INTRODUCTION TO THE ESSENTIALS OF TENSOR CALCULUS

****

****

I. Basic Principles

We shall treat only the basic ideas, which will suffice for much of physics. The objective is to analyze problems in any coordinate system, the variables of which are expressed as

$$q^j(x^i) \quad \text{or} \quad q'^j(q^i) \qquad\qquad \text{where } x^i : \text{ Cartesian coordinates, } i = 1,2,3, \ldots N$$

for any dimension N. Often N=3, but in special relativity, N=4, and the results apply in any dimension. Any well-defined set of q^j will do. Some explicit requirements will be specified later.

An invariant is the same in any system of coordinates. A vector, however, has components which depend upon the system chosen. To determine how the components change (transform) with system, we choose a prototypical vector, a small displacement dx^i. (Of course, a vector is a geometrical object which is, in some sense, independent of coordinate system, but since it can be prescribed or quantified only as components in each particular coordinate system, the approach here is the most straightforward.) By the chain rule, $dq^i = (\partial q^i / \partial x^j) \, dx^j$, where we use the famous summation convention of tensor calculus: each repeated index in an expression, here j, is to be summed from 1 to N. The relation above gives a prescription for transforming the (contravariant) vector dx^i to another system. This establishes the rule for transforming any contravariant vector from one system to another.

$$A^i (q) = \left(\frac{\partial q^i}{\partial x^j}\right) A^j (x)$$

$$A^i(q') = \left(\frac{\partial q'^i}{\partial q^j}\right) A^j(q) = \left(\frac{\partial q'^i}{\partial q^j}\right) \left(\frac{\partial q^j}{\partial x^k}\right) A^k(x) \equiv \left(\frac{\partial q'^i}{\partial x^k}\right) A^k(x)$$

$$\Lambda^i_j (q,x) \equiv \frac{\partial q^i}{\partial x^j} \qquad \text{Contravariant vector transform}$$

The (contravariant) vector is a mathematical object whose representation in terms of components transforms according to this rule. The conventional notation represents only the object, A^k, without indicating the coordinate system. To clarify this discussion of transformations, the coordinate system will be indicated by $A^k(x)$, but this should <u>not</u> be misunderstood as implying that the components in the "x" system are actually expressed as <u>functions</u> of the x^i. (The choice of variables to be used to express the results is totally independent of the choice coordinate system in which to express the components A^k. The $A^k(q)$ might still be expressed in terms of the x^i, or $A^k(x)$ might be more conveniently expressed in terms of some q^i.)

Distance is the prototypical invariant. In Cartesian coordinates, $ds^2 = \delta_{ij} \, dx^i \, dx^j$, where δ_{ij} is the Kroneker delta: unity if i=j, 0 otherwise. Using the chain rule,

$$dx^i = \left(\frac{\partial x^i}{\partial q^j}\right) dq^j$$

$$ds^2 = \delta_{ij} \left(\frac{\partial x^i}{\partial q^k}\right) \left(\frac{\partial x^j}{\partial q^l}\right) dq^k\, dq^l = g_{kl}(q)\, dq^k\, dq^l$$

$$g_{kl}(q) \equiv \left(\frac{\partial x^i}{\partial q^k}\right) \left(\frac{\partial x^j}{\partial q^l}\right) \delta_{ij} \qquad \text{(definition of the metric tensor)}$$

One is thus led to a new object, the metric tensor, a (covariant) tensor, and by analogy, the covariant transform coefficients:

$$\underline{\Lambda}_i^j(q,x) \equiv \left(\frac{\partial x^j}{\partial q^i}\right) \qquad \text{Covariant vector transform}$$

{More generally, one can introduce an arbitrary <u>measure</u> (a generalized notion of 'distance') in a chosen reference coordinate system by $ds^2 = g_{kl}(0)\, dq^k\, dq^l$, and that measure will be invariant if g_{kl} transforms as a covariant tensor. A space having a measure is a metric space.}

Unfortunately, the preservation of an invariant has required two different transformation rules, and thus two types of vectors, covariant and contravariant, which transform by definition according to the rules above. (The root of the problem is that our naive notion of 'vector' is simple and well-defined only in simple coordinate systems. The appropriate generalizations will all be developed in due course here.) Further, we define tensors as objects with arbitrary covariant and contravariant indices which transform in the manner of vectors with each index. For example,

$$T_k^{ij}(q) \equiv \Lambda_m^i(q,x)\, \Lambda_n^j(q,x)\, \underline{\Lambda}_k^l(q,x)\, T_l^{mn}(x)$$

The metric tensor is a special tensor. First, note that distance is indeed invariant:

$$ds^2(q') = g_{kl}(q')\, dq'^k\, dq'^l$$

$$= \left(\frac{\partial q^i}{\partial q'^k}\right) \left(\frac{\partial q^j}{\partial q'^l}\right) g_{ij}(q) \left(\frac{\partial q'^k}{\partial q^s}\right) dq^s \left(\frac{\partial q'^l}{\partial q^t}\right) dq^t$$

$$= g_{ij}(q) \underbrace{\left(\frac{\partial q^i}{\partial q'^k}\right)\left(\frac{\partial q'^k}{\partial q^s}\right)}_{\frac{\partial q^i}{\partial q^s} = \delta_{is}} \underbrace{\left(\frac{\partial q^j}{\partial q'^l}\right)\left(\frac{\partial q'^l}{\partial q^t}\right)}_{\delta_{jt}} dq^s\, dq^t$$

$$= g_{ij}(q)\, dq^i\, dq^j \equiv ds^2(q)$$

There is also a consistent and unique relation between the covariant and contravariant components of a vector. (There is indeed a single 'object' with two representations in each coordinate system.)

$$dq_j \equiv g_{ji}\, dq^i$$

$$dq'_j \equiv g_{ji}(q')\, dq'^i = g_{kl}(q) \left(\frac{\partial q^k}{\partial q'^j}\right) \underbrace{\left(\frac{\partial q^l}{\partial q'^i}\right) \left(\frac{\partial q'^i}{\partial q^p}\right)}_{\Downarrow\ \delta_{lp}} dq^p$$

$$= \left(\frac{\partial q^k}{\partial q'^j}\right) g_{kl}(q)\, dq^l = \left(\frac{\partial q^k}{\partial q'^j}\right) dq_k$$

Thus it transforms properly as a covariant vector.

These results are quite general; summing on an index (contraction) produces a new object which is a tensor of lower rank (fewer indices).

$$T^{ij}_k\, G^k_l = R^{ij}_l$$

The use of the metric tensor to convert contravariant to covariant indices can be generalized to 'raise' and 'lower' indices in all cases. Since $g_{ij} = \delta_{ij}$ in Cartesian coordinates, $dx^i = dx_i$; there is no difference between co- and contra-variant. Hence $g^{ij} = \delta_{ij}$, too, and one can thus define g^{ij} in other coordinates. {More generally, if an arbitrary measure and metric have been defined, the components of the contravariant metric tensor may be found by inverting the [N(N+1)/2] equations (symmetric g) of $g^{ij}(0)\, g_{ik}(0)\, g_{nj}(0) = g_{kn}(0)$. The matrices are inverses.}

$$A^i(q) \equiv g^{ij}(q)\, A_j(q)$$

$$g^i_j = g^{ik} g_{kj} = \left(\frac{\partial q^i}{\partial x^m}\right) \underbrace{\left(\frac{\partial q^k}{\partial x^n}\right) \delta_{mn} \left(\frac{\partial x^r}{\partial q^k}\right)}_{\Downarrow\ \delta_{rn}} \left(\frac{\partial x^s}{\partial q^j}\right) \delta_{rs}$$

$$= \left(\frac{\partial q^i}{\partial x^s}\right) \left(\frac{\partial x^s}{\partial q^j}\right) = \delta_{ij} = \delta^i_j$$

Thus g^i_j is a unique tensor which is the same in all coordinates, and the Kroneker delta is sometimes written as δ^i_j to indicate that it can indeed be regarded as a tensor itself.

Contraction of a pair of vectors leaves a tensor of rank 0, an invariant. Such a scalar invariant is indeed the same in all coordinates:

$$A^i(q')B_i(q') = \left(\frac{\partial q'^i}{\partial q^j}\right) A^j(q) \left(\frac{\partial q^k}{\partial q'^i}\right) B_k(q) = \delta_{jk}\, A^j(q)\, B_k(q)$$
$$= A^j(q)\, B_j(q)$$

It is therefore a suitable definition and generalization of the dot or scalar product of vectors.

Unfortunately, many of the other operations of vector calculus are not so easily generalized. The usual definitions and implementations have been developed for much less arbitrary coordinate systems than the general ones allowed here.

For example, consider the gradient of a scalar. One can define the (covariant) derivative of a scalar as

$$Ø(x)_{,i} \equiv \frac{\partial Ø}{\partial x^i} \qquad Ø(q)_{,i} \equiv \frac{\partial Ø}{\partial q^i} = \left(\frac{\partial Ø}{\partial x^j}\right)\left(\frac{\partial x^j}{\partial q^i}\right)$$

The (covariant) derivative thus defined does indeed transform as a covariant vector. The comma notation is a conventional shorthand. {However, it does not provide a direct generalization of the gradient operator. The gradient has special properties as a directional derivative which presuppose orthogonal coordinates and use a measure of physical length along each (perpendicular) direction. We shall return later to treat the restricted case of orthogonal coordinates and provide specialized results for such systems. All the usual formulas for generalized curvilinear coordinates are easily recovered in this limit.} A (covariant) derivative may be defined more generally in tensor calculus; the comma notation is employed to indicate such an operator, which adds an index to the object operated upon, but the operation is more complicated than simple differentiation if the object is not a scalar. We shall not treat the more general object in this section, but we shall examine a few special cases below.

II. Three Dimensional Spaces

For many physical applications, measures of area and volume are required, not only the basic measure of distance or length introduced above. Much of conventional vector calculus is concerned with such matters. Although it is quite possible to develop these notions generally for an N-dimensional space, it is much easier and quite sufficient to restrict ourselves to three dimensions. The appropriate generalizations are straightforward, fairly easy to perceive, and readily found in mathematics texts, but rather cumbersome to treat.

For writing compact expressions for determinants and various other quantities, we introduce the permutation symbol, which in three dimensions is

$$\underline{e}^{ijk} = \begin{cases} 1 & \text{for } i,j,k=1,2,3 \text{ or an even permutation thereof, i.e. } 2,3,1 \text{ or } 3,1,2 \\ -1 & \text{for } i,j,k= \text{ an odd permutation, i.e. } 1,3,2 \text{ or } 2,1,3 \text{ or } 3,2,1 \\ 0 & \text{otherwise, i.e. there is a repeated index: } 1,1,3 \text{ etc.} \end{cases}$$

The determinant of a 3x3 matrix can be written as

$$|a| = \underline{e}^{ijk}\, a_{1i}\, a_{2j}\, a_{3k}$$

Another useful relation for permutation symbols is

$$\underline{e}^{ijk}\, \underline{e}^{ilm} = \delta_{jl}\,\delta_{km} - \delta_{jm}\,\delta_{kl}$$

Furthermore,

$$\delta^{ijk}_{lmn} = \underline{e}^{ijk}\,\underline{e}^{lmn} \quad \text{and} \quad \delta^{ijk}_{ijk} = 3!$$

where δ^{ijk}_{lmn} is a multidimensional form of the Kroneker delta which is 0 except when ijk and lmn are each distinct triplets. Then it is +1 if lmn is an even permutation of ijk, -1 if it is an odd permutation. These symbols and conventions may seem awkward at first, but after some practice they become extremely useful tools for manipulations. Fairly complicated vector identities and rearrangements, as one often encounters in electromagnetism texts, are made comparatively simple.

Although the permutation symbol is not a tensor, two related objects are:

$$\varepsilon_{ijk} = \sqrt{g}\;\underline{e}^{ijk} \quad \text{and} \quad \varepsilon^{ijk} = \frac{1}{\sqrt{g}}\,\underline{e}^{ijk} \qquad \text{where} \quad g \equiv |\, g_{ij} \,|$$

with absolute value understood if the determinant in negative. This surprising result may be confirmed by noting that the expression for the determinant given above may also be written as

$$\underline{e}^{lmn}\,|a| = \underline{e}^{ijk}\ a_{li}\ a_{mj}\ a_{nk}$$

which is certainly true for l,m,n=1,2,3, and a little thought will show it to be true in all cases. The transformation law for g may then be obtained as

$$\underline{e}^{lmn}\,|g(q')| = \underline{e}^{ijk}\ g(q')_{li}\ g(q')_{mj}\ g(q')_{nk}$$

$$= \underline{e}^{ijk}\left(\frac{\partial q^p}{\partial q'^l}\right)\left(\frac{\partial q^q}{\partial q'^i}\right)\left(\frac{\partial q^r}{\partial q'^m}\right)\left(\frac{\partial q^s}{\partial q'^j}\right)\left(\frac{\partial q^u}{\partial q'^n}\right)\left(\frac{\partial q^v}{\partial q'^k}\right)$$

$$g_{pq}(q)\,g_{rs}(q)\,g_{uv}(q) \qquad \text{(tensor transform of metric tensors)}$$

$$= \underline{e}^{qsv}\left|\frac{\partial q}{\partial q'}\right|\left(\frac{\partial q^p}{\partial q'^l}\right)\left(\frac{\partial q^r}{\partial q'^m}\right)\left(\frac{\partial q^u}{\partial q'^n}\right) g_{pq}\,g_{rs}\,g_{uv}$$

(considering the terms with indices i,j,k)

$$= \underline{e}^{pru}\ g(q)\left|\frac{\partial q}{\partial q'}\right|(\partial q^p/\partial q'^l)(\partial q^r/\partial q'^m)(\partial q^u/\partial q'^n)$$

(considering the terms with indices q,s,v in constituting a determinant as above)

$$= \underline{e}^{lmn}\ g(q)\left|\frac{\partial q}{\partial q'}\right|^2 \qquad \text{(forming another determinant as above)}$$

thus establishing that g transforms with the square of the Jacobian determinant. For the putatively covariant form of the permutation tensor,

$$\varepsilon_{ijk}(q') = \sqrt{g(q)}\ \ \underline{e}^{rst}\left(\frac{\partial q^r}{\partial q'^i}\right)\left(\frac{\partial q^s}{\partial q'^j}\right)\left(\frac{\partial q^t}{\partial q'^k}\right)$$

$$= \underline{e}^{ijk}\left|\frac{\partial q}{\partial q'}\right|\sqrt{g(q)}\ = \underline{e}^{ijk}\sqrt{g(q')}, \text{ the form desired.}$$

Raising indices in the usual way will produce the contravariant form by arguments similar to those applied above.

The permutation tensors enable one to construct true vectors analogous to the familiar ones. The vector or cross product becomes

$$A_i = \varepsilon_{ijk}\ B^j\,C^k$$

although again we have both co and contravariant forms.

The invariant measure of volume is easily constructed as

$$\Delta V = \varepsilon_{ijk} \frac{dq^i \, dq^j \, dq^k}{(3!)}$$

which is explicitly an invariant by construction and can be identified as volume in Cartesian coordinates. (This is a general method of argument in tensor calculus. If a result is stated as an equation between tensors [or vectors or scalars], if it can be proven or interpreted in any coordinate system, it is true for all. That is the power of tensor calculus and its general properties of transformation between coordinates.)

Note that the application of this relation for ΔV in terms of dq^i and transforming directly from Cartesian dx^i gives immediately the familiar relation

$$\Delta V = J \; dq^1 dq^2 dq^3 \qquad J = \left| \frac{\partial x}{\partial q} \right| \text{ the Jacobian.}$$

For the volume integrals of interest, note that $\int I \; \varepsilon_{ijk} \; dq^i \, dq^j \, dq^k$, for I invariant, is invariant, but $\int T^v \; \varepsilon_{ijk} \; dq^i \, dq^j \, dq^k$ is <u>not</u> a vector, because the transformation law for T^v in general changes over the volume.

The operators of divergence and curl require more care. Just as the gradient has a direct physical significance, these operators are constructed to satisfy certain Green's theorems, Gauss' and Stokes law. These must be preserved if their utility is to continue. One can prove a beautiful general theorem in spaces of arbitrary dimension, from which all common vector theorems are simple corollaries, but the proof requires extensive formal preparation. Instead, we shall provide straightforward, if lengthy, proofs of the two specific results desired.

For Gauss' law, we require a relation which is a proper equation between invariants and further reduces to the usual result in Cartesian coordinates,

$$\int \mathrm{div}(T^m) \; \varepsilon_{ijk} \frac{dq^i \, dq^j \, dq^k}{3\,!} = \int T^i \; dS_i$$

the choice $dS_i = \varepsilon_{ijk} \; dq^j \, dq^k$ is explicitly a (covariant) vector, making the right integral invariant, and it gives the correct result in Cartesians. On the left, we require a suitable operator. We shall next prove that

$$\frac{1}{\sqrt{g}} \; \frac{\partial\left[\left(\sqrt{g}\;\right) T^i\right]}{\partial q_i}$$

is such an invariant. It certainly gives the usual Cartesian divergence, but the inspiration for this guess must remain obscure, for it is deep in the development of general covariant differentiation and Christofel symbols. Fortunately, that need not concern us. Proof that this expression is indeed invariant requires proving that the form is the same in any two systems:

$$\frac{1}{\sqrt{g'}} \; \frac{\partial\left[\left(\sqrt{g'}\;\right) T'^i\right]}{\partial q_i} = \frac{1}{\sqrt{g'}} \; \frac{\partial \left[\left(\sqrt{g'}\;\right) T^k \left(\frac{\partial q^i}{\partial x^k} \right)\right]}{\partial q_i} =$$

$$\frac{1}{\sqrt{g}} \; \frac{\partial\left[\left(\sqrt{g}\;\right) T^i\right]}{\partial x_i} \qquad \text{where } \sqrt{g'} = \left|\frac{\partial x}{\partial q}\right| \sqrt{g} \equiv J \sqrt{g}$$

as shown above, introducing J for the Jacobian determinant. The expression in the new coordinates can then be written

$$\frac{1}{J\sqrt{g}}\left\{\frac{\partial\left[\left(\sqrt{g}\ \right)T^k\right]}{\partial q_i}\right\} J\left(\frac{\partial q^i}{\partial x^k}\right) + \frac{T^k}{J}\left[\frac{\partial\left(J\ \frac{\partial q^i}{\partial x^k}\right)}{\partial q_i}\right]$$

where the first term is simply the desired expression in x^i by the chain rule, and we must show that the second term, the portion in brackets [], is then zero. That term may be written

$$\frac{\partial\left|\frac{\partial x}{\partial q}\right|}{\partial\ q^i}\left(\frac{\partial q^i}{\partial x^k}\right) + J\,\frac{\partial^2\ q^i}{\partial x^k\,\partial x^l}\left(\frac{\partial x^l}{\partial q^i}\right)$$

and the first term converted using $J' = \left|\frac{\partial q}{\partial x}\right| = 1/J$ to $\frac{\partial\ J'^{-1}}{\partial x^k} =$

$$-J^2\ \frac{\partial\left|\frac{\partial q}{\partial x}\right|}{\partial x^k} = -J^2\ \left|\frac{\partial q}{\partial x}\right|\ \frac{\partial^2\ q^i}{\partial x^k\,\partial x^l}\left(\frac{\partial x^l}{\partial q^i}\right)$$

thereby canceling the second term and proving the assertion. The last step requires some algebra to confirm, but it is straightforward using the methods used above for writing a determinant, considering all the terms present, and inserting a

$$\delta^i_j = \left(\frac{\partial q^i}{\partial x^s}\right)\left(\frac{\partial x^s}{\partial q^j}\right)$$

'

(with appropriate choice of indices), the inverse of the usual procedure. The 'tensorial' form of the divergence theorem is therefore an equality of invariants:

$$\int \frac{1}{\sqrt{g}}\ \frac{\partial\left[\left(\sqrt{g}\ \right)T^m\right]}{\partial q_m}\ \varepsilon_{ijk}\ \frac{dq^i\ dq^j\ dq^k}{3!} = \int T^i\,\varepsilon_{ijk}\ dq^j\,dq^k$$

Furthermore, the familiar result, div(Ø**A**) = Ødiv(**A**) + ∇Ø·**A** , remains as

$$\mathrm{div}(\text{Ø}\mathbf{A}) = \text{Ø}\,\mathrm{div}(\mathbf{A}) + \frac{\partial \text{Ø}}{\partial q_i}\ A^i$$

Fortunately, Stokes theorem is somewhat easier; there is only one subtlety. The naive generalization is

$$\int \varepsilon^{ijk}\ \frac{\partial T_k}{\partial q_j}\ \varepsilon_{ist}\,dq^s\,dq^t = \int T_i\ dq^i$$

which again obviously reduces to the usual result in Cartesian coordinates and would be explicitly a good 'tensor' equation between invariants if $\partial T_k/\partial q_j$ were indeed a covariant tensor of rank two. It is not, but the portion used in the equation above is. In general,

$$R_{ij} = \frac{R_{ij} + R_{ji}}{2} + \frac{R_{ij} - R_{ji}}{2}$$

the sum of a symmetric and antisymmetric part. For contractions with the anti-symmetric permutation symbol as used above, only the anti-symmetric part contributes; replacing

$$\frac{\partial T_k}{\partial q_j} = \frac{\left(\frac{\partial T_k}{\partial q_j} - \frac{\partial T_j}{\partial q_k}\right)}{2}$$

is equivalent and gives the identical Cartesian reduction. The antisymmetric expression is easily shown to be a tensor as follows:

$$R_{ij} \equiv \frac{\partial T_i}{\partial x_j} - \frac{\partial T_j}{\partial x_i} \quad \text{and} \quad R'_{ij} \equiv \frac{\partial T'_i}{\partial q_j} - \frac{\partial T'_j}{\partial q_i}$$

but by the laws of tensor transformation, this should also be

$$R'_{ij} = \frac{\partial\left[T_k(\frac{\partial x^k}{\partial q_i})\right]}{\partial q_j} - \frac{\partial\left[T_k(\frac{\partial x^k}{\partial q_j})\right]}{\partial q_i} = R_{kl}\,(\frac{\partial x^k}{\partial q_i})\,(\frac{\partial x^k}{\partial q_j})$$

$$= (\frac{\partial T_k}{\partial q_j})(\frac{\partial x^k}{\partial q_i}) - (\frac{\partial T_k}{\partial q_i})(\frac{\partial x^k}{\partial q_j}) + T_k\,\frac{\partial^2 x^k}{\partial q_j \partial q_i} - T_k\,\frac{\partial^2 x^k}{\partial q_i \partial q_j}$$

where the last two terms cancel and the first two, using the chain rule $(\partial/\partial q_i)=(\partial/\partial x_k)(\partial x^k/\partial q_i)$, give the required tensor transform of R_{ij}. We therefore have the desired tensor form of the divergence and curl operators and the corresponding integral theorems. Note also that the important results curl (grad Ø) = 0 and div (curl A_i) = 0 both follow easily from these forms by symmetry

$$\underline{e}^{ijk}\ \frac{\partial^2}{\partial q_i \partial q_j} = 0.$$

III. Physical Vectors

The distinction between covariant and contravariant vectors is essential to tensor analysis, but it is a complication which is unnecessary for elementary vector calculus. In fact, the usual formulation of vector calculus can be obtained from tensor calculus as a special case, that being one in which the coordinate system is orthogonal. Most practical coordinate systems are of this type, for which tensor analysis is not really necessary, but a few are not. (For example, in plasma physics, the natural coordinates may be ones determined by the magnetic geometry and not be orthogonal.) In orthogonal systems with positive metric, one can define 'physical' vectors, which are neither covariant nor contravariant. Nevertheless, they have well-defined transformation properties among orthogonal systems, and they have simple physical significance. For example, all components of a displacement vector have the dimensions of length. They are the vectors of traditional vector calculus. For orthogonal systems of this type,

$$g_{ij} = h^2_i\ \delta_{ij} \qquad \text{(h_i is not a vector; no summation)}$$

$$A(i) \equiv h_i A^i = \frac{A_i}{h_i} \text{ (no summation)}$$

for the components of the 'physical' vector. The usual dot or scalar product is simply A(i)A(i) and produces the same result as given above. (In this special case, the metric tensor can be 'put into' the vector in a natural manner.)

All the usual vector formulas can be obtained from the preceding tensor expressions by consistently converting to physical vectors. Note that $g = (h_1h_2h_3)^2$ and $\varepsilon_{ijk} = h\ \underline{e}^{ijk}$, using $h = (h_1h_2h_3)$.

$$C(i) = A(j)\ \mathbf{X}\ B(k) = \underline{e}^{ijk}\ A(j)\,B(k)$$

$$(\text{grad } \emptyset)(i) = (1/h_i)(\partial \emptyset/\partial q_i)$$

$$\text{div } A = (1/h)\{\partial[hA(i)/h_i]/\partial q_i\}$$

$$(\text{curl } A)(i) = (h_i/h)\ \underline{e}^{ijk}\ \partial[h_k A(k)]/\partial q_j$$

$$\text{Volume: } (d^3v) = h\ \underline{e}^{ijk}\ dq_i dq_j dq_k = d^3l = \underline{e}^{ijk}\ dl_i dl_j dl_k$$

Integrations are over physical volumes, areas, and lengths. If the integrals are set up in coordinates like dq, the necessary factors must be inserted to give the physical units as illustrated here for volume.

IV. Examples

<u>Cylindrical coordinates</u>

A simple example to illustrate the ideas is provided by cylindrical coordinates:

$$x = r\cos\theta \qquad r = \sqrt{x^2 + y^2}$$
$$y = r\sin\theta \qquad \theta = \tan^{-1}(y/x)$$
$$z = z$$

$$\Lambda^i_j \equiv \frac{\partial q^i}{\partial x^j} = \begin{array}{c} i \backslash j = \quad 1 \qquad\qquad 2 \qquad\quad 3 \\ \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -(\sin\theta)/r & (\cos\theta)/r & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

$$\underline{\Lambda}^j_i \equiv \frac{\partial x^j}{\partial q^i} = \begin{array}{c} i \backslash j \\ \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -r(\sin\theta) & r(\cos\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad\qquad g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$g = r^2 \qquad\qquad h_i = (1,r,1)$$

Spherical Coordinates

A second example of broad utility is spherical coordinates:

$$x = r \sin\theta \cos\phi \qquad r = \sqrt{x^2 + y^2 + z^2}$$

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

$$z = r \cos\theta \qquad \phi = \tan^{-1}(y/x)$$

$$\Lambda^i_j \equiv \frac{\partial q^i}{\partial x^j} = \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ (\cos\theta\cos\phi)/r & (\cos\theta\sin\phi)/r & -(\sin\theta)/r \\ \dfrac{-\sin\phi}{r\sin\theta} & \dfrac{\cos\phi}{r\sin\theta} & 0 \end{pmatrix}$$

$$\underline{\Lambda}^j_i \equiv \frac{\partial x^j}{\partial q^i} = \begin{array}{l} i \backslash j \\ \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ r\cos\theta\cos\phi & r\cos\theta\sin\phi & -r\sin\theta \\ -r\sin\theta\sin\phi & r\sin\theta\cos\phi & 0 \end{pmatrix} \end{array}$$

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix} \qquad g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & r^{-2}\sin^{-2}\theta \end{pmatrix}$$

$$g = r^4 \sin^2\theta \qquad h_i = (1, r, r\sin\theta) \qquad h = r^2 \sin\theta$$

V. Application: Special Relativity

Special relativity is generally introduced without tensor calculus, but the results often seem rather *ad hoc*. Einstein used the ideas of tensor calculus to develop the theory, and it certainly assumes its most natural and elegant formulation using tensors. The arguments are easily stated. The use of tensors is natural, for it guarantees that if the laws of physics are properly formulated as equations between scalars, vectors, or tensors, a result or equality in one coordinate system will be true in any.

Special relativity is based on only two postulates. The first is that all coordinate systems moving uniformly with respect to one another are equivalent, i.e. indistinguishable from one another. The second is that the speed of light is constant in all such systems. (The first was a long-standing principle. The second was the implication of the Michelson-Morley experiment.) These are easily phrased in tensor calculus. The first implies that <u>the metric tensor must be the same in all equivalent systems</u>, otherwise the differences would provide a basis for distinguishing among them. The second is achieved by introducing a space of four dimensions with Cartesian coordinates (x,y,z,ct) and choosing the metric tensor to be

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

[This is one of many equivalent choices, none of which has become standard. Sometimes the time is placed first, the indices may run from 0-3 instead of 1-4, and the factors of c can be put into g instead of into the coordinates.]

The resulting invariant measure "length" is $d^2\sigma = g_{\mu\nu}\ dx^\mu dx^\nu = -\ d^2s + c^2 d^2t$, introducing the usual convention that Greek indices range 1-4, whereas Latin indices range only over 1-3, the spatial dimensions: $d^2s = dx^i dx^j$; $x^\mu = (x,y,z,ct) = (x^i, ct)$. It is this measure of "length", sometimes called 'proper distance', no better a choice of words, which makes c a unique constant. (You may be more familiar with this invariant called 'proper time' $d\tau = d\sigma/c$.) Specifically, a disturbance propagating at c in one system (ds/dt=c in that system) will produce events in that system for which $d^2\sigma = 0$. Since this "length" is invariant, it will be the same in all systems: $d^2\sigma = 0$ for the events transformed to any other system, and they will thus also appear to move at ds'/dt'=c. For all equivalent uniformly moving systems, which have the metric above, a speed of c will be invariant. (This argument is carefully phrased to avoid "the speed of light", although "the speed of light in vacuum" would suffice. If light is observed in a medium, which is difficult to avoid, the medium introduces a preferred reference frame and the speed is no longer strictly invariant.)

It remains only to obtain the transformation law between uniformly moving coordinate systems which will preserve the metric. Let the origins coincide at t=0 and the origin of one system (0,ct) move with velocity v in the other along x. If one looks for the simplest (covariant) transform which could accomplish this

$$\underline{\Lambda}^{\alpha}_{\ \mu} = \begin{pmatrix} A & 0 & 0 & B \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ C & 0 & 0 & D \end{pmatrix} \qquad g'_{\mu\nu} = \underline{\Lambda}^{\alpha}_{\ \mu}\ \underline{\Lambda}^{\beta}_{\ \nu}\ g_{\alpha\beta}$$

$$g'_{\mu\nu} = \begin{pmatrix} B^2 - A^2 & 0 & 0 & BD - AC \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ BD - AC & 0 & 0 & D^2 - C^2 \end{pmatrix}$$

where one must be careful if one does the tensor contraction as matrix multiplication; transposes must sometimes be used to obtain the proper index matching. The requirements are thus

$$AC = BD \qquad B^2 - A^2 = -1 \qquad D^2 - C^2 = 1$$

$(0,ct) \rightarrow (-\ Bct,0,0,Dct) \Rightarrow B/D = \mathbf{v/c} \equiv \beta$, where the signs come from using covariant displacements to employ the transform law above, but one is not concerned about the sign of v. Note that co and contravariant vectors differ, but only in sign of the spatial part.) The unique solution to these four equations in four unknowns is

$$\begin{pmatrix} \gamma & 0 & 0 & \beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \beta\gamma & 0 & 0 & \gamma \end{pmatrix} = \underline{\Lambda}^{\alpha}{}_{\mu}$$

$$\begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} = \Lambda^{\alpha}{}_{\mu}$$

$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}}$$

which give the rules for transforming tensors between uniformly moving systems.

(Note that the metric is not positive definite here. The notion of physical vectors introduced in Section III cannot be employed to disguise a difference between co and contravariant. An attempt to do so introduces $\sqrt{(-1)}$, the origin of the ubiquitous i's which permeate non-tensor treatments of special relativity. It is ironic that the attempt to "hide" the metric by introducing "physical" vectors should result in the rather unphysical appearance of imaginary dimensions.)

Because the metric does not depend upon position, we have the useful generalization, already employed above, that not only is the displacement, dx^μ, a contravariant vector, as it always must be, but **the coordinates or vector position of a point, x^μ, is also a vector**, which is not true in general and constitutes a major conceptual subtlety in tensor calculus. This is a great simplification for special relativity, and it means that the law above for transformation of contravariant vectors is **<u>also</u>** the law for coordinate transformations.

Finally, note that $g^{\mu\nu} = g_{\mu\nu}$, which can be confirmed by direct calculation. (As noted earlier, the two must be matrix inverses of one another.)

All the usual relativistic effects follow in a straightforward manner from these equations. An event at x_o, ct_o occurs at $\gamma(x_o - \beta ct_o)$, $\gamma(ct_o - \beta x_o)$ in the moving system. The origin of the initial coordinates appears to be moving at -v in the new system, whereas the origin in the new system appears to be moving at v in the initial system. Events at the point x_o but separated by Δt_o occur at different points and different times, the time difference being $\gamma\,\Delta t_o$, the well-known time dilation. A stationary bar with ends $0,ct_o$ and L,ct_1 appears at

$$-\beta\gamma ct_o\,,\ \gamma\, ct_o \qquad\qquad \text{and} \qquad \gamma\,(L-\beta ct_1),\gamma(ct_1-\beta L)$$

Expressed in terms of a new t', $t' = \gamma\, t_o$ and $t' = \gamma(t_1 - \beta L/c)$

$$-\beta ct'\,,\ ct' \qquad\qquad \text{and} \qquad\qquad (L/\gamma)\ -\beta ct',\ ct'$$

which implies that the ends appear separated by a distance L/γ, the contraction of length, if they are observed (measured) simultaneously in the new system. The velocity addition formula follows simply by applying two successive transformations:

$$\begin{pmatrix} (1+\beta\beta')\gamma\gamma' & 0 & 0 & -(\beta+\beta')\gamma\gamma' \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -(\beta+\beta')\gamma\gamma' & 0 & 0 & (1+\beta\beta')\gamma\gamma' \end{pmatrix} = \begin{pmatrix} \gamma'' & 0 & 0 & -\beta''\gamma'' \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta''\gamma'' & 0 & 0 & \gamma'' \end{pmatrix}$$

$$\beta'' \equiv \frac{\beta+\beta'}{1+\beta\beta'} \qquad\qquad \gamma'' = (1 + \beta\beta')\gamma\gamma' = \gamma''(\beta'')$$

but note that the addition of two velocities in <u>different</u> directions gives much more complicated results; the transformations do not even commute.

If the physical laws are expressed in terms of relativistic vectors and tensors, they will transform properly with coordinate system and have the same form in any system, as desired. The analog of velocity is

$$v^\mu \equiv dx^\mu/d\tau \qquad d\sigma = c\,d\tau$$

$$v^\mu = (\gamma v^i, \gamma c) \qquad\qquad p^\mu = m v^\mu = (p^i, E/c)$$

These relations for the four-velocity follow directly if $x^i = v^i t$, $d^2x^i = v^2 d^2 t$

$$d^2\tau = d^2t - d^2x^i/c^2 = (1-\beta^2)d^2t = d^2t/\gamma^2$$

This is a well-formed vector which reduces to the usual velocity for v << c; it is the only useful relativistic expression for velocity, and thus momentum. The fourth component of the momentum vector is identified as E because it becomes $mc^2 + (1/2)mv^2$ = K.E. + constant in the usual limit. Because of the tensor transformation law, if $p_1{}^\mu = p_2{}^\mu$ in one system, $p'_1{}^\mu = p'_2{}^\mu$ in any other, and only momentum defined in this way will be conserved in all systems if it is conserved any system. Because the conserved momentum is that given by these expression, the relativistic equations are often described as giving a mass increase γm, because $p^i = \gamma m v^i$. The generalization of energy is $E = \gamma mc^2$. (Since only the rest mass ever appears, we shall omit m_o and keep all factors of γ explicit.)

The equations of mechanics are

$$f^\mu \equiv \frac{dp^\mu}{d\tau} = \gamma \frac{dp^\mu}{dt} = \gamma(F^i, P/c) \qquad F^i = dp^i/dt \text{ (Newtonian force)} \qquad P = \frac{dE}{dt}$$

$$a^\mu \equiv \frac{dv^\mu}{d\tau} = \gamma \frac{dv^\mu}{dt}$$

<u>Example: 'Uniform Acceleration'</u>

To illustrate the use of these equations, consider a particle subjected to a constant force, e.g. an electron in a constant electric field, starting from rest. The spatial part of Newton's law, canceling γ's, is simply $\frac{dp^i}{dt} = F^i = \frac{d(\gamma m v^i)}{dt}$. For motion in one dimension, $\alpha_o \equiv \frac{F}{mc}$, and $\frac{d}{dt}\left(\frac{\beta}{\sqrt{1-\beta^2}}\right) = \alpha_o$.

This may be integrated directly to give $\beta(t) = \frac{\alpha_o t}{\sqrt{1 + \alpha_o{}^2 t^2}}$, which has the necessary v=at behavior at small t and $\beta \sim 1$ at large t, and $\gamma(t) = \sqrt{1 + \alpha_o{}^2 t^2}$. This is a solution for the motion in a fixed reference frame in which the particle was originally at rest. From the view of the particle, things are more complicated, for the particle does not define an inertial frame. At best, one can consider a succession of inertial frames in which the particle is instantaneously at rest. From the solution, $v^\mu = \gamma(v,0,0,c)$ and $a^\mu = \gamma c(\alpha_o, 0, 0, \frac{d\gamma}{dt}) = \gamma c(\alpha_o, 0, 0, \beta\alpha_o)$, the (contravariant) vector transform to the particle 'rest' frame (at v) gives $v'^\mu = (0,0,0,c)$, as it should, and $a'^\mu = c(\alpha_o,0,0,0)$. The constant force in the laboratory frame implies a constant acceleration in the instantaneous rest frame; the power, $\frac{dE}{dt}$, is always zero in that frame because $\mathbf{F}\cdot\mathbf{v} = 0$ there.

Another very useful four-vector is the wave vector $k_\mu = (k^i, -\omega/c)$, such that $k_\mu x^\mu$ is an invariant, $\mathbf{k}\cdot\mathbf{r} - \omega t$, the phase of a wave. (The formal argument is just the reverse: The phase of a wave must be an invariant--all observers can identify a peak. Since $k_\mu x^\mu$ is the phase, $k_\mu x^\mu$ must be an invariant, and hence k_μ must transform as a [covariant] vector.) Transforming this as a four-vector easily gives the Doppler shift of frequency and the change in wavelength in a new system, accurate for all values of v.

Maxwell's equations and the equations of electromagnetism are comparatively straightforward in four-vector form. The current vector is

$$j^\mu \equiv (j^i, \rho c) \qquad \frac{\partial j^\mu}{\partial x^\mu} = \operatorname{div} j + \partial\rho/\partial t = 0,$$

the natural form of a conservation law. [Compare discussion above for case here where $\sqrt{|g|}=1$ and therefore there are no contributions from g to the derivatives. For a constant metric, covariant differentiation reduces to partial differentiation in the sense that $\partial/\partial x^i$ simply adds a well-formed covariant index.] This the unique well-formed tensor equation which guarantees that if charge is conserved in one reference frame, it is conserved in all. Charge conservation means that **charge** is an invariant, e.g. all observers agree on e for the electron, but note that j^μ transforms as a vector and that different observers measure different currents and charge densities.

The potentials also make a natural four-vector,

$$A^\mu \equiv (A^i, \text{Ø}/c) \qquad\qquad A_\mu \equiv (-A^i, \text{Ø}/c)$$

The argument is straightforward: A tensorial differential operator (an invariant) is easily formed as $-g^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}$ which is familiar as the operator of the wave equation, $\nabla^2 - \frac{\partial^2}{c^2\partial t^2}$. The usual equations for the potentials (in the Lorentz gauge) can therefore be expressed as

$$\left(g^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}\right) A^\mu = \mu_o j^\mu \qquad\qquad \frac{\partial A^\mu}{\partial x^\mu} = 0$$

with the choice of A^μ above, and these are proper tensor equations if A^μ is a four-vector. Furthermore,

$$T_{\mu\nu} = \frac{\partial A_\mu}{\partial x_\nu} - \frac{\partial A_\nu}{\partial x_\mu}$$

is, by the arguments above, also a good tensor, whose components are in fact

$$\mu\backslash\nu$$

$$\begin{pmatrix} 0 & B_z & -B_y & E_x/c \\ -B_z & 0 & B_x & E_y/c \\ B_y & -B_x & 0 & E_z/c \\ -E_x/c & -E_y/c & -E_z/c & 0 \end{pmatrix} = T_{\mu\nu}$$

$$\mu\backslash\nu$$

$$\begin{pmatrix} 0 & B_z & -B_y & -E_x/c \\ -B_z & 0 & B_x & -E_y/c \\ B_y & -B_x & 0 & -E_z/c \\ E_x/c & E_y/c & E_z/c & 0 \end{pmatrix} = T^{\mu\nu}$$

$$\frac{\partial T^{\mu\nu}}{\partial x_\nu} = \mu_o\, j^\mu$$

expresses the two Maxwell's equations with sources, Gauss and Ampere's Laws, directly. Since the fields are constructed from the potentials using the usual equations, the other two Maxwell's equations are automatically satisfied, but they can also be expressed as

$$\varepsilon^{\alpha\beta\gamma\delta}\, \frac{\partial T_{\beta\gamma}}{\partial x_\delta} = 0$$

noting that the simple permutation symbols are tensors when $||g||=1$ (absolute value of the determinant of the metric tensor), a simple generalization of the arguments of Section II. One can construct two interesting invariants from the fields as

$$T^{\mu\nu}\ T_{\nu\mu} = |B|^2 - |E/c|^2 \quad \text{and} \quad \varepsilon^{\alpha\beta\gamma\delta}\, T_{\alpha\beta}\, T_{\gamma\delta} = 2\, \mathbf{E\cdot B}$$

These have important physical consequences, implying that if the field is purely electric in one frame, there will be a dominant electric field in all frames, and vise versa. Conversely, if there are both electric and magnetic fields in some frame, it is possible to find a frame in which one vanishes. An important consequence of the second invariant is that if the fields are transverse in one frame (perpendicular to one another), they will be so in all frames.

The Lorentz force expressions may also be constructed:

$$f_\nu = T_{\mu\nu}\, j^\mu = -T_{\nu\mu}\, j^\mu \qquad \text{and} \qquad f_\nu = qT_{\mu\nu}\, v^\mu = -qT_{\nu\mu}\, v^\mu$$

The covariant force <u>density</u> f_ν, appropriate to a continuous system with a current-density, charge-density four-vector j^μ, is to be distinguished from the four-vector force f_ν, which acts on a particle of charge q. These expressions are explicitly formed as invariant (tensor) expressions and may be directly computed to verify that they give the familiar results of electromagnetism (for the contravariant form):

$$f^\mu = (\,\rho E^i + \mathbf{j \times B},\ \mathbf{j\cdot E}/c) \qquad\qquad f^\mu = q\gamma(E^i + \mathbf{v \times B},\ \mathbf{v\cdot E}/c)$$

The four-vector force f^μ, which appears here has the same factor of γ multiplying the familiar terms as did the corresponding four-vector in the tensor form of Newton's law.

These constructions of the tensor equivalents of mechanics and electromagnetism may appear to lack rigor, but that is not the case. If an equation is written as a proper equation among tensors, tensor calculus guarantees that it will remain true in all coordinate systems. Therefore an equation of the proper form which is correct in one coordinate system will be universal. You may find more detailed arguments helpful in understanding relativistic effects, but they are not necessary. For example, to prove that j^μ is a four-vector, it is not necessary to examine current densities and charge densities in one coordinate system and determine their complex transformations as velocities and volumes transform between systems. It suffices to declare that charge conservation is a physical law. Only $\frac{\partial j^\mu}{\partial x^\mu} = 0$ with $j^\mu = (j^i, \rho c)$ being a genuine four-vector is a proper tensor equation which provides the usual form of the charge conservation equation in a reference system. Therefore j^μ must be a four-vector. (It is a symptom of the Lorentz invariance of electromagnetism that the equation of charge conservation indeed has the familiar form in <u>all</u> inertial coordinate systems. However, the tensor equations for mechanics involving f^μ etc. include factors of γ and reduce to the familiar forms only for low velocity, $\gamma \sim 1$.)

The most important application of this argument is to the electromagnetic field, the tensor and transformation character of which would otherwise require considerable, tedious argument. The argument above shows that the fields are thoroughly linked, being components of a single tensor. Since **E** and **B** are conventionally vectors, one might have expected analogous four-vectors, but that would create a conceptual difficulty in expressing a four-vector force coupling four-vector fields and the four-vector velocity, a difficulty which is obviated by the tensor force expressions above. The field tensor transforms normally; for reference, the result is shown here:

$$\begin{array}{l} \mu\backslash\nu \\ \begin{pmatrix} 0 & \gamma(B_z - \beta E_y/c) & -\gamma(B_y + \beta E_z/c) & E_x/c \\ -\gamma(B_z - \beta E_y/c) & 0 & B_x & \gamma(-\beta B_z + E_y/c) \\ \gamma(B_y + \beta E_z/c) & -B_x & 0 & \gamma(\beta B_y + E_z/c) \\ -E_x/c & -\gamma(-\beta B_z + E_y/c) & -\gamma(\beta B_y + E_z/c) & 0 \end{pmatrix} = T_{\mu\nu} \end{array}$$

The familiar **vxB** contribution to the new E is present, but there are factors of γ and contributions to B as well.

The tensor form of energy conservation may be obtained by similar arguments, or it can be obtained as follows using methods analogous to those of the classical argument. (Since momentum-energy conservation already involves tensors, the four-vector analog is not particularly easy to construct.)

$$f_\nu = T_{\mu\nu}\, j^\mu = \frac{1}{\mu_o}\, T_{\mu\nu} \frac{\partial T^{\mu\alpha}}{\partial x_\alpha}$$

$$= \frac{1}{2\mu_o}\, T_{\mu\nu} \frac{\partial T^{\mu\alpha}}{\partial x_\alpha} + \frac{1}{2\mu_o}\left(\frac{\partial (T_{\mu\nu}\, T^{\mu\alpha})}{\partial x_\alpha} - T^{\mu\alpha} \frac{\partial T_{\mu\nu}}{\partial x_\alpha}\right)$$

Since $\dfrac{\partial T_{\mu\nu}}{\partial x_\alpha} + \dfrac{\partial T_{\alpha\mu}}{\partial x_\nu} + \dfrac{\partial T_{\nu\alpha}}{\partial x_\mu} = 0$ (if the indices are distinct, this is one of Maxwell's equations $\varepsilon^{\alpha\beta\gamma\delta} \dfrac{\partial T_{\beta\gamma}}{\partial x_\delta} = 0$, otherwise is it true by the antisymmetry of T),

$$f_\nu = \frac{1}{2\mu_o} T_{\mu\nu} \frac{\partial T^{\mu\alpha}}{\partial x_\alpha} + \frac{1}{2\mu_o}\left\{\frac{\partial (T_{\mu\nu} T^{\mu\alpha})}{\partial x_\alpha} + T^{\mu\alpha}\left(\frac{\partial T_{\alpha\mu}}{\partial x_\nu} + \frac{\partial T_{\nu\alpha}}{\partial x_\mu}\right)\right\}$$

$$= \frac{1}{2\mu_o} T_{\mu\nu} \frac{\partial T^{\mu\alpha}}{\partial x_\alpha} + \frac{1}{2\mu_o}\left\{\frac{\partial (T_{\mu\nu} T^{\mu\alpha})}{\partial x_\alpha} + \frac{1}{2}\frac{\partial (T^{\mu\alpha} T_{\alpha\mu})}{\partial x_\nu} + T^{\mu\alpha}\frac{\partial T_{\nu\alpha}}{\partial x_\mu}\right\}$$

$$= \frac{1}{2\mu_o} T_{\mu\nu} \frac{\partial T^{\mu\alpha}}{\partial x_\alpha} + \frac{1}{2\mu_o}\left\{\frac{\partial (T_{\mu\nu} T^{\mu\alpha})}{\partial x_\alpha} + \frac{1}{2}\frac{\partial (T^{\mu\alpha} T_{\alpha\mu})}{\partial x_\nu} + \frac{\partial (T^{\mu\alpha} T_{\nu\alpha})}{\partial x_\mu} - T_{\nu\alpha}\frac{\partial T^{\mu\alpha}}{\partial x_\mu}\right\}$$

By changing dummy indices and using the antisymmetry of T, the first and last terms cancel, and the second and fourth terms are identical, leaving

$$f_\nu = \frac{1}{\mu_o}\left\{\frac{\partial (T_{\mu\nu} T^{\mu\alpha})}{\partial x_\alpha} - \frac{1}{4}\frac{\partial (T^{\mu\alpha} T_{\mu\alpha})}{\partial x_\nu}\right\} = \frac{\partial G^{\mu}_{\ \nu}}{\partial x_\mu}$$

$$G^{\mu}_{\ \nu} = \frac{1}{\mu_o}\left\{T_{\alpha\nu} T^{\alpha\mu} - \frac{1}{4}(T^{\alpha\beta} T_{\alpha\beta})\,\delta^{\mu}_{\ \nu}\right\}$$

for the relativistic stress tensor. It can be converted to other forms, for example:

$$G^{\mu\nu} = \frac{1}{\mu_o}\left\{g_{\alpha\beta} T^{\beta\nu} T^{\alpha\mu} - \frac{1}{4}(T^{\alpha\beta} T_{\alpha\beta})\, g^{\mu\nu}\right\}$$

which is clearly symmetric, but the elements remain complicated functions of the fields. This completes the fundamental formulation of mechanics and electrodynamics in relativistic form.

VI. Covariant Differentiation

Differentiation of tensors is not simple. The partial derivatives of an invariant form a good (covariant) vector, and certain antisymmetric forms have been shown above to be tensors, but generally speaking, the partial derivatives of vectors (and perforce tensors) introduce derivatives of the transform law and metric. Only for constant g_{ij}, e.g. Cartesian coordinates and special relativity, but not even cylindrical or spherical coordinates, do partial derivatives produce tensors. The formulation of derivatives (i.e. finding definitions for derivatives of a tensor --- absolute and covariant differentiation) which do behave properly is subtle. Several approaches are possible; the one here is 'geometric' rather than formal and strives to provide a basis for and understanding of the complications which arise. Nevertheless, not all steps can be well motivated, and certain choices will become clear only in retrospect.

Since only derivatives of invariants have tensor character, we begin by considering a simple, fundamental object, the tangent to a curve $q^i(u)$:

$$p^i \equiv \frac{dq^i}{du} \qquad (1)$$

This is a well-formed contravariant vector, from which an invariant $w = g_{ij}p^ip^j$ may be constructed. Its derivative must likewise be an invariant

$$\frac{dw}{du} = 2\, g_{ij}p^i \frac{dp^j}{du} + \frac{\partial g_{ij}}{\partial q_k}\, p^ip^jp^k \qquad (2)$$

which can be written in this simple, symmetric form because of the definition of p^i above. A fundamental (and rather obvious) theorem of tensor calculus, sometimes called the quotient rule, implies that if $A_iB^i = Ø$ (an invariant) and B^i is an arbitrary contravariant vector, then A_i must be a covariant vector. One can thus factor out a term p^i from this expression and conclude that the remainder is a good covariant vector. However, i is a dummy index; any of the three p factors in the final product could be extracted. In fact, a particular combination is particularly useful: the sum of the two symmetric forms in ij, minus the form using k:

$$f_i = g_{ij}\frac{dp^j}{du} + [jk,i]\, p^jp^k \qquad [ij,k] \equiv \frac{1}{2}\left(\frac{\partial g_{jk}}{\partial q_i} + \frac{\partial g_{ik}}{\partial q_j} - \frac{\partial g_{ij}}{\partial q_k}\right) \qquad (3)$$

where the bracket defines a famous object, the Christoffel symbol of the first kind. It is clear that this f_i is not the only covariant vector involving $\frac{dp^j}{du}$, but the special symmetry of the Christoffel symbol makes it an advantageous choice. There is an obvious corresponding contravariant vector

$$f^i = \frac{dp^i}{du} + \left\{ {i \atop jk} \right\} p^jp^k \qquad \left\{ {i \atop jk} \right\} \equiv g^{il}\, [jk,l] \qquad (4)$$

which employs the Christoffel symbol of the second kind. These objects are <u>not</u> tensors, their transformation law remaining to be inferred from the known transformation character of the other terms in the equation, but raised and lowered indices are used to indicate the indices with which they are to be summed in the usual convention.

This leads one to define a derivative, the absolute derivative, of p^i as the contravariant vector

$$\frac{\delta p^i}{\delta u} = \frac{dp^i}{du} + \left\{ {i \atop jk} \right\} p^jp^k \qquad (5)$$

The significance of the Christoffel symbols may be understood as follows: If u is chosen to be length s, then a 'straight line' would have a constant tangent $\frac{\delta p^i}{\delta s} = 0$ as an invariant property. In Cartesian coordinates, that is equivalent to $\frac{dp^i}{ds} = 0$, but this definition implies otherwise if the Christoffel symbols are non-zero. In fact, in 'curved' coordinates, even ones as simple as cylindrical or spherical systems, $\frac{dp^i}{ds} \neq 0$ for a straight line, and a <u>constant</u> tangent vector has varying r,θ components along the line in general. The Christoffel symbols embody this curvature and introduce it into the equations, guaranteeing that only the proper $\frac{dp^i}{ds}$ will produce $\frac{\delta p^i}{\delta s} = 0$. (Very similar equations and calculations to those here appear in the rigorous generalization of a straight line, which is a geodesic, a curve of variationally stationary length or simply the 'shortest distance' if the metric is positive definite.)

As mentioned above, the transformation laws for Christoffel symbols may be adduced from the tensorial form of the terms in (3) and (4). Specifically from (4),

$$f'^i = \frac{d^2q^i}{d^2u} + \left\{ {i \atop jk} \right\}' \frac{dq^j}{du}\frac{dq^k}{du} = \frac{\partial q^i}{\partial x^j}\left(\frac{d^2x^j}{d^2u} + \left\{ {j \atop lm} \right\} \frac{dx^l}{du}\frac{dx^m}{du} \right) \qquad (6)$$

$$\frac{\partial q^i}{\partial x^j}\left(\frac{d^2x^j}{d^2u} \right) = \frac{\partial q^i}{\partial x^j}\frac{d}{du}\left(\frac{\partial x^j}{\partial q^l}\frac{dq^l}{du} \right) = \frac{\partial q^i}{\partial x^j}\frac{\partial x^j}{\partial q^l}\frac{d^2q^l}{d^2u} + \frac{\partial q^i}{\partial x^j}\frac{\partial^2 x^j}{\partial q^l \partial q^p}\frac{dq^l}{du}\frac{dq^p}{du}$$

$$= \frac{d^2q^i}{d^2u} + \frac{\partial q^i}{\partial x^t}\frac{\partial^2 x^t}{\partial q^j \partial q^k}\frac{dq^j}{du}\frac{dq^k}{du}$$

The second derivatives of q^i match, leaving

$$\left\{ {i \atop jk} \right\}' \frac{dq^j}{du}\frac{dq^k}{du} = \frac{\partial q^i}{\partial x^t}\frac{\partial^2 x^t}{\partial q^j \partial q^k}\frac{dq^j}{du}\frac{dq^k}{du} + \frac{\partial q^i}{\partial x^t}\left\{ {t \atop lm} \right\}\frac{\partial x^l}{\partial q^j}\frac{\partial x^m}{\partial q^k}\frac{dq^j}{du}\frac{dq^k}{du}$$

$$\left\{ {i \atop jk} \right\}' = \frac{\partial q^i}{\partial x^n}\frac{\partial^2 x^n}{\partial q^j \partial q^k} + \left\{ {n \atop lm} \right\}\frac{\partial x^l}{\partial q^j}\frac{\partial x^m}{\partial q^k}\frac{\partial q^i}{\partial x^n} \qquad (7)$$

This implies that the Christoffel symbols transform like tensors, but with an additional term, which involves the second derivatives of the coordinate transformations. They therefore remain zero for all linear transformations like rotation and the Lorentz group. They are non-zero in cylindrical and spherical coordinates, and the transformation law (7) from Cartesian can be as convenient for calculation as the definition (4). The same procedure may be applied to (3) to give

$$[ij,k]' = g_{lm}\frac{\partial x^l}{\partial q^i}\frac{\partial^2 x^m}{\partial q^j \partial q^k} + [lm,n]\frac{\partial x^l}{\partial q^i}\frac{\partial x^m}{\partial q^j}\frac{\partial x^n}{\partial q^k} \qquad (8)$$

An absolute derivative was defined for the contravariant vector $p^i \equiv \frac{dq^i}{du}$, but the calculation depended on the special properties of p. However, a straightforward generalization is possible, based on the invariant $Ø = g_{ij}p^iT^j$, for any vector field T^j defined along the curve $q^j(u)$:

$$\frac{dØ}{du} = g_{ij}\frac{dp^i}{du}T^j + g_{ij}\,p^i\frac{dT^j}{du} + \frac{\partial g_{ij}}{\partial q_k}\,p^iT^jp^k$$

using (3),

$$= (f_j - [ik,j]\,p^ip^k)T^j + g_{ij}\,p^i\frac{dT^j}{du} + \frac{\partial g_{ij}}{\partial q_k}\,p^iT^jp^k$$

$$\frac{dØ}{du} - f_jT^j = g_{ij}\,p^i\frac{dT^j}{du} + [jk,i]\,p^iT^jp^k \qquad (9)$$

Since the quantity on the left in an invariant, so is the right, and factoring out p^i implies that

$$g_{ij}\frac{dT^j}{du} + [jk,i]\,T^jp^k = f_i$$

a covariant vector. The corresponding contravariant vector is the appropriate absolute derivative:

$$\frac{\delta T^i}{\delta u} \equiv \frac{dT^i}{du} + \left\{ {i \atop jk} \right\} T^j \frac{dq^k}{du} \qquad (10)$$

[The vector character of (10) may also be confirmed by direct transformation using (7) and the procedure used to obtain (7).]

A similar procedure gives the form for the absolute derivative of a covariant vector R_i: An invariant may be formed with any T^i, and an additional derivative invariant likewise as

$$\frac{d(R_i T^i)}{du} = \frac{dR_i}{du} T^i + R_i \frac{dT^i}{du} = \frac{dR_i}{du} T^i + R_i \left(\frac{\delta T^i}{\delta u} - \left\{ {i \atop jk} \right\} T^j \frac{dq^k}{du} \right) \tag{11}$$

Choosing the arbitrary T^i such that $\frac{\delta T^i}{\delta u} = 0$ means that the coefficient of T^i is a vector:

$$\frac{\delta R_i}{\delta u} \equiv \frac{dR_i}{du} - \left\{ {j \atop ik} \right\} R_j \frac{dq^k}{du} \tag{12}$$

By forming an invariant with a collection of arbitrary vectors, each of which has zero absolute derivative, the absolute derivative of any tensor, defined by analogy with the form below, is easily shown to have the same tensor character as the tensor itself:

$$\frac{\delta T^{ij}_k}{\delta u} = \frac{dT^{ij}_k}{du} + \left\{ {i \atop ln} \right\} T^{lj}_k \frac{dq^n}{du} + \left\{ {j \atop ln} \right\} T^{il}_k \frac{dq^n}{du} - \left\{ {l \atop kn} \right\} T^{ij}_l \frac{dq^n}{du} \tag{13}$$

(This one could be proved using $Ø = T^{ij}_k A_i B_j C^k$.)

Mathematicians typically strive for the greatest generality, meaning minimal assumptions. In this case, the vectors and tensors need only be defined along the curve, e.g. $R_i(u)$. However, we are generally concerned with vector and tensor <u>fields</u>, meaning objects which are defined at all points in space. In this case, one can use $\frac{d}{du} = \frac{dq^k}{du} \frac{\partial}{\partial q^k}$, and since $\frac{dq^k}{du}$ is now an arbitrary contravariant vector factor in the absolute derivative, its coefficient must be a tensor and covariant in that index. One thereby defines the covariant derivative

$$T^i{}_{,j} = \frac{dT^i}{dq^j} + \left\{ {i \atop jk} \right\} T^k \tag{14}$$

$$T_{i,j} = \frac{dT_i}{dq^j} - \left\{ {k \atop ij} \right\} T_k \tag{15}$$

with the obvious generalization to tensors of higher rank. (Other common notations are $T^i{}_{;j}$ and $T^i{}_{|j}$.)

These results are susceptible to some helpful and intuitive interpretation. In general, if the derivative of a function is zero, the function is constant in some sense. This idea may be pursued by noting that $g_{ij,k} = 0$. [Verification is straightforward using the extension of (15) for two covariant indices with the definitions (3) and (4) of the Christoffel symbols; it is a further illustration of the advantage of choosing the particular symmetry for f_i in (3).] The sense in which g_{ij}, having zero covariant derivative, is constant is both special and significant. Since g_{ij} is a rather arbitrary symmetric tensor, it certainly varies with position in general, and none of its partial derivatives with respect to q^i need be zero. In fact, its covariant derivative, through the Christoffel symbols, has been implicitly constructed to be zero. The metric tensor defines the space; 'changes' in the metric tensor are changes in the space itself. The tensor derivatives show changes with respect to the space. Almost by definition, the space does not change with respect to itself, and g_{ij} should be a constant with respect to the space defined by g_{ij}.

The concept of 'constancy' may be developed by noting that (11) may be written as

$$\frac{d(R_i T^i)}{du} = R_i \frac{\delta T^i}{\delta u} + \frac{\delta R_i}{\delta u} T^i \tag{16}$$

Applying this to a single vector, if $\frac{\delta T^i}{\delta u} = 0$, then the length of T^i remains constant along u. Furthermore, the angle between two vectors may be defined in the usual sense as $R_i T^i = |R_i||T^i| \cos\theta$, $|T^i| = \sqrt{T_i T^i}$. If both vectors have zero absolute derivative along u, then their lengths and the angle between them remain constant. For this reason, if $\frac{\delta T^i}{\delta u} = 0$, the vector T^i is considered to be propagated parallel to itself along u. Parallelism is easily defined at a point in the usual sense that $R_i T^i = \pm |R_i||T^i|$, but vectors at different points cannot generally be compared. This offers a generalization which preserves most of the usual properties. [Unfortunately, uniqueness is not one of them; different curves u between a pair of points (A, B) may lead to different T^i at B starting from a given T^i at A.] An interpretation of Christoffel symbols can again be given from noting their role in a $\frac{\delta T^i}{\delta u} = 0$ condition as that of driving $\frac{dT^i}{du}$, causing T^i to change to compensate for the 'curvature' of the space.

VII. Geodesics and Lagrangians

As noted above, the concepts of parallelism, straight line, and really all non-local (global) comparisons require some specialization in general metric space. They cannot be carried over with all their familiar properties. A primitive (if 'correct') notion of straight line as $\frac{\delta p^i}{\delta s} = 0$ ($p^i = \frac{dq^i}{ds}$) was introduced in the previous section in interpreting the meaning of absolute differentiation, but a more general formulation is useful. The fundamental formulation is based on a variational principle, and such principles are also important for mechanics.

To review, if a definite integral **I**, whose value is expressed as a functional of functions of a parameter u between fixed end points, is to have an extremum (maximum, minimum, or possibly an inflection point),

$$\delta \mathbf{I} = \delta \left[\int_{u_1}^{u_2} L\left(\frac{dq^i}{du}, q^i(u), u\right) du \right] = 0 \tag{1}$$

By the usual argument in calculus of variations, if the set of functions $q_o^i(u)$ is a solution, then for a small variation about that, $q^i = q_o^i + \delta q^i$, $\delta \mathbf{I}$ must be second order in δq^i, and the first order variation is zero. (This is simply a generalization of the fact that the first derivative of a function is zero at extrema.) If L is then regarded as a function of the functions listed above <u>regarded as independent</u>,

$$\delta \mathbf{I} = \int_{u_1}^{u_2} du \left(\frac{\partial L}{\partial q'^i} \frac{d(\delta q^i)}{du} + \frac{\partial L}{\partial q^i} \delta q^i \right) = 0 \qquad \text{where } q'^i = \frac{dq^i}{du} \tag{2}$$

and a sum over the index i is understood. The first term may be integrated by parts, and if the end points are prescribed so that $\delta q^i(u_1) = \delta q^i(u_2) = 0$, the condition may be written as

$$\int_{u_1}^{u_2} du \left[\frac{d}{du}\left(\frac{\partial L}{\partial q'^i}\right) - \frac{\partial L}{\partial q^i} \right] \delta q^i = 0 \tag{3}$$

since the δq^i are arbitrary, the integral will be zero only if all its coefficients zero, which are the well-known Euler-Lagrange equations for the variational problem.

$$\frac{d}{du}\left(\frac{\partial L}{\partial q'^i}\right) - \frac{\partial L}{\partial q^i} = 0 \tag{4}$$

The application to straight lines arises because a straight line is, among other things, the shortest distance between two points, and this criterion can be formulated in any metric space. A geodesic is defined as a curve for which

$$\delta \mathbf{I} = \delta \left[\int_{u_1}^{u_2} \sqrt{\left| g_{ij} \frac{dq^i}{du} \frac{dq^j}{du} \right|}\, du \right] = 0 \tag{5}$$

and in cases like special relativity for which the metric is not positive definite and there are curves of zero length, the integral may be a maximum. In any case, the solutions $q^i(u)$ are geodesics and the best generalization of a "straight line" in a general metric space. The Euler equations are thus

$$\frac{d}{du}\left(\frac{\partial \sqrt{w}}{\partial q'^i}\right) - \frac{\partial \sqrt{w}}{\partial q^i} = 0 = \frac{d}{du}\left(\frac{1}{\sqrt{w}} \frac{\partial w}{\partial q'^i}\right) - \frac{1}{\sqrt{w}} \frac{\partial w}{\partial q^i} \qquad \text{for} \qquad w = g_{ij} \frac{dq^i}{du} \frac{dq^j}{du}$$

and if u is chosen to be the measure of distance $ds^2 = g_{ij} \frac{dq^i}{du} \frac{dq^j}{du} du^2$, $w = 1$ and $\frac{dw}{ds} = 0$ leaving

$$\frac{d}{ds}\left(\frac{\partial w}{\partial q'^i}\right) - \frac{\partial w}{\partial q^i} = 0 = \frac{d}{ds}\left(2 g_{ij} \frac{dq^j}{ds}\right) - \frac{\partial g_{jk}}{\partial q^i} \frac{dq^j}{ds} \frac{dq^k}{ds} \tag{6}$$

as the equation of the geodesic. Computing the derivative through the q^i dependence and rearranging dummy indices produces

$$2 g_{ij} \frac{d^2 q^j}{ds^2} + \frac{\partial g_{ij}}{\partial q^k} \frac{dq^j}{ds} \frac{dq^k}{ds} + \frac{\partial g_{ik}}{\partial q^j} \frac{dq^j}{ds} \frac{dq^k}{ds} - \frac{\partial g_{jk}}{\partial q^i} \frac{dq^j}{ds} \frac{dq^k}{ds} = 0 \tag{7}$$

which, through no accident, can be written

$$g_{ij} \frac{d^2 q^j}{ds^2} + [jk,i] \frac{dq^j}{ds} \frac{dq^k}{ds} = 0 \quad \text{or} \quad \frac{d^2 q^i}{ds^2} + \left\{ {i \atop jk} \right\} \frac{dq^j}{ds} \frac{dq^k}{ds} = 0 \tag{8}$$

which are the standard forms for these equations.

Variational principles are also used to form the Lagrangian and related equations of motion. The familiar results may be extended to construct relativistically proper forms, but somewhat indirectly. The normal construction of $L = T-V$ with $\int dt$ has no clear tensor equivalent. Instead, we must try to find an invariant L such that $\int L du$ generates the correct equations of motion. For example,

$$L = mc\sqrt{g_{\alpha\beta}\frac{dx^{\alpha}}{du}\frac{dx^{\beta}}{du}} \tag{9}$$

is manifestly invariant and also independent of position (only derivatives enter) and thus a possible starting point as the Lagrangian for a free particle. The Euler equations are

$$mc\frac{d}{du}\left(\frac{g_{\alpha\beta}\frac{dx^{\beta}}{du}}{\sqrt{g_{\alpha\beta}\frac{dx^{\alpha}}{du}\frac{dx^{\beta}}{du}}}\right) = 0 \tag{10}$$

If u is now chosen to be the invariant parameter τ, the radical becomes the invariant constant c, and the equations reduce to the standard equation of motion for a free particle:

$$m\frac{d^2x^{\alpha}}{d\tau^2} = 0 = \frac{dp^{\alpha}}{d\tau} \tag{11}$$

With this start, the Lagrangian for a particle in an electromagnetic field could be

$$L = mc\sqrt{g_{\alpha\beta}\frac{dx^{\alpha}}{du}\frac{dx^{\beta}}{du}} + q\, g_{\alpha\beta}\frac{dx^{\alpha}}{du}A^{\beta} \tag{12}$$

which is again an invariant and linear in q, v^{μ}, and A^{β}; one can argue the second term as the only plausible one. The equations of motion thereby implied are

$$\frac{d}{du}\left(\frac{mc\, g_{\alpha\beta}\frac{dx^{\beta}}{du}}{\sqrt{g_{\alpha\beta}\frac{dx^{\alpha}}{du}\frac{dx^{\beta}}{du}}} + q\, g_{\alpha\beta}A^{\beta}\right) - q\, g_{\mu\beta}\frac{dx^{\mu}}{du}\frac{\partial A^{\beta}}{\partial x^{\alpha}} = 0 \tag{13}$$

and with the same choice of u as τ and extraction of the τ dependence of A^{β} through the x^{μ},

$$g_{\alpha\beta}\, m\frac{d^2x^{\beta}}{d\tau^2} = q\frac{dx^{\mu}}{d\tau}\left(g_{\mu\beta}\frac{\partial A^{\beta}}{\partial x^{\alpha}} - g_{\alpha\beta}\frac{\partial A^{\beta}}{\partial x^{\mu}}\right) = q\frac{dx^{\mu}}{d\tau}\left(\frac{\partial A_{\mu}}{\partial x^{\alpha}} - \frac{\partial A_{\alpha}}{\partial x^{\mu}}\right)$$
$$= q\frac{dx^{\mu}}{d\tau}T_{\mu\alpha} = f_{\alpha} \tag{14}$$

the same equation as obtained previously, thus confirming the choice of L above.

The procedures of classical mechanics may be continued to construct a Hamiltonian from the conjugate momenta

$$P_\alpha = \frac{\partial L}{\partial\left(\frac{dx^\alpha}{du}\right)} = \frac{mc\ g_{\alpha\beta}\ \frac{dx^\beta}{du}}{\sqrt{g_{\alpha\beta}\ \frac{dx^\alpha}{du}\ \frac{dx^\beta}{du}}} + q\ g_{\alpha\beta}\ A^\beta \qquad \text{i.e., } P^\mu = mv^\mu + q\ A^\mu \tag{15}$$

(The partials of L are always taken with respect to a contravariant quantity and generate a covariant index in consistent analogy with the usual tensor derivatives with respect to coordinates, the dq^i being contravariant, although no real tensor character can be ascribed to the partial derivatives associated with the derivation of the Euler-Lagrange equations.)

The Hamiltonian is then

$$H = \frac{1}{2}\left(P_\alpha v^\alpha - L\right) \qquad \text{where} \quad mv^\mu = P^\mu - q\ A^\mu \tag{16}$$

is to be used to eliminate v^μ in favor of P^μ. Straightforward algebra produces the Hamiltonian as

$$H = \frac{g_{\alpha\beta}}{2m}\ (P^\alpha - q\ A^\alpha)\ (P^\beta - q\ A^\beta) - \frac{mc^2}{2} \tag{17}$$

and the Hamiltonian equations of motion:

$$\frac{dx_\alpha}{d\tau} = \frac{\partial H}{\partial P^\alpha} = \frac{g_{\alpha\beta}}{m}\ (P^\beta - q\ A^\beta) \qquad \text{or} \qquad \frac{dx^\alpha}{d\tau} = \frac{P^\alpha - q\ A^\alpha}{m} \tag{18}$$

$$\frac{dP_\mu}{d\tau} = -\frac{\partial H}{\partial x^\mu} = \frac{qg_{\alpha\beta}}{m}\ (P^\alpha - q\ A^\alpha)\ \frac{\partial A^\beta}{\partial x^\mu} \tag{19}$$

The first is the trivial $\frac{dx^\alpha}{d\tau} = v^\mu$, and the second, after elimination of P from (15) on both sides, leaves the same equation of motion as that obtained above from the Lagrangian (14), because $\frac{dA^\mu}{d\tau}$ expands into the remaining portion of the E-M coupling term.
