Tensor Analysis in Euclidean Space

James Emery

Revision 4/18/2007

Contents

1 Classical Tensor Notation

Given two sets of coordinates x_1, x_2, x_3 and y_1, y_2, y_3, where each y_i, for $i = 1, 2, 3$ is a function of the x_1, x_2, x_3, and where each x_i, for $i = 1, 2, 3$ is a function of the y_1, y_2, y_3. We have the following relationship between differentials

$$dy_i = \sum_{j=1}^{3} \frac{\partial y_i}{\partial x_j} dx_j = \frac{\partial y_i}{\partial x_j} dx_j,$$

where we use the Einstein summation convention in the last expression. This convention assumes that repeated indices are summed, so that we avoid writing the summation symbol. On the other hand given a function f, consider the relationships between the sets of partial derivatives of f in the two coordinate systems.

$$\frac{\partial f}{\partial y_i} = \frac{\partial x_j}{\partial y_i} \frac{\partial f}{\partial x_j}.$$

These two transformation rules are different. The coefficients

$$\frac{\partial y_i}{\partial x_j},$$

occur in the first transformation rule, but the coefficients

$$\frac{\partial x_j}{\partial y_i}$$

occur in the second rule.

Suppose some entity has coefficients $C(i, x), i = 1, 3$ in the x coordinate system, and coefficients $C(i, y), i = 1, 3$ in the y. Suppose these coefficients transform linearly so that

$$C(i, y) = a(i, j)C(j, x).$$

Suppose

$$a(i, j) = \frac{\partial y_i}{\partial x_j},$$

then the coefficients of the entity transform as in the differential example above. In this case the coefficients are said to be the coefficients of a rank one contravarient tensor. In this case the coefficients are written with superscript indices

$$C(i, x) = C^i(x), C(i, y) = C^i(y).$$

And the transformation law is written as

$$C^i(y) = \frac{\partial y_i}{\partial x_j} C^j(x).$$

The convention is that indices to be summed should occur once as a superscript and once as a subscript.
Suppose on the other hand that

$$a(i,j) = \frac{\partial x_j}{\partial y_i},$$

then the coefficients of the entity transform as in the partial derivative example above. In this case the coefficients are said to be the coefficients of a rank one covarient tensor. In this case the coefficients are written with subscript indices

$$C(i,x) = C_i(x), C(i,y) = C_i(y).$$

And the transformation law is written as

$$C_i(y) = \frac{\partial x_j}{\partial y_i} C_j(x).$$

Notice that in transforming from the x coefficients to a y coefficients suming is done over an index j on x. So the cases

$$a(i,j) = \frac{\partial y_j}{\partial x_i},$$

and

$$a(i,j) = \frac{\partial x_i}{\partial y_j},$$

do not make sense.

The names covarient and contravarient correspond to transformations that, go resectively, like the "ordinary" transformation, and against the "ordinary" transformation. But what is the "ordinary" transformation. I am not aware of a good argument for designating the "ordinary" transformation. I think that the names covarient and contravarient should be taken just as a convention. These designations seem to be applied consistently in all the books that I am aware of.

Higher rank tensors are defined similarly. So suppose we have a rank 5 tensor $C^{ijk}_{lm}(x)$, which is contravarient of rank 3 and covarient of rank 2. Then the transformation rule is

$$C^{nop}_{qr}(y) = \frac{\partial y_n}{\partial x_i}\frac{\partial y_o}{\partial x_j}\frac{\partial y_p}{\partial x_k}\frac{\partial x_l}{\partial y_q}\frac{\partial x_m}{\partial y_r}C^{ijk}_{lm}(x).$$

So we know how to transform the tensor components, but what are the actual tensors? Tensors were thought of originally as the infinite set of all possible coeffecients in all possible coordinate systems. This is not a very clear thought. It turns out that tensors can be interpreted as multilinear functionals defined on products of vector spaces. One such vector space is called the tangent space at a point. So this is something like the case of a 2d surface, say an ellipsoid, which at a point p has a tangent plane, which is a 2d vector space. This is for example where the velocity vectors of mechanics would live. The dual of this tangent space, in the linear algebra sense, is called the cotangent space. So contravarient and covarient tensors correspond to multilinear functionals on products of tangent and cotangent spaces.

2 Multilinear Functionals

Suppose $u_1, u_2, ..., u_n$ are a basis of a vector space V. Suppose $u^1, u^2, ..., u^n$ are duals of these vectors. Recall that the duals are linear functionals defined by

$$u^i(u_j) = \delta^i_j,$$

having value 0 if i and j differ, and 1 if $i = j$. They are a basis of the dual vector space. A linear functional f on the vector space is a real valued map that satisfies

$$f(u + v) = f(u) + f(v),$$

where u and v are vectors, and

$$f(\alpha u) = \alpha f(u),$$

where α is a scalar. A multilinear functional is a real valued mapping from products of the vector space and products of the dual space. So as an example, suppose u and v are in V, and ω is a vector in the dual of V. If a mapping

$$f(u, v, \omega)$$

is linear in variables u, v, ω, then f is a multilinear functional. It is a tensor contravariant of degree 2, and covariant of degree 1. The connection with the classically defined tensor components is that the coefficients are the values of the multilinear functional on basis vectors. Thus in this example we have say components

$$C_k^{ij} = f(u_i, u_j, u^k).$$

We see that a general tensor f, contravariant of degree m and covariant of degree n would have components

$$C_{j_1,\ldots,j_n}^{i_1,\ldots,i_m} = f(u_i, ...u_m, u^1, ..., u^n).$$

3 Operations With Tensors

See the Speigel reference, page 169, for these operations on tensors:

1. *Addition.*
2. *Outer Multiplication.*
3. *Contraction.*
4. *Inner Multiplication.*
5. *Quotient Law.*

4 The Directional Derivative

Let x_1, x_2, x_3 be the Cartesian coordinates of a point P. Let A be a coordinate vector attached to the point P that has coordinates A_i so that

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix}.$$

Let f be a real function defined in a neighborhood of the point P. We define a directional derivative of f in the direction of A by

$$A_P[f] = \frac{d}{dt} f(P + tA)|_0$$

$$= \frac{\partial f}{\partial x_1} A_1 + \frac{\partial f}{\partial x_2} A_2 + \frac{\partial f}{\partial x_3} A_3.$$

Suppose vector A is the ith unit coordinate vector u_i defined by

$$u_i = \begin{bmatrix} \delta_{i1} \\ \delta_{i2} \\ \delta_{i3} \end{bmatrix},$$

where δ is the Kronecker delta.

We may identify directional derivatives and tangent vectors. So the set of tangent vectors at a point is equivalent to the set of directional derivatives at the point. These vectors are called the tangent space at the point P.

The directional derivative in the ith coordinate direction is

$$(u_i)_P(f) = \frac{\partial f}{\partial x_1}.$$

Hence the directional derivative is the operator

$$\frac{\partial}{\partial x_i}.$$

We may use this operator for the ith unit coordinate vector u_i.

In the general case of a curved surface, or of a more general differential manifold, we can essentially identify the tangent space with the set of differentiable curves through a point. This is because such curves have derivatives which are tangents, which in turn define directional derivatives of functions. Of course distinct curves can define the same tangent. These operators on functions are derivations, and they constitute the tangent space. See **Differential Geometry** by James Emery.

5 Curvilinear Coordinates

Let q_i, q_2, q_3 be a new set of coordinates related to the Cartesian coordinates by a function g, where

$$x_i = g_i(q_1, q_2, q_3).$$

We assume that the Jacobian of g is not zero in the domain of the coordinate system.

Let f be a real valued function defined on 3-space. We have

$$\frac{\partial f}{\partial q_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial q_i} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial q_i} + \frac{\partial f}{\partial x_3}\frac{\partial x_3}{\partial q_i}.$$

If the vector A is defined as

$$A = \begin{bmatrix} \frac{\partial x_1}{\partial q_i} \\ \frac{\partial x_2}{\partial q_i} \\ \frac{\partial x_3}{\partial q_i} \end{bmatrix},$$

then the directional derivative of f in the direction A is

$$A_P[f] = \frac{\partial f}{\partial q_i}.$$

Hence we shall denote the ith curvilinear coordinate vector as

$$\frac{\partial}{\partial q_i} = \begin{bmatrix} \frac{\partial x_1}{\partial q_i} \\ \frac{\partial x_2}{\partial q_i} \\ \frac{\partial x_3}{\partial q_i} \end{bmatrix}.$$

Then from the equation above

$$\frac{\partial}{\partial q_i} = \frac{\partial x_1}{\partial q_i}\frac{\partial}{\partial x_1} + \frac{\partial x_2}{\partial q_i}\frac{\partial}{\partial x_2} + \frac{\partial x_3}{\partial q_i}\frac{\partial}{\partial x_3}.$$

As an example consider spherical coordinates (r, θ, ϕ) . They are defined by

$$x_1 = r\sin(\theta)\cos(\phi),$$
$$x_2 = r\sin(\theta)\sin(\phi),$$
$$x_3 = r\cos(\theta).$$

Then the spherical coordinate vectors are

$$\frac{\partial}{\partial r} = \begin{bmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix}$$

$$\frac{\partial}{\partial \theta} = \begin{bmatrix} r\cos(\theta)\cos(\phi) \\ r\cos(\theta)\sin(\phi) \\ -r\sin(\theta) \end{bmatrix}$$

$$\frac{\partial}{\partial \phi} = \begin{bmatrix} -r\sin(\theta)\sin(\phi) \\ r\sin(\theta)\cos(\phi) \\ 0 \end{bmatrix}$$

Let an ith coordinate curve be C_i. This is the curve with q_i as parameter, and where the other coordinates are held fixed. That is,

$$C_i(q_i) = \begin{bmatrix} x_1(q_i) \\ x_2(q_i) \\ x_3(q_i) \end{bmatrix} = \begin{bmatrix} g_1(q_1, q_2, q_3) \\ g_2(q_1, q_2, q_3) \\ g_3(q_1, q_2, q_3) \end{bmatrix},$$

where two of the coordinates are held constant. For example, if $i = 2$, then q_1 and q_3 are held fixed.

It is clear that the ith coordinate vector is tangent to this coordinate curve. The ith coordinate vector is the directional derivative operator in the direction of the ith coordinate curve tangent vector. Therefore there is a one to one correspondence between the tangent vectors and the directional derivative operators. These curvilinear vectors are not necessarily unit vectors, nor are they necessarily orthogonal. A generalization of the directional derivative is the derivation, which is introduced in the theory of differential manifolds. A derivation has certain linear and product properties when operating on real valued functions. A tangent space at a point can be defined as the set of derivations.

In classical mathematical physics a tangent vector was thought of as an infinitesimal difference between points. Hence the components of tangent vectors were written as differentials. So a tangent vector A would appear as

$$A = \begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix}.$$

We shall find the components of the tangent vector A in the curvilinear coordinate system q_1, q_2, q_3. Let us write

$$A = dx^1 \frac{\partial}{\partial x_1} + dx^2 \frac{\partial}{\partial x_2} + dx^3 \frac{\partial}{\partial x_3}$$

$$= dx^1 \sum_{j=1}^{3} \frac{\partial q_j}{\partial x_1} \frac{\partial}{\partial q_j} + dx^2 \sum_{j=1}^{3} \frac{\partial q_j}{\partial x_2} \frac{\partial}{\partial q_j} + dx^3 \sum_{j=1}^{3} \frac{\partial q_j}{\partial x_3} \frac{\partial}{\partial q_j}$$

$$= (\sum_{i=1}^{3} \frac{\partial q_1}{\partial x_i} dx^i) \frac{\partial}{\partial q_1} + (\sum_{i=1}^{3} \frac{\partial q_2}{\partial x_i} dx^i) \frac{\partial}{\partial q_2} + (\sum_{i=1}^{3} \frac{\partial q_3}{\partial x_i} dx^i) \frac{\partial}{\partial q_3}.$$

Then

$$dq^i = \sum_{i=1}^{3} \frac{\partial q_i}{\partial x_j} dx^j.$$

A vector in the tangent space has a representation in a coordinate system with coordinates q_1, q_2, q_3 as

$$c^1 \frac{\partial}{\partial q_1} + c^2 \frac{\partial}{\partial q_2} + c^3 \frac{\partial}{\partial q_3}$$

and in a coordinate system with coordinates $\bar{q}_1, \bar{q}_2, \bar{q}_3$ as

$$\bar{c}^1 \frac{\partial}{\partial \bar{q}_1} + \bar{c}^2 \frac{\partial}{\partial \bar{q}_2} + \bar{c}^3 \frac{\partial}{\partial \bar{q}_3}$$

.

For rather obscure reasons a tangent vector is called a contravariant vector. The coefficients c^i of a contravariant vector are traditionally written with superscripts. Using the Einstein summation convention the transformation law has the form

$$\bar{c}^i = \frac{\partial \bar{q}_i}{\partial q_j} c^j.$$

Classically, the square of the length of the tangent vector, expressed in Euclidean orthogonal coordinates, is written as

$$ds^2 = (A, A) = dx_1^2 + dx_2^2 + dx_3^2,$$

This is a diagonalized quadratic form in the coordinates of a tangent vector. In a curvilinear coordinate system the quadratic form will not necessarily be diagonal. The tensor that gives the length of a tangent vector is known as the metric tensor. A tensor is a multilinear functional defined on products of tangent spaces and tangent space duals.

The space dual to a given tangent space with basis

$$\frac{\partial}{\partial q_1}, \frac{\partial}{\partial q_1}, \frac{\partial}{\partial q_1}.$$

has a basis

$$\mathbf{dq}^1, \mathbf{dq}^2, \mathbf{dq}^3.$$

These are linear functionals defined on the tangent space and are defined by

$$\mathbf{dq}^i(\frac{\partial}{\partial q_j}) = \delta^i_j.$$

The modern interpretation of a differential takes it to be a functional on the space of tangent vectors. Thus the differential of a function $f(q_1, q_2, q_3)$ is

$$\mathbf{df} = \frac{\partial f}{\partial q_1}\mathbf{dq}^1 + \frac{\partial f}{\partial q_2}\mathbf{dq}^2 + \frac{\partial f}{\partial q_3}\mathbf{dq}^3.$$

It is in the dual vector space of the tangent space.

These are called covariant vectors. Their coefficient transformation law is

$$\bar{d}_i = \frac{\partial q_j}{\partial \bar{q}_i} d_j.$$

Their coefficients are written with subscripts.

If A is a tangent vector (contravariant vector), and dB is a differential form (covariant vector), then

$$dB(A)$$

is equal to sum of the products of the coefficients of the two vectors in a given coordinate system. That is, it is the inner product of the coefficients. Because we are dealing with a Euclidean space, the distinction between contravariant and covariant vectors is not so great as it would be on some surface or general manifold. Traditionally the concept of a dual basis was handled as a reciprocal basis in the same vector space. So given a set $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ of basis vectors, one could construct a reciprocal basis $\mathbf{b}^1, \mathbf{b}^2, \mathbf{b}^3$ with the property that

$$\mathbf{b}_i \cdot \mathbf{b}^j = \delta^j_i$$

Then the $\mathbf{b}^i$ play the role of the differential forms. A reciprocal basis may be constructed explicitely as

$$\mathbf{b}^1 = \frac{\mathbf{b}_2 \times \mathbf{b}_3}{\mathbf{b}_1 \cdot \mathbf{b}_2 \times \mathbf{b}_3},$$

$$\mathbf{b}^2 = \frac{\mathbf{b}_3 \times \mathbf{b}_1}{\mathbf{b}_1 \cdot \mathbf{b}_2 \times \mathbf{b}_3},$$

and

$$\mathbf{b}^3 = \frac{\mathbf{b}_1 \times \mathbf{b}_2}{\mathbf{b}_1 \cdot \mathbf{b}_2 \times \mathbf{b}_3}.$$

In our case we shall define the $\mathbf{b}_i$ as

$$\mathbf{b}_i = \frac{\partial}{\partial q_i}$$

$$= \frac{\partial x_j}{\partial q_i}\frac{\partial}{\partial x_j}.$$

and the $\mathbf{b}^i$ as

$$\mathbf{b}^i = \mathbf{dq}^i$$

$$= \frac{\partial q_i}{\partial x_j}\mathbf{dx}^j.$$

Because we are in Euclidean space, we can think of the coefficient vectors as being in the same space, and thus write

$$\mathbf{b}_i = \begin{bmatrix} \frac{\partial x_1}{\partial q_i} \\ \frac{\partial x_2}{\partial q_i} \\ \frac{\partial x_3}{\partial q_i} \end{bmatrix}$$

and

$$\mathbf{b}^j = \begin{bmatrix} \frac{\partial q_1}{\partial x_j} \\ \frac{\partial q_2}{\partial x_j} \\ \frac{\partial q_3}{\partial x_j} \end{bmatrix}$$

Then

$$\mathbf{dq}^j(\frac{\partial}{\partial q_i}) = \mathbf{b}^j \cdot \mathbf{b}_i.$$

The mapping between coordinate systems is invertible. So let

$$Q = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}$$

and

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Then we have functions f and g so that

$$X = f(Q)$$

and

$$Q = g(X)$$

Then $f \circ g$ is the identity. Then taking derivatives and using the chain rule, we have

$$J_f J_g = I$$

where J_f and J_g are the Jacobians of the derivatives. And I is the identity matrix. Therefore the product of the ith row and the jth column gives

$$\frac{\partial x_i}{\partial q_k}\frac{\partial q_k}{\partial x_j} = \delta^i_j.$$

That is

$$\mathbf{b}_i \cdot \mathbf{b}^j = \delta^j_i.$$

This shows explicitely that $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ and $\{\mathbf{b}^1, \mathbf{b}^2, \mathbf{b}^3\}$ are reciprocal systems, and also that

$$\mathbf{dq}^j(\frac{\partial}{\partial q_i}) = \delta^j_i.$$

We must introduce a metric tensor giving the length of tangent vectors with respect to a given curvalinear coordinate system. Suppose the coordinates are $q_1, q_2, ..., q_n$. The tangent space basis vectors for $i = 1, .., n$ are

$$\mathbf{b}_i = \frac{\partial}{\partial q_i}.$$

Given two tangent vectors,

$$\mathbf{u} = u^i \mathbf{b}_i$$

and

$$\mathbf{v} = v^i \mathbf{b}_i,$$

the metric g, a bilinear form, has value

$$g(\mathbf{u}, \mathbf{v}) = g_{ij} u^i v^j.$$

The bilinear form g, which is a rank 2 covariant tensor, is defined as the usual inner product in a Cartesian coordinate system. So that for the normal Euclidean coordinates x_1, x_2, x_3, metric g is given by

$$g_{ij} = \delta_{ij}.$$

For a general curvilinear system q_1, q_2, q_3, the basis vectors

$$\frac{\partial}{\partial q_i}$$

have Cartesian coordinates

$$\frac{\partial}{\partial q_i} = \begin{bmatrix} \frac{\partial x_1}{\partial q_i} \\ \frac{\partial x_2}{\partial q_i} \\ \frac{\partial x_3}{\partial q_i} \end{bmatrix}.$$

Therefore the coefficients of g for this coordinate system are

$$g_{ij} = \frac{\partial x_k}{\partial q_i} \frac{\partial x_k}{\partial q_j}.$$

The length of the tangent vector u is

$$\sqrt{g(u,u)}.$$

In the coordinate system q_1, q_2, q_3, we have g_{ii} as the square of the length of the ith basis vector

$$\frac{\partial}{\partial q_i}.$$

This metric is the Euclidean metric, so that if the basis vectors of the given curvilinear coordinate system are orthogonal, then

$$g_{ij} = 0.$$

For most of the common systems of coordinates, this is the case.
Consider spherical coordinates:

$$\frac{\partial}{\partial r} = \begin{bmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix}$$

$$\frac{\partial}{\partial \theta} = \begin{bmatrix} r\cos(\theta)\cos(\phi) \\ r\cos(\theta)\sin(\phi) \\ -r\sin(\theta) \end{bmatrix}$$

$$\frac{\partial}{\partial \phi} = \begin{bmatrix} -r\sin(\theta)\sin(\phi) \\ r\sin(\theta)\cos(\phi) \\ 0 \end{bmatrix}$$

Then

$$g_{11} = 1$$
$$g_{22} = r^2$$
$$g_{33} = r^2\sin^2(\theta).$$

In the case of an orthogonal coordinate system, it is common to use the notation

$$h_i = \sqrt{g_{ii}}.$$

Let $\mathbf{b}_i$ be the representation of the vector

$$\frac{\partial}{\partial q_i},$$

in the Cartesian coordinate system x_1, x_2, x_3. Let $\mathbf{b}^1, \mathbf{b}^2, \mathbf{b}^3$ be the reciprocal basis. For some a_{ij} we have

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

So

$$g_{ki} = \mathbf{b}_k \cdot \mathbf{b}_i = a_{ij}\mathbf{b}_k\mathbf{b}^j = a_{i,j}\delta_k^j = a_{i,k}.$$

And so

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

Similarly,

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

where

$$g^{ij} = \mathbf{b}^i \cdot \mathbf{b}^j.$$

Notice that because

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

we have

$$\delta_i^k = \mathbf{b}_i \cdot \mathbf{b}^k = g_{ij}\mathbf{b}^j \cdot \mathbf{b}^k = g_{ij}g^{jk}.$$

Let χ be a function. The gradient is

$$\nabla\chi = \frac{\partial\chi}{\partial q_k}\nabla q_k$$

$$= \mathbf{b}^k \frac{\partial\chi}{\partial q_k}$$

$$= g^{kj}\mathbf{b}_j \frac{\partial\chi}{\partial q_k}.$$

If the coordinate system q_1, q_2, q_3 is orthogonal, this becomes

$$\nabla\chi = g^{kk}\mathbf{b}_k \frac{\partial\chi}{\partial q_k}.$$

$$= \frac{1}{g_{kk}}\frac{\partial\chi}{\partial q_k}\frac{\partial}{\partial q_k}.$$

$$= \frac{1}{h_k}\frac{\partial\chi}{\partial q_k}u_k,$$

where u_k is a unit vector in the direction of

$$\frac{\partial}{\partial q_k}.$$

6 Polar Coordinates

As a very simple example of curvilinear coordinates, consider polar coordinates, ϕ, r in 2-space.

$$x = r\cos(\phi)$$

$$y = r\sin(\phi)$$

The unit coordinate vectors are

$$\mathbf{u}_\phi = \frac{\partial}{\partial \phi}$$

and

$$\mathbf{u}_r = \frac{\partial}{\partial r}.$$

The metric is

$$ds^2 = g_{11}d\phi^2 + g_{12}d\phi dr + g_{21}drd\phi + g_{22}dr^2$$

$$= r^2 d\phi^2 + dr^2.$$

As a quadratic form this may be written as

$$ds^2 = \begin{bmatrix} d\phi & dr \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} d\phi \\ dr \end{bmatrix},$$

where the matrix

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix},$$

is a positive definite symmetric matrix. A velocity $\mathbf{v}$ would be written as

$$\mathbf{v} = v_1 \mathbf{u}_\phi + v_2 \mathbf{u}_r.$$

To differentiate the velocity, we can not just differentiate the components v_1 and v_2, because the unit coordinate vectors $\mathbf{u}_\phi$, and $\mathbf{u}_r$, vary with position. This is why the Christoffel Symbols are needed for a covariant derivative, that is, a derivative independent of coordinate system.

7 The Derivative of a Curve

Suppose we have a curve $C(t)$ expressed in coordinates $q_1(t), q_2(t), q_3(t)$. The curve tangent, namely the derivative, is

$$\frac{dq_1}{dt}, \frac{dq_2}{dt}, \frac{dq_3}{dt}$$

The ith curve in a second coordinate system is

$$\bar{q}_i(t) = \bar{q}_i(q_1(t), q_2(t), q_3(t)).$$

So

$$\frac{d\bar{q}_i}{dt} = \frac{\partial \bar{q}_i}{\partial q_j}\frac{dq_j}{dt}.$$

We see that the velocity components transform according to the contravariant transformation law. Hence we get a covariant representation of the curve tangent by just differentiating the coefficients in any coordinate system. The term "covariant" used here was used by Einstein and means that a definition depends solely on the set of coordinates, and not on any underlying fundamental coordinate system. This is similar to the principle of relativity in physics, which among other things says that there is no absolute space. The meaning of "covariant" used here differs from the meaning in a phase such as "covariant" vector.

8 Properties of the Metric Tensor

Let restate for reference the properties of the metric tensor. Its components are g_{ij}, in a flat Euclidean space this is the partial derivative of the euclidean cordinate x_i with respect to the curvalinear coordinate q_j. It is a quadratic form represented by a symmetric matrix. We have

$$g_{ij} = (\mathbf{b}_i, \mathbf{b}_j),$$

the inner product of contravariant coordinate basis vectors. The matrix inverse is the matrix with components g^{ij}, and

$$g^{ij} = (\mathbf{b}^i, \mathbf{b}^j),$$

which is the inner product of the dual basis coordinate vectors. The basis vectors and the dual basis vectors are linearly related with the g_{ij} as coefficients. The Christoffel symbols are equal to a relation involving the g^{ij} and the coordinate derivatives of the g_{ij}. In a general Riemannian, or semiReimannian space, the metric coefficients define a unique covariant derivative. See the Emery Differential geometry reference, or the book by Hicks, **Notes on Differential Geometry**.

9 Velocity

As shown in the previous section, given the curve of a particle, its velocity is a tangent vector obtained by differentiating the coordinate functions with respect to time in any coordinate system.

10 Acceleration, Christoffel Symbols, Metric Coefficients

In a fixed orthogonal Cartesian coordinate system the acceleration is obtained as the second derivatives of the coordinate functions with respect to time. In more general curvilinear coordinate systems this is not a valid way of getting the acceleration. This is not valid method mainly because the coordinate frames are not fixed, but vary as the location is changed. Suppose the velocity vector is given as

$$\mathbf{v}(t) = \dot{q}^i \mathbf{b}_i,$$

where $\mathbf{b}_i$ is the coordinate vector

$$\frac{\partial}{\partial q_i}.$$

We can find the acceleration by differentiating $\mathbf{v}$ with respect to time t, but notice that not only is $\dot{q}^i$ a function of time, but $\mathbf{b}_i$ is a function of the coordinates q_1, q_2, q_3, which in turn are functions of the time t. So we must find a way to differentiate the vector $\mathbf{b}_i$.
We have

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

$$= \ddot{q}^i \mathbf{b}_i + \dot{q}^i (\frac{\partial \mathbf{b}_i}{\partial q^j}\frac{dq^j}{dt})$$

$$= \ddot{q}^i \mathbf{b}_i + \dot{q}^i \dot{q}^j \frac{\partial \mathbf{b}_i}{\partial q^j}.$$

Let us write

$$\frac{\partial \mathbf{b}_i}{\partial q^j} = \Gamma^k_{ij} \mathbf{b}_k,$$

where the Γ^k_{ij} are called the Christoffel symbols of the second kind. We have

$$\mathbf{b}_i = \frac{\partial x^k}{\partial q_i}\frac{\partial}{\partial x^k}.$$

So $\mathbf{b}_i$ is identified with the coordinate vector

$$(\frac{\partial x^1}{\partial q_i}, \frac{\partial x^2}{\partial q_i}, \frac{\partial x^3}{\partial q_i}).$$

Because reversing the order of differentiation does not change the value, we get the symmetry

$$\frac{\partial \mathbf{b}_i}{\partial q_j} = \frac{\partial \mathbf{b}_j}{\partial q_i}.$$

Or treating the derivative as a second order differential operator, we get

$$\frac{\partial \mathbf{b}_i}{\partial q_j} = \frac{\partial^2}{\partial q_j \partial q_i}$$

$$= \frac{\partial^2}{\partial q_i \partial q_j} = \frac{\partial \mathbf{b}_j}{\partial q_i}.$$

So again we obtain the symmetry

$$\frac{\partial \mathbf{b}_i}{\partial q_j} = \frac{\partial \mathbf{b}_j}{\partial q_i}.$$

So also

$$\Gamma^k_{ij} = \Gamma^k_{ji}.$$

If we differentiate

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

with respect to q_j then we get

$$\mathbf{b}_i \cdot \frac{\partial \mathbf{b}_k}{\partial q_j} + \mathbf{b}_k \cdot \frac{\partial \mathbf{b}_i}{\partial q_j} = \frac{\partial g_{ik}}{\partial q_j}.$$

We can interchange indices i and j, and also interchange indices k and j, and so get two more similar equations. We get

$$\mathbf{b}_j \cdot \frac{\partial \mathbf{b}_k}{\partial q_i} + \mathbf{b}_k \cdot \frac{\partial \mathbf{b}_j}{\partial q_i} = \frac{\partial g_{jk}}{\partial q_i},$$

and

$$\mathbf{b}_i \cdot \frac{\partial \mathbf{b}_j}{\partial q_k} + \mathbf{b}_j \cdot \frac{\partial \mathbf{b}_i}{\partial q_k} = \frac{\partial g_{ij}}{\partial q_k}.$$

Adding the first two equations, subtracting the third, and using the symmetry found above, we get

$$\frac{\partial g_{ik}}{\partial q_j} + \frac{\partial g_{jk}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_k}$$

$$= \mathbf{b}_i \cdot (\frac{\partial \mathbf{b}_k}{\partial q_j} - \frac{\partial \mathbf{b}_j}{\partial q_k}) + \mathbf{b}_j \cdot (\frac{\partial \mathbf{b}_k}{\partial q_i} - \frac{\partial \mathbf{b}_i}{\partial q_k}) + \mathbf{b}_k \cdot (\frac{\partial \mathbf{b}_i}{\partial q_j} + \frac{\partial \mathbf{b}_j}{\partial q_i})$$

$$= 0 + 0 + 2b_k \cdot \frac{\partial \mathbf{b}_i}{\partial q_j}$$

$$= 2\frac{\partial b_i}{\partial q_j} \cdot \mathbf{b}_k$$

$$= 2\Gamma^n_{ij} \mathbf{b}_n \cdot \mathbf{b}_k$$

$$= 2\Gamma^n_{ij} g_{nk}.$$

Continuing, we get

$$(\frac{\partial g_{ik}}{\partial q_j} + \frac{\partial g_{jk}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_k}) g^{km}$$

$$= 2\Gamma^n_{ij} g_{nk} g^{km}$$

$$= 2\Gamma^n_{ij} \delta^m_n$$

$$= 2\Gamma^m_{ij}.$$

Therefore

$$\Gamma_{ij}^{m} = \frac{1}{2}g^{km}(\frac{\partial g_{ik}}{\partial q_j} + \frac{\partial g_{jk}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_k})$$

Using this result, we can write the acceleration as

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

$$= \ddot{q}^i\mathbf{b}_i + \dot{q}^i\dot{q}^j\Gamma_{ij}^{k}\mathbf{b}_k.$$

$$= (\ddot{q}^k + \dot{q}^i\dot{q}^j\Gamma_{ij}^{k})\mathbf{b}_k.$$

So the coefficients of the contravariant acceleration are

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

We can use the metric tensor g to get a one-to-one onto map from the space of contravariant vectors to the space of covariant vectors.

A covariant vector is a linear functional on the tangent space of contravariant vectors. Thus, for A fixed,

$$A'(...) = g(A, ...),$$

is a linear functional and hence a covariant vector corresponding to A. If X is any contravariant vector with components X^i, then we have

$$A'(X) = g(A, X) = g_{ij}A^iX^j.$$

It follows that the components of the covariant acceleration are

$$A_j = g_{ij}A^i.$$

This technique of getting covariant components from contravariant components is called "lowering the index."

Now we will find an expression for the acceleration, and thus the generalized force. This comes from the Lagrangian. It relates to the covariant form of Newton's second law.

For the components of acceleration we have

$$A_l = g_{lk}(\ddot{q}^k + \dot{q}^i\dot{q}^j\Gamma_{ij}^{k})$$

$$= g_{lk}\ddot{q}^k + \dot{q}^i\dot{q}^j g_{lk}[\frac{1}{2}g^{km}(\frac{\partial g_{im}}{\partial q_j} + \frac{\partial g_{jm}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_m})]$$

$$= g_{lk}\ddot{q}^k + \dot{q}^i\dot{q}^j[\frac{1}{2}(\frac{\partial g_{il}}{\partial q_j} + \frac{\partial g_{jl}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_l})]$$

$$= g_{lk}\ddot{q}^k + \frac{1}{2}(\frac{\partial g_{il}}{\partial q_j} + \frac{\partial g_{jl}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_l})\dot{q}^i\dot{q}^j$$

$$= g_{lk}\ddot{q}^k + \frac{\partial g_{il}}{\partial q_j}\dot{q}^i\dot{q}^j - \frac{1}{2}\frac{\partial g_{ij}}{\partial q_l}\dot{q}^i\dot{q}^j$$

$$= \frac{d}{dt}(g_{li}\dot{q}^i) - \frac{1}{2}\frac{\partial g_{ij}}{\partial q_l}\dot{q}^i\dot{q}^j$$

11 The Covariant Form of Newton's Second Law

We see that the covariant form of Newton's second law is

$$Q_k = mA_k,$$

where the Q_k are the generalized forces, m is the mass, and the A_K are the generalized components of acceleration. The units of A_k are not necessarily length per time squared, but in the case of an orthogonal coordinate system, where $g_{ij} = 0$ when i is not equal to j, we see that

$$\sqrt{g^{kk}}A_k = \frac{1}{\sqrt{g_{kk}}}A_k$$

does have these units.

The kinetic energy of a particle is

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

It is a function of the independent variables $q_1, ..., q_n, \dot{q}_1, ..., \dot{q}_n$. So

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

and
$$\frac{\partial T}{\partial \dot{q}^k} = m g_{ij} \dot{q}^i.$$

Then we have
$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^k}\right) - \frac{\partial T}{\partial q^k}$$
$$= m\left[\frac{d}{dt}(g_{ij}\dot{q}^i) - \frac{1}{2}\frac{\partial g_{ij}}{\partial q^k}\dot{q}^i\dot{q}^j\right]$$
$$= mA_k = Q_k.$$

This allows us to find the forces of constraint when solving a mechanics problem using Lagrange's equations. If there is a potential function V, then the force due to the potential gets included in the Lagrangian, which is defined by
$$L = T - V.$$

Then we have
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) - \frac{\partial L}{\partial q^k} = Q_k.$$

And Q_k is the force of constraint as a generalized force. By definition of Q_k the increment of work is
$$dW = Q_k dq_k.$$

For more about generalized force and forces of constraint, see the **Emery** entry, and the **Goldstein** entry in the bibliography. Also see the spherical coordinate example in **Bradbury**.

12 Bibliography

[1]Emery J. D., **Physics**, JDE Press, 1996, (See the chapter on mechanics.)
[2]Bradbury T. C., **Theoretical Mechanics**, sections 2.15 to 2.21, John Wiley, 1968.
[3]Lass Harry, **Vector and Tensor Analysis**, McGraw-Hill, 1950.

[4]Goldstein, Herbert **Classical Mechanics**,Addison-Wesley, 1965.
[5]Bishop Richard L, Goldberg Samuel I, **Tensor Analysis on Manifolds**, MacMillan, 1968.
[6]Marion Marion B, **Classical Dynamics of Particles and Systems**, Academic Press, 1970.
[7]Spiegel Murray R, **Vector Analysis**, Schaum Publishing Co., 1959.