Generalized Coordinate Systems

Grant W. Mason

Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602

INTRODUCTION

Cartesian coordinates are the familiar rectangular coordinates x_i. They have the dimensions of $[length]$ so that speed has the dimensions of $[length]/[time]$ and acceleration has the dimensions of $[length]/[time]^2$. But what if you wish to use plane polar coordinates, r and θ? Theta does not have the dimension of length. How then should we write down an expression for acceleration in terms of polar coordinates? What we attempt below is to deal with so-called generalized coordinates so that we see how to write down correct expressions for positions, velocities and accelerations. We will try to do it in a general way so that we do it once and for all for plane polar coordinates, cylindrical coordinates, spherical coordinates, parabolic coordinates and some coordinate systems you haven't even made up yet.

Unless otherwise indicated, it is assumed that there is a summation on repeated indices.

Definitions

$\hat{\mathbf{e}}_i$	(unit vectors)
$\mathbf{a} = a_i\hat{\mathbf{e}}_i$	(physical components, a_i)
q^i, dq^i	(generalized coordinate, differential)
$d\mathbf{s}$	(differential line element)
$\mathbf{b}_i = \frac{\partial \mathbf{r}}{\partial q^i}$	(basis vectors, $\mathbf{b}_i$)
$\mathbf{a} = A^i\mathbf{b}_i$	(contravariant components, A^i)
$\mathbf{b}^i = \nabla q^i$	(reciprocal basis vectors, $\mathbf{b}^i$)
$\mathbf{a} = A_i\mathbf{b}^i$	(covariant components, A_i)
$g_{ij} = \mathbf{b}_i \cdot \mathbf{b}_j$	(covariant metric tensor, g_{ij})
$g^{ij} = \mathbf{b}^i \cdot \mathbf{b}^j$	(contravariant metric tensor, g^{ij})
$\Gamma^k_{ij} = \mathbf{b}^k \cdot \frac{\partial \mathbf{b}_i}{\partial q^j}$	(Christoffel symbols, Γ^k_{ij})
$T = 1/2m(ds/dt)^2$	(kinetic energy)
V	(potential energy)
$F_i = -\nabla_i V$	(covariant force component, F_i)
$L = T - V$	(Lagrangian)

Theorems

$d\mathbf{s} = dq^i \mathbf{b}_i$

$ds^2 = g_{ij} dq^i dq^j$

$\mathbf{b}^i \cdot \mathbf{b}_j = \delta^i_j$

$a_i = \sqrt{g_{(ii)}} A^i$ (no sum on (ii))

$g_{ij} g^{jk} = \delta^k_i$

$A_i = g_{ij} A^j$

$A^i = g^{ik} A_k$

$\mathbf{b}_i = g_{ij} \mathbf{b}^j$

$\mathbf{b}^i = g^{ij} \mathbf{b}_j$

$\frac{\partial \mathbf{b}_i}{\partial q^j} = \frac{\partial \mathbf{b}_j}{\partial q^i}$

$\Gamma^k_{ij} = g^{kl} \frac{1}{2} (\partial g_{il}/\partial q^j + \partial g_{jl}/\partial q^i - \partial g_{ij}/\partial q^l)$ (Christoffel symbols)

If $\mathbf{a} = a_i \hat{\mathbf{e}}_i = A^i \mathbf{b}_i = A_i \mathbf{b}^i$ is acceleration:

$A^k = (\ddot{q}^k + \dot{q}^i \dot{q}^j \Gamma^k_{ij})$

$A_l = \frac{d}{dt}(g_{lk} \dot{q}^k) - \frac{1}{2}(\partial g_{ij}/\partial q^l) \dot{q}^i \dot{q}^j$

$a_i = \sqrt{g_{(ii)}} A^i$ (no sum on (ii))

$a_i = \sqrt{g_{(ii)}} g^{ij} A_j$ (no sum on (ii))

$F_k = m A_k = \frac{d}{dt}(\frac{\partial T}{\partial \dot{q}^k}) - \frac{\partial T}{\partial q^k}$

As a practical matter, we use this machinery in the following manner. We write down an arbitrary d**s** from which we form $ds^2 \equiv d\mathbf{s} \cdot d\mathbf{s}$. We divide the expression for ds^2 by dt^2 to form $(ds/dt)^2$ and from this we form $T = (1/2)m(ds/dt)^2$. We then operate on T according to the prescription,

$$A_k = \frac{1}{m} \left(\frac{d}{dt} (\frac{\partial T}{\partial \dot{q}^k}) - \frac{\partial T}{\partial q^k} \right)$$

to obtain A_i. From the A_i we obtain the A^i and from the A^i we obtain the physical components of acceleration, a_i. These expressions tell us how to write down correctly the acceleration in our particular coordinate system. Since $F_k = mA_k$, we may also obtain the generalized forces if we know the acceleration *a priori*. In systems where a potential energy function, V, exists and we can write,

$$F_k = -\frac{\partial V}{\partial q^k}$$

we can dispense with most of the machinery by defining $L = T - V$ and operating on L according to the following prescription:

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

These latter equations are *Lagrange's Equations* and they are the differential equations of motion of the system.

In problems with particle motion in electric and magnetic fields, the Lagrangian, L, becomes:

$$L = T - V - e\psi + (e/c)(\mathbf{A}\cdot\mathbf{v})$$

where e is the charge on the particle, ψ is the electric potential, $\mathbf{A}$ is the magnetic vector potential and $\mathbf{v}$ is the velocity of the particle.

Example: Plane polar coordinates

$$\begin{array}{ll}
\hat{\mathbf{e}}_1 = \hat{\mathbf{r}} & \hat{\mathbf{e}}_2 = \hat{\theta} \\
x_1 = r\cos\theta & r = \sqrt{(x_1^2 + x_2^2)} \\
x_2 = r\sin\theta & \theta = \arctan(x_2/x_1) \\
\partial r/\partial x_1 = x_1/\sqrt{(x_1^2 + x_2^2)} = \cos\theta & \partial r/\partial x_2 = x_2/\sqrt{(x_1^2 + x_2^2)} = \sin\theta \\
\partial\theta/\partial x_1 = -x_2/(x_1^2 + x_2^2) = -\sin\theta/r & \partial\theta/\partial x_2 = x_1/(x_1^2 + x_2^2) = \cos\theta/r \\
\partial x_1/\partial r = \cos\theta & \partial x_1/\partial\theta = -r\sin\theta \\
\partial x_2/\partial r = \sin\theta & \partial x_2/\partial\theta = r\cos\theta
\end{array}$$

An Example of Generalized Coordinates

Coordinates are used to specify the position of a particle in space. In a three-dimensional space, we need three such coordinates. The three Cartesian coordinates (x_1, x_2, x_3) are an example. Spherical coordinates, cylindrical coordinates and plane polar coordinates are alternatives that are sometimes used. Consider, for example, spherical coordinates (r, θ, ϕ). Observe that while the Cartesian coordinates all have dimensions of length, only r in the set of spherical coordinates has dimensions of length. Newton's Second Law of Motion can easily be expressed in Cartesian coordinates as

$$F_i = m\frac{d^2x_i}{dt^2} = m\ddot{x}_i$$

where the second derivative of x_i with respect to time will have the dimensions of acceleration if Cartesian coordinates are used. But, how should we express the Second Law if we use coordinates that may not even have dimensions of length? In what follows we shall try to solve this problem once and for all for all admissible *generalized coordinates*. The development is somewhat abstract and general, and it will probably be useful to keep a concrete example in mind.

Consider spherical coordinates,

$$r = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

$$\theta = \arctan(\frac{\sqrt{x_1^2 + x_2^2}}{x_3})$$

$$\phi = \arctan(\frac{x_2}{x_1}).$$

Think of each of these as an example of the form $\psi(x_1, x_2, x_3) = c$ discussed in connection with the directional derivative. If r is a constant, then a spherical surface is defined. If θ is constant, then a conical surface is defined. If ϕ is constant, then a half-plane is defined. These surfaces intersect at a point whose coordinates are (r, θ, ϕ). Define unit vectors to lie along the lines of intersection of these surfaces and pointing in the direction of increasing coordinate. For example, $\hat{\mathbf{r}}$ lies along the line of intersection of the cone and the plane and points away from the origin, i.e. in the direction of increasing r if θ and ϕ are held constant to define the constant surfaces (cone and plane). Similarly, $\hat{\theta}$ lies along the intersection of sphere and plane in the direction of increasing θ. Finally, $\hat{\phi}$ lies along the intersection of sphere and cone in the direction of increasing ϕ. See Fig. 1.

Observe the following:

1. $\hat{\mathbf{r}}$, $\hat{\theta}$, $\hat{\phi}$ are unit vectors. They have unit length. In general, we will denote unit vectors associated with generalized coordinates as $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$,$\hat{\mathbf{e}}_3$.
2. (r, θ, ϕ) are an example of generalized coordinates. In general, we will denote generalized coordinates as (q^1, q^2, q^3).
3. A small displacement of the particle is denoted $d\mathbf{s}$. If θ and ϕ are held constant in spherical coordinates, a small displacement resulting from a change in r would be

   $$d\mathbf{s}(1) = dr\hat{\mathbf{r}}.$$

   Similarly, if r and ϕ are held constant,

   $$d\mathbf{s}(2) = rd\theta\hat{\theta}.$$

   If r and θ are held constant,

   $$d\mathbf{s}(3) = r\sin\theta d\phi\hat{\phi}.$$

4. A completely arbitrary displacement in spherical coordinates would be a vector sum of these three,

   $$d\mathbf{s} = d\mathbf{s}(1) + d\mathbf{s}(2) + d\mathbf{s}(3) = dr\hat{\mathbf{r}} + rd\theta\hat{\theta} + r\sin\theta d\phi\hat{\phi}.$$

   If we form the dot product, $d\mathbf{s} \cdot d\mathbf{s}$, we obtain

   $$ds^2 = dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2.$$

   In general, we denote this quantity as $ds^2 = g_{ij}dq^i dq^j$, where a double sum is intended.
5. The velocity of the particle is immediately found in spherical coordinates to be,

   $$\mathbf{v} = \frac{d\mathbf{s}}{dt} = \frac{dr}{dt}\hat{\mathbf{r}} + r\frac{d\theta}{dt}\hat{\theta} + r\sin\theta\frac{d\phi}{dt}\hat{\phi}.$$

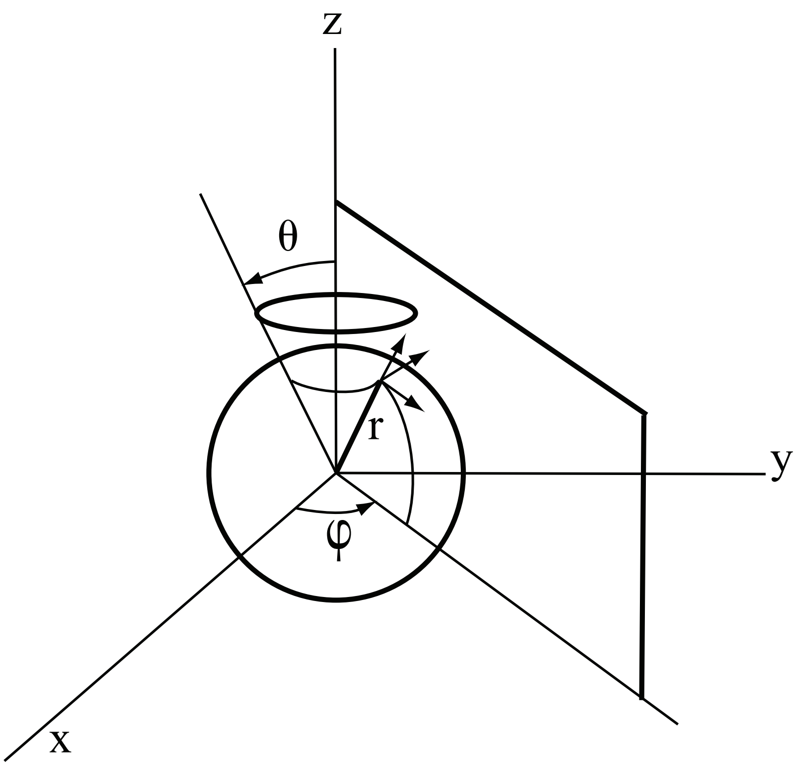

FIGURE 1. The intersection of constant surfaces (sphere, cone, and half-plane) that define the spherical coordinates of a point.

The square of the speed is,

$$v^2 = \dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\dot{\phi}^2.$$

In general, this latter quantity will be written,

$$v^2 = g_{ij}\dot{q}^i\dot{q}^j.$$

6. Finally, the kinetic energy of the particle in spherical coordinates may be written,

$$T = \frac{1}{2}mv^2 = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2).$$

   In general, we will write,

$$T = \frac{1}{2}mg_{ij}\dot{q}^i\dot{q}^j.$$

   The pattern which is illustrated here for finding the kinetic energy will turn out to be very important.

Generalized Coordinates

Imagine a set of generalized coordinates, $q^i(x_1, x_2, x_3)$. When each is set equal to a constant, a set of surfaces are defined for which the point of intersection, P, has coordinates (q_p^1, q_p^2, q_p^3). For spherical coordinates, the surface of constant r is a sphere, the surface of constant θ is a cone, and the surface of constant ϕ is a half-plane. The cone of angle θ_p intersects the sphere of radius r_p in a circle. The half-plane of constant ϕ_p intersects the circle at a point. The generalized coordinates that uniquely define this point, P, are (r_p, θ_p, ϕ_p).

Now, define a set of unit vectors, $\hat{\mathbf{e}}_i$, which are tangent to the lines of intersection of the various surfaces taken two at a time in the same way that the unit vectors in spherical coordinates were defined. Thus defined, the unit vectors must form a linearly independent set for the coordinates system to be "admissible." The unit vectors are not necessarily orthogonal, although, like the unit vectors in spherical coordinates, they often are. Unlike the unit vectors in a Cartesian system, the directions of the unit vectors depend on the point at which they are defined. For example, in spherical coordinates, the unit vector $\hat{\mathbf{r}}$ has a different direction for different points in space. Its direction is given by the tangent line along the intersection of the cone and the half-plane, but as θ or ϕ varies, this tangent line varies in space.

An arbitrary vector may be expressed as a linear combination of admissible unit vectors,

$$\mathbf{a} = a_1\hat{\mathbf{e}}_1 + a_2\hat{\mathbf{e}}_2 + a_3\hat{\mathbf{e}}_3 = a_i\hat{\mathbf{e}}_i.$$

The set of numbers, a_i, are said to be the *physical components* of **a**. The set of numbers, q^i, are said to be the *generalized coordinates* of a particle in the system, but they do not necessarily have dimensions of length. The coordinates must uniquely define the position of the particle.

Generalized Basis Vectors

We now introduce a new set of basis vectors that have the same direction as the unit vectors, but are not necessarily of unit length nor dimensionless. Consider a point of intersection of the level surfaces such that q_p^2 and q_p^3 are held fixed and only q^1 is allowed to vary. This defines the line of intersection of the 2-surface and the 3-surface which we can describe as a path $\mathbf{s}(1)$. A tangent vector to this path, $\mathbf{b}_1$ is obtained by differentiation,

$$\mathbf{b}_1 \equiv \frac{\partial \mathbf{s}(1)}{\partial q^1}.$$

The partial derivative indicates that q^2 and q^3 are being held constant. The faster q^1 changes in space along $s(1)$, the smaller will be the magnitude of $\mathbf{b}_1$. Two other generalized basis vectors, $\mathbf{b}_2$ and $\mathbf{b}_3$, are defined in a similar way.

In this context, $d\mathbf{s} = d\mathbf{s}(1) + d\mathbf{s}(2) + d\mathbf{s}(3)$ represents a differential change, $d\mathbf{r}$, to the position of a particle at position, $\mathbf{r}$. Thus, the partial derivatives imply that,

$$\frac{\partial \mathbf{r}}{\partial q^1} = \frac{\partial \mathbf{s}(1)}{\partial q^1},$$

$$\frac{\partial \mathbf{r}}{\partial q^2} = \frac{\partial \mathbf{s}(2)}{\partial q^2},$$

$$\frac{\partial \mathbf{r}}{\partial q^3} = \frac{\partial \mathbf{s}(3)}{\partial q^3}.$$

Therefore, we may write,

$$\mathbf{b}_i = \frac{\partial \mathbf{r}}{\partial q^i}.$$

If these new basis vectors are "admissible," then we may use them as a basis for expressing an arbitrary vector, $\mathbf{a}$,

$$\mathbf{a} = \mathbf{b}_1 A^1 + \mathbf{b}_2 A^2 + \mathbf{b}_3 A^3 = \mathbf{b}_i A^i.$$

The coefficients, A^i, are called the *contravariant components* of the vector $\mathbf{a}$. They are denoted with a superscript to establish a pattern for the Einstein summation convention: A repeated index, once as a subscript and once as a superscript, denotes summation. Unless otherwise indicated, this convention applies henceforth throughout this chapter.

Theorem

Theorem: $d\mathbf{s} = \mathbf{b}_i dq^i$.

Proof: Let $d\mathbf{s} = d\mathbf{s}(1) + d\mathbf{s}(2) + d\mathbf{s}(3)$ be the infinitesimal displacement from (q^1, q^2, q^3) to $(q^1 + dq^1, q^2 + dq^2, q^3 + dq^3)$. By the chain rule,

$$d\mathbf{s} = \frac{\partial \mathbf{s}(i)}{\partial q^i} dq^i.$$

(Sum on i.) But, by definition, $\mathbf{b}_i = \partial \mathbf{s}(i)/\partial q^i$, so $d\mathbf{s} = \mathbf{b}_i dq^i$. QED.

The Metric Tensor

The inner (dot) product of two arbitrary vectors, **a** and **c** may now be computed using the generalized basis vectors:

$$\mathbf{a}\cdot\mathbf{c} = (\mathbf{b}_i A^i)\cdot(\mathbf{b}_j C^j) = (\mathbf{b}_i\cdot\mathbf{b}_j)A^i C^j.$$

Thus, any inner product is characterized by the nine products, $\mathbf{b}_i\cdot\mathbf{b}_j$. These may be grouped together into a 3×3 matrix G called the *metric tensor* with components g_{ij} such that,

$$g_{ij} = \mathbf{b}_i\cdot\mathbf{b}_j.$$

Observe that the matrix is symmetric, $g_{ij} = g_{ji}$. (A tensor is a generalization of a vector. Vectors and tensors are defined by their transformation patterns when one changes from one coordinate system to another. In our present application, the tensor property of the metric tensor is not important.)

Theorem

Theorem: $ds^2 = g_{ij}dq^i dq^j$.

Proof: Since we have shown that we may write $d\mathbf{s} = \mathbf{b}_i dq^i$, it follows that,

$$d\mathbf{s}\cdot d\mathbf{s} = ds^2 = (\mathbf{b}_i\cdot\mathbf{b}_j)dq^i dq^j = g_{ij}dq^i dq^j.$$

QED.

As a practical matter, we obtain the metric tensor by inspection of the form of ds^2. For example, an arbitrary displacement in spherical coordinates is written,

$$d\mathbf{s} = dr\hat{\mathbf{r}} + rd\theta\hat{\theta} + (r\sin\theta)\hat{\phi}.$$

Taking the dot product with itself, we have,

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

Comparing to the form $ds^2 = g_{ij}dq^i dq^j$, we obtain, by inspection,

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}$$

We are now in a position to establish the relationship between the unit vectors and our generalized basis vectors. The length of one of the basis vectors is obtained from the inner product of the vector with itself,

$$|\mathbf{b}_i| = \sqrt{\mathbf{b}_i\cdot\mathbf{b}_i} = \sqrt{g_{(ii)}}.$$

(No sum on (ii).) Since the unit vectors are already known to have the same direction as the basis vectors, we immediately have

$$\hat{\mathbf{e}}_i = \frac{\mathbf{b}_i}{|\mathbf{b}_i|} = \frac{\mathbf{b}_i}{\sqrt{g_{(ii)}}}.$$

Theorem

Theorem: $a_i = \sqrt{g_{(ii)}} A^i$. (No sum on (ii).)
Proof:

$$\mathbf{a} = \mathbf{b}_i A^i = (\sqrt{g_{(ii)}}\hat{\mathbf{e}}_i) A^i = (\sqrt{g_{(ii)}} A^i)\hat{\mathbf{e}}_i = a_i \hat{\mathbf{e}}_i.$$

Thus, $a_i = \sqrt{g_{(ii)}} A^i$. QED.

Reciprocal Basis Vectors

Set each of the generalized coordinates, $q^i(x_1, x_2, x_3) = q^i(x_i) = c$, where each constant, c, defines a different surface. The gradients of these functions of the Cartesian coordinates are vectors that are normal to the respective surfaces at the point (q_p^1, q_p^2, q_p^3), i.e.,

$$\nabla q^i = \mathbf{b}^i.$$

The vectors, $\mathbf{b}^i$, are called *reciprocal basis vectors*. Two sets of basis vectors, $\mathbf{b}_i$ and $\mathbf{b}^i$, are “reciprocal” if they satisfy the relationship, $\mathbf{b}_i \cdot \mathbf{b}^j = \delta_i^j$. This relationship says that $\mathbf{b}^1$ is orthogonal to both $\mathbf{b}_2$ and $\mathbf{b}_3$. It also says that if $\mathbf{b}^1$ has a large magnitude, then its reciprocal, $\mathbf{b}_1$, has a small magnitude. However, we must show that the reciprocal relationship is, indeed, satisfied.

Theorem

Theorem: $\mathbf{b}_i \cdot \mathbf{b}^j = \delta_i^j$.
Proof: The proof is a simple application of the chain rule. The generalized coordinates are independent of one another. Hence,

$$\mathbf{b}_i \cdot \mathbf{b}^j = \frac{\partial \mathbf{r}}{\partial q^i} \cdot \nabla q^j = \frac{\partial x^k}{\partial q^i}\frac{\partial q^j}{\partial x^k} = \frac{\partial q^j}{\partial q^i} = \delta_i^j.$$

QED.

An arbitrary vector, $\mathbf{a}$, can be written in terms of reciprocal basis vectors just as well as it can be written in terms of unit vectors,

$$\mathbf{a} = A_1\mathbf{b}^1 + A_2\mathbf{b}^2 + A_3\mathbf{b}^3 = A_i\mathbf{b}^i.$$

The set of numbers, A_i, are called the *covariant components* of **a**. The reciprocal basis vectors do not necessarily have unit length nor are they necessarily parallel to the unit vectors, but they must be linearly independent if the coordinate system is to be admissible. The covariant components of **a** are not the same as the physical components of **a** nor are they the same as the contravariant components.

Theorem

Theorem: $A^i = \mathbf{b}^i \cdot \mathbf{a}$.
Proof: We use the relationship, $\mathbf{b}^i \cdot \mathbf{b}_j = \delta^i_j$.

$$\mathbf{b}^i \cdot \mathbf{a} = \mathbf{b}^i \cdot (A^j \mathbf{b}_j) = (\mathbf{b}^i \cdot \mathbf{b}_j) A^j = \delta^i_j A^j = A^i.$$

QED. In similar fashion, we can also show $A_i = \mathbf{b}_i \cdot \mathbf{a}$.

Define $g^{ij} \equiv \mathbf{b}^i \cdot \mathbf{b}^j$. By doing so, we maintain a parallelism between the reciprocal basis vectors and the basis vectors. The elements, g^{ij}, form the matrix G^{-1} and are called the contravariant components of the metric tensor to distinguish them from the covariant components of the metric tensor, g_{ij}. Observe that G^{-1}, like G, is symmetric, i.e. $g^{ij} = g^{ji}$.

The covariant and contravariant forms of the metric tensor are useful in relating covariant and contravariant components of vectors. Note that the pattern of notation is consistent: an index repeated once as a subscript and once as a superscript denotes summation.

Theorem

Theorem: $A_i = g_{ij} A^j$.
Proof:

$$A_i = \mathbf{b}_i \cdot \mathbf{a} = \mathbf{b}_i \cdot (\mathbf{b}_j A^j) = (\mathbf{b}_i \cdot \mathbf{b}_j) A^j = g_{ij} A^j.$$

QED. This important operation is called “lowering the index.”

Theorem

Theorem: $A^i = g^{ij} A_j$.
Proof:

$$A^i = \mathbf{b}^i \cdot \mathbf{a} = \mathbf{b}^i \cdot (\mathbf{b}^j A_j) = (\mathbf{b}^i \cdot \mathbf{b}^j) A_j = g^{ij} A_j.$$

QED. This important operation is called “raising the index.”

The covariant and contravariant forms of the metric tensor are inverses of each other,

$$G^{-1} G = I,$$

or,

$$g_{ij}g^{jk} = \delta_i^k.$$

If you apply the combination $G^{-1}G$ in succession to an arbitrary vector, G will first "lower the index" and G^{-1} will turn around and "raise the index," returning you to where you began, so that the combination of the two, one after the other, acts exactly as the identity operation. In practice, this property is used to find the contravariant components of the metric tensor after the covariant components are extracted from the form, $ds^2 = g_{ij}dq^i dq^j$.

Theorem

Theorem: $\mathbf{b}_i = g_{ik}\mathbf{b}^k$.
Proof:

1. In an admissible generalized coordinate system, the reciprocal basis vectors must form a linearly independent set of vectors. Thus, the set of basis vectors, $\mathbf{b}_i$, can be expressed as a linear combination of the reciprocal basis vectors,

   $$\mathbf{b}_i = a_{ij}\mathbf{b}^j.$$

2. We have previously established that,

   $$g_{ik} = g_{ki} = \mathbf{b}_k \cdot \mathbf{b}_i$$

   so,

   $$\begin{aligned} g_{ik} &= \mathbf{b}_k \cdot (a_{ij}\mathbf{b}^j) = a_{ij}(\mathbf{b}_k \cdot \mathbf{b}^j) \\ &= a_{ij}\delta_k^j = a_{ik}. \end{aligned}$$

3. We conclude,

   $$\mathbf{b}_i = g_{ij}\mathbf{b}^j.$$

   QED.

Theorem

Theorem: $\mathbf{b}^i = g^{ij}\mathbf{b}_j$.
Proof:

1. We have established that,

   $$g_{ij}g^{kj} = \delta_i^k$$

   $$\mathbf{b}_i = g_{ij}\mathbf{b}^j.$$

2. Hence, multiplying both sides of the latter by g^{ki} (and forming a sum indicated by the repeated index, i), we have,

   $$g^{ki}\mathbf{b}_i = g^{ki}(g_{ij}\mathbf{b}^j) = \delta_j^k\mathbf{b}^j = \mathbf{b}^k.$$

3. We conclude,
$$\mathbf{b}^i = g^{ij}\mathbf{b}_j.$$
QED.

Theorem

Theorem: $\partial\mathbf{b}_i/\partial q^j = \partial\mathbf{b}_j/\partial q^i$.

Proof: This symmetry of pattern merely reflects the fact that the order in which partial derivatives are taken does not matter:
$$\frac{\partial\mathbf{b}_i}{\partial q^j} = \frac{\partial}{\partial q^j}\frac{\partial\mathbf{r}}{\partial q^i} = \frac{\partial}{\partial q^i}\frac{\partial\mathbf{r}}{\partial q^j} = \frac{\partial\mathbf{b}_j}{\partial q^i}.$$
QED.

Since each of these nine partial derivatives, $\partial\mathbf{b}_i/\partial q^j$ are themselves vectors, we can express each as some linear combination of either the unit vectors, reciprocal basis vectors or basis vectors. We choose to express them as a linear combination of basis vectors,
$$\frac{\partial\mathbf{b}_i}{\partial q^j} = \Gamma^k_{ij}\mathbf{b}_k.$$
The numbers which form the coefficients of the sum generated by the dummy index k are called Christoffel symbols of the first kind.

Theorem

Theorem:
$$\Gamma^l_{ij} = g^{lk}\frac{1}{2}\left(\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k}\right)$$
Proof:

1. We have, by definition,
$$\mathbf{b}_i\cdot\mathbf{b}_k = g_{ik}$$
$$\mathbf{b}_j\cdot\mathbf{b}_k = g_{jk}$$
$$\mathbf{b}_i\cdot\mathbf{b}_j = g_{ij}.$$
2. Differentiate each of these expressions in turn to form,
$$\frac{\partial g_{ik}}{\partial q^j} = \mathbf{b}_i\cdot\frac{\partial\mathbf{b}_k}{\partial q^j} + \frac{\partial\mathbf{b}_i}{\partial q^j}\cdot\mathbf{b}_k$$
$$\frac{\partial g_{jk}}{\partial q^i} = \mathbf{b}_j\cdot\frac{\partial\mathbf{b}_k}{\partial q^i} + \frac{\partial\mathbf{b}_j}{\partial q^i}\cdot\mathbf{b}_k$$

$$\frac{\partial g_{ij}}{\partial q^k} = \mathbf{b}_i \cdot \frac{\partial \mathbf{b}_j}{\partial q^k} + \frac{\partial \mathbf{b}_i}{\partial q^k} \cdot \mathbf{b}_j$$

and form the combination,

$$\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k}.$$

3. We have proved as an earlier theorem that

$$\frac{\partial \mathbf{b}_k}{\partial q^j} = \frac{\partial \mathbf{b}_j}{\partial q^k}$$

and,

$$\frac{\partial \mathbf{b}_k}{\partial q^i} = \frac{\partial \mathbf{b}_i}{\partial q^k}.$$

Thus, when we form the combination

$$\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k},$$

four terms add out and two combine. We are left with the result,

$$\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k} = 2\frac{\partial \mathbf{b}_i}{\partial q^j} \cdot \mathbf{b}_k.$$

4. We may now use our assumption that

$$\frac{\partial \mathbf{b}_i}{\partial q^j} = \Gamma^l_{ij}\mathbf{b}_l.$$

Then,

$$\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k} = 2[\Gamma^l_{ij}\mathbf{b}_l] \cdot \mathbf{b}_k = 2\Gamma^l_{ij}(\mathbf{b}_l \cdot \mathbf{b}_k) = 2\Gamma^l_{ij}g_{lk}.$$

5. It is tempting to solve for Γ^l_{ij} by dividing both sides of this expression by $2g_{lk}$, but one must remember that the repeated index l indicates a sum of terms, not an isolated term. In fact, what we are dealing with here are twenty-seven equations, one for each of the combinations of i, j, and k! The correct way to proceed is to multiply both sides by g^{mk}, thus introducing an additional sum indicated by the dummy index k. We then use a result that we have already proved,

$$g^{mk}g_{lk} = \delta^m_l.$$

Thus,

$$2g^{mk}\Gamma^l_{ij}g_{lk} = 2\delta^m_l\Gamma^l_{ij} = 2\Gamma^m_{ij}.$$

6. We conclude,

$$\Gamma^m_{ij} = g^{mk}\frac{1}{2}\left(\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k}\right).$$

QED.

In principle, once one knows the covariant and contravariant forms of the metric tensor, the Christoffel symbols can be calculated. In three dimensional space there are twenty-seven of them. You will be relieved to know that for our purposes this is a formal result essential to the derivation of Lagrange's equations that follow, but not one that we will use as a practical tool.

APPLICATION TO ACCELERATION

The mathematics we have introduced can now be extended in several directions. It was used by Albert Einstein as the original language of his General Theory of Relativity and a simple modification of it is also the basis of the most elegant formalism of the Special Theory of Relativity (see Chapter 10). The mathematical formalism can also be used to give general expressions for the divergence, curl and Laplacian in generalized coordinates. However, our immediate purpose is to apply the formalism to a particular vector quantity, the acceleration of a particle, and thus to reformulate Newton's Second Law of Motion, $\mathbf{F} = m\mathbf{a}$, into a new form. The advantage of this new form is that it is done in terms of generalized coordinates, so that the reader gets to choose the most convenient coordinates for a problem without undue concern about how to write down acceleration properly in terms of the coordinates that have been chosen. Newton's Second Law is a vector equation which, when resolved into components, becomes a set of ordinary differential equations. The equations which result from the reformulation of Newton's Second Law are called Lagrange's Equations, but they are completely equivalent ordinary differential equations which describe the motion of a particle. Our immediate purpose is to show the connection between Newton's Second Law and Lagrange's equations. We shall first find the contravariant components of acceleration, then "lower the index" to find the covariant components. Finally, we will write the covariant form of Newton's Second Law and show that for a large class of problems, the covariant components of Newton's Second Law are Lagrange's equations.

Contravariant Components of Acceleration

1. We begin with the expression for an arbitrary displacement: $d\mathbf{s} = dq^i\mathbf{b}_i$.
2. We divide both sides by dt. Observe that this is a division and not the process of differentiation, but the outcome is to turn $d\mathbf{s}$ into a velocity,

$$\mathbf{v} = \frac{d\mathbf{s}}{dt} = \dot{q}^i\mathbf{b}_i.$$

3. To obtain acceleration, we must differentiate $\mathbf{v}$ with respect to time, but we must note that as the particle moves, the basis vectors change with its position and must therefore be considered functions of time. Hence,

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \ddot{q}^i\mathbf{b}_i + \dot{q}^i\frac{d\mathbf{b}_i}{dt}.$$

4. To compute $d\mathbf{b}_i/dt$, we can apply the chain rule,

$$\frac{d\mathbf{b}_i}{dt} = \frac{\partial \mathbf{b}_i}{\partial q^j}\frac{dq^j}{dt} = \dot{q}^j\frac{\partial \mathbf{b}_i}{\partial q^j}.$$

   From this, we conclude,

$$\mathbf{a} = \ddot{q}^i\mathbf{b}_i + \dot{q}^i\dot{q}^j\frac{\partial \mathbf{b}_i}{\partial q^j} = \ddot{q}^i\mathbf{b}_i + \dot{q}^i\dot{q}^j\Gamma^k_{ij}\mathbf{b}_k,$$

   or,

$$\mathbf{a} = [\ddot{q}^k + \dot{q}^i\dot{q}^j\Gamma^k_{ij}]\mathbf{b}_k = A^k\mathbf{b}_k.$$

   From this expression, we identify the contravariant components of acceleration as,

$$A^k = \ddot{q}^k + \dot{q}^i\dot{q}^j\Gamma^k_{ij}.$$

Covariant Components of Acceleration

1. To obtain the covariant components of acceleration, we lower the index using $A_l = g_{lk}A^k$.

$$A_l = g_{lk}A^k = g_{lk}\ddot{q}^k + g_{lk}\Gamma^k_{ij}\dot{q}^i\dot{q}^j.$$

2. We may use the expression which we have just derived for the Christoffel symbols to write,

$$g_{lk}\Gamma^k_{ij} = \frac{1}{2}g_{lk}g^{kn}\left(\frac{\partial g_{in}}{\partial q^j} + \frac{\partial g_{jn}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^n}\right).$$

   Observe,

$$g_{lk}g^{kn} = \delta^n_l,$$

   so that we can write,

$$g_{lk}\Gamma^k_{ij} = \frac{1}{2}\delta^n_l\left(\frac{\partial g_{in}}{\partial q^j} + \frac{\partial g_{jn}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^n}\right)$$

$$= \frac{1}{2}\left(\frac{\partial g_{il}}{\partial q^j} + \frac{\partial g_{jl}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^l}\right).$$

3. Thus,

$$A_l = g_{lk}\ddot{q}^k + \frac{1}{2}\left(\frac{\partial g_{il}}{\partial q^j} + \frac{\partial g_{jl}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^l}\right)\dot{q}^i\dot{q}^j.$$

   In principle, the covariant components of the acceleration can be obtained from this expression once the covariant form of the metric tensor is known. Again, this is a formal result that is essential to our derivation, but not one that we will use as a practical tool.

A Better Form for the Covariant Components of Acceleration

We now begin to change the form of the covariant components of acceleration to put them into a form that is a practical tool. Consider the first term in the expression we have just derived for A_l, namely,

$$g_{lk}\ddot{q}^k.$$

Observe that this term arises in the expression,

$$\frac{d}{dt}(g_{lk}\dot{q}^k) = g_{lk}\ddot{q}^k + \frac{\partial g_{lk}}{\partial q^m}\dot{q}^m\dot{q}^k,$$

where we have used the chain rule in the second term on the right to obtain

$$\frac{d}{dt}(g_{lk}) = \frac{\partial g_{lk}}{\partial q^m}\dot{q}^m.$$

A renaming of dummy indices (in the second term below) immediately yields,

$$\frac{d}{dt}(g_{lk}\dot{q}^k) = g_{lk}\ddot{q}^k + \frac{\partial g_{lj}}{\partial q^i}\dot{q}^i\dot{q}^j.$$

We may now write,

$$A_l = \frac{d}{dt}(g_{lk}\dot{q}^k) - \frac{\partial g_{lj}}{\partial q^i}\dot{q}^i\dot{q}^j$$
$$+\frac{1}{2}\left(\frac{\partial g_{il}}{\partial q^j} + \frac{\partial g_{jl}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^l}\right)\dot{q}^i\dot{q}^j.$$

Observe that by simply switching dummy indices and using the fact that $g_{il} = g_{li}$, we can show that

$$\frac{\partial g_{lj}}{\partial q^i}\dot{q}^i\dot{q}^j = \frac{\partial g_{il}}{\partial q^j}\dot{q}^j\dot{q}^i = \frac{\partial g_{jl}}{\partial q^i}\dot{q}^i\dot{q}^j.$$

These equalities may be used to eliminate the middle three terms and to reduce our expression for A_l to

$$A_l = \frac{d}{dt}(g_{lk}\dot{q}^k) - \frac{1}{2}\frac{\partial g_{ij}}{\partial q^l}\dot{q}^i\dot{q}^j.$$

Newton's Second Law in Covariant Form

Finally, we arrive at the point of our entire excursion into differential geometry! The covariant components of Newton's Second Law are $F_k = mA_k$. In some texts, the covariant components of force, F_k, are called generalized forces. In terms of physical components, of course, Newton's Second Law is the familiar $f_k = ma_k$.

1. Observe that

$$ds^2 = g_{ij}dq^idq^j$$

and that, therefore, we may write the kinetic energy, T, of a particle as,

$$T = \frac{1}{2}m(\frac{ds}{dt})^2 = \frac{1}{2}mg_{ij}\dot{q}^i\dot{q}^j.$$

2. Observe further that if we take the view of generalized coordinates and generalized velocities as independent variables for purposes of taking partial derivatives,

$$\frac{\partial T}{\partial q^k} = \frac{1}{2}m\frac{\partial g_{ij}}{\partial q^k}\dot{q}^i\dot{q}^j,$$

$$\frac{\partial T}{\partial \dot{q}^k} = mg_{ik}\dot{q}^i,$$

and,

$$\frac{d}{dt}(\frac{\partial T}{\partial \dot{q}^k}) = m\frac{d}{dt}(g_{ik}\dot{q}^i) = m\frac{d}{dt}(g_{ki}\dot{q}^i)$$

(The factor of one-half disappears in the latter two expressions because there is a double sum in the expression for T. If you have trouble seeing this, you should write out a short double sum so that you see it works.)

3. We may therefore write,

$$F_l = mA_l = m\frac{d}{dt}(g_{lk}\dot{q}^k) - \frac{1}{2}m\frac{\partial g_{ij}}{\partial q^l}\dot{q}^i\dot{q}^j = \frac{d}{dt}(\frac{\partial T}{\partial \dot{q}^l}) - \frac{\partial T}{\partial q^l}.$$

Newton's Second Law, written in covariant form is,

$$F_l = \frac{d}{dt}(\frac{\partial T}{\partial \dot{q}^l}) - \frac{\partial T}{\partial q^l}.$$

This is a perfectly general result. The primary difficulty with it is knowing how to write down F_l.

Lagrange's Equations

If we narrow ourselves to a class of problems that satisfies two additional conditions we arrive at a remarkable result. If the system we are describing is "conservative" so that we may define a potential energy function, V, such that

$$F_l = -\frac{\partial V}{\partial q^l}$$

and V depends only on the generalized coordinates and not on the generalized velocities, $\dot{q}^i$, then we may simplify the covariant form of Newton's Second Law. We may write,

$$\frac{d}{dt}(\frac{\partial T}{\partial \dot{q}^l}) - \frac{\partial T}{\partial q^l} = -\frac{\partial V}{\partial q^l}$$

$$\frac{d}{dt}\frac{\partial(T-V)}{\partial\dot{q}^l}-\frac{\partial(T-V)}{\partial q^l}=0.$$

(Since V does not depend on the generalized velocities, it may formally be included in the first term without making any difference.) We now define the so-called “Lagrangian”, L,

$$L \equiv T-V$$

and write

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}^l}\right)-\frac{\partial L}{\partial q^l}=0.$$

These are Lagrange’s equations and they are very important in theoretical mechanics.

1. Lagrange’s equations are just the components of Newton’s Second Law for conservative systems written in a different and probably unfamiliar form.
2. The equations that are obtained by operating on the Lagrangian according to the prescription of Lagrange’s equations are ordinary differential equations. They are the “equations of motion” of the system. Solutions of the equations are models of the motion of the system.
3. Much of the intervening machinery used to derive the equations disappears and does not return. For example, we do not directly use (for our purposes) $\mathbf{b}_i$, $\mathbf{b}^i$, Γ^k_{ij}, or g^{ij}. These were introduced in order to derive Lagrange’s equations and to show their connection to Newton’s Second Law.
4. T and V are usually easy to write down correctly and once written down lead to the differential equations of the system in a very methodical way.

PROBLEMS

1. Do the following:
    - Draw a diagram (in two dimensions) of a particle whose position is specified by plane polar coordinates (r,θ). Show the unit vectors $\hat{\mathbf{r}}$ and $\hat{\theta}$ on your diagram and clearly label them.
    - Express an arbitrary differential displacement $d\mathbf{s}$ in terms of these coordinates and show this displacement on your diagram. Find $ds^2 = d\mathbf{s}\cdot d\mathbf{s}$ in terms of dr and $d\theta$ and extract from ds^2 the metric tensor g_{ij}. Display the metric tensor as a 2×2 matrix. Using ds^2, write down the kinetic energy of the particle in terms of r and θ and their time derivatives.
    - Determine g^{ij} and display as a 2×2 matrix.
    - Determine the covariant, contravariant and physical components of the acceleration of the particle. Show how these might appear on a diagram. (The diagram is only to be qualitative, i.e. it need only indicate directions of the vectors and whether the vectors have unit length or not.)

- Let a particle move on a circle of constant radius b. What is the radial component of acceleration (physical)? What is the transverse component?
- A bug crawls outward with constant speed v_0 along the spoke of a wheel which is rotating with constant angular speed ω. Find the radial and transverse components of the physical acceleration as functions of time. Assume $r = 0$ at $t = 0$. (Ans: $a_r = -v_0 t\omega^2$, $a_\theta = 2v_0\omega$)

2. Cylindrical coordinates are defined by $x = r\cos\theta$, $y = r\sin\theta$, $z = z$. Calculate the metric tensor, physical velocity and physical acceleration in cylindrical coordinates.
3. Do the following:
   - Beginning with Newton's Second Law,

     $$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^k}\right) - \frac{\partial T}{\partial q^k} = F_k,$$

     show that if F_k can be expressed in the form,

     $$F_k = -\frac{\partial V}{\partial q^k}(q^1, q^2, q^3),$$

     then,

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

   - For the one-dimensional harmonic oscillator, we can write $F_x = -kx$. Show that $V = \frac{1}{2}kx^2$. Define the function $L = T - V$ and obtain the differential equation of the simple harmonic oscillator, $m\ddot{x} + kx = 0$.
4. Obtain the differential equations of motion for a tether ball. (Spherical coordinates. Observe that the tether line has constant length. The system has only two degrees of freedom and there will only be two differential equations of motion. The length of the tether appears as a constant in the Lagrangian.)
5. Obtain the differential equations of motion for a projectile on a flat earth. (Cartesian coordinates)
6. Obtain the differential equations of motion for the earth's motion about the sun. (Plane polar coordinates)
7. Obtain the differential equations of motion for a particle moving without friction on the inside of a cone. The cone is parameterized in cylindrical coordinates as $z = \alpha r$. Write the equations in terms of r, $\dot{r}$ and $\ddot{r}$ by eliminating z.
8. Obtain the differential equations of motion for a particle sliding without friction from the top of a hemispherical dome. (Plane polar coordinates. The particle does not separate from the dome. The system has only one degree of freedom since the radius of the motion appears as a constant in the Lagrangian.)
9. Consider an inclined plane consisting of a right-triangular shaped block that is free to slide on its bottom along a frictionless horizontal surface. If the incline slopes toward the right, the left-hand edge of the block is vertical. Define a coordinate X that keeps track of the position of this vertical side. Now, put another block on the inclined plane so that it slides down the plane (without friction) under the

influence of gravity. Keep track of the second block's position with a coordinate x measured from the top of the inclined plane. Obtain the differential equations of the motions of both blocks and solve them directly for the accelerations $(\ddot{x}, \ddot{X})$ of each. (Unusual, nonorthogal coordinates. This problem introduces a problem not found in the previous ones. Here there are two masses. In such a case, the Lagrangian is written simply as the sum of the the two independent kinetic energies of the two objects minus the sum of the two independent potential energies. The system has two degrees of freedom.)