More on Free Vector Space PhL 9.20.15

In my Notes on Tensor Products doc, I quote from about 4 sources on free vector spaces. But when I try to "write up" the concept, things just don't work right.

- Is it really f: S→R, reals ?
- What does an f(s) look like? Would $f(s') = \Sigma_s f_s(s')\, s = \Sigma_i\, f_i(s')\, s$ be in f(s)?
- If elements of S are not reals, then what is meant by the product $f_s(s')\, s$ where f_s is real?
- Must the elements of S be reals?
- Where does "basis" fit into this picture?

None of my sources really answers these questions, so I cannot do a presentation without a clearer view. My sources dance around some of these questions.

So here I will just blindly continue to look for new sources in hopes of finding one that answers these questions.

1. Roman on the subject of Free Modules

What about Roman? He talks about a "free module" on page 116. But he uses the term R-module, so I have to go look that up first, and that is on page 109.

Modules

Definition *Let R be a commutative ring with identity, whose elements are called* **scalars**. *An R-***module** (*or a* **module over** R) *is a nonempty set M, together with two operations. The first operation, called* **addition** *and denoted by $+$, assigns to each pair $(u, v) \in M \times M$, an element $u + v \in M$. The second operation, denoted by juxtaposition, assigns to each pair $(r, v) \in R \times M$, an element $rv \in M$. Furthermore, the following properties must hold:*

1) M is an abelian group under addition.

2) For all $r, s \in R$ and $u, v \in M$

$$r(u+v) = ru + rv$$
$$(r+s)u = ru + su$$
$$(rs)u = r(su)$$
$$1u = u$$

The ring R is called the **base ring** *of M.* □

First, the only things that distinguishes are ring from a field are:
possibly non-abelian, possible no identity, possible no inverses.

But in the above, Roman requires that his ring be abelian *and* have an identity. So the only missing ingredient is that R might not contain all its inverses, in which case R just misses being a field. The elements of his ring R are the "scalars" of this definition.

So what is this module M with elements m_i? Here are my comments

1. The ring R has operations called $+_R$ and $\bullet_R$. For example, $r_1 +_R r_2$ lies within R.

2. Addition of elements of M is defined, so that $m_1 +_M m_2$ lies within M, and this is abelian addition.

3. You are allowed to multiply an R element by a M element, and the product r*m, lies in M. Formally, R * M $\subset$ M. You always write the R element first.

4. The * multiplication by a ring element r is distributive over module addition:

$$r*(m_1 +_M m_2) = r*m_1 +_M r*m_2$$

5. The * multiplication by a module element m is distributive over ring addition in this way

$$(r_1 +_R r_2)*m = r_1*m +_M r_2*m$$

6. Elements of the form RRM are associative in this sense

$$(r_1 \bullet r_2)*m = r_1*(r_2*m)$$

What is all this really doing? You start with a ring and its $+_R$ and $\bullet_R$ operations. You then postulate a cross multiplication operation between R and M. You also postulate an addition operation $+_M$ for M. There is no definition of any kind of $\bullet_M$ operation between elements of M. So this R-module M is a fancier object that is in some sense "built on to of R". I guess as a simple case you could have M = R and then the + and • operations would be the same, AND you would then be able to multiply elements of M, though that is not part of the module definition.

Now that we know all about R modules, consider his definition of a free module

Free Modules

The definition of a basis for a module parallels that of a basis for a vector space.

Definition *Let M be an R-module. A subset $\mathcal{B}$ of M is a* **basis** *if $\mathcal{B}$ is linearly independent and spans M. An R-module M is said to be* **free** *if $M = \{0\}$ or if M has a basis. If $\mathcal{B}$ is a basis for M, we say that M is* **free on** *$\mathcal{B}$.* □

This seems to say the following:

A free module is an R-module M which has a basis, which we shall call B. M is then "free on B".
(If M has no elements, then it is also regarded as free. This would certainly be a trivial module.)

This seems a far cry from the definitions of "free vector space" that I quoted from random web locations! Some of those made no mention of a basis!

Example: Consider the ring of integers Z which is a viable R for the above: commutes and has identity. Then you could talk about a Z-module.

Example: Consider R = ZxZ with elements (m,n). This is a ring where $(m,n)\bullet_R(i,j) = (mi, nj)$ and $(m,n) +_R (i,j) = (m+i,n+j)$. You can write *any* element as $(m,n) = (m,n)\bullet_R(1,1)$, so you can regard the single element $\{(1,1)\}$ as forming a basis (I think this is what he later claims). Maybe ZxZ is a direct product ring, he does not use that term.

Roman continues with no gap

We have the following analog of part of Theorem 1.7.

Theorem 4.3 *A subset $\mathcal{B}$ of a module M is a basis if and only if every nonzero $v \in M$ is an essentially unique linear combination of the vectors in $\mathcal{B}$.*□

This certainly seems reasonable to me. He goes on

In a vector space, a set of vectors is a basis if and only if it is a minimal spanning set, or equivalently, a maximal linearly independent set. For modules, the following is the best we can do in general. We leave proof to the reader.

Theorem 4.4 *Let $\mathcal{B}$ be a basis for an R-module M. Then*
1) $\mathcal{B}$ is a minimal spanning set.
2) $\mathcal{B}$ is a maximal linearly independent set.□

Again, seems very reasonable. He is just laying these things out.

He then shows that R = ZxZ provides an example of a module with a basis, but it has a submodule which has no basis, and so modules and vector spaces have different "behavior". Continuing after this:

Theorem 4.5 *Let M and N be R-modules where M is free with basis $\mathcal{B} = \{b_i \mid i \in I\}$. Then we can define a unique R-map $\tau: M \to N$ by specifying the values of τb_i arbitrarily for all $b_i \in \mathcal{B}$ and then extending τ to M by linearity, that is,*

$$\tau(a_1 v_1 + \cdots + a_n v_n) = a_1 \tau v_1 + \cdots + a_n \tau v_n \qquad \square$$

Here we have two R-modules M and N, where M is "free" and so has a basis $\{b_i\}$. He then defines a mapping τ: M→N in which he first defines $\tau(b_i)$ in some arbitrary fashion, where $\tau(b_i) \in N$. Then he does the trick I saw elsewhere of simply "declaring" by fiat that τ is linear (extending τ) so then you could say that $\tau(v) = \tau(\Sigma_i v_i b_i) = \Sigma_i v_i \tau(b_i)$ and then you have τ defined on all of M, and then in general this linearity says that $\tau(a_1 v_1 + a_2 v_2) = a_1\tau(v_1) + a_2\tau(v_2)$. The term "R-map" is not defined anywhere. I guess it is a map between two R-modules.

Roman has other things to say about free modules. On page 131 for example,

Theorem 5.2 *Let B be any set and let R be a commutative ring with identity. The set $(R^B)_0$ of all functions from B to R that have finite support is a free R-module of rank $|B|$ with basis $\mathcal{B} = \{\delta_b\}$ where*

$$\delta_b(x) = \begin{cases} 1 & \text{if } x = b \\ 0 & \text{if } x \neq b \end{cases}$$

This basis is referred to as the **standard basis** for $(R^B)_0$.□

So here we assume some ring R, and B is "any set" . We can consider f such that f: B → R (R = ring, not reals!!) where we restrict f so that is must have "finite support" -- I think this means as elsewhere that f is a function of only a finite number of elements in the set B. This set of functions {f} is given the name $(R^B)_0$ which no doubt has some later extension. The elements of set S are called b or x. He then defines a set of functions as shown, whose values are 0 or 1. Notice that 0 and 1 are always elements of a ring R which has an identity. He claims these functions $\delta_b(x)$ form a basis for a free R-module M, though he does not use the letter M. A free module must have a basis, so here B = $\{\delta_b\}$. In fact M = $(R^B)_0$ which is the set of functions f:B→R with finite support.

So why is this set of $\delta_b(x)$ a basis? Well I guess for any f in $(R^B)_0$ you can say

$f(x) = \Sigma_b\, f(b)\, \delta_b(x) = \Sigma_b\, f(b)\, \delta_{b,x} = f(x)$ // f_b and $\delta_b(x)$ and f(x) are all elements of R

The other author used the word Dirac, and perhaps the above is like

$f(x) = \int dy\, f(y)\, \delta(y-x) = f(x)$

Comment: This discussion sounds a lot more like the free vector space discussions I quoted in my Notes on Tensor Products doc. If R is a field, this would be defining a "free vector space" as a set of functions which map f: B→F where F is that field. Notice in the above equation that f(b), $\delta_b(x)$ and f(x) would all be elements of this field F, and we never have to talk about whether or not b is an element of F.

Question: is the function f(b) = b included in Roman's R-module $(R^B)_0$? If so, then you would have to have b ϵ F. But this subject never arises at this point in Roman.

Thinking of B as a set, the function f merely has to assign a real number to each element b of B. That is exactly what the above does

$f(x) = \Sigma_b\, f_b\, \delta_b(x)$ for all x in B

The real number assigned to b is called f_b or f(b). You don't have to have the b be real in order to assign them to real numbers. They don't have to be anything! They are just set elements.

So I guess f(b) = b does not really make sense because such an f is not an element of F. So I guess strictly you would rule out such a function. You could on the other hand say $f(b_i) = i$ where the ring R is now Z.

So this does render suspect my author who write f(b) as linear combinations of the b.

Let's now review all my previous free vector space section notes.

4.1 wiki. This quote explicitly says that F(S) "has a basis", so I sort of missed that fact. I see now that this fact is Roman's main idea for a free thing. Wiki in fact quotes this basis as exactly the $\delta_b(x)$ discussed above. Wiki ends with the statement "F(S) is often called a formal sum of symbols in S" . So they are calling the elements of the set S "symbols". And the sum above does not exhibit a bare "b" in the sum.

Note that $f(x) = \delta_b(x)$ is a function which maps b to 1, and all other elements of B to 0, so it is a selector. My idea of $b = e_i$ now seems a little far-fetched but it might be viable. But I would not then have a raw e_i appear in any sum, just as b does not appear above in raw form. A set function is a bit different from an ordinary function I guess.

4.2 Drexel. This guy writes F[X] where X is a finite set, and he DOES write sums in which x is exposed in a raw sense. Drexel's F[X] is a set of sums, not a set of functions, so he does not have any f: X→ T. He never mentions that. He does not mention "basis" at all.

However, Drexel in fact has a lot to say about "free modules" which I did not notice earlier. One point he makes is that not all R-modules have a basis, whereas all vector spaces do, so modules are different. The free module is an R-module which DOES have a basis. In fact Drexel manages to connect the "free module world" to the "triangle world" as in the following section taken from web page here

https://drexel28.wordpress.com/2011/11/16/free-modules-pt-i/

Let S be a set and R a unital ring. The ordered pair (M, i) is a *free left R-module over S* if M is a left R-module and $i : S \to M$ is a set map which satisfies the following universal property: for any left R-module N and and set map $f : S \to N$ there exists a unique R-map $g : M \to N$ such that the following diagram commutes

$$\begin{array}{ccc} S & \xrightarrow{i} & M \\ & {\scriptstyle f}\searrow & \downarrow{\scriptstyle g} \\ & & N \end{array}$$

We say that an R-module F is *free* if there exists some free module (M, i) over a set S with $F \cong M$.

I guess a "unital ring" is a ring with an identity, but web suggests a more general definition.

Look at the above triangle where S is a set and i:S→M is a "set map" and M is an R-module. He gives this R-module the name (M,i) and calls it an "R-module over S". Clearly S and R are different from the first sentence. So we really have here an object which is "a module over the ring R and over the set S" or "an R-module over the set S". The connection to the set S is that there is some mapping i which maps S into M.

Drexel then says that the R-module M is a *free* R-module IF the following is true: for any *other* R-module N over S, (N,f) indicated by f: S→N, one can find a *unique* mapping g (an R-map since it maps between two R-modules) which links M to N such that the triangle commutes. This IF clause he calls "the

universal property", similar to my other authors. So in the end he says an R-module F on S is *free* if you can just find *some* free module i:S→M such that F is isomorphic to M in the sense of the above commuting picture (unique g). Note that the symbol ≅ is traditionally used to indicate an isomorphism, though sometimes ≈ is used instead. Note also that he does not say that g is linear, but I suspect it can always be made to be linear.

Well, I am not sure this helps me much. He goes on with other fancy related topics. He gives some references:

References:

[1] Dummit, David Steven., and Richard M. Foote. *Abstract Algebra*. Hoboken, NJ: Wiley, 2004. Print.

[2] Rotman, Joseph J. *Advanced Modern Algebra*. Providence, RI: American Mathematical Society, 2010. Print.

[3] Blyth, T. S. *Module Theory.* Clarendon, 1990. Print.

So the key idea is that a free module is one with a basis. This key fact seems irrelevant in the parallel topic of a free vector space, since any vector space has a basis. So free vector space must really be a different animal. I guess I will try to zero in on people talking about such things rather than free modules. It was perhaps good to do a little tour in the module area. In my tensor product definition doc I am now trying to write, I guess I will not mention Drexel's linear sum approach. I do, however, have to say something about a basis for the vector space and right now I have no info really on this subject. Well, I have an opening which so far reads this way:

1.1 Tensor Product as a Quotient Space.

It does seem odd that one might think of a product V⊗W in terms of a quotient. We shall outline how this path goes in a set of steps, omitting technical details or relegating them to "notes".

1. If V and W are two sets (which could be vector spaces), then VxW is called the **Cartesian product** of the two sets. The elements of this Cartesian product set have the form (v,w) where v ϵ V and w ϵ W. No structure of any kind is implied. For example, (v,w)+(v',w') and α(v,w) are undefined. The object VxW is *not* a tensor product of V and W.

2. Holding that idea for a moment, one can define F(S) to be the set of all functions f which map a set S into the reals. The elements of F(S) are functions f where f : S → R. One can declare that within F(S), one can add two functions so that f(s) + g(s) = h(s) = (f+g)(s) and one can scale a function by a real so that (αf)(s) = α f(s), where here (αf) is a function in F(S) different from f. The set of functions F(S) is called a **free vector space**.

Each function f in F(S) just assigns a real number to each point in the set S. One can define a basis for F(S) as the set of basis functions $\delta_s(x) = 1$ if x = s, otherwise 0. Then the most general f in F(S) has the form

$f(x) = \Sigma_s f_s \delta_s(x)$ for all x in S

where f_s is the real number assigned to point s in S. If the set S is infinite (like Z, the integers) this sum must only have a finite number of terms for which $f_s \neq 0$.

The above construction works for any field K, not just the reals R. Note that the values of $\delta_s(x)$, 0 and 1, belong to any field K.

What happens when I try to set the set S equal to VxW, the product of two vector spaces? Read it off.

The free vector space F(VxW) consists of maps f: (v,w) → R. If V has a basis e_i and W a basis e'_j, we can represent a most general element f of F(VxW) as

$f = \Sigma_{ij} f_{ij} (e_i \times e_j)$

STOP. This is supposed to be a real number, but I don't see that $(e_i \times e_j)$ is a real number. Maybe back up to a simpler case.

Suppose S = V, a vector space with some basis e_i. This set S is continuous with an infinite number of elements. How do I apply the free vector space idea from above? The elements of S are

$s = \Sigma_i f_i e_i$

We are supposed to assign each of these points s to a real number. Can I still say

$\delta_s(x) = 1$ if x = s, otherwise 0 ?

But now s and x are vectors in V! I guess I could say

$\delta_s(\mathbf{x}) = \delta^{(n)}(\mathbf{x}-\mathbf{s})$ = a Dirac n-fold delta function.

But this takes values 0 and "undefined", so does not quite work right. Then you would write $f(x) = \Sigma_s f_s \delta_s(x)$ in this way

$f(x) = \int d^n s \; f_s \; \delta^{(n)}(\mathbf{x}-\mathbf{s})$

My conclusion is that my presentation of a free vector space given above for set S does not apply very well to the case that S is a vector space! If I can't make F(V) fly, I don't have a chance for F(VxW) where we then have *two* vector spaces!

Suppose I try to "roll my own" without using the official free vector space concept. I start over:

Let F(V) be the set of all functions f : V → V so that $v_2 = f(v_1)$.

But why do I even talk about a space of functions. Start over

1. Start with the inert Cartesian space VxW of elements (v,w).

2. Endow VxW with the ability to add linear combinations of elements with scalar coefficients. Let's define F(VxW) to be a space which contains any such linear combinations. A typical element of F(VxW) space might be $3(v_1,w_3) - 2(v_2,w_5)$. Again, such a linear combination is an element of the space F(VxW). Of course (v_1,w_3) also lies in F(VxW) and one might call this a pure element, whereas $3(v_1,w_3) - 2(v_2,w_5)$ is a mixed element.

3. Now define the following set of "equivalence relations"

$(v_1+v_2, w) \sim (v_1,w) + (v_2,w)$	for all $v_1,v_2 \in V$ and all $w \in W$
$(v, w_1+w_2) \sim (v,w_1) + (v,w_2)$	for all $v \in V$ and all $w_1, w_2 \in W$
$\alpha(v,w) \sim (\alpha v,w)$	for all $v \in V$ and all $w \in W$ and all $\alpha \in R$
$\alpha(v,w) \sim (v,\alpha w)$	for all $v \in V$ and all $w \in W$ and all $\alpha \in R$

where ~ means "is equivalent to". Rewrite these as

$(v_1+v_2, w) - (v_1,w) + (v_2,w) \sim 0$
$(v, w_1+w_2) - (v,w_1) + (v,w_2) \sim 0$
$\alpha(v,w) - (\alpha v,w) \sim 0$
$\alpha(v,w) - (v,\alpha w) \sim 0$

We are declaring here that lots of linear combinations are equivalent to 0. These functions taken together define an "equivalence class" which is equivalent to 0. Call this class N (for null). There then exists a space which we shall call F(VxW)/N or F(VxW) "mod" N. This is a standard structure in equivalence class theory where one takes the quotient of one space S divided by another space of equivalent items in space S, often written S/~. The upshot is that the elements of the new quotient space F(VxW)/N consist of all linear combinations of F(VxW) *except* any linear combination which has one of the four forms shown above is filtered out ("modded out") by setting it equal to 0.

Example: $3(w_3,v_4) + (v_2, w_1+w_2) - (v_2,w_1) + (v,w_2)$ = an element of F(VxW)
$3(w_3,v_4)$ = the corresponding element of F(VxW)/N.

4. We now give this space F(VxW)/N a new name:

F(VxW)/N = V⊗W = the tensor product space of V and W

The elements of V⊗W are linear combinations of elements called v⊗w instead of (v,w) as a reminder that the equivalence class N must be respected. The net effect of all this song and dance is the following:

The tensor product space V⊗W is the set of all linear combinations of elements (v,w) of the Cartesian product space VxW, written as v⊗w, where the following rules are *declared by fiat*:

$(v_1+v_2) \otimes w = (v_1 \otimes w) + (v_2 \otimes w)$	for all $v_1,v_2 \in V$ and all $w \in W$

$v\otimes (w_1+w_2) = (v\otimes w_1) + (v\otimes w_2)$ for all $v \in V$ and all $w_1, w_2 \in W$

$\alpha(v\otimes w) = (\alpha v)\otimes w$ for all $v \in V$ and all $w \in W$ and all $\alpha \in R$

$\alpha(v\otimes w) = v\otimes(\alpha w)$ for all $v \in V$ and all $w \in W$ and all $\alpha \in R$

If these rules were declared for a *function* f(v,w), they would appear as

$f(v_1+v_2,w) = f(v_1,w) + f(v_2,w)$

$f(v,w_1+w_2) = f(v,w_1) + f(v,w_1)$

$\alpha\ f(v,w) = f(\alpha v,w)$

$\alpha\ f(v,w) = f(v,\alpha w)$

Such a function would then be described as being "**bilinear**" because it is linear separately in each argument. One can then regard the rules shown above for ⊗ as expressing bilinearity for the tensor product space V⊗W.

Notes:

- One can assume the scalars α are elements of any field K, not just the field of reals R.
- Sometimes the space F(VxW) is described as a "free vector space". Usually such spaces are defined over a discrete set S of "objects", and it is not clear how this works when the set S is continuous, this being the case for S = VxW where V and W are entire vector spaces. For that reason, we did not use the term free vector space in the above, but the term is freely used by other sources.

5. The above development is easily generalized to the tensor product of any finite number of vector spaces. One first defines F(V,W,....Z) as linear combinations of elements of the Cartesian product space VxWx..xZ , which elements have the form (v,x,...z). One then defines a large set of equivalence relations analogous to those described above, but for each variable. One then ends up with a large set of linear combinations which are all equivalent to 0, and this defines the equivalence class N. One then creates F(V,W,....Z)/N as the space of linear combinations where any pieces which are equivalent to 0 are filtered out. One then defines

$F(VxWx...Z)/N = V\otimes W\otimes...\otimes Z$ = the tensor product of spaces V and W and... and Z.

The upshot is now that the tensor product space V⊗W⊗...⊗Z is the set of all linear combinations of elements (v,w,...z) of the Cartesian product space VxWx...xZ, written as v⊗w⊗...⊗z, where the following rules are *declared by fiat*:

$(v_1+v_2)\otimes w\otimes .... \otimes z = v_1\otimes w\otimes .... \otimes z + v_2\otimes w\otimes .... \otimes z$

$v\otimes(w_1+w_2)\otimes .... \otimes z = v\otimes w_1\otimes .... \otimes z + v\otimes w_2\otimes .... \otimes z$

and so on for all positions in the ⊗ product. In addition,

$\alpha(v\otimes w\otimes...\otimes z) = (\alpha v)\otimes w\otimes...\otimes z$

$\alpha(v\otimes w\otimes...\otimes z) = v\otimes(\alpha w)\otimes...\otimes z$

When these rules are written for a function f(v,w,....z) one has,

$f(v_1+v_2,w,...z) = f(v_1,w,...z) + f(v_2,w,...z)$
$f(v,w_1+w_2,...z) = f(v,w_1,...z) + f(v,w_2,...z)$
etc

$\alpha\, f(v,w,...z) = f(\alpha v,w...z) = f(v,\alpha w,...) =$ etc.

If there are k factors in the tensor product $V\otimes W\otimes ...\otimes Z$, then the function f has k arguments, and a function obeying all of the above rules is said to be **k-multilinear**. For k = 2 we have bilinear, for k = 3 we have trilinear, and so on. One can mix in the scalar rule by saying for example

$f(\alpha v_1+\beta v_2,w,...z) = \alpha f(v_1,w,...z) + \beta f(v_2,w,...z)$ and similarly for all positions

but we have kept the scalar rules separate. Either statement of the rules is equivalent.

We can then regard the set of $\otimes$ rules shown above as describing k-multilinearity for the tensor product space $V\otimes W\otimes ...\otimes Z$. Written in the second form we would say

$(\alpha v_1+\beta v_2)\otimes w\otimes .... \otimes z = \alpha\,(v_1\otimes w\otimes .... \otimes z) + \beta\,(v_2\otimes w\otimes .... \otimes z)$

and similarly for all positions.

Question 1: What would a basis be for F(VxW) which I say has elements like $3(v_1,w_3) - 2(v_2,w_5)$? At this point in the discussion, we have no "rules" for simplifying such an expression. For example suppose it happened that all the elements shown were basis elements. You then have $3(e_1,e'_3) - 2(e_2,e'_5)$ but we have no way to write this as $\Sigma_{ij}F_{ij}(e_i, e'_j)$ because

Where would I inject the idea that (v,w) + (v',w') = (v+v',w+w') ? I think this is wrong!

Question 2: If $V\otimes W$ is a vector space, what is the inverse of $v\otimes w$? How about $(-1)(v\otimes w)$.

Question 3: What is the additive identity? Need something so that $0 + v\otimes w = v\otimes w$. I don't know what to do here! If I try $0\otimes 0$ then I don't know how to show that $0\otimes 0 + v\otimes w = v\otimes w$. We need "0" element in order to have a $V\otimes W$ be a vector space! I do know that $0\otimes w$ would act as a zero element for $v\otimes w$, but $0\otimes w$ only works for elements of the form $v\otimes w$, not $v\otimes w'$. Ouch. I keep finding problems everywhere I look! I will have to go ask the world for help on this question. I thought $V\otimes W$ was a vector space, am I wrong about that!