An Attempt to Integrate the T/g tensor discussion into a Coherent Whole

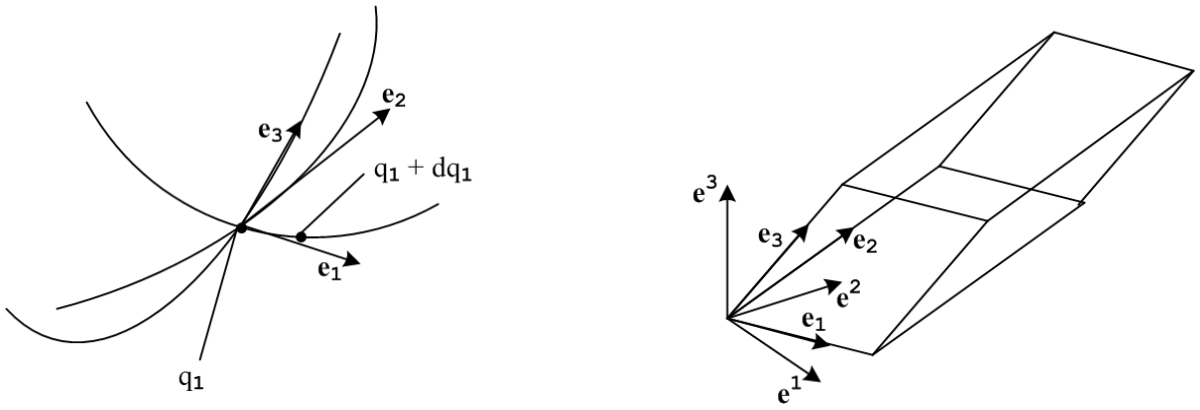
[Read this entire document, added one note, made a few fixes, seems OK. PhL 11.7.09]

Contents:

Review and complaints. 1
The Transformation T and its Tensors: possible existence of a g_{ab} index-lowering tensor 2
Notion of a distance metric; def. of the metric tensor g_{ij}; its use as the index-lowering tensor. 4
Special Case #1: Transformation from Cartesian x to general x' = q coordinates 5
Special Case #2: Transformation of Special Relativity 6
Theorem. 6
Generators $M^{\mu\nu}$ and a boost example. 6

Review and complaints.

The subject of "tensors" and "metric tensor" needs some serious "integration". I had a major problem with the M&M "order of presentation", I suspect multiple authors were involved. M&M start with tensors in Chapter 4 section 4.20, then we have to jump far ahead to Chapter 5 section 5.16 to learn about the *metric* tensor and about basis vectors and relationships in the space x' = q which you "arrive at" by a transformation from Cartesian space. But the fact that this q-space is related to the original x-space by a transformation T is not even mentioned. We treat q-space completely on its own and the subject is suddenly called "curvilinear coordinates". A horrible omission is that there is no picture of the volume element in q-space with all the basis vectors, so I had to draw my own and clarify things in my divergence etc document, which drawing I copy here now:



I have to admit that, at first, I saw no use for all 6 of these basis vectors, but later I came to understand that they are very important calculationally and give another perspective on the metric tensor in that $g_{ij} = \mathbf{e}_i \bullet \mathbf{e}_j$ for example. This provides a "geometric" understanding of things in that the metric tensor is seen to be the set of "unnormalized" direction cosines or dot products of these basis vectors. After considering line, area and volume elements in this Section 5.16, M&M then move on to Section 5.17 where they have what I consider a weak discussion of what the "differential operators" look like in this general curvilinear q-space -- which I had to amp up in my divergence document. We are distracted here by the Christoffel symbols and covariance of the "covariant derivative", but we are given no explanation as to why this

should be the "divergence". Then after all this we have to jump backwards to Section 5.1 where we have a simplified rehash of the metric tensor, but here it is halfway called Q_{ij} and contravariant coordinate indices are not used at all, and on top of that we have diagonal derivatives called $g_{ii} = Q_{ii}^2$ but off diagonal ones called $g_{ij} = Q_{ij}$. I guess they wanted to avoid using g without dealing with the index locations. This was a horrible section IMHO. Nothing is said that "normal coordinates" are contravariant, even in those later sections.

Then in Section 5.2 we are back computing the differential operators again, instead of just specializing the general formulas we got in section 5.17. But this section fails to state that it is doing this only for orthogonal curvilinear coordinates, though at the bottom of page 173 they say that in fact all sections 5.2 through 5.15 will be orthogonal only. In this section, the $\mathbf{u}_i$ are unit vectors, but no connection is made to the fact that $\mathbf{u}_i = \hat{\mathbf{e}}_i$ of section 5.16. Again, there is no use of proper contravariant indices here because we are now doing "analysis" work instead of "tensor" work. I did not bother to read this section's derivations, but instead used it to verify the specializations of the general formulas in section 5.17. Not one peep is said about whether the vector identities are valid in curvilinear coordinates.

Then they start to review specific systems. The first is sphericals, but all they do is state the Q_i and nothing more, whereas they could have used this as an example of all the horsepower they developed in the later sections.

They then go on to describe all kinds of oddball orthogonal curvilinear coordinates. Here is a list of the 12 coordinates systems treated:

5.4 spherical	177	(page)
5.5 cylindrical	178	
5.6 confocal ellipsoidal	178	
5.7 prolate spheroidal	180	
5.8 oblate spheroidal	182	
5.9 elliptical cylindrical	182	
5.10 conical	183	
5.11 confocal paraboloidal	184	
5.12 parabolic	185	
5.13 parabolic cylindrical	186	
5.14 bipolar	187	
5.15 toroidal	190	

Some have pictures, but there are no great pictures.

So all in all, getting this information out of M&M was a bit like pulling teeth. It is all there somewhere, but the presentation order is not what I would have done. Murphy is a chemist, and this book was written in 1943 and updated in 1956 (my printing is 1967) when drawings were expensive. Still, having said all this, I find myself frequently going to this book for various math topics because it has so much detail on so many topics in its ample 600 pages.

So here is my attempt at integration.

The Transformation T and its Tensors: possible existence of a g_{ab} index-lowering tensor

We start out in Section 4.20 "tensors". I invented a shorthand T notation since it is easy for me to type, compared to the partial derivative notation, so M&M don't use my T. T is the "Partial Derivative Tensor" or "Transformation Matrix",

$$T^{a}{}_{b} \equiv \frac{\partial x'^{a}}{\partial x^{b}} \qquad T_{a}{}^{b} \equiv \frac{\partial x^{b}}{\partial x'^{a}} \qquad \delta^{a}{}_{c} = T_{c}{}^{b}\, T^{a}{}_{b}$$

This is based on some *arbitrary* $\mathbf{x'} = \mathbf{f(x)}$ **coordinate transformation** and the f_i could be any set of functions you might make up, as complex and tangled as you want. The orthonormality shown just follows from the chain rule of partial derivatives.

[Emphasis: This thing T is an ***arbitrary*** Coordinate Transformation based on some x' = f(x) which can be *anything you want*, does not even have to be linear, probably should be one to one in some sense. The Lorentz Transformation is just a particular case which happens to be *linear* x' = Λx with dimension 4. The linearity is what forces preservation of world-scalars as follows:

$$x'^{\mu} x'_{\mu} = \Lambda^{\mu}{}_{\alpha} x^{\alpha}\, \Lambda_{\mu}{}^{\beta} x_{\beta} = (\Lambda^{\mu}{}_{\alpha} \Lambda_{\mu}{}^{\beta})\, x^{\alpha}\, x_{\beta} = \delta_{\alpha}{}^{\beta}\, x^{\alpha}\, x_{\beta} = x^{\alpha}\, x_{\alpha}$$

So, a linear coordinate transformation preserves the quantity $x^{\alpha}\, x_{\alpha}$]

We learn about arbitrarily **mixed tensors** of any rank, how they transform. We learn how to make larger tensors doing **outer product** of smaller ones. We learn how to reduce rank by 2 doing **contraction**, and when this happens between two multiplied tensor objects it is an **inner product**.

In my "**facts added**" section (MM Chap 4_5 doc), I note that

$$(\partial_a\, T^{b}{}_{c}) = \partial_a \frac{\partial x'^{b}}{\partial x^{c}} = \frac{\partial^2 x'^{b}}{\partial x^{a} \partial x^{c}} = S^{b}{}_{ac} \qquad \text{displays a,c symmetry} \qquad \text{Fact}$$

$$(\partial_s\, T_{c}{}^{e}) = -\, T_{a}{}^{e}\, T_{c}{}^{b}\, S^{a}{}_{sb} \qquad \text{where recall } S^{a}{}_{sb} = (\partial_s T^{a}{}_{b}) = \text{symmetric } s \leftrightarrow b \qquad \text{Fact 2}$$

$$(\partial_s\, g^{dc}) = -\, g^{db}\, g^{c}{}_{a} (\partial_s\, g^{a}{}_{b}) \qquad \text{Fact 3}$$

[Note added 11.7.09: What is this g thing? A few paragraphs below, we are going to define the matrix g_{ij} (the metric tensor), so we are a little out of order here. If x^i are Cartesian coordinates, and if we start at some point $\mathbf{r} = \{x^i\}$ and make some small change in all the coordinates so $d\mathbf{r} = \{dx^i\}$, we end up at some new point $\mathbf{r}+d\mathbf{r}$. The distance^2 between this new vector and the original vector is $dr^2 \equiv \| d\mathbf{r} \|^2 = dx^i dx^i$. If we change to some non-Cartesian coordinate system where the coordinates are $x'^a = T^a{}_b\, x^b$, we find that the (normal Cartesian space) distance^2 between a vector $\mathbf{r'} = \{x'^i\}$ and $\mathbf{r'} + d\mathbf{r'}$ is given by $dr'^2 \equiv \| dr' \|^2 = g_{ij} dx'^i dx'^j$. In 2D polar coordinates for example we have $\|dr'\|^2 = (dr)^2 + r^2 (d\theta)^2$ so $g_{\theta\theta} = r^2$. The matrix g_{ij} which appears here is called "the metric tensor". With this as *definition*, we are then able to show many things about this g_{ij} object. (1) It is given by $g_{ij} = \Sigma_m\, T^m{}_i\, T^m{}_j$; (2) g_{ij} transforms under T as a rank 2 tensor; (3) g_{ij} is the thing that raises and lowers indices, changing covariant vectors to contravariant, etc. (4) g_{ij} is symmetric; (5) when it raises its own index, we get the identity matrix $g_{ab}\, g^{ac} = \delta_b{}^c$. Differentiating this equation is what gives Fact 3 above. The metric d is given by

[d(**r'**,**r'**+d**r'**)]^2 = $g_{ij}dx'^i dx'^j$. Since g determines the metric (= distance) between differentially separated points, and since g is a rank 2 tensor with respect to the transformation T, it is called "the metric tensor".]

We then learn that ***if*** there exists a symmetric **index-lowering tensor** g_{ab}, then we can construct an index-raising tensor according to

g^{ab} = cofactor(g_{ab})/det(g_{ab}) = G^{ab}/g // notation M&M page 196

One says that A^μ and A_μ are **associated** with respect to $g_{\mu\nu}$ if such a lowering tensor exists. *In this case*, we can do **arbitrary "tilt reversal"** on contracted indices (as long as not acted upon by a derivative!).

Then in Section 4.22 we learn that $\partial_a V_b$ is in general not a tensor, but $\partial_a V_b - \{ab,c\}V_c$ is. I then did a long proof of this last fact.

Nowhere in the above sections did we study the subject of differential distance, area or volume. And nowhere did we have a "metric tensor", but we did talk about possible index raising and lowering tensors.

Here are a few more facts we learned:

$T_a{}^b = (T^{-1})^b{}_a$
$T^a{}_b = (T^{-1})_b{}^a$

$V'^a = T^a{}_b\ V^b$ $\qquad$ $V'_a = T_a{}^b\ V_b$ $\qquad$ // fwd and inverse vector
$V^a = T_b{}^a\ V'^b$ $\qquad$ $V_a = T^b{}_a\ V'_b$

$Q'^{ab} = T^a{}_c\ T^b{}_d\ Q^{cd}$ $\qquad$ $Q'_{ab} = T_a{}^c\ T_b{}^d\ Q_{cd}$ $\qquad$ // fwd and inverse tensor
$Q^{ab} = T_c{}^a\ T_d{}^b\ Q'^{cd}$ $\qquad$ $Q_{ab} = T^c{}_a\ T^d{}_b\ Q'_{cd}$

General Comment: notice that the whole notion of something being a "tensor" means it is a tensor *with respect to* some transformation you specify! Consider T = $R_z(15^0)$ which is a certain 3x3 matrix. Or consider the most general rotation T = R(φ,θ,φ') with some Euler angles. This too is a 3x3 matrix, and this matrix R is what appears in all the tensor transformation formulas. Everything looks the same no matter what angles you choose, so this tensor world would be associated with T ε SO(3). For the Lorentz group, we have SO(1,3) Minkowski. For the transformation from Cartesian to spherical, we get a certain T which we will be seeing eventually here.

Notion of a distance metric; def. of the metric tensor g_{ij}; its use as the index-lowering tensor.

At this point, we move the Chapter 5 Section 5.16 and we start pondering differential distance, area and volume. But the topic is "curvilinear coordinates" q^k which we "get to" starting with some Cartesian coordinates x^k. I see no reason why we cannot think of q^k as being x'^k of the huge tensor discussion above. So, let's reformulate Section 5.1 in this language. Here is our inversion formula:

$$dx^i = \Sigma_k\ T_k{}^i\ dx'^k = \Sigma_k \frac{\partial x^i}{\partial x'^k}\ dx'^k$$

def: metric tensor: the thing g that appears here: $ds^2 = \Sigma_{ij}\ g_{ij}\ dx^i dx^j$

Notice that if you started with a mixed sym+antisym g_{ij} , only the sym part does anything in terms of distance, so throw out any antisym part if present!

[ds^2 must have dimensions of L^2 and must be a function of $\{dx^i\}$. What other form could it have?]

If we have two points that differ in position by $d\mathbf{r} = \{dx^i\}$, then the "distance squared" between those points is given by $ds^2 = \Sigma_{ij}\, g_{ij}\, dx^i dx^j$ and this is clearly a scalar or rank 0 tensor! Recall from Stakgold that if you have a linear vector space with a **norm** $\| \|$, you can define a "natural" **metric** for that space by d(x,y) = ||x-y|| and then you have a **metric space**. So we can state the above as

$$\| d\mathbf{r} \| = d(\mathbf{r}+d\mathbf{r}, \mathbf{r}) = ds = \sqrt{\Sigma_{ij}\, g_{ij}\, dx^i dx^j}$$

So we have a metric space, and it has a metric tensor called $\mathbf{g}_{ij}$.

Now consider a transformation between two spaces. Then our inverse transform is:

$$dx^a = T_b{}^a\, dx'^b \quad \text{so}$$

$$ds^2 = \Sigma_{ij}\, g_{ij}\, dx^i dx^j = \Sigma_{ij}\, g_{ij}\, (\Sigma_k T_k{}^i\, dx'^k)(\Sigma_p T_p{}^j\, dx'^p)$$
$$= \Sigma_k \Sigma_p\, (\Sigma_{ij}\, g_{ij}\, T_k{}^i\, T_p{}^j\,) dx'^k\, dx'^p$$
$$= \Sigma_k \Sigma_p\, g'_{kp}\, dx'^k\, dx'^p \qquad \text{where } g'_{kp} = T_k{}^i\, T_p{}^j\, g_{ij} \qquad \text{//implied sums}$$

and all this shows is the fact that **g transforms as a covariant rank 2 tensor**, in case we did not already know that fact -- here was the proof!

Just for fun, let's try to go "the other direction" and see what happens:

$$ds'^2 = \Sigma_{ij}\, g'_{ij}\, dx'^i dx'^j = \Sigma_{ij}\, g'_{ij}\, (\Sigma_k T^i{}_k\, dx^k)(\Sigma_p T^j{}_p\, dx^p)$$
$$= \Sigma_k \Sigma_p\, (\Sigma_{ij}\, g'_{ij}\, T^i{}_k\, T^j{}_p\,) dx^k\, dx^p$$
$$= \Sigma_k \Sigma_p\, g_{kp}\, dx^k\, dx^p \qquad \text{where } g_{kp} = T^i{}_k\, T^j{}_p\, g'_{ij} \qquad \text{//implied sums}$$
$$= ds^2$$

Here we see that g satisfies the correct "inverse formula" rule noted above. Also, ds^2 is of course a scalar or rank 0 tensor.

Now, if we do a general transform between two spaces, all we learn is that g transforms as it should.

Special Case #1: Transformation from Cartesian x to general x' = q coordinates

Consider the first space to be Cartesian so that $g_{ij} = \delta_{ij}$. In this case we find from the above that

$$g'_{kp} = T_k{}^i\, T_p{}^i = \frac{\partial x^i}{\partial x'^k}\frac{\partial x^i}{\partial x'^p} \qquad \text{but normally written:} \qquad g_{kp} = T_k{}^i\, T_p{}^i = \frac{\partial x^i}{\partial q^k}\frac{\partial x^i}{\partial q^p}$$

So this let's us actually *compute* the metric tensor in the primed new space (q space) from the transformation matrix. I think this is what we will be doing when we study spherical coordinates, and all the other cases.

Special Case #2: Transformation of Special Relativity

The requirement of special relativity is this:

$$g_{ij} = \text{diag}(1,-1,-1,-1)$$
$$g'_{ij} = \text{diag}(1,-1,-1,-1)$$

Review of this special relativity case:

(1) the conditions like $g_c{}^a = T_c{}^b\, T^a{}_b$ have nothing to do with the specialness of this case. These conditions are true for ANY transformation you can come up with! BDI p 16 writes this as $a_c{}^b\, a^a{}_b = \delta_c{}^a$ and $a_b{}^c\, a^b{}_a = \delta^c{}_a$

(2) the requirement that $g_{ab} = g'_{ab} = (1,-1,-1,-1)$ *seems* to impose a severe set of extra conditions on the transformation matrix elements, namely this tensor equation:

$$g'^{kp} = g^{kp} = T^k{}_i\, T^p{}_j\, g^{ij}.$$

which is equivalent to this

$$g^k{}_p = T^k{}_i\, T_p{}^i$$

But this is in fact the same as the orthonormality condition, so I guess the only conditions are the orthonormality ones.

Theorem. If you insist that $g_{ab} = g'_{ab}$ (regardless of what they actually are), then the equation which says that the metric tensor transforms as a rank 2 tensor is the same as the equation which imposes the orthonormality conditions.

Special relativity says that $g_{ab} = g'_{ab}$ AND it says that they are (1,-1,-1,-1).

You then figure out what the z-boost matrix is, for example, by imposing all the orthonormality conditions *and* the fact that $x'^3 = x^3 + (v/c)x^0$, $x'^1 = x$, $x'^2 = x^3$. Goldstein shows how this information then suffices to determine the boost matrix $T^k{}_i$ which he quickly reduces to a small 2x2 problem since x and y don't change.

Generators $M^{\mu\nu}$ and a boost example.

There are many ways to write the transformations in special relativity, here is one way, where antisym $M^{\mu\nu}$ represent the 6 Lorentz boost and rotation generators, and antisym $\sigma_{\mu\nu}$ the six corresponding parameters,

$T^k{}_i = \exp(\pm\, i\, \sigma_{\mu\nu}\, M^{\mu\nu})^k{}_i$ where $(M^{\mu\nu})^{kp} = g^{\mu k} g^{\nu p} - g^{\nu k} g^{\mu p}$

If you write $g'_{kp} = T_k{}^i\, T_p{}^j\, g_{ij}$ using the above and consider small $\sigma_{\mu\nu}$, it is easy to show that $g'_{kp} = g_{ij}$ to first order. In fact this is true to all orders, a harder fact to verify.

An interesting fact here is that the Lorentz generators are constructed from the metric tensor. I am not sure what this implies if the metric tensor is a function of x as in *general* relativity. Then we have a set of generators that can be different at each point.

Here is an example of a finite x-boost matrix as shown on my hand-written yellow checked paper notes, where s = sinh(u) and c = cosh(u), with u the boost parameter. The row/col ordering is the standard one where $x^\mu = (ct,x,y,z)$ for $\mu = 0,1,2,3$.

$$T^k{}_i = \begin{pmatrix} c & -s & & \\ -s & c & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$ // notice that $T^T = T$ for this particular matrix.

Now we can check whether $g'^{kp} = T^k{}_i\, T^p{}_j g^{ij} = T^k{}_i\, g^{ij} (T^T)^j{}_p$ or $g' = T\, g\, T^T = T\, g\, T$. Here is what Maple has to say:

```
T := matrix([[c,-s,0,0],[-s,c,0,0],[0,0,1,0],[0,0,0,1]]);
```

$$T := \begin{bmatrix} c & -s & 0 & 0 \\ -s & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

```
G := matrix([[1,0,0,0],[0,-1,0,0],[0,0,-1,0],[0,0,0,-1]]);
```

$$G := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

```
evalm(T&*G&*T);
```

$$\begin{bmatrix} c^2 - s^2 & 0 & 0 & 0 \\ 0 & s^2 - c^2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

and of course $c^2 - s^2 = 1$, so the result is the same as G, meaning $g'^{kp} = g^{kp}$. Thus, for an x-boost at least, we have shown that $g'^{kp} = g^{kp}$ "to all orders" in the boost parameter.