

Quadratic Forms and the Clifford Product

From the diagonal of a bilinear form to the algebra it specifies

James Pustejovsky

Brandeis University
Department of Computer Science

CS 135: Computational Semantics

Fall 2026

Roadmap

I	One vector in both slots	what a quadratic form keeps; polarization
II	Quadratic forms at work	Mahalanobis, label propagation, Lorentz; signature
III	From a quadratic form to a product	declare Q , demand $\mathbf{v}^2 = Q(\mathbf{v})$; both halves follow
IV	What the signature decides	complex, split-complex, dual numbers; rotors; mirrors
V	Toward FGA	a declared Q for meaning; grades and slots; binding
VI	Close	summary, exercises, reading

Sessions 10 and 11 split every bilinear form into a symmetric half and an antisymmetric half. Today we feed the same vector to both slots of a form, see which half survives, and follow that survivor all the way to the Clifford algebra.

Every number in this deck is arithmetic you can check by hand.

Where We Are

Recall the attention head from Session 11: two tokens $\mathbf{x}_1 = (3, 1)$ and $\mathbf{x}_2 = (1, 3)$, and a learned form $G = W_Q^\top W_K$, split into its two halves.

$$G = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}}_{S, \text{ symmetric}} + \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{A, \text{ antisymmetric}}$$

	1 $\rightarrow$ 2	2 $\rightarrow$ 1	
$\mathbf{x}^\top S \mathbf{y}$	16	16	how alike
$\mathbf{x}^\top A \mathbf{y}$	+8	-8	which way

The fold of HRR keeps S and destroys A ; a VSA puts order back by stipulating role atoms. Session 11 derived a product that keeps both, $\mathbf{ab} = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b}$ – but only for the ordinary dot product.

Session 10 listed three sources for the matrix in the middle: *learn* it, *assume* it, or *declare* it. Geometric algebra declares it. Today: what exactly is declared.

The Question for Today

A puzzle

Declaring a geometric algebra means declaring a signature, and a signature fixes only a *symmetric* form – the half that is blind to order. Yet the algebra built from it keeps order: $\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}$. **Where does the other half come from?**

The answer comes in three steps, one per part of the deck:

- 1 Put one vector into both slots of a bilinear form. What comes out is a **quadratic form**, and it sees only the symmetric half.
- 2 A quadratic form is the symmetric half, no more and no less: **polarization** recovers it exactly.
- 3 Demand a product in which every vector squares to its quadratic form. Polarization then fixes the symmetric half of the product, and the antisymmetric half comes along unconstrained, as a new kind of object.

So declaring Q is enough, because the algebra supplies the half that Q cannot see.

Part I

One Vector in Both Slots

Feed the Same Vector to Both Slots

A bilinear form takes two vectors. Give it the same vector twice:

$$B(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top G \mathbf{y} \quad \rightsquigarrow \quad Q(\mathbf{x}) = B(\mathbf{x}, \mathbf{x}) = \mathbf{x}^\top G \mathbf{x}$$

Q is the **quadratic form** of B : a function of one vector, quadratic in its coordinates. For the head:

	$G\mathbf{x}$	$Q(\mathbf{x}) = \mathbf{x}^\top (G\mathbf{x})$
$\mathbf{x}_1 = (3, 1)$	$(5, 1)$	$3 \cdot 5 + 1 \cdot 1 = 16$
$\mathbf{x}_2 = (1, 3)$	$(7, 3)$	$1 \cdot 7 + 3 \cdot 3 = 16$

Written as a polynomial, $Q(\mathbf{x}) = x_1^2 + 2x_1x_2 + x_2^2 = (x_1 + x_2)^2$. Under this head both tokens have squared length 16.

You have met this already: $\mathbf{x}^\top \mathbf{x} = \|\mathbf{x}\|^2$ is the quadratic form of the dot product, and every squared length in the course has been a quadratic form.

The Antisymmetric Half Disappears

For any antisymmetric A , the number $\mathbf{x}^\top A \mathbf{x}$ equals its own transpose:

$$\mathbf{x}^\top A \mathbf{x} = (\mathbf{x}^\top A \mathbf{x})^\top = \mathbf{x}^\top A^\top \mathbf{x} = -\mathbf{x}^\top A \mathbf{x} \implies \mathbf{x}^\top A \mathbf{x} = 0$$

So $Q(\mathbf{x}) = \mathbf{x}^\top S \mathbf{x} + \mathbf{x}^\top A \mathbf{x} = \mathbf{x}^\top S \mathbf{x}$. The quadratic form sees only the symmetric half, and different forms with the same S give one and the same Q :

matrix	$Q(3, 1)$	$Q(1, -1)$	$Q(2, 5)$
$G = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, the head	16	0	49
$G^\top = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$, the head reversed	16	0	49
$S = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, its symmetric half	16	0	49
$\begin{pmatrix} 1 & 7 \\ -5 & 1 \end{pmatrix} = S + 6A$	16	0	49

A head and its reverse attend in opposite directions, and their quadratic forms are identical. Whatever a quadratic form records, it is not direction.

Polarization: The Symmetric Half Comes Back

Nothing has been lost from S , though. Given only the function Q , the symmetric half can be rebuilt:

$$S(\mathbf{x}, \mathbf{y}) = \frac{1}{2} [Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})]$$

This is the **polarization identity**. Run it on the two tokens:

$Q(\mathbf{x}_1 + \mathbf{x}_2) = Q(4, 4)$	$(4 + 4)^2 = 64$
$Q(\mathbf{x}_1), Q(\mathbf{x}_2)$	16, 16
$\frac{1}{2} (64 - 16 - 16)$	16 $= \mathbf{x}_1^\top S \mathbf{x}_2$

The 16 comes back. The 8 does not, and no formula in Q alone could return it: G and $G^\top$ have the same Q but opposite A .

So a quadratic form is a symmetric bilinear form seen along the diagonal – and the diagonal is enough to rebuild the rest of the symmetric half.

Why Polarization Works

Expand $Q(\mathbf{x} + \mathbf{y})$ with any G . The cross terms arrive in *both* orders:

$$Q(\mathbf{x} + \mathbf{y}) = \mathbf{x}^\top G \mathbf{x} + \underbrace{\mathbf{x}^\top G \mathbf{y} + \mathbf{y}^\top G \mathbf{x}}_{= 2 \mathbf{x}^\top S \mathbf{y}} + \mathbf{y}^\top G \mathbf{y}$$

Adding a quantity to its order-swapped copy is the half-sum of Session 11, doubled: the antisymmetric parts cancel and the symmetric parts add. With the head, $24 + 8 = 32 = 2 \cdot 16$.

The same information, two ways

A quadratic form determines exactly one symmetric bilinear form, by polarization, and a symmetric bilinear form determines its quadratic form, by putting $\mathbf{x}$ in both slots. [Declaring one is declaring the other.](#)

The $\frac{1}{2}$ needs division by 2, which the real numbers allow; over arithmetic modulo 2 the two notions come apart, a case this course never needs.

What Counts as a Quadratic Form

A function Q on vectors is a quadratic form when (i) $Q(c\mathbf{x}) = c^2 Q(\mathbf{x})$, and (ii) the polarization $Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})$ is bilinear. Equivalently: $Q(\mathbf{x}) = \mathbf{x}^\top S \mathbf{x}$ for some symmetric S .

function	quadratic form?	why
$x_1^2 + x_2^2$	yes	$S = I$: squared length
$(x_1 + x_2)^2$	yes	$S = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$: the head
$x_1^2 + 4x_1x_2 + x_2^2$	yes	$S = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$: Session 10's indefinite form
x_1x_2	yes	$S = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$
$\sqrt{x_1^2 + x_2^2}$	no	scales by $ c $, not c^2
$x_1^2 + x_2$	no	not homogeneous
$ x_1 + x_2 $	no	scales by $ c $; polarization not bilinear

The length is not a quadratic form; its square is. That is why squared lengths, not lengths, carry the algebra.

Reading the Matrix Off the Polynomial

Every quadratic form is a polynomial with only degree-2 terms. Its symmetric matrix is read off directly: the coefficients of the squares go on the diagonal, and half of each cross coefficient goes in each off-diagonal cell.

$$Q(\mathbf{x}) = \sum_i S_{ii} x_i^2 + \sum_{i < j} 2 S_{ij} x_i x_j$$

polynomial	symmetric matrix	a few values
$x_1^2 + 4x_1x_2 + x_2^2$	$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$	$Q(1, 1) = 6, \quad Q(1, -1) = -2, \quad Q(1, \sqrt{3} - 2) = 0$
$3x_1^2 - 2x_1x_2 + 5x_2^2$	$\begin{pmatrix} 3 & -1 \\ -1 & 5 \end{pmatrix}$	$Q(1, 0) = 3, \quad Q(1, 1) = 6$
x_1x_2	$\begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$	$Q(1, 1) = 1, \quad Q(1, -1) = -1$

Many matrices produce the same polynomial – add any antisymmetric A – but **exactly one of them is symmetric**. That one is what a quadratic form *is*, written as a matrix.

The first row is Session 10's indefinite form: positive, negative, and zero values, all from one polynomial.

Part II

Quadratic Forms at Work

Mahalanobis: Length Measured Against the Data

Given data with mean μ and covariance Σ , measure how far a point lies from the data with the quadratic form of Σ^{-1} :

$$Q_{\Sigma}(\mathbf{x}) = (\mathbf{x} - \mu)^{\top} \Sigma^{-1} (\mathbf{x} - \mu)$$

With $\Sigma = I$ this is squared Euclidean distance. Take instead Session 10's correlated matrix as the covariance, $\Sigma = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$, so $\Sigma^{-1} \approx \begin{pmatrix} 2.78 & -2.22 \\ -2.22 & 2.78 \end{pmatrix}$:

offset from the mean	Euclidean, squared	Mahalanobis, squared
(1, 1), along the correlation	2	1.11
(1, -1), against it	2	10.00

Same Euclidean distance, yet the second point is three times as far (nine times, squared), because it breaks the pattern the data follows.

In NLP: Podolskiy et al. (2021) fit class means and a shared covariance to transformer embeddings of in-domain utterances, then flag out-of-domain intents by this distance, and report new state-of-the-art results on all the datasets they tested.

The Graph Laplacian: How Much a Labeling Disagrees

On a graph, the Laplacian $L = D - W$ has a quadratic form that sums squared disagreements across edges, $\mathbf{x}^\top L \mathbf{x} = \sum_{(i,j) \in \text{edges}} (x_i - x_j)^2$. A toy synonym chain:



labeling	good	decent	okay	mediocre	bad	$\mathbf{x}^\top L \mathbf{x}$
abrupt	1	1	1	-1	-1	4.00
smooth	1	0.5	0	-0.5	-1	1.00
constant	1	1	1	1	1	0.00

Label propagation clamps seed words (good = +1, bad = -1) and minimizes this quadratic form over the rest. The minimum is the smooth row: okay lands at 0, neutral. Rao and Ravichandran (2009) induce polarity lexicons this way, over WordNet synonym graphs.

The constant labeling costs nothing: the form measures disagreement, never level. That is why seeds are needed.

The Lorentz Form: Room for a Taxonomy

Hyperbolic embeddings of hierarchies use an *indefinite* quadratic form, one minus sign and the rest plus:

$$Q_L(\mathbf{x}) = -x_0^2 + x_1^2 + x_2^2, \quad \mathbb{H}^2 = \{\mathbf{x} : Q_L(\mathbf{x}) = -1, x_0 > 0\}$$

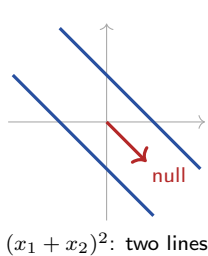
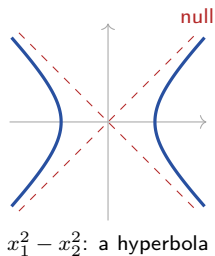
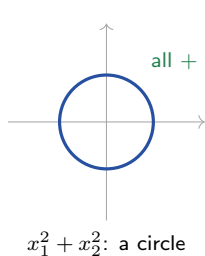
Distance on the surface is $d(\mathbf{x}, \mathbf{y}) = \operatorname{arcosh}(-B_L(\mathbf{x}, \mathbf{y}))$, where B_L is the polarization of Q_L . Worked points:

point	coordinates	Q_L	distance
$\mathbf{o}$, the origin	(1, 0, 0)	-1	-
$\mathbf{p}$	(cosh 2, sinh 2, 0) $\approx$ (3.76, 3.63, 0)	-1	$d(\mathbf{o}, \mathbf{p}) = 2$
$\mathbf{q}$, the mirror image of $\mathbf{p}$	(cosh 2, -sinh 2, 0)	-1	$d(\mathbf{p}, \mathbf{q}) = 4$
a null vector	(1, 1, 0)	0	not on the surface

In NLP: Nickel and Kiela (2018) embed the WordNet noun hierarchy on this surface in a handful of dimensions. General concepts such as entity sit near $\mathbf{o}$, specific ones such as chihuahua far out, and the room available grows exponentially with distance, as a branching taxonomy needs.

Three Kinds of Direction

Draw the set where $Q(\mathbf{x}) = 1$. Its shape *is* the quadratic form:



a direction with	meaning	examples from Part II
$Q(\mathbf{v}) > 0$	positive squared length	any Euclidean or Mahalanobis offset
$Q(\mathbf{v}) < 0$	negative squared length	the Lorentz axis x_0
$Q(\mathbf{v}) = 0, \mathbf{v} \neq \mathbf{0}$	null : nonzero, length zero	a constant labeling; $(1, 1, 0)$; the head's $(1, -1)$

Sylvester: Only the Count of Signs Survives

Any symmetric S can be brought, by a change of basis, to a diagonal matrix with entries $+1$, -1 , and 0 only. **Sylvester's law of inertia**: the counts (p, q, r) of each do not depend on the basis chosen. They are the **signature** of the form, written $\text{Cl}(p, q, r)$.

quadratic form	eigenvalues of S	signature
squared Euclidean length on $\mathbb{R}^n$	all 1	$\text{Cl}(n, 0, 0)$
Mahalanobis, Σ^{-1} above	0.56, 5.00	$\text{Cl}(2, 0, 0)$
$x_1^2 + 4x_1x_2 + x_2^2$	3, -1	$\text{Cl}(1, 1, 0)$
x_1x_2	$\frac{1}{2}$, $-\frac{1}{2}$	$\text{Cl}(1, 1, 0)$
Lorentz, $-x_0^2 + x_1^2 + x_2^2$	1, 1, -1	$\text{Cl}(2, 1, 0)$
Laplacian of the five-word chain	0, 0.38, 1.38, 2.62, 3.62	$\text{Cl}(4, 0, 1)$
the head, $(x_1 + x_2)^2$	2, 0	$\text{Cl}(1, 0, 1)$
a tied head, $W_Q^\top W_Q$ of rank k	k positive, the rest 0	$\text{Cl}(k, 0, n - k)$

The signature is everything about a quadratic form that no change of coordinates can alter. Session 10 introduced $\text{Cl}(p, q, r)$ as a name for a form; Part III shows why the letters are Cl.

The Head's Quadratic Form Is Degenerate

The head's $S = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ has eigenvalues 2 and 0: signature $\text{Cl}(1, 0, 1)$, with null direction $(1, -1)$. Now look at the two tokens:

$$\mathbf{x}_1 - \mathbf{x}_2 = (2, -2), \quad Q(\mathbf{x}_1 - \mathbf{x}_2) = (2 - 2)^2 = 0$$

The two tokens differ by a null vector. Under the symmetric half they cannot be told apart by any comparison at all:

$$S(\mathbf{x}_1, \mathbf{y}) = 4(y_1 + y_2) = S(\mathbf{x}_2, \mathbf{y}) \text{ for every } \mathbf{y}, \quad \cos_S(\mathbf{x}_1, \mathbf{x}_2) = \frac{16}{\sqrt{16 \cdot 16}} = 1$$

Everything that distinguishes token 1 from token 2 for this head lives in A – exactly the half its quadratic form cannot see.

Why this matters

Session 11 found that a head with tied weights, $W_Q = W_K$, cannot break a tie between two equally similar tokens. The quadratic form shows how complete the tie is: for the symmetric half, these two tokens are one direction.

In-Class Exercise 1

- 1 For $G = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix}$, write $Q(\mathbf{x}) = \mathbf{x}^\top G \mathbf{x}$ as a polynomial, and give the symmetric matrix that represents it.
- 2 Using *only* values of that Q , compute $S((1, 0), (0, 1))$ by polarization. Check it against your matrix.
- 3 Find the signature of $Q(\mathbf{x}) = x_1 x_2$ and two null directions.
- 4 ★ On the chain good–decent–okay–mediocre–bad, why can the energy $\mathbf{x}^\top L \mathbf{x}$ alone never decide whether okay is positive or negative? What do the seeds change, in terms of the signature?

Six minutes. Item 4 is the starred one.

In-Class Exercise 1: Answers

1. $Q = 2x_1^2 + (3 - 1)x_1x_2 + x_2^2 = 2x_1^2 + 2x_1x_2 + x_2^2$, so $S = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. The 3 and the -1 are averaged; their difference, the antisymmetric part, is gone.
2. $Q(1, 1) = 5$, $Q(1, 0) = 2$, $Q(0, 1) = 1$, so $S = \frac{1}{2}(5 - 2 - 1) = 1$, the off-diagonal entry of the matrix.
3. $S = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$ has eigenvalues $\frac{1}{2}$ and $-\frac{1}{2}$: signature $\text{Cl}(1, 1, 0)$. The axes are null, $Q(1, 0) = Q(0, 1) = 0$, while $Q(1, 1) = 1$ and $Q(1, -1) = -1$.
4. ★ Adding the same constant to every label leaves every difference $x_i - x_j$ unchanged, so the energy cannot see the level: the constant vector is a null direction, and the chain's signature is $\text{Cl}(4, 0, 1)$. The seeds fix two values, which removes that null direction; the remaining problem has a unique minimum, with okay at 0.

Part III

From a Quadratic Form to a Product

Declare Q , Then Ask for a Product

Session 11 asked for a product whose two halves were the dot product and the wedge. Ask instead for something that mentions only the quadratic form:

The specification

Find an associative product of vectors, distributive over addition, such that **every vector squares to its quadratic form**:

$$\mathbf{v} \mathbf{v} = Q(\mathbf{v}) \quad \text{for every vector } \mathbf{v}$$

That single rule is the whole definition of the **Clifford algebra** $\text{Cl}(V, Q)$, which Clifford introduced in 1878 by combining Grassmann's algebra of the wedge with Hamilton's quaternions.

Notice what is *not* in the specification: no antisymmetric form, no wedge, no orientation. [In the programming picture we will use later in the course](#), Q is the specification and everything else is derived from it.

The left side is a product of vectors; the right side is a number. A product of two vectors is therefore allowed to be a number plus something else – and the next two slides find out what else.

Polarization, Inside the Product

Apply the rule to a sum, $\mathbf{u} + \mathbf{v}$, and expand by distributivity, keeping the order of each product:

$$(\mathbf{u} + \mathbf{v})(\mathbf{u} + \mathbf{v}) = \mathbf{u}\mathbf{u} + \mathbf{u}\mathbf{v} + \mathbf{v}\mathbf{u} + \mathbf{v}\mathbf{v} = Q(\mathbf{u} + \mathbf{v})$$

Replace $\mathbf{u}\mathbf{u}$ and $\mathbf{v}\mathbf{v}$ by $Q(\mathbf{u})$ and $Q(\mathbf{v})$, and solve:

$$\mathbf{u}\mathbf{v} + \mathbf{v}\mathbf{u} = Q(\mathbf{u} + \mathbf{v}) - Q(\mathbf{u}) - Q(\mathbf{v}) = 2S(\mathbf{u}, \mathbf{v})$$

The **symmetric half of the product is forced**: it is the polarization of the declared Q . With the head's form and its two tokens, $(\mathbf{x}_1 + \mathbf{x}_2)^2 = 64$ and $\mathbf{x}_1^2 = \mathbf{x}_2^2 = 16$, so $\mathbf{x}_1\mathbf{x}_2 + \mathbf{x}_2\mathbf{x}_1 = 32$.

The polarization identity of Part I has reappeared inside the product. Squaring a sum produces both orders of the cross term, and adding the two orders is how a quadratic form returns its symmetric half.

The Other Half Is Left Free

The rule says nothing about the difference $\mathbf{uv} - \mathbf{vu}$. The algebra keeps it, as an object of a new kind, the **wedge**:

$$\mathbf{u} \wedge \mathbf{v} := \frac{1}{2} (\mathbf{uv} - \mathbf{vu}),$$

$$\mathbf{uv} = S(\mathbf{u}, \mathbf{v}) + \mathbf{u} \wedge \mathbf{v}$$

half	determined by	grade	answers
symmetric, $S(\mathbf{u}, \mathbf{v})$	the declared Q , by polarization	0, a number	how alike
antisymmetric, $\mathbf{u} \wedge \mathbf{v}$	the two vectors alone: no metric needed	2, a bivector	which way

Session 11 derived $\mathbf{ab} = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b}$ for the ordinary dot product. Nothing in the derivation above used the dot product: **the same formula holds for every quadratic form**, with S in place of the dot.

This answers the puzzle. The declaration fixes only Q . The antisymmetric half was never declared, because it never needed declaring: the wedge of two vectors does not depend on Q at all.

The Head's Two Halves, Recomputed

Declare the head's quadratic form, $Q(\mathbf{x}) = (x_1 + x_2)^2$, and multiply the two tokens in its Clifford algebra:

$$\mathbf{x}_1 \mathbf{x}_2 = \underbrace{16}_{S(\mathbf{x}_1, \mathbf{x}_2)} + \underbrace{8 \mathbf{e}_1 \wedge \mathbf{e}_2}_{\mathbf{x}_1 \wedge \mathbf{x}_2}, \quad \mathbf{x}_2 \mathbf{x}_1 = 16 - 8 \mathbf{e}_1 \wedge \mathbf{e}_2$$

since $\mathbf{x}_1 \wedge \mathbf{x}_2 = (3 \cdot 3 - 1 \cdot 1) \mathbf{e}_1 \wedge \mathbf{e}_2$. Set that beside what the head computed:

	the head's score	the Clifford product
symmetric half, $1 \rightarrow 2$ and $2 \rightarrow 1$	16, 16	16, 16 (declared, via Q)
antisymmetric half, $1 \rightarrow 2$ and $2 \rightarrow 1$	+8, -8	+8, -8 (computed, not declared)

The head had to learn both halves. The algebra was told only the first. The second agrees because in the plane every antisymmetric form is a multiple of the wedge, and this head happened to learn the multiple 1.

Under this Q , $\mathbf{e}_1$ and $\mathbf{e}_2$ are not orthogonal, so $\mathbf{e}_1 \mathbf{e}_2 \neq \mathbf{e}_1 \wedge \mathbf{e}_2$; the wedge is written out explicitly for that reason.

A Basis Gives the Whole Multiplication Table

Choose a basis that is orthogonal for S . Polarization in the product gives the two rules that generate everything:

$$\mathbf{e}_i \mathbf{e}_j = -\mathbf{e}_j \mathbf{e}_i \quad (i \neq j), \quad \mathbf{e}_i \mathbf{e}_i = Q(\mathbf{e}_i)$$

By Sylvester, the basis can be scaled so that each $Q(\mathbf{e}_i)$ is $+1$, -1 , or 0 . With p , q , and r of each, the algebra is $\text{Cl}(p, q, r)$ – the name Session 10 gave the form, now the name of the algebra it generates.

The head's algebra. $\mathbf{g}_1 = (\frac{1}{2}, \frac{1}{2})$ and $\mathbf{g}_2 = (1, -1)$ are orthogonal for S , with $Q(\mathbf{g}_1) = 1$ and $Q(\mathbf{g}_2) = 0$: the algebra is $\text{Cl}(1, 0, 1)$.

$$\mathbf{x}_1 = 4\mathbf{g}_1 + \mathbf{g}_2, \quad \mathbf{x}_2 = 4\mathbf{g}_1 - \mathbf{g}_2$$

coordinates in the new basis

$$\mathbf{x}_1 \mathbf{x}_2 = 16 \mathbf{g}_1 \mathbf{g}_1 - 4 \mathbf{g}_1 \mathbf{g}_2 + 4 \mathbf{g}_2 \mathbf{g}_1 - \mathbf{g}_2 \mathbf{g}_2 = 16 - 8 \mathbf{g}_1 \mathbf{g}_2$$

the same product, from the table

$$(\mathbf{x}_1 - \mathbf{x}_2)^2 = (2\mathbf{g}_2)^2 = 0$$

the null difference again

And $\mathbf{g}_1 \mathbf{g}_2 = \mathbf{g}_1 \wedge \mathbf{g}_2 = -\mathbf{e}_1 \wedge \mathbf{e}_2$, so $-8 \mathbf{g}_1 \mathbf{g}_2$ is the $+8 \mathbf{e}_1 \wedge \mathbf{e}_2$ of the last slide.

What the Tensor Algebra Would Have Kept

The outer product never forgets: $\mathbf{e}_2 \otimes \mathbf{e}_1$, $\mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2$, and so on are all independent, and products of degree k need n^k coordinates – the order growth of Session 11b, and the cost of TPR.

The Clifford rule cuts this down. Every word in the basis vectors reduces by two moves: [swap neighbors](#) (a sign, $\mathbf{e}_j \mathbf{e}_i = -\mathbf{e}_i \mathbf{e}_j$) and [cancel equal neighbors](#) ($\mathbf{e}_i \mathbf{e}_i = Q(\mathbf{e}_i)$). Take the word $\mathbf{e}_2 \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_2 \mathbf{e}_1$:

declared signature	reduction	value
$\text{Cl}(2, 0, 0): \mathbf{e}_1^2 = \mathbf{e}_2^2 = +1$	$\mathbf{e}_2 \mathbf{e}_1 (\mathbf{e}_2 \mathbf{e}_2) \mathbf{e}_1 = \mathbf{e}_2 (\mathbf{e}_1 \mathbf{e}_1)$	$\mathbf{e}_2$
$\text{Cl}(1, 1, 0): \mathbf{e}_2^2 = -1$	$\mathbf{e}_2 \mathbf{e}_1 (-1) \mathbf{e}_1 = -\mathbf{e}_2 (\mathbf{e}_1 \mathbf{e}_1)$	$-\mathbf{e}_2$
$\text{Cl}(1, 0, 1): \mathbf{e}_2^2 = 0$	$\mathbf{e}_2 \mathbf{e}_1 (0) \mathbf{e}_1$	0

One word, three values. The tensor algebra cannot evaluate it; the Clifford algebra can, and what it returns depends on nothing but the declared quadratic form.

Formally, $\text{Cl}(V, Q)$ is the tensor algebra with $\mathbf{v} \otimes \mathbf{v} - Q(\mathbf{v})$ set to zero for every $\mathbf{v}$. Every word reduces to $\pm$ a sorted product of *distinct* basis vectors, or to 0.

Tensors Grow; the Clifford Algebra Stays Put

A reduced word is a choice of *which* basis vectors appear: 2^n of them, sorted by **grade**, the number of vectors in each.

$n = 4$, degree or grade k	0	1	2	3	4	total
tensor products, n^k	1	4	16	64	256	keeps growing
Clifford grade k , $\binom{n}{k}$	1	4	6	4	1	$16 = 2^4$

A product of two vectors lands in grades 0 and 2: **the symmetric half in grade 0**, **the antisymmetric half in grade 2**, each in a place of its own. At $n = 300$, grades 0 through 2 hold $1 + 300 + 44,850 = 45,151$ coordinates, against 90,000 for a 300×300 matrix. **The saving has a price**: grade 2 is alternating, so it drops the symmetric and diagonal structure that a general relation needs (Part V).

	size	antisymmetric half	unbinding
TPR	grows as n^k	kept	exact, with orthonormal roles
VSA (HRR, MAP)	fixed, d	folded away	approximate; needs cleanup
Clifford algebra	fixed, 2^n , graded	kept, in grade 2	exact, with orthonormal roles

Part IV

What the Signature Decides

One Generator, Three Algebras

Start with a single basis vector $\mathbf{e}$. The algebra has elements $a + b\mathbf{e}$, and the declared sign of $\mathbf{e}^2$ is the only choice there is:

$$(a + b\mathbf{e})(c + d\mathbf{e}) = (ac + bd\mathbf{e}^2) + (ad + bc)\mathbf{e}$$

$\mathbf{e}^2$	algebra	worked	known as
+1	$\text{Cl}(1, 0, 0)$	$(2 + 3\mathbf{e})(1 + 4\mathbf{e}) = 14 + 11\mathbf{e}$	split-complex numbers
-1	$\text{Cl}(0, 1, 0)$	$(2 + 3\mathbf{e})(1 + 4\mathbf{e}) = -10 + 11\mathbf{e}$	the complex numbers
0	$\text{Cl}(0, 0, 1)$	$(3 + \mathbf{e})(3 + \mathbf{e}) = 9 + 6\mathbf{e}$	dual numbers

One sign, three different worlds. The next two slides show what each world is good for: +1 splits things into two clean pieces, -1 turns them, and 0 carries a rate of change.

The complex numbers are not postulated here. They are the Clifford algebra of a single direction declared to have squared length -1 .

$e^2 = +1$: Two Complementary Projectors

When $e^2 = +1$, the algebra contains two elements that split it cleanly in two:

$$\mathbf{p}_+ = \frac{1}{2}(1 + \mathbf{e}), \quad \mathbf{p}_- = \frac{1}{2}(1 - \mathbf{e})$$

property	check	meaning
$\mathbf{p}_\pm^2 = \mathbf{p}_\pm$	$\frac{1}{4}(1 \pm 2\mathbf{e} + \mathbf{e}^2) = \frac{1}{2}(1 \pm \mathbf{e})$	each is a projector
$\mathbf{p}_+\mathbf{p}_- = 0$	$\frac{1}{4}(1 - \mathbf{e}^2) = 0$	exclusive
$\mathbf{p}_+ + \mathbf{p}_- = 1$	by inspection	exhaustive

In FGA: the position paper analyzes a polar pair such as alive/dead with exactly these idempotents on a positive-signature axis. The two predicates partition the domain, and nothing belongs to both.

The sign matters. With $e^2 = -1$ the same formulas give $\mathbf{p}_+\mathbf{p}_- = \frac{1}{2}$ and $\mathbf{p}_+^2 = \frac{1}{2}\mathbf{e}$: no projectors at all. A semantic space that needs exclusive, exhaustive pairs has to declare that axis positive.

$\mathbf{e}^2 = -1$ and $\mathbf{e}^2 = 0$: Turning and Rate

Turning. With $\mathbf{e}^2 = -1$, multiplying by $\cos \theta + \mathbf{e} \sin \theta$ rotates by θ :

$$(\cos \theta + \mathbf{e} \sin \theta)(a + b \mathbf{e}) = (a \cos \theta - b \sin \theta) + (a \sin \theta + b \cos \theta) \mathbf{e}$$

Repeat it and you come back where you started: a direction of squared length -1 generates cycles.

Rate. With $\mathbf{e}^2 = 0$, write ε for $\mathbf{e}$. Expanding any polynomial at $a + \varepsilon$, every term with ε^2 vanishes, and what survives is the value and the derivative:

$$f(a + \varepsilon) = f(a) + f'(a) \varepsilon, \quad \text{e.g.} \quad (2 + \varepsilon)^3 = 8 + 12 \varepsilon$$

and indeed $\frac{d}{dx} x^3 = 3x^2 = 12$ at $x = 2$. This is **forward-mode automatic differentiation**: PyTorch's interface for it calls these objects dual tensors.

A null direction is a nonzero quantity whose square is zero – an infinitesimal, which records the direction of a change without contributing to its size.

Two Generators: The Sign of the Bivector

With two orthogonal basis vectors, the bivector $\mathbf{e}_1\mathbf{e}_2$ appears, and its square is decided by the declared signs:

$$(\mathbf{e}_1\mathbf{e}_2)^2 = \mathbf{e}_1\mathbf{e}_2\mathbf{e}_1\mathbf{e}_2 = -\mathbf{e}_1\mathbf{e}_1\mathbf{e}_2\mathbf{e}_2 = -Q(\mathbf{e}_1)Q(\mathbf{e}_2)$$

signature	$\mathbf{e}_1^2, \mathbf{e}_2^2$	$(\mathbf{e}_1\mathbf{e}_2)^2$	what lives there
Cl(2, 0, 0)	+1, +1	-1	rotations of the plane; the complex numbers inside it
Cl(1, 1, 0)	+1, -1	+1	boosts, as in special relativity
Cl(0, 2, 0)	-1, -1	-1	the quaternions
Cl(1, 0, 1)	+1, 0	0	translations

Quaternions, for free. In Cl(0, 2, 0) put $i = \mathbf{e}_1$, $j = \mathbf{e}_2$, $k = \mathbf{e}_1\mathbf{e}_2$. Then $i^2 = j^2 = k^2 = ijk = -1$ and $ij = k = -ji$: Hamilton's rules of 1843 are the multiplication table of a declared quadratic form with two minus signs.

Exponentials of Bivectors

Exponentiate a bivector B by its power series, $e^{B\theta} = \sum_k (B\theta)^k / k!$. The even powers are numbers and the odd powers are multiples of B , so the result depends only on B^2 :

	$B^2 = -1$	$B^2 = +1$	$B^2 = 0$
$e^{B\theta}$	$\cos \theta + B \sin \theta$	$\cosh \theta + B \sinh \theta$	$1 + B\theta$
at $\theta = 0.7$	$0.765 + 0.644 B$	$1.255 + 0.759 B$	$1 + 0.7 B$
kind	elliptic	hyperbolic	parabolic

The even/odd split of Session 11, once more: $\cosh$ and $\sinh$ were the even and odd halves of e^x , and here they are the scalar and bivector parts of $e^{B\theta}$ when $B^2 = +1$.

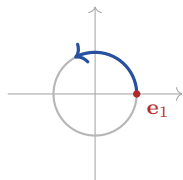
A **rotor** is such an exponential, $R = e^{-B\theta/2}$. It acts on a vector by the sandwich

$$\mathbf{x} \longmapsto R \mathbf{x} \tilde{R},$$

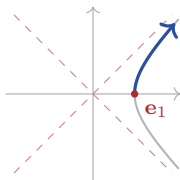
where $\tilde{R}$, the **reverse**, writes every product in the opposite order; for $R = \cos \frac{\theta}{2} - B \sin \frac{\theta}{2}$ it is $\cos \frac{\theta}{2} + B \sin \frac{\theta}{2}$. The next slide shows where the three kinds of rotor send $\mathbf{e}_1$.

Rotors Move Vectors Along the Unit Set

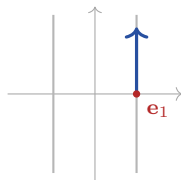
Take $B = \mathbf{e}_1 \mathbf{e}_2$ in the three two-dimensional signatures, and follow $\mathbf{e}_1$ as θ grows:



$\text{Cl}(2, 0, 0)$: a circle



$\text{Cl}(1, 1, 0)$: a hyperbola



$\text{Cl}(1, 0, 1)$: a line

signature	$R \mathbf{e}_1 \tilde{R}$	its quadratic form
$\text{Cl}(2, 0, 0)$	$\cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2$	$\cos^2 \theta + \sin^2 \theta = 1$
$\text{Cl}(1, 1, 0)$	$\cosh \theta \mathbf{e}_1 + \sinh \theta \mathbf{e}_2$	$\cosh^2 \theta - \sinh^2 \theta = 1$
$\text{Cl}(1, 0, 1)$	$\mathbf{e}_1 + \theta \mathbf{e}_2$	$1 + 0 = 1$

Each rotor slides vectors along the set $Q = 1$ of its own quadratic form. The null case is a translation, which is how the projective algebra $\text{Cl}(3, 0, 1)$ – the basis of the Geometric Algebra Transformer (Brehmer et al. 2023) – writes translations as products.

Worked: A Rotation in $Cl(2, 0, 0)$

Take $B = \mathbf{e}_1\mathbf{e}_2$, with $B^2 = -1$. The rotor for angle θ is

$$R = e^{-B\theta/2} = \cos \frac{\theta}{2} - \mathbf{e}_1\mathbf{e}_2 \sin \frac{\theta}{2}, \quad \tilde{R} = \cos \frac{\theta}{2} + \mathbf{e}_1\mathbf{e}_2 \sin \frac{\theta}{2}$$

and the sandwich gives $R\mathbf{e}_1\tilde{R} = \cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2$. Worked cases:

θ	R	input	output
90°	$\frac{1}{\sqrt{2}}(1 - \mathbf{e}_1\mathbf{e}_2)$	$\mathbf{e}_1$	$\mathbf{e}_2$
90°	$\frac{1}{\sqrt{2}}(1 - \mathbf{e}_1\mathbf{e}_2)$	$3\mathbf{e}_1 + \mathbf{e}_2$	$-\mathbf{e}_1 + 3\mathbf{e}_2$
60°	$0.866 - 0.5 \mathbf{e}_1\mathbf{e}_2$	$\mathbf{e}_1$	$0.5 \mathbf{e}_1 + 0.866 \mathbf{e}_2$

Why $\theta/2$: the vector is multiplied twice, once on each side, and each side turns it by half. Every output has the same squared length as its input, $Q = 1$ or $Q = 10$, because the rotor was built from the declared Q .

The same sandwich rotates any multivector, not only vectors – including the bivectors and higher-grade objects that FGA uses for relations and role bindings.

Worked: A Boost in $Cl(1, 1, 0)$

Declare $\mathbf{e}_1^2 = +1$ and $\mathbf{e}_2^2 = -1$. Now $B = \mathbf{e}_1\mathbf{e}_2$ has $B^2 = +1$, and the same recipe gives hyperbolic functions:

$$R = e^{-B\varphi/2} = \cosh \frac{\varphi}{2} - \mathbf{e}_1\mathbf{e}_2 \sinh \frac{\varphi}{2}, \quad R\mathbf{e}_1\tilde{R} = \cosh \varphi \mathbf{e}_1 + \sinh \varphi \mathbf{e}_2$$

Take $\varphi = \ln 2$, so that $\cosh \varphi = 1.25$ and $\sinh \varphi = 0.75$:

R	$1.0607 - 0.3536 \mathbf{e}_1\mathbf{e}_2$
$R\mathbf{e}_1\tilde{R}$	$1.25 \mathbf{e}_1 + 0.75 \mathbf{e}_2$
its squared length	$1.25^2 (+1) + 0.75^2 (-1) = 1.5625 - 0.5625 = 1$

The Euclidean length of the output is 1.46, not 1. **What the boost preserves is the declared length**, not the length your eye measures on the page.

In FGA: the position paper models good knife with this boost, $\mathbf{e}_1$ a form axis and $\mathbf{e}_2$ a telic axis declared negative. Boosting the form factor and rebuilding the binding gives $(R\mathbf{e}_1\tilde{R}) \wedge \mathbf{e}_2 = \cosh \varphi (\mathbf{e}_1 \wedge \mathbf{e}_2)$: the same form-purpose plane, bound 1.25 times as strongly. The paper offers the elliptic/hyperbolic contrast as an encoding hypothesis for extensional versus intensional modification.

Mirrors Are Vectors, and Two Mirrors Make a Rotation

A vector $\mathbf{a}$ with $Q(\mathbf{a}) \neq 0$ has an inverse, $\mathbf{a}^{-1} = \mathbf{a}/Q(\mathbf{a})$, since $\mathbf{a}\mathbf{a} = Q(\mathbf{a})$. Then the sandwich

$$\mathbf{v} \longmapsto -\mathbf{a}\mathbf{v}\mathbf{a}^{-1}$$

reflects $\mathbf{v}$ in the line (or hyperplane) orthogonal to $\mathbf{a}$. Worked in $\text{Cl}(2, 0, 0)$:

step	mirror	result
start	–	$3\mathbf{e}_1 + \mathbf{e}_2$
reflect	$\mathbf{a} = \mathbf{e}_1$	$-3\mathbf{e}_1 + \mathbf{e}_2$
reflect again	$\mathbf{b} = \frac{1}{\sqrt{2}}(\mathbf{e}_1 + \mathbf{e}_2)$	$-\mathbf{e}_1 + 3\mathbf{e}_2$ the 90° rotation

Two reflections compose to a rotation by twice the angle between the mirrors, and the rotor is their product: $\mathbf{b}\mathbf{a} = \frac{1}{\sqrt{2}}(1 - \mathbf{e}_1\mathbf{e}_2)$, the rotor of the last slide but one. The quadratic form is the specification; products of vectors are the programs.

A null vector cannot be a mirror: with $Q(\mathbf{a}) = 0$ there is no $\mathbf{a}^{-1}$. The head's $\mathbf{g}_2$ is such a vector.

What the Quadratic Form Decided

declared	consequence in the algebra	example
$Q(\mathbf{a}) \neq 0$	$\mathbf{a}$ is invertible, so it can be a mirror	any Euclidean vector
$Q(\mathbf{a}) = 0, \mathbf{a} \neq 0$	no inverse; change without size	dual numbers; the head's $\mathbf{g}_2$
one axis, $\mathbf{e}^2 = +1$	projectors $\frac{1}{2}(1 \pm \mathbf{e})$	a polar pair, alive/dead
one axis, $\mathbf{e}^2 = -1$	cycles	the complex numbers
a bivector, $B^2 = -1$	elliptic rotors: rotations	$\text{Cl}(2, 0, 0), \text{Cl}(0, 2, 0)$
a bivector, $B^2 = +1$	hyperbolic rotors: boosts	$\text{Cl}(1, 1, 0)$
a bivector, $B^2 = 0$	translations	$\text{Cl}(1, 0, 1), \text{Cl}(3, 0, 1)$

Every row is a consequence of one declaration. [Choosing a signature is choosing which of these behaviors a space will have](#) – for physics, for graphics, or, in FGA, for meaning.

And none of it needed an antisymmetric form to be declared. The orientation in every rotor and bivector came from the product.

In-Class Exercise 2

- 1 In $\text{Cl}(2, 0, 0)$, compute $(\mathbf{e}_1 + 2\mathbf{e}_2)(3\mathbf{e}_1 - \mathbf{e}_2)$ by the two rules, and split the result into its symmetric and antisymmetric halves.
- 2 Repeat the computation in $\text{Cl}(1, 1, 0)$, where $\mathbf{e}_2^2 = -1$. Which half changed, and why that one?
- 3 Show that a vector with $Q(\mathbf{a}) \neq 0$ has inverse $\mathbf{a}/Q(\mathbf{a})$. What goes wrong when $Q(\mathbf{a}) = 0$?
- 4 ★ Verify that $\mathbf{p}_{\pm} = \frac{1}{2}(1 \pm \mathbf{e})$ satisfy $\mathbf{p}_+\mathbf{p}_- = 0$ when $\mathbf{e}^2 = +1$. A colleague proposes modeling a polar pair on an axis with $\mathbf{e}^2 = -1$ instead. What would go wrong?

Seven minutes. Item 2 is the one to discuss.

In-Class Exercise 2: Answers

1. $3\mathbf{e}_1\mathbf{e}_1 - \mathbf{e}_1\mathbf{e}_2 + 6\mathbf{e}_2\mathbf{e}_1 - 2\mathbf{e}_2\mathbf{e}_2 = 3 - \mathbf{e}_1\mathbf{e}_2 - 6\mathbf{e}_1\mathbf{e}_2 - 2 = 1 - 7\mathbf{e}_1\mathbf{e}_2$. Symmetric half 1 (the dot product, $3 - 2$); antisymmetric half $-7\mathbf{e}_1\mathbf{e}_2$ (the wedge, $1 \cdot (-1) - 2 \cdot 3$).
2. Now $\mathbf{e}_2\mathbf{e}_2 = -1$, so the result is $3 + 2 - 7\mathbf{e}_1\mathbf{e}_2 = 5 - 7\mathbf{e}_1\mathbf{e}_2$. Only the symmetric half changed: it is the polarization of Q , and Q is what was changed. The wedge does not depend on Q , so it stayed at -7 .
3. $\mathbf{a}(\mathbf{a}/Q(\mathbf{a})) = Q(\mathbf{a})/Q(\mathbf{a}) = 1$. If $Q(\mathbf{a}) = 0$, then $\mathbf{a}\mathbf{a} = 0$; were there an inverse $\mathbf{b}$, we would have $\mathbf{a} = \mathbf{a}(\mathbf{a}\mathbf{b}) = (\mathbf{a}\mathbf{a})\mathbf{b} = 0$, a contradiction.
4. $\star \quad \frac{1}{4}(1 + \mathbf{e})(1 - \mathbf{e}) = \frac{1}{4}(1 - \mathbf{e}^2) = 0$. With $\mathbf{e}^2 = -1$ the same product is $\frac{1}{2}$, and $\mathbf{p}_+^2 = \frac{1}{2}\mathbf{e} \neq \mathbf{p}_+$: the two predicates would overlap and neither would be a projector. That axis generates rotation, which suits a cyclic domain – the seasons, say – and not an exclusive pair.

Part V

Toward FGA

FGA Declares a Quadratic Form for Meaning

Functional Geometric Algebra (FGA) builds semantic representations in a Clifford algebra $\text{Cl}(V, Q)$. Its semantic space includes the quadratic form as a declared component, along with a decomposition of V into mutually orthogonal subspaces – for entities, their subject and object slots, predicate content, kinds, roles, states, and control – and a set of qualia rotors.

The position paper treats the signature as [a design parameter of the semantic space](#), and proposes:

squares to proposed use		behavior, from Part IV
+1	spatial, extensional, similarity-preserving	projectors; rotations, in pairs
-1	intensional, modal, contrastive	cycles; boosts, with a +1 axis
0	not used for meaning; projective geometry (GATr)	translations

For a first application it uses the Euclidean signature $\text{Cl}(n, 0, 0)$, with entity vectors normalized to unit length, so that the symmetric half of a product of two entities is their cosine.

This is the third option from Session 10, taken seriously: neither learn the metric nor assume it, but *declare* it, and let the declaration decide what the algebra can do.

Grades and Subspaces as Types

Session 11b mapped types to tensors of rising order. FGA maps them to **grades** of one algebra, and adds **subspace support**, because grade alone cannot separate e from $\langle e, t \rangle$:

type	example	tensor framework	FGA: grade, and where it lives
t	a truth value	order 0	grade 0, a scalar
e	Mary	order 1	grade 1, in V_E
$\langle e, t \rangle$, verb	laughs	order 1	grade 1, in $V_E \oplus V_P$
$\langle e, t \rangle$, noun	dog	order 1	grade 1, in $V_E \oplus V_K \oplus V_P$
$\langle e, \langle e, t \rangle \rangle$	loves	order 2, n^2 numbers	grade 2, in $V_E^s \wedge V_E^o$, n^2 numbers
k -place	gives, $k = 3$	order k , n^k numbers	grade k , one slot per argument

Application is still contraction. A predicate's V_E part meets the entity, as $\mathbf{p}^\top \mathbf{x}$ did in Session 6; its V_P content drops out, since the paper takes all the subspaces mutually orthogonal. A noun's V_K part, which a verb lacks, carries its criterion of individuation – what licenses *three dogs*. **Grade and subspace are both visible in the object**, read off by projection.

Why a Relation Needs Two Slots

Return to the head of Session 11. It scores $\mathbf{x}^\top G \mathbf{y}$ with $G = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = S + A$, where $S = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, on the tokens $\mathbf{x}_1 = (3, 1)$ and $\mathbf{x}_2 = (1, 3)$. Three ways to store such a relation:

store	holds	numbers	$(\mathbf{x}_1, \mathbf{x}_2)$	$(\mathbf{x}_2, \mathbf{x}_1)$	$(\mathbf{x}_1, \mathbf{x}_1)$
a symmetric form, all Q can see	S	$\frac{1}{2}n(n+1)$	16	16	16
a bivector in one space, $\text{Cl}^2(V_E)$	A	$\binom{n}{2}$	8	-8	0
a cross-slot bivector, $V_E^s \wedge V_E^o$	G	n^2	24	8	16

One entity space is not enough. A bivector there holds only the antisymmetric half: the swap merely flips the sign, and a reflexive scores $\mathbf{x}^\top A \mathbf{x} = 0$, so *John loves himself* would vanish. FGA gives each argument an isometric copy of the entity space, a subject slot V_E^s and an object slot V_E^o , entered through injections ι_s and ι_o :

$$L = \sum_{i,k} G_{ik} \mathbf{e}_i^s \wedge \mathbf{e}_k^o, \quad \iota_s(\mathbf{x}) \rfloor (L \rfloor \iota_o(\mathbf{y})) = \mathbf{x}^\top G \mathbf{y}$$

The quadratic form sees only S ; one space holds only A ; two slots hold both. The price is n^2 numbers, as for a tensor; the paper notes that the $\binom{n}{2}$ saving came from keeping only the alternating sector. The algebra adds the multiplication law and the typing.

Binding With the Wedge, Unbinding Exactly

Roles get a subspace of their own, V_R , and the paper's decomposition makes it orthogonal to the entities, $V_R \perp V_E$. A worked case: roles $\mathbf{r}_{\text{ag}} = \mathbf{e}_1$ and $\mathbf{r}_{\text{th}} = \mathbf{e}_2$; entities $\text{john} = \mathbf{e}_3$ and $\text{mary} = 0.6 \mathbf{e}_3 + 0.8 \mathbf{e}_4$, with $\cos(\text{john}, \text{mary}) = 0.6$.

$$\mathbf{E} = \mathbf{r}_{\text{ag}} \wedge \text{john} + \mathbf{r}_{\text{th}} \wedge \text{mary} = \mathbf{e}_1 \mathbf{e}_3 + 0.6 \mathbf{e}_2 \mathbf{e}_3 + 0.8 \mathbf{e}_2 \mathbf{e}_4$$

Since $V_R \perp V_E$, $\mathbf{r} \mathbf{f} = \mathbf{r} \wedge \mathbf{f}$: a role times a filler is pure wedge. To unbind, contract with the inverse role, $\mathbf{r}^{-1} = \mathbf{r}/Q(\mathbf{r})$:

$$\mathbf{r}^{-1} \rfloor (\mathbf{r} \wedge \mathbf{f}) = \mathbf{f} - \frac{\mathbf{r} \cdot \mathbf{f}}{Q(\mathbf{r})} \mathbf{r} = \mathbf{f} \quad \text{exactly when } \mathbf{r} \perp \mathbf{f}$$

query	John loves Mary, $\mathbf{E}$	Mary loves John, $\mathbf{E}'$
$\mathbf{r}_{\text{ag}}^{-1} \rfloor (\cdot)$	$\mathbf{e}_3 = \text{john}$, exactly	$0.6 \mathbf{e}_3 + 0.8 \mathbf{e}_4 = \text{mary}$, exactly
$\mathbf{r}_{\text{th}}^{-1} \rfloor (\cdot)$	mary , exactly	john , exactly

$\cos(\mathbf{E}, \mathbf{E}') = 0.6$, which is the similarity of the two fillers and nothing more. **No cleanup step**: the MAP record of Session 12 returned its fillers at cosine 0.50 and needed one. The paper states this as a proposition: orthonormal roles and $V_R \perp V_E$ suffice, whatever the fillers.

What Changes From the VSA

	MAP, from Session 12	FGA
the metric	assumed: random atoms, $G \approx I$	declared: a signature
role versus filler	the role inventory; $\odot$ commutes	the subspace: $\mathbf{r} \in V_R$, $\mathbf{f} \in V_E$
roles	random atoms in the shared space	an orthonormal set in V_R
unbinding	noisy; cleanup required	exact, by contraction with $\mathbf{r}^{-1}$
how many roles fit	many, up to a capacity set by d	at most $\dim V_R$
types	outside the vectors	visible as grade and subspace

Roles are still declared in FGA, as they were in the VSA. What the algebra adds is exact recovery and a type for every object; what it costs is a dedicated dimension for each role, where a VSA packs many quasi-orthogonal roles into one shared space. The paper states the trade as exact recovery for a bounded number of bindings, against approximate recovery for an unbounded number.

Each side pays a real cost. The paper argues that argument structure, whose arity is small and bounded, suits the exact regime; how far that holds beyond argument structure is an empirical question.

Coercion as a Rotor: A Preview

Mary began the book. Begin selects an event, and book denotes an object; Generative Lexicon resolves the clash through the noun's qualia – reading (telic) or writing (agentive). FGA models this as **Inner Application**: a qualia rotor turns the argument, and the predicate then contracts it,

$$P(arg) \longmapsto (R_q arg \tilde{R}_q) \rfloor P.$$

A toy version. Take a *formal* direction $\mathbf{f} \in V_E$ and a *telic* direction $\mathbf{t} \in V_P$, orthonormal and both declared positive. Let book = $\mathbf{f}$, and let begin look for $\mathbf{t}$:

the rotor, as two mirrors	$R = \frac{1}{\sqrt{2}}(\mathbf{f} + \mathbf{t}) \mathbf{f} = \frac{1}{\sqrt{2}}(1 - \mathbf{f}\mathbf{t})$
outer application, no rotor	$\text{begin} \cdot \text{book} = \mathbf{t} \cdot \mathbf{f} = 0$ (a type clash)
inner application	$R \text{book} \tilde{R} = \mathbf{t}$, then $\text{begin} \cdot \mathbf{t} = 1$

This is Part IV's 90° rotation, axes renamed. In the position paper, outer application is the case where R_q is the identity, and inner application the general case.

The declared signature limits coercion. A rotor preserves Q , and the paper notes that no real rotor maps a positive-square vector entirely into a negative-definite subspace: with $\mathbf{t}$ declared negative, the plane holds a boost, $R\mathbf{f}\tilde{R} = \cosh \varphi \mathbf{f} + \sinh \varphi \mathbf{t}$, which adds a telic part but keeps the formal one.

Where Each Ingredient Comes From

declared	computed by the algebra	added by FGA
the quadratic form Q	S , by polarization: the symmetric half	a signature chosen for meaning
–	$\mathbf{u} \wedge \mathbf{v}$: the antisymmetric half, metric-free	subspaces for roles and slots
an orthogonal basis	$\text{Cl}(p, q, r)$: the whole table	grade and subspace as type
–	inverses $\mathbf{a}/Q(\mathbf{a})$; mirrors; rotors	wedge binding; exact unbinding
–	projectors, cycles, rates	polar pairs; coercion by rotor

Read the first column: [the only thing declared is a symmetric form, which is to say a quadratic form.](#) Everything in the second column follows from $\mathbf{v}^2 = Q(\mathbf{v})$. The third column is what one semantic theory builds on top.

Session 10 asked what the matrix in the middle of $\mathbf{x}^\top G \mathbf{y}$ should be, and Session 11 split it in two. [The Clifford algebra is the answer that declares one half and derives the other.](#)

Summary

- 1 A **quadratic form** is a bilinear form with the same vector in both slots, $Q(\mathbf{x}) = \mathbf{x}^\top G \mathbf{x}$. Since $\mathbf{x}^\top A \mathbf{x} = 0$, it sees only the symmetric half.
- 2 **Polarization**, $S(\mathbf{x}, \mathbf{y}) = \frac{1}{2}[Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})]$, recovers that half exactly. A quadratic form and a symmetric bilinear form are the same information.
- 3 Quadratic forms are everywhere in NLP: Mahalanobis distance for out-of-domain detection, the Laplacian energy of label propagation, the Lorentz form of hyperbolic embeddings.
- 4 By Sylvester, a quadratic form is classified by its **signature** (p, q, r) . The attention head of Session 11 has signature $(1, 0, 1)$, and its two tokens differ by a null vector.
- 5 The **Clifford algebra** is defined by one rule, $\mathbf{v}^2 = Q(\mathbf{v})$. Polarization forces the symmetric half of the product; the antisymmetric half, the wedge, comes free: $\mathbf{u}\mathbf{v} = S(\mathbf{u}, \mathbf{v}) + \mathbf{u} \wedge \mathbf{v}$.
- 6 The signature decides what the algebra can do: projectors, cycles, rates; rotations, boosts, translations; which vectors can be mirrors.
- 7 **FGA** declares a quadratic form for meaning, reads type from grade and subspace, gives each argument of a relation its own slot, binds roles with the wedge, and unbinds exactly by contraction.

Coming up: the multiplication table of $\text{Cl}(3, 0, 0)$, and rotors at work.

Exercises

- ① Prove that $\mathbf{x}^\top A \mathbf{x} = 0$ for every antisymmetric A , and conclude that no function of Q can recover the antisymmetric half of a form.
- ② Show that polarizing $Q(\mathbf{x}) = \mathbf{x}^\top S \mathbf{x}$ returns $S(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top S \mathbf{y}$ for every symmetric S .
- ③ Find the signatures of the Laplacians of a triangle (three words, all linked) and of two separate edges (four words). Relate the number of zeros to the number of connected components.
- ④ In $\text{Cl}(3, 0, 0)$, show that $(\mathbf{e}_1 \mathbf{e}_2)^2 = (\mathbf{e}_2 \mathbf{e}_3)^2 = (\mathbf{e}_3 \mathbf{e}_1)^2 = -1$. Compute $(\mathbf{e}_2 \mathbf{e}_3)(\mathbf{e}_3 \mathbf{e}_1)$, and find signs for i, j, k that make these bivectors obey $ij = k$.
- ⑤ Use dual numbers to compute the derivative of $f(x) = x^3 - 2x$ at $x = 2$. Check it by hand.
- ⑥ ★ You are designing a semantic space that must hold a polar pair (alive/dead) and a cyclic domain (the seasons). Which sign would you declare for each axis, and why? Use the one-generator algebras of Part IV.

Reading

- W. K. Clifford, *Applications of Grassmann's extensive algebra*, American Journal of Mathematics 1, 1878 – the original construction.
- D. Hestenes and G. Sobczyk, *Clifford Algebra to Geometric Calculus*, Reidel, 1984.
- L. Dorst, D. Fontijne and S. Mann, *Geometric Algebra for Computer Science*, Morgan Kaufmann, 2007.
- P. Lounesto, *Clifford Algebras and Spinors*, 2nd ed., Cambridge University Press, 2001.
- A. Podolskiy et al., *Revisiting Mahalanobis distance for transformer-based out-of-domain detection*, AAAI 2021.
- X. Zhu, Z. Ghahramani and J. Lafferty, *Semi-supervised learning using Gaussian fields and harmonic functions*, ICML 2003.
- D. Rao and D. Ravichandran, *Semi-supervised polarity lexicon induction*, EACL 2009.
- M. Nickel and D. Kiela, *Learning continuous hierarchies in the Lorentz model of hyperbolic geometry*, ICML 2018.
- J. Brehmer, P. de Haan, S. Behrends and T. Cohen, *Geometric Algebra Transformer*, NeurIPS 2023.
- J. Pustejovsky, *Toward a Functional Geometric Algebra for natural language semantics*, arXiv:2604.25902v2, 2026.