Groups, Fields, and Vector Spaces, part 2

Fields

Field axioms

A field is a set of elements $\alpha, \beta, ...$ along with two operations, + and $\cdot$.

For the operation +, the elements form a commutative group. The identity is denoted by 0. The inverse of α is denoted $-\alpha$.

For the operation $\cdot$ (typically denoted by juxtaposition), the elements other than 0 form a commutative group, and the identity is denoted by 1. The inverse of *a* is denoted $1/\alpha$ or α^{-1}.

The operations + and $\cdot$ are linked by the distributive law, $\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma$.

Fields can be finite or infinite.

Field examples

The real numbers $\Re$ and the complex numbers $\mathbb{C}$ are the familiar ones, and the ones we typically use to represent scalar quantities

$\mathbb{C}$ has a very non-obvious property that $\Re$ does not have: in $\mathbb{C}$, every polynomial equation has roots. ("$\mathbb{C}$ is an algebraically closed field.") Not the case in $\Re$: for example, $x^2+1=0$ has no roots. On the other hand, $\Re$ is an "ordered field" ("order" here meaning size rank, not the group-theoretic meaning of "order"): if $\alpha \neq \beta$, then either $\alpha < \beta$ or $\beta < \alpha$, but not both.

The integers $\mathbb{Z}$ do not form a field, since there are no inverses.

There are many other fields, including the rational numbers ($\mathbb{Q}$) and finite fields.

Finite fields

For each size that is a power of a prime number, p^n, there is exactly *one* finite field, known as a Galois field. Finite fields are important for experimental design.

Simplest case is $n=1$, i.e., a field of prime size. This is denoted $\mathbb{Z}_p$ or $GF(p,1)$, and consists of the integers $\{0,1,...,p-1\}$. Addition and multiplication are "mod *p*". That is, carry out addition and multiplication in the ordinary fashion, and then find the remainder after dividing by *p*. We'll look at $n > 1$ later; that construction generalizes the relationship of the complex numbers to the real numbers.

To see that $\mathbb{Z}_p$ (the $n = 1$-case) is in fact a field:

The additive group is the cyclic group, generated by 1. But are there multiplicative inverses? I.e., given an α in $\{0,1,...,p-1\}$, are we guaranteed to find a β such that $\alpha\beta = 1$ (mod *p*)?

Two very different ways to see that multiplicative inverses exist.

Method 1: use the fact that automorphisms form a group, and apply this to $\mathbb{Z}_p$. Viewed abstractly, the additive group of $\mathbb{Z}_p$ is a cyclic group. We are guaranteed that the order of every nonzero element of $\mathbb{Z}_p$ (under addition) is *p*, since the order of every element must be a factor of $\#(\mathbb{Z}_p) = p$, and *p* is prime. We also know that the map $\varphi_\alpha(x) = \alpha x$ is a homomorphism of the additive group of $\mathbb{Z}_p$. This follows from the distributive law: $\varphi_\alpha(x+y) = \alpha(x+y) = \alpha x + \alpha y = \varphi_\alpha(x) + \varphi_\alpha(y)$. Now choose *x* to be any element that is not the identity. If $\alpha < p$, then αx cannot be the identity, since the order of *x* is *p*. (Note that αx is NOT the group operation for the additive group; it means $x + x + \ldots + x$ a total of α times.) Since αx is not the identity, its order must be *p* (since the only possible orders are factors of *p*, and *p* is prime). Therefore, successive applications of + to αx will produce all of the members of the group $\mathbb{Z}_p$. Therefore $\varphi_\alpha(x) = \alpha x$ is an isomorphism, not just a homomorphism. Since isomorphisms form a group, φ_α must have an inverse. Call it ψ. $\psi(x)$ must be something in $\mathbb{Z}_p$, so let's call $\psi(x) = \beta x$. Since ψ preserves structure, $\psi(2x) = \psi(x+x) = \psi(x) + \psi(x) = \beta x + \beta x = \beta(2x)$, and similarly for 3*x*, … so $\psi(y) = \beta y$ for any *y*. Finally, since ψ and φ_α are inverses, $x = \psi(\varphi_\alpha(x)) = \psi(\alpha x) = \beta\alpha x$, so $x = \beta\alpha x$. That is, *x* added to itself $\beta\alpha$ times is *x*, i.e., $\beta\alpha - 1$ is a multiple of *p*. And β is a multiplicative inverse for α (mod *p*).

Method 2: use the Euclidean algorithm to construct an inverse.

$\alpha\beta = 1$ (mod *p*) is equivalent to $\alpha\beta + pq = 1$, for some integer *q*.

In general: for a given *A* and *P*, $AB + PQ = 1$ has a solution in integers when, and only when, *A* and *P* are relatively prime. (This is the classic Euclidean algorithm). If *p* is prime, α and *p* are guaranteed to be relatively prime, and consequently, $\alpha\beta + pq = 1$ has a solution. The solution is obtained by "descent", using the Euclidean algorithm: The Euclidean algorithm is easiest to explain by example. Say we want to find the multiplicative inverse of 18 in $\mathbb{Z}_{79}$. That is, we want to find integers β and *q* that solve $\alpha\beta + pq = 1$ for $p = 79$ and $\alpha = 18$. To solve $18\beta + 79q = 1$ in integers:

Step 1: Note that $79 = 4\cdot 18 + 7$. So,
$18\beta + 79q = 1$ is equivalent to $18\beta + (4\cdot 18 + 7)q = 1$, or, $18(\beta + 4q) + 7q = 1$.
So if $18\beta' + 7q' = 1$, we can solve $18\beta + 79q = 1$ with $\beta = \beta' - 4q$ and $q = q'$.

Step 2: Note that $18 = 2\cdot 7 + 4$. So,
$18\beta' + 7q' = 1$ is equivalent to $(2\cdot 7 + 4)\beta' + 7q' = 1$, or, $4\beta' + 7(2\beta' + q') = 1$.
So if $4\beta'' + 7q'' = 1$, we can solve $18\beta' + 7q' = 1$ with $\beta' = \beta''$ and $q' = q'' - 2\beta'$.

Step 3: Note that $7 = 1\cdot 4 + 3$. So,
$4\beta'' + 7q'' = 1$ is equivalent to $4\beta'' + (1\cdot 4 + 3)q'' = 1$, or, $4(\beta'' + q'') + 3q'' = 1$.
So if $4\beta''' + 3q''' = 1$, we can solve $4(\beta'' + q'') + 3q'' = 1$ with $\beta'' = \beta''' - q''$ and $q'' = q'''$.

We are guaranteed to have smaller and smaller coefficients, since at each stage we reduce one coefficient to its remainder when divided by the other.
An integer solution of $4\beta''' + 3q''' = 1$ is "obvious" – if not, we could go one more stage.
$\beta''' = 1$, $q''' = -1$; working backwards yields
$\beta'' = 2$, $q'' = -1$; then
$\beta' = 2$, $q' = -5$; then
$\beta = 22$, $q = -5$. So, $18\cdot 22 - 79\cdot 5 = 1$, i.e., $18\cdot 22 = 1$ (mod 79), i.e., 22 is the inverse of 18 in $\mathbb{Z}_{79}$.

Relationships between fields

Familiar example: the real numbers $\Re$ and the complex numbers $\mathbb{C}$
We write complex numbers as $z = x + yi$, where $i^2 = -1$ and *x* and *y* are reals.

If all we are told is that the α's and β's are drawn from a field *k*, and that θ is a symbol that can be added and multiplied by the α's and β's, in a manner that follows the distributive law, we would know how to add quantities like $\alpha_0 + \alpha_1\theta$. For example,
$(\alpha_0 + \alpha_1\theta) + (\beta_0 + \beta_1\theta) = (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)\theta$. We could also try to multiply them, but we would find:

$$\begin{aligned}(\alpha_0 + \alpha_1\theta)\cdot(\beta_0 + \beta_1\theta) &= \alpha_0\cdot(\beta_0 + \beta_1\theta) + \alpha_1\theta\cdot(\beta_0 + \beta_1\theta) \\ &= \alpha_0\beta_0 + (\alpha_0\beta_1 + \alpha_1\beta_0)\theta + \alpha_1\beta_1\theta^2\end{aligned}$$

which is a "problem", since there is a θ^2-term. So, to ensure that the result of the multiplication is of the same form ($\gamma_0 + \gamma_1\theta$), we need to have a way to write θ^2 in terms of θ and field elements, i.e., an equation of the form $\theta^2 + c_1\theta + c_0 = 0$. We want to choose c_1 and c_0 so that there is no solution of $x^2 + c_1x + c_0 = 0$ in *k*, since if there *was* a solution, i.e, $x^2 + c_1x + c_0 = 0$ for *x* in *k*, then we'd have two ways of writing the same thing (*x* and θ). So this process of "extension" only works if we choose a polynomial that does not have a root in the starting field.

Extending from the real numbers $\Re$ and the complex numbers $\mathbb{C}$ chooses the simplest such polynomial: $\theta^2 = -1$.

$$\begin{aligned}&(\alpha_0 + \alpha_1\theta)\cdot(\beta_0 + \beta_1\theta) = \alpha_0\cdot(\beta_0 + \beta_1\theta) + \alpha_1\theta\cdot(\beta_0 + \beta_1\theta)\\ &= \alpha_0\beta_0 + (\alpha_0\beta_1 + \alpha_1\beta_0)\theta + \alpha_1\beta_1\theta^2 = \alpha_0\beta_0 - \alpha_1\beta_1 + (\alpha_0\beta_1 + \alpha_1\beta_0)\theta\end{aligned}$$

Choosing other polynomials $x^n + c_{n-1}x^{n-1} + \ldots c_1x^1 + c_0 = 0$ (x in $\Re$) does not yield anything useful. And attempting to extend $\mathbb{C}$ by this strategy does not yield anything useful – since $\mathbb{C}$ is "algebraically closed." (We can get some nontrivial fields by choosing x in $\mathbb{Q}$ and $n \ge 2$. These are very interesting to number theorists, not so useful for us.)

(Digression: But if we choose polynomials such as $x^n + c_{n-1}x^{n-1} + \ldots c_1x^1 + c_0 = 0$ with x in $\mathbb{Z}_p$, we will get the Galois fields of size p^n. For p=2, this is the algebraic structure underlying "m-sequences", a way of producing pseudorandom cyclic binary sequences that have some very nice properties for experimental design. The main property of an m-sequence is that a cyclic shift of the sequence is very nearly orthogonal to the original sequence is a consequence of the field axioms.)

The extension process guarantees that the extension field has a nontrivial automorphism; in the familiar case of extending from $\Re$ to $\mathbb{C}$, this automorphism is complex conjugation. To see this: Equations of the form $\theta^2 + c_1\theta + c_0 = 0$ are expected to have two roots, so which one do we choose when we use $\theta^2 = -1$ to extend from $\Re$ to $\mathbb{C}$? Obviously it can't matter, since reducing θ^2 did not depend on this choice. But we know that the two roots, though members of $\mathbb{C}$, are related by elements in $\Re$. This is because $x^2 + c_1x + c_0 = (-c_1 - x)^2 + c_1(-c_1 - x) + c_0$. So if θ solves $x^2 + c_1x + c_0 = 0$, so does $\theta^* = -c_1 - \theta$.

So we can replace θ by $\theta^* = -c_1 - \theta$ and leave the rules of the extension field unchanged. I.e., we have found an automorphism of the extension field, *conj*, which maps $x + \theta y$ into $conj(x + \theta y) = (x + \theta^* y)$.

Note also that $(x + \theta y)(x + \theta^* y)$ is something that is unchanged by *conj*. Therefore, when it is written in the form $a + b\theta$, it cannot involve θ (i.e., $b = 0$). So the mapping from it is a mapping from $x + \theta y$ into $(x + \theta y)(x + \theta^* y)$ is a mapping from the extension field back to the base field that preserves multiplication.

The mapping *conj* lets us see that the extension field has multiplicative inverses: $\dfrac{1}{(x + \theta y)} = \dfrac{(x + \theta^* y)}{(x + \theta y)(x + \theta^* y)} = \dfrac{x}{D} + \dfrac{y}{D}\theta^*$, for $D = (x + \theta y)(x + \theta^* y)$ which we know is in the base field.

In the familiar case of $\Re$ and $\mathbb{C}$, *i* solves $x^2 = -1$ (so, $c_1 = 0$ and $i^* = -i$), so $conj$ is the familiar $conj(x+iy) = (x-iy)$, complex conjugation, more typically written z^*. For $z = x+iy$, we write $|z|^2 = zz^* = x^2 + y^2$.

In the general case ($x^n + c_{n-1}x^{n-1} + \ldots c_1 x^1 + c_0 = 0$) there are *n* roots that are abstractly indistinguishable in the base field, and therefore more automorphisms of the extension field. Given an element of the extension field $z = a_0 + a_1\theta + \ldots + a_{n-1}\theta^{n-1}$, the product of the elements obtained by replacing θ by all other roots of $x^n + c_{n-1}x^{n-1} + \ldots c_1 x^1 + c_0 = 0$ is invariant under these automorphisms, and thus must be in the base field, analogous to $(x+\theta y)(x+\theta^* y)$.

Vector Spaces

Vector space axioms

Vector spaces have two kinds of elements: vectors (*v*, *w*, …) drawn from a set *V* and scalars (α, β, …) drawn from a set *k*.

The scalars form a field *k*, operations are scalar addition, +, and multiplication, $\cdot$.

The vectors form a commutative group under addition, operation is vector addition, +. The additive inverse of v is $-v$.

There is an operation "scalar multiplication" that maps a scalar α and a vector *v* into a vector αv. It satisfies two kinds of distributive laws,

$\alpha(v+w) = \alpha v + \alpha w$ and
$(\alpha+\beta)v = \alpha v + \beta v$. As a consequence of the latter, $0v$ must be the identity for vector addition, since $\alpha v = (\alpha + 0)v = \alpha v + 0v$ (and the identity is unique). And the additive inverse of αv, $-(\alpha v)$, is the same as $(-\alpha)v$, since inverses are unique, and $0v = (\alpha + (-\alpha))v = \alpha v + (-\alpha)v$.

and one kind of associative law that relates scalar and field multiplication:
$(\alpha\beta)v = \alpha(\beta v)$. (The multiplication $\alpha\beta$ is in the field.) As a consequence of this, $1v = v$.
(Calculate $1(1v - v) = 1(1v) - 1v = (1\cdot 1)v - 1v = 1v - 1v = 0$.)

Unless stated otherwise, we work with the field $k = \mathbb{C}$.

Nothing is said about length, angles, or dimension.

Vector space examples

A field can always be considered as a vector space over itself, with the vector operations identical to the field operations.

Ordered *n*-tuples of field elements form a vector space. To define the operations: say $v=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ and $w=(\beta_1,\beta_2,\ldots,\beta_n)$. Operations work component-by-component.
Vector addition: $v+w=(\alpha_1+\beta_1,\alpha_2+\beta_2,\ldots,\alpha_n+\beta_n)$.
Scalar multiplication: $\lambda v=(\lambda\alpha_1,\lambda\alpha_2,\ldots,\lambda\alpha_n)$.

An extension field is a vector space over the base field. The vector operations are identical to the field operations. (But it has a lot more structure too.) One can think of $z=a_0+a_1\theta+\ldots+a_{n-1}\theta^{n-1}$ as an ordered *n*-tuple $(a_0,a_1,...,a_{n-1})$, with the powers of θ merely tagging the places.

The set of functions from any set *S* to the field *k* forms a vector space. This generalizes the ordered *n*-tuple example; the subscripts were merely placeholders, i.e., $S=\{1,2,...,n\}$. To be explicit: Say *f* and g are functions on *S*. To define $f+g$, we need to define how it acts on elements of *S*: $(f+g)(s)=f(s)+g(s)$, where the addition is in *k*. To define αf, we need to define its action: $(\alpha f)(s)=\alpha\cdot(f(s))$ where the "outer" multiplication on the right is in the field *k*. This construction is known as the "free vector space" on *S*.

Linear independence, span, and basis sets

Definition of linear independence: A set of vectors $\{v_1,\ldots,v_h\}$ is linearly independent if $\sum_{k=1}^{h}\alpha_k v_k=0$ implies that each $\alpha_k=0$.

Definition of linear span (or just span): The span of a set of vectors $\{v_1,\ldots,v_h\}$ is the set of all vectors *v* that can be written as a linear combination $v=\sum_{k=1}^{r}\alpha_k v_k$ of members of the set. Note that the span of a set of vectors is always a vector space.

Definition of a basis: A set of vectors $\{v_1,\ldots,v_h\}$ is a basis for a vector space *V* if (i) it is linearly independent, and (ii) its span is the entire vector space, i.e, any vector *v* in *V* can be written as $v=\sum_{k=1}^{h}\alpha_k v_k$. The α_k are the "coordinates" for *v*, with respect to the basis set $\{v_1,\ldots,v_h\}$.

Once we have chosen the basis set, the coordinates are unique. That is, if

$v=\sum_{k=1}^{h}\alpha_k v_k$ and also $v=\sum_{k=1}^{h}\beta_k v_k$, then $\sum_{k=1}^{h}\alpha_k v_k-\sum_{k=1}^{h}\beta_k v_k=0$, so $\sum_{k=1}^{h}(\alpha_k-\beta_k)v_k=0$, and each $\alpha_k-\beta_k=0$ (since the basis set is linearly independent).

As an example: For the free vector space on *S*, a basis set consists of the vectors δ_s, one for each element *s* of *S* , defined as follows: $\delta_s(s)=1$, and $\delta_s(t)=0$ if $t\neq s$. To see that it is a basis, note that $f=\sum_{s\in S}f(s)\delta_s$.

Note that if a set of vectors $\{v_1,\ldots,v_r\}$ is linearly independent, then it is always a basis set for something, namely, its span.

Slightly less obviously: if a set of vectors $\{v_1,\ldots,v_r\}$ spans a vector space *V*, then it always has a subset that is a basis for *V*. To see this: The original set could fail to be a basis if its members are not linearly independent, i.e., that there is some set of field elements β_k for which $\sum_{k=1}^{r}\beta_k v_k=0$, with at least one $\beta_j\neq 0$. This means that $v_j=-\sum_{k=1,k\neq j}^{r}\frac{\beta_k}{\beta_j}v_k$. This in turn allows us to eliminate v_j from the set, and still be able to represent any vector:

$$v=\sum_{k=1}^{r}\alpha_k v_k=\sum_{k=1,k\neq j}^{r}\alpha_k v_k+\alpha_j v_j=\sum_{k=1,k\neq j}^{r}\alpha_k v_k-\alpha_j\sum_{k=1,k\neq j}^{r}\frac{\beta_k}{\beta_j}v_k=\sum_{k=1,k\neq j}^{r}\left(\alpha_k-\alpha_j\frac{\beta_k}{\beta_j}\right)v_k$$

We then continue eliminating until we can no longer find a relationship of linear dependence. The resulting set is the required basis. Note that the field properties (existence of a multiplicative inverse) play an important role.

Dimension

If the size of a basis set is finite, then this size is an intrinsic characteristic of the vector space, namely, its dimension.

We need to see that any two basis sets for a vector space have the same size (if the set size is finite). Suppose to the contrary, and that we've found the smallest such example. To be specific, say that $S_v=\{v_1,\ldots,v_h\}$ is a basis for *V*, and so is $S_w=\left\{w_1,\ldots,w_r\right\}$, with $h>r$. We need to show that this situation does not allow the elements of S_w to be linearly independent (but we don't want to resort to coordinates, counting degrees of freedom, etc.)

The proof is surprisingly tricky, and the reason is that the finiteness of the set size is critical.

We begin by adjoining a vector from S_v to S_w. Since $S_w=\left\{w_1,\ldots,w_r\right\}$ is a basis, we can write v_1 as a linear combination of S_w , $v_1=\sum_{k=1}^{r}\beta_k w_k$ At least one of the coefficients, say β_j, must be nonzero (since otherwise we would have found a linear relationship among the elements of S_w).

So we can use this to eliminate w_j from the adjoined set. This results in a new basis set, containing v_1 and all the w_k except w_j. By our construction, the new set still spans *V*, and it is linearly independent. The latter follows because if there were a linear dependence, we could eliminate yet another vector, resulting in a smaller example. (We assumed we were dealing with the smallest example.)

We can now continue the swapping, each time bringing in another element of S_v. We can always eliminate a "*w*" from the augmented S_w, since, if there were a linear dependence among just the *v*'s, then they could not have been linearly independent. After *r* steps, we've replaced all of the *w*'s in S_w by a v, but there are still $h-r$ *v*'s in S_v. So now there is a contradiction: at each stage, we showed that S_w is a basis, so now there must be a way to write one of these remaining *v*'s as a linear combination of the ones we swapped into S_w.

The definitions of linear independence and of a basis set make sense for infinite sets $\{v_1,\ldots,v_h,\ldots\}$, i.e., for infinite-dimensional vector spaces. But one cannot claim that basis sets have a definite "size".

Combining vector spaces

General set-up here: *V* and *W* are vector spaces over the same field *k*.
$\{v_1,\ldots,v_m\}$ is a basis for *V*, and $\{w_1,\ldots,w_n\}$ is a basis for *W*.

Purposes: (a) a review of linear algebra, (b) setting up the material we need to describe how groups transform vectors (data), (c) a coordinate-free definition of the determinant.

Direct sum

The direct sum of *V* and *W*, $V\oplus W$, is a vector space consisting of ordered pairs of elements from V and *W*, i.e., (v,w).

Vector-space operations are defined as component by component: vector addition, $(v,w)+(v',w')=(v+v',w+w')$, and scalar multiplication, $\lambda(v,w)=(\lambda v,\lambda w)$.
A basis for $V\oplus W$ is the $(m+n)$-element set, $\{(v_1,0),\ldots,(v_m,0),(0,w_1),\ldots,(0,w_n)\}$.

Homomorphism

$Hom(V,W)$ indicates the set of homomorphisms (structure-preserving maps) from *V* to *W*.

In this context, "φ preserves structure" is equivalent to "φ is a linear mapping", i.e.,
$\varphi(\alpha_1 v_1+\alpha_2 v_2)=\alpha_1\varphi(v_1)+\alpha_2\varphi(v_2)$,
and also to "φ obeys superposition."

The set of homomorphisms between two vector spaces is, itself, a vector space over *k*. To see this, we need to define the vector space operations among the homomorphisms. That is, if φ and ψ are both homomorphisms from *V* to *W* and λ is a scalar, we need to define vector addition, $\varphi+\psi$, and scalar multiplication, $\lambda\varphi$. Since both of these need to be in $Hom(V,W)$, we define them by their actions on V.

Addition: $\varphi+\psi$ is defined by $(\varphi+\psi)(v)=\varphi(v)+\psi(v)$.
(Right hand side is addition in *W*).

Scalar multiplication: $\lambda\varphi$ is defined by $(\lambda\varphi)(v)=\lambda(\varphi(v))$. (Right hand side is scalar multiplication in *W*).

The definitions are almost automatic, but it is worthwhile seeing how it is guaranteed that (a) they preserve the structure of *V* (and therefore are members of $Hom(V,W)$), and (b) that the vector space axioms are obeyed by the above definitions.

Given a basis $\{v_1,\ldots,v_m\}$ for *V*, and $\{w_1,\ldots,w_n\}$ for *W*, we can build a basis for $Hom(V,W)$:

Consider the mapping φ_{ij} from *V* to *W* defined by $\varphi_{ij}(\sum_{k=1}^{m}\alpha_k v_k)=\alpha_i w_j$. In other words, let φ_{ij} (a) map the *i*th element of $\{v_1,\ldots,v_m\}$ to the *j*th element of $\{w_1,\ldots,w_n\}$, (b) map every other basis vector to zero, and (c) extend to the rest of *V* as required by linearity.

The set $S_{VW}=\{\varphi_{11},\ldots,\varphi_{1n},\varphi_{21},\ldots,\varphi_{2n},\ldots,\varphi_{m1},\ldots,\varphi_{mn}\}$ form a basis set for $Hom(V,W)$. (Also makes sense if either *V* or *W* is infinite-dimensional.)

To express an arbitrary φ in terms of S_{VW}, we note that we only have to express how φ acts on each basis element of V (since the fact that φ is linear allows us to extend the action of φ from any basis of IV to the whole space). So we simply express the action of φ on a basis element v_i. in terms of the basis $\{w_1,\ldots,w_n\}$ of *W*: $\varphi(v_i)=\sum_{j=1}^{n}\gamma_{ji}w_j$. Then, it follows from the definition of φ_{ij} that $\varphi=\sum_{i=1}^{m}\sum_{j=1}^{n}\gamma_{ji}\varphi_{ij}$.

To see that S_{VW} has no linear dependencies, we suppose that we had some linear combination of its elements that is zero: $\psi=\sum_{i=1}^{m}\sum_{j=1}^{n}c_{ji}\varphi_{ij}=0$. We need to show that this forces each of the c_{ji} to be 0. (The order of the subscripts of c_{ji} is reversed, because of matrix conventions, see below.) If $\psi=0$, then its action on every element of V must yield 0. Specifically, its action on any basis element v_k of V must be 0. $0=\psi(v_k)=\sum_{i=1}^{m}\sum_{j=1}^{n}c_{ji}\varphi_{ij}(v_k)=\sum_{j=1}^{n}c_{jk}\varphi_{kj}(v_k)=\sum_{j=1}^{n}c_{jk}w_j$, where the last two equalities follow from the definition of φ_{ij} . Since $\{w_1,\ldots,w_n\}$ is a basis set for W, $\sum_{j=1}^{n}c_{jk}w_j=0$ can only happen if all $c_{jk}=0$.

Coordinates

This should begin to look a lot like matrices. The connection is explicit if we choose specific basis sets $\{v_1,\ldots,v_m\}$ and $\{w_1,\ldots,w_n\}$.

We think of V as a set of m numbers in a column, and choose as a basis for V the following:

$v_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},v_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\cdots,v_n=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$. So a vector $v=\sum_{k=1}^{m}\alpha_k v_k$ corresponds to $v=\begin{pmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_m\end{pmatrix}$

Similarly, think of W as a set of n numbers in a column, and choose $w_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},w_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\cdots,w_m=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$. So a vector $w=\sum_{j=1}^{n}\beta_j w_j$ corresponds to $w=\begin{pmatrix}\beta_1\\\beta_2\\\vdots\\\beta_n\end{pmatrix}$.

Transformations from V to W can now be thought of as arrays of n rows, m columns:

$\varphi=\begin{pmatrix}\gamma_{11}&\cdots&\gamma_{1m}\\\vdots&\ddots&\vdots\\\gamma_{n1}&\cdots&\gamma_{nm}\end{pmatrix}$. A basis element φ_{ij}, which maps v_i to w_j, corresponds to a matrix in which the element in the ith column and jth row (γ_{ji}) is equal to 1, and all other elements are 0. With these correspondences, $\varphi v=w$ is equivalent to

$\begin{pmatrix}\gamma_{11}&\cdots&\gamma_{1m}\\\vdots&\ddots&\vdots\\\gamma_{n1}&\cdots&\gamma_{nm}\end{pmatrix}\begin{pmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_m\end{pmatrix}=\begin{pmatrix}\beta_1\\\beta_2\\\vdots\\\beta_n\end{pmatrix}$, with the usual rules of matrix multiplication: $\beta_j=\sum_{i=1}^{m}\gamma_{ji}\alpha_i$.

Note that the above would hold no matter what basis we choose – and there is no reason at this point to choose this particular basis. In another basis, the coordinates α_i, β_j, and γ_{ij} that represent particular vectors and transformations would be different.

If we fix a particular vector *v*, its representation in coordinates is *completely arbitrary* – i.e., by changing the basis, we could have used *any* set of numbers to represent it. Put another way, if we are given the coordinates of a vector, we know nothing about its intrinsic properties, since the same set of numbers could represent any vector.

To what extent is this also true of the numerical representation of a member of $Hom(V,W)$? It can't be completely true, since the dimension of the range of φ is an intrinsic property, so it must constrain the numerical representation. And there are other constraints that arise, when we consider $Hom(V,V)$.

The dual space

An important special case of $Hom(V,W)$ is that of $W = k$, the base field. $Hom(V,k)$ is thus the set of linear mappings from V to the scalars, and is also known as the dual space of *V*, V^*.

If *V* has some finite dimension *m*, then the dimension of V^* is also *m* (since the dimension of *k* is 1).

Importantly, and perhaps non-obviously, there is no natural relationship between *V* and its dual. This seems surprising because in some sense, V^* and *V* have the same intrinsic structure – they are abstractly the same as a free vector space on a set of size *m*. The problem appears when we try to set up a correspondence between *V* and V^*. The obvious way to proceed is to take a vector in *V*, determine its coordinates with respect to some basis set $\{v_1,\ldots,v_m\}$, and then find the element in of V^* that has the same coordinates. The problem with this construction is that when we change coordinates, the vectors in *V* and in V^* change in different ways. Let's say we happened to have an element in *v* in *V* and an element φ in V^* that had an *intrinsic* relationship, for example, $\varphi(v) = 1$. For $v = \sum_{k=1}^{m} \alpha_k v_k$ and $\varphi = \sum_{j=1}^{m} \gamma_{1j} \varphi_{j1}$,
$\varphi(v) = \sum_{j=1}^{m} \gamma_{1j} \varphi_{j1}\left(\sum_{k=1}^{m} \alpha_k v_k\right) = \sum_{j=1}^{m} \gamma_{1j} \sum_{k=1}^{m} \alpha_k \varphi_{j1}(v_k) = \sum_{j=1}^{m} \gamma_{1j} \alpha_j$. If you transform the α- and γ-coordinates in the same way (e.g., double them, because you halved the lengths of the basis vectors), you would change the value of $\varphi(v)$, and this shouldn't happen – you would want $\varphi(v) = 1$ in all coordinate systems.

Note that in the above "matrix" model, if elements of *V* are represented by column vectors of length *m*, elements of V^* are represented by *row* vectors of length *m*. This allows the coordinates in *V* and V^* to change in different ways when you change basis sets.

This problem – the lack of a natural correspondence of *V* and V^* -- is fixed by adding a little more structure to *V*, namely, an inner product (or dot product). The dot-product implicitly defines distances, perpendicularity, projection, angles, etc. We can always (for finite-dimensional *V*) impose a dot-product once we have chosen coordinates, but it is important to recognize when this is arbitrary.

When dealing with a vector space of signals or stimuli, the dot-product is typically arbitrary. This means that the distinction between *V* (data) and V^* (mappings from data to the field) are different kinds of objects.

An important example is imaging data. Considering an image as a set of values at pixels, $x^{[i]} = \sum_{k=1}^{m} \alpha_k{}^{[i]} x_k$, elements of *V*. (α_k is the value at pixel *k*, x_k is an image consisting of a unit intensity at pixel *k*, and 0 elsewhere): One general task is to describe a set of images $\left\{x^{[i]}\right\}$; here, the simplest kind of solution would consist of an average $\frac{1}{N}\sum_{i=1}^{N} x^{[i]}$ image, which is an element of *V*.

A second task is to distinguish one set of images from another, e.g., to distinguish $\left\{x^{[i]}\right\}$ from some other set $\left\{y^{[i]}\right\}$. Here, a useful strategy is to identify a "decision function" φ, for which the values of $\varphi(x^{[i]})$ are different from the values of $\varphi(y^{[i]})$. Simple (linear) decision functions are part of V^*, not *V*, even though you can describe them by their weights at each pixel. Put another way, there is no reason to assume that the best decision function is the "contrast" (the difference between the means $\frac{1}{N}\sum_{i=1}^{N} x^{[i]} - \frac{1}{N}\sum_{i=1}^{N} y^{[i]}$ -- they're different kinds of objects, and they solve different problems. Making this assumption is the same as making an assumption about the natural "distance" in *V*, effectively, whether pixels are independent.

Another example is the distinction between lights (described by an intensity at each wavelength) and neural mechanisms for color (described by mappings from lights into responses). Here, the vector space of lights is infinite-dimensional, $I(\lambda)$. Vector-space operations in the space of lights can be defined (e.g., superposition), but there is no first-principles way to make a correspondence between lights and mechanisms.

Tensor products

One more way to combine vector spaces. Strange at first, but this is the foundation for (a) finding the intrinsic properties of $Hom(V,V)$, and (b) making a bridge between linear procedures and nonlinear ones.

Same set-up: *V* and *W* are vector spaces over the same field *k*.

The “tensor product” of V and W, $V\otimes W$, is the set of elements $v\otimes w$ and all of their formal linear sums, e.g., $\lambda(v\otimes w)+\lambda'(v'\otimes w')$, along with the following rules for reduction:
$(v\otimes w)+(v\otimes w')=v\otimes(w+w')$,
$(v\otimes w)+(v'\otimes w)=(v+v')\otimes w$,
$\lambda(v\otimes w)=(\lambda v\otimes w)=(v\otimes\lambda w)$.

$v\otimes w$ is an “elementary tensor product.”

Intuitively, the tensor product space is the substrate for functions that act bi-linearly on V and W. Put another way, if a function $f(v,w)$ acts linearly on each argument v and w separately, then it can always be extended to a function that acts linearly on $V\otimes W$. The first of the above rules ensure linearity when components are added in V, the second ensures linearity when components are added in W, and the third, ensures linearity for scalar multiplication.

When V and W are both finite-dimensional (with $\{v_1,\ldots,v_m\}$ is a basis for V, and $\{w_1,\ldots,w_n\}$ is a basis for W), then $V\otimes W$ is of dimension mn, and it has a basis consisting of the elementary tensor products $v_m\otimes w_n$. Writing q in $V\otimes W$ as $q=\sum_{i=1}^{m}\sum_{j=1}^{n}q_{ij}(v_i\otimes w_j)$, we can see that elements in
$V\otimes W$ can be thought of as rectangular arrays, and they are added coordinate-by-coordinate.

The three laws, together, enable us to rewrite any elementary tensor product $v\otimes w$ in this basis. For if $v=\sum_{i=1}^{m}a_iv_i$ and $w=\sum_{j=1}^{n}b_jw_j$, then $v\otimes w=\sum_{i=1}^{m}\sum_{j=1}^{n}q_{ij}(v_i\otimes w_j)$ for $q_{ij}=a_ib_j$. Note, though, that generic elements q of $V\otimes W$ are not elementary tensor products, i.e., cannot be written as just one term $v\otimes w$, since this requires that $q=\sum_{i=1}^{m}\sum_{j=1}^{n}q_{ij}(v_i\otimes w_j)$ where q_{ij} is “separable”, i.e., $q_{ij}=a_ib_j$.

$V\otimes W$ has the same dimension as $Hom(V,W)$, but (just like for V and V^*), there is no intrinsic relationship between them. But see the homework: there is a coordinate-free relationship between $(V\otimes W)^*$ and $Hom(V,W^*)$.

We can extend the above construction to $\left(V\otimes W\right)\otimes X$, etc. We can also verify that $\left(V\otimes W\right)\otimes X$ and $V\otimes\left(W\otimes X\right)$ are abstractly identical, and that there *is* a coordinate-free correspondence, namely, $(v\otimes w)\otimes x\leftrightarrow v\otimes(w\otimes x)$. (The way to show this is that both $\left(V\otimes W\right)\otimes X$ and $V\otimes\left(W\otimes X\right)$ are the same – they are the substrates for the tri-linear functions on V, W, and X.) This extends to h-fold tensor products. Since the associative law holds, we don't need to pay attention to the parentheses when we write out multiple tensor products.

Relationship to familiar (physical) tensors

What is the relationship to more familiar "tensors", such as the diffusion tensor and the conductivity tensor? The short, informal answer is that these objects are elements of a tensor product space in which *V* is 3-dimensional, and their entries transform like the above tensors, under coordinate transformation.

The diffusion tensor is based on a model that particles diffuse via Brownian motion in a medium that may be anisotropic. For ordinary Brownian motion in a 1-dimensional medium, the expected mean-squared distance moved by a particle in time *t* is proportional to time, and the proportionality is the diffusion constant, namely, $< x^2 >= Dt$. In a 3-d medium, if diffusion along each coordinate axis is independent, this generalizes to $< x^2 >= D_{xx}t$, $< y^2 >= D_{yy}t$, and $< z^2 >= D_{zz}t$. But one could imagine that the fastest axis of diffusion is along some oblique axis, i.e., that a cohort of particles released at the origin would tend to form a cloud that is elongated along an oblique axis. So the position of a typical particle along each axis need not be independent; i.e., $< xy >\neq 0$. Working out the physics leads to $< xy >= D_{xy}t$.

So for the "standard" *x*, *y*, *z* coordinate system for *V*, we can characterize the variances and covariances of the particle position by an array $D = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}$. This array must be symmetric, since $D_{xy}t =< xy >=< yx >= D_{yx}t$.

To see that *D* conforms to our more abstract notion of a tensor, we have to verify that *D* transforms in the proper way when we change coordinates. It does. For example, if we choose units in *V* that leads to *b* times the numerical value for a vector that has the same physical length, then we multiply the numerical values of the variances by b^2, and hence, the entries of *D* by b^2. We also need to check that *D* transforms properly when we choose oblique axes, but this works out too.

In other physical situations, the tensor need not be symmetric. For example, a "conductivity tensor" *M* is (abstractly) a quantity in $(V^* \otimes V)^*$ such that $M(E \otimes v)$ gives the current in the direction *v* induced by an electric field E in V^*. An electric field is considered to be a member of the dual space since it is a way of assigning a value (the potential) to an exploring vector. The clue is that electric field have units that include a *reciprocal* length (volts/cm), while vectors have units of length (cm).

The determinant

Here we use these ingredients to define the determinant and derive its properties.

Permutations act on tensor products of a vector space with itself

$V\otimes W$ has additional structure if $V=W$: This arises because of some homomorphisms on $V\otimes V$. We build these homomorphisms by permuting the copies of *V*.

In this simplest case, there is only one nontrivial permutation: a permutation τ that takes 1 to 2, and 2 to 1.This provides a homomorphism on $V\otimes V$ that swaps the first and second copies. Say $q=v^{[1]}\otimes v^{[2]}$, an elementary tensor product. Define $\sigma_\tau(q)=v^{[2]}\otimes v^{[1]}$ for elementary tensor proeducts, and use linearity to extend σ_τ to all of $V\otimes V$. σ_τ is a homomorphism on $V\otimes V$.

For some tensors *x*, $\sigma_\tau(x)=x$ -- these are the "symmetric" tensors (such as the diffusion tensor). The symmetric tensors form a subspace of $V\otimes V$ -- one way to see this is that the symmetric tensors are the kernel of $\sigma_\tau-I$, as $(\sigma_\tau-I)(x)=0$ is equivalent to $\sigma_\tau(x)=Ix=x$. So any transformation *A* that acts on *V* can be thought of as acting linearly in $V\otimes V$, and hence, in the symmetric subspace of $V\otimes V$. We want to generalize this idea, to find a one-dimensional subspace, derived from *V*, in which *A* acts linearly. In this example, if *V* has dimension *n*, then $V\otimes V$ has dimension n^2 (it has basis elements $v_i\otimes v_j$). The symmetric part of $V\otimes V$ has dimension $n+\frac{n(n-1)}{2}=\frac{n(n+1)}{2}$, namely, the elements $v_i\otimes v_i$ and $\frac{1}{2}\left(v_i\otimes v_j+v_j\otimes v_i\right)$ (for $i\neq j$).

The first step in the generalization is that the above construction extends to *h*-fold tensor products $V^{\otimes h}=V\otimes V\otimes\cdots\otimes V$, and to any permutation τ on the set $\{1,2,\ldots,h\}$. For example, if $\tau(1)=4$, $\tau(2)=2$, $\tau(3)=1$, and $\tau(4)=3$ and $q=v^{[1]}\otimes v^{[2]}\otimes v^{[3]}\otimes v^{[4]}$, then $\sigma_\tau(q)=v^{[\tau(1)]}\otimes v^{[\tau(2)]}\otimes v^{[\tau(3)]}\otimes v^{[\tau(4)]}=v^{[4]}\otimes v^{[2]}\otimes v^{[1]}\otimes v^{[3]}$.
Thus, for each permutation τ, we have a homomorphism (actually, an isomorphism) on $V^{\otimes h}$.

The plan is to show that when we choose *h* to be the dimension *m* of *V*, that there is a unique one-dimensional subspace of $V^{\otimes h}$ that we can identify *without resorting to coordinates*. (This is the *n*-fold antisymmetrized tensor product space $anti(V^{\otimes m})$, whose elements are $anti(v^{\otimes m})$, which we will define below.)

Now, assume that we had some linear transformation *A* in $Hom(V,V)$. We could have carried out the above procedure on vectors *v* in *V*, or, on the transformed space vectors Av. Since $anti(v^{\otimes m})$ is one-dimensional, i.e., the field *k*, the two results $anti(v^{\otimes m})$ and $anti((Av)^{\otimes m})$, would have to differ by some scalar factor, i.e., by some multiplier in *k*. This number (the ratio) is the determinant of *A*, $\det(A)$. Put another way, $\det(A)$ indicates how much $anti(V^{\otimes m})$ expands, when *V* is transformed by *A*.

Since we defined the determinant without resorting to coordinates for *V*, we are guaranteed that the determinant is independent of coordinates. The other key properties of the determinant follow immediately.

The determinant of a product is the product of the determinants: $\det(BA) = \det(B)\det(A)$. Since applying *BA* to *v* is the same as applying *B* to *Av*, which is in turn the same as applying *A* and then *B*, we can calculate the expansion of $anti(V^{\otimes m})$ induced by *BA* two ways:

(1) apply *BA* to *v*; this yields $\det(BA)$.
(2) in stages: apply *A* to *v*, yielding a factor of $\det(A)$; apply *B* to Av, yielding a factor of $\det(B)$. Since the two results must be identical, $\det(BA) = \det(B)\det(A)$. Symbolically,

$$\det(BA) = \frac{anti((BAv)^{\otimes m})}{anti(v^{\otimes m})} = \frac{anti((BAv)^{\otimes m})}{anti((Av)^{\otimes m})}\frac{anti((Av)^{\otimes m})}{anti(v^{\otimes m})} = \det(B)\det(A).$$

The determinant of a mapping to a lower-dimensional space is zero. This will follow because we will show that for a vector space of dimension $\leq m-1$, the dimension of $anti(V^{\otimes m})$ is 0.

There is also a very nice, simple geometric view of the determinant as the quotient of $anti((Av)^{\otimes m})$ and $anti(v^{\otimes m})$: it is the amount that the volume of a parallelepiped spanned by *v* expands, when *v* is transformed to *Av*. The fact that this expansion factor is independent of the choice of *v* (i.e,. the choice of paralellopiped) also has a simple geometric interpretation (from Bruce Knight): one could always space-fill one parallelepiped with smaller copies of another one of a different shape, and (just by counting) see that the volume expansion ratio has to be independent of shape.

Construction of the antisymmetrized tensor product

We want to generalize the following from 2 copies of *V* to multiple copies:
For any $q = v^{[1]} \otimes v^{[2]}$ in $V \otimes V$, we have a homomorphism $\sigma_\tau(q) = v^{[2]} \otimes v^{[1]}$ based on the permutation τ that takes 1 to 2, and 2 to 1. Similarly, we can write $\sigma_e(q) = q$, where *e* is the trivial permutation (that takes 1 to 1, and 2 to 2). Now define

$sym = \frac{1}{2}\left(\sigma_e + \sigma_\tau\right)$ and $anti = \frac{1}{2}\left(\sigma_e - \sigma_\tau\right)$. So $sym(v^{[1]} \otimes v^{[2]}) = \frac{1}{2}\left(v^{[1]} \otimes v^{[2]} + v^{[2]} \otimes v^{[1]}\right)$, and $anti(v^{[1]} \otimes v^{[2]}) = \frac{1}{2}\left(v^{[1]} \otimes v^{[2]} - v^{[2]} \otimes v^{[1]}\right)$. The homomorphisms *sym* and *anti* can be thought of as symmetrizing, and antisymmetrizing, the tensors $q = v^{[1]} \otimes v^{[2]}$. That is, $sym(q)$ is unchanged by swapping the components of *q*, and $anti(q)$ is negated by swapping the components of *q*. (Note also that $sym(sym(q)) = sym(q)$, $anti(anti(q)) = anti(q)$, and $sym(anti(q)) = anti(sym(q)) = 0$).

For three copies, *sym* and *anti* are:

$$\begin{gathered} sym(v^{[1]}\otimes v^{[2]}\otimes v^{[3]})=\frac{1}{6}(v^{[1]}\otimes v^{[2]}\otimes v^{[3]}+v^{[2]}\otimes v^{[1]}\otimes v^{[3]}+ \\ v^{[3]}\otimes v^{[1]}\otimes v^{[2]}+v^{[1]}\otimes v^{[3]}\otimes v^{[2]}+v^{[2]}\otimes v^{[3]}\otimes v^{[1]}+v^{[3]}\otimes v^{[2]}\otimes v^{[1]}) \end{gathered}$$

and

$$\begin{gathered} anti(v^{[1]}\otimes v^{[2]}\otimes v^{[3]})=\frac{1}{6}(v^{[1]}\otimes v^{[2]}\otimes v^{[3]}-v^{[2]}\otimes v^{[1]}\otimes v^{[3]}+ \\ v^{[3]}\otimes v^{[1]}\otimes v^{[2]}-v^{[1]}\otimes v^{[3]}\otimes v^{[2]}+v^{[2]}\otimes v^{[3]}\otimes v^{[1]}-v^{[3]}\otimes v^{[2]}\otimes v^{[1]}) \end{gathered} .$$

The general form is

$$sym(v^{[1]}\otimes v^{[2]}\otimes\cdots\otimes v^{[h]})=\frac{1}{h!}\sum_{\tau}v^{[\tau(1)]}\otimes v^{[\tau(2)]}\otimes\cdots\otimes v^{[\tau(h)]}$$

and

$$anti(v^{[1]}\otimes v^{[2]}\otimes\cdots\otimes v^{[h]})=\frac{1}{h!}\sum_{\tau}parity(\tau)\left(v^{[\tau(1)]}\otimes v^{[\tau(2)]}\otimes\cdots\otimes v^{[\tau(h)]}\right)$$

where the summation is over all permutations τ of $\{1,\ldots,h\}$, and $parity(\tau)$ is +1 or -1, depending on whether the number of pairwise swaps required to make τ is even or odd.

More compact form, writing $z=v^{[1]}\otimes v^{[2]}\otimes\cdots\otimes v^{[h]}$:

$sym(z)=\frac{1}{h!}\sum_{\tau}\sigma_{\tau}(z)$ and $anti(z)=\frac{1}{h!}\sum_{\tau}parity(\tau)\sigma_{\tau}(z)$.

The above makes it explicit that *sym* and *anti* are averages over a group (here, the permutation group).

The fact that *sym* and *anti* are averages over a group leads to two properties: if ρ is a permutation that swaps a single pair of indices, then $sym(\sigma_{\rho}z)=sym(z)$ and $anti(\sigma_{\rho}z)=-anti(z)$.

The *sym* property is straightforward:

$$sym(\sigma_{\rho}z)=\frac{1}{h!}\sum_{\tau}\sigma_{\tau}(\sigma_{\rho}z)=\frac{1}{h!}\sum_{\tau}\sigma_{\tau\circ\rho}(z)=\frac{1}{h!}\sum_{\tau}\sigma_{\tau}(z)=sym(z) .$$

The first equality follows from the definition of *sym*, the second from the definition of the group operation (composition) for permutations. The third equality, which is the critical one, from the fact that, since composition by ρ is invertible (because of the group properties), a sum over all permutations $\tau\circ\sigma$ is the same as a sum over all permutations τ.

For the *anti* property, we also need to notice that $parity(\tau_1\circ\tau_2)=parity(\tau_1)parity(\tau_2)$. (That is, *parity* is a homomorphism from the permutation group to $\{+1,-1\}$ under multiplication.) This is because *parity* counts the number of pair-swaps, and $\tau_1\circ\tau_2$ can always be constructed by first applying the pairs needed to make τ_2, and then the pairs needed to make τ_1. So if ρ is a pairwise swap, $parity(\rho) = -1$, and $parity(\tau) = -parity(\tau \circ \rho)$. Consequently,

$$anti(\sigma_\rho z) = \frac{1}{h!}\sum_\tau parity(\tau)\sigma_\tau(\sigma_\rho z) = \frac{1}{h!}\sum_\tau parity(\tau)\sigma_{\tau\circ\rho}(z) = -\frac{1}{h!}\sum_\tau parity(\tau\circ\rho)\sigma_{\tau\circ\rho}(z) \\ = -anti(z)$$
.

To complete the construction of the determinant, we need to (a) show that $parity(\tau)$ is well-defined, and (b) count the dimensions of $anti(V^{\otimes h})$.

To show that $parity(\tau)$ is well-defined, we use a classic trick. Define a polynomial $P(X_1, X_2, \ldots, X_h) = (X_2 - X_1)(X_3 - X_1)(X_3 - X_2) \cdot \ldots \cdot (X_h - X_{h-1})$. This has one term $X_c - X_a$ for each pair of indices (a,c) with $a < c$. We will show that $P(X_{\tau(1)}, X_{\tau(2)}, \ldots, X_{\tau(h)}) = parity(\tau) \cdot P(X_1, X_2, \ldots, X_h)$. First, observe that when we apply τ to the subscripts, we simply scramble the order of the terms, and we also may change some terms into their negatives (if $a < c$ but $\tau(c) < \tau(a)$). So $P(X_{\tau(1)}, X_{\tau(2)}, \ldots, X_{\tau(h)}) = \pm P(X_1, X_2, \ldots, X_h)$. To show that the $\pm$ factor is $parity(\tau)$, we observe that if τ is a single pair-swap (say, of *a* and *c*, with $a < c$), then the sign of *P* is inverted. This is because we can catalog the effects of τ on *P*: the sign of $X_c - X_a$ is inverted (leading to a factor of -1), and, for all *b* between *a* and *c*, pairs of terms $(X_c - X_b)(X_b - X_a)$ become $(X_a - X_b)(X_b - X_c)$, which contributes no net sign change.

Dimension count

To count the dimensions of $anti(V^{\otimes h})$, we count the size of a basis. We start with a basis for $V^{\otimes h}$, and let *anti* act on it. A basis for $V^{\otimes h}$ can be built from the basis $\{v_1, \ldots, v_m\}$ for *V*: select one element of $\{v_1, \ldots, v_m\}$ for each of the *h* copies in the tensor product space. So a typical basis element is $v_{i_1} \otimes v_{i_2} \otimes \ldots \otimes v_{i_h}$, where each of the subscripts $i_1, \ldots, i_h$ is drawn from $\{1, \ldots, m\}$. We can write this compactly as $z_{\vec{i}} = v_{i_1} \otimes v_{i_2} \otimes \ldots \otimes v_{i_h}$.

What happens when *anti* acts on $z_{\vec{i}}$? If any of the subscripts $i_1, \ldots, i_h$ match, then we have to get 0. This is for the following reason. Say τ is a permutation that swaps two of the identical subscripts. On the one hand, $anti(z_{\vec{i}}) = anti(\sigma_\tau(z_{\vec{i}}))$, since $z_{\vec{i}}$ and $\sigma_\tau(z_{\vec{i}})$ are identical. But since τ is a pair-swap, we also have $anti(z_{\vec{i}}) = -anti(\sigma_\tau(z_{\vec{i}}))$. So both quantities must be 0.

So, *anti* maps a basis element $z_{\vec{i}}$ of $V^{\otimes h}$ to 0 if any of its subscripts match. If none of the subscripts match, *anti* maps $z_{\vec{i}}$ to a linear combination of distinct basis elements of $V^{\otimes h}$, which therefore cannot be 0. (For the same reason, $z_{\vec{i}}$'s with distinct subscripts must be linearly indepdenent.) The dimension of $anti(V^{\otimes h})$, which is the count of the number of basis elements that do not map to 0, is the number of ways of choosing *h* distinct elements out of *m* – which is $\begin{pmatrix} m \\ h \end{pmatrix} = \frac{m!}{(m-h)!h!}$.

For *sym* – which we don't need for the determinant, but is useful for other purposes -- the dimension count is the number of ways of choosing *h* elements out of *m* that need not be distinct. This is $\begin{pmatrix} m-1+h \\ m-1 \end{pmatrix} = \frac{(m-1+h)!}{(m-1)!h!}$. (This follows from a standard counting argument, sketched here: To choose a list of *h* items out of the numbers $\{1, \ldots m\}$: imagine you start a counter at 1. At each instant, you take one of two options: either record ("R") the value on the counter or increment it ("I"). After you have made *h* "record" moves and $m-1$ "increment" moves, you will have chosen *h* numbers, possibly with repetition, and the counter will now read *m*, so the process terminates. Every unique set of choices corresponds to a unique sequence of *h* "R" moves and *m*-1 "I" moves. The number of ways of labeling a sequence of $m-1+h$ steps as either R's or I's is the above binomial coefficient.)

Note that for $h=2$, the dimension of the antisymmetric space is $\frac{m(m-1)}{2}$, and the dimension of the symmetric space is $\frac{m(m+1)}{2}$, which adds up to m^2, the dimension of $V \otimes V$. So for $h=2$, we have completely decomposed $V^{\otimes 2} = V \otimes V$ into two parts: $sym(V^{\otimes 2})$ and $anti(V^{\otimes 2})$, and there is nothing left over. For $h \geq 3$, there is a similar decomposition of $V^{\otimes h}$; it involves these two parts and additional parts with more complex symmetries.