AST 231 (E. Blackman) SUPPLEMENT:

Clarification of Index/Tensor Notation, Basis Vectors, and the Minus Sign in Equation 5.83 of the Textbook

I wanted to clarify some of the basics of index. 4-vector, and tensor notation used in the text and explain why there is a minus sign in computing the energy of a particle as seen by a moving observer (equation 5.82, 5.83 of textbook). The book is not optimally clear on these points and hopefully the exposition below will help.

First note that in the text when 4-vectors such as space-time position x^μ or energy-momentum p^μ are written with the indices as superscripts, are called "contravariant" whereas when the indices are written as subscripts, the quantites are said to be "covariant". The two are "dual" spaces to each other in that there is a 1 to 1 correspondence between the information contained. Using the spacetime position vector as an example, the two forms can be related in special relativity by

$$x_\mu = \eta_{\mu\nu} x^\nu, \tag{1}$$

where repeated indices are summed and $\eta_{\mu\nu}$ is our metric tensor for Minkowski space used to calculate distances and scalar products. This tensor is of "rank 2" because it has two subscript indices and is here essentially just a symmetric 4 x 4 matrix with $\eta_{11} = \eta_{22} = -\eta_{33} = 1 = -\eta_{00}$ and all other components zero.

The metric tensor $\eta_{\mu\nu}$ is used to calculate scalar products in the following way: For example, the line element is a scalar product

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = dx_\nu dx^\nu = dx_\lambda dx^\lambda \tag{2}$$

remembering that repeated indices are summed (so the replacement of ν by λ in the last equality simply gives equivalent sums). In (2) the contravariant differential spacetime position 4-vector $dx^\mu = (dt, dx, dy, dz)$. The metric tensor $\eta_{\mu\nu}$ can be said to "lower" indices and thereby convert contravariant to covariant forms as in Eq. (1) so that e.g. covariant 4-vector $dx_\mu = \eta_{\mu\nu} dx^\nu = (-dt, dx, dy, dz)$. We can also "raise" indices by the tensor $\eta^{\mu\nu}$. This is related to $\eta_{\mu\nu}$ by

$$\delta_\mu^{\ \nu} = \eta_{\mu\lambda}\eta^{\lambda\nu}, \tag{3}$$

where repeated indices are summed and the left side is the 4 x 4 identity matrix. This shows that $\eta_{\mu\lambda} = \eta^{\mu\lambda}$, and so multiplying equation (1) by $\eta^{\lambda\mu}$ will return x^λ on the right side.

When we write a 4-vector in index form, (i.e. x^μ) this represents an unspecified component of the 4-vector $\mathbf{x}$ much like in 3-D if we write x_i, we mean any one of the x, y, z

components of a vector. The book introduces the basis notation where the 4-vector can be written as

$$\mathbf{x} = x^\alpha \mathbf{e}_\alpha, \tag{4}$$

where $\mathbf{e}_\alpha$ is a set of four covariant 4-vectors. That is, e.g. $\mathbf{e}_0$ is a time-like 4-vector, and $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are the spacelike 4-vectors. The subtlety of this notaton is that $\mathbf{e}_\alpha$ is really a matrix because it is a set of four 4-vectors, each of which have 4 components. This is why (4) is not a scalar: although repeated indices are summed in (4) there are still 4 unsummed indices. To see this more explicitly, for practical purposes we can write (4) in index notation as

$$x^\mu = x^\alpha {e_\alpha}^\mu \tag{5}$$

Presenting (5) in this index notation format helps explain why there is a minus sign in the first equality of equation 5.82 and 5.83 of the text as I now discuss.

Equation 5.82 of the text represents the projection of the energy momentum 4-vector of a particle in an inertial frame, onto the local orthonormal frame of an arbitrarily moving observer and therefore represents the value of the energy-momentum 4-vector as measured by this observer. Using index notation we can write Equation 5.82 as a 4-vector

$${p_{obs}}^\mu = \eta^{\mu\nu} \eta_{\lambda\sigma} p^\lambda {e_\nu}^\sigma, \tag{6}$$

remembering that repeated indices are summed, and that the energy momentum 4-vector on the right side is that measured in the inertial frame, while the basis "matrix" ${e_\nu}^\sigma$ corresponds to that of the moving observer as also measured in the inertial frame.

Now consider the 0 component of (6) (the timelike component):

$${p_{obs}}^0 = \eta^{0\nu} \eta_{\lambda\sigma} p^\lambda {e_\nu}^\sigma. \tag{7}$$

Since $\eta^{0\nu}$ is zero unless $\nu = 0$ we can replace ν by 0 in (7). This gives

$${p_{obs}}^0 = \eta^{00} \eta_{\lambda\sigma} p^\lambda {e_0}^\sigma. \tag{8}$$

But $\eta^{00} = -1$ so this becomes

$${p_{obs}}^0 = -\eta_{\lambda\sigma} p^\lambda {e_0}^\sigma. \tag{9}$$

But this is just MINUS the scalar product between the energy momentum 4-vector in the inertial frame and the time-like basis vector of the moving observer. Because the σ index is a superscript this basis vector is contravariant and, as discussed in class and the text, $e_0^\sigma = u_{obs}^\sigma$, the 4-velocity of the moving observer. Thus (9) is

$${p_{obs}}^0 = -\mathbf{p} \cdot \mathbf{u}_{obs}, \tag{10}$$

which is exactly equation 5.83 of the text since $E = p^0_{obs}$.

For the 1,2,3 components of (6) (the space like components) we do not obtain a minus sign as in 10. The reason can be seen using the 1 component of (6) as an example. That component is

$$p_{obs}{}^1 = \eta^{1\nu}\eta_{\lambda\sigma}p^\lambda e_\nu{}^\sigma. \tag{11}$$

Here only the $\nu = 1$ contribution contributes because $\eta_{\mu\nu}$ is diagonal. Since $\eta^{11} = 1$ we then have:

$$p_{obs}{}^1 = \eta^{11}\eta_{\lambda\sigma}p^\lambda e_1{}^\sigma = \eta_{\lambda\sigma}p^\lambda e_1{}^\sigma = \mathbf{p}\cdot\mathbf{e}_1, \tag{12}$$

which corresponds to the second equation in 5.82. The 2 and 3 components (y,z components) follow similarly.

One final point on the meaning of equations 5.81-5.83: The energy and momentum measured by a specified observer with a specified orthonormal basis is the SAME when the right side of 5.81-5.83 are calculated in ANY FRAME. This is same concept as the fact that all observers agree what the time measured on a specified observer's clock reads once that observers frame is specified. (E.g. in the twin paradox, the stationary twin agrees that moving twin's clock should measure less time passed given her different path in spacetime). So once you specify the orthonormal basis of a specific observer using some set of coordinates, you can calculate what that specific observer measures using the coordinates in any other frame.