Chapter 1. Four-Dimensional Vector Spaces and Linear Mappings

1.1. The Minkowski Vector Space $\mathbb{V}_4$

Note. In relativity, we study *spacetime*, which consists of points called *events*. Events consist of three spatial coordinates (x, y, z) and one time coordinate t. Our spacetime will be a vector space over the field $\mathbb{R}$.

Note. We initially denote our vector space as $\mathbb{V}$ and require several properties (axioms).

Addition Axioms.

A1. $\mathbf{a} + \mathbf{b} \in \mathbb{V}$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{V}$.

A2. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{V}$.

A3. $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{V}$.

A4. There exists $\mathbf{0} \in \mathbb{V}$ such that $\mathbf{a} + \mathbf{0} = \mathbf{a}$ for all $\mathbf{a} \in \mathbb{V}$.

A5. For each $\mathbf{a} \in \mathbb{V}$ there exists $-\mathbf{a} \in \mathbb{V}$ such that $(-\mathbf{a}) + \mathbf{a} = \mathbf{0}$.

Scalar Multiplication Axioms.

M1. $\alpha\mathbf{a} \in \mathbb{V}$ for all $\alpha \in \mathbb{R}$ and for all $\mathbf{a} \in \mathbb{V}$.

M2. $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$ for all $\alpha, \beta \in \mathbb{R}$ and for all $\mathbf{a} \in \mathbb{V}$.

M3. $1\mathbf{a} = \mathbf{a}$ for all $\mathbf{a} \in \mathbb{V}$.

M4. $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$ for all $\alpha \in \mathbb{R}$ and for all $\mathbf{a}, \mathbf{b} \in \mathbb{V}$.

M5. $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$ for all $\alpha, \beta \in \mathbb{R}$ and for all $\mathbf{a} \in \mathbb{V}$.

Note. Axioms A1–A5 and M1–M5 are sufficient to define a vector space over $\mathbb{R}$.

Dimensionality Axiom.

D1. $\dim(\mathbb{V}) = 4$.

Note. Axioms A1–A5, M1–M5, and D1 define $\mathbb{R}^4$.

<u>Inner Product Axioms.</u>

We postulate an operation $\cdot$ which takes a pair of vectors to a scalar. $\cdot$ is called a *dot product, scalar product*, or *inner product*. Some desirable axioms for $\cdot$ are

I1. $\mathbf{a} \cdot \mathbf{b} \in \mathbb{R}$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{V}$.

I2. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{V}$.

I3. $(\lambda \mathbf{A} + \mu \mathbf{b}) \cdot \mathbf{c} = \lambda(\mathbf{a} \cdot \mathbf{c}) + \mu(\mathbf{b} \cdot \mathbf{c})$ for all $\lambda, \mu \in \mathbb{R}$ and for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{V}$.

I4. (Axiom of Nondegeneracy) $\mathbf{a} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in \mathbb{V}$ if and only if $\mathbf{a} = \mathbf{0}$.

I5. (Axiom of Positive Definite Inner Product) $\mathbf{a} \cdot \mathbf{a} \geq 0$ for all $\mathbf{a} \in \mathbb{V}$, and $\mathbf{a} \cdot \mathbf{a} = 0$ if and only if $\mathbf{a} = \mathbf{0}$.

Note. Axioms A1-A5, M1-M5, D1, I1-I3, and I5 define $\mathbb{R}^4$ with the usual dot product as the inner product.

Note. *I*5 implies *I*4.

Proof. First, suppose $\mathbf{a} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in \mathbb{V}$. Then $\mathbf{a} \cdot \mathbf{a} = 0$. So $\mathbf{a} = \mathbf{0}$. Second, suppose $\mathbf{a} = \mathbf{0}$. Then $\mathbf{a} \cdot \mathbf{x} = (\mathbf{a} + \mathbf{a}) \cdot \mathbf{x} = \mathbf{a} \cdot \mathbf{x} + \mathbf{a} \cdot \mathbf{x} = 2\mathbf{a} \cdot \mathbf{x}$ So $\mathbf{a} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in \mathbb{V}$. ∎

Note. In defining the Minkowski vector space, we will exclude I5 and take the weaker I4. We will add one final axiom shortly.

Definition. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ be an ordered basis for a vector space satisfying A1-A5, M1-M5, I1-I5, and D1. Define a 4×4 matrix G of *metric tensor components* relative to the ordered basis as $G = [g_{ij}]$ where $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j$ for $i, j \in \{1, 2, 3, 4\}$.

Note. Under Axiom I2, $g_{ij} = g_{ji}$ and so G is symmetric. By the "Fundamental Theorem of Symmetric Matrices" (see, for example, Fraleigh and Beauregard's *Linear Algebra*), square real matrices which are symmetric have real eigenvalues. Hence the eigenvalues of G are real.

Note. Assume the Axiom of Nondegeneracy (I4), along with A1-A5, M1-M5, D1, and I1-I3. We show that 0 cannot be an eigenvalue of G. Suppose, to the contrary, that $\lambda = 0$ is an eigenvalue of G and suppose $\mathbf{a}$ is a corresponding (nonzero) eigenvector. Then for any $\mathbf{x} \in \mathbb{V}$ we have $\mathbf{x} \cdot \mathbf{a} = \mathbf{x} G \mathbf{a}^T = \mathbf{x} \cdot \mathbf{0} = 0$. (Notice $\mathbf{x} \cdot \mathbf{0} = \mathbf{x} \cdot (\mathbf{0} + \mathbf{0}) = 2\mathbf{x} \cdot \mathbf{0}$ implies $\mathbf{x} \cdot \mathbf{0} = 0$. If we assume an ordered basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ for $\mathbb{V}$ and write $\mathbf{a}$ and $\mathbf{x}$ in terms of this basis, then we find $\mathbf{x} \cdot \mathbf{a} = \mathbf{x} G \mathbf{a}^T$ — see the proof of Theorem 1.1.2.) However, $\mathbf{a} \neq \mathbf{0}$ contradicting I4. Hence 0 cannot be an eigenvalue of G.

Note. We have one last axiom/property in the definition of the Minkowski vector space:

S1. (Axiom of Lorentz Signature) The eigenvalues of G satisfy $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0, \lambda_4 < 0$.

Definition. A vector space satisfying A1-A5, M1-M5, I1-I4 (not I5), D1, and S1 is called the *Minkowski vector space*, denoted $\mathbb{V}_4$.

Definition. $\mathbf{a}, \mathbf{b} \in \mathbb{V}_4$ are *Minkowski orthogonal* (or *M-orthogonal*) if $\mathbf{a} \cdot \mathbf{b} = 0$.

Theorem 1.1.1 There exists an M-orthogonal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ for $\mathbb{V}_4$ such that $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = d_{ij}$ for $i, j \in \{1, 2, 3, 4, \}$ where

$$D = [d_{ij}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Note. The text states that "the proof [of Theorem 1.1.1] is rather involved." However, we can satisfy the desired properties by taking $\mathbf{e} = [1, 0, 0, 0]$, $\mathbf{e}_2 = [0, 1, 0, 0]$, $\mathbf{e}_3 = [0, 0, 1, 0]$, and $\mathbf{e}_4 = [0, 0, 0, i]$ where $i^2 = -1$. This is somewhat illegitimate since we have stated that $\mathbb{V}_4$ is a vector space over $\mathbb{R}$, and we have the length of $\mathbf{e}_4$ is i (using the norm induced by the inner product).

Definition. We define M as the *matrix of the Lorentz metric* and the d_{ij} as the *coefficients of the Lorentz metric*.

Note. The text comments (in terms of trace and signature) that some sources replace D with $-D$.

Note. We adopt the *Einstein summation notation* in which any repeated Roman index is understood to be summed over for that index taking the values $1, 2, 3, 4$. For example,

$$u^k v_k = \sum_{k=1}^{4} u^k v_k = \sum_{l=1}^{4} u^l v_l = u^l v_l$$

and

$$g_{ij} u^i v^j = \sum_{i=1}^{4} \sum_{j=1}^{4} g_{ij} u^i v^j = \sum_{k=1}^{4} \sum_{l=1}^{4} g_{kl} u^k v^l = g_{kl} u^k v^l.$$

Such indices cannot be repeated more than twice.

Definition. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ be an M-orthonormal basis (or *tetrad*) for $\mathbb{V}_4$. The unique scalars u^i where

$$\mathbf{u} = \sum_{i=1}^{4} u^i \mathbf{e}_i = u^i \mathbf{e}_i$$

are the *Minkowski components* of $\mathbf{u}$ with respect to the (ordered) basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$.

Theorem 1.1.2. The Minkowski components u^i and v^j for $\mathbf{u}$ and $\mathbf{v}$ are related to the inner product of $\mathbf{u}$ and $\mathbf{v}$ as

$$\mathbf{u} \cdot \mathbf{v} = d_{ij} u^i v^j.$$

Proof. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ be an M-orthonormal basis for $\mathbb{V}_4$ where $\mathbf{u} = u^i \mathbf{e}_i$ and $\mathbf{v} = v^k \mathbf{e}_j$. Then

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{v} &= (u^i \mathbf{e}_i) \cdot (v^j \mathbf{e}_j) = \left(\sum_{i=1}^{4} u^i \mathbf{e}_i \right) \cdot \left(\sum_{j=1}^{4} v^j \mathbf{e}_j \right) \\
&= \sum_{i=1}^{4} \sum_{j=1}^{4} (u^i \mathbf{e}_i) \cdot (v^j \mathbf{e}_j) = (u^i \mathbf{e}_i) \cdot (v^j \mathbf{e}_j) \\
&= (\mathbf{e}_i \mathbf{e}_j) u^i v^j = d_{ij} u^i v^j.
\end{aligned}$$

∎

Note. Given the values of d_{ij} from Theorem 1.1.1, we see that

$$\mathbf{u} \cdot \mathbf{u} = d_{ij}u^i u^j = (u^1)^2 + (u^2)^2 + (u^3)^2 - (u^4)^2.$$

So we cannot employ the usual definition of length of a vector as $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$, since $\mathbf{u} \cdot \mathbf{u}$ may be negative (unless we are willing to allow imaginary "lengths"). We also cannot define the angle between two vectors $\mathbf{u}$ and $\mathbf{v}$ as less than -1 (unless we allow complex angles). Therefore, we abandon the ideas of length and angle in $\mathbb{V}_4$. However, the subspace

$$\mathbb{V}_3 = \{\mathbf{v} \in \mathbb{V}_3 \mid v^4 = 0\}$$

has all the usual properties we are familiar with, since it is isomorphic to $\mathbb{R}^3$.

Definition. Let $\mathbf{u} \in \mathbb{V}_4$.

(i) if $\mathbf{u} \cdot \mathbf{u} > 0$, then $\mathbf{u}$ is *spacelike*,

(ii) if $\mathbf{u} \cdot \mathbf{u} < 0$, then $\mathbf{u}$ is *timelike*, and

(iii) if $\mathbf{u} \cdot \mathbf{u} = 0$, then $\mathbf{u}$ is a *null vector* or is *lightlike*.

Definition. We define a *point* in $\mathbb{V}_4$ as an ordered 4-tuple $U = (u^1, u^2, u^3, u^4)$ where $u^i \in \mathbb{R}$. Points in $\mathbb{V}_4$ are called *events*. The *vector* from $U = (u^1, u^2, u^3, u^4)$ to $V = (v^1, v^2, v^3, v^4)$ where U and V are points in $\mathbb{V}_4$, is the vector space with Minkowski components $v^1 - u^1$, $v^2 - u^2$, $v^3 - u^3$, $v^4 - u^4$.

Note. If there is a lightlike (null) vector from one event to another, then it is possible for the same photon to be present at both events, If there is a timelike vector from one event to another, then it is possible for a material particle (such as an observer) to be present at both event (that is, there is enough *time* between the two events for an observer to traverse the *space* between them). If there is a spacelike vector from one event to another, then it is not possible for a material particle or photon to be present at both events (that is, there is too much *space* between the events for a particle or photon to traverse the space in the "limited" amount of *time*).

Example. For $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ an M-orthonormal basis for $\mathbb{V}_4$, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are spacelike vectors since for $i \in \{1, 2, 3\}$ $\mathbf{e}_i \cdot \mathbf{e}_i = d_{ii} = 1$. However, $\mathbf{e}_4 \cdot \mathbf{e}_4 = d_{44} = -1$ and so $\mathbf{e}_4$ is timelike. For $i \in \{1, 2, 3\}$ $\mathbf{u} = \mathbf{e}_i + \mathbf{e}_4$ is a null vector since

$$\mathbf{u} \cdot \mathbf{u} = (\mathbf{e}_i + \mathbf{e}_4) \cdot (\mathbf{e}_i + \mathbf{e}_4) = \mathbf{e}_i \cdot \mathbf{e}_i + 2\mathbf{e}_i \cdot \mathbf{e}_4 + \mathbf{e}_4 \cdot \mathbf{e}_4$$
$$= d_{ii} + 2d_{i4} + d_{44} = 1 + 2(0) + (-1) = 0.$$

Definition. The *separation number* between two points in $\mathbb{V}_4$ is $\sigma(\mathbf{u}) = \sqrt{|\mathbf{u} \cdot \mathbf{u}|}$ where $\mathbf{u}$ is a vector between the two points. A vector $\mathbf{e} \in \mathbb{V}_4$ is a *unit vector* if $\sigma(\mathbf{e}) = 1$.

Notation. We will use small Greek indices $(\alpha, \beta, \gamma, \delta, \ldots)$ when we want to cover values $1, 2, 3$, and use small Roman indices $(i, j, k, l, \ldots)$ when we want to cover the values $1, 2, 3, 4$.

Theorem 1.1.3. Schwarz Inequality.

For $\mathbf{u}, \mathbf{v} \in \mathbb{V}_3$, we have

$$|\mathbf{u} \cdot \mathbf{v}|^2 = |u^\alpha v^\alpha|^2 \leq u^\alpha u^\alpha v^\beta v^\beta = (\mathbf{u} \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{v}) = |\mathbf{u}|^2 |\mathbf{v}|^2.$$

Equality holds if and only if $u^\alpha = \lambda v^\alpha$ for some $\lambda \in \mathbb{R}$.

Proof. Suppose that $\mathbf{u} = \mathbf{0}$ and $u^\alpha = 0$. Then the result holds trivially. Now suppose $\mathbf{u} \neq \mathbf{0}$ and so $u^\alpha u^\alpha > 0$. Then for any $\lambda \in \mathbb{R}$ we have

$$(\lambda u^\alpha + v^\alpha)(\lambda u^\alpha + v^\alpha) = \lambda^2 u^\alpha u^\alpha + v^\alpha v^\alpha + 2\lambda u^\alpha v^\alpha \geq 0.$$

With λ set equal to $-(u^\alpha v^\alpha)/(u^\gamma u^\gamma)$ we get

$$[-(u^\alpha v^\alpha)^2 + (u^\alpha u^\alpha)(v^\beta v^\beta)]/(u^\gamma u^\gamma) \geq 0.$$

The result follows from this. The necessary and sufficient conditions for equality are left as Exercise 1.1.4.   ∎

Note. We now derive some properties of M-orthogonality and try to give a geometric interpretation at the end of the section.

Theorem 1.1.4. No two timelike vectors in $\mathbb{V}_4$ can be M-orthogonal.

Proof. Let $\mathbf{u}, \mathbf{v}$ be timelike. Then

$$\mathbf{u} \cdot \mathbf{u} = d_{ij}u^i u^j = u^\alpha u^\alpha - (u^4)^2 < 0$$

$$\mathbf{v} \cdot \mathbf{v} = d_{ij}v^i v^j = v^\beta v^\beta - (v^4)^2 < 0.$$

Therefore

$$u^\alpha u^\alpha < (u^4)^2$$
$$v^\beta v^\beta < (v^4)^2$$

and

$$u^\alpha u^\alpha v^\beta v^\beta < (u^4 v^4)^2,$$

or

$$\sqrt{u^\alpha u^\alpha v^\beta v^\beta} < |u^4 v^4|.$$

By the Schwarz Inequality, $|u^\alpha v^\alpha| \leq \sqrt{u^\alpha u^\alpha v^\beta v^\beta}$, so we conclude that $|u^\alpha v^\alpha| < |u^4 v^4|$. Hence we cannot have

$$\mathbf{u} \cdot \mathbf{v} = u^i v^i = u^\alpha u^\alpha - u^4 u^4 = 0.$$

∎

Note. Of course, two spacelike vectors can be M-orthogonal. Consider $\mathbf{e}_1$ and $\mathbf{e}_2$.

Corollary 1.1.1. For two timelike vectors $\mathbf{u}, \mathbf{v}$ such that $u^4 > 0$, $v^4 > 0$, we have $\mathbf{u} \cdot \mathbf{v} < 0$. ($\mathbf{u}$ and $\mathbf{v}$ are called "future pointing.")

Proof. Since $\mathbf{u}$ and $\mathbf{v}$ are timelike, then

$$\mathbf{u} \cdot \mathbf{u} = u^i u^i = u^\alpha u^\alpha - u^4 u^4 < 0$$

and

$$\mathbf{v} \cdot \mathbf{v} = v^i v^i = v^\beta v^\beta - v^4 v^4 < 0,$$

or

$$u^\alpha u^\alpha < u^4 u^4 \text{ and } v^\beta v^\beta < v^4 v^4$$

or

$$\sqrt{u^\alpha u^\alpha} < |u^4| \text{ and } \sqrt{v^\beta v^\beta} < |v^4|.$$

Now

$$\begin{aligned} u^\alpha v^\alpha &\leq |u^\alpha v^\alpha| \leq \sqrt{u^\alpha u^\alpha v^\beta v^\beta} \text{ by the Schwarz Inequality} \\ &< |u^4 v^4| = u^4 v^4. \end{aligned}$$

Hence $\mathbf{u} \cdot \mathbf{v} = u^\alpha v^\alpha - u^4 v^4 < 0$. ∎

Theorem 1.1.5. (Synge)

Let $\hat{\mathbf{t}}, \mathbf{t}$ be two timelike, future pointing (i.e. $\hat{t}^4 > 0$ and $t^4 > 0$) unit vectors. Then $-\infty < \hat{\mathbf{t}} \cdot \mathbf{t} \leq -1$ ("future pointing" may be replaced with "past pointing").

Proof. Since $\hat{\mathbf{t}}$ and $\mathbf{t}$ are future pointing unit vectors,

$$\sigma(\mathbf{t}) = 1, \quad t^\alpha t^\alpha - (t^4)^2 = -1, \quad t^4 > 0$$

$$\sigma(\hat{\mathbf{t}}) = 1, \quad \hat{t}^\beta \hat{t}^\beta - (\hat{t}^4)^2 = -1, \quad \hat{t}^4 > 0.$$

Now consider two 3-dimensional unit vectors $\mathbf{a}$ and $\hat{\mathbf{a}}$. Then $a^\alpha a^\alpha = 1$ and $\hat{a}^\alpha \hat{a}^\alpha = 1$. We can express the components of these unit vectors (with respect to the standard basis) in spherical coordinates as

$$
\begin{aligned}
a^1 &= \sin\theta\cos\phi & \hat{a} &= \sin\hat{\theta}\cos\hat{\phi} \\
a^2 &= \sin\theta\sin\phi & \hat{a}^2 &= \sin\hat{\theta}\sin\hat{\phi} \\
a^3 &= \cos\theta & \hat{a}^3 &\cos\hat{\theta}
\end{aligned}
$$

where $\theta, \hat{\theta} \in [0, \pi]$ and $\phi, \hat{\phi} \in [-\pi, \pi)$. If $\psi \in [0, \pi]$ is the angle between $\mathbf{a}$ and $\hat{\mathbf{a}}$ then since these are unit vectors

$$
\begin{aligned}
\cos\psi &= \mathbf{a} \cdot \hat{\mathbf{a}} = a^\alpha \hat{a}^\alpha \\
&= \sin\theta\cos\phi\sin\hat{\theta}\cos\hat{\phi} + \sin\theta\sin\phi\sin\hat{\theta}\sin\hat{\phi} + \cos\theta\cos\hat{\theta} \\
&= \cos\theta\cos\hat{\theta} + \sin\theta\sin\hat{\theta}(\cos\phi\cos\hat{\phi} + \sin\phi\sin\hat{\phi}) \\
&= \cos\theta\cos\hat{\theta} + \sin\theta\sin\hat{\theta}\cos(\phi - \hat{\phi}). \quad (*)
\end{aligned}
$$

Now since $t^\alpha t^\alpha - (t^4)^2 = -1$ and $t^4 > 0$, then $t^4 \geq$ (and similarly for $\hat{t}^4 \geq 1$). So we can find χ and $\hat{\chi}$ such that $t^4 = \cosh\chi$ and $\hat{t}^4 = \cosh\hat{\chi}$.

Next, for a^α components of a (3-D) unit vector, we have

$$(\sinh^2 \chi) a^\alpha a^\alpha - \cosh^2 \chi = \sinh^2 \chi - \cosh^2 \chi = -1.$$

Therefore such a^α's can be chosen for a given $\mathbf{t}$ with

$$t^\alpha = (\sinh \chi) a^\alpha$$
$$t^4 = \cosh \chi$$

and similarly for $\hat{\mathbf{t}}$:

$$\hat{t}^\alpha = (\sinh \hat{\chi}) \hat{a}^\alpha$$
$$\hat{t}^4 = \cosh \hat{\chi}.$$

Therefore

$$
\begin{aligned}
-\mathbf{t} \cdot \hat{\mathbf{t}} &= -\sinh \chi \sinh \hat{\chi} \, a^\alpha \hat{a}^\alpha + \cosh \chi \cosh \hat{\chi} \\
&= -\sinh \chi \sinh \hat{\chi} (\cos \psi) + \cosh \chi \cosh \hat{\chi} \text{ from } (*) \\
&= \cosh \chi \cosh \hat{\chi} (\cos^2(\psi/2) + \sin^2(\psi/2)) \\
&\quad - \sinh \chi \sinh \hat{\chi} (\cos^2(\psi/2) - \sin^2(\psi/2)) \\
&\quad \text{since } \cos \psi = \cos^2(\psi/2) - \sin^2(\psi/2) \\
&= \cos^2(\psi/2)(\cosh \chi \cosh \hat{\chi} - \sinh \chi \sinh \hat{\chi}) \\
&\quad + \sin^2(\psi/2)(\cosh \chi \cosh \hat{\chi} + \sinh \chi \sinh \hat{\chi}) \\
&= \cos^2(\psi/2) \cosh(\chi - \hat{\chi}) + \sin^2(\psi/2) \cosh(\chi + \hat{\chi}) \\
&> 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (**)
\end{aligned}
$$

since $\cosh(x \pm y) = \cosh x \cosh y \pm \sinh x \sinh y$. Since $\psi \in [0, \pi]$, then we can find $x \in [0, 1]$ such that $x = \sin(\psi/2)$. Then from $(**)$ we get

$$y \equiv -\mathbf{t} \cdot \hat{\mathbf{t}} = \cosh(\chi - \hat{\chi})(1 - x^2) + \cosh(\chi + \hat{\chi})(x^2)$$

$$= [\cosh(\chi + \hat{\chi}) - \cosh(\chi - \hat{\chi})]x^2 + \cosh(\chi - \hat{\chi})$$
$$\equiv f(x).$$

So we are interested in the graph of f for $x \in [0, 1]$. We consider three cases:

Case 1. If $\cosh(\chi + \hat{\chi}) - \cosh(\chi - \hat{\chi}) > 0$ then $f'(x) > 0$ and $f(0) \leq f(x) \leq f(1)$, or

$$\cosh(\chi - \hat{\chi}) \leq y \leq \cosh(\chi + \hat{\chi}).$$

Case 2. If $\cosh(\chi + \hat{\chi}) - \cosh(\chi - \hat{\chi}) < 0$ then $f'(x) < 0$ and $f(0) \geq f(x) \geq f(1)$ or

$$\cosh(\chi - \hat{\chi}) \leq y \leq \cosh(\chi + \hat{\chi}).$$

Case 3. If $\cosh(\chi + \hat{\chi}) - \cosh(\chi - \hat{\chi}) = 0$, then $f'(x) = 0$ and $f(0) = f(x) = f(1)$ or

$$\cosh(\chi + \hat{\chi}) = y = \cosh(\chi - \hat{\chi}).$$

In each case,

$$\min\{\cosh(\chi + \hat{\chi}), \cosh(\chi - \hat{\chi})\} \leq y \leq \max\{\cosh(\chi + \hat{\chi}), \cosh(\chi - \hat{\chi})\}.$$

Since the range of $\cosh x$ is $[1, \infty)$, then $y \in [1, \infty)$ and so $-y \in (-\infty, -1]$ or $-\infty < \mathbf{t} \cdot \hat{\mathbf{t}} \leq -1$. ∎

Corollary 1.1.2. Let $\mathbf{t}, \hat{\mathbf{t}}$ be two timelike, future pointing unit vectors such that $\mathbf{t} \cdot \hat{\mathbf{t}} = -1$. Then $\mathbf{t} = \hat{\mathbf{t}}$.

Proof. We use the notation of Theorem 1.1.5 and consider the cases given there again (respectively).

Case 1. Suppose $\cosh(\chi + \hat{\chi}) > \cosh(\chi - \hat{\chi})$. Then the minimum value of $y = f(x) = -\mathbf{t} \cdot \hat{\mathbf{t}}$ occurs at $x = 0$ and the minimum is $\cosh(\chi - \hat{\chi}) = 1$ (since we hypothesized $\mathbf{t} \cdot \hat{\mathbf{t}} = -1$), and $\psi = 0$ and $\chi = \hat{\chi}$. Since ψ is the angle between $\mathbf{a}$ and $\hat{\mathbf{a}}$ then $a^\rho = \hat{a}^\rho$ (i.e. the spatial components of $\mathbf{t}$ and $\hat{\mathbf{t}}$ are the same). Since $t^4 = \cosh \chi$ and $\hat{t}^4 = \cosh \hat{\chi}$, then $t^4 = \hat{t}^4$ and $\mathbf{t} = \hat{\mathbf{t}}$.

Case 2. Suppose $\cosh(\chi + \hat{\chi}) < \cosh(\chi - \hat{\chi})$. Then the minimum value of $y = f(x) = -\mathbf{t} \cdot \hat{\mathbf{t}}$ occurs at $x = 1$ and the minimum is $\cosh(\chi + \hat{\chi}) = 1$. So $\chi = -\hat{\chi}$ and since $x = \sin(\psi/2) = 1$, then $\psi = \pi$. Since ψ is the angle between $\mathbf{a}$ and $\hat{\mathbf{a}}$, then $\hat{\theta} = \pi - \theta$ and $\phi = \hat{\phi} + \pi$ (since $\cos(\pi - \theta) = -\cos \theta$, $\sin(\pi - \theta) = \sin \theta$, $\cos(\phi + \pi) = -\cos \phi$, and $\sin(\phi + \pi) = -\sin \phi$, $a^\rho = -a^\rho$ and from the equations on page 5). Thus

$$
\begin{aligned}
t^\alpha &= (\sinh \chi) a^\rho \\
&= (\sinh \chi)(-\hat{a}^\alpha) \text{ since } a^\alpha = -\hat{a}^\alpha \\
&= (-\sinh \hat{\chi})(-\hat{a}^\alpha) \text{ since } \sinh \text{ is an odd function} \\
&= (\sinh \hat{\chi})(\hat{a}^\alpha) = \hat{t}^\alpha
\end{aligned}
$$

and $t^4 = \cosh \chi = \cosh \hat{\chi} = \hat{t}^4$ since $\cosh$ is an even function. So

$$\mathbf{t} = \hat{\mathbf{t}}.$$

Case 3. Suppose $\cosh(\chi + \hat{\chi}) = \cosh(\chi - \hat{\chi})$. Then the minimum value of $y = f(x) = -\mathbf{t} \cdot \hat{\mathbf{t}}$ occurs for all $x \in [0, 1]$ since f is constant. Now $f(0) = \cosh(\chi - \hat{\chi}) = 1$, so $\chi = \hat{\chi}$, and $f(1) = \cosh(\chi + \hat{\chi}) \equiv 1$, so $\chi + \hat{\chi} = 0$ and it follows that $\chi = \hat{\chi} = 0$. So $t^\alpha = (\sinh 0)a^\alpha = 0 = (\sinh 0)\hat{a}^\alpha = \hat{t}^\alpha$ and $t^4 = \cosh 0 = 1 = \cosh 0 = \hat{t}^4$. Hence $\mathbf{t} = \hat{\mathbf{t}}$.

∎

Theorem 1.1.6. A timelike vector cannot be M-orthogonal to a nonzero null vector.

Proof. Suppose $\mathbf{t}$ is timelike. Then $\mathbf{t} \cdot \hat{\mathbf{t}} = t^\alpha t^\alpha - (t^4)^2 < 0$ and $t^\alpha t^\alpha < (t^4)^2$, where $t^4 \neq 0$. Suppose $\mathbf{n}$ is a nonzero null vector. Then $\mathbf{n} \cdot \mathbf{n} = n^\alpha n^\alpha - (n^4)^2 = 0$ and $n^\alpha n^\alpha = (n^4)^2$, where $n^4 \neq 0$. So

$$\begin{aligned} (t^\alpha n^\alpha)^2 &\leq t^\alpha t^\alpha n^\beta n^\beta \text{ by the Schwarz Inequality} \\ &< (t^4 n^4)^2. \end{aligned} \tag{1.1.15}$$

Suppose $\mathbf{t} \cdot \mathbf{n} = 0$, or $t^\alpha n^\alpha - t^4 n^4 = 0$. Then $(t^\alpha n^\alpha)^2 = (t^4 n^4)^2$, a contradiction to (1.1.15). Hence $\mathbf{t} \cdot \mathbf{n} \neq 0$. ∎

Note. A spacelike vector *can* be M-orthogonal to a nonzero null vector. Consider $\mathbf{e}_2$ and $\mathbf{e}_1 + \mathbf{e}_4$.

Theorem 1.1.7. Two nonzero null vectors are M-orthogonal if and only if they are scalar multiples of each other.

Proof. First, if $\mathbf{m}$ and $\mathbf{n}$ are null vectors and $\mathbf{m} = \lambda\mathbf{n}$ for some $\lambda \in \mathbb{R}$, then $\mathbf{m} \cdot \mathbf{n} = (\lambda\mathbf{n}) \cdot \mathbf{n} = \lambda\mathbf{n} \cdot \mathbf{n} = 0$.

Second, suppose $\mathbf{m}$ and $\mathbf{n}$ are nonzero M-orthogonal null vectors. Then

$$\mathbf{m} \cdot \mathbf{m} = m^\alpha m^\alpha - (m^4)^2 = 0$$

$$\mathbf{n} \cdot \mathbf{n} = n^\alpha n^\alpha - (n^4)^2 = 0$$

$$\mathbf{m} \cdot \mathbf{n} = m^\alpha n^\alpha - (m^4 n^4) = 0, \quad m^4 \neq 0, \quad n^4 \neq 0.$$

Therefore $(m^\alpha n^\alpha)^2 = (m^4 n^4)^2 = m^\alpha m^\alpha n^\beta n^\beta$ and $|m^\alpha n^\alpha| = \sqrt{m^\alpha m^\alpha n^\beta n^\beta}$. But then we have equality in the Schwarz Inequality for $\mathbf{m}$ and $\mathbf{n}$. Hence $m^\alpha = \lambda n^\alpha$ for some scalar $\lambda \neq 0$. Since $n^4 \neq 0$, we have $m^4 = m^\alpha n^\alpha / n^4 = (\lambda n^\alpha) n^\alpha / n^4 = \lambda (n^4)^2 / n^4 = \lambda n^4$. Thus $\mathbf{m} = \lambda\mathbf{n}$. ∎

Note. Let's try to give some geometric interpretation to the above results. First, consider the 3-dimensional subspace span$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4\}$ of $\mathbb{V}_4$. Suppose $\mathbf{n}$ is a null vector in this subspace, $\mathbf{n} = n^1 \mathbf{e}_1 + n^2 \mathbf{e}_2 + n^4 \mathbf{e}_4$. Then $\mathbf{n} \cdot \mathbf{n} = n^1 n^1 + n^2 n^2 - n^4 n^4 = 0$ and $(n^1)^2 + (n^2)^2 = (n^4)^2$. If we plot such a vector in a 3-D coordinate system in standard position (i.e. with tail at the origin), then the head of the vector (with coordinates (n^1, n^2, n^4) — that is, the vector points from the event $(0, 0, 0)$ to the event (n^1, n^2, n^4)) lies on the surface $x^2 + y^2 = z^2$, or $\sqrt{x^2 + y^2} = |z|$. That is, the set of all null vectors lie on these infinite cones:
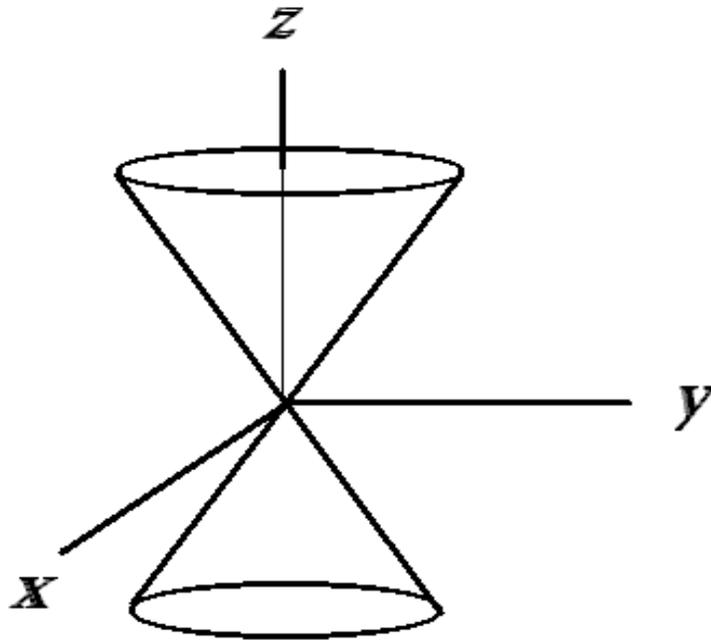
If $\mathbf{t}$ is timelike and forward pointing, then $\mathbf{t} \cdot \mathbf{t} = t^1 t^1 + t^2 t^2 - t^4 t^4 < 0$ and $(t^1)^2 + (t^2)^2 < (t^4)^2$ and the vector lies inside the upper cone. If $\mathbf{t}$ is timelike and backward pointing, then it lies in the lower cone. Any spacelike vector lies outside of the cones.

Note. The above comment suggests why Theorem 1.1.4 holds: two time-like vectors cannot be M-orthogonal since they must either lie in the same cone (and be separated by less than 90°) or lie in different cones (and be separated by more than 90°). This also explains Theorem 1.1.6 (a time-like vector cannot be M-orthogonal to a nonzero null vector) since the null vectors lie on the surfaces of the cones and the timelike vectors lie on the interiors.

Note. The above model must still be taken rather lightly. It does not give insight as to why null vectors are M-orthogonal if and only if they are scalar multiples of one another. In fact, $\mathbf{e}_1 + \mathbf{e}_4$ and $-\mathbf{e}_1 + \mathbf{e}_4$ are perpendicular in our model (the text uses the term "Euclidean orthogonal"), but not M-orthogonal since:

$$(\mathbf{e}_1 + \mathbf{e}_4) \cdot (-\mathbf{e}_1 + \mathbf{e}_4) = (1)(-1) - (1)(1) = -2.$$