1. Minkowski Spaces

<u>Def</u> Define the <u>Minkowski product</u> $\circ$ on $\mathbb{R}^n$ by $x \circ y = x_1 y_1 + \cdots + x_{n-1} y_{n-1} - x_n y_n$ for $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Call the pair $(\mathbb{R}^n, \circ)$ <u>Minkowski n-space</u> and denote it by $\mathbb{M}^n$.

Observe that the Minkowski product is a <u>symmetric bilinear form</u>. In other words, $(ax + by) \circ z = a(x \circ z) + b(y \circ z)$, $x \circ (ay + bz) = a(x \circ y) + b(x \circ z)$ and $x \circ y = y \circ x$ for all $x, y, z \in \mathbb{R}^n$ and all $a, b \in \mathbb{R}$.

<u>Def</u> For $x \in \mathbb{M}^n$, call x <u>timelike</u> if $x \circ x < 0$, <u>spacelike</u> if $x \circ x > 0$, and <u>null</u> if $x \circ x = 0$. Call x <u>degenerate</u> if $x \circ y = 0$ for all $y \in \mathbb{M}^n$.

<u>Lemma 1.1.</u> 0 is the only degenerate element of $\mathbb{M}^n$.

<u>Proof</u> If $x = (x_1, \ldots, x_n) \in \mathbb{M}^n$, let $\bar{x} = (x_1, \ldots, x_{n-1}, -x_n)$. If $x \neq 0$, then $x \circ \bar{x} = \|x\|^2 \neq 0$. $\square$

Lemma 1.2 If $x, y \in \mathbb{M}^n$ are non-zero and non-spacelike and $x \circ y = 0$, then x and y are null and $y_n x = x_n y$.

Proof Let $x = (x_1 \cdots x_n)$, $y = (y_1 \cdots y_n)$, $u = (x_1 \cdots x_{n-1}\ 0)$ and $v = (y_1 \cdots y_{n-1}\ 0)$. Then $0 \geq x \circ x = \|u\|^2 - x_n^2$, $0 \geq y \circ y = \|v\|^2 - y_n^2$ and $0 = x \circ y = u \circ v - x_n y_n$. Hence,

$|x_n| \geq \|u\|$, $|y_n| \geq \|v\|$ and $x_n y_n = u \circ v$.

Thus, $|u \circ v| = |x_n|\,|y_n| \geq \|u\|\,\|v\|$.

The Cauchy Inequality implies $|u \circ v| \leq \|u\|\,\|v\|$. Therefore $|x_n|\,|y_n| = |u \circ v| = \|u\|\,\|v\|$.

Since $|x_n| \geq \|u\|$ and $|y_n| \geq \|v\|$, then

$|x_n|\,|y_n| \geq \|u\|\,|y_n| \geq \|u\|\,\|v\| = |x_n|\,|y_n|$ and

$|x_n|\,|y_n| \geq |x_n|\,\|v\| \geq \|u\|\,\|v\| = |x_n|\,|y_n|$.

Hence $|x_n|\,|y_n| = \|u\|\,|y_n| = |x_n|\,\|v\| = \|u\|\,\|v\|$.

Therefore, $(|x_n| - \|u\|)(|y_n| - \|v\|) =$

$(|x_n|\,|y_n| - \|u\|\,y_n) - (|x_n|\,\|v\| - \|u\|\,\|v\|) = 0$.

Thus, either $|x_n| - \|u\| = 0$ or $|y_n| - \|v\| = 0$.

So either $|x_n| = \|u\|$ or $|y_n| = \|v\|$.

Suppose $|x_n| = \|u\|$. Since $x \neq 0$ and $x = u + x_n e_n$, then $x_n \neq 0$. Since $|x_n||y_n| = |x_n|\|v\|$, then $|y_n| = \|v\|$. Similarly, $|y_n| = \|v\|$ implies $|x_n| = \|u\|$. Thus, both $|x_n| = \|u\|$ and $|y_n| = \|v\|$. Hence, $x \circ x = \|u\|^2 - x_n^2 = 0$ and $y \circ y = \|v\|^2 - y_n^2 = 0$. Therefore, x and y are null.

Since $|u \cdot v| = \|u\|\|v\|$, then there is an $\varepsilon = \pm 1$ such that $u \cdot v = \varepsilon \|u\|\|v\|$. Then the Cauchy Inequality implies $\|v\|u = \varepsilon\|u\|v$. Thus, $(u \cdot v)u = \varepsilon\|u\|\|v\|u = \varepsilon^2\|u\|^2 v = \|u\|^2 v$. Since $x_n y_n = u \cdot v$ and $x_n^2 = \|u\|^2$, we have $x_n y_n u = x_n^2 v$. Since $x_n \neq 0$, then $y_n u = x_n v$. Since $y_n x_n = x_n y_n$, it follows that $y_n x = x_n y$. $\square$

<u>Corollary 1.3.</u> If $x, y \in M^n$ such that $x \circ x < 0$, $x \circ y = 0$ and $y \neq 0$, then $y \circ y > 0$.

<u>Proof</u> Suppose $x \circ x < 0$, $x \circ y = 0$ and $y \neq 0$. If $y \circ y \leq 0$, then Lemma 1.2 implies $x \circ x = 0$, a contradiction. Hence, $y \circ y > 0$. $\square$

<u>Def</u> Let $T^n = \{x \in \mathbb{M}^n : x \circ x < 0\}$.
Call T^n the <u>open time cone</u> in $\mathbb{M}^n$.
Observe that T^n has two components:
$T_+^n = \{x \in T^n : x_n > 0\}$ and $T_-^n = \{x \in T^n : x_n < 0\}$.
T_+^n and T_-^n are disjoint non-empty open convex subsets of $\mathbb{M}^n$ and $T^n = T_+^n \cup T_-^n$.

<u>Def</u> For $x \in \mathbb{M}^n$, let $|||x||| = \sqrt{|x \circ x|}$.

<u>Theorem 1.4</u> <u>The Cauchy Inequality for T^n</u>.

a) If x, y belong to the same component of T^n, then $x \circ y \leq -|||x|||\,|||y|||$. Furthermore, if $x \circ y = -|||x|||\,|||y|||$, then $|||y|||x = |||x|||y$.

b) If x, y belong to different components of T^n, then $x \circ y \geq |||x|||\,|||y|||$. Furthermore, if $x \circ y = |||x|||\,|||y|||$, then $|||y|||x = -|||x|||y$.

<u>Proof of a)</u>. Let x, y belong to the same component of T^n. Let $u = x/|||x|||$ and $v = y/|||y|||$. Then $u \circ u = v \circ v = -1$ and u and v belong to the same component of T^n as x and y.
Write $u = (u_1, \dots u_n)$ and $v = (v_1, \dots v_n)$.
Let $w = (u_1, \dots u_{n-1}, 0)$ and $z = (v_1, \dots v_{n-1}, 0)$.
Then $-1 = u \circ u = \|w\|^2 - u_n^2$ and
$-1 = v \circ v = \|z\|^2 - v_n^2$. So

1.5

$u_n^2 = \|w\|^2 + 1$ and $v_n^2 = \|z\|^2 + 1$. Thus,

$$u_n^2 v_n^2 = \|w\|^2 \|z\|^2 + (\|w\|^2 + \|z\|^2) + 1.$$

The Cauchy Inequality implies $\|w\| \|z\| \geq w \cdot z$. Also since $0 \leq \|w - z\|^2 = \|w\|^2 - 2w\cdot z + \|z\|^2$, then $\|w\|^2 + \|z\|^2 \geq 2w\cdot z$. Hence,

$$u_n^2 v_n^2 \geq (w\cdot z)^2 + 2(w\cdot z) + 1 = (w\cdot z + 1)^2.$$

Since u and v belong to the same component of T^n, then $u_n v_n > 0$.

Therefore, $u_n v_n \geq w\cdot z + 1$.

Thus $u \circ v = w\cdot z - u_n v_n \leq -1$.

Hence, $x \circ y \leq -|||x|||\,|||y|||$.

Now assume $x \circ y = -|||x|||\,|||y|||$.

Then $u \circ v = -1$. Thus, $w\cdot z - u_n v_n = -1$.

Therefore $(u_n v_n)^2 = (w\cdot z)^2 + 2w\cdot z + 1$.

On the other hand, since $u_n^2 = \|w\|^2 + 1$ and $v_n^2 = \|z\|^2 + 1$, then

$$(u_n v_n)^2 = \|w\|^2 \|z\|^2 + (\|w\|^2 + \|z\|^2) + 1$$

Thus $(w\cdot z)^2 + 2w\cdot z = \|w\|^2 \|z\|^2 + (\|w\|^2 + \|z\|^2)$.

Since $(w\cdot z)^2 \leq \|w\|^2 \|z\|^2$ by the Cauchy Inequality, then $2w\circ z \geq \|w\|^2 + \|z\|^2$. Thus

$$0 \geq \|w\|^2 - 2w\circ z + \|z\|^2 = \|w-z\|^2 \geq 0.$$

Hence, $w = z$. Therefore $u_n^2 = v_n^2$. Since $u_n v_n > 0$, then $u_n = v_n$. Hence $u = v$.

Therefore, $|||y|||\,x = |||x|||\,y$. □

Homework Problem 1.1 - Prove Theorem 1.4.b.

Homework Problem 1.2

Prove Corollary 1.5

Corollary 1.5. Let $x, y \in T^n$. If x, y belong to the same component / opposite component of T^n, then $x \circ y < 0$ / $x \circ y > 0$.

Homework Problem 1.3. Suppose $x, y \in \mathbb{M}^n$ are space like. Let V be the vector subspace of $\mathbb{M}^n$ spanned by x and y.

a) Show it is possible for V to contain a null vector.

b) Prove that if V contains no non-zero null vectors, then the restriction of the Minkowski product $\circ$ to V is positive definite. In other words, prove that if V contains no non-zero null vectors, then for every $z \in V$, $z \circ z \geq 0$ and $z \circ z = 0$ if and only if $z = 0$.

<u>Remark</u> There is a weak version of the Cauchy Inequality for spacelike points. If $x, y \in \mathbb{M}^n$ are spacelike and have the property that the vector subspace V of $\mathbb{M}^n$ spanned by x and y contains no null vectors, then $|x \circ y| \leq |||x|||\,|||y|||$ (and $x \circ y = \varepsilon |||x|||\,|||y|||$ where $\varepsilon = \pm 1$ implies $|||y|||x = \varepsilon |||x|||y$).

~~Hence, $\|x\| = \sqrt{x \circ x}$ [illegible]~~

The justification for this assertion is that if V contains no null vectors, then according to Homework Problem 1.2.b, the restriction of the Minkowski product $\circ$ to V is positive definite. It follows that the Euclidean version of the Cauchy Inequality holds in V.

<u>Def</u> An element $u \in M^n$ is a <u>unit vector</u> if $u \circ u = \pm 1$. A sequence $u_1, \dots, u_k \in M^n$ is <u>orthogonal</u> if $u_i \circ u_j = 0$ for $i \neq j$. An orthogonal sequence of unit vectors in M^n is called an <u>orthonormal sequence</u>. An orthonormal sequence in M^n which is also a basis for the vector space structure on $\mathbb{R}^n$ is called an <u>orthonormal basis</u> for M^n.

Observe that the standard basis $e_1, \dots, e_n$ for $\mathbb{R}^n$ is an orthonormal basis for M^n.

<u>Def</u> If $u_1, \dots, u_n$ is an orthonormal basis for M^n such that $u_i \circ u_i = +1$ for $1 \le i \le n-1$ and $u_n \circ u_n = -1$, then we say that $u_1, \dots, u_n$ is <u>standardly ordered</u>.

<u>Observation</u> An orthonormal basis with respect to the dot product on $\mathbb{R}^n$ need not be orthonormal in M^n, and vice versa. For instance, $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ is an orthonormal basis for $\mathbb{R}^2$ with the dot product, but $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) \circ (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) = 1$. Also, $(2, \sqrt{3}), (\sqrt{3}, 2)$ is an orthonormal basis for M^2, but $(2, \sqrt{3}) \cdot (\sqrt{3}, 2) = 4\sqrt{3}$.

Lemma 1.6. Every orthonormal sequence in $\mathbb{M}^n$ is linearly independent.

Proof. Suppose $u_1, \dots, u_k$ is an orthonormal sequence in $\mathbb{M}^n$ and $a_1, \dots, a_k \in \mathbb{R}$ such that $\sum_{i=1}^k a_i u_i = 0$. Then for $1 \le j \le k$,

$$0 = 0 \circ u_j = \left(\sum_{i=1}^k a_i u_i\right) \circ u_j = \sum_{i=1}^k a_i (u_i \circ u_j) = a_j(\pm 1)$$

Hence $a_j = 0$ for $1 \le j \le k$. $\square$

Theorem 1.7. a) Every orthonormal sequence in $\mathbb{M}^n$ has at most one timelike element.

b) Every orthonormal basis for $\mathbb{M}^n$ has exactly one timelike element.

Proof of a). Suppose $u_1, \dots, u_k$ is an orthonormal sequence in $\mathbb{M}^n$ such that u_1 is timelike. Let $2 \le i \le k$. Since $u_1 \circ u_1 < 0$, $u_1 \circ u_i = 0$ and $u_i \neq 0$, then Corollary 1.3 implies $u_i \circ u_i > 0$. So u_i is spacelike. $\square$

Proof of b). Let $u_1, \dots, u_n$ be an orthonormal basis for $\mathbb{M}^n$. Part a) implies at most one u_i is timelike. Assume no u_i is timelike. Then $u_i \circ u_i = +1$ for $1 \le i \le n$. There are $a_1, \dots, a_n \in \mathbb{R}$ such that

$e_n = \sum_{i=1}^n a_i u_i$. Then $-1 = e_n \circ e_n = \left(\sum_{i=1}^n a_i u_i\right) \circ \left(\sum_{j=1}^n a_j u_j\right)$
$= \sum_{i=1}^n a_i^2 (u_i \circ u_i) = \sum_{i=1}^n a_i^2 \geq 0$. We've reached a contradiction. Hence, at least one u_i is timelike. $\square$.

<u>Def</u> The <u>orthogonal group</u> of $\mathbb{M}^n$, denoted $O(\mathbb{M}^n)$, is the set of all Minkowski product preserving functions from $\mathbb{M}^n$ to itself. Thus, $f: \mathbb{M}^n \to \mathbb{M}^n$ is an element of $O(\mathbb{M}^n)$ if and only if $f(x) \circ f(y) = x \circ y$ for all $x, y \in \mathbb{M}^n$.

<u>Theorem 1.8.</u> Let $f: \mathbb{M}^n \to \mathbb{M}^n$. Then $f \in O(\mathbb{M}^n)$ if and only if f is linear and $f(e_1), \ldots, f(e_n)$ is a standardly ordered orthonormal basis for $\mathbb{M}^n$.

<u>Proof</u> Assume $f \in O(\mathbb{M}^n)$. Since $e_1, \ldots, e_n$ is a standardly ordered orthonormal basis for $\mathbb{M}^n$ and f preserves the Minkowski product, then $f(e_1), \ldots, f(e_n)$ is clearly a standardly ordered orthonormal basis for $\mathbb{M}^n$.

We assert: if $x = \sum_{i=1}^n x_i e_i \in \mathbb{M}^n$, then $f(x) = \sum_{i=1}^n x_i f(e_i)$. Since $f(e_1), \ldots, f(e_n)$ is a basis for $\mathbb{M}^n$, then there are $y_1, \ldots, y_n \in \mathbb{R}$

such that $f(x) = \sum_{i=1}^n y_i f(e_i)$. Hence, for $1 \le j \le n$:

$$f(x) \circ f(e_j) = x \circ e_j = \left(\sum_{i=1}^n x_i e_i\right) \circ e_j = \sum_{i=1}^n x_i (e_i \circ e_j) = x_j (e_j \circ e_j)$$

Also

$$f(x) \circ f(e_j) = \left(\sum_{i=1}^n y_i f(e_i)\right) \circ f(e_j) = \sum_{i=1}^n y_i (f(e_i) \circ f(e_j)) =$$

$$\sum_{i=1}^n y_i (e_i \circ e_j) = y_j (e_j \circ e_j).$$

Thus, $x_j (e_j \circ e_j) = y_j (e_j \circ e_j)$. Since $e_j \circ e_j \ne 0$, then $x_j = y_j$ for $1 \le j \le n$. This proves $f(x) = \sum_{i=1}^n x_i f(e_i)$.

To prove f is linear, let $x = \sum_{i=1}^n x_i e_i$ and $y = \sum_{i=1}^n y_i e_i \in M^n$ and let $a \in R$. Then $x + y = \sum_{i=1}^n (x_i + y_i) e_i$ and $ax = \sum_{i=1}^n (a x_i) e_i$. Hence, $f(x+y) = \sum_{i=1}^n (x_i + y_i) f(e_i) = \sum_{i=1}^n x_i f(e_i) + \sum_{i=1}^n y_i f(e_i)$ $= f(x) + f(y)$, and $f(ax) = \sum_{i=1}^n (a x_i) f(e_i) = a\left(\sum_{i=1}^n x_i f(e_i)\right) = a f(x)$.

Now assume f is linear and $f(e_1), \dots, f(e_n)$ is a standardly ordered orthonormal basis for M^n. Let $x = \sum_{i=1}^n x_i e_i$ and $y = \sum_{i=1}^n y_i e_i \in M^n$. Then:

$$f(x) \circ f(y) = \left(\sum_i x_i f(e_i)\right) \circ \left(\sum_j y_j f(e_j)\right) = \sum_{i,j} x_i y_j \, f(e_i) \circ f(e_j) =$$

$$\sum_{i,j} x_i y_j (e_i \circ e_j) = \left(\sum_i x_i e_i\right) \circ \left(\sum_j y_j e_j\right) = x \circ y.$$ Thus, $f \in O(M^n)$. $\square$

Theorem 1.9. $O(\mathbb{M}^n)$ is a group with respect to composition. In other words,

a) $id_{\mathbb{M}^n} \in O(\mathbb{M}^n)$

b) $f, g \in O(\mathbb{M}^n)$ implies $g \circ f \in O(\mathbb{M}^n)$

c) $f \in O(\mathbb{M}^n)$ implies f is invertible and $f^{-1} \in O(\mathbb{M}^n)$.

Proof of a) For $x, y \in \mathbb{M}^n$, $id(x) \circ id(y) = x \circ y$. Thus $id \in O(\mathbb{M}^n)$

Proof of b) Let $f, g \in O(\mathbb{M}^n)$. For $x, y \in \mathbb{M}^n$, $(g \circ f)(x) \circ (g \circ f)(y) = g(f(x)) \circ g(f(y)) = f(x) \circ f(y) = x \circ y$. Thus $g \circ f \in O(\mathbb{M}^n)$.

Proof of c) Let $f \in O(\mathbb{M}^n)$. Since $f : \mathbb{M}^n \to \mathbb{M}^n$ is linear by Theorem 1.7, it suffices to prove $Ker(f) = \{0\}$ to show that f is invertible. Let $x \in Ker(f)$. Then for every $y \in \mathbb{M}^n$, $x \circ y = f(x) \circ f(y) = 0 \circ f(y) = 0$. Thus, x is degenerate. Hence $x = 0$ by Lemma 1.1. Therefore, $Ker(f) = \{0\}$. So f is invertible.

To prove $f^{-1} \in O(\mathbb{M}^n)$, let $x, y \in \mathbb{M}^n$. Then $f^{-1}(x) \circ f^{-1}(y) = f(f^{-1}(x)) \circ f(f^{-1}(y)) = x \circ y$. Thus $f^{-1} \in O(\mathbb{M}^n)$. $\square$.

~~**Def.** Let $T^n = \{x \in \mathbb{M}^n : x \circ x < 0\}$. Call T^n the *open time cone* of $\mathbb{M}^n$. Observe that T^n has two components: $T^n_+ = \{x \in T^n : x_n > 0\}$ and $T^n_- = \{x \in T^n : x_n < 0\}$. T^n_+ and T^n_- are disjoint non-empty open convex subsets of $\mathbb{M}^n$, and $T^n = T^n_+ \cup T^n_-$.~~

¶ Observe that if $f \in O(\mathbb{M}^n)$, then $f(T^n) = T^n$. Hence, either $f(T^n_+) = T^n_+$ and $f(T^n_-) = T^n_-$, or $f(T^n_+) = T^n_-$ and $f(T^n_-) = T^n_+$.

Def Let $f \in O(\mathbb{M}^n)$. If $f(T^n_+) = T^n_+$ and $f(T^n_-) = T^n_-$, we say f is *time orientation preserving* (or more briefly *time preserving*). If $f(T^n_+) = T^n_-$ and $f(T^n_-) = T^n_+$, we say f is *time orientation reversing* (or more briefly *time reversing*).

Let $O^+(\mathbb{M}^n) = \{f \in O(\mathbb{M}^n) : f \text{ is time preserving}\}$.

Corollary 1.10. $O^+(\mathbb{M}^n)$ is an index 2 subgroup of $O(\mathbb{M}^n)$.

Homework Problem 1.4. Prove Corollary 1.10.

<u>Theorem 1.10.</u> Let $f : \mathbb{M}^n \to \mathbb{M}^n$ be a linear function. Let A be the $n \times n$ matrix representing f. (For $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $f(x) = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.)

Let $J = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}$. Then the following are equivalent.

a) $f \in O(\mathbb{M}^n)$

b) The columns of A are a standardly ordered orthonormal basis for $\mathbb{M}^n$.

c) $A^t J A = J$

d) $A J A^t = J$

e) The rows of A are a standardly ordered orthonormal basis for $\mathbb{M}^n$.

Also: $f \in O^+(\mathbb{M}^n)$ if and only if $a_{nn} > 0$.

<u>Homework Problem 1.5.</u> Prove Theorem 1.10.

<u>Hint</u> Recall the following fact about $n \times n$ matrices: if $AB = I$, then $BA = I$.

Def Let u be a unit vector in $\mathbb{M}^n$. Define the linear reflection $Z_u : \mathbb{M}^n \to \mathbb{M}^n$ by

$$Z_u(x) = x - 2\left(\frac{x \circ u}{u \circ u}\right)u$$

for $x \in \mathbb{M}^n$.

Theorem 1.12. Let u be a unit vector in $\mathbb{M}^n$. Then:

a) $Z_u \in O(\mathbb{M}^n)$

b) $Z_u^{-1} = Z_u$

c) For $x \in \mathbb{M}^n$, $Z_u(x) = x$ if and only if $z \circ u = 0$.

d) For any unit vector v in $\mathbb{M}^n$, $Z_u = Z_v$ if and only if $u = \pm v$.

e) If $y, z \in \mathbb{M}^n$ such that $y \circ y = z \circ z$ and $|||y - z||| \neq 0$, and if $v = (y-z)/|||y-z|||$, then v is a unit vector and $Z_v(y) = z$.

f) If $y, z \in \mathbb{M}^n$ such that $y \neq z$ and $Z_u(y) = z$, then $|||y - z||| \neq 0$ and $u = \pm(y-z)/|||y-z|||$.

g) $Z_u \in O^+(\mathbb{M}^n)$ if and only if u is spacelike ($u \circ u > 0$). Hence, $Z_u \notin O^+(\mathbb{M}^n)$ if and only if u is timelike ($u \circ u < 0$).

Proof of a) For $x, y \in \mathbb{M}^n$,

$$Z_u(x) \circ Z_u(y) = \left(x - 2\left(\frac{x \circ u}{u \circ u}\right)u\right) \circ \left(y - 2\left(\frac{y \circ u}{u \circ u}\right)u\right) =$$

$$x \circ y - 4\,\frac{(x \circ u)(y \circ u)}{u \circ u} + 4\,\frac{(x \circ u)(y \circ u)}{(u \circ u)^2}(u \circ u) = x \circ y.$$

Thus, $Z_u \in O(\mathbb{M}^n)$. $\square$

Proof of b) For $x \in \mathbb{M}^n$:

$$Z_u \circ Z_u(x) = Z_u(x) - 2\left(\frac{Z_u(x) \circ u}{u \circ u}\right)u =$$

$$\left(x - 2\left(\frac{x \circ u}{u \circ u}\right)u\right) - 2\left(\frac{\left(x - 2\left(\frac{x \circ u}{u \circ u}\right)u\right) \circ u}{u \circ u}\right)u =$$

$$x - 4\left(\frac{x \circ u}{u \circ u}\right)u + 4\,\frac{(x \circ u)(u \circ u)}{(u \circ u)^2}\,u = x.$$

Thus, $Z_u \circ Z_u = id$. Hence, $Z_u^{-1} = Z_u$. $\square$

Proof of c) For $x \in \mathbb{M}^n$, the following statements are equivalent. $Z_u(x) = x$,
$x - 2\left(\frac{x \circ u}{u \circ u}\right)u = x$, $2\left(\frac{x \circ u}{u \circ u}\right)u = 0$, $x \circ u = 0$. $\square$

Proof of d) First observe that

$$Z_{-u}(x) = x - 2\left(\frac{x \circ (-u)}{(-u) \circ (-u)}\right)(-u) = x - 2\left(\frac{x \circ u}{u \circ u}\right)u = Z_u(x).$$

Next assume v is a unit vector such that $Z_u = Z_v$. Then $Z_u(u) = Z_v(u)$.

$$Z_u(u) = u - 2\left(\frac{u \circ u}{u \circ u}\right)u = u - 2u = -u.$$

$$Z_v(u) = u - 2\left(\frac{u \circ v}{v \circ v}\right)v.$$ Hence, $$-u = u - 2\left(\frac{u \circ v}{v \circ v}\right)v.$$

Thus, $2u = 2\left(\frac{u \circ v}{v \circ v}\right)v$. So $u = \left(\frac{u \circ v}{v \circ v}\right)v$.

Therefore $u \circ u = \left(\frac{u \circ v}{v \circ v}\right)^2 (v \circ v)$. Hence, $\frac{u \circ v}{v \circ v} = \pm 1$.

So $u = \pm v$. $\square$

Proof of e) Suppose $y, z \in \mathbb{M}^n$ such that $y \circ y = z \cdot z$ and $|||y-z||| \neq 0$, and suppose $v = (y-z)/|||y-z|||$. Clearly,

$$v \circ v = \frac{(y-z) \circ (y-z)}{|||y-z|||^2} = \pm 1.$$ So v is a unit vector. Since $y \circ y = z \circ z$, then

$$(y-z) \circ (y-z) = y \circ y - 2\, y \circ z + z \circ z = 2\, y \circ y - 2\, y \circ z = 2\, y \circ (y-z).$$ Thus,

$$\frac{y \circ v}{v \circ v} = \frac{(y \circ (y-z)) / |||y-z|||}{(y-z) \circ (y-z) / |||y-z|||^2} = \frac{|||y-z|||\,(y \circ (y-z))}{2\, y \circ (y-z)} = |||y-z|||/2.$$ Hence, $$Z_v(y) = y - 2\left(\frac{y \circ v}{v \circ v}\right)v = y - 2\left(\frac{|||y-z|||}{2}\right)\left(\frac{y-z}{|||y-z|||}\right) = y - (y-z) = z. \quad \square$$

Proof of f). Suppose $y, z \in \mathbb{M}^n$ such that $y \neq z$ and $Z_u(y) = z$. Then $y - 2\left(\frac{y \circ u}{u \circ u}\right) u = z$. Thus, $y - z = 2\left(\frac{y \circ u}{u \circ u}\right) u$.

Therefore, $(y-z)\circ(y-z) = 4\left(\frac{y \circ u}{u \circ u}\right)^2 (u \circ u) = 4\frac{(y \circ u)^2}{u \circ u}$

Since $Z_u(y) = z \neq y$, then $y \circ u \neq 0$, by c).

Hence, $(y-z)\circ(y-z) \neq 0$. Also

$|||y - z||| = \sqrt{|(y-z)\circ(y-z)|} = 2\,|y \circ u|$.

Thus, $\frac{y-z}{|||y-z|||} = \frac{2\left(\frac{y \circ u}{u \circ u}\right) u}{2\,|y \circ u|} = \left(\frac{y \circ u}{|y \circ u|}\right)\left(\frac{1}{u \circ u}\right) u = \pm u$. $\square$

Proof of g) Assume u is spacelike. Then $u \circ u = 1$. So $Z_u(e_n) \circ e_n = \left(e_n - 2\left(\frac{e_n \circ u}{u \circ u}\right) u\right) \circ e_n = -1 - 2(e_n \circ u)^2 < 0$. Thus, the nth coordinate of $Z_u(e_n)$ is positive. Since $e_n \in T^n$, then $Z_u(e_n) \in T^n$. Therefore $Z_u(e_n) \in T_+^n$. Hence, $Z_u(T_+^n) = T_+^n$. So $Z_u \in O^+(\mathbb{M}^n)$.

Now assume u is timelike. Then $u \circ u = -1$.

So $Z_u(e_n) \circ e_n = \left(e_n - 2\left(\frac{e_n \circ u}{u \circ u}\right)u\right) \circ e_n = 2(e_n \circ u)^2 - 1$.

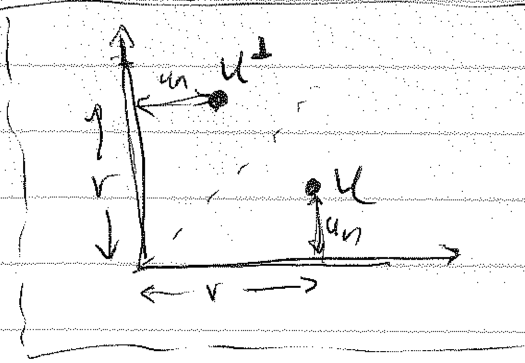
Let $u = (u_1 \cdots u_n)$ and $v = (u_1 \cdots u_{n-1}\, 0)$. Then $u = v + u_n e_n$. Hence, $-1 = u \circ u = \|v\|^2 - u_n^2$. Hence, $u_n^2 = \|v\|^2 + 1$. Also $e_n \circ u = -u_n$.

Thus, $Z_u(e_n) \circ e_n = 2u_n^2 - 1 = 2(\|v\|^2 + 1) - 1 = 2\|v\|^2 + 1 > 0$. Hence, the nth coordinate of $Z_u(e_n)$ is negative. So $Z_u(e_n) \in T^n_-$.

Thus $Z_u(T^n_+) = T^n_-$. Hence, $Z_u \notin O^+(M^n)$.

<u>Alternate Proof of g)</u>. Assume u is spacelike. Then $u \circ u = 1$. Let $u = (u_1 \cdots u_n)$ and $v = (u_1 \cdots u_{n-1}, 0)$. Let $r = \|v\|$. Hence $u = v + u_n e_n$.

Let $u_\perp = u_n\left(\frac{v}{r}\right) + r e_n$.



Then $u_\perp \circ u_\perp = u_n^2 \frac{\|v\|^2}{r^2} - r^2$

$= u_n^2 - r^2 = -(\|v\|^2 - u_n^2)$

$= -u \circ u = -1$. Also the nth coordinate of $u_\perp$ is $r > 0$. Hence, $u_\perp \in T^n_+$.

$u \circ u_\perp = v \cdot \left(u_n\left(\frac{v}{r}\right)\right) - u_n r = u_n \frac{\|v\|^2}{r} - u_n r = 0$.

Hence, $Z_u(u_\perp) = u_\perp$ by c). Since Z_u fixes a point of T^n, then $Z_u(T^n_+) = T^n_+$. Thus $Z_u \in O^+(M^n)$

Assume u is timelike. Thus $u \circ u = -1$. Hence, $u \in T^n$. $Z_u(u) = u - 2\left(\frac{u \circ u}{u \circ u}\right)u = -u$. Hence, $u \in T^n_+$ implies $Z_u(u) \in T^n_-$ and $u \in T^n_-$ implies $Z_u(u) \in T^n_+$. Thus, $Z_u(T^n_+) = T^n_-$. So $Z_u \notin O^+(M^n)$. $\square$

Technical Lemma 1.13.

a) If $x, y \in T^n$, $x \neq y$ and $x \circ x = y \circ y$, then $(x-y) \circ (x-y) \neq 0$. In particular, if x and y lie in the same component of T^n, then $(x-y) \circ (x-y) > 0$; and if x and y lie in different components of T^n, then $(x-y) \circ (x-y) < 0$.

b) If $x, y \in M^n$ are both spacelike, then at least one of $(x-y) \circ (x-y)$ and $(x+y) \circ (x+y)$ is positive.

c) If $x, y \in M^n$, $x \neq y$ and $x \circ e_n = y \circ e_n$, then $(x-y) \circ (x-y) > 0$.

Example Let $x = (1, 0, 0)$ and $y = (1, 1, 1) \in M^3$. Then x and y are spacelike, $x \circ x = y \circ y = 1$, $x \circ e_n \neq y \circ e_n$ and $(x-y) \circ (x-y) = 0$.

Proof of a) Let $x, y \in T^n$ such that $x \neq y$ and $x \circ x = y \circ y$. First assume x and y lie in the same component of T^n. Let $r = |||x||| = |||y|||$. Then $x \circ x = y \circ y = -r^2$. Since $x \neq y$, then $|||y|||\, x \neq |||x|||\, y$. Hence, Theorem 1.4.a implies $x \circ y < -|||x|||\,|||y||| = -r^2$. Therefore, $(x-y)\circ(x-y) = x\circ x - 2x\circ y + y\circ y = -2r^2 - 2x\circ y > -2r^2 + 2r^2 = 0$.

Second assume x and y lie in different components of T^n. Again let $r = |||x||| = |||y|||$. Then $x \circ x = -r^2 = y \circ y$. Here Theorem 1.4.b implies $x \circ y \geq |||x|||\,|||y||| = r^2$. Hence,
$(x-y)\circ(x-y) = x\circ x - 2x\circ y + y\circ y = -2r^2 - 2x\circ y < -2r^2 - 2r^2 < 0$. □

Proof of b) Assume x and y are both spacelike and $(x-y)\circ(x-y) \leq 0$. Then $x\circ x + y\circ y \leq 2x\circ y$. Thus $x\circ x$, $y\circ y$ and $2x\circ y$ are all positive. Hence, $(x+y)\circ(x+y) = x\circ x + 2x\circ y + y\circ y > 0$. □

Proof of c) Assume $x, y \in M^n$, $x \neq y$ and $x\circ e_n = y\circ e_n$. Then $(x-y)\circ e_n = 0$. Hence, $(x-y)\circ(x-y) = \|x-y\|^2 > 0$ because $x \neq y$. □

<u>Corollary 1.14.</u> a) If x and y lie in the same component of T^n, $x \neq y$ and $x \circ x = y \circ y$, then there is a spacelike unit vector u such that $Z_u(x) = y$. (Hence, $Z_u \in O^+(\mathbb{M}^n)$).

b) If x and y lie in different components of T^n and $x \circ x = y \circ y$, then there is a timelike unit vector u such that $Z_u(x) = y$. (Hence, $Z_u \notin O^+(\mathbb{M}^n)$.)

c) If x and $y \in \mathbb{M}^n$ are both spacelike and $x \circ x = y \circ y$, then there is a spacelike unit vector u such that either $Z_u(x) = y$ or $Z_u(x) = -y$. (Hence, $Z_u \in O^+(\mathbb{M}^n)$.)

d) If $x, y \in \mathbb{M}^n$, $x \neq y$, $x \circ x = y \circ y$ and $x \circ e_n = y \circ e_n$, then there is a spacelike unit vector u such that $Z_u(x) = y$. (Hence, $Z_u \in O^+(\mathbb{M}^n)$.)

<u>Proof of a)</u> Lemma 1.13.a implies $(x-y) \circ (x-y) > 0$. Let $u = (x-y)/|||x-y|||$. Then $u \circ u > 0$. So u is spacelike. Theorem 1.12.e implies $Z_u(x) = y$. Theorem 1.12.g implies $Z_u \in O^+(\mathbb{M}^n)$. $\square$

~~Proof of b) Lemma 1.13.a implies $(x-y)\circ(x-y) < 0$. Let $u = (x-y)/|||x-y|||$. Then $u \circ u < 0$, so u is timelike.~~

Homework Problem 1.6.

Prove Corollary 1.14.b, c and d.

Theorem 1.15. Every orthonormal sequence in $\mathbb{M}^n$ extends to an orthonormal basis for $\mathbb{M}^n$.

Proof. Let $u_1, \dots, u_k$ be an orthonormal sequence in $\mathbb{M}^n$. We induct on k.

Let $k=1$. If u_1 is timelike, then Corollary 1.14.a and b imply there is a unit vector v such that $Z_v(u_1) = e_n$. Since $e_1, \dots, e_n$ is an orthonormal basis for $\mathbb{M}^n$ and $Z_v \in O(\mathbb{M}^n)$, then $Z_v(e_n) = u_1, Z_v(e_2), \dots, Z_v(e_{n-1})$ is an orthonormal basis for $\mathbb{M}^n$.

If u_1 is spacelike, then Corollary 1.14.c implies there is a unit vector v such that $Z_v(u_1) = \varepsilon e_1$ where $\varepsilon = \pm 1$. Since $\varepsilon e_1, e_2, \dots, e_n$ is an orthonormal basis for $\mathbb{M}^n$ and $Z_v \in O(\mathbb{M}^n)$, then $Z_v(\varepsilon e_1) = u_1, Z_v(e_2), \dots, Z_v(e_n)$ is an orthonormal basis for $\mathbb{M}^n$.

Assume Theorem 1.15 holds for all orthonormal sequences of length k, and suppose $u_1, \dots, u_k, u_{k+1}$ is an orthonormal sequence. Then $u_1, \dots, u_k$ extends to an orthonormal basis $u_1, \dots, u_k, v_{k+1}, \dots, v_n$. If u_{k+1} is spacelike/timelike, then one of the v_i's, $k+1 \le i \le n$, must also be spacelike/timelike. Reorder $v_{k+1}, \dots, v_n$ if necessary so that v_{k+1} is of the same type (spacelike or timelike) as u_{k+1}. Corollary 1.14 implies there is a unit vector w such that $Z_w(u_{k+1}) = \varepsilon v_{k+1}$ where $\varepsilon = \pm 1$. Theorem 1.12.f implies $w = (u_{k+1} - \varepsilon v_{k+1}) / |||u_{k+1} - \varepsilon v_{k+1}|||$. For $1 \le i \le k$, since $u_i \circ u_{k+1} = 0 = u_i \circ v_{k+1}$, then $u_i \circ w = 0$. Hence, Theorem 1.12.e implies $Z_w(u_i) = u_i$ for $1 \le i \le k$. Since $u_1, \dots, u_k, \varepsilon v_{k+1}, v_{k+2}, \dots, v_n$ is an orthonormal basis for M^n and $Z_w \in O(M^n)$, then $Z_w(u_1) = u_1, \dots, Z_w(u_k) = u_k, Z_w(\varepsilon v_{k+1}) = u_{k+1}, Z_w(v_{k+2}), \dots, Z_w(v_n)$ is an orthonormal basis for M^n. $\square$

<u>Remark</u> Theorem 1.15 can also be proved by a variant of the proof of the Gram Schmidt Orthogonalisation Process.

<u>Theorem 1.16</u> Every element of $O(\mathbb{M}^n)$ equals the composition of n or fewer linear reflections at most one of which is time reversing (and all the others are time preserving). Every element of $O^+(\mathbb{M}^n)$ equals the composition of n or fewer time preserving linear reflections.

<u>Proof</u> Let $f \in O(\mathbb{M}^n)$. We will prove there are $g_1, \ldots, g_n \in O(\mathbb{M}^n)$ such that

- each g_i is either id or a linear reflection,
- $g_1, \ldots, g_{n-1}$ are time preserving,
- g_n is time preserving if $f \in O^+(\mathbb{M}^n)$, and
- for $1 \le k \le n$, $g_k \circ \cdots \circ g_n \circ f(e_i) = e_i$ for $k \le i \le n$.

(We construct the g_i's in reverse order: $g_n, g_{n-1}, \ldots, g_1$.)

We begin by constructing g_n. If $f(e_n) = e_n$ let $g_n = id$. Assume $f(e_n) \neq e_n$. Then Corollary 1.14.a and b imply there is a unit vector u_n such that $Z_{u_n}(f(e_n)) = e_n$.

Furthermore, if $f \in O^+(M^n)$, then $f(e_n)$ and e_n lie in T^n_+, and u_n can be chosen to be spacelike. So, in that case, $Z_{u_n} \in O^+(M^n)$. Let $g_n = Z_{u_n}$.

Let $1 \le k \le n-1$ and assume $g_{k+1}, \ldots, g_n$ have been constructed as prescribed. Let $F = g_{k+1} \circ \cdots \circ g_n \circ f$. Then $F \in O(M^n)$ and $F(e_i) = e_i$ for $k+1 \le i \le n$. If $F(e_k) = e_k$, let $g_k = id$. Assume $F(e_k) \ne e_k$. Let $y = F(e_k)$. Then $y \circ y = e_k \circ e_k$ and $y \circ e_n = F(e_k) \circ F(e_n) = e_k \circ e_n$. Therefore, Corollary 1.14.d implies there is a spacelike unit vector u such that $Z_u(y) = e_k$. Let $g_k = Z_u$. Then g_k is a time preserving linear reflection. Theorem 1.12.f implies $u = \pm(y - e_k)/|||y - e_k|||$. For $k+1 \le i \le n$,

$$e_i \circ u = \pm(e_i \circ y - e_i \circ e_k)/|||y - e_k||| =$$

$$\pm(F(e_i) \circ F(e_k) - e_i \circ e_k)/|||y - e_k||| = 0$$

Hence $g_k(e_i) = Z_u(e_i) = e_i$ by Theorem 1.12.c.

Consequently, for $k+1 \le i \le n$,

$$g_k \circ g_{k+1} \circ \cdots \circ g_n \circ f(e_i) = Z_u \circ F(e_i) = Z_u(e_i) = e_i$$

and

$$g_k \circ g_{k+1} \circ \cdots \circ g_n \circ f(e_k) = Z_u \circ f(e_k) = Z_u(y) = e_k .$$

It follows that $g_1, g_2, \ldots, g_n$ can be constructed as specified.

Since $g_1 \circ \cdots \circ g_n \circ f$ is linear and $g_1 \circ \cdots \circ g_n \circ f(e_i) = e_i$ for $1 \le i \le n$, then $g_1 \circ \cdots \circ g_n \circ f = id$. Since $g_i^{-1} = g_i$ for $1 \le i \le n$, then $f = g_n \circ \cdots \circ g_1$. $\square$