THE TENSOR PRODUCT OF VECTOR SPACES

TOM COATES

Abstract. My presentation of tensor products in class today was extremely confusing. Here is a (hopefully) clearer version.

1. Motivation

One of the most powerful ideas in 20th-century mathematics — an idea which will come up in a lot of classes as you study more math — is that one can study the geometry of a space X (which could be a metric space, or a topological space, or a manifold, or $\ldots$) by studying the functions on X. For example, one could study the space

$$[0,1]$$

by studying the vector space[1] of bounded continuous functions

$$\mathcal{C}([0,1]).$$

This suggests a question: we know how to take the product $X \times Y$ of two spaces, but how is the vector space of functions on $X \times Y$ related to the vector spaces of functions on X and functions on Y. The tensor product is the answer to this question: roughly speaking, we will define the tensor product of two vector spaces so that

$$Functions(X \times Y) = Functions(X) \otimes Functions(Y).$$

The "roughly speaking" in the last sentence is because this statement will be true only for X and Y finite sets.

Exercise. Once you have read this note, read through it again and work out why the statement isn't true for infinite sets. What happens if we replace *Functions* by *Functions-which-are-non-zero-at-only-finitely-many-points*? Or by *Continuous-functions*?

One can define the tensor product of vector spaces in a number of different ways — Halmos uses a different definition, for example. His definition is significantly simpler than the one we are about to develop. The reason that I want to use this definition is that it works in a very general setting: the same construction gives the tensor product of infinite-dimensional

Date: February 17, 2005.

[1]This is slightly misleading: one should study $\mathcal{C}([0,1])$ not as a vector space but as an *algebra*. This means that we should think of $\mathcal{C}([0,1])$ as a vector space equipped with a multiplication map

$$\begin{array}{rcl} \mathcal{C}([0,1]) \times \mathcal{C}([0,1]) & \longrightarrow & \mathcal{C}([0,1]) \\ (f(t), g(t)) & \longmapsto & f(t)g(t). \end{array}$$

But the basic point remains: one can study $[0,1]$ by looking at $\mathcal{C}([0,1])$.

vector spaces, the tensor product of modules over a ring (once one knows what modules and rings are), etc.

2. Construction

From now on, think about two finite dimensional vector spaces V and W. We will regard V as the vector space of functions on some finite set S, and W as the vector space of functions on some finite set T.

Example. In all our examples, we will take

$$\begin{aligned} S &= \{1,2,3,4\} \\ T &= \{1,2,3\} \end{aligned}$$

so we can think of $S \times T$ as a 4×3 grid

$$\begin{array}{lcccc} T & & & & \\ 1 & \bullet & \bullet & \bullet & \bullet \\ 2 & \bullet & \bullet & \bullet & \bullet \\ 3 & \bullet & \bullet & \bullet & \bullet \\ & & & & \\ S & 1 & 2 & 3 & 4 \end{array}$$

2.1. Functions on the product $S \times T$.

Given a function $f(s)$ on S and a function $g(t)$ on the set T, we can form a function

$$h(s,t) = f(s)g(t)$$

on the set $S \times T$.

Example. If

$$f(s) = \begin{cases} 1 & s = 2 \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad g(t) = \begin{cases} 1 & t = 1 \\ 0 & \text{otherwise} \end{cases}$$

then we get

$$h(s,t) = \begin{cases} 1 & (s,t) = (2,1) \\ 0 & \text{otherwise} \end{cases}$$

Every function on $S \times T$ can be written as a linear combination of functions which are 1 in exactly one place and 0 everywhere else, so we see that every function on $S \times T$ can be written as a linear combination of products of functions on S and functions on T:

$$H(s,t) = a^{\alpha\beta} f_\alpha(s) g_\beta(t) \tag{1}$$

for any function $H : S \times T \to k$ and some choice of scalars $a^{\alpha\beta}$ and functions $f_\alpha : S \to k$, $g_\beta : T \to k$.

Exercise. Show that

$$H(s,t) = \begin{cases} 1 & (s,t) = (2,1) \\ 5 & (s,t) = (1,3) \\ 0 & \text{otherwise} \end{cases}$$

is not the product of a function on S and a function on T. Express H as a linear combination of products, as in (1).

Since we are thinking about V as functions on S and W as functions on W, and we want $V \otimes W$ to be functions on $S \times T$, this suggests that elements of $V \otimes W$ should be built as linear combinations of pairs of elements (f,g), where $f \in V$ and $g \in W$. Here we think of the pair (f,g) as representing the function

$$(s,t) \longmapsto f(s)g(t)$$

on $S \times T$.

2.2. **The free vector space generated by $V \times W$.** The free vector space generated by $V \times W$ is a precise version of "all linear combinations of pairs of elements (f,g), where $f \in V$ and $g \in W$". It is defined to be the vector space over k with basis

$$\{\delta_{(f,g)} : (f,g) \in V \times W\}$$

So in other words, elements of the free vector space F generated by $V \times W$ have the form

$$\alpha_1 \delta_{(v_1,w_1)} + \ldots + \alpha_n \delta_{(v_n,w_n)}$$

for some n, some choice of scalars $\alpha_1, \ldots, \alpha_n$, and some choice of n distinct elements $(v_1,w_1), \ldots, (v_n,w_n) \in V \times W$.

Example. Let v, v' be distinct elements of V and w, w' be distinct elements of W. Then

$$\left(3\delta_{(v,w)} - \delta_{(v,w')}\right) + 6\left(\delta_{(v,w')} + \delta_{(v',w')}\right) = 3\delta_{(v,w)} + 5\delta_{(v,w')} + 6\delta_{(v',w')}$$

To find out why I am using the notation $\delta_{(v,w)}$, read the footnote[2]. Note that we could use this construction to make "the free vector space generated by the set P" for any set P — we never needed to use the fact that V and W are vector spaces.

[2]As Toly pointed out at the end of class, another way to think about the free vector space generated by $V \times W$ is as a vector space of functions

$$\{F : V \times W \to k \mid F \text{ is non-zero at only finitely many points of } V \times W\}.$$

A basis for this vector space is given by the "delta functions"

$$\begin{aligned} \delta_{(v,w)} : V \times W &\to k \\ (x,y) &\mapsto \begin{cases} 1 & \text{if } (x,y) = (v,w) \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

These match up with the basis elements $\delta_{(v,w)}$ used above.

2.3. **The subspace of relations** Z**.** To reiterate, we are thinking of elements $f \in V$ as functions on the set S and elements $g \in W$ as functions on the set T. So the free vector space F generated by $V \times W$ looks rather like the functions on $S \times T$, where we regard the element

$$\alpha_1 \delta_{(f_1,g_1)} + \ldots + \alpha_n \delta_{(f_n,g_n)} \in F$$

as representing the function on $S \times T$ which is

$$(s,t) \mapsto \alpha_1 f_1(s) g_1(t) + \ldots + \alpha_n f_n(s) g_n(t).$$

But F is not quite the functions on $S \times T$, because

$$(af_1(s) + bf_2(s))g(t) = af_1(s)g(t) + bf_2(s)g(t)$$

as functions on $S \times T$, but

$$\delta_{(af_1+bf_2,g)} \neq a\delta_{(f_1,g)} + b\delta_{(f_2,g)}$$

in F.

In other words, if we want to turn F into the vector space of functions on $S \times T$ then we need to impose the relations

$$\begin{aligned} \delta_{(af_1+bf_2,g)} &= a\delta_{(f_1,g)} + b\delta_{(f_2,g)} \\ \delta_{(f,ag_1+bg_2)} &= a\delta_{(f,g_1)} + b\delta_{(f,g_2)} \end{aligned}$$

But we know how to do this from class: we make a subspace Z of F which contains all the things that we want to be zero (*i.e.* all the relations that we want to hold)

$$Z = span \left\{ \delta_{(af_1+bf_2,g)} - a\delta_{(f_1,g)} - b\delta_{(f_2,g)}, \delta_{(f,ag_1+bg_2)} - a\delta_{(f,g_1)} - b\delta_{(f,g_2)} : a, b \in k, f, f_1, f_2 \in V, g, g_1, g_2 \in\right.$$

and then define

$$V \otimes W = F/Z$$

2.4. **Why this does exactly what we want.** Write

$$v \otimes w = \delta_{(v,w)} + Z$$

Then the fact that

$$\delta_{(af_1+bf_2,g)} - a\delta_{(f_1,g)} - b\delta_{(f_2,g)} \in Z$$

means that

$$\delta_{(af_1+bf_2,g)} + Z = a\delta_{(f_1,g)} + b\delta_{(f_2,g)} + Z$$

or in other words that

$$(af_1 + bf_2) \otimes g = a(f_1 \otimes g) + b(f_2 \otimes g). \tag{2}$$

Similarly,

$$f \otimes (ag_1 + bg_2) = a(f \otimes g_1) + b(f \otimes g_2). \tag{3}$$

2.5. **A basis for the tensor product.** At the end of class, I claimed that if $\{x_1, \ldots, x_n\}$ is a basis for V and $\{y_1, \ldots, y_m\}$ is a basis for W then

$$B = \{x_i \otimes y_j : 1 \le i \le n, 1 \le j \le m\}$$

is a basis for $V \otimes W$. To get some practice working with tensor products, let us first see why the set B spans $V \otimes W$. We need to take an arbitrary element of $V \otimes W$ and write it as a linear combination of elements of B. But, since $V \otimes W = F/Z$, we know that anything in $V \otimes W$ is of the form

$$\alpha_1 \delta_{(v_1, w_1)} + \ldots + \alpha_n \delta_{(v_n, w_n)} + Z.$$

Put another way, anything in $V \otimes W$ is of the form

$$\alpha_1 v_1 \otimes w_1 + \ldots + \alpha_n v_n \otimes w_n$$

But we can write[3]

$$\begin{aligned} v_i &= b_i^\gamma x_\gamma \\ w_i &= c_i^\epsilon y_\epsilon \end{aligned}$$

and so our element of $V \otimes W$ is

$$\alpha_1 (b_1^\gamma x_\gamma) \otimes (c_1^\epsilon y_\epsilon) + \ldots + \alpha_n (b_n^\gamma x_\gamma) \otimes (c_n^\epsilon y_\epsilon)$$

Now we can apply equations (2) and (3) repeatedly to get

$$\alpha_1 b_1^\gamma c_1^\epsilon (x_\gamma \otimes y_\epsilon) + \ldots + \alpha_n b_n^\gamma c_n^\epsilon (x_\gamma \otimes y_\epsilon)$$

This is a linear combination of the $x_i \otimes y_j$'s, so we're done.

In class on Friday, we will prove that the set B is LI.

[3]Summation convention!