TENSORS AT TWILIGHT

RUSHIL MALLARAPU

Contents

To borrow a phrase, the goal of these notes is to teach tensors as a (component of a) mathematical lifestyle. A general theme is that even when you a concrete construction of an object, it takes some playing around to really appreciate it. Time and exposure will help *you* believe you understand tensor products, but it's easier to learn tensors by understanding their behavior. So, we'll do lots of examples, and try to not shy away from using categories to make our lives easier.

All mistakes in these notes are my own. Email me if you spot anything!

1. Universal Properties and "Definitions"

To start, recall the *category of vector spaces over a field k*, which I'll denote $\mathbf{Vect}_k$. It'll be convenient to generalize and consider the *category of modules over a (commutative unital) ring R*, which I'll denote $\mathbf{Mod}_R$.[1] Some examples to keep in mind are $\mathbf{Mod}_k = \mathbf{Vect}_k$, for k a field, and $\mathbf{Mod}_\mathbb{Z} = \mathbf{Ab}$, the category of abelian groups (you proved this on a past homework). I won't spell out a definition of a category or functor or natural transformation because (a) you know what these are already and (b) you should expect every nice collection of mathematical objects with a notion of morphisms between objects to be a category anyway.

To start, here's a definition you've seen in class:

Date: October 25, 2022.

[1]You might have been told that modules are like badly-behaved vector spaces, but this is a lie. Really, vector spaces are *extremely well-behaved* modules. This is why the theory of tensor products over vector spaces is in some sense really boring.

Definition 1.1. Let M, N, L be R-modules. An *R-bilinear map* $f: M \times N \to L$ is a map which is R-linear in each variable separately. Specifically, we have

(a) $f(m + m', n) = f(m, n) + f(m', n)$,

(b) $f(m, n + n') = f(m, n) + f(m, n')$,

(c) $f(mr, n) = rf(m, n) = f(m, rn)$,

for $m, m' \in M$, $n, n' \in N$ and $r \in R$.

For $M, N, L \in \mathbf{Mod}_R$, let $\mathrm{Bilin}_R(M, N; L)$ denote the set of R-bilinear maps $M \times N \to L$. In particular, postcomposition by maps $L \to L'$ makes $\mathrm{Bilin}_R(M, N; -): \mathbf{Mod}_R \to \mathbf{Set}$ a functor from R-modules to sets. With this in hand, we have the best definition of the tensor product you'll ever need.

Definition 1.2. A *tensor product* $M \otimes_R N$ of M and N as R-modules represents the functor $\mathrm{Bilin}_R(M, N; -)$. I.e. it is the "universal" R-module with a bilinear map $M \times N \to M \otimes_R N$.

The goal of this section is to unpack that definition. First, a little unbridled abstraction:

Definition 1.3. Let $\mathcal{C}$ be a category, and $F: \mathbf{C} \to \mathbf{Set}$ a functor. An object $c \in \mathcal{C}$ *represents* F if there exists a natural isomorphism $\alpha: \mathcal{C}(c, -) \cong F$, where $\mathcal{C}(c, -)$ is the functor sending an object $d \in \mathcal{C}$ to the set of morphisms $c \to d \in \mathcal{C}(c, d)$. We say that this natural isomorphism expresses the *universal property* of the object c.

What does this mean? It means every map out of c corresponds uniquely to some element of the image of this functor, which we think of as parametrizing sets of things we care about. For instance, the functor $\mathrm{Bilin}_R(M, N; -)$ records the data of *all* bilinear maps out of $M \times N$. Thus, saying that $M \otimes_R N$ "represents" this functor means that there is a natural (in L) isomorphism $\alpha: \mathrm{Hom}_R(M \otimes_R N, L) \cong \mathrm{Bilin}_R(M, N; L)$, so the tensor product *represents* bilinear maps.

This is our first motivation for *why* tensor products. In general, the functor we're trying to represent might not be as innocent as $\mathrm{Bilin}_R(M, N; -)$; we might care about much more involved and intricate data that depends functorially on the objects in our category! However, if its possible to represent a functor by an object, we're really in luck – somehow, we think we "know" more about maps between objects natively to our category of choice than we know about these weird functors.

Now some desiderata:

Proposition 1.4. *Let c be a representing object for a functor $F: \mathbf{C} \to \mathbf{Set}$. Recall that means we have a natural isomorphism $\alpha: \mathcal{C}(c, -) \cong F$.*

(a) The isomorphism α is determined entirely by where it sends the identity on c, i.e. the element $x := \alpha_c(\mathrm{id}_c) \in Fc$. In particular, the image of $f: c \to d$ is given by $F(f)x \in Fd$, and $\forall y \in Fd$, there exists a unique $f: c \to d$ such that $F(f)x = y$.

(b) Any two such representing objects are isomorphic via a unique isomorphism.

The first part of this proposition is a form of the Yoneda lemma, which is the most important triviality in mathematics, and it's this existence of maps $c \to d$ corresponding to elements of Fd that is usually referred to as a *universal property*, at least colloquially.

Proof. (a) The Yoneda lemma is really something you should learn about and prove for yourself, but here's a basic sketch in the case of interest. The first claim follows from naturality of the isomorphism α applied to the map $f: c \to d$, as follows by tracing $\mathrm{id}_c \in \mathcal{C}(c,c)$ around the naturality square:

$$\begin{tikzcd} \mathcal{C}(c,c) \arrow[r, "\alpha_c"] \arrow[d, "f_*"'] & Fc \arrow[d, "Ff"] \\ \mathcal{C}(c,d) \arrow[r, "\alpha_d"'] & Fd \end{tikzcd}$$

Across the bottom composite, we get $1_c \to f \to \alpha_d(f) \in Fd$, but across the top composite we get $F(f)(\alpha_c(\mathrm{id}_c)) = F(f)x \in Fd$. Commutativity of the diagram tells us that these must coincide. The second claim follows by a similar argument: existence of the unique $f: c \to d$ corresponding to $y \in Fd$ follows from the isomorphism $\alpha_d: \mathcal{C}(c,d) \cong Fd$, and the identification $F(f)x = y$ comes from a similar diagram chase.

(b) The cleanest proof of this (see [9], 2.4.9) requires some technology I don't have time to set up. The basic idea is that if c, c' both represent F, then there is a composite natural isomorphism $\mathcal{C}(c,-) \cong \mathcal{C}(c',-)$, which by some *more* abstract nonsense must come from a unique isomorphism $c \to c'$ (the uniqueness is forced by requiring compatibility with the isomorphism to F). □

Let's specialize to the case of tensor products: by part (b) we know that any possible construction of a tensor product – i.e. any object which has this universal property of representing bilinear maps – will yield a uniquely isomorphic end product, which is why we can speak of *the* tensor product and pick any construction we want (or find convenient in a given situation). This is why Defn. 2.2 is an actual definition – to specify an object, it suffices to specify it up to unique isomorphism. Of course, you need to show that such an object exists, i.e. to construct it, and that's what concrete constructions of the tensor product allow you to do (as well as enabling element-theoretic proofs).

Part (a) is a bit more complicated. First, recall that the action of $\mathrm{Bilin}_R(M, N; -)$ on maps of R-modules $g: L \to L'$ is to send a bilinear map $B: M \times N \to L$ to the postcomposition $g_*B := g \circ B = M \times N \to L \to L'$. Part (a) tells us that the element corresponding to $\mathrm{id}: M \otimes_R N \to M \otimes_R N$ is the "universal bilinear map" $\otimes: M \times N \to M \otimes_R N$, in the sense that any bilinear map $f: M \times N \to L$ factors uniquely through $M \times N \to M \otimes_R N$. This is usually depicted in the following diagram:

$$\begin{tikzcd} M \times N \arrow[r, "\otimes"] \arrow[dr, "f"'] & M \otimes_R N \arrow[d, "\exists! f'"] \\ & L \end{tikzcd}$$

It's likely this is what you were sold as the "universal property of the tensor product", but hopefully it's clear why there's something far richer going on.

Problem 1.5. *Fix any category $\mathcal{C}$ and let $x, y \in \mathcal{C}$ be any two objects. Think through what a representation of the functor $F(c) = \mathcal{C}(x, c) \times \mathcal{C}(y, c)$ is, where the product on the right is the cartesian product in sets. Maybe start by thinking about what the action of F on maps $c \to d$ is. Hint: you've definitely seen this concept before.*

2. Constructions of a Tensor Product

Anyway, I should probably construct a tensor product. Of course, if I work with vector spaces and can pick a basis, I can take the tensor product to be the thing with basis given by the tensors of basis elements, but over a general module I'm not even guaranteed a basis. But, our universal definition gives us an idea of how to construct it: take the most general thing we can think of, and impose all the relations (i.e. take a quotient) such that bilinearity of $\otimes$ holds. This is where the following construction comes from:

Construction 2.1. Let $M, N \in \mathbf{Mod}_R$. Define $F := R\{m \otimes n\}_{m \in M, n \in N}$ to be the free R-module on the formal symbols $m \otimes n$, and quotient out by the relations

$$\begin{gathered} mr \otimes n - r(m \otimes n) \\ m \otimes rn - r(m \otimes n) \\ (m + m') \otimes n - (m \otimes n + m' \otimes n) \\ m \otimes (n + n') - (m \otimes n + m \otimes n'), \end{gathered}$$

i.e. let U be the submodule of F spanned by elements of this form, and take the quotient F/U. Then, F/U satisfies the universal property of the tensor product.

Problem 2.2. *Check this. In particular, show that if $f: M \times N \to L$ is a bilinear map, then there's a well-defined map $f': F/U \to L$, and that there's a "universal bilinear map" $\otimes M \times N \to F/U$ such that $f' \circ \otimes = f$.*

This description gives an idea of what *elements* of the tensor product $M \otimes_R N$ look like – they're linear combinations of things like $m \otimes n$. In fact, we have the following:

Theorem 2.3. *Let $\{x_i\}$ and $\{y_j\}$ be spanning sets of $M, N \in \mathbf{Mod}_R$. Then $M \otimes_R N$ is spanned by the elementary tensors $x_i \otimes y_j$.*

Proof. It suffices to show every elementary tensor $m \otimes n$ is a linear combination of $x_i \otimes y_j$'s. Write $m = \sum a_i x_i$ and $n = \sum b_j y_j$. Then by bilinearity we have

$$m \otimes n = \sum_i (a_i x_i \otimes n) = \sum_{ij} a_i b_j (x_i \otimes y_j),$$

and we are done. □

Corollary 2.4. *Let V and W be finite-dimensional vector spaces over a field k, and pick bases $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$. Then, $\{v_i \otimes w_j\}$ is a basis for $V \otimes_k W$; in particular* $\dim V \otimes_k W = \dim V \cdot \dim W$.

Proof. By the previous theorem, ll that needs to be shown is linear independence. So, let $a_{ij} \in k$ be such that $\sum_{ij} a_{ij} v_i \otimes w_j = 0$. To show a_{ij} is zero, define a map

$$\varphi_{ij} \colon V \otimes_k W \to k.$$

which, for $v = \sum b_i v_i$ and $w = \sum c_j w_j$, maps $\varphi_{ij}(v \otimes w) = b_i c_j$, and extends by linearity; this is well-defined precisely because $\{v_i\}$ and $\{w_j\}$ are linearly independent! In particular, $\varphi_{ij}(v_{i'} \otimes w_{j'}) = 1$ if $(i, j) = (i', j')$ and 0 otherwise. Then, applying φ_{ij} to $\sum_{ij} a_{ij} v_i \otimes w_j = 0$, we get $a_{ij} = 0$, so all the a_{ij} must be zero, and we are done. □

This corollary tells us that, in some sense, tensor products of *just* vector spaces, which are always finite free, are really boring. That's why I'm setting this all up for arbitrary modules, although I could be far more brutally general. Perhaps for another time...

By the way, this proof doesn't try to divide by coefficients in k anywhere, so it's entirely valid for free modules over a ring. Note what we did here – instead of defining a map on a tensor product $V \otimes_k W$ by writing down a bilinear map $V \times W$, we just define it on pure tensors and extend by linearity. This is only valid when we've checked that the map we've written down, when treated as a map from $V \times W$, is actually bilinear, but usually this is either clear or not worth explicitly checking. At the end of the day, everything goes through the universal property of the tensor product!

3. Examples, Examples, Examples

This is the part where we do examples.

Theorem 3.1. $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \cong \mathbf{Z}/\gcd(a, b)\mathbf{Z}$, *for* $a, b \in \mathbf{Z}$.

Proof. We know by Theorem 2.3 that $1 \otimes 1$ spans $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z}$, and a quick calculation shows that it's order is a divisor of both a and b. However, we can do better and give an explicit isomorphism. Let $f \colon \mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \to \mathbf{Z}/\gcd(a, b)\mathbf{Z}$ be the map sending $x \otimes y$ to $xy \mod d$ (you should check that multiplication is bilinear – while addition is not – and that this is well-defined). Furthermore, let $g \colon \mathbf{Z}/\gcd(a, b)\mathbf{Z} \to \mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z}$ be the map sending a to $a(1 \otimes 1)$. To see that this is well-defined, use Bezout's lemma: if $a, b \in \mathbf{Z}$, then there exist integers x, y such that $ax + by = \gcd(a, b)$ (equivalently, $a\mathbf{Z} + b\mathbf{Z} = \gcd(a, b)\mathbf{Z}$). Now, we need only check that $d := \gcd(a, b)$ gets sent to 0:

$$d \mapsto d(1 \otimes 1) = d \otimes 1 = ax + by \otimes 1 = ax \otimes 1 + by \otimes 1 = 0 + y \otimes b = 0.$$

Now, you can check that f and g are inverses:

$$f(g(x)) = f(x \otimes 1) = x \quad \text{and} \quad g(f(x \otimes y) = g(xy) = xy \otimes 1 = x \otimes y,$$

thus giving the desired isomorphism. ◻

This is a generalization of the following theorems:

Theorem 3.2. *Let I, J be ideals of a ring R. Then $R/I \otimes_R R/J \cong R/(I + J)$.*

Proof. This is exactly the same as before. Define $f: R/I \otimes_R R/J \to R/(I+J)$ as $f(x \otimes y) = xy$ and $g: R/(I + J) \to R/I \otimes_R R/J$ as $g(x) = x \otimes 1$. You need to do some algebra to check that these are well-defined, but once that's done, it's clear that they're inverses, giving the desired isomorphism. ◻

Alternatively, given an ideal $I \leq R$ and an R-module M, we can define M/IM as the quotient module of M by the submodule of the form $IM = \{\sum_i a_i m_i \mid a_i \in I, m_i \in M\}$, where the sum is of finitely many nonzero terms.

Theorem 3.3. *$R/I \otimes_R M \cong M/IM$.*

Proof. To write down the map, let $f: R/I \otimes_R M \to M/IM$ be given by $f(a \otimes m) = am$.To define the inverse, let $g': M \to R/I \otimes_R M$ be defined by $g'(m) = (1 \otimes m)$. Note this sends things like $a_i m_i$, with $a_i \in I$ and $m_i \in M$, to $1 \otimes a_i m_i = a_i \otimes m_i = 0$, so it passes to the quotient to give a map $g: M/IM \to R/I \otimes_R M$. You can check these maps are inverses. ◻

To see why this recovers the previous theorem, take $M = R/I$ and observe that

$$(R/I)/J(R/I) = (R/I)/(I + J/I) \cong R/I + J,$$

by the third isomorphism theorem!

Theorem 3.4. *$\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Z} \cong \mathbf{Q}$ and $\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Z}/a\mathbf{Z} \cong 0$.*

Proof. The first is a direct consequence of Proposition 3.7(d). For the second, we can compute on elementary tensors $x \otimes y \in \mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Z}/a\mathbf{Z}$:

$$x \otimes y = x \cdot \frac{a}{a} \otimes y = \frac{x}{a} \otimes ay = \frac{x}{a} \otimes 0 = 0.$$

◻

This says that tensoring with $\mathbf{Q}$ turns free abelian groups into $\mathbf{Q}$-vector spaces, and kills torsion abelian groups. This has tons of ramifications for computational methods in algebraic topology and algebraic geometry and algebraic number theory!

A fun example which I won't prove (and you should ponder) is this:

Proposition 3.5. $R[x] \otimes_R R[y] \cong R[x, y]$.

Finally, I want to talk about a very general setup. Let $\varphi: R \to S$ be a homomorphism of unital rings. This means that $\varphi(x + y) = \varphi(x) + \varphi(y)$, $\varphi(xy) = \varphi(x)\varphi(y)$, and $\varphi(1) = 1$. Now, if we have an S-module N, we can consider this as an R-module, where the R-scaling is given by $r \cdot n = \varphi(r)n$. Turning an S-module into an R-module like this is called *restriction of scalars*. For example, this is how you can treat a $\mathbf{C}$-vector space as an $\mathbf{R}$-vector space, by restriction of scalars along the natural embedding $\mathbf{R} \hookrightarrow \mathbf{C}$.

However, if we start with an R-module M, we can also turn this into an S-module in the following way: treat S as an R-module via φ and form the tensor product $S \otimes_R M$. This is an S module via $s \cdot (s' \otimes m) = ss' \otimes m$. In fact, we even have the following "universal characterization" of this extension of scalars, which is that given an R-module M, an S-module N, and an R-linear map $f: M \to N$, then this factors uniquely as the embedding $M \to S \otimes_R M$ sending $m \mapsto 1 \otimes m$ followed by an S-linear map $f': S \otimes_R M \to N$.

$$\begin{tikzcd} M \arrow[r, "m \mapsto 1 \otimes m"] \arrow[dr, "f"'] & S \otimes_R M \arrow[d, "f'"] \\ & N \end{tikzcd}$$

The map f' is defined by sending $s \otimes m$ to $s \otimes f(m)$, and then using the S-module structure on N to send this to $sf(m) \in N$. This fact is often expressed as a natural isomorphism

$$\mathrm{Hom}_S(S \otimes_R M, N) \cong \mathrm{Hom}_R(M, N),$$

called the *extension-restriction adjunction*, between the extension and restriction functors:

$$\begin{tikzcd} \mathbf{Mod}_R \arrow[r, bend left, "S \otimes_R -"] \arrow[r, phantom, "\perp"] & \mathbf{Mod}_S \arrow[l, bend left, "\text{restrict}"] \end{tikzcd}$$

Proposition 3.6. *The following are some examples of extension of scalars:*

(a) Complexification: if V is an $\mathbf{R}$-vector space, then $V_{\mathbf{C}} := V \otimes_{\mathbf{R}} \mathbf{C}$ is a complex vector space known as the complexification of $\mathbf{R}$. If $\{e_1, \ldots, e_n\}$ is a basis of V, then $e_i \otimes 1$ and $e_i \otimes i$ is a $\mathbf{R}$-basis of $V_{\mathbf{C}}$. However, $e_i \otimes 1$ is a $\mathbf{C}$-basis of $V_{\mathbf{C}}$, as $i(e_i \otimes 1) = e_i \otimes i$. In particular, $V_{\mathbf{C}}$ has complex dimension n, but real dimension $2n$.

(b) Polynomial coefficients: for any ring R, $R[X] \cong R \otimes_{\mathbf{Z}} \mathbf{Z}[X]$, using that every (commutative unital) ring has a unique map from $\mathbf{Z} \to R$.

(c) *Reduction of coefficients: if $I \le R$ is an ideal, then $M/IM \cong R/I \otimes_R M$ is the extension of scalars of M from $\mathbf{Mod}_R$ to $\mathbf{Mod}_{R/I}$ (even though we're reducing mod I).*

Another common use of such things is extending modules over an integral domain to vector spaces over its fraction field, which can be used to prove that if $R^m \cong R^n$, then $m = n$. We won't get into this however. Note also that with complexification, if V is a complex vector space of dimension n, it can be treated as a real vector space of dimension $2n$; if I complexify that back to a complex vector space, I'll actually get $V \oplus iV \cong V \otimes_{\mathbf{R}} \mathbf{C}$, not just V, because restriction of scalars forgets about the complex linear structure on V.

I want to end with the following list of propositions about how tensor products interact with other things. In particular, there are no assumptions being made here about the modules involved, and all of these are *natural*, in an appropriate sense. I'm not going to prove them, because (1) great proofs already exist in numerous places in writing and online and (2) I recommend thinking them over for yourself. The last one, often referred to as the "tensor-hom adjunction" might show up on pset 11, but the rest boil down to showing that both of the displayed things represent the same set of bilinear maps!

Proposition 3.7. *Let M, N, L be R-modules. There are unique isomorphisms*

(a) $M \otimes_R N \to N \otimes_R M$, *with* $m \otimes n \mapsto n \otimes m$,

(b) $(M \otimes_R N) \otimes_R L \to M \otimes_R (N \otimes_R L) \to M \otimes_R N \otimes_R L$,

(c) $\bigoplus_{i \in I} (M_i \otimes_R N) \to (\bigoplus_{i \in I} M_i) \otimes_R N$, *where I could be infinite,*

(d) $R \otimes_R M \to M$, *with* $r \otimes m \mapsto rm$,

(e) $\mathrm{Hom}_R(M \otimes_R N, L) \cong \mathrm{Hom}_R(M, \mathrm{Hom}_R(N, L))$.

In particular, isomorphism (3) should tell you that the tensor product "distributes" over direct sum, much like how the product of numbers distributed over addition. In accordance with the principle that an object's behavior should tell you what it looks like, think of this as telling you why the tensor product is like a "product" in the first place.[2]

Finally, here are two isomorphisms that are in some sense *special* to the case of finite free modules, e.g. vector spaces.

Proposition 3.8. *Let M, N be R-modules. There is a natural map $\varphi\colon M^\vee \otimes_R N \to \mathrm{Hom}_R(M, N)$ sending $\ell \otimes n$ to $m \mapsto \ell(m)n$. If M, N are finite free, this map is an isomorphism.*

Proof. It is possible to directly check φ is injective and surjective. For surjectivity in particular, you should that if you let $\{e_1, \dots, e_n\}$ be a basis of M and $f_1, \dots, f_m$ be a basis of N, then the image of φ on the basis $\{e_i^\vee \otimes f_j\}$ in $\mathrm{Hom}_R(M, N)$ is the matrix with a 1 in the (j, i)'th position and 0's elsewhere – this is a basis for all $m \times n$ matrices! However, given

[2]The correct statement is that $(\mathbf{Mod}_R, \otimes_R, R)$ is a (closed) symmetric monoidal category, but don't worry about what this means – it effectively says that all is well in the world of $\mathbf{Mod}_R$.

a basis, it is straightforward to write down an inverse map. Let $T \in \mathrm{Hom}_R(M, N)$ be an R-linear map, and check that the map

$$\psi : T \mapsto \sum_i e_i^\vee \otimes T(e_i)$$

is an inverse to φ, using properties of the dual basis. □

Remark 3.9. This isn't important, but this is true whenever M is *finitely generated projective*; because $\mathbf{Mod}_R$ is closed symmetric monoidal, $M^\vee = \mathrm{Hom}_R(M, R)$ is always the *categorical* dual, not just the algebraic dual, if M is finitely generated projective; these are precisely the dualizable objects in $\mathbf{Mod}_R$. If M is f.g. projective, then using it's splitting as a summand of a finite free, we can get a "dual generating set" for $M^\vee$, which lets us write down the inverse map. Maybe there's some connection here to why f.g. projective things show up and the algebraic K-theory of R, but take this with a grain of salt.

Proposition 3.10. *Let M, N be R-modules. There is a natural map $\varphi : M^\vee \otimes_R N^\vee \to (M \otimes_R)^\vee$ sending $f \otimes g$ to $m \otimes n \mapsto f(m)g(n)$. If M, N are finite free, this map is an isomorphism.*

Proof. As before, one can pick a basis and write down an inverse in the finite free case. However, we can also use the previous propositions to string together the isomorphisms

$$M^\vee \otimes_R N^\vee \cong \mathrm{Hom}_R(M, N^\vee) \cong \mathrm{Hom}_R(M \otimes_R N, R) = (M \otimes_R N)^\vee.$$

Try tracing through the maps here to see that it does what you expect. □

Problem 3.11. *Convince yourself that, in principle, using the structure theorem for finitely generated abelian groups, you now know how to compute all tensor products of such things.*

4. Tensor Algebras and Friends

I want to discuss "tensor algebras" and some related constructions, namely symmetric, exterior, and divided powers. Throughout all of this, we'll work with (finite-dimensional) vector spaces over a field k, which you should *not* assume to be characteristic 0. First, some background:

Definition 4.1. Fix a ring R. An R-algebra A is one of the following equivalent things:

(1) A ring A with a ring homomorphism $f : R \to A$ such that $f(R)$ is contained in the center of A; i.e. if I identify R with it's image in A, then $ra = ar$ for all $r \in R, a \in A$, even if A is not commutative.

(2) An R-module A with a unital multiplicative structure, i.e. an R-linear map $\mu : A \otimes_R A \to A$ satisfying some axioms. Namely, we require that μ is associative and unital, so that writing $x \cdot y = \mu(x, y)$, we have $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ and $1 \cdot x = x \cdot 1 = x$.

In both cases, you can think of R as the "scalars" of A.

Some examples of an $\mathbf{R}$-algebra are the algebra of $n \times n$ matrices, $M_n(\mathbf{R})$, as matrices can be added and scaled, but also multiplied, or the polynomial algebra $\mathbf{R}[x]$, as polynomials can be multiplied. In fact, $\mathbf{R}[x]$ is a *commutative* $\mathbf{R}$-algebra; i.e. it is a commutative ring, or via definition 4.1.(2), we have $x \cdot y = y \cdot x$ for all $x, y \in A$. An important case to keep in mind is that $\mathbf{Z}$-algebras are the same as associative rings, and commutative $\mathbf{Z}$-algebras are the same as commutative rings.

Definition 4.2. Let $V \in \mathbf{Vect}_k^{fd}$ be a finite-dimensional k-vector space. The *tensor algebra* on V is the free associative k-algebra on V. In particular, letting $V^{\otimes n}$ be the n-fold tensor product of V, where $V^{\otimes 0} = k$, the tensor algebra is, as a vector space,

$$TV := \bigoplus_{n \geq 0} V^{\otimes n}.$$

The algebra structure comes from the bilinear pairing $V^{\otimes n} \otimes_k V^{\otimes m} \to V^{\otimes(n+m)}$ sending

$$(v_1 \otimes \ldots \otimes v_n) \otimes (w_1 \otimes \ldots \otimes w_m) \mapsto v_1 \otimes \ldots \otimes v_n \otimes w_1 \otimes \ldots \otimes w_m.$$

Note that this is a noncommutative algebra in general. If $e_1, e_2, , e_3$ is a basis of V, then an example element in TV looks like $(2e_1 + e_2) \otimes e_3 + e_1 \otimes e_2 \otimes e_3 + e_2 \otimes e_1 \otimes e_3$, etc. It's worth pointing out that this is a *graded algebra*, which means that (1) it decomposes into a direct sum of pieces of different *gradings*, which in this case are the pieces $V^{\otimes n}$, i.e. length n tensors, and (2) addition and multiplication respect the grading, which really means that the product of something in degree n and degree m lives in degree $n + m$.

This might seem like a big and unwieldy object, but it has some nice properties. In particular, it's the *universal* associative algebra on V, which means that if $f : V \to A$ is any k-linear map from V to an associative algebra A, then it extends to an k-algebra map $TV \to A$ which factors the inclusion of $V = V^{\otimes 1} \hookrightarrow TV$ as the degree 1 component.

With this in mind, the *symmetric algebra* is a natural thing to define next: see if you can guess what I'm going to say.

Definition 4.3. The *symmetric algebra* on V is the free commutative k-algebra on V. The graded components are written $\mathrm{Sym}^d(V) \subset V^{\otimes d}$, and are called the d-th symmetric powers of V. This is the quotient of $V^{\otimes d}$ by all commutators, i.e. symbols of the form $v \otimes w - w \otimes v$. The whole algebra is

$$\mathrm{Sym}\, V := \bigoplus_{n \geq 0} \mathrm{Sym}^d(V).$$

The algebra structure comes from the concatenation pairing $\mathrm{Sym}^n V \otimes_k \mathrm{Sym}^m V \to \mathrm{Sym}^{n+m} V$.

I want to emphasize that the symmetric algebra is a *quotient* algebra, not a subspace algebra. Fulton-Harris, the representation theory book you'll read this November, likes to conflate the two because basically everything happens in characteristic zero, but this will mess up your intuitions.

A good exercise is to spell out what the universal property of the symmetric algebra on V actually is, in terms of maps from V to a commutative k-algebra. More concretely, however, Sym is like a basis-free polynomial algebra. In particular, if V has dimension n, then choosing a basis $e_1, \ldots, e_n$ of V, we get an isomorphism of commutative algebras

$$\operatorname{Sym} V \cong k[e_1, \ldots, e_n],$$

where the d-th graded pieces on the left correspond to homogenous polynomials of degree d on the right. Note that by our conventions, $\operatorname{Sym} k \cong k[x]$ and $\operatorname{Sym} 0 \cong k$; that last one says that a polynomial k-algebra on 0 variables is just your ground field.

One way to describe $\operatorname{Sym}^d V$ is that there is an action of $\Sigma_d \curvearrowright V^{\otimes d}$ given by permuting factors. An element of $\operatorname{Sym}^d V$ then is an *orbit* of this action, i.e. we identify two tensors which are related by a permutation. Now the dual concept to orbits are fixed points, so we might ask what the fixed points are.

Definition 4.4. The *divided power* $\Gamma^d V$ is the subspace of symmetric tensors in $V^{\otimes d}$, i.e. fixed points of the Σ_d action.

The algebra structure with this is quite involved, so I won't get into it. You might see these things come up honestly if you ever read about *crystalline cohomology*, but literally do not worry about that. The important thing is that while we can give a basis for $\operatorname{Sym}^d V$ by taking degree d monomials in $e_1, \ldots, e_n$, a basis for $\Gamma^d V$ comes by considering, for $a_1 + \ldots + a_n = d$, the *divided power monomials*

$$e_1^{(a_1)} \cdots e_n^{(a_n)} = \sum_{\sigma \in \Sigma_d} \sigma \cdot e_1^{\otimes a_1} \otimes \cdots e_n^{\otimes a_n}.$$

A combinatorics argument (count the number of ways to put d unlabeled balls into n boxes) tells us that

$$\dim \operatorname{Sym}^d V = \dim \Gamma^d V = \binom{n+d-1}{d}.$$

Now, there's a natural map $\alpha \colon \Gamma^d V \to \operatorname{Sym}^d V$ given by including into $V^{\otimes d}$ and projecting down to $\operatorname{Sym}^d V$, which is defined elementwise on a basis as

$$\alpha \colon e_1^{(a_1)} \cdots e_n^{(a_n)} \mapsto \frac{d!}{a_1! \cdots a_n!} e_1^{a_1} \cdots e_n^{a_n}.$$

This is invertible when the coefficients $d!$ are invertible, so in particular when $\operatorname{char} k = 0$ or $\operatorname{char} k > d$. Otherwise, it is neither injective or surjective in general. The natural map in the other direction is often called the *norm* map, which is usually given for any group action from orbits to fixed points (or as is usually stated, from coinvariants to invariants):

$$\begin{aligned} \beta \colon \operatorname{Sym}^d V &\longrightarrow \Gamma^d V \\ v_1 \cdots v_d &\longmapsto \frac{1}{d!} \sum_{\sigma \in \Sigma_d} v_{\sigma(1)} \cdots v_{\sigma(d)}. \end{aligned}$$

To see when this map is not an isomorphism, consider what happens over $\mathbf{F}_2$; here, the image of $x \otimes y + y \otimes x$ under α is $2xy = 0$.

The final character I want to introduce is the exterior algebra.

Definition 4.5. The *d-th exterior power* $\bigwedge^d V$ is the quotient of $V^{\otimes d}$ by elements of the form $v \otimes w + w \otimes v$. Writing the image of $v \otimes w$ as $v \wedge w$, or "v wedge product w", we have $v \wedge w = -w \wedge v$, which implies (when char $k \neq 2$) that $v \wedge v = 0$. The *exterior algebra* is the free alternating graded k-algebra, i.e. is given by

$$\bigwedge V \coloneqq \bigoplus_{d \geq 0} \bigwedge^d V.$$

This is an algebra under wedge product.

Now, we can again cook up a basis by some combinatorics. You should check via symbol-pushing that if $v_1 \wedge \ldots \wedge v_d$ is an element of $\bigwedge^d V$ with any $v_i = v_j$ for $i \neq j$, then $v_1 \wedge \ldots \wedge v_d = 0$, and because reordering just introduces a sign, we can always put elements into a canonical order. Thus, if $e_1, \ldots, e_n$ is a basis of V, then a basis for $\bigwedge^d V$ is given by wedges of the form $e_{i_1} \wedge \ldots \wedge e_{i_d}$, for $1 \leq i_1 < \ldots < i_d \leq n$. Thus, we have that

$$\dim \bigwedge^d V = \binom{n}{d} \text{ if } d < n, \text{ and } 0 \text{ otherwise.}$$

Thus, while $\operatorname{Sym} V$ is infinite dimensional, $\dim \bigwedge V = 2^n$, and $\dim \bigwedge^n V = 1$. However, there is famously no canonical isomorphism from $\bigwedge^n V$ to k; any such isomorphism requires a basis to define. By the way, we can extend maps of vector spaces to maps of tensor products of vector spaces, by acting separately in each variable. This passes to quotients, so we can define the exterior power of a linear map $f \colon V \to W$ as

$$\begin{aligned} f^{\wedge d} \colon \bigwedge^d V &\longrightarrow \bigwedge^d W \\ v_1 \wedge \ldots \wedge v_d &\longmapsto f(v_1) \wedge \ldots \wedge f(v_d). \end{aligned}$$

Because $\bigwedge^n V \cong k$ (noncanonically), this means that if f is an endomorphism of V, then $f^{\wedge n}$ is given by multiplication by a scalar. This is the *determinant* of f.

Also, while bilinear forms, i.e. $\operatorname{Bilin}_k(V, V; k) = \operatorname{Hom}_k(V \otimes V, k) = (V \otimes V)^\vee$ decompose naturally into symmetric and skew-symmetric forms (for V finite dimensional), this isn't true for k-forms, i.e. elements of $(V^{\otimes k})^\vee$. The fundamental reason comes down to the *representation theory of* Σ_k, but a quick sanity check is that for $d = 2$,

$$\dim V^{\otimes 2} = \dim \operatorname{Sym}^2 V + \dim \bigwedge^2 V.$$

Two facts which I won't prove are the following isomorphisms. They won't be useful immediately, but you should ponder them all the same. The nicest proofs involve using the notion of a *perfect pairing*, which you should google.

Theorem 4.6. *Let V be a finite dimensional vector space and $d \geq 1$. There are canonical isomorphisms*

$$(\operatorname{Sym}^d V)^\vee \cong \Gamma^d(V^\vee) \quad \textit{and} \quad \bigwedge^d (V^\vee) \cong \left(\bigwedge^d V \right)^\vee,$$

the left one when char $k > d$.

5. Dimension and Trace

Here's some additional cool things that come naturally out of the theory of tensor products. I want to shout out Keeley Hoek's linear algebra notes [7] for being an absolutely amazing presentation of this material.

Fix $\mathcal{C}$ a *closed symmetric monoidal category.* There's a formal definition of this in terms of some axioms and commutative diagrams, but here's the key idea. The data of such a thing is a category $\mathcal{C}$, bifunctors[3] $- \otimes -: \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ and $\mathrm{Hom}(-,-): \mathcal{C}^{\mathrm{op}} \times \mathcal{C} \to \mathcal{C}$, called the "tensor" and "internal hom," and a distinguished object $1 \in \mathcal{C}$ called the unit. These satisfy:

(a) $\otimes$ is an associative product and 1 is its unit. This means that $X \otimes Y \otimes Z \cong X \otimes (Y \otimes Z) \cong (X \otimes Y) \otimes Z$, all via specified isomorphisms, and $X \cong 1 \otimes X \cong X \otimes 1$. These isomorphisms are also required to be natural and "nice" in some technical ways.

(b) $\otimes$ is a symmetric product – i.e. there are natural isomorphisms $X \otimes Y \cong Y \otimes X$.

(c) The internal hom is right adjoint to the tensor – this means that for all objects $X, Y, Z \in \mathcal{C}$, there is a natural isomorphism (of sets)

$$\mathrm{Mor}_{\mathcal{C}}(X \otimes Y, Z) \cong \mathrm{Mor}_{\mathcal{C}}(X, \mathrm{Hom}(Y \otimes Z))$$

(which is currying), and this internalizes to an isomorphism in $\mathcal{C}$ as

$$\mathrm{Hom}(X \otimes Y, Z) \cong \mathrm{Hom}(X, \mathrm{Hom}(Y, Z)).$$

(d) By some abstract nonsense, this implies that $\otimes$ preserves "colimits" separately in each variable. Practically speaking, this means it preserves (1) coproducts, i.e. direct sums, and (2) cokernels, i.e. quotients.

Now, here's your favorite example of a closed symmetric monoidal category: $\mathbf{Mod}_R$. The tensor is $- \otimes_R -$, the unit is R, and the internal hom is $\mathrm{Hom}_R(-,-)$, which is itself an R-module. We mentioned that $\otimes_R$ distributes over direct sums, but it's also true that if

$$M \xrightarrow{f} N \xrightarrow{g} L \to 0$$

is an exact sequence, i.e.

$$L \cong \frac{N}{\mathrm{im}(f: M \to N)},$$

then for any $P \in \mathbf{Mod}_R$, we get an exact sequence

$$M \otimes_R P \xrightarrow{f \otimes \mathrm{id}_P} N \otimes_R P \xrightarrow{g \otimes \mathrm{id}_P} L \otimes_R P \to 0.$$

[3]That is, if you fix one of the inputs, you get a functor in the other input; also the notation $\mathcal{C}^{\mathrm{op}}$ means that $\mathrm{Hom}(-,-)$ is contravariant in the first variable.

This is what it means to "preserve cokernels", and actually lets us compute things like $\mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} P$ in a neat way. In this case, we can write $\mathbf{Z}/2\mathbf{Z}$ in the following exact sequence

$$\mathbf{Z} \xrightarrow{2} \mathbf{Z} \to \mathbf{Z}/2\mathbf{Z} \to 0.$$

Then, tensoring with P and applying the natural isomorphism $\mathbf{Z} \otimes_{\mathbf{Z}} P \cong P$ (under this isomorphism, $2 \otimes \mathrm{id}_P$ gets sent to the map of multiplication by 2 on P), we get

$$P \xrightarrow{2} P \to \mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} P \to 0.$$

Thus, we get that

$$\mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} P \cong \frac{P}{\{2x | x \in P\}} = P/2P.$$

Anyway, being closed symmetric monoidal is a fancy way of saying that you have something that behaves like a tensor product, and if I was a mean person, I could make this entire talk 1 sentence by saying that "$\mathbf{Vect}_k$ is a closed symmetric monoidal category under the tensor product" and just leaving. The point of this is to define a dualizable object.

Definition 5.1. Fix a closed symmetric monoidal category $\mathcal{C}$. An object $X \in \mathcal{C}$ is *dualizable* if there exists a "dual" $X^\vee \in \mathcal{C}$ and maps $\eta\colon 1 \to X^\vee \otimes X$ (the "unit" or "coevaluation") and $\epsilon\colon X \otimes X^\vee \to 1$ (the "counit" or "evaluation") such that the following diagrams commute:

$$\begin{tikzcd} V \cong V \otimes 1 \arrow[r, "\mathrm{id} \otimes \eta"] \arrow[dr, equal, "\mathrm{id}_V"'] & V \otimes V^\vee \otimes V \arrow[d, "\epsilon \otimes 1"] \\ & V \cong 1 \otimes V \end{tikzcd} \qquad \begin{tikzcd} V^\vee \cong 1 \otimes V^\vee \arrow[r, "\eta \otimes 1"] \arrow[dr, equal, "\mathrm{id}_{V^\vee}"'] & V^\vee \otimes V \otimes V^\vee \arrow[d, "1 \otimes \epsilon"] \\ & V^\vee \cong V^\vee \otimes 1 \end{tikzcd}$$

These diagrams are often called the "triangle identities."

Some consequences of this are natural isomorphisms for any dualizable X and objects $Y, Z \in \mathcal{C}$ given as

$$\mathrm{Hom}(X \otimes Y, Z) \cong \mathrm{Hom}(Y, X^\vee \otimes Z) \quad \text{and} \quad \mathrm{Hom}(X^\vee \otimes Y, Z) \cong \mathrm{Hom}(Y, X \otimes Z).$$

The first isomorphism is given by taking a map $f\colon X \otimes Y \to Z$, and mapping this to

$$\hat{f}\colon Y \cong 1 \otimes Y \xrightarrow{\eta \otimes \mathrm{id}_Y} X^\vee \otimes X \otimes Y \xrightarrow{\mathrm{id}_{X^\vee} \otimes f} X^\vee \otimes Z.$$

I'll leave it as an exercise to define the rest of the maps here.

The most important example of a dualizable object is a *finite-dimensional* vector space V. In general, the "algebraic dual" of an object P in a closed symmetric monoidal category is defined to be $P^\vee := \mathrm{Hom}(P, 1)$, but this is often not a *categorical* dual. We can check these axioms for V and $V^\vee$ however. Fix a basis $e_1, \ldots, e_n$ of V, and let e_i^* be the corresponding dual basis. The evaluation map is

$$\begin{aligned} \epsilon\colon V \otimes V^\vee &\longrightarrow k \\ v \otimes \ell &\longmapsto \ell(v). \end{aligned}$$

and the unit map is

$$\begin{aligned} \eta\colon k &\longrightarrow V^\vee \otimes V \\ 1 &\longmapsto \sum_{i=1}^n e_i^* \otimes e_i. \end{aligned}$$

In particular, we need finite dimensionality of V for the sum on the right to be well-defined. You should write out what the triangle identities look like – because of this well-definedness issue, $V^\vee$ isn't a categorical dual if V is infinite-dimensional.

The general intuition of dualizability is that of "finiteness" – specifically, an object is dualizable if its size is smaller than the "additivity" of the latent category. By the way, dualizable objects in $\mathbf{Mod}_R$ are precisely finitely generated projective modules, which means a module P such that there is a surjection $r\colon R^n \to P$ which is split by a map $s\colon P \to R^n$ such that $rs = 1_P$ (but maybe not $sr = 1_{R^n}$). This means that P is a submodule of a finite free module. Because over a field (or at least a PID) submodules of free modules are free, we know that the dualizable objects in $\mathbf{Mod}_k = \mathbf{Vect}_k$ are precisely finite dimensional vector spaces. Really this is getting too far into Math 221 and Math 225 material though; email me if you want some more details on this, or see [12].[4]

The immediate payoff of all of this is an even better abstract definition of the trace, which applies for endomorphisms of a dualizable object in any closed symmetric monoidal category. Let $f\colon V \to V$ be an endomorphism of a dualizable object V. Then, we can form the following composite:

$$\operatorname{tr}(f) := 1 \xrightarrow{\eta} V^\vee \otimes V \xrightarrow{\mathrm{id}_{V^\vee} \otimes f} V^\vee \otimes V \cong V \otimes V^\vee \xrightarrow{\epsilon} 1.$$

This is the *trace* of f. In many cases, this map is given by multiplication by some scalar, which is also called the trace.

[4]It's also worth pointing out that any dualizable object supports an isomorphism as in Proposition 3.8: all we need is the mate isomorphisms mentioned above. Take any other object $W \in \mathcal{C}$, and do

$$\operatorname{Hom}(V, W) \cong \operatorname{Hom}(1 \otimes V, W) \cong \operatorname{Hom}(1, V^\vee \otimes W) \cong V^\vee \otimes W.$$

The last isomorphism is something you should expect from the fact that R-linear maps into a module M are in bijection with M; i.e. $\operatorname{Hom}_R(R, M) \cong M$, via evaluating at 1. Thanks to Natalie Stewart for pointing out this argument to me.

Let's work this out for the case of an endomorphism T of a finite-dimensional vector space V. Let $\{e_i\}$ be a basis of V, and $\{e_i^*\}$ be the dual basis. We have the following:

$$\begin{array}{ccccccccc} k & \xrightarrow{\eta} & V^\vee \otimes_k V & \xrightarrow{\mathrm{id}\otimes f} & V^\vee \otimes_k V & = & V \otimes_k V^\vee & \xrightarrow{\epsilon} & k \\ 1 & \longmapsto & \sum_i e_i^* \otimes e_i & \longmapsto & \sum_i e_i^* \otimes f(e_i) & \longmapsto & \sum_i f(e_i) \otimes e_i^* & \longmapsto & \sum_i e_i^*(f(e_i)) \end{array}$$

So, $\mathrm{tr}(f) = \sum_i e_i^*(f(e_i))$. What is this? Well, $f(e_i)$ is the image of the i-th basis vector, and taking e_i^* of this pulls out the i-th component of the image of the i-th basis vector. This is effectively how much f scales e_i along it's span, and if we were to write f as a matrix with respect to $\{e_i\}$, this would be the i-th diagonal entry. Summing over these is precisely the usual definition of the trace! You can also check that this is what you get via the more pedestrian definition using the isomorphism $V^\vee \otimes_k W \cong \mathrm{Hom}(V, W)$; the map back gets you to the third thing in this diagram, and then you apply the evaluation map (which in this context is often called contraction).

With that, here's the coolest and perhaps most useful definition of the dimension of a vector space you'll ever see.

Definition 5.2. Let V be a dualizable k-vector space. Then $\dim_k V := \mathrm{tr}(\mathrm{id}_V)$.

Acknowledgments

Thanks to Peter Luo for helping me organize this, and to Lev Kruglyak and Reuben Stern for some helpful thoughts on the pedagogy of tensor products.

References

[1] M.F. Atiyah, I.G. MacDonald. Introduction to Commutative Algebra. Addison-Wesley Publishing Company. 1969.

[2] Brian Conrad. Tensor algebras, tensor pairings, and duality. http://virtualmath1.stanford.edu/ conrad/diffgeomPage/handouts/tensor.pdf

[3] Keith Conrad. Complexification. https://kconrad.math.uconn.edu/blurbs/linmultialg/complexification.

[4] Keith Conrad. Tensor Products. https://kconrad.math.uconn.edu/blurbs/linmultialg/tensorprod.pdf

[5] William Fulton, Joe Harris. Representation Theory: A First Course. Springer. 2004.

[6] Joe Harris. Tensor Products and Multilinear Algebra.

[7] Keeley Hoek. Crash course: constructions in point-set topology and linear algebra. https://hoek.io/res/2022.vbkt/warmup-crash-course-constructions.pdf

[8] Claudiu Raicu. Koszul Modules: Appendix. http://www.math.utah.edu/ca/Raicu_appendix.pdf

[9] Emily Riehl. Category Theory in Context. Dover Mathematics. 2016.

[10] Jonathan M. Rosenberg. Problems on Tensor Products.

[11] https://ncatlab.org/nlab/show/dualizable+object

[12] https://math.stackexchange.com/questions/1988633/an-isomorphism-concerned-about-any-finitely-generated-projective-module

[13] https://math.stackexchange.com/questions/81640/proving-that-the-tensor-product-is-right-exact

[14] https://math.stackexchange.com/questions/108475/question-about-perfect-pairings

[15] https://en.wikipedia.org/wiki/Tensor_product

[16] https://en.wikipedia.org/wiki/Tensor_algebra

[17] https://en.wikipedia.org/wiki/Symmetric_algebra

Email address: rushil_mallarapu@college.harvard.edu