is

$$a^1 = \frac{\delta \mathrm{v}^1}{\delta t} = \ddot{r} + \begin{Bmatrix} 1 \\ 22 \end{Bmatrix} \mathrm{v}^2 \dot{\theta} = \ddot{r} - r\dot{\theta}^2,$$

$$a^2 = \frac{\delta \mathrm{v}^2}{\delta t} = \ddot{\theta} + \begin{Bmatrix} 2 \\ 12 \end{Bmatrix} \mathrm{v}^1 \dot{\theta} + \begin{Bmatrix} 2 \\ 21 \end{Bmatrix} \mathrm{v}^2 \dot{r} = \ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} = \frac{1}{r^2}\frac{d}{dt}(r^2\dot{\theta}).$$

The generalized acceleration a^i includes in its first component a term $-r\dot{\theta}^2$ due to change of reference system; this is a **centrifugal** term. The second component includes a term $\frac{2}{r}\dot{r}\dot{\theta}$ due to the rate of change of scale **and** reference system; this is sometimes called a **Coriolis** term. The terminology for these terms obviously derives from dynamics, where they are of special interest.

Ex. (2.7) Determine $g_{ij}\, e^j_{(k),l}$.

Ans. Since $e^j_{(k),l} = \begin{Bmatrix} j \\ lk \end{Bmatrix}$, *we must have that*

$$g_{ij}\, e^j_{(k),l} = g_{ij} \begin{Bmatrix} j \\ lk \end{Bmatrix} = [lk, i].$$

We can thus interpret the Christoffel symbols of the first kind.

Ex. (2.8) The Cartesian components of the velocity and acceleration vectors are $\dot{x}^i$ *and* $\ddot{x}^i$. *By the vector transformation law, determine the components of these vectors in plane polar coordinates. Check your answers against the first and second intrinsic derivatives of the position vector in polar coordinates.*

Ex. (2.9) (a) Show by calculation in plane polar coordinates that $p^i_{,j} = \delta^i_j$ *where* p^i *is the position vector of any point. (b) Verify that* $p^i_{,j} = \delta^i_j$ *in all coordinates by applying the tensor transformation to the results of (a).*

Ex. (2.10) Show that $\dfrac{\partial g_{ij}}{\partial x^k} = [ik, j] + [jk, i]$.

Ex. (2.11) Show that $\dfrac{1}{2} g^{ij} \dfrac{\partial g_{ij}}{\partial x^k} = \begin{Bmatrix} i \\ ki \end{Bmatrix}$. *(Hint: use the results of Ex. (2.10).)*

3. Intrinsic and Covariant Differentiation of Vectors and Tensors in General

We have seen how we may differentiate contravariant vectors intrinsically to obtain new vectors. It might reasonably be supposed that a similar process could be devised for differentiating covariant vectors intrinsically. We proceed in analogous fashion by considering a rectilinear coordinate system $\bar{x}^i$ and a curvilinear coordinate system x^i in which are defined a covariant vector whose components in the two coordinate systems satisfy the law of vector transformation, namely

$$\mathrm{v}_j = \frac{\partial \bar{x}^i}{\partial x^j} \bar{\mathrm{v}}_i .$$

Differentiating with respect to some parameter t gives

$$\begin{aligned} \frac{d\mathrm{v}_j}{dt} &= \frac{\partial \bar{x}^i}{\partial x^j} \frac{d\bar{\mathrm{v}}_i}{dt} + \frac{\partial^2 \bar{x}^i}{\partial x^j \partial x^k} \bar{\mathrm{v}}_i \frac{dx^k}{dt} \\ &= \frac{\partial \bar{x}^i}{\partial x^j} \frac{d\bar{\mathrm{v}}_i}{dt} + \frac{\partial x^l}{\partial \bar{x}^i} \frac{\partial^2 \bar{x}^i}{\partial x^j \partial x^k} \mathrm{v}_l \frac{dx^k}{dt} . \end{aligned} \tag{3.1}$$

We now *define* a set of non-tensor quantities Γ'^l_{jk}, possibly different from the previously defined Γ^l_{jk}, which transform according to the rule*

$$-\Gamma'^l_{jk} \mathrm{v}_l \frac{dx^k}{dt} + \frac{\partial x^l}{\partial \bar{x}^i} \frac{\partial^2 \bar{x}^i}{\partial x^j \partial x^k} \mathrm{v}_l \frac{dx^k}{dt} = -\frac{\partial \bar{x}^i}{\partial x^j} \bar{\Gamma}'^q_{ip} \bar{\mathrm{v}}_q \frac{d\bar{x}^p}{dt} . \tag{3.2}$$

Then, adding equations (3.1) and (3.2) and cancelling like terms on right and left sides, we get

$$\frac{d\mathrm{v}_j}{dt} - \Gamma'^l_{jk} \frac{dx^k}{dt} = \frac{\partial \bar{x}^i}{\partial x^j} \left(\frac{d\bar{\mathrm{v}}_i}{dt} - \bar{\Gamma}'^q_{ip} \bar{\mathrm{v}}_q \frac{d\bar{x}^p}{dt} \right) .$$

Clearly, the quantity

$$\frac{\delta \mathrm{v}_j}{\delta t} \equiv \frac{d\mathrm{v}_j}{dt} - \Gamma'^l_{jk} \mathrm{v}_l \frac{dx^k}{dt} \tag{3.3}$$

is a covariant vector. It is called the **intrinsic derivative of v_i with respect to t**. The definition (3.2) is again a transparent artifice to enable us to construct a true covariant vector from the derivative of a covariant vector.

---

* The minus signs appear in equation (3.2) with malice aforethought; they will guarantee that $\Gamma'^l_{jk} = \Gamma^l_{jk}$ in Riemannian geometries. For instances where this is not so, see Chapter 7.

As before, we may write $\dfrac{d\mathbf{v}_j}{dt} = \dfrac{\partial \mathbf{v}_j}{\partial x^k}\dfrac{dx^k}{dt}$ when the components of $\mathbf{v}_j$ do not depend explicitly on the parameter t. Then

$$\frac{\delta \mathbf{v}_j}{\delta t} = \left(\frac{\partial \mathbf{v}_j}{\partial x^k} - \Gamma'^{l}_{jk}\mathbf{v}_1\right)\frac{dx^k}{dt} = \mathbf{v}_{j,k}\frac{dx^k}{dt}.$$

The quantity

$$\mathbf{v}_{j,k} \equiv \frac{\partial \mathbf{v}_j}{\partial x^k} - \Gamma'^{l}_{jk}\mathbf{v}_l \tag{3.4}$$

is called the **covariant derivative of the covariant vector $\mathbf{v}_j$ with respect to x^k.**

Again, there remains the need for choosing the quantities Γ'^{l}_{jk} in some coordinate system. Again, we *elect* to make the intrinsic derivative identical with the ordinary derivative and the covariant derivative identical with the partial derivative in rectilinear coordinates. This implies that

$$\overline{\Gamma}'^{q}_{ip} = 0$$

in equation (3.2), whence it follows that

$$\Gamma'^{l}_{jk} = \frac{\partial x^l}{\partial \bar{x}^i}\frac{\partial^2 \bar{x}^i}{\partial x^j \partial x^k}.$$

This is seen to be identical with the expression in equation (2.7). Therefore the Γ'^{l}_{jk} are simply the Christoffel symbols of the second kind. For comparison's sake, we therefore write the intrinsic derivatives of $\mathbf{v}^i$ and $\mathbf{v}_i$ in their final form, namely

$$\frac{\delta \mathbf{v}^i}{\delta t} = \frac{d\mathbf{v}^i}{dt} + \begin{Bmatrix} i \\ jk \end{Bmatrix}\mathbf{v}^j\frac{dx^k}{dt}, \quad \frac{\delta \mathbf{v}_i}{\delta t} = \frac{d\mathbf{v}_i}{dt} - \begin{Bmatrix} j \\ ik \end{Bmatrix}\mathbf{v}_j\frac{dx^k}{dt}. \tag{3.5}$$

Ex. (3.1) From equation (3.5), show that

$$\mathbf{v}_i\frac{\delta u^i}{\delta t} + u^i\frac{\delta \mathbf{v}_i}{\delta t} = \mathbf{v}_i\frac{du^i}{dt} + u^i\frac{d\mathbf{v}_i}{dt} = \frac{d}{dt}(u^i\mathbf{v}_i).$$

Hence show that

$$\frac{d}{dt}(u^i\mathbf{v}_i) = \frac{\delta}{\delta t}(u^i\mathbf{v}_i)$$

provided intrinsic differentiation follows the same rules for differentiation of a product as does ordinary differentiation.

The differentiation of tensors is now a relatively simple matter, for as we have seen on numerous previous occasions, the products of vectors are typical tensors. Let us therefore note how a tensor $\mathbf{T}^i_j = u^i\mathbf{v}_j$ could be expected to be differentiated. Since the usual rules of differential calculus are to be obeyed in rectilinear systems, we require that intrinsic differentiation follow the same rules.

In particular, we require satisfaction of the rule for the differentiation of a product (see Ex. (3.1)). Hence

$$\frac{\delta}{\delta t}(u^i \mathrm{v}_j) = \frac{\delta u^i}{\delta t}\mathrm{v}_j + u^i \frac{\delta \mathrm{v}_j}{\delta t}$$

$$= \mathrm{v}_j\left(\frac{du^i}{dt} + \begin{Bmatrix} i \\ kl \end{Bmatrix} u^k \frac{dx^l}{dt}\right) + u^i\left(\frac{d\mathrm{v}_j}{dt} - \begin{Bmatrix} k \\ jl \end{Bmatrix} \mathrm{v}_k \frac{dx^l}{dt}\right)$$

$$= \left(\mathrm{v}_j \frac{du^i}{dt} + u^i \frac{d\mathrm{v}_j}{dt}\right) + \begin{Bmatrix} i \\ kl \end{Bmatrix}(u^k \mathrm{v}_j)\frac{dx^l}{dt} - \begin{Bmatrix} k \\ jl \end{Bmatrix}(u^i \mathrm{v}_k)\frac{dx^l}{dt}$$

or

$$\frac{\delta \mathrm{T}^i_j}{\delta t} = \frac{d\mathrm{T}^i_j}{dt} + \begin{Bmatrix} i \\ kl \end{Bmatrix} \mathrm{T}^k_j \frac{dx^l}{dt} - \begin{Bmatrix} k \\ jl \end{Bmatrix} \mathrm{T}^i_k \frac{dx^l}{dt}. \tag{3.6}$$

Plainly, every additional superscript requires the inclusion of an additional term analogous to the second term on the right, whereas every additional subscript requires the inclusion of an additional term analogous to the third term on the right. Thus we are able to differentiate intrinsically a tensor of any order and type.

At the same time, we can also differentiate covariantly a tensor of any order and type. By an almost trivial generalization similar to that by which we derived equation (3.6), it is very simple to show that

$$\begin{aligned} \mathrm{T}^{ik\cdots}_{jl\cdots,p} = \frac{\partial \mathrm{T}^{ik\cdots}_{jl\cdots}}{\partial x^p} &+ \begin{Bmatrix} i \\ qp \end{Bmatrix} \mathrm{T}^{qk\cdots}_{jl\cdots} + \begin{Bmatrix} k \\ qp \end{Bmatrix} \mathrm{T}^{iq\cdots}_{jl\cdots} + \cdots \\ &- \begin{Bmatrix} q \\ jp \end{Bmatrix} \mathrm{T}^{ik\cdots}_{ql\cdots} - \begin{Bmatrix} q \\ lp \end{Bmatrix} \mathrm{T}^{ik\cdots}_{jq\cdots} - \cdots \end{aligned} \tag{3.7}$$

is the **covariant derivative of** $\mathrm{T}^{ik\cdots}_{jl\cdots}$ **with respect to** x^p. Covariant differentiation may be repeated without limit. All subscripts following a comma are differentiation indices.

One of the several advantages which follows from our electing to make identical the ordinary and intrinsic derivatives and the partial and covariant derivatives in rectilinear coordinates is that tensors which are constant in such coordinate systems, and therefore have zero ordinary and partial derivatives, will necessarily have zero intrinsic and covariant derivatives in all coordinate systems; this is clear if one simply applies the tensor transformation. Important among these are, for example, the tensors g_{ij}, δ^i_j, g^{ij} and the ε_{ij}, ε^{ij}. Hence in any intrinsic or covariant differentiation, they may be treated as constants would be treated in ordinary differentiation. For example,

$$\mathrm{v}_{i,j} = (g_{ik}\mathrm{v}^k)_{,j} = g_{ik,j}\mathrm{v}^k + g_{ik}\mathrm{v}^k_{,j} = g_{ik}\mathrm{v}^k_{,j}.$$

Consequently, *indices may be raised or lowered at will without regard to intrinsic or covariant differentiation.*

Ex. (3.2) By writing out the covariant derivative of g_{ij} *in full, show that it vanishes identically.*

Ans. We have

$$g_{ij,k} = \frac{\partial g_{ij}}{\partial x^k} - \begin{Bmatrix} l \\ ik \end{Bmatrix} g_{lj} - \begin{Bmatrix} l \\ jk \end{Bmatrix} g_{il} = \frac{\partial g_{ij}}{\partial x^k} - [ik,j] - [jk,i]$$

$$= \frac{\partial g_{ij}}{\partial x^k} - \frac{1}{2}\left[\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{kj}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i}\right] = 0.$$

Ex. (3.3) Show that $\begin{Bmatrix} i \\ ki \end{Bmatrix} = -\frac{1}{2} g_{ij} \frac{\partial g^{ij}}{\partial x^k}$.

Ans. We have $g^{ij}_{,k} = \frac{\partial g^{ij}}{\partial x^k} + \begin{Bmatrix} i \\ kl \end{Bmatrix} g^{lj} + \begin{Bmatrix} j \\ kl \end{Bmatrix} g^{il} = 0$.

Multiplying through by g_{ij} *gives*

$$g_{ij} \frac{\partial g^{ij}}{\partial x^k} + \begin{Bmatrix} i \\ kl \end{Bmatrix} g^{lj} g_{ij} + \begin{Bmatrix} j \\ kl \end{Bmatrix} g^{il} g_{ij} = g_{ij} \frac{\partial g^{ij}}{\partial x^k} + \begin{Bmatrix} i \\ ki \end{Bmatrix} + \begin{Bmatrix} j \\ kj \end{Bmatrix} = 0,$$

whence the above result follows.

Ex. (3.4) Show that $\begin{Bmatrix} i \\ ki \end{Bmatrix} = \frac{1}{2} g^{ij} \frac{\partial g_{ij}}{\partial x^k}$. *Thus verify Ex. (2.11). (Hint: use the fact that* $g^{ij} g_{ij} = 2$ *and the results of Ex. (3.3).)*

Consider at any point a unit vector λ^i. **Along a curve to which is a unit** tangent λ^i, the coordinate differentials must be $dx^i = \lambda^i ds$. If a vector field v^i is defined along the curve to which λ^i is the tangent, then

$$v^i_{,j} \frac{dx^j}{ds} = v^i_{,j} \lambda^j = \frac{\delta v^i}{\delta s}.$$

If we choose, we can interpret this equation as a statement that the covariant derivative $v^i_{,j}$ of the vector v^i is an operator which maps the unit vector λ^j into the intrinsic derivative $\delta v^i / \delta s$ along the curve. This derivative is called the **directional derivative** of v^i in the direction λ^j.

Ex. (3.5) Show that the directional derivative of the contravariant basis vector is

$$\frac{\delta e^{i}_{(k)}}{\delta s_{(m)}} = \frac{\left\{ {i \atop mk} \right\}}{\sqrt{g_{mm}}}$$

along the x^m*– coordinate curve.*

Ans. Since $e^{i}_{(k)} = \left\{ {i \atop jk} \right\}$ *and* $\lambda^{j}_{(m)} = \dfrac{e^{j}_{(m)}}{\sqrt{g_{mm}}}$, , *we have*

$$\frac{\delta e^{i}_{(k)}}{\delta s_{(m)}} = e^{i}_{(k),j}\,\lambda^{j}_{(m)} = \left\{ {i \atop jk} \right\} \frac{e^{j}_{(m)}}{\sqrt{g_{mm}}} = \frac{\left\{ {i \atop mk} \right\}}{\sqrt{g_{mm}}}.$$

Ex. (3.6) Find the directional derivatives of the contravariant basis vectors in polar coordinates.

Ans.

$$\frac{\delta e^{i}_{(1)}}{\delta s_{(1)}} = (0, 0), \quad \frac{\delta e_{(1)}}{\delta s_{(2)}} = (0, \frac{1}{r^2}),$$

$$\frac{\delta e^{i}_{(2)}}{\delta s_{(1)}} = (0, \frac{1}{r}), \quad \frac{\delta e^{i}_{(2)}}{\delta s_{(2)}} = (-1, 0).$$

Ex. (3.7) Show that the directional derivative of the covariant basis vector is

$$\frac{\delta e^{(i)}_{k}}{\delta s_{(m)}} = -\frac{\left\{ {i \atop mk} \right\}}{\sqrt{g_{mm}}}$$

along the x^m*– coordinate curve.*

Ex. (3.8) A particularly useful and important directional derivative is that in the direction of the normal to a curve. This is called the ***normal derivative.*** *If* n^i *is the unit normal to a curve, then* $\mathrm{v}^{i}_{,l}\,n^l$ *is the normal derivative of* v^i. *Find the normal derivative to the curve* $r^2(1 - e^2\cos^2\theta)$ = *constant in plane polar coordinates for the vector* $\mathrm{v}^i = (r, 1)$. *(Hint: use the results of Ex. (1.11) and (1.12).)*

Ans. Since $\mathrm{v}^1_{,1} = 1$, $\mathrm{v}^1_{,2} = -r$, $\mathrm{v}^2_{,1} = \dfrac{1}{r}$, $\mathrm{v}^2_{,2} = 1$, *we have*

$$\mathrm{v}^1_{,l}\,n^l = \frac{1}{\mathrm{R}}[1 - e^2\cos\theta(\cos\theta - \sin\theta)],$$

$$\mathrm{v}^2_{,l}\,n^l = \frac{1}{r\mathrm{R}}[1 - e^2\cos\theta(\cos\theta + \sin\theta)],$$

where $\mathrm{R} = \sqrt{1 - (2e^2 - e^4)\cos\theta}$.

Ex. (3.9) The covariant derivative of a vector is a tensor of second rank. As such, it possesses a covariant derivative. Find the covariant derivative of the covariant derivative of the covariant basis vector $e_i^{(k)}$.

Ans. $e_{i,jm}^{(k)} = -\dfrac{\partial}{\partial x^m}\begin{Bmatrix} k \\ ij \end{Bmatrix} + \begin{Bmatrix} l \\ im \end{Bmatrix}\begin{Bmatrix} k \\ lj \end{Bmatrix} + \begin{Bmatrix} l \\ jm \end{Bmatrix}\begin{Bmatrix} k \\ il \end{Bmatrix}.$

Ex. (3.10) Show that $\begin{Bmatrix} i \\ ki \end{Bmatrix} = \dfrac{\partial \log\sqrt{g}}{\partial x^k}$. *(Hint: use the results of Ex. (3.4), written out in full, remembering that* $g^{ij} = G^{ij}/g$, *where* G^{ij} *is the cofactor of* g_{ij}.

4. Plane Curves in Generalized Coordinates

The treatment of plane curves in generalized coordinates is formally identical with the treatment we have already considered in rectilinear coordinates except that ordinary derivatives are to be replaced by intrinsic derivatives. Thus, let $x^i(t)$ be a curve with parameter t in a generalized coordinate system in which the square of the line element is

$$(ds)^2 = g_{ij}\,dx^i\,dx^j. \tag{4.1}$$

The fundamental tensor may now be a function of position. Then the vector

$$\lambda^i = dx^i/ds \tag{4.2}$$

is a unit tangent to the curve, i. e., $g_{ij}\lambda^i\lambda^j = 1$. Hence

$$\frac{\delta}{\delta s}(g_{ij}\lambda^i\lambda^j) = 2g_{ij}\lambda^i\frac{\delta\lambda^j}{\delta s} = 0,$$

so that $\delta\lambda^j/\delta s$ must be orthogonal to λ^i. We set it equal to

$$\frac{\delta\lambda^i}{\delta s} = \kappa\mu^i, \tag{4.3}$$

where μ^i is a unit vector in the direction of $\delta\lambda^i/\delta s$ and where κ is the magnitude of $\delta\lambda^i/\delta s$; i.e.,

$$(\kappa)^2 = g_{ij}\frac{\delta\lambda^i}{\delta s}\frac{\delta\lambda^j}{\delta s}. \tag{4.4}$$

The quantity κ is called the **curvature** of the curve and μ^i is the **unit normal**.

Ex. (4.1) Show that

$$\frac{\delta \lambda^i}{\delta s} = \left(\frac{dt}{ds}\right)^2 \left[\frac{\delta}{\delta t}\left(\frac{dx^i}{dt}\right) + \left(\frac{\frac{d^2 s}{dt^2}}{\frac{ds}{dt}} \right) \frac{dx^i}{dt} \right].$$

Ex. (4.2) Find (a) the unit tangent, (b) the unit normal and (c) the curvature κ *for the ellipse* $r(1 + e\cos\theta) = a(1 - e^2)$.

Ans. (a) $\lambda^i = \frac{1}{R}\left(e\sin\theta, \frac{1 + e\cos\theta}{r}\right)$, *where*

$$R = [(1 + e^2) + 2e\cos\theta]^{1/2};$$

(b) $\mu^i = \frac{1}{R}\left(1 + e\cos\theta, \frac{-e\sin\theta}{r}\right)$; (c) $\kappa = -\frac{a^2(1 - e^2)^2}{(rR)^3}$.

Ex. (4.3) Verify Ex. (4.2b) by finding $\mu_i = \varepsilon_{ij}\lambda^j$, $\mu^i = g^{ij}\mu_j$.

Ex. (4.4) Find (a) the unit tangent, (b) the unit normal and (c) the curvature to the curve $\theta = \alpha \log r$ *in polar coordinates. (Be certain that* λ^i *and* μ^i *form a right-handed system.)*

Ans. (a) $\lambda^i = \frac{1}{\sqrt{1 + \alpha^2}}\left(1, \frac{\alpha}{r}\right)$;

(b) $\mu^i = \frac{1}{\sqrt{1 + \alpha^2}}\left(\alpha, -\frac{1}{r}\right)$; (c) $\kappa = -\frac{\alpha}{r\sqrt{1 + \alpha^2}}$.

Ex. (4.5) Show that $\mu_i = \varepsilon_{ij}\lambda^j$ *is orthogonal to* λ^i *(Hint: find the inner product* $\lambda^i \mu_i$.*)*

Ex. (4.6) Show that $\lambda^k = -\varepsilon^{ki}\mu_i$.

Ex. (4.7) Show that

$$\frac{\delta \mu^i}{\delta s} = -\kappa \lambda^i. \tag{4.5}$$

(Hint: use Ex. (4.5) and equation (4.3), then the results of Ex. (4.6).) Equations (4.3) and (4.5) constitute the ***Frenet formulae for a surface.***

Ex. (4.8) Verify the Frenet formulae for Ex. (4.4).

A kind of curve of particular importance is, of course, the straight line. If we define it as a curve whose curvature is everywhere zero, then we see that its differential equation is

(4.6) $$\frac{\delta \lambda^i}{\delta s} = \frac{\delta}{\delta s}\left(\frac{dx^i}{ds}\right) = \frac{d^2 x^i}{ds^2} + \left\{ {i \atop jk} \right\} \frac{dx^j}{ds}\frac{dx^k}{ds} = 0 .$$

The same condition may be expressed directly in terms of the curve parameter t. To do so, we note that

(4.7)
$$\frac{d^2x^i}{ds^2} = \frac{d}{ds}\left(\frac{dx^i}{ds}\right) = \frac{d}{ds}\left(\frac{dx^i}{dt}\frac{dt}{ds}\right) = \frac{dt}{ds}\left[\frac{d}{ds}\left(\frac{dx^i}{dt}\right)\right] + \frac{dx^i}{dt}\frac{d}{ds}\left(\frac{dt}{ds}\right)$$
$$= \frac{dt}{ds}\left[\frac{dt}{ds}\frac{d}{dt}\left(\frac{dx^i}{dt}\right)\right] + \frac{dx^i}{dt}\frac{d^2t}{ds^2} = \left(\frac{dt}{ds}\right)^2 \frac{d^2x^i}{dt^2}\frac{d^2x^i}{dt^2} + \frac{dx^i}{dt}\frac{d^2t}{ds^2} .$$

Making these substitutions in equation (4.6), we get

(4.8) $$\frac{d^2x^i}{dt^2} + \left\{ {i \atop jk} \right\} \frac{dx^j}{dt}\frac{dx^k}{dt} = \frac{\delta}{\delta t}\left(\frac{dx^i}{dt}\right) = \alpha \frac{dx^i}{dt} ,$$

where

(4.9) $$\alpha = -\frac{\dfrac{d^2t}{ds^2}}{\left(\dfrac{dt}{ds}\right)^2} = \frac{d}{ds}\frac{1}{\left(\dfrac{dt}{ds}\right)} = \frac{d}{ds}\left(\frac{ds}{dt}\right) = \frac{\dfrac{d^2s}{dt^2}}{\dfrac{ds}{dt}} .$$

Note that when $t \propto s$, $\alpha = 0$.

These results bear an interesting relation to a particularly important type of vector field, the **parallel vector field**. In the plane, vectors are parallel in the ordinary sense if their rectilinear components differ at most by a common factor. Thus in Fig. 41 the vectors $\mathbf{A}^i$, $\mathbf{B}^i$, $\mathbf{C}^i$ and $\mathbf{D}^i$, though unequal in length, have the same direction and are therefore parallel. In the rectilinear system of coordinates, therefore,

$$\varphi(x^j_{(\mathrm{A})})\mathbf{A}^i = \varphi(x^j_{(\mathrm{B})})\mathbf{B}^i = \varphi(x^j_{(\mathrm{C})})\mathbf{C}^i = \varphi(x^j_{(\mathrm{D})})\mathbf{D}^i, \text{ etc.,}$$

where $\varphi(x^j)$ is an appropriate invariant function of the coordinates, evaluated at the points $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$, respectively; φ is in effect a scale factor which compensates for the variety of the lengths of $\mathbf{A}^i$, $\mathbf{B}^i$, $\mathbf{C}^i$ and $\mathbf{D}^i$. Without loss of generality, we take $\varphi = 1$ at some convenient point.

Now suppose that C(t) is some continuous curve with parameter t and that φ is a continuous and differentiable function. Then if we let p^i represent not only $\mathbf{A}^i$, $\mathbf{B}^i$, $\mathbf{C}^i$ and $\mathbf{D}^i$ but all members of the parallel vector field, we can say that

(4.10) $$\varphi(x^j)p^i(x^j) = \mathbf{A}^i ,$$

where A^i is fixed. In particular, since $\delta A^i / \delta t = 0$ in a rectilinear system, along C(t) it will be true that $\frac{\delta}{\delta t}[\varphi p^i] = 0$ or

$$\varphi \frac{\delta p^i}{\delta t} + p^i \frac{d\varphi}{dt} = 0 . \tag{4.11}$$

The existence of a scalar field φ which satisfies equation (4.11) is evidently the necessary and sufficient condition that a vector field p^i be a parallel vector field.

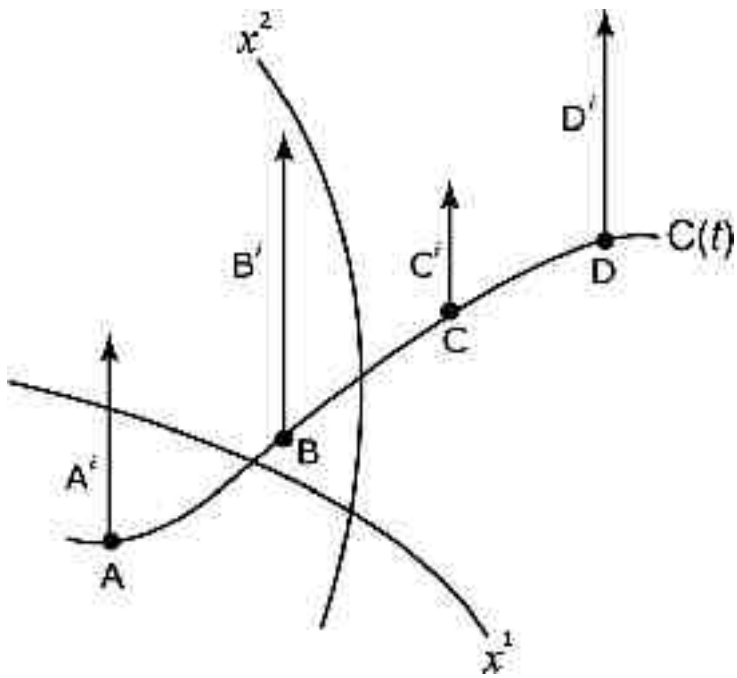

Figure 41

Note that when φ = constant, the vector p^i has a constant magnitude, namely, $|p^i| = \frac{1}{\varphi}|A^i|$, according to equation (4.10). Then $\frac{\delta p^i}{\delta t} = 0$. Otherwise the magnitude of the parallel field p^i varies along the curve.

The relation to straight lines comes in evidence from a comparison of equations (4.8) and (4.11). If

$$p^i = \frac{dx^i}{dt} \text{ and } \alpha = -\frac{1}{\varphi}\frac{d\varphi}{dt},$$

the two equations are identical. Geometrically, this signifies that the tangents

$$\lambda^i = \frac{dx^i}{ds} = \frac{dx^i}{dt}\frac{dt}{ds} = p^i \frac{dt}{ds}$$

form a parallel vector field. Hence either definition of a straight line leads to the same differential equation, as it should.

Ex. (4.9) Show that the vector field

$$p^1 = \alpha\cos\theta + \beta\sin\theta, \quad p^2 = \frac{-\alpha\sin\theta + \beta\cos\theta}{r}$$

in polar coordinates is a parallel vector field with constant magnitude. (Hint: show that $\delta p^i/\delta t = 0$*.) Here* α *and* β *are constants.*

Ex. (4.10) (a) Is the vector field

$$p^1 = \gamma r^2 \sin(\theta - \theta_0), \quad p^2 = \gamma r\cos(\theta - \theta_0)$$

a parallel vector field in polar coordinates? Here γ *is a constant. (b) Is it of constant magnitude? (c) If not, what function of position is the magnitude?*

Ans. (a) Yes, since $\delta p^i/\delta t = \frac{2\dot{r}}{r}p^i$*. (b) No, since*

$$-\frac{1}{\varphi}\frac{d\varphi}{dt} = \frac{2}{r}\frac{dr}{dt}, \quad \varphi = r^{-2}.$$

(c) Therefore $|p^i| = r^2|p^i|_{r=1}$*.*

Ex. (4.11) Determine the condition for the parallelism of a unit vector in a plane polar coordinate system.

Ans. A vector u^i *of constant magnitude is parallel along a curve C if* $\frac{\delta u^i}{\delta\sigma} = 0$ *along C, where* $d\sigma$ *is the element of arc length. In plane polar coordinates* (r, θ)*, the conditions of parallelism are thus*

$$\frac{du^1}{d\sigma} - r u^2\frac{d\theta}{d\sigma} = 0, \quad r\frac{du^2}{d\sigma} + u^1\frac{d\theta}{d\sigma} + u^2\frac{dr}{d\sigma} = 0.$$

Without loss of generality (see Ex. (6.11)), we may take u^i *to be of the form* $u^i = (\cos\alpha, \frac{\sin\alpha}{r})$*. Then from the first of the equations above we get*

$$-(\sin\alpha)\frac{d\alpha}{d\sigma} - r\left(\frac{\sin\alpha}{r}\right)\frac{d\theta}{d\sigma} = 0,$$

$$\frac{d\alpha}{d\theta} = -1, \quad \alpha = \alpha_0 - \theta.$$

Hence $u^i = \left(\cos[\alpha_0 - \theta], \frac{\sin[\alpha_0 - \theta]}{r}\right)$*. Evidently* α_0 *is determined by the components of* u^i *in the direction* $\theta = 0$*. Note that in this instance it was not necessary, after all, to specify a curve along which the vectors are parallel. This is the case only because the plane is a flat space; it would otherwise have been necessary (see Ex. (7.15)). Observe that the second equation for parallelism is identical with the first.*

5. Vectors and Tensors on Curved Surfaces

Our discussion thus far has been confined exclusively to vectors and tensors in a plane. This might have been anticipated, since a vector was initially conceived as a directed straight line segment in a plane. Our first generalization, to curvilinear coordinates in a plane, required no modification other than that of introducing possible rotations of axes (tangents to the coordinate curves) and changes of scale from point to point. A potentially more difficult generalization is in prospect when one contemplates vectors or tensors in a two-dimensional space which is not a plane but a curved surface. Can vectors be associated in some way with the points of a surface? If so, can they be differentiated? These and related questions are some of the problems which we now address.

The first step is to define a set of coordinates in the surface in question. Generally, this requires two families of curves, the members of which intersect each other once but not themselves. As a familiar example, we cite the meridians and parallels on the surface of the earth (except at the geographic poles). The x^1– curves are those along which only the parameter x^1 varies, whereas the x^2– curves are those along which only the parameter x^2 varies. Thus x^2 is constant along an x^1– curve and vice versa. The curves through any point $\mathbf{P}$ on the surface $\mathbf{S}$ (see Fig. 42) are identified by particular values of x^1 and x^2. These are therefore the coordinates of the point $\mathbf{P}$.

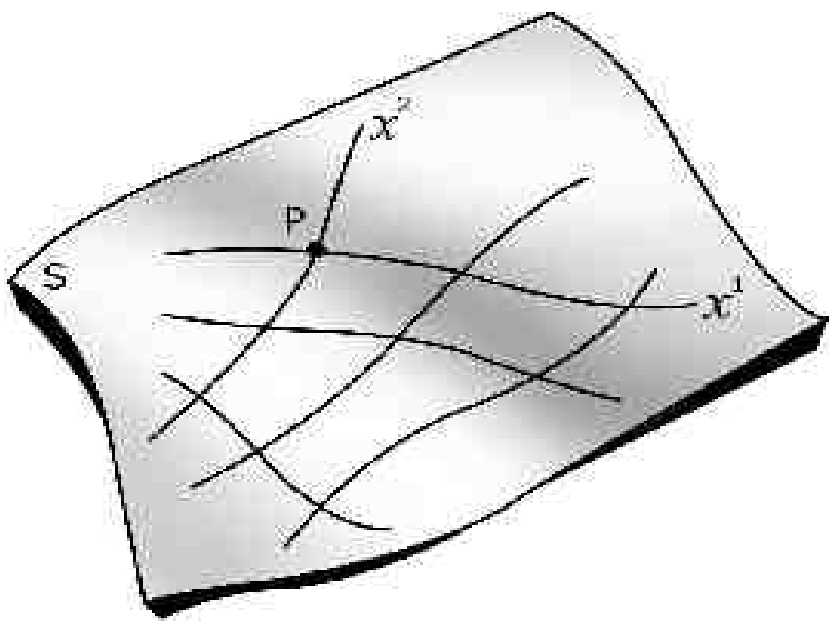

Figure 42

The first step in defining vectors in or on the surface $\mathbf{S}$ is to construct the plane $\mathbf{T}$ tangent to $\mathbf{S}$ at $\mathbf{P}$ (see Fig. 43). This is reminiscent of the state of affairs when we introduced the tangents to coordinate curves at a point in the plane. The parallel may be extended by drawing in $\mathbf{T}$ as rectilinear axes the straight lines $\mathbf{PX^1}$ and $\mathbf{PX^2}$ which are tangent to the coordinate curves at the point $\mathbf{P}$. A difference between this and the previous case, however, is that in general the coordinate curves do not lie in the plane of the tangents.

Now it must be clear that the only sense in which a vector can be associated with a point P of a surface S is that it lies in the tangent plane T, else it must have a component in at least a third dimension. We therefore say that *any vector in* S *at the point* P *is simply a plane vector in the plane* T *tangent to* S *at* P. It is difficult to see how the association could be made more naturally.

It follows from such a definition that all the *algebraic* results for plane vectors may be taken over at once as they were in generalized plane coordinates. Sums, differences, products and transformations of coordinates have exactly the same meanings as before inasmuch as they pertain solely to vectors or tensors at one and the same point in a plane. We therefore need not repeat the development of these details.

Somewhat less straightforward than the definition of vectors and tensors at the same point in a surface is the comparison of vectors and tensors at different points, even points common to a small region. Since a vector is characterized by a direction and magnitude, this means that the problem of comparing vectors at different points is one of defining quality of magnitudes and directions at different points. This is readily done when the vectors lie in a common plane, as has been the case thus far. How may it be done over a surface? We divide the problem into two parts: (1) defining an invariant standard of length at every point of the surface, and (2) defining a condition of parallelism between vectors in the same small neighborhood.

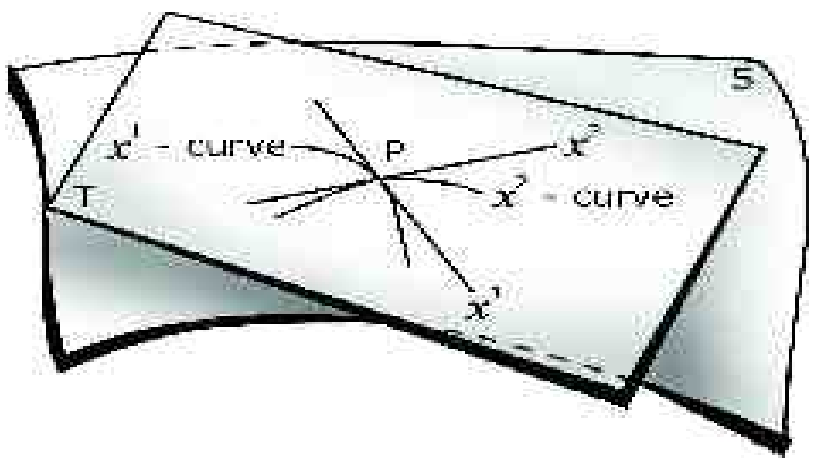

Figure 43

In considering the first problem, that of defining a ***congruence relation,*** we may be guided by a special case. Suppose, for example, that the surface is a sphere. In spherical coordinates, the equation of a sphere with center at the origin is r = constant = a. If this condition is applied to the expression for the square of the line element in the full three dimensions, namely

$$(ds)^2 = (dr)^2 + (r\cos\varphi)^2(d\theta)^2 + r^2(d\varphi)^2,$$

where θ is the longitude, φ the latitude, and r the radius vector, then the square of the line element *in the surface of the sphere* $r = a$ is evidently

(5.1) $$(ds)^2 = a^2 [\cos^2\varphi\,(d\theta)^2 + (d\varphi)^2].$$

Thus, by imposing some defining condition upon the square of the line element for the space in which the surface is embedded, this line element has been specialized to a form which is valid only on the surface. We have thus identified an infinitesimal invariant standard of length on the surface of the sphere. Writing it as

$$(ds)^2 = g_{ij}\,dx^i\,dx^j,$$

it determines the tensor g_{ij}, the fundamental tensor on the curved surface. As in the plane, the quantities g_{ij} therefore determine the length of any infinitesimal displacement. In this way, a unit length is defined at each point and in every direction by the unit vector $\lambda^i = dx^i/ds$. The fundamental tensor for the surface thus provides the solution of the first of our two problems. Once it is defined, by whatever means, quantities derived from it, such as the Christoffel symbols, are also defined.

Let us note expressly that a surface line element need not have been obtained by specializing the line element of some higher-dimensional space. Once the fundamental tensor is assigned, the history of its derivation becomes irrelevant. This may be emphasized by noting that the procedure we followed in the example of a spherical surface is not uniquely reversible; given equation (5.1), we cannot say that it *must* have been derived as we derived it — neither from spherical coordinates nor from a space of three dimensions. The fundamental tensor of the surface thus pertains to the surface and is characteristic of the surface independent of all other considerations. We may underscore this fact by writing equation (5.1) in the more routine tensor index notation as

(5.2) $$(ds)^2 = a^2\left[\cos^2 x^2\,(dx^1)^2 + (dx^2)^2\right],$$

thus divesting the coordinates of any inferential identification as longitude and latitude. We may even anticipate that if one were given equation (5.2) without any hint as to what x^1 and x^2 are construed to be, they might equally well be taken to be rectilinear coordinates in a plane, though to be sure, distances in the plane would not and could not be reckoned as we have heretofore reckoned distances in a plane when the fundamental tensor had constant components. This and other differences are usually summed up by saying that the *geometries* of the two spaces differ. It will be one of our purposes to make this somewhat vague and intuitive argument more precise and more explicit. Our present discussion of this matter is meant only to help make the ensuing developments seem more purposeful as well as to alert the student to some of the distinctions which will arise.

Ex. (5.1) Determine the fundamental tensor on the surface of a spheroid of eccentricity e. (Hint: eliminate dr from the three-dimensional line element by imposing the constraint which defines the spheroidal surface, namely

$$\rho = r\sqrt{1 - e^2\cos^2\varphi} = a\sqrt{1 - e^2},\ 1 > e \geq 0.$$

Ans. $$g_{11} = \frac{\rho^2\cos^2\varphi}{1 - e^2\cos^2\varphi},\ g_{22} = \frac{\rho^2[1 - (2e^2 - e^4)\cos^2\varphi]}{(1 - e^2\cos^2\varphi)^3},$$

all others zero.

Ex. (5.2) Verify your answer to Ex. (5.1) by applying the tensor transformation to the fundamental tensor in the case

$$x^1 = r = \frac{\rho}{\sqrt{1 - e^2 \cos^2 \varphi}} = \frac{\bar{x}^1}{\sqrt{1 - e^2 \cos^2 \bar{x}^3}},$$

$$x^2 = \theta = \bar{x}^2, \quad x^3 = \varphi = \bar{x}^3.$$

(Note that the g_{11} *and* g_{22} *of Ex. (5.1) are now* $\bar{g}_{22}$ *and* $\bar{g}_{33}$*, respectively.)*

Let us turn now to the second of our two problems, namely, defining a condition of parallelism between vectors in the same small neighborhood. The condition, whatever it is, should be one which may be expressed so as to be true in any or all possible coordinate systems. Such would be the case if it were stated as a vector or tensor equation. For the sake of consistency, we wish also to require that the condition reduce to the condition for parallelism in the limit as the surface becomes a plane, wholly or in part. These conditions together lead us to choose the condition that a set of vectors p^i form a parallel vector field if

$$\frac{\delta p^i}{\delta t} = \alpha p^i, \tag{5.3}$$

where t is some invariant parameter. If we wish to dispense with unnecessary generalities and confine our attention to vector fields of constant magnitude (a well-defined notion once the fundamental tensor has been defined), we then have $\alpha = 0$ in equation (5.3). Hence, taking $dt = ds$,

$$\frac{dp^i}{ds} = -\left\{ {i \atop j\ k} \right\} p^j \frac{dx^k}{ds} \tag{5.4}$$

prescribes the change dp^i in the components of a vector when displaced parallely from x^i to $x^i + dx^i$, $ds = |dx^i|$. Increment by increment, a one-parameter vector field may be constructed from a vector at any given point by using equation (5.4). This solves our second problem.

It should be clear in retrospect that we are now able to define at every point a standard of length and a criterion of parallelism with respect to vectors *in the same neighborhood*. It is imperative to recognize, however, that the qualifying phrase "in the same neighborhood" has consequences not to be foreseen at a glance. Though it is an unnecessary limitation when the surface is a Euclidean plane, it is an indispensable qualifier otherwise. The sections to follow will attempt to bring out its full significance. Before undertaking to do so, however, we may be content in having defined intrinsic and covariant differentiation on a surface; we first give them application in a famous problem, one whose solution will further ratify the assumptions by which we have generalized, step by step, the algebra and calculus of vectors from rectilinear coordinates in the plane to generalized coordinates on a surface.

Ex. (5.3) (a) On the unit sphere, find the components of the unit tangent λ^i *along a parallel of latitude. (b) Determine whether or not these tangents constitute a parallel field along a parallel of latitude. (Hint: calculate* $\frac{\delta \lambda^i}{\delta s}$ *.)*

Ans. (a) $\lambda^i = (\sec\varphi, 0)$ *. (b)* $\frac{\delta \lambda^i}{\delta s} = (0, \tan\varphi)$ *. Since this is not zero, the unit tangents are not a parallel vector field. The sole exception is the case* $\varphi = 0$ *.*

Ex. (5.4) (a) On the unit sphere, find the components of the unit tangent along a meridian of longitude. (b) Determine whether or not these tangents constitute a parallel vector field.

Ans. (a) $\lambda^i = (0, 1)$ *. (b)* $\frac{\delta \lambda^i}{\delta s} = (0, 0)$ *, hence the* λ^i *are a parallel vector field.*

Ex. (5.5) Verify the Frenet formulae of Ex. (4.7) for curves on curved surfaces.

6. Geodesics

Consider a surface on which coordinates x^i **have been defined (Fig. 44). Let** A and B be two points on the surface, connected by a curve $C(t)$, where t is a curve parameter. We propose to find the equation satisfied by $x^i(t)$ the coordinates of that curve C for which the distance from A to B is a minimum.

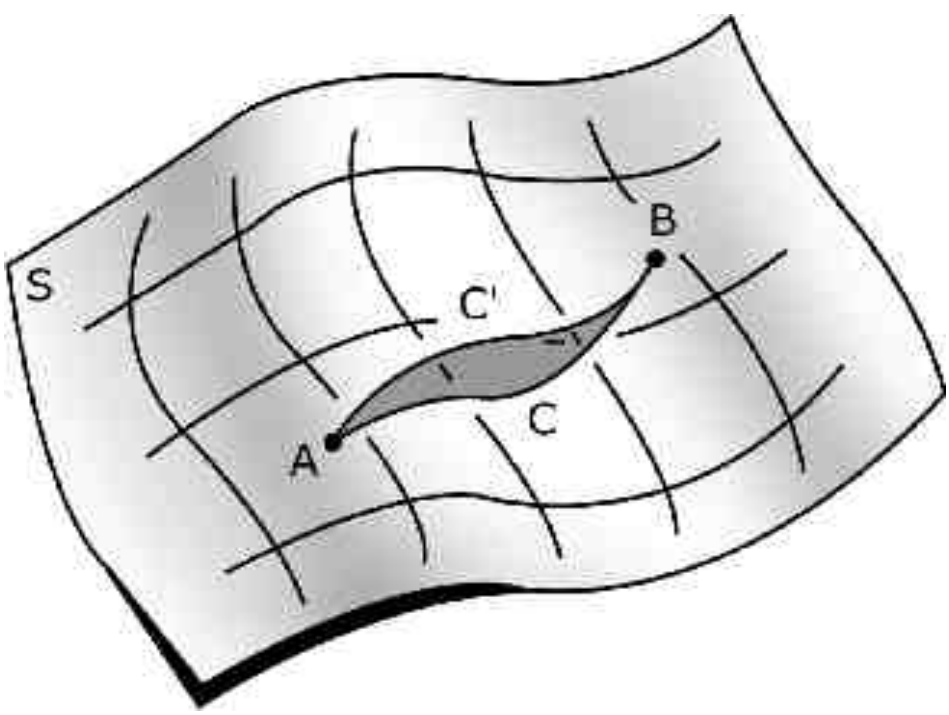

Figure 44

For this purpose, consider any other nearby curve along which the coordinates are $x'^i(t) = x^i(t) + \varepsilon y^i(t)$, the quantities $y^i(t)$ vanishing at $\mathbf{A}$ and $\mathbf{B}$; here ε is a parameter independent of t. The arc length from $\mathbf{A}$ to $\mathbf{B}$ along any curve $\mathbf{C}'$ is then

$$s' = \int_{\mathbf{A}}^{\mathbf{B}} ds' = \int_{\mathbf{C}'} \left(g_{ij}\, dx'^i\, dx'^j\right)^{1/2} = \int_{\mathbf{C}'} \left(g_{ij}\, \dot{x}'^i \dot{x}'^j\right)^{1/2} dt,$$

where dots denote derivatives with respect to t. Now by varying ε, we may vary s'; that is, $s' = s'(\varepsilon)$. By hypothesis, $s(0)$ is the arc length of $\mathbf{C}$, the curve of minimum length. But at the same time, this must be the value of $s'(\varepsilon)$ where

$$\frac{ds'}{d\varepsilon} = 0.$$

Since we require ε to be small, it should be possible to expand $s'(\varepsilon)$ in a Taylor's series. Putting

$$\begin{aligned} \mathrm{L}(\varepsilon) &= \left(g_{ij}\, \dot{x}'^i \dot{x}'^j\right)^{1/2} = \mathrm{L}_0 + \varepsilon \left(\frac{\partial \mathrm{L}}{\partial \varepsilon}\right)_0 + \frac{\varepsilon^2}{2}\left(\frac{\partial^2 \mathrm{L}}{\partial \varepsilon^2}\right)_0 + \cdots \\ &= \left(g_{ij}\, \dot{x}^i \dot{x}^j\right)^{1/2} + \varepsilon \left[\frac{\partial \mathrm{L}}{\partial x^i} y^i + \frac{\partial \mathrm{L}}{\partial \dot{x}^i} \dot{y}^i\right] + \cdots, \end{aligned}$$

we have at $\varepsilon = 0$

$$\begin{aligned} \left(\frac{ds}{d\varepsilon}\right)_0 &= \int \left(\frac{\partial \mathrm{L}}{\partial \varepsilon}\right)_0 dt = \int \left[\frac{\partial \mathrm{L}}{\partial x^i} y^i + \frac{\partial \mathrm{L}}{\partial \dot{x}^i} \dot{y}^i\right] dt \\ &= \frac{\partial \mathrm{L}}{\dot{x}^i} y^i \Big|_{\mathbf{A}}^{\mathbf{B}} + \int \left[\frac{\partial \mathrm{L}}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial \mathrm{L}}{\partial \dot{x}^i}\right)\right] y^i\, dt, \end{aligned}$$

where we have carried out one integration by parts. Since $y^i = 0$ at $\mathbf{A}$ and $\mathbf{B}$, the first term vanishes. For the curve $\mathbf{C}$ of minimum length, the remainder must be zero no matter what the functions y^i may be. This can be true only if the quantities in the square brackets vanish, i. e., if

$$\frac{\partial \mathrm{L}}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial \mathrm{L}}{\partial \dot{x}^i}\right) = 0. \tag{6.1}$$

This, then, is the differential equation of the curve $\mathbf{C}$ of minimal length between $\mathbf{A}$ and $\mathbf{B}$. Such a curve is called a **geodesic** of the surface. Equation (6.1) is the equation of the geodesics. The expression on the left is also sometimes called the **Lagrangian** or **Eulerian derivative** of L.

Now consider the special case when $dt = ds$, the differential of arc length. Then

$$\mathrm{L} = \left(g_{ij} \frac{dx^i}{ds} \frac{dx^j}{ds}\right)^{1/2} = 1.$$

In this particularly simple instance we have

$$\frac{\partial L}{\partial x^r} = \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial x^r}\dot{x}^i\dot{x}^j\right)L^{-1} = \frac{1}{2}\frac{\partial g_{ij}}{\partial x^r}\dot{x}^i\dot{x}^j,$$

$$\frac{\partial L}{\partial \dot{x}^r} = \left(\frac{1}{2}2\,g_{rj}\dot{x}^j\right)L^{-1} = g_{rj}\dot{x}^j.$$

Therefore

$$-\frac{\partial L}{\partial x^r} + \frac{d}{ds}\left(\frac{\partial L}{\partial \dot{x}^r}\right) = g_{rj}\ddot{x}^j + \frac{\partial g_{rj}}{\partial x^n}\dot{x}^j\dot{x}^n - \frac{1}{2}\frac{\partial g_{ij}}{\partial x^r}\dot{x}^i\dot{x}^j$$

$$= g_{rj}\ddot{x}^j + \frac{1}{2}\left[\frac{\partial g_{rj}}{\partial x^n} + \frac{\partial g_{nr}}{\partial x^j} - \frac{\partial g_{nj}}{\partial x^r}\right]\dot{x}^j\dot{x}^n = g_{rj}\ddot{x}^j + [nj,r]\dot{x}^n\dot{x}^j.$$

Multiplying both sides by g^{pr} gives

$$g^{pr}\left[\frac{d}{ds}\left(\frac{\partial L}{\partial x^r}\right) - \frac{\partial L}{\partial x^r}\right] = \delta^p_j\ddot{x}^j + \left\{{p \atop nj}\right\}\dot{x}^n\dot{x}^j$$

$$= \ddot{x}^p + \left\{{p \atop nj}\right\}\dot{x}^n\dot{x}^j = \frac{\delta\dot{x}^p}{\delta s} = 0.$$

Hence the equation of a geodesic in the surface S is

$$\frac{\delta}{\delta s}\left(\frac{dx^p}{ds}\right) = 0, \tag{6.2}$$

the same equation as the condition for parallelism of the unit tangent vectors and for zero curvature of the curve. Thus a geodesic is (1) a curve of minimum length between two given points or (2) a curve whose tangents form a parallel field. We can show also that it is (3) a curve of zero curvature; the demonstration is identical with that for curves in a plane.

Ex. (6.1) Find the equations of a geodesic on the surface of a unit sphere.

Ans. $L = (\cos^2\varphi)\dot{\theta}^2 + \dot{\varphi}^2$, $\dfrac{\partial L}{\partial \dot{\theta}} = 2(\cos^2\varphi)\dot{\theta}$, $\dfrac{\partial L}{\partial \theta} = 0$,

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

$$\frac{\partial L}{\partial \dot{\varphi}} = 2\dot{\varphi},\quad \frac{\partial L}{\partial \varphi} = -2[\sin\varphi\cos\varphi]\dot{\theta}^2,$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\varphi}}\right) - \frac{\partial L}{\partial \varphi} = 2[\ddot{\varphi} + (\sin\varphi\cos\varphi)\dot{\theta}^2] = 0.$$

Ex. (6.2) Compare the results of Ex. (6.1) with the equations $\dfrac{\delta \lambda^i}{\delta t} = 0$, *where* λ^i *is a unit tangent.*

Ex. (6.3) Find the equations of a geodesic on the surface of a spheroid. (Hint: use the results of Ex. (5.1) to determine L, *then apply equation (6.1).)*

Ex. (6.4) Show that the equations of Ex. (6.1) are satisfied by the solutions

$$\sin\varphi = \beta \sin s, \quad \cos\varphi \sin(\theta - \theta_0) = \sqrt{1 - \beta^2}\sin s.$$

When the surface is a plane and the coordinates are rectilinear, equation (6.2) becomes $d^2x^i/ds^2 = 0$, the solution of which is a straight line. Therefore geodesics in the plane are straight lines, as we would expect. (On this account, geodesics are sometimes loosely referred to as "straight lines" even in surfaces which are not plane.) This result adds the weight of *ex post facto* justification for our assumptions concerning the relations of vectors in curved surfaces to those in the plane.

Geodesics also afford the means of carrying through the proofs of many results in a much more elegant manner than might otherwise be possible. In many cases, the means for doing so is the employment of a special class of coordinates known as **geodesic coordinates**. These are coordinates in which the Christoffel symbols are zero at a particular point O. Only in rectilinear plane coordinates are the fundamental tensors constant everywhere and therefore the Christoffel symbols zero everywhere. By a particular transformation, however, the Christoffel symbols may all be made zero at O, in the neighborhood of which, therefore, the coordinates x^i approximate plane coordinates.

Thus, consider two coordinate systems x^i and $\bar{x}^i$. The system $\bar{x}^i$ is given and the system x^i is to be found. By equation (2.5), the relation of their Christoffel symbols is

$$\left\{ {m \atop pq} \right\} = \frac{\partial x^m}{\partial \bar{x}^j}\frac{\partial^2 \bar{x}^j}{\partial x^p \partial x^q} + \overline{\left\{ {j \atop kl} \right\}}\frac{\partial x^m}{\partial \bar{x}^j}\frac{\partial \bar{x}^k}{\partial x^p}\frac{\partial \bar{x}^l}{\partial x^q}. \tag{6.3}$$

If we are to have $\left\{ {m \atop pq} \right\} = 0$ at a point O, then

$$\frac{\partial^2 \bar{x}^j}{\partial x^p \partial x^q} + \overline{\left\{ {j \atop kl} \right\}}\frac{\partial x^m}{\partial \bar{x}^j}\frac{\partial \bar{x}^k}{\partial x^p}\frac{\partial \bar{x}^l}{\partial x^q} = 0.$$

If we take

$$\bar{x}^i = x^i - \frac{1}{2}\overline{\left\{ {i \atop kl} \right\}}_{(0)}(x^k - x^k{}_{(0)})(x^l - x^l{}_{(0)}),$$

where $x^i = x^i_{(0)}$ corresponds to the point O and where the subscripts (0) denote evaluation at the point O, then

$$\frac{\partial \bar{x}^i}{\partial x^m} = \delta^i_m \quad \text{and} \quad \left(\frac{\partial^2 \bar{x}^j}{\partial x^p \partial x^q} \right)_{(0)} = -\overline{\left\{ {j \atop pq} \right\}}_{(0)}$$

to the first order in x^i. Inserted into equation (6.3), this makes all $\left\{ {m \atop pq} \right\} = 0$ at O. Hence the coordinate system x^i is a geodesic coordinate system.

Ex. (6.5) Show that for any orthogonal coordinate system

(a) $g_{12} = 0, \quad g^{12} = 0, \quad g^{11} = \dfrac{1}{g_{11}}, \quad g^{22} = \dfrac{1}{g_{22}}.$

(b) $[11,1] = \dfrac{1}{2}\dfrac{\partial g_{11}}{\partial x^1}, \; [12,1] = [21,1] = \dfrac{1}{2}\dfrac{\partial g_{11}}{\partial x^2} = -[11,2],$

$$[22,1] = -\frac{1}{2}\frac{\partial g_{22}}{\partial x^1} = -[12,2] = -[21,2], \quad [22,2] = \frac{1}{2}\frac{\partial g_{22}}{\partial x^2}.$$

$$\left\{ {1 \atop 11} \right\} = \frac{1}{2g_{11}}\frac{\partial g_{11}}{\partial x^1}, \quad \left\{ {1 \atop 21} \right\} = \left\{ {1 \atop 12} \right\} = \frac{1}{2g_{11}}\frac{\partial g_{11}}{\partial x^2}, \quad \left\{ {1 \atop 22} \right\} = -\frac{1}{2g_{11}}\frac{\partial g_{22}}{\partial x^1},$$

$$\left\{ {2 \atop 11} \right\} = -\frac{1}{2g_{22}}\frac{\partial g_{11}}{\partial x^2}, \quad \left\{ {2 \atop 12} \right\} = \left\{ {2 \atop 21} \right\} = \frac{1}{2g_{22}}\frac{\partial g_{22}}{\partial x^1}, \quad \left\{ {2 \atop 22} \right\} = \frac{1}{2g_{22}}\frac{\partial g_{22}}{\partial x^2}.$$

Ex. (6.6) Find the Christoffel symbols of the second kind on the surface of a spheroid. Use longitude θ *and geocentric latitude* φ *as variables. (Hint: see Ex. (5.1).)*

Ans.

$$\left\{ {1 \atop 12} \right\} = \left\{ {1 \atop 21} \right\} = -\tan\varphi,$$

$$\left\{ {2 \atop 11} \right\} = \sin\varphi\cos\varphi\left[\frac{1 - e^2\cos^2\varphi}{1 - (2e^2 - e^4)\cos^2\varphi}\right],$$

$$\left\{ {2 \atop 22} \right\} = e^2\sin\varphi\cos\varphi\left\{\frac{(2 - e^2)(1 + 2e^2\cos^2\varphi) - 3}{(1 - e^2\cos^2\varphi)[1 - (2e^2 - e^4)\cos^2\varphi]}\right\}.$$

Ex. (6.7) Determine the geodesic coordinates $\bar{x}^i$ *in the neighborhood of a point* $\bar{x}^i_{(0)} = (\theta_{(0)}, \varphi_{(0)})$ *on the surface of a spheroid.*

Ans. $\bar{x}^1 - \bar{x}^1_{(0)} = \theta - \theta_{(0)} = x^1 + (\tan\varphi_{(0)})x^1x^2,$

$$\bar{x}^2 - \bar{x}^{(0)} = \varphi - \varphi_{(0)} = x^2 - \frac{1}{2}\frac{\sin\varphi_{(0)}\cos\varphi_{(0)}}{1 - (2e^2 - e^4)\cos^2\varphi_{(0)}}\Big[(1 - e^2\cos^2\varphi_{(0)})(x^1)^2$$
$$+ \frac{e^2}{1 - e^2\cos^2\varphi_{(0)}}\left\{(2 - e^2)(1 + 2e^2\cos^2\varphi_{(0)}) - 3\right\}(x^2)^2\Big].$$

Slightly more general than geodesic coordinates is a particular class called geodesic polar coordinates; they are not limited to the neighborhood of a particular point. To construct geodesic polar coordinates at a point O (see Fig. 45), take the x^1– curves to be all geodesics through O and let x^1 be the arc length along any one, measured from O. Let x^2 be the angle at O between two geodesics, measured from some one x^1– curve designated as $x^2 = 0$. The similarity to plane polar coordinates is clear.

Along any particular x^1– curve we have $dx^1 = ds$, $dx^2 = 0$. From this we infer that

(6.4) $$g_{11} = 1 .$$

Since the x^1– curves are also geodesics, at every point on them it must be true that

$$\frac{d^2x^1}{ds^2} + \begin{Bmatrix} 1 \\ jk \end{Bmatrix} \frac{dx^j}{ds}\frac{dx^k}{ds} = 0, \quad \frac{d^2x^2}{ds^2} + \begin{Bmatrix} 2 \\ jk \end{Bmatrix} \frac{dx^j}{ds}\frac{dx^k}{ds} = 0 .$$

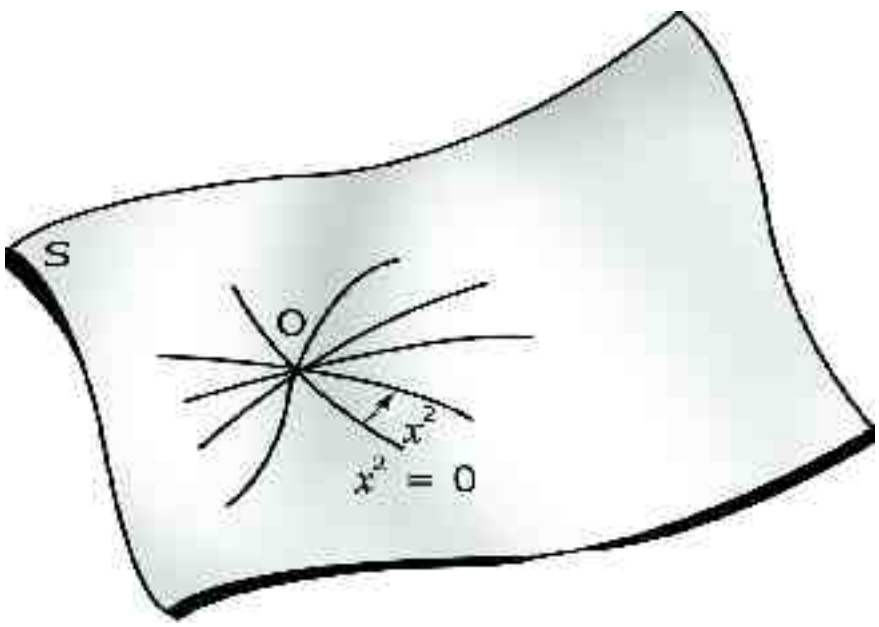

Figure 45

From these we infer also that $\begin{Bmatrix} 1 \\ 11 \end{Bmatrix} = 0$, $\begin{Bmatrix} 2 \\ 11 \end{Bmatrix} = 0$, whence $[11, i] = g_{ij}\begin{Bmatrix} j \\ 11 \end{Bmatrix} = 0$. In particular,

$$[11, 2] = \frac{1}{2}\left\{\frac{\partial g_{12}}{\partial x^1} + \frac{\partial g_{21}}{\partial x^1} - \frac{\partial g_{11}}{\partial x^2}\right\} = \frac{\partial g_{12}}{\partial x^1} = 0$$

everywhere. Therefore by considering the function g_{12} along any x^1– curve, we see that it has a constant value everywhere. In particular, it has whatever value it assumes at O. But here $(ds)^2 = (dx^1)^2$ no matter what the value of dx^2.

Therefore $g_{22} = 0$ at O and $g_{12} = 0 = g_{21}$ everywhere; the x^1-curves are therefore orthogonal to the x^2-curves. In geodesic polar coordinates, then, the line element always has the form

$$(ds)^2 = (dx^1)^2 + g_{22}(dx^2)^2$$

where g_{22} vanishes at the origin.

Ex. (6.8) Show that in a geodesic polar coordinate system at most $\begin{Bmatrix} 1 \\ 22 \end{Bmatrix}$, $\begin{Bmatrix} 2 \\ 12 \end{Bmatrix} = \begin{Bmatrix} 2 \\ 21 \end{Bmatrix}$, *and* $\begin{Bmatrix} 2 \\ 22 \end{Bmatrix}$ *are not zero.*

Ex. (6.9) Show that on the unit sphere $x^1 = 90° - \varphi = \chi$ *(the co-latitude) and* $x^2 = \theta$ *(longitude) constitute a geodesic polar coordinate system.*

Ans. Through the pole, the curves $x^2 =$ ***constant*** *are great circles, hence geodesics on the sphere. In addition,*

$$g_{ij} = \begin{Vmatrix} 1 & 0 \\ 0 & \sin^2 x^1 \end{Vmatrix},$$

whence $g_{22}(0) = 0$.

Ex. (6.10) Show that in a geodesic polar coordinate system the vector $u^i = (1, 0)$ *is at every point a unit vector tangent to the* x^1-*curve.*

Ex. (6.11) Show that in a geodesic polar coordinate system the vector $u^i = \left(\cos\alpha, \dfrac{\sin\alpha}{\sqrt{g_{22}}} \right)$ *is at every point a unit vector at an angle* α *to the* x^1-*curve.*

7. The Riemann-Christoffel Tensor

We have seen that covariant and intrinsic differentiation of vectors and tensors provide the vector and tensor counterparts of partial and ordinary differentiation of functions. These counterparts obey the familiar rules for the differentiation of scalars, sums, differences, and products without sacrificing vectorial or tensorial character. As already noted, however, each component of covariant and intrinsic derivatives is a function of *all* components, unlike derivatives of sets of functions. With such a precedent in mind, it should not be greatly surprising, then, to discover yet another difference, one of great consequence: *the order of covariant differentiation is in general not interchangeable*, even when the components have continuous derivatives to any required order.

To demonstrate this result, consider the covariant derivative of some vector X_r. It is

$$X_{r,s} = \frac{\partial X_r}{\partial x^s} - \begin{Bmatrix} m \\ rs \end{Bmatrix} X_m .$$

The covariant derivative of $\mathbf{X}_{r,s}$ is in turn

$$\mathbf{X}_{r,st} = \frac{\partial \mathbf{X}_{r,st}}{\partial x^t} - \begin{Bmatrix} m \\ rt \end{Bmatrix} \mathbf{X}_{m,s} - \begin{Bmatrix} m \\ st \end{Bmatrix} \mathbf{X}_{r,m}$$

$$= \frac{\partial^2 \mathbf{X}_r}{\partial x^s \partial x^t} - \begin{Bmatrix} m \\ rs \end{Bmatrix} \frac{\partial \mathbf{X}_m}{\partial x^t} - \begin{Bmatrix} m \\ rt \end{Bmatrix} \frac{\partial \mathbf{X}_m}{\partial x^s} - \begin{Bmatrix} m \\ st \end{Bmatrix} \frac{\partial \mathbf{X}_r}{\partial x^m}$$

$$- \mathbf{X}_m \left[\frac{\partial}{\partial x^t} \begin{Bmatrix} m \\ rs \end{Bmatrix} - \begin{Bmatrix} m \\ rp \end{Bmatrix} \begin{Bmatrix} p \\ st \end{Bmatrix} - \begin{Bmatrix} m \\ sp \end{Bmatrix} \begin{Bmatrix} p \\ rt \end{Bmatrix} \right].$$

We now find the part of $\mathbf{X}_{r,st}$ antisymmetric in s and t ; if it is not identically zero, then the order of differentiation must be consequential. The part of $\mathbf{X}_{r,st}$ antisymmetric in s and t is

$$\mathbf{X}_{r,st} - \mathbf{X}_{r,ts} = \delta^{pq}_{st} \mathbf{X}_{r,pq} = \left(\frac{\partial^2 \mathbf{X}_r}{\partial x^s \partial x^t} - \frac{\partial^2 \mathbf{X}_r}{\partial x^t \partial x^s} \right)$$

$$- \left(\begin{Bmatrix} m \\ rs \end{Bmatrix} \frac{\partial \mathbf{X}_m}{\partial x^t} - \begin{Bmatrix} m \\ rt \end{Bmatrix} \frac{\partial \mathbf{X}_m}{\partial x^s} + \begin{Bmatrix} m \\ rt \end{Bmatrix} \frac{\partial \mathbf{X}_m}{\partial x^s} - \begin{Bmatrix} m \\ rs \end{Bmatrix} \frac{\partial \mathbf{X}_m}{\partial x^t} \right)$$

$$- \left(\begin{Bmatrix} m \\ st \end{Bmatrix} - \begin{Bmatrix} m \\ ts \end{Bmatrix} \right) \frac{\partial \mathbf{X}_r}{\partial x^m} - \left[\frac{\partial}{\partial x^t} \begin{Bmatrix} m \\ rs \end{Bmatrix} - \frac{\partial}{\partial x^s} \begin{Bmatrix} m \\ rt \end{Bmatrix} \right] \mathbf{X}_m$$

$$- \left[\begin{Bmatrix} m \\ rq \end{Bmatrix} \begin{Bmatrix} q \\ st \end{Bmatrix} - \begin{Bmatrix} m \\ rq \end{Bmatrix} \begin{Bmatrix} q \\ ts \end{Bmatrix} + \begin{Bmatrix} m \\ sq \end{Bmatrix} \begin{Bmatrix} q \\ rt \end{Bmatrix} - \begin{Bmatrix} m \\ tq \end{Bmatrix} \begin{Bmatrix} q \\ rs \end{Bmatrix} \right] \mathbf{X}_m .$$

Because the order of partial differentiation is immaterial for functions with continuous derivatives and because $\begin{Bmatrix} m \\ st \end{Bmatrix} = \begin{Bmatrix} m \\ ts \end{Bmatrix}$, the only terms which survive on the right hand side are

$$\left[\frac{\partial}{\partial x^s} \begin{Bmatrix} m \\ rt \end{Bmatrix} - \frac{\partial}{\partial x^t} \begin{Bmatrix} m \\ rs \end{Bmatrix} + \begin{Bmatrix} m \\ sq \end{Bmatrix} \begin{Bmatrix} q \\ rt \end{Bmatrix} - \begin{Bmatrix} m \\ tq \end{Bmatrix} \begin{Bmatrix} q \\ rs \end{Bmatrix} \right] \mathbf{X}_m = \mathbf{R}^m_{rst} \mathbf{X}_m .$$

Since this is true for *any* vector $\mathbf{X}_m$ and *any* coordinate system, the quantity

$$\mathbf{R}^m_{rst} = \frac{\partial}{\partial x^s} \begin{Bmatrix} m \\ rt \end{Bmatrix} - \frac{\partial}{\partial x^t} \begin{Bmatrix} m \\ rs \end{Bmatrix} + \begin{Bmatrix} m \\ sq \end{Bmatrix} \begin{Bmatrix} q \\ rt \end{Bmatrix} - \begin{Bmatrix} m \\ tq \end{Bmatrix} \begin{Bmatrix} q \\ rs \end{Bmatrix} \tag{7.1}$$

must characterize what is independent of both of these, the surface itself. By the Quotient Theorem, $\mathbf{R}^m_{rst}$ is a tensor of the fourth order; it is the **Riemann Christoffel tensor**. It will be seen from equation (7.1) to depend only on the components of the fundamental tensor and their first and second derivatives.

It will be one of our ultimate purposes to explore the meaning of the Riemann-Christoffel tensor. This will require a development of some length. As a first step, we carry through several algebraic preliminaries. First, we form the wholly covariant form of the Riemann-Christoffel tensor, namely, $\mathrm{R}_{prst} = g_{pm}\mathrm{R}^{m}_{rst}$. Since

$$g_{pm}\frac{\partial}{\partial x^s}\begin{Bmatrix} m \\ rt \end{Bmatrix} = \frac{\partial}{\partial x^s}\left(g_{pm}\begin{Bmatrix} m \\ rt \end{Bmatrix}\right) - \begin{Bmatrix} m \\ rt \end{Bmatrix}\frac{\partial g_{pm}}{\partial x^s}$$
$$= \frac{\partial}{\partial x^s}[rt,p] - \begin{Bmatrix} m \\ rt \end{Bmatrix}\big([ps,m] + [ms,p]\big),$$

we get on substituting expressions of this form into (7.1) the result

(7.2)
$$\mathrm{R}_{prst} = \frac{\partial}{\partial x^s}[rt,p] - \frac{\partial}{\partial x^t}[rs,p] + \begin{Bmatrix} m \\ rs \end{Bmatrix}[pt,m] - \begin{Bmatrix} m \\ rt \end{Bmatrix}[ps,m].$$

Inserting the expressions for the Christoffel symbols and collecting terms, we have finally that

(7.3)
$$\mathrm{R}_{prst} = \frac{1}{2}\left(\frac{\partial^2 g_{pt}}{\partial x^r \partial x^s} + \frac{\partial^2 g_{rs}}{\partial x^p \partial x^t} - \frac{\partial^2 g_{ps}}{\partial x^r \partial x^t} - \frac{\partial^2 g_{rt}}{\partial x^p \partial x^s}\right)$$
$$+ g^{mn}\big([rs,m][pt,n] - [rt,m][ps,n]\big).$$

Ex. (7.1) Show that

(7.4) $$\mathrm{R}^{k}_{rst} = \delta^{jm}_{st}\, e^{(k)}_{r,jm},$$

where $e^{(k)}_{r}$ is the covariant basis vector along the x^k– curve. (Hint: refer to the results of Ex. (3.9).)

From the original definition we could have seen that R_{prst} is antisymmetric in *s* and *t*, as equation (7.3) shows. We can see further from equation (7.3) that

(7.5) $$\mathrm{R}_{prst} = -\mathrm{R}_{rpst} = -\mathrm{R}_{prts} = \mathrm{R}_{stpr}.$$

Thus R_{prst} is antisymmetric to interchange of *p* and *r* and of *s* and *t* , though symmetric to the interchange of the pair *p r* with the pair *s t*.

Let us now define from R_{prst} the invariant

(7.6) $$\mathrm{K} = \frac{1}{4}\varepsilon^{pr}\varepsilon^{st}\mathrm{R}_{prst}.$$

The invariant **K** is called the **Gaussian curvature.** We may easily recover from it the Riemann-Christoffel tensor R_{prst}, for

$$K\,\varepsilon_{ij}\,\varepsilon_{kl} = \left(\frac{1}{4}\,\varepsilon^{pr}\,\varepsilon^{st}\,R_{prst}\right)\varepsilon_{ij}\,\varepsilon_{kl} = \frac{1}{2}\,\varepsilon^{pr}\,\varepsilon^{ij}\left(\frac{1}{2}\,\varepsilon^{st}\,\varepsilon_{kl}\,R_{prst}\right)$$

$$= \delta^{pr}_{ij}\left(\frac{1}{2}\,\delta^{st}_{kl}\,R_{prst}\right) = \frac{1}{2}\,\delta^{pr}_{ij}\cdot\frac{1}{2}\left(R_{prkl} - R_{prlk}\right)$$

$$= \frac{1}{2}\,\delta^{pr}_{ij}\,R_{prkl} = \frac{1}{2}\left(R_{ijkl} - R_{jikl}\right) = R_{ijkl}.$$

Hence

$$R_{ijkl} = K\,\varepsilon_{ij}\,\varepsilon_{kl} = g\,K\,e_{ij}\,e_{kl}. \tag{7.7}$$

With the Riemann-Christoffel tensor in this form, we can derive a result of much interest. We first write

$$\varepsilon_{ij}\,\varepsilon_{kl} = (g_{ip}\,g_{jq}\,\varepsilon^{pq})\,\varepsilon_{kl} = g_{ip}\,g_{jq}\,\delta^{pk}_{kl} = g_{ik}\,g_{jl} - g_{il}\,g_{jk}.$$

Thus

$$R_{ijkl} = K\,(g_{ik}\,g_{jl} - g_{il}\,g_{jk}). \tag{7.8}$$

We see, therefore, that *in two dimensions, the Riemann-Christoffel tensor is always isotropic*.

The Riemann-Christoffel tensor in its general forms (7.1), (7.2) or (7.3) gives little hint as to its geometrical significance, nor is it especially convenient for the purposes of actual calculation. It is very useful, therefore, to consider the form which it takes in an orthogonal coordinate system. In such a system, we have (Ex. (6.5))

$$\begin{Bmatrix}1\\11\end{Bmatrix} = \frac{1}{2\,g_{11}}\frac{\partial g_{11}}{\partial x^1},\quad \begin{Bmatrix}1\\12\end{Bmatrix} = \frac{1}{2\,g_{11}}\frac{\partial g_{11}}{\partial x^2}, \tag{7.9}$$

$$\begin{Bmatrix}1\\22\end{Bmatrix} = -\frac{1}{2\,g_{11}}\frac{\partial g_{22}}{\partial x^1},\quad \begin{Bmatrix}2\\11\end{Bmatrix} = -\frac{1}{2\,g_{22}}\frac{\partial g_{22}}{\partial x^2},$$

$$\begin{Bmatrix}2\\21\end{Bmatrix} = \frac{1}{2\,g_{11}}\frac{\partial g_{22}}{\partial x^1},\quad \begin{Bmatrix}2\\22\end{Bmatrix} = \frac{1}{2\,g_{22}}\frac{\partial g_{22}}{\partial x^2}.$$

Noting that apart from sign, the only non-zero components of the Riemann-Christoffel tensor (in two dimensions) are equal to R_{1212}, we find by substituting (7.9) into equation (7.2) that

$$\mathrm{R}_{1212} = -\frac{1}{2}\left[\frac{\partial^2 g_{22}}{(\partial x^1)^2} + \frac{\partial^2 g_{11}}{(\partial x^2)^2}\right] \tag{7.10}$$

$$+\frac{1}{4}\left\{\frac{1}{g_{11}}\left(\frac{\partial g_{11}}{\partial x^2}\right)^2 + \frac{1}{g_{22}}\left(\frac{\partial g_{22}}{\partial x^1}\right)^2 + \frac{1}{g_{11}}\frac{\partial g_{22}}{\partial x^1}\frac{\partial g_{11}}{\partial x^1} + \frac{1}{g_{22}}\frac{\partial g_{22}}{\partial x^2}\frac{\partial g_{11}}{\partial x^2}\right\}$$

$$-\frac{1}{2}\left[\frac{\partial^2 g_{22}}{(\partial x^1)^2} - \frac{1}{2g_{11}g_{22}}\frac{\partial g_{22}}{\partial x^1}\left\{g_{11}\frac{\partial g_{22}}{\partial x^1} + g_{22}\frac{\partial g_{11}}{\partial x^1}\right\}\right.$$

$$\left.+\frac{\partial^2 g_{11}}{(\partial x^2)^2} - \frac{1}{2g_{11}g_{22}}\frac{\partial g_{11}}{\partial x^2}\left\{g_{22}\frac{\partial g_{11}}{\partial x^2} + g_{11}\frac{\partial g_{22}}{\partial x^2}\right\}\right]$$

$$= -\frac{\sqrt{g}}{2}\left[\frac{1}{\sqrt{g}}\frac{\partial^2 g_{22}}{(\partial x^1)^2} - \frac{\partial g_{22}}{\partial x^1}\left\{\frac{1}{2g^{3/2}}\frac{\partial g}{\partial x^1}\right\} + \frac{1}{\sqrt{g}}\frac{\partial^2 g_{11}}{(\partial x^2)^2} - \frac{\partial g_{11}}{\partial x^2}\left\{\frac{1}{2g^{3/2}}\frac{\partial g}{\partial x^2}\right\}\right]$$

$$= -\frac{\sqrt{g}}{2}\left[\frac{\partial}{\partial x^1}\left(\frac{1}{\sqrt{g}}\frac{\partial g_{22}}{\partial x^1}\right) + \frac{\partial}{\partial x^2}\left(\frac{1}{\sqrt{g}}\frac{\partial g_{11}}{\partial x^2}\right)\right].$$

The simplest possible application of equation (7.10) is to the surface of a sphere of radius *a*, where from equation (5.1) it is evident that

$$x^1 = \theta, \quad x^2 = \varphi, \quad g_{11} = (a)^2\cos^2 x^2, \quad g_{22} = (a)^2.$$

Then for the sphere

$$\mathrm{R}_{1212} = -\frac{(a)^2\cos x^2}{2}\frac{\partial}{\partial x^2}\left(\frac{1}{(a)^2\cos x^2}\frac{\partial[(a)^2\cos^2 x^2]}{\partial x^2}\right)$$

$$= (a)^2\cos^2 x^2 = (a)^2\cos^2\varphi.$$

From equation (7.8), we now see further that

$$\mathrm{K} = \frac{\mathrm{R}_{1212}}{g} = \frac{1}{a^2}. \tag{7.11}$$

Ex. (7.2) Show that the Gaussian curvature on the surface of a spheroid is

$$\mathrm{K} = \frac{(1-e^2)}{\rho^2}\frac{(1-e^2\cos^2\varphi)^2}{[1-(2e^2-e^4)\cos^2\varphi]^2}.$$

Ex. (7.3) Show that

$$\left|\mathrm{R}_{1212}\right| = \left[g^{ip}g^{jq}g^{kr}g^{ls}\mathrm{R}_{ijkl}\mathrm{R}_{pqrs}\right]^{1/2} = |\mathrm{K}|.$$

(Hint: use equation (7.8).)

Ex. (7.4) Calculate $\mathbf{R}_{1212}$ *in geodesic polar coordinates on the surface of a sphere. (Hint: see Ex. (6.9).)*

Ex. (7.5) Show that if the coordinate curves on a surface are both orthogonal and geodesic, then the curvature tensor vanishes. (Hint: select a geodesic polar coordinate system and use equation (7.10).)

Ans. Since only g_{22} *may not be constant in a geodesic polar coordinate system, we consider the field of tangent vectors* $\lambda^i = (0, \frac{1}{\sqrt{g_{22}}})$ *to the* x^2*-curve. Since it is geodesic,*

$$\frac{\delta\lambda^1}{\delta s} = \begin{Bmatrix} 1 \\ 2\,2 \end{Bmatrix} \frac{1}{g_{22}} = 0, \quad \begin{Bmatrix} 1 \\ 2\,2 \end{Bmatrix} = -\frac{1}{2g_{11}} \frac{\partial g_{22}}{\partial x^1} = 0, \quad \textit{etc.}$$

It may be recognized from this that the invariant $\mathbf{K}$ **is equal to the** curvature of the surface of the sphere. This is the simplest identification of $\mathbf{K}$ and one which prompts its being designated as the **total curvature** or **Gaussian curvature** of the surface. Such an identification is not limited to the surface of a sphere. The intimate relation of $\mathbf{R}_{1212}$ with $\mathbf{K}$ then naturally favors the terminology that $\mathbf{R}_{1212}$ is the **curvature tensor.**

Further results of immediate interest may be extracted from equation (7.10) when it is applied to a geodesic polar coordinate system. Here $g_{11} = 1$. Then

(7.12)
$$\mathbf{K} = \frac{\mathbf{R}_{1212}}{g} = -\frac{1}{2\sqrt{g_{22}}} \frac{\partial}{\partial x^1}\left(\frac{1}{\sqrt{g_{22}}} \frac{\partial g_{22}}{\partial x^1} \right) = -\frac{1}{\sqrt{g_{22}}} \frac{\partial^2 \sqrt{g_{22}}}{(\partial x^1)^2}.$$

In a geodesic polar coordinates system, since we have

$$g_{ij} = \begin{Vmatrix} 1 & 0 \\ 0 & g_{22} \end{Vmatrix},$$

the only Christoffel symbols of the second kind which do not vanish are

$$\begin{Bmatrix} 2 \\ 2\,2 \end{Bmatrix} = -\frac{1}{\sqrt{g_{22}}} \frac{\partial \sqrt{g_{22}}}{\partial x^2}, \quad \begin{Bmatrix} 1 \\ 2\,2 \end{Bmatrix} = -\frac{\partial g_{22}}{\partial x^1}, \text{ and}$$

$$\begin{Bmatrix} 2 \\ 2\,1 \end{Bmatrix} = \begin{Bmatrix} 2 \\ 1\,2 \end{Bmatrix} = \frac{1}{\sqrt{g_{22}}} \frac{\partial \sqrt{g_{22}}}{\partial x^1}.$$

In these coordinates, consider a multiple η of the basis vector $\eta^i = (0, \eta)$ (η a constant), called the **geodesic deviation**. Let us determine the intrinsic derivative of η^i with respect to $s_{(1)} = |x^1|$. Then

$$\frac{\delta\eta^i}{\delta s_{(1)}} = \frac{d\eta^i}{ds_{(1)}} + \begin{Bmatrix} i \\ jk \end{Bmatrix} \eta^j \frac{dx^k}{ds_{(1)}} = \begin{Bmatrix} i \\ 2\,1 \end{Bmatrix} \eta = \left(0, \frac{\eta}{\sqrt{g_{22}}} \frac{\partial \sqrt{g_{22}}}{\partial x^1} \right).$$

The intrinsic derivative of this, in turn, is

$$\frac{\delta}{\delta s_{(1)}}\left(\frac{\delta \eta^i}{\delta s_{(1)}}\right) = \frac{d}{ds}\left(\frac{\delta \eta^i}{\delta s_{(1)}}\right) + \left\{ {i \atop j\ k} \right\} \frac{\delta \eta^j}{\delta s_{(1)}} \frac{dx^k}{ds_{(1)}}$$

$$= \frac{\partial}{\partial x^1}\left(\frac{\delta \eta^i}{\delta s_{(1)}}\right) + \left\{ {i \atop 21} \right\} \frac{\delta \eta^2}{\delta s_{(1)}}.$$

From this we obtain

$$\frac{\delta^2 \eta^1}{\delta s_{(1)}^2} = 0,$$

$$\frac{\delta^2 \eta^2}{\delta s_{(1)}^2} = \eta \left\{ \frac{\partial}{\partial x^1}\left[\frac{1}{\sqrt{g_{22}}}\frac{\partial \sqrt{g_{22}}}{\partial x^1}\right] + \left(\frac{1}{\sqrt{g_{22}}}\frac{\partial g_{22}}{\partial x^1}\right)^2 \right\}$$

$$= \frac{\eta}{\sqrt{g_{22}}}\left\{ \frac{\partial^2 \sqrt{g_{22}}}{(\partial x^1)^2} - \frac{1}{2 g_{22}}\frac{\partial g_{22}}{\partial x^1} + \left[\frac{1}{\sqrt{g_{22}}}\frac{\partial \sqrt{g_{22}}}{\partial x^1}\right]^2 \right\} = -\mathrm{K}\eta$$

by equation (7.12). Therefore

$$\frac{\delta^2 \eta^i}{\delta s_{(1)}^2} + \mathrm{K}\eta^i = 0. \tag{7.13}$$

This is the **equation of geodesic deviation**.

Let us now define ζ^i to be the unit vector

$$\zeta^i = \frac{\eta^i}{|\eta^i|} = \left(0, \frac{1}{\sqrt{g_{22}}}\right)$$

in the direction of η^i. It is easy to show by direct calculation that

$$\frac{\delta \zeta^i}{\delta s_{(1)}} = 0, \quad \frac{\delta^2 \zeta^i}{\delta s_{(1)}^2} = 0.$$

Therefore, since $\eta^i = |\eta^j|\zeta^i$, and since

$$\frac{\delta \eta^i}{\delta s_{(1)}} = |\eta^j|\frac{\delta \zeta^i}{\delta s_{(1)}} + \zeta^i \frac{d|\eta^j|}{ds_{(1)}} = \zeta^i \frac{d|\eta^j|}{ds_{(1)}} \text{ and}$$

$$\frac{\delta^2 \eta^i}{\delta s_{(1)}^2} = \frac{\delta}{\delta s_{(1)}}\left(\frac{\delta \eta^i}{\delta s_{(1)}}\right) = \frac{d|\eta^i|}{ds_{(1)}}\frac{\delta \zeta^i}{\delta s_{(1)}} + \zeta^i \frac{d^2|\eta^j|}{ds_{(1)}^2},$$

equation (7.13) requires that

$$\zeta^i \left(\frac{d^2 |\eta^j|}{ds_{(1)}^2} + \mathrm{K} |\eta^j| \right) = 0 .$$

Since ζ^i is not identically zero, the quantity in parenthesis must vanish. In other words, since $|\eta^i| = \sqrt{g_{22}}$,

$$\frac{d^2 \sqrt{g_{22}}}{ds_{(1)}^2} + \mathrm{K} \sqrt{g_{22}} = 0 . \tag{7.14}$$

When K is constant, the solutions of equation (7.14) are

$$\sqrt{g_{22}} = \begin{cases} \mathrm{A} \sin(\sqrt{\mathrm{K}}\, s_{(1)}), & \mathrm{K} > 0 \\ \mathrm{A}\, s_{(1)}, & \mathrm{K} = 0 \\ \mathrm{A} \sinh(\sqrt{-\mathrm{K}}\, s_{(1)}), & \mathrm{K} < 0 \end{cases} \tag{7.15}$$

where A is at most a function of x^2. Therefore, recalling that $s_{(1)} = x^1$ along the x^1-curves,

$$g_{22} = \begin{cases} [\mathrm{A}(x^2)]^2 \sin^2(\sqrt{\mathrm{K}}\, x^1), & \mathrm{K} > 0, \\ [\mathrm{A}(x^2)]^2 (x^1)^2, & \mathrm{K} = 0, \\ [\mathrm{A}(x^2)]^2 \sinh^2(\sqrt{-\mathrm{K}}\, x^1), & \mathrm{K} < 0 . \end{cases}$$

Consider the case $\mathrm{K} > 0$, for example. Then

$$(ds)^2 = g_{ij}\, dx^i\, dx^j = (dx^1)^2 + g_{22} (dx^2)^2$$

$$= (dx^1)^2 + [\mathrm{A}(x^2)\, dx^2]^2 \sin^2(\sqrt{\mathrm{K}}\, x^1) .$$

We may absorb the factor $\mathrm{A}(x^2)$ into the differential by the transformation

$$\bar{x}^1 = x^1, \quad \bar{x}^2 = \int \mathrm{A}(x^2)\, dx^2 .$$

Then

$$(ds)^2 = \bar{g}_{ij}\, d\bar{x}^i\, d\bar{x}^j ,$$

where

$$\bar{g}_{ij} = \begin{Vmatrix} 1 & 0 \\ 0 & \sin^2\left(\sqrt{\mathrm{K}}\, \bar{x}^1\right) \end{Vmatrix} .$$

A corresponding procedure is clear for the case $\mathrm{K} < 0$. When $\mathrm{K} = 0$, these coordinates clearly reduce to plane polar coordinates.

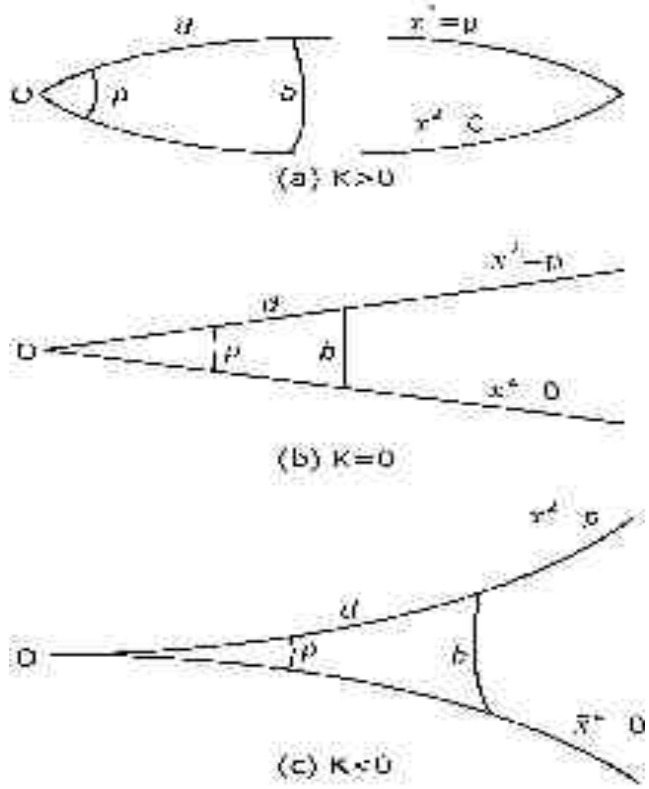

Figure 46

The meaning of geodesic deviation may now be made clear for a space of constant curvature. Consider a geodesic polar coordinate system with origin at $\mathbf{O}$. Draw a geodesic through $\mathbf{O}$ identified as $\bar{x}^2 = 0$ and an adjacent geodesic through $\mathbf{O}$ which is the curve $\bar{x}^2 = p << 1$. Then along the $\bar{x}^2$– curve through any point $(d, 0)$, the distance between the two geodesics is

$$b = \sqrt{g_{22}}\,p\,.$$

When $\mathbf{K} > 0$, $b = [\sin(\sqrt{\mathbf{K}}\,d)]\,p$, when $\mathbf{K} < 0$, $b = [\sinh(\sqrt{-\mathbf{K}}\,d)]\,p$, and when $\mathbf{K} = 0$, $b = d \times p$. These results are shown in Fig. 46.

Now consider a closed curve $\mathbf{C}$ in the surface S (Fig. 47). Let p^i be a unit parallel vector field about $\mathbf{C}$. Let $\mathbf{O}$ be the origin of a geodesic polar coordinate system in $\mathbf{S}$. If θ is the angle between p^i and the x^i– curve through any point of $\mathbf{C}$, then p^i must have the components

$$p^1 = \cos\theta\,,\ p^2 = \frac{\sin\theta}{\sqrt{g_{22}}}\,.$$

Now by hypothesis, p^i is a parallel field about $\mathbf{C}$. Therefore, using the values given in equation (7.9), with $g_{11} = 1$,

$$\frac{\delta p^1}{\delta s} = \frac{dp^1}{ds} + \left\{ {1 \atop mn} \right\} p^m \frac{dx^n}{ds} = -\sin\theta\,\frac{d\theta}{ds} - \frac{1}{2}\,\frac{\partial g_{22}}{\partial x^1}\,\frac{\sin\theta}{\sqrt{g_{22}}}\,\frac{dx^2}{ds} = 0\,.$$

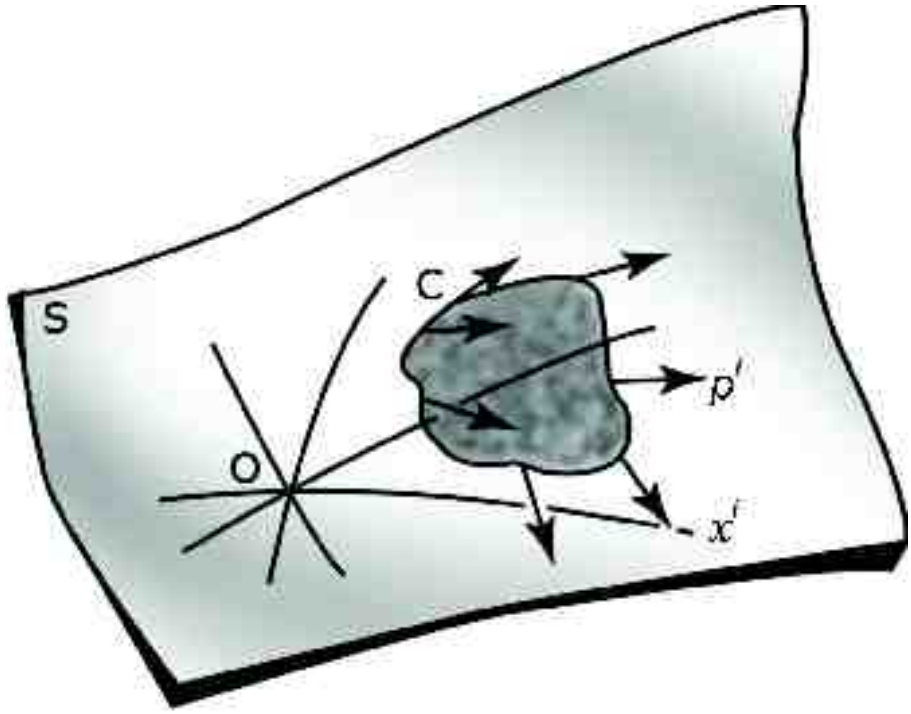

Figure 47

Hence

$$\frac{d\theta}{ds} = \frac{\partial\sqrt{g_{22}}}{\partial x^1}\frac{dx^2}{ds}.$$

Since both $\frac{d\theta}{ds}$ and ds are invariants, $d\theta = \frac{d\theta}{ds}\cdot ds$ may be integrated entirely around $\mathbf{C}$, giving

$$\mathrm{E} = \int_C d\theta = \int \frac{d\theta}{ds}\,ds = -\int \frac{\partial\sqrt{g_{22}}}{\partial x^1}\frac{dx^2}{ds}\,ds = \iint \frac{\partial^2\sqrt{g_{22}}}{(\partial x^1)^2}\,dx^1\,dx^2.$$

But since $\sqrt{g} = \sqrt{g_{22}}$ in a geodesic polar coordinate system, and since the element of area in $\mathbf{S}$ is $d\mathbf{S} = \varepsilon_{ij}\,dx^i_{(1)}\,dx^j_{(2)} = \sqrt{g}\,dx^1\,dx^2$, where $dx^i_{(1)}$ is a displacement along the x^1-curve and $dx^j_{(2)}$ is a displacement along the x^2-curve, we see that

$$\mathrm{E} = -\iint \frac{1}{\sqrt{g_{22}}}\frac{\partial^2\sqrt{g_{22}}}{(\partial x^1)^2}\,d\mathbf{S} = \iint_{\mathbf{S}} \mathrm{K}\,d\mathbf{S} \tag{7.16}$$

by equation (7.12).

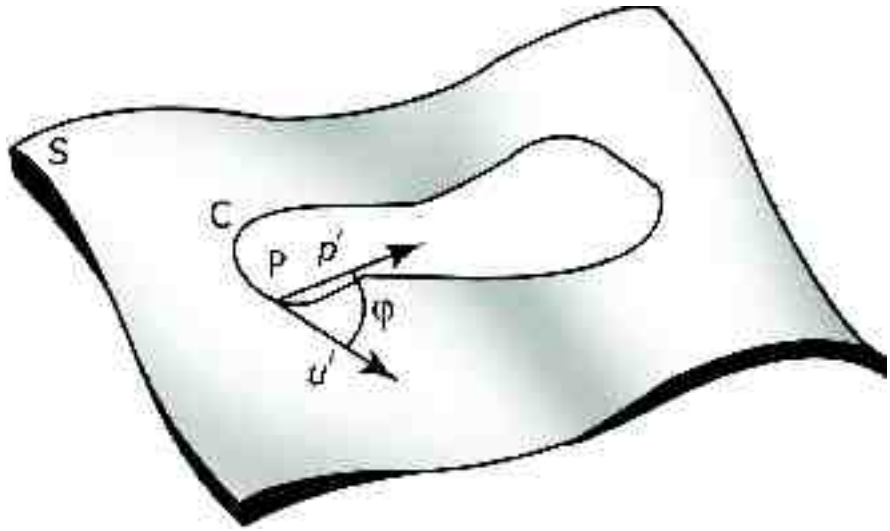

Figure 48

Additional insight into these results is to be had by an alternative computation of E. In this case, let u^i be the unit tangent vector to C (see Fig. 48). Then, as we have seen,

$$u^i = \frac{dx^i}{ds},\ u_i u^i = 1,\ \frac{\delta u_i}{\delta s} = \kappa \mu_i,\ \mu_i = \varepsilon_{ij} u^i,$$

where the invariant κ is the curvature of C. If, now, φ is the angle between u^i and p^i (rather than between the x^1-curve and p^i), we must have

$$\cos\varphi = u^i p_i,\ \sin\varphi = \varepsilon_{ji} p^i u^i.$$

Therefore

$$\frac{d\sin\varphi}{ds} = \cos\varphi \frac{d\varphi}{ds} = \frac{\delta}{\delta s}(\varepsilon_{ji} p^i u^j) = \varepsilon_{ji}\left[\frac{\delta p^i}{\delta s} u^j + p^i \frac{\delta u^j}{\delta s}\right] = \varepsilon_{ji} \frac{\delta u^j}{\delta s} p^i$$

since p^i is a parallel vector field. Hence

$$\frac{d\varphi}{ds} = \frac{\varepsilon_{ji} \frac{\delta u^j}{\delta s} p^i}{u^k p_k} = \frac{\kappa \varepsilon_{ji} \mu^j p^i}{u^k p_k} = -\frac{\kappa p^i u_i}{u^k p_k} = -\kappa.$$

Now, **if we integrate entirely around C**, the integral

$$\int_C d\varphi = \int \frac{d\varphi}{ds} ds$$

is evidently the same as the integral $\int_C d\theta$ *except* that the reference direction upon completion of the circuit has rotated by 2π. Hence

$$(7.17)\quad \int \frac{d\varphi}{ds}\,ds = -\int_C \kappa\,ds = E - 2\pi \quad \text{or} \quad \iint_S K\,dS + \int_C \kappa\,ds = 2\pi .$$

This result is known as **Bonnet's theorem**.

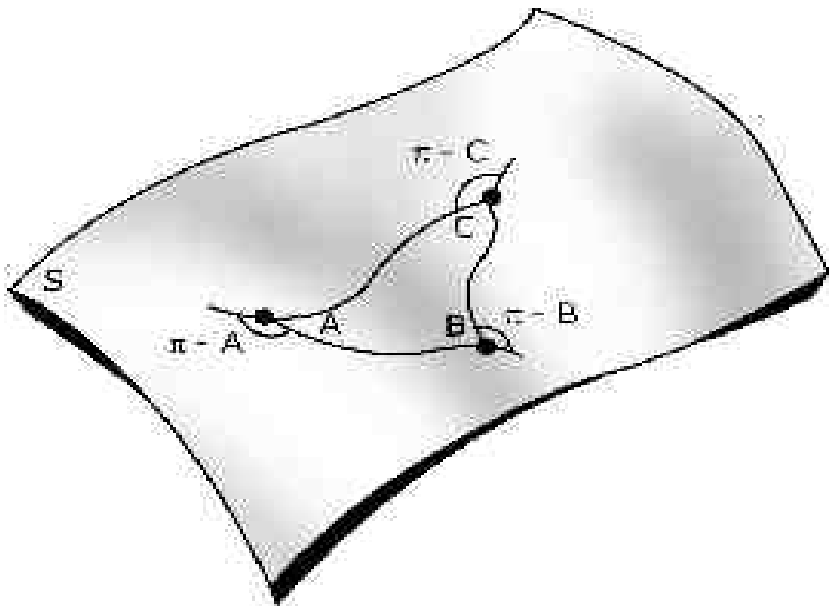

Figure 49

Bonnet's theorem has an interesting application to geodesic triangles. In a surface S let A, B, and C be three points not on a common geodesic (see Fig. 49). When each pair is connected by the geodesic between them, the resultant figure is a **geodesic triangle**. Consider the integral $\int_C d\varphi$ about the perimeter C (not to be confused with the vertex C). Along each side of the triangle $\kappa = 0$ because it is a geodesic arc. There is therefore no contribution to the integral along the sides. At the vertices, however, the direction of u^i changes by the complement of the vertex angle. Hence the integral is discontinuously diminished by this amount. Therefore

$$\iint_S K\,dS - 2\pi = -[(\pi - A) + (\pi - B) + (\pi - C)] \text{ or}$$

$$(7.18)\qquad E = A + B + C - \pi = \iint_S K\,dS .$$

This result is **Gauss's theorem**. From Gauss's theorem it is clear that for a geodesic triangle on a surface of positive curvature $E > 0$, $A + B + C > \pi$. The simplest such surface is a sphere, whence the quantity is called the **spherical excess**. On a surface of negative curvature, on the other hand, $E < 0$ and $A + B + C < \pi$.

It is clear that by combining the arguments which led to Bonnet's theorem and Gauss's theorem, we may consider on S a closed curve whose arcs need not be geodesic and which has vertex angles $A_{(i)}$. For such a curve the result

$$\int\int_S K\,dS + \int_C \kappa\,ds + \sum_i \left(\pi - A_{(i)}\right) = 2\pi \tag{7.19}$$

must evidently be true. Bonnet's theorem and Gauss's theorem are evidently special cases of this result.

Ex. (7.6) Show that plane geometry holds only when the Riemann-Christoffel tensor vanishes. (Hint: From equation (7.18) and the fact that E = 0 *in a plane triangle, we see that* K = 0 *in a plane. This, with equation (7.11), implies* $R_{1212} = 0$.

Ex. (7.7) On the unit sphere, what is the area of the spherical triangle whose angles are $A = 113°51'22''$, $B = 66°17'20''$, $C = 96°0'18''$?

Ans. Since $K = 1$, area = $S = E$. *Since* $E = 96.15° = 0.53417\,\pi$ *radians, the area of the spherical triangle is* $0.53417\,\pi = 1.6783$ *square units.*

Ex. (7.8) Show that in a space for which the metric is

$$g_{ij} = \frac{\delta_{ij}}{[1 + c\,\delta_{mn} x^m x^n]^2} \tag{7.20}$$

where c is a constant, the curvature is K = 4c, *hence a constant. (Hint: use equations (7.7) and (7.10).)*

(We note without proof that the converse is also true. That is, in a space of constant curvature, there exists a coordinate system x^i *such that the fundamental tensor has the form given in equation (7.20).)*

Ex. (7.9) Consider three points on a surface but not on a common geodesic. Let geodesics connect them by pairs. The resultant figure is a ***geodesic triangle****. Derive the fundamental trigonometric relations in flat space by (a) choosing one vertex of a geodesic triangle as the origin of a geodesic polar coordinate system, and (b) solving the geodesic equations for the equation of the side opposite the origin. Thus show that the fundamental tensor does indeed determine the geometry of a surface of zero curvature by determining its trigonometry.*

Ans. Let the plane geodesic triangle be ABC. *Choose* C *to be the pole of a geodesic coordinate system* $x^i = (r, \theta)$. *Then since*

$$g_{ij}\frac{dx^i}{ds}\frac{dx^j}{ds} = \left(\frac{dr}{ds}\right)^2 + r^2\left(\frac{d\theta}{ds}\right)^2 = 1, \tag{7.21}$$

the geodesic equations are

$$\begin{cases} \ddot{r} - r\dot{\theta}^2 = 0, \\ r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0 \end{cases} \tag{7.22}$$

where dots denote differentiation with respect to s. *We multiply the second of equations (7.22) by* r *and obtain at once the integral*

$$r^2\dot{\theta} = \mathrm{H} = \text{constant}. \tag{7.23}$$

We may interpret H *by combining equations (7.21) and (7.23); they give*

$$\dot{r}^2 + \frac{\mathrm{H}^2}{r^2} = 1. \tag{7.24}$$

Evidently $\mathrm{r} = \mathrm{H}$ *where* $\dot{r} = 0$. *Hence* H *is the length of the geodesic to the point on* AB *(possibly extended) which is nearest* C. *Let this point be a distance* s_0 *from* A *along the geodesic* AB; *it is a point whose geodesic polar coordinates are* (H, θ_0).

We can readily see that $\mathrm{C}s_0$ *is perpendicular to* AB, *for the unit tangent along* $\mathrm{C}s_0$ *at* s_0 *is* $\lambda^i_{(1)} = (1, 0)$ *whereas the unit tangent along* AB *at* s_0 *is* $\lambda^i_{(2)} = (0, \frac{1}{\mathrm{H}})$. *Their inner product clearly vanishes and the curves are thus orthogonal at* s_0.

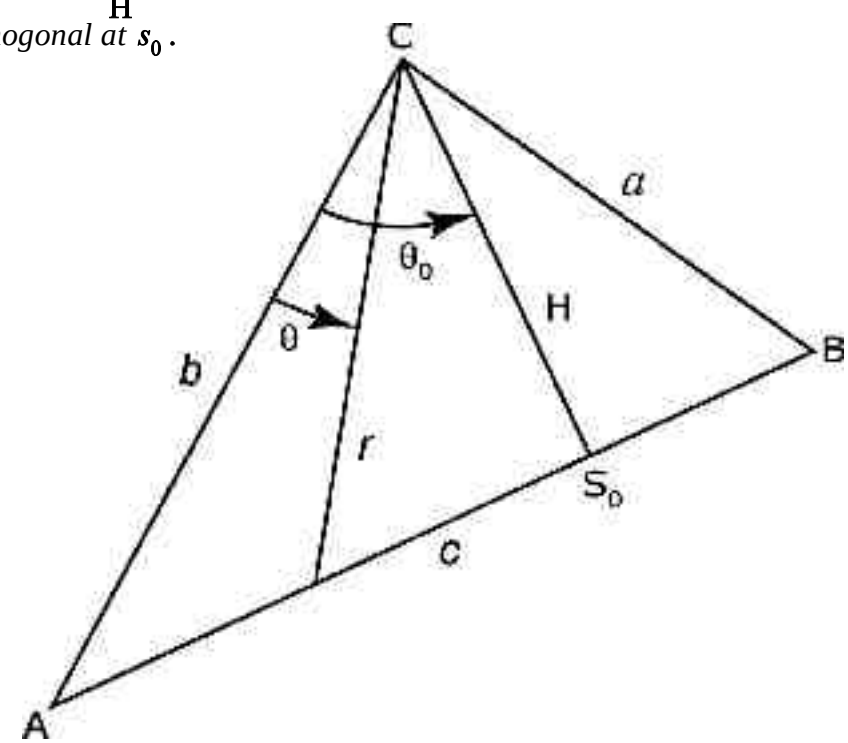

Figure 50

Equations (7.24) and (7.21) provide two of the four integrals needed for a complete solution. It is most interesting that they also suffice for a derivation of the law of sines. To derive this law, consider unit tangents $\lambda^i_{(1)}$ *and* $\lambda^i_{(2)}$ *at* A, *directed toward vertices* B *and* C, *respectively. It is clear that* $\lambda^i_{(2)} = (-1, 0)$ *and from equations (7.24) and (7.21) we find that*

$$\lambda^i_{(1)} = \left(-\left[1 - \frac{H^2}{b^2}\right]^{1/2}, \frac{H}{b^2} \right).$$

Hence

$$\cos A = g_{ij}\lambda^i_{(1)}\lambda^j_{(2)} = \left[1 - \frac{H^2}{b^2}\right]^{1/2},$$

(7.25) $$\sqrt{b^2 - H^2} = b\cos A, \quad H = b\sin A.$$

Similarly, at vertex B,

(7.26) $$\cos B = \left[1 - \frac{H^2}{a^2}\right]^{1/2}, \quad H = a\sin B.$$

Equating the common values of H, *we get*

(7.27) $$H = b\sin A = a\sin B, \quad \frac{\sin A}{a} = \frac{\sin B}{b},$$

which is the ***law of sines*** *of plane trigonometry.*

Let us now seek the remaining integrals. From equation (7.24) we get

whence $$ds = \pm\frac{r\,dr}{\sqrt{r^2 - H^2}}$$

(7.28) $$(s - s_0)^2 + H^2 = r^2,$$

where we have imposed the condition that $s = s_0$ *when* $r = H$. *We now combine equations (7.23) and (7.28) to obtain*

(7.29) $$d\theta = \frac{H\,ds}{H^2 + (s - s_0)^2}$$

$$\theta - \theta_0 = \tan^{-1}\left(\frac{s - s_0}{H}\right), \quad s - s_0 = H\tan(\theta - \theta_0),$$

$$r\cos(\theta - \theta_0) = H,$$

which we recognize as the equation of the straight line AB.

Let us now evaluate equation (7.28) at the point B, *where* $s = c$ *and* $r = a$.

Then

(7.30) $$a^2 = \mathrm{H}^2 + (c - s_0)^2 = (\mathrm{H}^2 + s_0^2) + c^2 - 2\,c\,s_0 .$$

However, equation (7.28) evaluated at A, *where* $s = 0$ *and* $r = b$, *shows that*

$$\mathrm{H}^2 + s_0^2 = b^2$$

and from this and equation (7.25) it follows that $s_0 = b \cos \mathrm{A}$. *Making these substitutions into equation (7.30) gives us*

(7.31) $$a^2 = b^2 + c^2 - 2\,b\,c \cos \mathrm{A},$$

which is the ***law of cosines*** *of plane trigonometry.*

Ex. (7.10) By a method analogous to that of the preceding exercise, derive the trigonometric relations upon a surface of positive constant curvature.

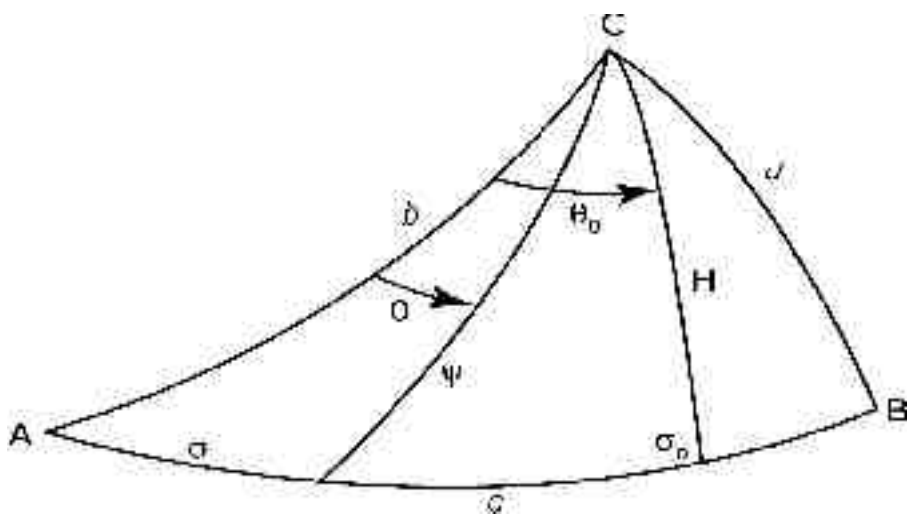

Figure 51

Ans. Let ABC *be the geodesic triangle. Choose* C *to be the origin of a geodesic polar coordinate system* $x^i = (\psi, \theta)$, *To normalize the independent variable, we set*

$$\sigma = s\sqrt{\mathrm{K}} = \frac{s}{\mathrm{R}},$$

where R *is the radius of the sphere. Then*

$$g_{ij}\frac{dx^i}{d\sigma}\frac{dx^j}{d\sigma} = \left(\frac{d\psi}{d\sigma}\right)^2 + (\sin^2\psi)\left(\frac{d\theta}{d\sigma}\right)^2 = 1\,. \tag{7.32}$$

The geodesic equations are then

$$\begin{cases} \dfrac{d^2\psi}{d\sigma^2} - (\sin\psi\cos\psi)\left(\dfrac{d\theta}{d\sigma}\right)^2 = 0\,, \\[2ex] \dfrac{d}{d\sigma}\left[(\sin^2\psi)\dfrac{d\theta}{d\sigma}\right] = 0\,. \end{cases} \tag{7.33}$$

From the second of these we obtain at once the integral

$$(\sin^2\psi)\,\dot{\theta} = \sqrt{1-\gamma^2} = \textit{constant}\,. \tag{7.34}$$

We can identify the constant of integration γ *as* $\cos\mathrm{H}$, *where* H *is the value of* ψ *at the point* σ_0, *the nearest point to* C *on* AB. *Again, the angle at* σ_0 *is a right angle.*

By taking unit tangent vectors at A *and* B *along the sides of the triangle, it is easy to show that*

$$\sin\mathrm{A}\sin b = \sin\mathrm{H} = \sin a\sin\mathrm{B}\,,$$

whence

$$\frac{\sin\mathrm{A}}{\sin a} = \frac{\sin\mathrm{B}}{\sin b}\,, \tag{7.35}$$

the ***law of sines*** *of spherical trigonometry. Again, as in plane trigonometry, the law of sines follows from two integrals (7.32) and (7.34) without the complete solution.*

Combining equations (7.33) and (7.34), we obtain

$$(\sin\psi)\frac{d\psi}{d\sigma} = \pm\sqrt{\gamma^2 - \cos^2\psi}$$

whose solution is*

$$\cos\psi = \gamma\cos(\sigma-\sigma_0) = \cos\mathrm{H}\cos(\sigma-\sigma_0)\,. \tag{7.36}$$

* Note that the minus sign above holds between A and σ_0, the plus sign from σ_0 to B.

We have here satisfied the condition that $\psi = \mathrm{H}$ *when* $\sigma = \sigma_0$. *Next, let us combine equations (7.34) and (7.36) to obtain*

$$\frac{d\theta}{d\sigma} = \frac{\sqrt{1-\gamma^2}}{\sin^2\psi} = \frac{\sqrt{1-\gamma^2}}{1-\cos^2(\sigma-\sigma_0)}$$

$$= \sqrt{1-\gamma^2}\times\left[\frac{\sin(\sigma-\sigma_0)}{\sqrt{1-\cos^2(\sigma-\sigma_0)}}\right]\times\frac{\sqrt{1-\gamma^2\cos^2(\sigma-\sigma_0)}}{[1-\gamma^2\cos^2(\sigma-\sigma_0)]^{3/2}}$$

$$= \frac{\sqrt{1-\gamma^2}\sin(\sigma-\sigma_0)}{\sqrt{[1-\gamma^2\cos^2(\sigma-\sigma_0)]-(1-\gamma^2)\cos^2(\sigma-\sigma_0)}}\times\sqrt{1-\gamma^2\cos^2(\sigma-\sigma_0)}$$

$$\times\frac{[1-\gamma^2\cos^2(\sigma-\sigma_0)]+\gamma^2\cos^2(\sigma-\sigma_0)}{[1-\gamma^2\cos^2(\sigma-\sigma_0)]^{3/2}}$$

$$= \sqrt{1-\gamma^2}\,\frac{\left\{\dfrac{\sin(\sigma-\sigma_0)}{\sqrt{1-\gamma^2\cos^2(\sigma-\sigma_0)}}+\dfrac{\cos(\sigma-\sigma_0)\times\gamma^2\sin(\sigma-\sigma_0)\cos(\sigma-\sigma_0)}{[1-\gamma^2\cos^2(\sigma-\sigma_0)]^{3/2}}\right\}}{\left[1-\dfrac{(1-\gamma^2)\cos^2(\sigma-\sigma_0)}{1-\gamma^2\cos^2(\sigma-\sigma_0)}\right]^{1/2}}$$

Therefore

$$\text{(7.37)}\qquad \begin{cases} d\theta = d\cos^{-1}\dfrac{\sqrt{1-\gamma^2\cos(\sigma-\sigma_0)}}{\sqrt{1-\gamma^2\cos^2(\sigma-\sigma_0)}} \\[2ex] \cos(\theta-\theta_0) = \dfrac{\sqrt{1-\gamma^2\cos(\sigma-\sigma_0)}}{\sqrt{1-\gamma^2\cos^2(\sigma-\sigma_0)}} \end{cases}.$$

If we combine equations (7.36) and (7.37), we obtain the alternate equations

$$\text{(7.38)}\qquad \begin{cases} \sin\psi\cos(\theta-\theta_0) = \sin\mathrm{H}\cos(\sigma-\sigma_0), \\ \cos(\theta-\theta_0) = \tan\mathrm{H}\cot\psi. \end{cases}$$

From equation (7.36) evaluated at point B, *we have*

$$\text{(7.39)}\qquad \cos a = \gamma\cos(c-\sigma_0) = (\gamma\cos\sigma_0)\cos c + (\gamma\sin\sigma_0)\sin c.$$

But at point A *the equation (7.36) gives*

$$\text{(7.40)}\qquad \gamma\cos\sigma_0 = \cos b.$$

Therefore

$$\gamma \sin \sigma_0 = \sqrt{\gamma^2 - \cos^2 b} = \sqrt{(1 - \cos^2 b) - (1 - \gamma^2)}$$

$$= \sin b \left[1 - \frac{\sin^2 H}{\sin^2 b} \right]^{1/2} .$$

However, in the right triangle $C\,\sigma_0\,A$, $\dfrac{\sin H}{\sin b} = \sin A$ *by the law of sines.*

Hence

$$\gamma \sin \sigma_0 = \sin b \cos A .$$

Substituting this result and that of equation (7.40) into equation (7.39) gives

(7.41) $$\cos a = \cos b \cos c + \sin b \sin c \cos A$$

which is the ***law of cosines*** *of spherical trigonometry.*

Ex. (7.11) By a method analogous to that of the preceding exercise, derive the trigonometric relations upon a surface of constant negative curvature. (Hint: set $s\sqrt{-K} = \sigma$*).*

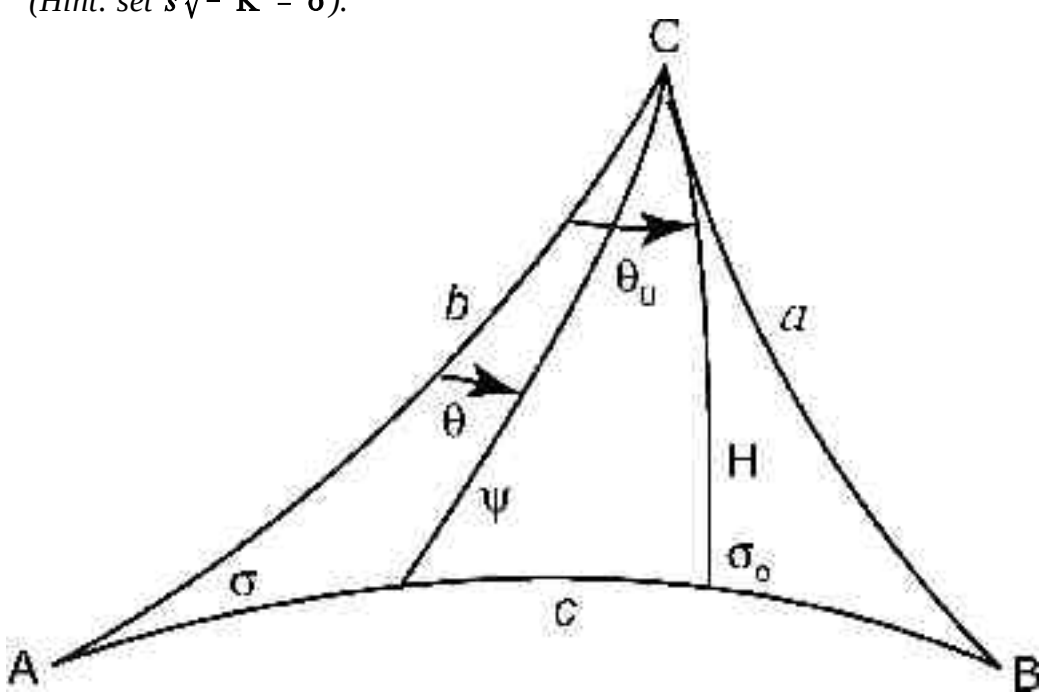

Figure 52

Ans. The integrals appropriate to Fig. 52 are:

$$(a) \qquad \left(\frac{d\psi}{d\sigma} \right)^2 + (\sinh^2 \psi) \left(\frac{d\theta}{d\sigma} \right)^2 = 1 ,$$

$$(b) \qquad \frac{d\theta}{d\sigma} = \frac{\sqrt{\gamma^2 - 1}}{\sinh^2 \psi}, \quad \gamma = \cosh \mathrm{H},$$

$$(c) \qquad \cosh \psi = \gamma \cosh(\sigma - \sigma_0),$$

$$(d) \qquad \cos(\theta - \theta_0) = \frac{\sqrt{\gamma^2 - 1}\cosh(\sigma - \sigma_0)}{\sqrt{\gamma^2 \cosh^2(\sigma - \sigma_0) - 1}}.$$

The ***law of sines*** *is*

$$(7.43) \qquad \frac{\sin \mathrm{A}}{\sinh a} = \frac{\sin \mathrm{B}}{\sinh b} = \frac{\sin \mathrm{C}}{\sinh c}.$$

The ***law of cosines*** *is*

$$(7.44) \qquad \cosh a = \cosh b \cosh c - \sinh b \sinh c \cos \mathrm{A}.$$

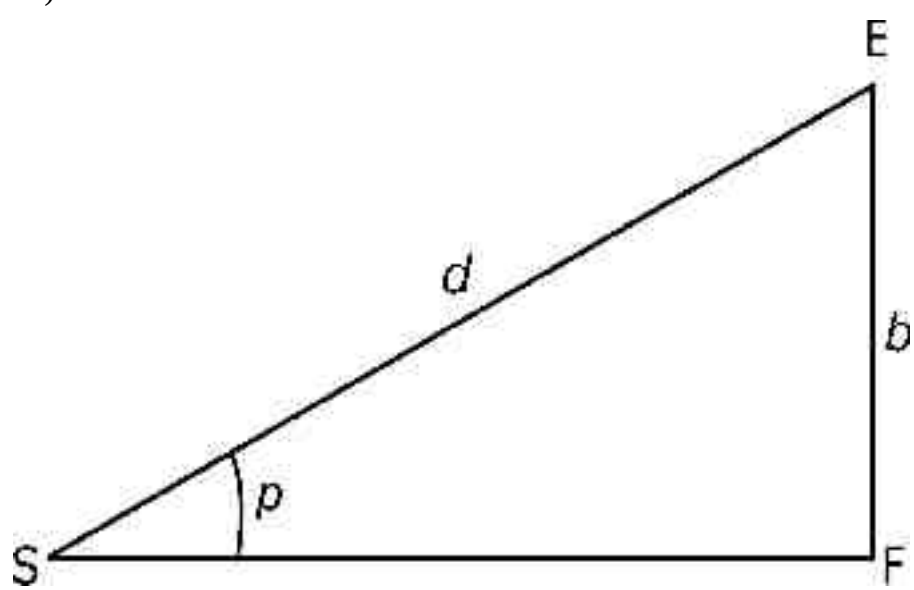

Figure 53

Ex. (7.12) In the geodesic triangle of Fig. 53, the angle at F *is a right angle. The side* ***d*** *is the distance of a star* S *from the earth* E*. The star is said to have a* ***parallax p*** *when side* ***b*** *is used as a* ***base line****. Using the results of the three preceding exercises, determine the parallax as a function of distance and base line for a space of (a) zero curvature, (b) constant positive curvature and (c) constant negative curvature.*

Ans. $(a) \quad \sin p = \left(\frac{b}{d}\right), \quad \mathrm{K} = 0;$

$(b) \ \sin p = \left(\frac{\sin b}{\sin d}\right), \quad \mathrm{K} > 0, \quad (c) \ \sin p = \left(\frac{\sinh b}{\sinh d}\right), \quad \mathrm{K} < 0.$

Ex. (7.13) Find the area of a geodesic triangle in a space of (a) zero curvature, (b) constant positive curvature, and (c) constant negative curvature. By the results of parts (b) and (c), verify Gauss's theorem. (Hint: the element of area on any surface is $dA = \sqrt{g}\, dx^1\, dx^2$ *).*

Ans. (a) Using the labelling of Fig. 50, the area of the plane triangle is

$$Area = \int \left[\int_0^{r(\theta)} r\, dr \right] d\theta = \frac{1}{2} \int_0^C \left[r^2 \right]_0^{H \sec(\theta - \theta_0)} d\theta$$

$$= \frac{1}{2} H^2 \int_0^C \sec^2(\theta - \theta_0)\, d\theta = \frac{1}{2} H^2 \tan(\theta - \theta_0) \Big|_0^C$$

$$= \frac{1}{2} H^2 \left[\tan(C - \theta) + \tan\theta_0 \right] = \frac{1}{2} H \left[(c - s_0) + s_0 \right]$$

$$= \frac{1}{2} H c = \frac{1}{2} (a \sin B)\, c = \frac{1}{2} a c \sin B .$$

(b) Using the labelling of Fig. 51, the area of the spherical triangle is

$$Area = \int \left[\int_0^{\psi(\theta)} (\sin\psi)\, d\psi \right] d\theta = \int_0^C \left[1 - \cos\psi(\theta) \right] d\theta = C - \int_0^C \left[\cos\psi(\theta) \right] d\theta .$$

Now

$$\int_0^C \left[\cos\psi(\theta) \right] d\theta = \gamma \int_0^C \cos(\sigma - \sigma_0) \times \frac{\sqrt{1 - \gamma^2}}{\sin^2\psi}\, d\sigma$$

$$= \frac{\gamma}{\sqrt{1 - \gamma^2}} \int \frac{\cos(\sigma - \sigma_0)\, d\sigma}{1 + \left[\dfrac{\gamma}{\sqrt{1 - \gamma^2}} \sin(\sigma - \sigma_0) \right]^2} = \tan^{-1} \left[\frac{\gamma}{\sqrt{1 - \gamma^2}} \sin(\sigma - \sigma_0) \right]_0^C$$

$$= \tan^{-1} \left[\frac{\gamma}{\sqrt{1 - \gamma^2}} \sin(c - \sigma_0) \right] + \tan^{-1} \left[\frac{\gamma}{\sqrt{1 - \gamma^2}} \sin\sigma_0 \right].$$

However,

$$\tan^{-1} \left[\frac{\gamma}{\sqrt{1 - \gamma^2}} \sin(c - \sigma_0) \right] = \sin^{-1} \left\{ \frac{\dfrac{\gamma \sin(c - \sigma_0)}{\sqrt{1 - \gamma^2}}}{\left[1 + \dfrac{\gamma^2}{1 - \gamma^2} \sin^2(c - \sigma_0) \right]^{1/2}} \right\}$$

$$= \sin^{-1} \left[\frac{\gamma \sin(c - \sigma_0)}{\sqrt{1 - \gamma^2 \cos^2(c - \sigma_0)}} \right] = \sin^{-1} \left[\frac{\cos H \sin(c - \sigma_0)}{\sin a} \right].$$

But in the triangle $C\sigma_0 B$ *we have*

$$\cos H = \cos(c - \sigma_0)\cos a + \sin(c - \sigma_0)\sin a \cos B$$

$$= \cos(c - \sigma_0)[\cos H \cos(c - \sigma_0)] + \sin(c - \sigma_0)\sin a \cos B$$

$$\cos H[1 - \cos^2(c - \sigma_0)] = \sin(c - \sigma_0)\sin a \cos B,$$

$$\frac{\cos H \sin(c - \sigma_0)}{\sin a} = \cos B.$$

Therefore

$$\tan^{-1}\left[\frac{\gamma}{\sqrt{1 - \gamma^2}}\sin(c - \sigma_0)\right] = \sin^{-1}[\cos B] = \frac{\pi}{2} - B.$$

Similarly,

$$\tan^{-1}\left[\frac{\gamma}{\sqrt{1 - \gamma^2}}\sin\sigma_0\right] = \frac{\pi}{2} - A.$$

Therefore the area of the spherical triangle is

$$\textit{Area} = C - [(\frac{\pi}{2} - B) + (\frac{\pi}{2} - A)] = A + B + C - \pi.$$

(c) $\textit{Area} = \pi - (A + B + C)$.

Ex. (7.14) On the surface of a unit sphere, let a unit vector p^i *be given at some point* A. *Find the form which* p^i *must have if it is to be a parallel vector field along the geodesic connecting* A *with* B.

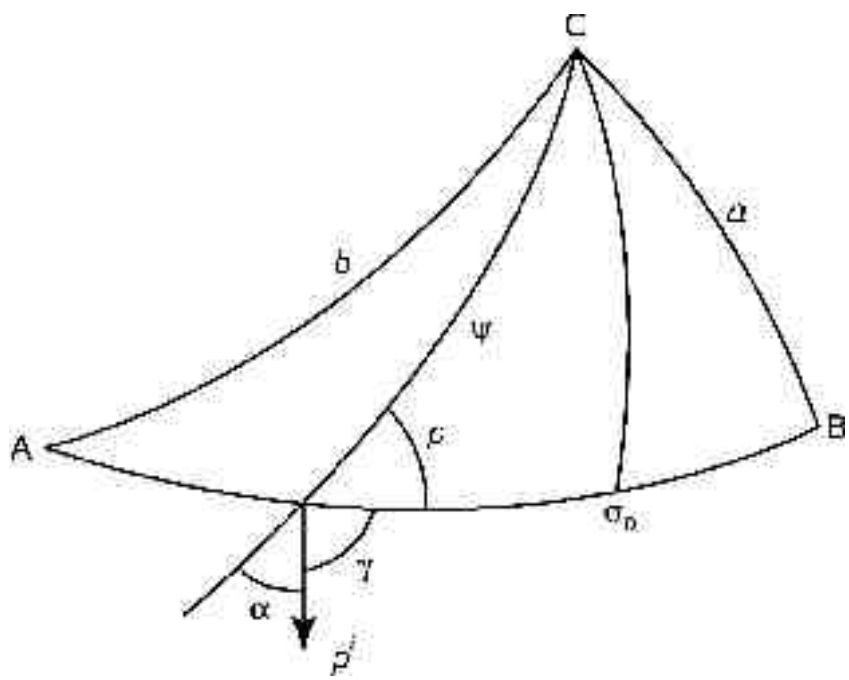

Figure 54

Ans. Let C *be the origin of a geodesic polar coordinate system with "radii"* $CA = b$ *and* $CB = a$ *to* A *and* B*, respectively (see Fig. 54). Let* AB *be the geodesic between* A *and* B *(see Ex. (7.10)). Then without loss of generality, one may take* p^i *to be of the form* $p^i = (\cos\alpha, \frac{\sin\alpha}{\sin\psi})$*, where* α *is the angle between the* ψ*-curve (*x^1*-curve) and* p^i*. In order that* p^i *be a parallel vector field along* AB*, we must have* $\frac{\delta p^i}{\delta s} = 0$ *or*

$$\begin{cases} \dfrac{dp^1}{d\sigma} - (\sin\psi\cos\psi)p^2\dfrac{d\theta}{d\sigma} = 0, \\ \dfrac{dp^2}{d\sigma} + \dfrac{\cos\psi}{\sin\psi}\left[p^1\dfrac{d\theta}{d\sigma} + p^2\dfrac{d\psi}{d\sigma}\right] = 0. \end{cases}$$

Since p^1 *and* p^2 *are expressible parametrically in terms of* α *it suffices to solve just the first of these equations. Substituting the parametric equations for* p^1 *and* p^2*, we get*

$$-(\sin\alpha)\frac{d\alpha}{d\sigma} = (\sin\psi\cos\psi)\frac{\sin\alpha}{\sin\psi}\frac{d\theta}{d\sigma}.$$

But since $\frac{d\theta}{d\sigma} = \sin H / \sin^2\psi$*, this becomes*

$$\frac{d\alpha}{d\sigma} = -\frac{\sin H\cos\psi}{\sin^2\psi}.$$

Again, since

$$d\sigma = -\frac{\sin\psi\, d\psi}{\sqrt{\cos^2 H - \cos^2\psi}} = -\frac{\sin\psi\, d\psi}{\sqrt{\sin^2\psi - \sin^2 H}}$$

we have that

$$d\alpha = \frac{\left[\dfrac{\sin H\cos\psi}{\sin^2\psi}\right]d\psi}{\sqrt{1 - \left(\dfrac{\sin H}{\sin\psi}\right)^2}} = -d\sin^{-1}\left[\frac{\sin H}{\sin\psi}\right].$$

Hence

$$\alpha - \alpha_0 = -\sin^{-1}\left[\frac{\sin H}{\sin\psi}\right] + \sin^{-1}\left[\frac{\sin H}{\sin b}\right] = -P + A.$$

Evidently $\alpha + P = \alpha_0 + A = \pi - \gamma =$ ***constant****. In other words, the parallel field*

$$p^i = \left(-\cos(P + \gamma), \frac{\sin(P + \gamma)}{\sin\psi}\right)$$

is a unit vector which makes a constant angle γ *with the geodesic between* A *and* B.

Ex. (7.15) Suppose that the vector p^i of Ex. (7.14) has components at A *which are* (1, 0). *Show that by parallel transport of p^i to* B, *the components become*

$$v^i = \left(\cos[A - B], \frac{\sin[A - B]}{\sin a} \right).$$

Ans. Since at A, $\pi - (A + \gamma) = 0$, we have from Ex. (7.14) that $\pi - (P + \gamma) = A - P$, whence the result follows when evaluated at point B.

Ex. (7.16) Show that if the equation for p^2 in Ex. (7.14) had been used, it would have led to the same results as the equation for p^1.

Ex. (7.17) Derive the form of a parallel unit vector field along a geodesic in a space of constant negative curvature.

Ans. Let the unit parallel vector field have the form

$$p^i = \left(\cos\alpha, \frac{\sin\alpha}{\sin\psi} \right),$$

in terms of a parameter α. *Then the first of the equations of parallelism is*

$$\frac{dp^1}{d\sigma} = (\sinh\psi \cosh\psi)\, p^2 \frac{d\theta}{d\sigma}.$$

Substituting for p^1 and p^2 leads to

$$\frac{d\alpha}{d\sigma} = \frac{\sinh H \cosh\psi}{\sinh^2\psi}.$$

Replacing $d\sigma$ by

$$d\sigma = -\frac{\dfrac{\sinh\psi\, d\psi}{\cosh H}}{\sqrt{\left(\dfrac{\cosh\psi}{\cosh H}\right)^2 - 1}}$$

(see Ex. (7.11)), we get

$$d\alpha = -d\sin^{-1}\left[\frac{\sinh H}{\sinh\psi}\right], \quad \alpha - \alpha_0 = \sin^{-1}\left[\frac{\sinh H}{\sinh a}\right] - \sin^{-1}\left[\frac{\sinh H}{\sinh\psi}\right] = A - P.$$

Therefore $$p^i = \left(-\cos[P + \gamma], \frac{\sin[P + \gamma]}{\sinh\psi} \right),$$

where $$\alpha + P = \alpha_0 + A = \pi - \gamma = \textit{constant}.$$

Ex. (7.19) Show that the second condition of parallelism leads to the same differential equation as the first.

Ex. (7.20) Show that quite generally a parallel vector field along a geodesic makes a constant angle with the geodesic.

Ex. (7.21) The angle $\alpha - \alpha_0$ *may be called the* ***rotation*** *of the parallel vector* p^i *in Ex. (7.14) and (7.15). Show that the rotation during parallel transport from* A *to* B *increases from zero to* $E - C$ *(Ex. (7.14)) or* $C - E$. *(Ex. (7.17)). Using the results of Ex. (7.13), show that the rotation of* p^i *is equal to*

$$\Delta\alpha = \mp\left(\frac{area}{R^2} - C\right) = K \times (area) + C,$$

where the area is that of the geodesic triangle and K *is the curvature of the space,* R *the radius of curvature. The upper sign is associated with spaces of positive curvature, the lower with spaces of negative curvature. Hence show that for geodesic triangles of small area, the rotation approaches the angle* C.

8. Mappings

Consider two surfaces S and $\bar{S}$. Let coordinates be defined on each — x^i on S and $\bar{x}^i$ on $\bar{S}$. Suppose that for some regions or perhaps all of S and $\bar{S}$ there is a reversible transformation of coordinates

(8.1) $$x^i = x^i(\bar{x}^j),\ \bar{x}^j = \bar{x}^j(x^i).$$

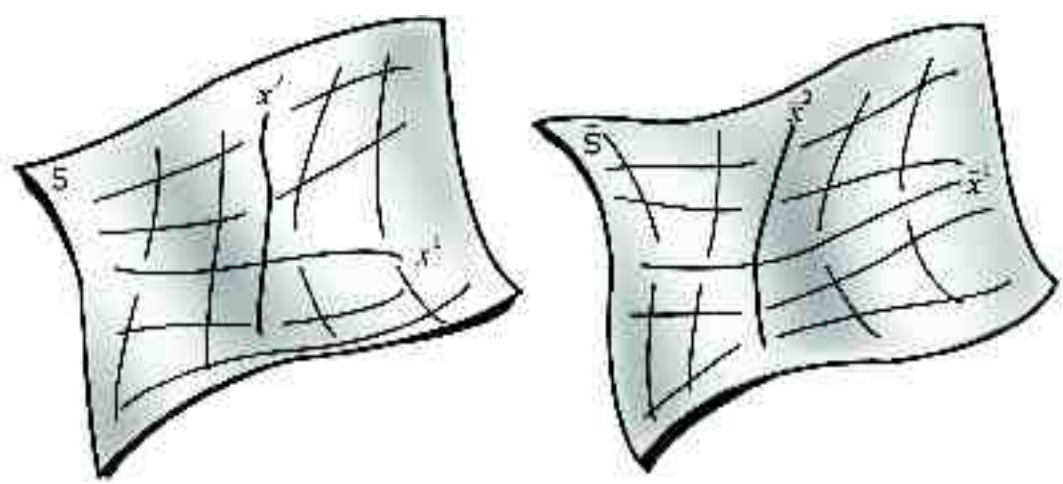

Figure 55

This transformation defines a **mapping** of S upon $\bar{S}$ and conversely. A specific example of mapping is, of course, the transfer of features of the surface of the earth to a sheet of paper, a "map" of the earth. It is of much interest to consider various mappings and to note the conditions under which certain properties of figures in S may be preserved in $\bar{S}$.

It is assumed that in each surface there is defined a fundamental tensor and, thereby, a congruence relation which determines infinitesimal arc lengths from infinitesimal coordinate differences. It should be clearly understood that the mapping exists regardless of what the congruence relations may be. The congruence relations, however, will determine what properties, if any, may be common to $\mathbf{S}$ and $\bar{\mathbf{S}}$.

We therefore have *two* relations

$$(8.2) \qquad (ds)^2 = g_{ij}\,dx^i\,dx^j \text{ and } (d\bar{s})^2 = \bar{g}_{ij}\,d\bar{x}^i\,d\bar{x}^j$$

in addition to equation (8.1). We regard them, however, as possibly independent. Therefore, although

$$(8.3) \qquad d\bar{x}^i = \frac{\partial \bar{x}^i}{\partial x^k}\,dx^k,$$

we cannot assume that

$$\bar{g}_{ij} = \frac{\partial x^k}{\partial x^l}\frac{\partial x^l}{\partial \bar{x}^j}\,g_{kl},$$

since we are dealing with different surfaces $\mathbf{S}$ and $\bar{\mathbf{S}}$, not merely with different coordinate systems on the same surface. To emphasize this point, we may map one surface upon the other, as $\bar{\mathbf{S}}$ onto $\mathbf{S}$. Then we take equations (8.2) and (8.3) to define on $\mathbf{S}$ a second fundamental tensor, which for the sake of differentiation we may designate as $\mathbf{G}_{ij}$. Thus, on $\mathbf{S}$ we have

$$(ds)^2 = g_{ij}\,dx^i\,dx^j \text{ and } (d\bar{s})^2 = G_{ij}\,dx^i\,dx^j = \left(\bar{g}_{kl}\frac{\partial \bar{x}^k}{\partial x^i}\frac{\partial \bar{x}^l}{\partial x^j}\right)dx^i\,dx^j,$$

where $d\bar{s}\ (\neq ds)$ is the value of the element of arc length measured on $\bar{\mathbf{S}}$ corresponding to the coordinate differentials defined by equation (8.3) at the point $\bar{\mathbf{P}}$ which is the counterpart on $\bar{\mathbf{S}}$ of $\mathbf{P}$ on $\mathbf{S}$. The mapping therefore serves (1) to associate points of $\mathbf{S}$ to points of $\bar{\mathbf{S}}$ or, alternatively, (2) to define a second fundamental tensor on $\mathbf{S}$.

Consider a familiar example. Let $\mathbf{S}$ be a plane (sheet of paper) and $\bar{\mathbf{S}}$ a sphere (globe of the earth). Let x^1 and x^2 be Cartesian coordinates on $\mathbf{S}$; let $\bar{x}^1$ and $\bar{x}^2$ be the longitude and latitude on the globe. Let the x^1- axis $(x^2 = 0)$ correspond to the curve $\bar{x}^2 = 0$ (equator) and the x^2-axis $(x^1 = 0)$ correspond to the curve $\bar{x}^1 = 0$ (prime meridian). In the plane, the square of the line element is given by

$$(ds)^2 = (dx^1)^2 + (dx^2)^2$$

whereas on the sphere, assumed to have unit radius,

$$(d\bar{s})^2 = (\cos\bar{x}^2)^2(d\bar{x}^1)^2 + (d\bar{x}^2)^2.$$

In the plane and on the sphere, ds and $d\bar{s}$ could be measured directly by applying a steel tape along the geodesic between the points x^i and $x^i + dx^i$ or $\bar{x}^i$ and $\bar{x}^i + d\bar{x}^i$, respectively.

Let us now associate by means of some transformation the various points in the plane to counterparts on the globe. One possible way of doing this is by the reversible transformation

$$\text{(8.4)} \qquad \begin{cases} x^1 = \bar{x}^1, \quad x^2 = \log(\sec\bar{x}^2 + \tan\bar{x}^2), \\ \bar{x}^1 = x^1, \quad \bar{x}^2 = \tan^{-1}(\sinh x^2). \end{cases}$$

This transformation constitutes a mapping. With it, let us see what $d\bar{s}$ becomes in terms of the differentials of dx^i. Thus

$$d\bar{x}^1 = dx^1, \quad d\bar{x}^2 = (\operatorname{sech} x^2)\, dx^2,$$

$$(d\bar{s})^2 = (\cos\bar{x}^2)^2 (d\bar{x}^1)^2 + (d\bar{x}^2)^2 = (\operatorname{sech} x^2)^2 [(dx^1)^2 + (dx^2)^2],$$

whence (see Fig. 56)

$$(\sec\bar{x}^2)^2 (d\bar{s})^2 = (d\sigma)^2 = (dx^1)^2 + (dx^2)^2 = (ds)^2.$$

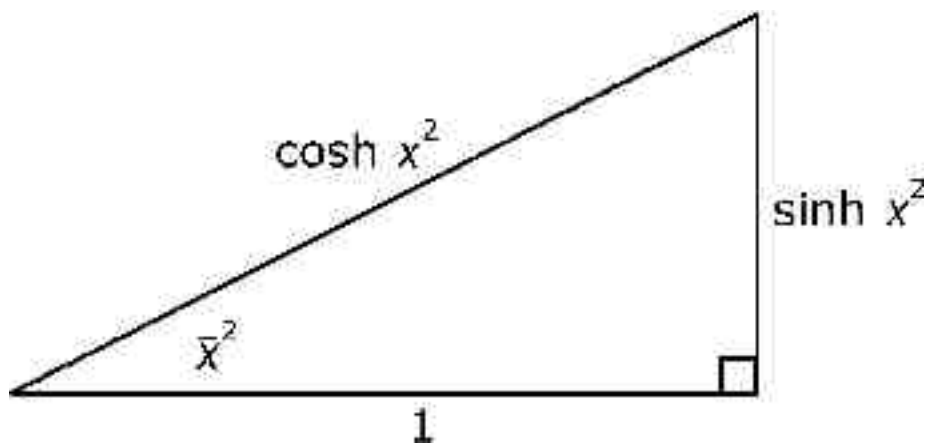

Figure 56

In other words, a new congruence relation which defines $d\sigma$ as $(\operatorname{sech}\bar{x}^2)\, d\bar{s}$ is the same as the ds originally defined in the plane. Therefore, whereas formerly steel tape could be used to measure either ds on the plane or $d\bar{s}$ on the sphere, now we use the steel tape exclusively in the plane for measurement of either ds or $d\sigma$. For measurments of $d\bar{s}$, we use a sliding scale on the map of $\bar{\mathrm{S}}$ onto S which differs from the steel tape by a factor $\cos\bar{x}^2$. This particular mapping is called a **Mercator mapping**. Mercator maps are provided with comparative scales for making differential distance measures.

Let us understand clearly what these manipulations signify. If we measure elements of arc length on the globe $\bar{\mathbf{S}}$ with the steel tape measure, then in terms of $\bar{x}^i$ (the longitude and latitude), infinitesimal lengths are given by

$$(d\bar{s})^2 = (\cos \bar{x}^2)^2 (d\bar{x}^1)^2 + (d\bar{x}^2)^2 .$$

The "map" of this same infinitesimal arc has a length in the plane which is

$$(ds)^2 = (dx^1)^2 + (dx^2)^2 = (\sec \bar{x}^2)^2 (d\bar{s})^2 = (d\bar{x}^1)^2 + (\sec \bar{x}^2)^2 (d\bar{x}^2)^2$$

as measured by the same steel tape measure. There are thus *two* ways of measuring arc lengths in the plane: (1) as

$$(ds)^2 = (dx^1)^2 + (dx^2)^2 = g_{ij}\, dx^i dx^j ;$$

and (2) as

$$(d\bar{s})^2 = (\operatorname{sech} x^2)^2 [(dx^1)^2 + (dx^2)^2] = \mathrm{G}_{ij}\, dx^i dx^j .$$

Thereby we define on $\mathbf{S}$ *two* fundamental tensors — g_{ij} and G_{ij}. The geometries associated with each will in general differ — different geodesics, curvatures, areas, trigonometries, etc. This serves to drive home the point that a "geometry" is not determined alone by the surface with which it is associated; it is only the *combination* of surface *and* congruence relation which fixes the "curvature" and "geometry". Thus a plane may be the locus of a "curved" geometry or, as by mapping a plane onto a sphere, a sphere may be the locus of a "flat" geometry.

In any case, a mapping may be used to define two fundamental tensors, g_{ij} and G_{ij}, from which all the metric properties of the space and its map may be inferred. If, for example,

$$g_{ij} = \mathrm{G}_{ij} \equiv \bar{g}_{kl} \frac{\partial \bar{x}^k}{\partial x^i} \frac{\partial \bar{x}^l}{\partial \bar{x}^j} ,$$

the transformation is the usual tensor transformation. Then all invariants of the space and its map are identical — geodesics, curvatures, areas, angles, etc. Such a mapping is an **isometric mapping**. A surface $\bar{\mathbf{S}}$ may be mapped isometrically onto a surface $\mathbf{S}$ only when the curvature of $\mathbf{S}$ and $\bar{\mathbf{S}}$ are the same. This is in general a trivial mapping, amounting only to a change of coordinates in the same space.

Consider next a **conformal** or **isogonal mapping**. This is a mapping such that the angle between two vectors u^i and v^i is the same in the space and its map.

Therefore

$$\frac{g_{ij} u^i \mathrm{v}^j}{\sqrt{(g_{kl} u^k u^l)(g_{mn} \mathrm{v}^m \mathrm{v}^n)}} = \cos\theta = \frac{\mathrm{G}_{ij} u^i \mathrm{v}^j}{\sqrt{(\mathrm{G}_{kl} u^k u^l)(\mathrm{G}_{mn} \mathrm{v}^m \mathrm{v}^n)}} .$$

It is clearly sufficient, and it can be shown to be necessary as well, that

$$\mathrm{G}_{ij} = f(x^k)\, g_{ij} . \tag{8.5}$$

The Mercator map is clearly conformal.

For the sake of illustration, let us consider a different conformal mapping of the sphere onto a plane. In this case, let us choose polar coordinates $x^1 = r$, $x^2 = \theta$ in the plane, where

$$g_{ij} = \begin{Vmatrix} 1 & 0 \\ 0 & r^2 \end{Vmatrix}.$$

Let the origin be the point of tangency to the sphere. On the sphere (assumed to be of unit radius), the coordinates are $\bar{x}^1 = \frac{\pi}{2} - \varphi$, $\bar{x}^2 = \theta$, φ being the latitude (see Fig. 57). The mapping is to be carried out by choosing $r = r(\varphi)$ in such a way as to satisfy equation (8.5). On the sphere

$$g_{ij} = \begin{Vmatrix} 1 & 0 \\ 0 & \cos^2\varphi \end{Vmatrix}.$$

Transforming from (r, θ) to variables (φ, θ) in the plane, and imposing the condition that the mapping be conformal, we have

$$G_{ij} = \frac{\partial \bar{x}^k}{\partial x^i}\frac{\partial \bar{x}^l}{\partial x^j} g_{kl} = \begin{Vmatrix} \left[\frac{dr}{d\varphi}\right]^2 & 0 \\ 0 & r^2 \end{Vmatrix} = \bar{f} g_{ij} = \bar{f} \begin{Vmatrix} 1 & 0 \\ 0 & \cos^2\varphi \end{Vmatrix}.$$

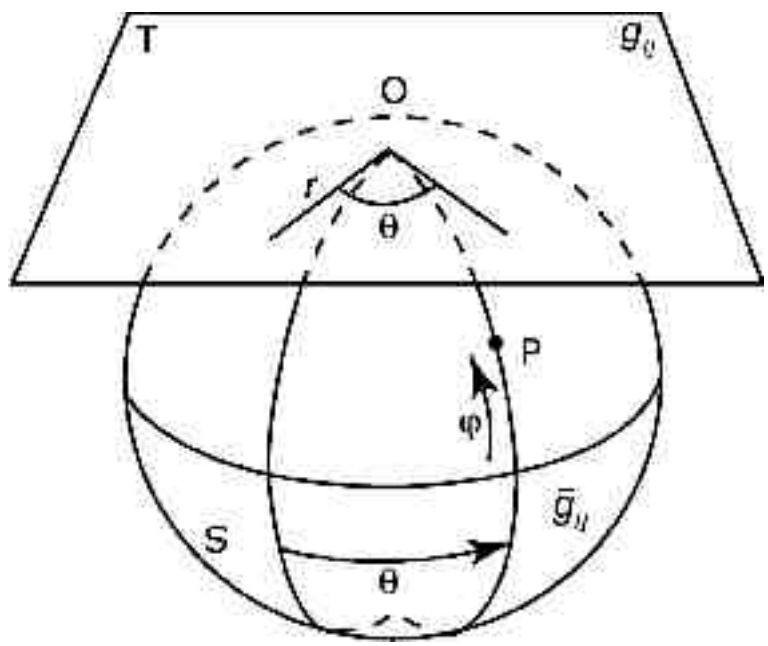

Figure 57

Evidently we may choose

$$f = \left(\frac{dr}{d\varphi}\right)^2, \quad f\cos^2\varphi = r^2 \text{ or } \frac{dr}{d\varphi} = \pm\frac{r}{\cos\varphi}$$

in order to satisfy the condition of conformality. Then, recognizing that an increase in r entails a decrease of φ,

$$\frac{dr}{r} = -\sec\varphi, \quad r = C\left(\frac{\pi}{4} - \frac{\varphi}{2}\right).$$

When $C = 2$, this mapping becomes the **stereographic projection** of a unit sphere from its south pole onto the plane tangent to the north pole (see Fig. 58).

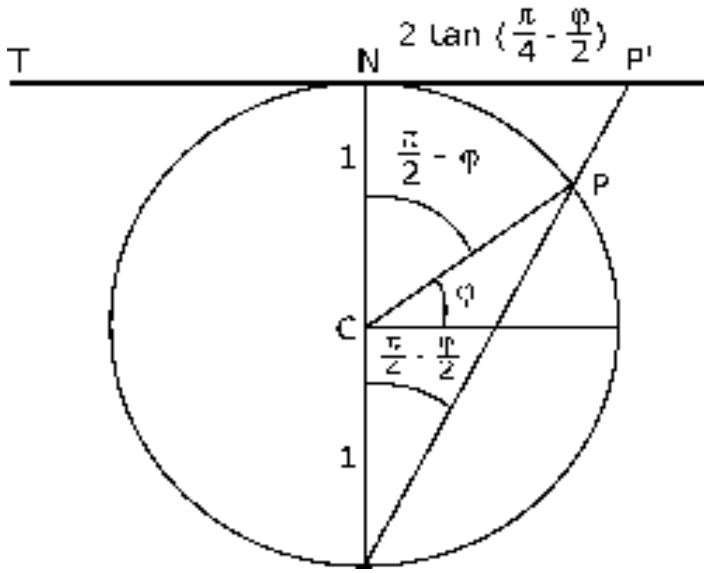

Figure 58

Still another method of mapping a sphere onto a plane is the **equi-areal mapping**. As the term implies, the areas of all figures on the sphere are equal to or proportional to the areas of their maps upon the plane. Now the element of area in any surface is $dA = \sqrt{g}\,dx^1\,dx^2$. Hence, using the same coordinates as before in the plane and on the sphere, we must require that

$$\sqrt{|g_{ij}|}\,dr\,d\theta = dA = \sqrt{|\bar{g}_{ij}|}\,d\varphi\,d\theta,$$

$$\frac{dr}{d\varphi} = \sqrt{\frac{|\bar{g}_{ij}|}{|g_{ij}|}} = \left[\frac{\cos^2\varphi}{r^2}\right]^{1/2} = \pm\frac{\cos\varphi}{r}.$$

Since increase of φ implies decrease of r, we choose the negative sign. Then

$$2\,r\,dr = -2\cos\varphi\,d\varphi, \quad r^2 = C - 2\sin\varphi.$$

Then in order that $r = 0$ when $\varphi = \frac{\pi}{2}$, we must choose C to be 2. Hence

$$r = \sqrt{2(1 - \sin\varphi)} = 2\sin\left(\frac{\pi}{4} - \frac{\varphi}{2}\right).$$

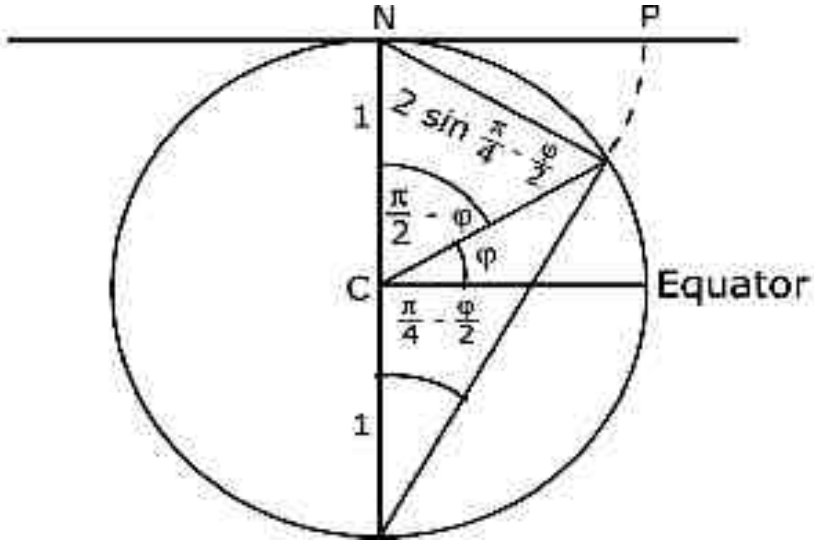

Figure 59

Still another type of mapping is the **geodesic mapping**, by which the geodesics of one surface go over into the geodesics of the other, or by which the geodesics of one metric on S become the geodesics of another. An example of such a mapping is the **gnomonic projection** of a sphere onto a tangent plane. In this, the sphere is projected through its center onto the plane tangent at the north pole. Since all planes through the center intersect the sphere in great circles and the tangent plane in straight lines, the geodesics of the sphere must map into those of the plane.

The transformation from ordinary polar coordinates (r, θ) in the tangent plane to the coordinates of the gnomonic projection can be seen from Fig. 60 to be

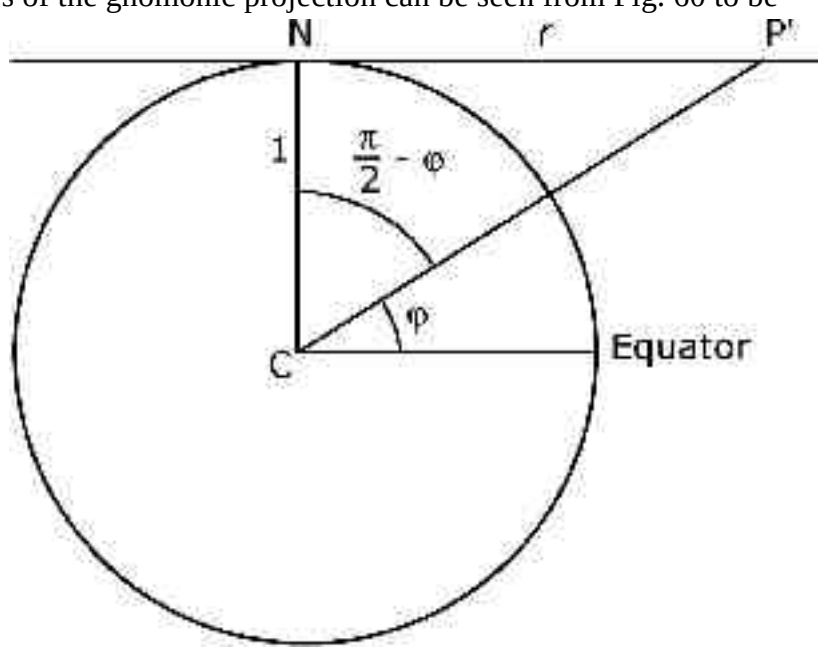

Figure 60

$$\bar{x}^1 = r = \tan\left(\frac{\pi}{2} - \varphi\right), \quad x^2 = \theta .$$

Geodesic mappings are generally difficult to find. However, we can state a condition which they must satisfy. Thus on a surface $\mathbf{S}$ let there be two fundamental tensors g_{ij} and $\bar{g}_{ij}$ to which are associated two line elements ds and $d\bar{s}$, respectively. If the geodesics of the metric g_{ij} and those of the metric $\bar{g}_{ij}$ are the same, then along such a geodesic $\mathrm{C}(x^i) \equiv \mathrm{C}(\bar{x}^i)$ we have

$$\text{(8.6)} \qquad \begin{cases} \dfrac{d^2x^i}{ds^2} + \Gamma^i_{jk}\dfrac{dx^j}{ds}\dfrac{dx^k}{ds} = 0 \\[2ex] \dfrac{d^2x^i}{d\bar{s}^2} + \bar{\Gamma}^i_{jk}\dfrac{dx^j}{d\bar{s}}\dfrac{dx^k}{d\bar{s}} = 0 \end{cases}$$

simultaneously. But, as we have seen in equations (4.8) and (4.9), changing the parameter from $\bar{s}$ to s in the second of these equations leads to the equivalent equation

$$\text{(8.7)} \qquad \frac{d^2x^i}{ds^2} + \bar{\Gamma}^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds} = \alpha\frac{dx^i}{ds},$$

where

$$\alpha = \frac{\dfrac{d^2s}{d\bar{s}^2}}{\left(\dfrac{ds}{d\bar{s}}\right)^2} = \frac{\dfrac{d^2\bar{s}}{ds^2}}{\dfrac{d\bar{s}}{ds}} = \frac{d}{ds}\ln\left(\frac{d\bar{s}}{ds}\right).$$

From equation (8.7) and the first of equations (8.6) we then have by subtraction that

$$\left(\Gamma^i_{jk} - \bar{\Gamma}^i_{jk}\right)\frac{dx^j}{ds}\frac{dx^k}{ds} = \alpha\frac{dx^i}{ds}.$$

We form from this the invariant

$$g_{il}\left(\Gamma^i_{jk} - \bar{\Gamma}^i_{jk}\right)\frac{dx^j}{ds}\frac{dx^k}{ds}\frac{dx^l}{ds} = -\alpha\, g_{il}\frac{dx^i}{ds}\frac{dx^l}{ds} = -\alpha .$$

Substituting this expression for α into equation (8.8) and setting

$$\text{(8.9)} \qquad \Gamma^i_{jk} - \bar{\Gamma}^i_{jk} \equiv \mathrm{A}^i_{jk},$$

we have

$$(\mathrm{A}^i_{jk}\dot{x}^j\dot{x}^k)(g_{rl}\dot{x}^r\dot{x}^l) = (\mathrm{A}^r_{jk}g_{rl}\dot{x}^j\dot{x}^k\dot{x}^l)\dot{x}^i,$$

$$g_{rl}\left[\mathrm{A}^k_{jk}\dot{x}^r - \mathrm{A}^r_{jk}\dot{x}^i\right]\dot{x}^j\dot{x}^k\dot{x}^l = 0$$

for arbitrary $\mathrm{v}^j = \dot{x}^j = dx^j/ds$. This therefore requires that

$$[\mathrm{A}^i_{jk}\mathrm{v}^r - \mathrm{A}^r_{jk}\mathrm{v}^i]\,\mathrm{v}^j\mathrm{v}^k\mathrm{v}_r = 0$$

for an arbitrary vector v^i. This is equivalent to the condition that

$$\left(A^i_{jk}\,\delta^r_l - A^r_{jk}\,\delta^i_l\right) v^j\, v^k\, v^l\, v_r = 0$$

for all permutations of j, k and l. Therefore

$$\left(A^i_{jk}\,\delta^r_l - A^r_{jk}\,\delta^i_l\right) + \left(A^i_{kl}\,\delta^r_j - A^r_{kl}\,\delta^i_j\right) + \left(A^i_{lj}\,\delta^r_k - A^r_{lj}\,\delta^i_k\right) = 0\,.$$

Setting $l = r$, we get

$$2\,A^i_{jk} - A^i_{jk} + A^i_{kj} - A^r_{kr}\,\delta^i_j + A^i_{kj} - A^r_{rj}\,\delta^i_k = 0\,,$$

$$3\,A^i_{jk} - \delta^i_j A^r_{kr} - \delta^i_k A^r_{jr} = 0\,.$$

Therefore it is necessary for a geodesic mapping that

(8.10) $$A^i_{jk} \equiv \Gamma^i_{jk} - \bar{\Gamma}^i_{jk} = \delta^i_j\,\varphi_k + \delta^i_k\,\varphi_j\,,$$ where

(8.11) $$\varphi_k = \frac{1}{3}\,A^r_{kr}\,.$$

But φ_k is a covariant vector, for by equation (8.8) and the Quotient Theorem A^i_{jk} is a tensor. Therefore, given any covariant vector φ_k, one may generate the functions $\bar{\Gamma}^i_{jk}$ from the Γ^i_{jk} and φ_k according to equation (8.10). Finding the $\bar{g}_{ij}$ for which these are the Christoffel symbols is yet another problem. Clearly, however, equations (8.10) and (8.11) represent necessary conditions which must be satisfied by any geodesic mapping and which may therefore be used to test any given mapping to ascertain whether or not it is geodesic.

One clearly geodesic mapping is given by the choice $\varphi_k = 0$. Then from equation (8.10) it follows that $\Gamma^i_{jk} = \bar{\Gamma}^i_{jk}$. Such a mapping is called an **affine mapping**. (See Ch. 7, §2.) Conceivably, different metrics could have the same Christoffel symbols, as when the g_{ij} and $\bar{g}_{ij}$ differ at most by an additive constant tensor γ_{ij}.

Of the several special properties of mapping – equi-areal, geodesic, conformal, or affine – usually at most one such property may be imposed upon any mapping unless that mapping is also **isometric** $(ds = d\bar{s})$. Thus, it can be shown that a mapping which is both conformal and equi-areal is also isometric; a mapping which is both geodesic and conformal is a **similarity mapping**, and differs at most by a constant from an isometric mapping; a mapping which is both equi-areal and geodesic is also affine and a similarity mapping. In general, therefore, a non-isometric mapping has at most one of the properties of being conformal, equi-areal or geodesic.

Notes — Chapter 2

With the background provided by Chapter 1, the student interested in an algebraic, rather than geometric, generalization, will find an excellent expository review in Davis (4), Ch. 5. This concludes with an illuminating summary of the history of the subject.

§2.6 The tedious direct method for calculating Christoffel symbols may often be avoided by using the Eulerian derivative

$$\frac{d}{ds}\left(\frac{\partial \mathrm{L}}{\partial \dot{x}^r}\right) - \frac{\partial \mathrm{L}}{\partial x^r} = g_{rj}\ddot{x}^j + [nj, r]\dot{x}^n\dot{x}^j$$

of the function $\mathrm{L} = \frac{1}{2}(g_{ij}\dot{x}^i\dot{x}^j)$. The Christoffel symbols of the first kind may then be identified as the coefficients of the $\dot{x}^n\dot{x}^j$, remembering that when $n \neq j$ the coefficient is really that of $([nj, r] + [jn, r])$. The contravariant form of the Eulerian derivative will give the Christoffel symbols of the second kind in similar fashion.

§2.8 Mappings are discussed at greater length in Kreyszig (5), Ch. IV, and Laugwitz (8), §13.