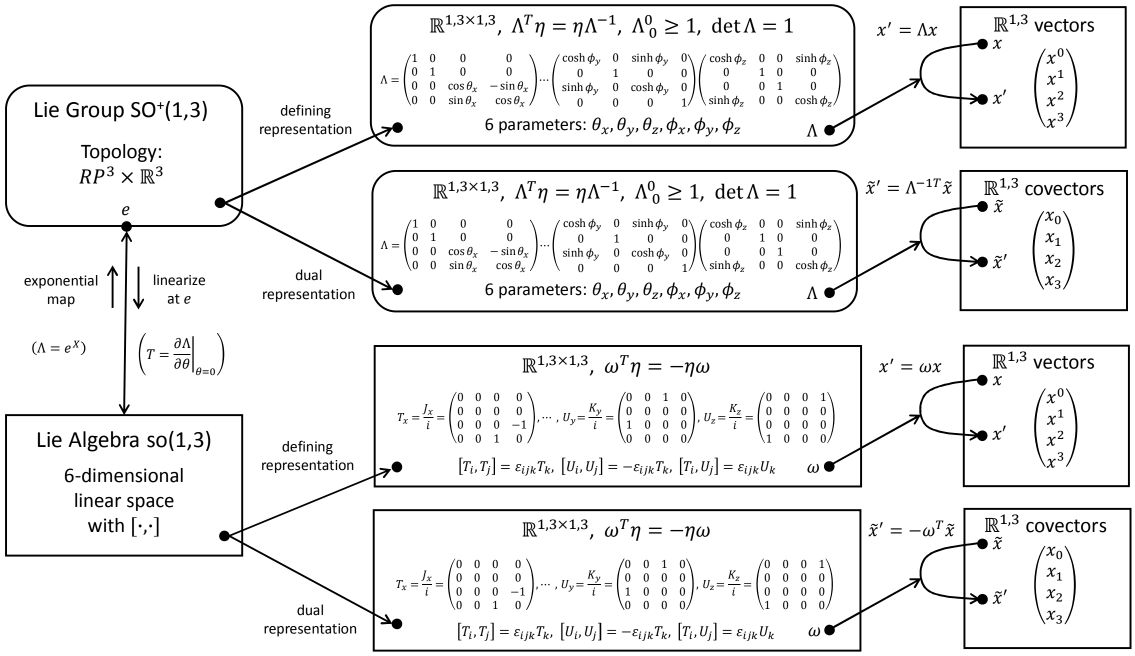
5.12 SO⁺(1,3): Dual Representation; Vectors and Covectors



In some respects, Euclidean space and Minkowski space are very similar, but in others they are rather different. Maybe the most dramatic difference is the following: when two points in Euclidean space are separated by zero distance, they are necessarily on top of each other, but when two events in Minkowski space are separated by zero proper time, they can be anywhere on the same light cone! Another difference, which is the topic of this example, is that whereas in Euclidean space there is no need to distinguish between vectors and *covectors* (= *dual vectors*), we do have to make that distinction in Minkowski space because the two objects transform differently under SO⁺(1,3).

We know from our discussion of SU(2) that a covector is a linear map from vectors to scalars. The prototypical space-time covector is the differential operator $\nabla = (\partial/\partial t, \partial/\partial x, \partial/\partial y, \partial/\partial z)^T$, which acts on the vector $x = (t, x, y, z)^T$ like $\nabla^T x = \partial t/\partial t + \partial x/\partial x + \partial y/\partial y + \partial z/\partial z = 4$ producing a scalar (invariant). Note that we do *not* need a metric when letting a covector act on a vector as we do when taking the scalar product of two vectors. So, how do covectors transform? If a vector transforms like $x' = \Lambda x$ and we have the invariance $\tilde{y}'^T x' = \tilde{y}^T x$, where $\tilde{y}$ is a covector, then $\tilde{y}^T\Lambda^{-1}\Lambda x = \tilde{y}^T x$. Thus, a covector transforms like $\tilde{y}'^T = \tilde{y}^T\Lambda^{-1}$ or, after transposing, like $\tilde{y}' = \Lambda^{-1T}\tilde{y}$ (where Λ can be any linear, invertible transformation). The diagram shows the defining representation of SO⁺(1,3), which acts on vectors, in the upper branch and its *dual representation*, which acts on covectors, in the lower branch.

Up to this point we did not make use of the fact that Λ is a Lorentz transformation. If Λ satisfies the defining condition of the Lorentz transformation, $\Lambda^T\eta = \eta\Lambda^{-1}$, we can write $\Lambda^{-1T} = \eta\Lambda\eta$ (note that $\eta = \eta^{-1} = \eta^T$). Thus, vectors and covectors in Minkowski space-time do transform differently! Nevertheless, the two transformations are related by the similarity transformation $\Lambda^{-1T} = \eta^{-1}\Lambda\eta$. In other words, the representations of SO⁺(1,3) on vectors and covectors are *equivalent*; we can think of them as two different *flavors* of the same representation. (Remember, something similar happened for SU(2).) Letting $\eta \to I$, confirms that in Euclidean space vectors and covectors transform in the same way.

It is conventional to write vector components with *upstairs* (a.k.a. *contravariant*) indices and covector components with *downstairs* (a.k.a. *covariant*) indices (see the diagram). Given that we learn in high school to write vector components with downstairs indices, this is the opposite of what we might expect, but unfortunately, we are stuck with this convention. Furthermore, it is conventional is to use Greek indices if they run over 0 to 3 (and Latin indices if they run over 1 to 3). Thus, a space-time vector, an energy-momentum vector, and the del covector (∇) are written as follows ($c = 1$) [QFTGA, Ch. 0.4]:

$$(x^\mu) = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}, \quad (p^\mu) = \begin{pmatrix} p^0 \\ p^1 \\ p^2 \\ p^3 \end{pmatrix} = \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix}, \quad (\partial_\mu) = \begin{pmatrix} \partial_0 \\ \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix} = \begin{pmatrix} \partial/\partial x^0 \\ \partial/\partial x^1 \\ \partial/\partial x^2 \\ \partial/\partial x^3 \end{pmatrix} = \begin{pmatrix} \partial/\partial t \\ \partial/\partial x \\ \partial/\partial y \\ \partial/\partial x \end{pmatrix}.$$

To avoid confusion, we will distinguish between "high-school notation", where we write vector components with downstairs indices and "tensor-index notation", where we write vectors with upstairs and covectors with downstairs indices. Using the tensor-index notation, we can rewrite the vector transformation $x' = \Lambda x$ as $x'^\mu = \Lambda^\mu{}_\nu x^\nu$ and the covector transformation $\tilde{x}' = \Lambda^{-1T}\tilde{x}$ as $x'_\mu = \Lambda_\mu^\nu x_\nu$, where a summation over the repeated index ν is implied [TM, Vol. 3, Ch. 6.1.4-5; QFTGA, Ch. 0.4]. Tensor-index notation always implies the Einstein summation convention. Note that although $\Lambda^\mu{}_\nu$ and Λ_μ^ν have the same main symbol, they represent two different matrices: the index positions do matter!

Given a metric η to take the scalar product of two vectors, we can identify covector $\tilde{y}$ with vector y such that $\tilde{y}^T x = y^T \eta x$ holds for any vector x. This identification can then be used to "hide" the metric: instead of taking the scalar product of two vectors (which requires a metric), we convert one vector into a covector (using the metric) and then act with it on the second vector (which does not require a metric). Using the tensor-index notation, this means that we identify the vector y^μ with the covector $y_\mu = y^\nu \eta_{\nu\mu}$. Then, the scalar product of two vectors, $\eta_{\nu\mu} y^\nu x^\mu$, can be written more elegantly as $y_\mu x^\mu$. For example, space-time and energy-momentum covectors are written as follows:

$$(x_\mu) = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} = \begin{pmatrix} t \\ -x \\ -y \\ -z \end{pmatrix}, \qquad (p_\mu) = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} p^0 \\ -p^1 \\ -p^2 \\ -p^3 \end{pmatrix} = \begin{pmatrix} E \\ -p_x \\ -p_y \\ -p_z \end{pmatrix}.$$

With this trick of pulling down an index, we can rewrite the proper time $\tau^2 = x^T \eta x = \eta_{\nu\mu} x^\nu x^\mu$ more elegantly as $\tau^2 = x_\mu x^\mu$ and the rest mass $m^2 = p^T \eta p = \eta_{\nu\mu} p^\nu p^\mu$ as $m^2 = p_\mu p^\mu$.

Two more comments about conventions: First, some authors (e.g. [TM]) define the Minkowski metric with the opposite sign than we do: $\eta = \mathrm{diag}(-1, +1, +1, +1)$. Second, physicists like to include a factor i as part of the generators and then put a $-i$ into the exponent of the exponential map to keep things consistent. This is analogous to what we did for SU(2) when discussing quantum-mechanical applications. The standard names for the modified basis generators are $J_i = iT_i$ and $K_i = iU_i$ (sometimes $K_i = -iU_i$). The commutation relations of the modified generators include a factor i on the right-hand side: $[J_i, J_j] = i\varepsilon_{ijk} J_k$, $[K_i, K_j] = -i\varepsilon_{ijk} J_k$, and $[J_i, K_j] = i\varepsilon_{ijk} K_k$ (regardless of whether $K_i = \pm iU_i$) [PfS, Ch. 3.7.3; QFTGA, Ch. 9.5]. For our purposes, however, it is more convenient to stay with the original basis generators T_i and U_i. The diagram shows both forms of the generators for comparison.