6. wedge product & volume revisited
7. inner products on V, extended to the tensor algebra.
8. tangent space & multivariable calculus.

Wedge product and Volume

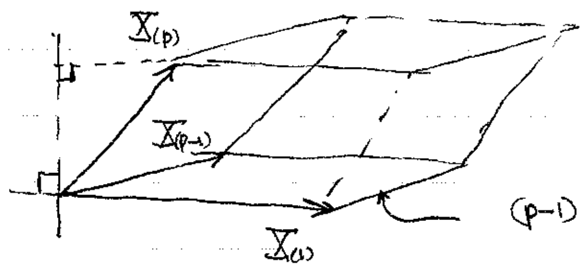
Linearly independent ordered set $\{X_{(1)},\dots,X_{(p)}\}$

Span $\{X_{(1)},\dots,X_{(p)}\}$ $\longrightarrow$ $X_{(1)}\wedge\dots\wedge X_{(p)} \neq 0$

The alternating property of the wedge product shows that one can add linear combinations of the remaining vectors in a set to a given vector without changing the wedge product, since by linearity the added terms will result in wedge products with repeated factors.

$$\Big(X_{(1)} + \sum_{j=2}^{p} C^j X_{(j)}\Big)\wedge X_{(2)}\wedge\dots\wedge X_{(p)}$$
$$= X_{(1)}\wedge\dots\wedge X_{(p)} + \sum_{j=2}^{p} C^j \; X_{(j)}\wedge X_{(2)}\wedge\dots\wedge X_{(p)} \quad = \quad X_{(1)}\wedge\dots\wedge X_{p}.$$

It is exactly this property which enables the wedge product to extend the notion of a length of a vector determined by an inner product to the definition of a p-measure for p-parallellopipeds. Geometrically this is defined inductively by multiplying the (orthogonal) height of the opposing face from a (p−1)-parallellopiped base by the area of that base. Tilting the p-parallelopiped along the directions of the (p−1)-parallellopiped while maintaining the height doesnt change the p-measure.



(p−1)-parallelopiped ~ $\{X_{(1)},\dots,X_{(p-1)}\}$.

This tilting corresponds exactly to the operation of adding a linear combination of the remaining edge vectors to the one singled out.

The Gram-Schmidt orthonormalization process consists of two parts: orthogonalization and normalization. The orthogonalization steps are special choices of the above tilting operation. One begins by retaining $X_{(1)}$ as the first vector in the set. Next one replaces $X_{(2)}$ by its projection perpendicular to $X_{(1)}$

$$Y_{(1)} = X_{(1)}$$
$$Y_{(2)} = X_{(2)} - \frac{g(X_{(2)},X_{(1)})\,X_{(1)}}{g(X_{(1)},X_{(1)})}$$

At each successive step one replaces the next vector from the original set by its projection perpendicular to the span of the preceding set.

$$Y_{(i)} = X_{(i)} - \sum_{j=1}^{i-1} \frac{g(X_{(i)}, X_{(j)})}{g(X_{(j)}, X_{(j)})} X_{(j)} \qquad 1 \leq i \leq p$$

The new set of orthogonal vectors $\{Y_{(1)}, \ldots, Y_{(p)}\}$ is such that the span of the first i vectors coincides with the span of the first i vectors of the original set.

By construction $\qquad Y_{(1)} \wedge \ldots \wedge Y_{(p)} = X_{(1)} \wedge \ldots \wedge X_{(p)}$

Geometrically the p-~~volume~~ measure of the parallellopipeds they determine is equal.

The second step of the Gram-Schmidt procedure tells us what that ~~vol~~ p-measure is. Normalization of the basis $\{Y_{(1)}, \ldots, Y_{(p)}\}$ for the given subspace is a simple step

$$\hat{Y}_{(i)} = |g(Y_{(i)}, Y_{(i)})|^{-1/2}\, Y_{(i)} \qquad 1 \leq i \leq p.$$

or $\qquad Y_{(i)} = |g(Y_{(i)}, Y_{(i)})|^{1/2}\, \hat{Y}_{(i)}.$

By linearity of the wedge product

$$Y_{(1)} \wedge \ldots \wedge Y_{(p)} = \left(\prod_{i=1}^{p} |g(Y_{(i)}, Y_{(i)})|^{1/2} \right) \hat{Y}_{(1)} \wedge \ldots \wedge \hat{Y}_{(p)} .$$

The coefficient in parentheses is the product of the lengths of the orthogonal sides of the rectangular p-parallellopiped, clearly the number that should be assigned as the p-measure of this p-parallellopiped and hence to the original.

This number can be obtained by introducing an extention of the inner product to p-vectors which assigns ~~with~~ the value ± 1 to the inner product with itself of ~~a set of~~ the wedge product of a set of p orthonormal vectors. The "length" of a p-vector defined by such an inner product will then produce the p-measure of the p-parallelopiped formed by the factors $\{\mathbb{X}_{(1)}, \dots, \mathbb{X}_{(p)}\}$ in the wedge product.

Thus the wedge product divides up the set of p-parallellopipeds into equivalence classes with the same p-measure. An inner product then assigns a value to each equivalence class.

To understand this ~~word~~ inner product for p-vectors, one must do a computation.

$\binom{p}{0}$-tensors already have a natural inner product obtained from the inner product g on the vector space. One simply uses the inner product on each index

$$
\begin{aligned}
S\cdot T &= g_{\alpha_1\beta_1}\cdots g_{\alpha_p\beta_p}\, S^{\alpha_1\cdots\alpha_p}\, T^{\beta_1\cdots\beta_p} \\
&= S^{\alpha_1\cdots\alpha_p}\, T_{\alpha_1\cdots\alpha_p} = S_{\beta_1\cdots\beta_p}\, T^{\beta_1\cdots\beta_p}
\end{aligned}
$$

where index raising and lowering enables this result to be expressed in many ways.

For p-vectors one divides by a factorial factor to avoid overcounting so that the sum is equivalent to the sum over ordered indices

$$\langle S,T\rangle = \frac{1}{p!}\, S\cdot T = S^{|\alpha_1\cdots\alpha_p|}\, T_{\alpha_1\cdots\alpha_p}.$$

For wedge products of p vectors

$$(X_{(1)}\wedge\cdots\wedge X_{(p)})^{\alpha_1\cdots\alpha_p} = \delta^{\alpha_1\cdots\alpha_p}_{\gamma_1\cdots\gamma_p}\, X_{(1)}^{\gamma_1}\cdots X_{(p)}^{\gamma_p}$$

$$(Y_{(1)}\wedge\cdots\wedge Y_{(p)})_{\alpha_1\cdots\alpha_p} = \delta^{\delta_1\cdots\delta_p}_{\alpha_1\cdots\alpha_p}\, Y_{(1)\delta_1}\cdots Y_{(p)\delta_p} \quad \leftarrow \text{?!}$$

?! index raising, lowering commutes with wedging

$$
\begin{aligned}
\langle X_{(1)}\wedge\cdots\wedge X_{p},\, Y_{(1)}\wedge\cdots\wedge Y_{(p)}\rangle &= \frac{1}{p!}\,\delta^{\alpha_1\cdots\alpha_p}_{\gamma_1\cdots\gamma_p}\,\delta^{\delta_1\cdots\delta_p}_{\alpha_1\cdots\alpha_p}\, X_{(1)}^{\gamma_1}\cdots X_{(p)}^{\gamma_p}\, Y_{(1)\delta_1}\cdots Y_{(p)\delta_p} \\
&= \delta^{\delta_1\cdots\delta_p}_{\gamma_1\cdots\gamma_p}\, X_{(1)}^{\gamma_1}\cdots X_{(p)}^{\gamma_p}\, Y_{(1)\delta_1}\cdots Y_{(p)\delta_p} \\
&\cancel{= p!\,\delta^{[\delta_1}_{\gamma_1}\cdots\delta^{\delta_p]}_{\gamma_p}\, X_{(1)}^{\gamma_1}\cdots X_{(p)}^{\gamma_p}\, Y_{(1)\delta_1}\cdots Y_{(p)\delta_p}} \\
&= p!\, X_{(1)}^{[\delta_1}\cdots X_{(p)}^{\delta_p]}\, Y_{(1)\delta_1}\cdots Y_{(p)\delta_p} \\
&= p!\, X_{[(1)}^{\delta_1}\cdots X_{(p)]}^{\delta_p}\, Y_{(1)\delta_1}\cdots Y_{(p)\delta_p} \\
&= \epsilon^{i_1\cdots i_p}\,(X_{(i_1)}^{\delta_1} Y_{(1)\delta_1})\cdots(X_{(i_p)}^{\delta_p} Y_{(p)\delta_p}) \\
&= \det(X_{(i)}\cdot Y_{(j)})
\end{aligned}
$$

verbal explanation

$\mathrm{ALT}\, T(X_{(1)},\ldots,X_{(p)})$

$\sum_\sigma T(X_{(\sigma 1)},\ldots,X_{(p)})$

$\downarrow$

$T_{[\alpha_1\cdots\alpha_p]}$

transfers perm from vector label to component index

Thus for an orthonormal set of vectors $\{\hat{Y}_{(1)},\dots,\hat{Y}_{(p)}\}$

$$\langle \hat{Y}_{(1)}\wedge\dots\wedge\hat{Y}_{(p)},\ \hat{Y}_{(1)}\wedge\dots\wedge\hat{Y}_{(p)}\rangle = \det(\hat{Y}_{(i)}\cdot\hat{Y}_{(j)}) = \pm 1$$

yields the determinant of a diagonal matrix whose entries are ± 1. The absolute value then gives the result 1. The p-vector inner product therefore yields the length of a wedge product of p vectors as the p-measure of the associated p-parallelopiped.

What is the interpretation of the "angle" between two such wedge products of p vectors? For $p=n-1$, the answer is simple but for other values, our 3-dimensional minds have a bit of visualization with the geometry (at least mine does).

The duals of the p-vectors are then vectors:

$$n_{(1)} = {}^*(X_{(1)}\wedge\dots\wedge X_{(p)}),\quad n_{(2)} = {}^*(Y_{(1)}\wedge\dots\wedge Y_{(p)})$$

and because of the invariance of the inner product under the dual operation

$$\langle S,T\rangle = \langle {}^*S, {}^*T\rangle,$$

one has

$$\langle X_{(1)}\wedge\dots\wedge X_{(p)},\ Y_{(1)}\wedge\dots\wedge Y_{(p)}\rangle = \langle n_{(1)}, n_{(2)}\rangle = n_{(1)}\cdot n_{(2)} \ \cancel{\cos\theta}$$

$$= \|n_{(1)}\|\,\|n_{(2)}\|\cos\theta$$

The angle between the normals to the hyperplanes is also the angle between the hyperplanes. The cosine of the angle is just the inner product of the normalized p-vectors.

$$\|n_{(1)}\| = \|X_{(1)}\wedge\dots\wedge X_{(p)}\|,\quad \|n_{(2)}\| = \dots$$

$$\widehat{X_{(1)}\wedge\dots\wedge X_{(p)}} = \|X_{(1)}\wedge\dots\wedge X_{p}\|^{-1/2}\, X_{(1)}\wedge\dots\wedge X_{(p)}$$

$${}^*(\widehat{X_{(1)}\wedge\dots\wedge X_{(p)}}) = \hat{n}_{(1)},\ \dots$$

An inner product and norm on a vector space and its extension to the tensor algebra over the vector space.

The Euclidean inner product is incredibly useful in elementary linear algebra and its hidden presence in many circumstances enables one to hide the rich ~~structure~~ mathematical structure sitting on top of $\mathbb{R}^n$. Having exorcized this presence, ~~we must~~ and developed this rich structure we must now reinsert it into the formalism simply because it does play an important role.

Recall that the key to avoiding this rich structure was the identification of the dual ~~vector~~ of $\mathbb{R}^n$ with $\mathbb{R}^n$ itself by using the Euclidean inner product to express linear functions. This identification can be made ~~by any inner~~ with any (nondegenerate) inner product (often referred to as a metric) and its extension to the tensor algebra falls under the classical terminology of "raising and lowering indices with the metric." The idea of lengths ~~and angles of~~ of vectors and angles between vectors is also important in aiding physical intuition and it too can be extended to the tensor algebra.

Given an n-dimensional vector space V with basis $\{e_\alpha\}$ and dual basis $\{\omega^\alpha\}$, we introduce the idea of an inner product and its associated norm on V as a symmetric $\binom{0}{2}$-tensor which is nondegenerate

$$g = g_{\alpha\beta}\,\omega^\alpha \otimes \omega^\beta \quad , \quad g_{\alpha\beta} = g_{\beta\alpha} \quad , \quad \det(g_{\alpha\beta}) \neq 0 .$$

The inner product of two vectors X and Y is defined to be the evaluation of g on these vectors

$$g(X,Y) = g_{\alpha\beta} X^\alpha Y^\beta = X \cdot Y \quad , \quad g_{\alpha\beta} = e_\alpha \cdot e_\beta ,$$

and occasionally written as a "dot product", which is a bilinear binary operator on the vector space with real values.

The norm of a single vector is its self inner product

$$N(X) = g(X,X) = g_{\alpha\beta} X^\alpha X^\beta = X \cdot X$$

and its length is the square root of the absolute value of its norm

$$\|X\| = |N(X)|^{1/2} = |g(X,X)|^{\frac{1}{2}} .$$

For vectors with nonzero norm, one can introduce a normalization operation $\wedge$ as follows

$$\hat{X} = \|X\|^{-1} X \quad , \quad \|X\| \neq 0 \;, \longrightarrow |\hat{X} \cdot \hat{X}| = 1$$

and has the interpretation of the "direction" of the vector X. The vector $\hat{X}$ is called a unit vector and is said to have been normalized. One may define an angle θ between unit vectors by

$$\hat{X} \cdot \hat{Y} = \begin{cases} \cos\theta & , |\hat{X} \cdot \hat{Y}| \leq 1 \\ \cosh\theta & , |\hat{X} \cdot \hat{Y}| \geq 1 \quad \text{(hyperbolic angle)} \end{cases}$$

and define this to be the angle between the original unnormalized vectors.

Two vectors for which the inner product vanishes are said to be orthogonal. A basis of mutually orthogonal vectors is called an orthogonal basis and a basis of mutually orthogonal unit vectors is called an orthonormal basis. The nonzero components of the metric in such a basis are diagonal and of absolute value unity.

If $V = \mathbb{R}^n$ and $\{e_\alpha\}$ is the standard basis, then the standard Euclidean inner product has components

$$g_{\alpha\beta} = \delta_{\alpha\beta} = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases},$$

i.e. the symmetric matrix of components $(g_{\alpha\beta})$ is the unit matrix and the standard basis is an orthonormal basis. In this case all of the above definitions reduce to familiar concepts. The norm of a vector is just the sum of the squares of its components

$$N(X) = \delta_{\alpha\beta} X^\alpha X^\beta = \sum_{\alpha=1}^{n} (X^\alpha)^2 \geq 0$$

and is positive definite, i.e. ~~produces~~ the norm produces only nonnegative values and the value zero occurs only for the zero vector.

Starting from any basis $\{E_\alpha\}$ of $\mathbb{R}^n$, one can always obtain an orthogonal and then orthonormal basis by a procedure called the Gram-Schmidt procedure. This involves the operation of orthogonal projection. Given any vector X of nonzero norm one can always project any other vector along X and orthogonal to X

$$Y_{\parallel} = \frac{(Y \cdot \hat{X})\hat{X}}{\hat{X} \cdot \hat{X}} = \frac{(Y \cdot X) X}{X \cdot X},$$

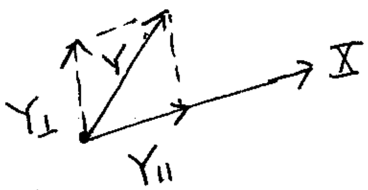
$$Y_{\perp} = Y - Y_{\parallel} = (1 - \hat{X}\hat{X}\cdot) Y .$$

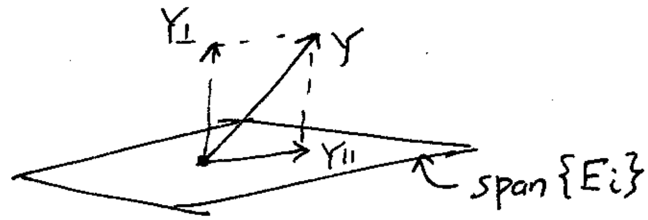
This projection holds for any inner product on a general vector space and can be extended to a projection along and orthogonal to any subspace which itself has an orthonormal basis.

If $\{E_i\}_{a=1,\dots,p}$ is an orthonormal basis of a p-dimensional subspace, then one simply sums up the projections along each basis vector

$$Y_{\parallel} = \sum_{i=1}^{p} \frac{(Y\cdot E_i)\, E_i}{E_i\cdot E_i}$$

$$Y_{\perp} = Y - Y_{\parallel}\,.$$





The Gram-Schmidt procedure then takes an arbitrary basis with at least 1 vector of nonzero norm, let's assume $\bar{e}_1\cdot\bar{e}_1 \neq 0$, and generates first an orthogonal basis by a series of projections, and then an orthonormal basis by normalizing the orthogonal basis.

Let $\bar{\bar{e}}_1 = \bar{e}_1$.

Next replace $\bar{e}_2$ by its projection orthogonal to $\bar{\bar{e}}_1$

$$\bar{\bar{e}}_2 = \bar{e}_2 - \frac{(\bar{e}_2\cdot\bar{e}_1)\,\bar{e}_1}{\bar{e}_1\cdot\bar{e}_1}\,.$$

Assuming $\bar{\bar{e}}_2$ has nonzero norm, replace $\bar{e}_3$ by its projection to the plane of the orthogonal vectors $\{\bar{\bar{e}}_1, \bar{\bar{e}}_2\}$.

$$\bar{\bar{e}}_3 = \bar{e}_3 - \sum_{i=1}^{2} \frac{(\bar{e}_3\cdot\bar{\bar{e}}_i)\,\bar{\bar{e}}_i}{\bar{\bar{e}}_i\cdot\bar{\bar{e}}_i}\,.$$

Continue in this way

$$\bar{\bar{e}}_j = \bar{e}_j - \sum_{i=1}^{j-1} \frac{(\bar{e}_j\cdot\bar{\bar{e}}_i)\,\bar{\bar{e}}_i}{\bar{\bar{e}}_i\cdot\bar{\bar{e}}_i}\,.$$

The hitch when the inner product is not positive-definite like the standard Euclidean inner product is that one can run into vectors of zero norm and the projection operation is then not defined at the next step. If this doesnt occur, including at the last step, then the basis $\{\bar{\bar{E}}_a\}_\bullet$ is an orthogonal basis of nonzero norm vectors which may be normalized to produce an orthonormal basis $\{\hat{\bar{\bar{E}}}_a\}$.

FACT. The numbers of positive norm vectors and of negative norm vectors are always the same for an orthonormal basis with respect to a fixed inner product. The difference (positive number − negative number) is called the signature of the inner product.

A positive-definite or Riemannian inner product has signature $s=n$; ~~while a Lorentz~~ the remaining cases are indefinite or pseudo-Riemannian. The case of only one positive norm vector or only one negative norm vector (basically equivalent as we will later see) is called Lorentzian and is the case of special and general relativity; this has signature $s=(n-1)-1=n-2$ or $-(n-2)$.

This means one has a standard set of components of the inner product in an orthonormal basis if we agree to put all the positive norm vectors first

$$(\eta_{\alpha\beta}) = \operatorname{diag}(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{q})$$

$$p+q = n = \dim V$$
$$p-q = s = \text{signature } g\,,$$

although in the Lorentzian case it is often conventional to emphasize the distinction between the single vector of the opposite signed norm by using the indices $0, 1, \ldots, n-1$, agreeing to use the zero index for the exceptional basis vector. Zero is refered to as a "timelike" index, while $1, \ldots, n-1$ are called "spatial"

$$(\eta_{\alpha\beta}) = \mathrm{diag}(-1, 1, \ldots, 1) \quad \text{or} \quad \mathrm{diag}(1, -1, \ldots, -1)$$

It helps to introduce an index convention for the spatial indices, agreeing to use Latin letters $a, b, c, \ldots = 1, \ldots, n-1$, so for example

$$\eta_{00} = -1 \,, \quad \cancel{\eta_{ab} = \delta_{ab}} \quad \eta_{0a} = 0 \,, \quad \eta_{ab} = \delta_{ab} .$$

in the first convention. In this case the positive norm is associated with the Euclidean inner product matrix — the Kronecker delta. Before developing this ~~analogy fu~~ further, let us return to the general case, but first an example:

<u>EX</u>. Let $gl(n,\mathbb{R})$ be the real n^2-dimensional vector space of $n\times n$ matrices. Introduce the following trace inner product

$$\langle \underline{A}, \underline{B} \rangle \equiv g(\underline{A}, \underline{B}) \equiv \mathrm{Tr}\, AB = \mathrm{Tr}\, BA$$

This is a pseudo-Riemannian or indefinite inner product. What is its signature? Well, any matrix can be uniquely decomposed into the sum of a symmetric matrix and an antisymmetric matrix

$$\underline{A} = \underline{A}_{+} + \underline{A}_{-} \,, \quad \underline{A}_{\pm} = \tfrac{1}{2}(\underline{A} \pm \underline{A}^{t}) .$$

This is a orthogonal direct sum decomposition of $gl(n,\mathbb{R})$ since the trace of the product of a symmetric and antisymmetric matrix is zero:

$$gl(n,\mathbb{R}) = \underbrace{sym(n,\mathbb{R})}_{\dim = \frac{n(n+1)}{2}} \oplus \underbrace{asym(n,\mathbb{R})}_{\frac{n(n-1)}{2}}$$

Furthermore, $\mathrm{Tr}\,\underline{A}^2$ equals the sum of the squares of the entries of a symmetric matrix and minus the sum of the squares for an antisymmetric matrix [VERIFY] so the trace inner product is positive definite on $sym(n,\mathbb{R})$ and negative definite on $asym(n,\mathbb{R})$. On each subspace we can find an orthonormal basis of vectors with the same sign norm by the Gram-Schmidt procedure, so the signature of $\langle , \rangle$ or g is $s = \frac{n(n+1)}{2} - \frac{n(n-1)}{2} = n$.
This inner product turns out to be quite important.

PROBLEM $n=2$.

Verify that
$$E_0 = 2^{-1/2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad E_1 = 2^{-1/2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$E_2 = 2^{-1/2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E_3 = 2^{-1/2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
is an orthonormal basis of $gl(2,\mathbb{R})$. This is a 4-dimensional vector space with a Lorentz inner product. What is the orthogonal projection of the matrix $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ along $\begin{pmatrix} -2 & 0 \\ 1 & 1 \end{pmatrix}$?

PROBLEM

Suppose we declare the basis $\overline{e}_0 = (1,1)$, $\overline{e}_1 = (2,1)$ to be a standard orthonormal basis of an inner product g on $\mathbb{R}^2$, with $-g(\overline{e}_0,\overline{e}_0) = +1 = g(\overline{e}_1,\overline{e}_1)$.
Express $g = g_{ab}\,\omega^a \otimes \omega^b$ in terms of the standard basis.

The dot product on $\mathbb{R}^n$ enabled us to avoid introducing 1-forms & the dual space by ~~writing~~ representing 1-forms in terms of the dot product of a vector. One can do this with any inner product on V. With each vector X we associate a 1-form σ_X by partially evaluating the metric

$$X \mapsto \sigma_X = g(X, \ ) \quad , \text{ i.e., } \sigma_X(Y) = g(X,Y)$$

or in components

$$(\sigma_X)_\bullet = g_{\alpha\beta}\,\omega^\alpha(X)\,\omega^\beta = g_{\alpha\beta} X^\alpha \omega^\beta = \underbrace{g_{\alpha\beta} X^\beta}_{(\sigma_X)_\alpha} \omega^\alpha$$

$$(\sigma_X)_\alpha = g_{\alpha\beta} X^\beta \equiv X_\alpha$$

In component form this is called "lowering the index" for obvious reasons.

Any $\binom{0}{2}$-tensor may be interpreted as a linear map from V to V^*. It is invertible only if its matrix of components has nonzero determinant, as assumed for $(g_{\alpha\beta})$. With the inverse map

$$X^\alpha = g^{\alpha\beta}(\sigma_X)_\beta = g^{\alpha\beta} X_\beta \qquad \text{("raising the index")}$$

associated with the inverse matrix $(g^{\alpha\beta})$, is a ~~$\binom{2}{0}$ symmet~~ nondegenerate symmetric $\binom{2}{0}$-tensor

$$g^{-1} = g^{\alpha\beta} e_\alpha \otimes e_\beta \quad , \quad g^{\alpha\beta} = g^{-1}(\omega^\alpha, \omega^\beta)$$

which can be used as an inner product on the dual space

$$\sigma \cdot \rho = g^{-1}(\sigma, \rho) = g^{\alpha\beta} \sigma_\alpha \rho_\beta \quad , \quad g^{\alpha\beta} = \omega^\alpha \cdot \omega^\beta$$

and accomplishes index raising by partial evaluation

$$X = g^{-1}(\sigma_X, \ ) .$$

Now it turns out that this invertible relationship between V and its dual is an "isometry", i.e., the maps commute with the inner product

$$g^{-1}(\sigma_X, \sigma_Y) = g^{\alpha\beta}(\sigma_X)_\alpha(\sigma_Y)_\beta = g^{\alpha\beta} g_{\alpha\gamma} X^\gamma g_{\beta\delta} Y^\delta$$
$$= g_{\beta\delta} X^\beta Y^\delta = g(X, Y) .$$

If we start with a 1-form, we can "raise its index" to get a vector

$$X_\sigma = g^{-1}(\sigma, \ ) \, .$$

$$(X_\sigma)^\alpha = g^{\alpha\beta}\sigma_\beta \equiv \sigma^\alpha .$$

Rather than introducing complicated notation like X_σ or σ_X or whatever, one uses the convention of keeping the same kernel symbol in the index notation and simply changing the position of the index. In the index free notation, one needs another way to distinguish the 1-form and vector symbol, so the sharp sign (raise) and the flat sign (lower) are used

$$X = X^\alpha e_\alpha \rightarrow X^\flat = X_\alpha \omega^\alpha \, , \quad X_\alpha \equiv g_{\alpha\beta} X^\beta$$

$$\sigma = \sigma_\alpha \omega^\alpha \rightarrow \sigma^\# = \sigma^\alpha e_\alpha \, , \quad \sigma^\alpha \equiv g^{\alpha\beta}\sigma_\beta \, .$$

Note that the ~~standard~~ matrix of inner product components in a standard orthonormal basis

$$(\eta_{\alpha\beta}) = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1) = (\eta^{\alpha\beta})$$

is its own inverse, so the dual basis to an orthonormal basis is itself an orthonormal basis with respect to the dual inner product. In fact index raising relates a vector and its corresponding ~~dual~~ dual covector in such a basis, multiplied by the sign of their norm.

In general by linearity

$$\left.\begin{aligned} X^\flat &= (X^\alpha e_\alpha)^\flat = X^\alpha e_\alpha{}^\flat = \\ &= g_{\beta\alpha} X^\alpha \omega^\beta \end{aligned}\right\} \rightarrow e_\alpha{}^\flat = g_{\beta\alpha}\omega^\beta = g_{\alpha\beta}\omega^\beta$$

and

$$\left.\begin{aligned} \sigma^\# &= (\sigma_\alpha \omega^\alpha)^\# = \sigma_\alpha \omega^{\alpha\#} \\ &= \sigma_\alpha g^{\alpha\beta} e_\beta \end{aligned}\right\} \rightarrow \omega^{\alpha\#} = g^{\alpha\beta} e_\beta \, .$$

Since $\det(g_{\alpha\beta}) \neq 0$, $\{\omega^{\alpha\#}\}$ is also a basis of V, called the basis reciprocal to $\{e_\alpha\}$. Inner products with the reciprocal basis are equivalent to evaluation of the dual covectors

$$X^\alpha = \omega^\alpha(X) = \omega^{\alpha\#} \cdot X$$

and the evaluation of any covector on a vector is equivalent to the ~~dot pr~~ inner product with the vector obtained from the covector by index raising.

$$\sigma(\underline{X}) = \sigma_\alpha X^\alpha = g_{\alpha\beta}\,\sigma^\beta X^\alpha = \sigma^\# \cdot \underline{X}\,.$$

On $\mathbb{R}^n$ index raising and lowering components in the standard basis is trivial since the metric matrix and its inverse are both the identity matrix and the contravariant and covariant components are the same. Thus the Euclidean dot product, familiar, was used to express covectors in elementary linear algebra to avoid having to develop the unfamiliar idea of the dual space.

In general, in any orthonormal basis, index raising and lowering simply changes the sign of a component by the sign of the norm of the basis vector it is associated with.

PROBLEM. For $gl(2,\mathbb{R})$ with the trace inner product, what is the basis reciprocal to the standard basis $\{\underline{e}^\alpha{}_\beta\}$ defined by

$\underline{A} = A^\alpha{}_\beta\, \underline{e}^\beta{}_\alpha$, $e^\beta{}_\alpha$ has a single nonzero entry 1 in the αth row and βth column (note the switch of rows & columns from the entry $A^\alpha{}_\beta$ to the basis vector $e^\beta{}_\alpha$).

?

So we have an inner product on V which extended easily to one on V^* and provided a vector space isomorphism between the two. The inner product can be easily extended to the entire tensor algebra over V. On any given tensor product space one defines the tensor product inner product on the basis by

$$\underbrace{[e_\alpha \otimes \cdots \otimes \omega^\beta]}\cdot \overbrace{[e_\gamma \otimes \cdots \otimes \omega^\delta \otimes \cdots]} \equiv (e_\alpha \cdot e_\gamma)\cdots(\omega^\beta\cdot\omega^\delta)\cdots$$
$$= g_{\alpha\gamma}\cdots g^{\beta\delta}\cdots$$

and extend by linearity

$$S \cdot T = [S^{\alpha\cdots}{}_{\beta\cdots}\, e_\alpha \otimes \cdots \otimes \omega^\beta \otimes \cdots] \cdot [T^{\gamma\cdots}{}_{\delta\cdots}\, e_\gamma \otimes \cdots \otimes \omega^\beta \otimes \cdots]$$
$$= S^{\alpha\cdots}{}_{\beta\cdots}\, T^{\gamma\cdots}{}_{\delta\cdots}\, [e_\alpha \otimes \cdots \otimes \omega^\beta \otimes \cdots] \cdot [e_\gamma \otimes \cdots \otimes \omega^\delta \otimes \cdots]$$
$$= g_{\alpha\gamma}\cdots\, g^{\beta\delta}\cdots\, S^{\alpha\cdots}{}_{\beta\cdots}\, T^{\gamma\cdots}{}_{\delta\cdots}$$

One simply takes the inner product of each index pair independently. This may be rewritten using the index raising and lowering conventions extended to each index separately. Namely a given $\binom{p}{q}$ - tensor

$$S = S^{\alpha\cdots}{}_{\beta\cdots}\, e_\alpha \otimes \cdots \otimes \omega^\beta \otimes \cdots$$

can be associated with a family of 2^{p+q} different tensors of total rank $p+q$ by raising & lowering each index independently. These index raising & lowering operations are all isometries, so it doesnt matter which member of the family one uses in evaluating the inner product. For example

$$S \cdot T = g_{\alpha\gamma}\cdots\, g^{\beta\delta}\, S^{\alpha\cdots}{}_{\beta\cdots}\, T^{\gamma\cdots}{}_{\delta\cdots}$$
$$= S^{\alpha\cdots}{}_{\beta\cdots}\, T_{\alpha\cdots}{}^{\beta\cdots}$$
$$= S_{\alpha\cdots}{}^{\beta\cdots}\, T^{\gamma\cdots}{}_{\delta\cdots}$$
$$= S^{\alpha\cdots\beta\cdots}\, T_{\alpha\cdots\beta\cdots}$$
$$= S_{\alpha\cdots\beta\cdots}\, T^{\alpha\cdots\beta\cdots} \,.$$

If $\{e_\alpha\}$ is an orthonormal basis, then so is $\{e_\alpha \otimes \cdots \otimes \omega^\beta \otimes \cdots\}$ since ~~its the norm of~~ $(g_{\alpha\beta})$ is diagonal with $\not{\Sigma}\, g_{\alpha\alpha} = \pm 1$. The norm is the product of the norms of the factors. For a positive definite inner product on V, all of the tensor product ~~are also pos~~ inner products are also positive-definite. For the standard dot product on $\mathbb{R}^n$, the norm consists of the sum of the squares of all the components of a tensor

$$T \cdot T = \delta_{\alpha\gamma} \cdots \delta^{\beta\delta} \cdots T^{\alpha \cdots}{}_{\beta \cdots} T^{\gamma \cdots}{}_{\delta \cdots}$$
$$= \sum_{\alpha=1}^{n} \cdots \sum_{\beta=1}^{n} \cdots \left(T^{\alpha \cdots}{}_{\beta \cdots}\right)^2 .$$

For $\binom{1}{1}$-tensors this is ~~just~~ closely related to the trace inner product:
$$T = T^\alpha{}_\beta \, e_\alpha \otimes \omega^\beta$$
$$T \cdot T = \delta_{\alpha\gamma} \delta^{\beta\delta} T^\alpha{}_\beta T^\gamma{}_\delta = \underbrace{\delta^{\beta\delta}}_{\text{takes trace}} \overbrace{(\underbrace{T^\alpha{}_\beta \delta_{\alpha\gamma}}_{\substack{\text{need transpose} \\ \text{to get index} \\ \text{adjacent for matrix} \\ \text{mult.}}} T^\gamma{}_\delta)}^{\text{matrix mult}}$$
$$= \mathrm{Tr}\, \underline{T}^t \, \underline{T}$$

where $(T^\alpha{}_\beta) \equiv \underline{T}$.

For symmetric tensors $\underline{T}^t = \underline{T}$, it is the trace inner product, but for antisymmetric tensors $\underline{T}^t = -\underline{T}$ and it is the sign reversal. Thus the transpose changes the indefinite trace inner product into a positive definite one.

For ~~p-form~~ p-vectors or p-covectors (p-forms), (totally antisymmetric $\binom{p}{0}$-tensors or $\binom{0}{p}$-tensors), there is a slight change we can make to avoid overcounting.

$$S = S_{\alpha_1 \cdots \alpha_p} \omega^{\alpha_1} \otimes \cdots \otimes \omega^{\alpha_p} = \frac{1}{p!} S_{\alpha_1 \cdots \alpha_p} \omega^{\alpha_1 \cdots \alpha_p}$$
$$= S_{\alpha_1 \cdots \alpha_p} \omega^{|\alpha_1 \cdots \alpha_p|}$$

The basis $\{\omega^{|\alpha_1 \cdots \alpha_p|}\}$ for p-forms is orthogonal but not normalized since

$$S \cdot T = [S_{\alpha_1 \cdots \alpha_p} \omega^{|\alpha_1 \cdots \alpha_p|}] \cdot [T_{\beta_1 \cdots \beta_p} \omega^{|\beta_1 \cdots \beta_p|}]$$
$$= S_{\alpha_1 \cdots \alpha_p} T^{\alpha_1 \cdots \alpha_p} = p! \; S_{\alpha_1 \cdots \alpha_p} T^{|\alpha_1 \cdots \alpha_p|}$$

so $\omega^{|\alpha_1 \cdots \alpha_p|} \cdot \omega^{|\beta_1 \cdots \beta_p|} = p! \, g^{\alpha_1\beta_1} \cdots g^{\alpha_p\beta_p}$.

We can make the p-form basis orthonormal by redefining the inner product to eliminate the $p!$ factor

$$\langle S, T \rangle \equiv \frac{1}{p!} S \cdot T \ .$$

~~This makes the~~

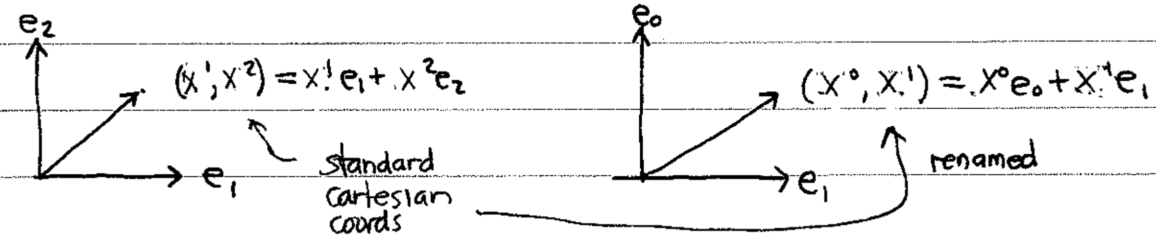
For a p-form on $\mathbb{R}^n$, the new norm $\langle S, S \rangle$ reduces to the sum of the squares of the ordered components of the p-form. We make the same redefinition for p-vectors.

Okay, so now we understand how to extend an inner product on V to the tensor algebra over V and use it to establish an isomorphism between the different tensor product spaces of the same total rank. For indefinite metrics we need to understand how to interpret the resulting geometry. The difference between Euclidean geometry and pseudo-Euclidean geometry can be understood in the lowest nontrivial dimension $n=2$.

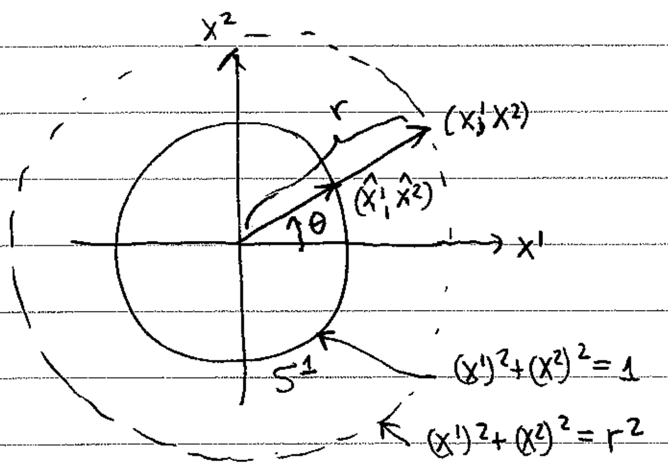
To follow conventions introduce the distinct notations for the standard basis of $\mathbb{R}^2$, an orthonormal basis with respect to the following two inner products

$e_1=(1,0) \quad e_2=(0,1)$ Euclidean case

$e_0=(1,0) \quad e_1=(0,1)$ Lorentz case $(-e_0\cdot e_0 = 1 = e_1\cdot e_1)$



To further complicate matters we always draw the pictures with the single timelike vector e_0 vertical instead of horizontal to conform with its interpretation as the time.



For the Euclidean $\mathbb{R}^2$, all nonzero vectors can be normalized and one can introduce projective cartesian coordinates

$$\hat{x}^1 = x^1/r$$
$$\hat{x}^2 = x^2/r$$
$$(\hat{x}^1)^2+(\hat{x}^2)^2 = 1 \rightarrow \text{ let } \hat{x}^1=\cos\theta,\ \hat{x}^2=\sin\theta$$
$$x^1 = r\hat{x}^1 = r\cos\theta$$
$$x^2 = r\hat{x}^2 = r\sin\theta$$

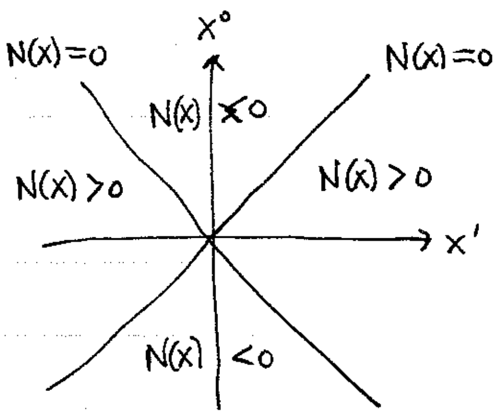
and then spherical coordinates r (distance from the origin, defined as the length of a vector) and θ (angle from positive axis in counterclockwise direction. The identification of the constraint on the projective coordinates with the fundamental identity of trigonometry leads to their parametrization by the cosine and sine. The radial coordinate is just the radius of the circle centered at the origin; the angular coordinate parametrizes the directions, ie., the unit vectors, which lie on the unit circle S^1. Translations in θ lead to rotations of the plane about the origin.

The inner product of two unit vectors

$$\hat{u}\cdot\hat{v} = (\cos\theta_1, \sin\theta_1)\cdot(\cos\theta_2, \sin\theta_2) = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2$$
$$= \cos(\theta_1-\theta_2)$$

is just the cosine of the angle between them, $|\theta_1-\theta_2|$.

The Lorentz $\mathbb{R}^2$ is quite different because of the indefinite signature.



$N(x) = -(x^0)^2+(x^1)^2$

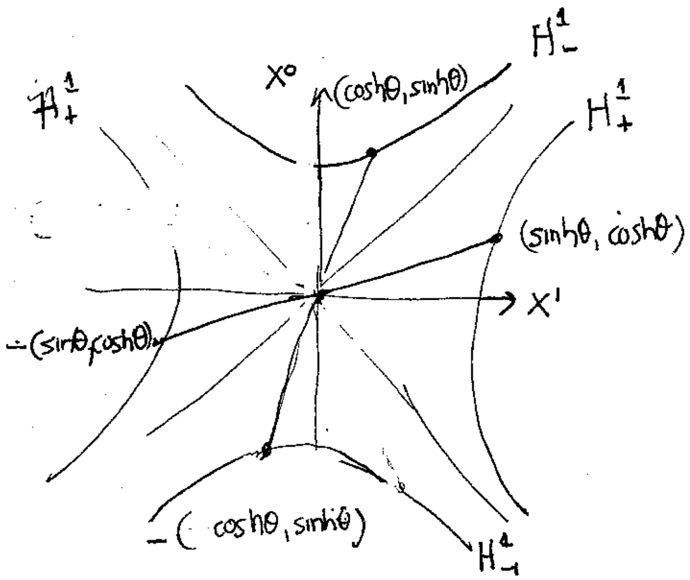
The "lightcone" of "null vectors" $x^0 = \pm x^1$ with zero norm divides the plane into four disjoint ~~subsq~~ regions:
two with $|x^0| > |x^1|$ and negative norm ("timelike vectors")
and two with $|x^0| < |x^1|$ and positive norm (spacelike vectors).

The length of a vector defines its distance from the origin.
Thus we have a timelike distance for the timelike vectors

$$-\tau^2 = -(x^0)^2+(x^1)^2 \rightarrow \tau = |-(x^0)^2+(x^1)^2|^{1/2}$$

and a spacelike distance for the timelike vectors

$$\sigma^2 = -(x^0)^2+(x^1)^2 \rightarrow \sigma = |-(x^0)^2+(x^1)^2|^{1/2}.$$



The unit "pseudosphere", all points at a unit distance from the origin, parametrizing the nonnull directions, now consists of 4 disjoint pieces, requiring 4 different parametrizations.

timelike: $\hat{x}^0 = x^0/\tau = \pm\cosh\theta$
$\hat{x}^1 = x^1/\tau = \sinh\theta$

$$-\underbrace{(\hat{x}^0)^2}_{\cosh^2\theta} + \underbrace{(\hat{x}^1)^2}_{\sinh^2\theta} = -1 \quad \left(\begin{array}{c}\text{unit hyperbola} \\ H^1_- \\ \uparrow \\ \text{timelike directions}\end{array}\right)$$

spacelike: $\hat{x}^0 = x^0/\sigma = \sinh\theta$
$\hat{x}^1 = x^1/\sigma = \pm\cosh\theta$

$$-\underbrace{(\hat{x}^0)^2}_{\sinh^2\theta} + \underbrace{(\hat{x}^1)^2}_{\cosh^2\theta} = 1 \quad \left(\begin{array}{c}\text{unit hyperbola} \\ H^1_+ \\ \uparrow \\ \text{spacelike directions}\end{array}\right)$$

The possibilities for inner products of unit vectors are:

1) two timelike unit vectors ($\epsilon_i = \pm 1$)

$$\epsilon_1(\cosh\theta_1, \sinh\theta_1) \cdot \epsilon_2(\cosh\theta_2, \sinh\theta_2) = \epsilon_1\epsilon_2[-\cosh\theta_1\cosh\theta_2 + \sinh\theta_1\sinh\theta_2]$$

$$= -\epsilon_1\epsilon_2 \cosh(\theta_1 - \theta_2) = \cancel{\epsilon_1\epsilon_2 \cosh(\theta_1 - \theta_2)}$$

relative angle: $|\theta_1 - \theta_2| = \cancel{|\epsilon_1\theta_1 - \epsilon_2\theta_2|}$

2) two spacelike unit vectors

$$\epsilon_1(\sinh\theta_1, \cosh\theta_1) \cdot \epsilon_2(\sinh\theta_2, \cosh\theta_2) = (\epsilon_1\epsilon_2)[-\sinh\theta_1\sinh\theta_2 + \cosh\theta_1\cosh\theta_2]$$

$$= \epsilon_1\epsilon_2 \cosh(\theta_1 - \theta_2) = \cancel{\epsilon_1\epsilon_2 \cosh(\epsilon_1\theta_1 - \epsilon_2\theta_2)}$$

relative angle: $|\theta_1 - \theta_2|$.

3) timelike and spacelike unit vector:

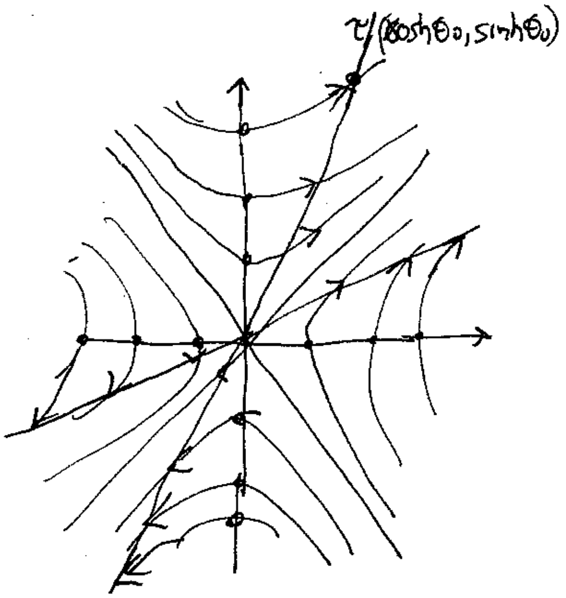
$\epsilon_1\epsilon_2$ $\cancel{-\cosh\theta_1 \sinh\theta_1}$

$$\epsilon_1(\cosh\theta_1, \sinh\theta_1) \cdot \epsilon_2(\sinh\theta_2, \cosh\theta_2) = [-\cosh\theta_1\sinh\theta_2 + \sinh\theta_1\cosh\theta_2]$$

$$= \epsilon_1\epsilon_2(\sinh(\theta_1 - \theta_2)) = \cancel{\sinh(\epsilon_2\theta_1 - \epsilon_1\theta_2) = \epsilon_1\epsilon_2 \sinh(\epsilon_1\theta_1 - \epsilon_2\theta_2)}$$

relative angle: $|\theta_1 - \theta_2|$

A translation in the angle θ of the 4 different parametrizations is a "pseudo-rotation".



The pseudorotations of the plane leave the pseudospheres (hyperbolas) invariant and push the points along them as indicated.

EXERCIZE. Using the hyperbolic addition formulas, show that a translation by θ_0 in the angular coordinate θ used in the four parametrizations above leads to the linear transformation

$$\begin{pmatrix} x^0 \\ x^1 \end{pmatrix} \to \begin{pmatrix} \cosh\theta_0 & \sinh\theta_0 \\ \sinh\theta_0 & \cosh\theta_0 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \end{pmatrix} = \begin{pmatrix} \bar{\bar{x}}^0 \\ \bar{\bar{x}}^1 \end{pmatrix} . \equiv \underline{B}(\theta) \begin{pmatrix} x^0 \\ x^1 \end{pmatrix} \qquad \left(\begin{matrix}\text{"active" or} \\ \text{point transf}\end{matrix}\right)$$

show that this preserves the norm. Note that the inverse transformation is $B(-\theta_0)$.
Repeat for the Euclidean case to find

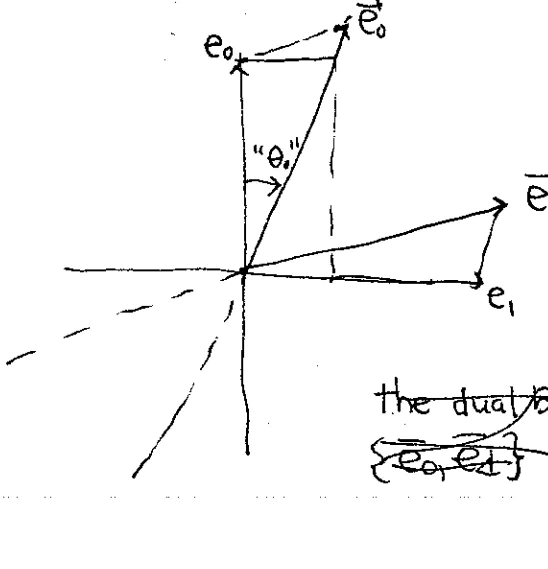
$$\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \to \begin{pmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} \bar{\bar{x}}^1 \\ \bar{\bar{x}}^2 \end{pmatrix} \equiv \underline{R}(\theta) \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \qquad (\text{ditto})$$

which preserves the Euclidean norm. Note that the inverse transformation.

The ray $\{\tau(\cosh\theta_0, \sinh\theta_0) \mid \tau \geq 0\}$ is the image of the positive x^0 axis under this pseudorotation. What is the slope $\frac{dx^1}{dx^0}$ of this ray?

If we interpret x^1 as a distance in the usual sense and x^0 as a time, what name do we give $\frac{dx^1}{dx^0}$?

This gives a physical interpretation of the angle θ_0, often called the boost parameter. The pseudorotation in this case is called a (Lorentz) boost.



Let $(\bar{e}_0, \bar{e}_1)$ be the images of (e_0, e_1) under the boost.
The cartesian coordinates $\{x^0, x^1\}$ viewed as real valued linear functions are just the dual basis $\{\omega^0, \omega^1\}$ to the standard basis, ~~while the new cartesian coordinates $\{\bar{x}^0, \bar{x}^1\}$ are the dual basis covectors $\{\bar{\omega}^0, \bar{\omega}^1\}$ to the new basis $\{\bar{e}_0, \bar{e}_1\}$~~

and the images under the boost has coordinates

$$\left.\begin{aligned}\bar{\bar{X}}^0 &= \omega^0(\underline{B}(\theta)\underline{X}) \\ \bar{\bar{X}}^1 &= \omega^1(\underline{B}(\theta)\underline{X})\end{aligned}\right\} \cdot \ \bar{\bar{X}}^\alpha = B(\theta)^\alpha{}_\beta X^\beta$$

with respect to the original basis. We can introduce new coordinates defined with respect to the image basis which are just the new dual basis covectors thought of simply as functions.

$$\bar{e}_a = B(\theta) e_a = e_a \underbrace{B(\theta)^a{}_b}$$

matrix of transformation has columns equal to old components of new vectors

$$\bar{\omega}^a = B(\theta)^{-1}\omega^a = B(\theta)^{-1a}{}_b\,\omega^b$$

so $\quad \bar{X}^a = B(\theta)^{-1a}{}_b X^b = B(-\theta)^a{}_b X^b$

$$\begin{pmatrix}\bar{X}^0 \\ \bar{X}^1\end{pmatrix} = \begin{pmatrix}\cosh\theta_0 & -\sinh\theta_0 \\ -\sinh\theta_0 & \cosh\theta_0\end{pmatrix}\begin{pmatrix}X^0 \\ X^1\end{pmatrix}$$

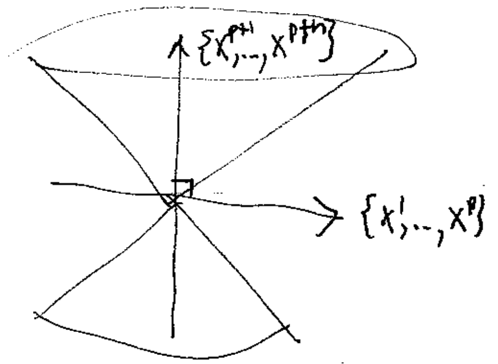
("passive" or coordinate transf.)

So the associated coordinate transformation is the inverse of the active point transformation. This is easy to see since the new coordinates of a point are just the coordinates of the point from which it came under the point transformation. The coordinates are said to be "dragged along" by the point transformation.

What is the situation for $n > 2$? Take $\mathbb{R}^n$ with the standard basis taken to be orthonormal

$$(\eta_{\alpha\beta}) = \operatorname{diag}(\underbrace{1,\dots,1}_{p},\underbrace{-1,\dots,-1}_{q}) \qquad p+q=n.$$

$$\underbrace{\{x^1,\dots,x^p\}}_{\text{spacelike coordinates}} \quad \underbrace{\{x^{p+1},\dots,x^n\}}_{\text{timelike coordinates.}}$$

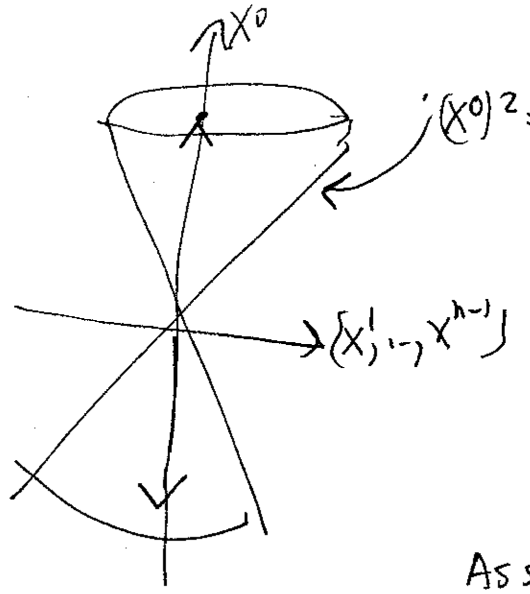


The space decomposes into an orthogonal direct sum of two subspaces – the spacelike directions (positive norm) and the timelike directions (negative norm)

$$N(x) = \underbrace{\sum_{\alpha=1}^{p} (x^\alpha)^2}_{T^2} - \underbrace{\sum_{\alpha=p+1}^{n} (x^\alpha)^2}_{S^2}$$

Within each subspace the distance function $|N(x)|$ is just the Euclidean distance and one can rotate the corresponding coordinates among themselves. Note that as far as distance is concerned, the overall sign of the norm is irrelevant since the absolute value is used so two metrics which differ only ~~with~~ by interchanging the ~~number~~ individual numbers of positive and negative normed vectors in an orthonormal basis do not have different geometries. Also, except in the Lorentz case, the choice of ~~positive~~ terminology spacelike & timelike is arbitrary.

The Lorentz case is different since the timelike direction is 1-dimensional so one has a set of two timelike unit vectors in opposing directions, allowing the definition of a future direction & a past direction.

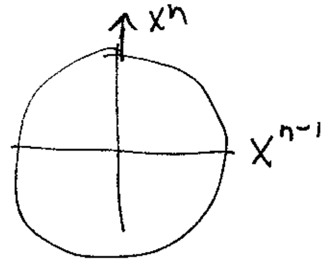


$(X^0)^2 = \sum_{\alpha=1}^{n-1} (X^\alpha)^2$ "null cone" (~~hyper~~

Only in this case does the set of timelike directions consist of two disjoint subsets which are ~~the~~ "inside" the null cone.

As soon as we have a metric with a second

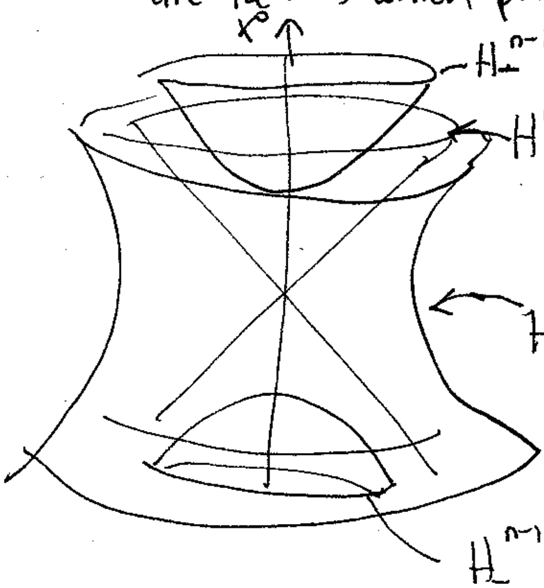
timelike direction



$$N(X)^2 = \underbrace{-(X^{n-1})^2 - (X^n)^2}_{} + \sum_{\alpha=1}^{n-2} (X^\alpha)^2$$

one can rotate ~~from~~ around, ie the timelike directions form a single connected subset & one loses the notion of CAUSALITY.

This is why Lorentz metrics, of all the possible indefinite metrics, are the ones which prove useful in physics.



H_-^{n-1} ← unit hyperboloid of future timelike directions (~~disconne~~

H_0^{n-1} null cone

connected hyperboloid of one sheet of spacelike directions. H_+^{n-1}

disconnected hyperboloid of two sheets of timelike directions

H_-^{n-1} ← unit hyperboloid of past timelike directions

unit pseudospheres

enough for now.