TENSOR PRODUCTS

KEITH CONRAD

1. Introduction

Let R be a commutative ring and M and N be R-modules. (We always work with rings having a multiplicative identity and modules are assumed to be unital: $1 \cdot m = m$ for all $m \in M$.) The direct sum $M \oplus N$ is an addition operation on modules. We introduce here a product operation $M \otimes_R N$, called the tensor product. We will start off by describing what a tensor product of modules is supposed to look like. Rigorous definitions are in Section 3.

Tensor products first arose for vector spaces, and this is the only setting where tensor products occur in physics and engineering, so we'll describe the tensor product of vector spaces first. Let V and W be vector spaces over a field K, and choose bases $\{e_i\}$ for V and $\{f_j\}$ for W. The tensor product $V \otimes_K W$ is defined to be the K-vector space with a basis of formal symbols $e_i \otimes f_j$ (we declare these new symbols to be linearly independent by definition). Thus $V \otimes_K W$ is the formal sums $\sum_{i,j} c_{ij} e_i \otimes f_j$ with $c_{ij} \in K$, which are called tensors. Moreover, for all $v \in V$ and $w \in W$ we define $v \otimes w$ to be the element of $V \otimes_K W$ obtained by writing v and w in terms of the original bases of V and W and then expanding out $v \otimes w$ as if $\otimes$ were a noncommutative product (allowing scalars to be pulled out).

For example, let $V = W = \mathbf{R}^2 = \mathbf{R}e_1 + \mathbf{R}e_2$, where $\{e_1, e_2\}$ is the standard basis. (We use the same basis for both copies of $\mathbf{R}^2$.) Then $\mathbf{R}^2 \otimes_{\mathbf{R}} \mathbf{R}^2$ is a 4-dimensional space with basis $e_1 \otimes e_1$, $e_1 \otimes e_2$, $e_2 \otimes e_1$, and $e_2 \otimes e_2$. If $v = e_1 - e_2$ and $w = e_1 + 2e_2$, then

$$v \otimes w = (e_1 - e_2) \otimes (e_1 + 2e_2) := e_1 \otimes e_1 + 2e_1 \otimes e_2 - e_2 \otimes e_1 - 2e_2 \otimes e_2. \tag{1.1}$$

Does $v \otimes w$ depend on the choice of a basis of $\mathbf{R}^2$? As a test, pick another basis, say $e_1' = e_1 + e_2$ and $e_2' = 2e_1 - e_2$. Then v and w can be written as $v = -\frac{1}{3}e_1' + \frac{2}{3}e_2'$ and $w = \frac{5}{3}e_1' - \frac{1}{3}e_2'$. By a formal calculation,

$$v \otimes w = \left(-\frac{1}{3}e_1' + \frac{2}{3}e_2'\right) \otimes \left(\frac{5}{3}e_1' - \frac{1}{3}e_2'\right) = -\frac{5}{9}e_1' \otimes e_1' + \frac{1}{9}e_1' \otimes e_2' + \frac{10}{9}e_2' \otimes e_1' - \frac{2}{9}e_2' \otimes e_2',$$

and if you substitute into this last linear combination the definitions of e_1' and e_2' in terms of e_1 and e_2, expand everything out, and collect like terms, you'll return to the sum on the right side of (1.1). This suggests that $v \otimes w$ has a meaning in $\mathbf{R}^2 \otimes_{\mathbf{R}} \mathbf{R}^2$ that is independent of the choice of a basis, although proving that might look daunting.

In the setting of modules, a tensor product can be described like the case of vector spaces, but the properties that $\otimes$ is supposed to satisfy have to be laid out in general, not just on a basis (which may not even exist): for R-modules M and N, their tensor product $M \otimes_R N$ (read as "M tensor N" or "M tensor N over R") is an R-module spanned – not as a basis, but just as a spanning set[1] – by all symbols $m \otimes n$, with $m \in M$ and $n \in N$, and these

[1] Recall a spanning set for an R-module is a subset whose *finite* R-linear combinations fill up the module. They always exist, since the entire module is a spanning set.

symbols satisfy distributive laws:

$$(m+m')\otimes n = m\otimes n + m'\otimes n, \quad m\otimes(n+n') = m\otimes n + m\otimes n'. \tag{1.2}$$

Also multiplication by $r \in R$ can be put into either side of $\otimes$: for $m \in M$ and $n \in N$,

$$r(m\otimes n) = (rm)\otimes n = m\otimes(rn). \tag{1.3}$$

(The notation $rm \otimes n$ is unambiguous: it is both $r(m\otimes n)$ and $(rm)\otimes n$.)

The formulas (1.2) and (1.3) in $M \otimes_R N$ should be contrasted with those for the direct sum $M \oplus N$, where

$$(m+m', n) = (m,n) + (m',0), \quad r(m,n) = (rm, rn).$$

In $M \oplus N$, an element (m,n) decomposes as $(m,0)+(0,n)$, but $m\otimes n$ in $M\otimes_R N$ does not break apart. While every element of $M \oplus N$ is a pair (m,n), there are usually *more* elements of $M \otimes_R N$ than the products $m \otimes n$. The general element of $M \otimes_R N$, which is called a tensor, is an R-linear combination[2]

$$r_1(m_1\otimes n_1) + r_2(m_2\otimes n_2) + \cdots + r_k(m_k\otimes n_k),$$

where $k \geq 1$, $r_i \in R$, $m_i \in M$, and $n_i \in N$. Since $r_i(m_i\otimes n_i) = (r_im_i)\otimes n_i$, we can rename r_im_i as m_i and write the above linear combination as a sum

$$m_1\otimes n_1 + m_2\otimes n_2 + \cdots + m_k\otimes n_k. \tag{1.4}$$

In the direct sum $M \oplus N$, equality is easy to define: $(m,n) = (m',n')$ if and only if $m = m'$ and $n = n'$. When are two sums of the form (1.4) equal in $M \otimes_R N$? This is not easy to say in terms of the description of a tensor product that we have given, except in one case: M and N are free R-modules with bases $\{e_i\}$ and $\{f_j\}$. In this case, $M \otimes_R N$ is free with basis $\{e_i \otimes f_j\}$, so every element of $M\otimes_R N$ is a (finite) sum $\sum_{i,j} c_{ij}e_i\otimes f_j$ with $c_{ij} \in R$ and two such sums are equal only when coefficients of like terms are equal.

To describe equality in $M \otimes_R N$ when M and N don't have bases, we will use a universal mapping property of the tensor product. In fact, the tensor product is the first concept in algebra whose applications in math make consistent sense *only* through a universal mapping property, which is: $M \otimes_R N$ is the universal object that turns bilinear maps on $M \times N$ into linear maps. What that means will become clearer later.

After a discussion of bilinear (and multilinear) maps in Section 2, the definition and construction of the tensor product is presented in Section 3. Examples of tensor products are in Section 4. In Section 5 we will show how the tensor product interacts with some other constructions on modules. Section 6 describes the important operation of base extension, which is a process of using tensor products to turn an R-module into an S-module where S is another ring. Finally, in Section 7 we describe the notation used for tensors in physics.

Here is a brief history of tensors and tensor products. Tensor comes from the Latin *tendere*, which means "to stretch." In 1822 Cauchy introduced the Cauchy stress tensor in continuum mechanics, and in 1861 Riemann created the Riemann curvature tensor in geometry, but they did not use those names. In 1884, Gibbs [5, Chap. 3] introduced tensor products of vectors in $\mathbf{R}^3$ with the label "indeterminate product"[3] and applied it to study

[2]Compare with the polynomial ring $R[X,Y]$, whose elements are not only products $f(X)g(Y)$, but sums of such products $\sum_{i,j} a_{ij}X^iY^j$. It turns out that $R[X,Y] \cong R[X]\otimes_R R[Y]$ as R-modules (Example 4.12).

[3]The label indeterminate was chosen because Gibbs considered this product to be, in his words, "the most general form of product of two vectors," as it is subject to no laws except bilinearity, which must be satisfied by any operation on vectors that deserves to be called a product.

strain on a body. He extended the indeterminate product to n dimensions in 1886 [6]. Voigt used tensors to describe stress and strain on crystals in 1898 [16], and the term tensor first appeared with its modern physical meaning there.[4] In geometry Ricci used tensors in the late 1800s and his 1901 paper [14] with Levi-Civita (in English in [9]) was crucial in Einstein's work on general relativity. Wide use of the term "tensor" in physics and math is due to Einstein; Ricci and Levi-Civita called tensors by the bland name "systems". The notation $\otimes$ is due to Murray and von Neumann in 1936 [11, Chap. II] for tensor products (they wrote "direct products") of Hilbert spaces.[5] The tensor product of abelian groups A and B, with that name but written as $A \circ B$ instead of $A \otimes_{\mathbf{Z}} B$, is due to Whitney [18] in 1938. Tensor products of modules over a commutative ring are due to Bourbaki [2] in 1948.

2. Bilinear Maps

We already described the elements of $M \otimes_R N$ as sums (1.4) subject to the rules (1.2) and (1.3). The intention is that $M \otimes_R N$ is the "freest" object satisfying (1.2) and (1.3). The essence of (1.2) and (1.3) is bilinearity. What does that mean?

A function $B \colon M \times N \to P$, where M, N, and P are R-modules, is called *bilinear* when it is linear in each argument with the other one fixed:

$$B(m_1 + m_2, n) = B(m_1, n) + B(m_2, n), \quad B(rm, n) = rB(m, n),$$

$$B(m, n_1 + n_2) = B(m, n_1) + B(m, n_2), \quad B(m, rn) = rB(m, n).$$

So $B(-, n)$ is a linear map $M \to P$ for each n and $B(m, -)$ is a linear map $N \to P$ for each m. In particular, $B(0, n) = 0$ and $B(m, 0) = 0$. Here are some examples of bilinear maps.

(1) The dot product $\mathbf{v} \cdot \mathbf{w}$ on $\mathbf{R}^n$ is a bilinear function $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$. More generally, for $A \in \mathrm{M}_n(\mathbf{R})$ the function $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v} \cdot A\mathbf{w}$ is a bilinear map $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$.
(2) Matrix multiplication $\mathrm{M}_{m,n}(\mathbf{R}) \times \mathrm{M}_{n,p}(\mathbf{R}) \to \mathrm{M}_{m,p}(\mathbf{R})$ is bilinear. The dot product is the special case $m = p = 1$ (writing $\mathbf{v} \cdot \mathbf{w}$ as $\mathbf{v}^\top \mathbf{w}$).
(3) The cross product $\mathbf{v} \times \mathbf{w}$ is a bilinear function $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}^3$.
(4) The determinant $\det \colon \mathrm{M}_2(R) \to R$ is a bilinear function of matrix columns.
(5) For an R-module M, scalar multiplication $R \times M \to M$ is bilinear.
(6) Multiplication $R \times R \to R$ is bilinear.
(7) Set the dual module of M to be $M^\vee = \mathrm{Hom}_R(M, R)$. The dual pairing $M^\vee \times M \to R$ given by $(\varphi, m) \mapsto \varphi(m)$ is bilinear.
(8) For $\varphi \in M^\vee$ and $\psi \in N^\vee$, the product function $M \times N \to R$ given by $(m, n) \mapsto \varphi(m)\psi(n)$ is bilinear.
(9) If $M \times N \xrightarrow{B} P$ is bilinear and $P \xrightarrow{L} Q$ is linear, the composite $M \times N \xrightarrow{L \circ B} Q$ is bilinear. (This is a very important example. Check it!)
(10) From Section 1, the expression $m \otimes n$ is supposed to be bilinear in m and n. That is, we want the function $M \times N \to M \otimes_R N$ given by $(m, n) \mapsto m \otimes n$ to be bilinear.

Here are a few examples of functions of two arguments that are *not* bilinear:

(1) For an R-module M, addition $M \times M \to M$, where $(m, m') \mapsto m + m'$, is usually not bilinear: it is usually not additive in m when m' is fixed (that is, $(m_1 + m_2) + m' \neq (m_1 + m') + (m_2 + m')$ in general) or additive in m' when m is fixed.

[4]Writing $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ for the standard basis of $\mathbf{R}^3$, Gibbs called a sum $a\mathbf{i} \otimes \mathbf{i} + b\mathbf{j} \otimes \mathbf{j} + c\mathbf{k} \otimes \mathbf{k}$ with positive a, b, and c a *right tensor* [5, p. 57], but I don't know if this had an influence on Voigt's terminology.

[5]I thank Jim Casey for bringing [11] to my attention.

(2) For $\varphi \in M^\vee$ and $\psi \in N^\vee$, the sum $M \times N \to R$ given by $(m,n) \mapsto \varphi(m) + \psi(n)$ is usually not bilinear.
(3) Treat $\mathrm{M}_n(\mathbf{C})$ as a $\mathbf{C}$-vector space. The function $\mathrm{M}_n(\mathbf{C}) \times \mathrm{M}_n(\mathbf{C}) \to \mathrm{M}_n(\mathbf{C})$ given by $(A, B) \mapsto A\overline{B}$ is not bilinear. It is biadditive (*i.e.*, additive in each component when the other one is fixed) but look at how scalar multiplication behaves in the second component: for $z \in \mathbf{C}$, $A\overline{zB}$ is $\overline{z}(A\overline{B})$ rather than $z(A\overline{B})$.

For two R-modules M and N, $M \oplus N$ and $M \times N$ are the same sets, but $M \oplus N$ is an R-module and $M \times N$ doesn't have a module structure. For example, addition on R is a linear function $R \oplus R \to R$, but addition on R is not a bilinear function $R \times R \to R$, as we saw above. Multiplication as a function $R \times R \to R$ is bilinear, but as a function $R \oplus R \to R$ it is not linear (*e.g.*, $(r + r')(s + s') \neq rs + r's'$ in general). Linear functions are generalized additions and bilinear functions are generalized multiplications. Don't confuse a bilinear function on $M \times N$ with a linear function on $M \oplus N$.

An extension of bilinearity is multilinearity. For R-modules $M_1, \dots, M_k$, a function $f \colon M_1 \times \cdots \times M_k \to M$ is called *multilinear* or *k-multilinear* when $f(m_1, \dots, m_k)$ is linear in each m_i with the other coordinates fixed. So 2-multilinear means bilinear. Here are a few examples of multilinear functions:

(1) The scalar triple product $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ is trilinear $\mathbf{R}^3 \times \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$.
(2) The function $f(\mathbf{u}, \mathbf{v}, \mathbf{w}) = (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$ is trilinear $\mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$.
(3) The function $M^\vee \times M \times N \to N$ given by $(\varphi, m, n) \mapsto \varphi(m)n$ is trilinear.
(4) If $B \colon M \times N \to P$ and $B' \colon P \times Q \to T$ are bilinear then $M \times N \times Q \to T$ by $(m, n, q) \mapsto B'(B(m, n), q)$ is trilinear.
(5) Multiplication $R \times \cdots \times R \to R$ with k factors is k-multilinear.
(6) The determinant $\det \colon \mathrm{M}_n(R) \to R$, as a function of matrix columns, is n-multilinear.
(7) If $M_1 \times \cdots \times M_k \xrightarrow{f} M$ is k-multilinear and $M \xrightarrow{L} N$ is linear then the composite $M_1 \times \cdots \times M_k \xrightarrow{L \circ f} N$ is k-multilinear.

The R-linear maps $M \to N$ form an R-module $\mathrm{Hom}_R(M, N)$ under addition of functions and R-scaling. The R-bilinear maps $M \times N \to P$ form an R-module $\mathrm{Bil}_R(M, N; P)$ in the same way. However, unlike linear maps, bilinear maps are *missing* some features:

(1) There is no "kernel" of a bilinear map $M \times N \to P$ since $M \times N$ is not a module.
(2) The image of a bilinear map $M \times N \to P$ need not form a submodule.

Example 2.1. Define $B \colon \mathbf{R}^n \times \mathbf{R}^n \to \mathrm{M}_n(\mathbf{R})$ by $B(\mathbf{v}, \mathbf{w}) = \mathbf{v}\mathbf{w}^\top$, where $\mathbf{v}$ and $\mathbf{w}$ are column vectors, so $\mathbf{v}\mathbf{w}^\top$ is $n \times n$. For example, when $n = 2$,

$$B\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right) = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}(b_1, b_2) = \begin{pmatrix} a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_2 \end{pmatrix}.$$

Generally, if $\mathbf{v} = \sum a_i e_i$ and $\mathbf{w} = \sum b_j e_j$ in terms of the standard basis of $\mathbf{R}^n$, then $\mathbf{v}\mathbf{w}^\top$ is the $n \times n$ matrix $(a_i b_j)$. The formula for $B(\mathbf{v}, \mathbf{w})$ is $\mathbf{R}$-bilinear in $\mathbf{v}$ and $\mathbf{w}$, so B is bilinear. For $n \geq 2$ the image of B isn't closed under addition, so the image isn't a subspace of $\mathrm{M}_n(\mathbf{R})$. Why? Each matrix $B(\mathbf{v}, \mathbf{w})$ has rank 1 (or 0) since its columns are scalar multiples of $\mathbf{v}$. The matrix

$$B(e_1, e_1) + B(e_2, e_2) = e_1 e_1^\top + e_2 e_2^\top = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

has a 2-dimensional image, so $B(e_1, e_1) + B(e_2, e_2) \neq B(\mathbf{v}, \mathbf{w})$ for all $\mathbf{v}$ and $\mathbf{w}$ in $\mathbf{R}^n$. (Similarly, $\sum_{i=1}^n B(e_i, e_i)$ is the $n \times n$ identity matrix, which is not of the form $B(\mathbf{v}, \mathbf{w})$.)

3. Construction of the Tensor Product

Any bilinear map $M \times N \to P$ to an R-module P can be composed with a linear map $P \to Q$ to get a map $M \times N \to Q$ that is bilinear.

$$\begin{tikzcd}
 & P \arrow[dd, "\text{linear}"] \\
M \times N \arrow[ur, "\text{bilinear}"] \arrow[dr, "\text{composite is bilinear!}"'] & \\
 & Q
\end{tikzcd}$$

We will construct the tensor product of M and N as a solution to a universal mapping problem: find an R-module T and bilinear map $b \colon M \times N \to T$ such that *every* bilinear map on $M \times N$ is the composite of the bilinear map b and a unique linear map out of T.

$$\begin{tikzcd}
 & T \arrow[dd, dotted, "\exists\ \text{linear?}"] \\
M \times N \arrow[ur, "b"] \arrow[dr, "\text{bilinear}"'] & \\
 & P
\end{tikzcd}$$

This is analogous to the universal mapping property of the abelianization $G/[G, G]$ of a group G: homomorphisms $G \longrightarrow A$ with abelian A are "the same" as homomorphisms $G/[G, G] \longrightarrow A$ because every homomorphism $f \colon G \to A$ is the composite of the canonical homomorphism $\pi \colon G \to G/[G, G]$ with a unique homomorphism $\widetilde{f} \colon G/[G, G] \to A$.

$$\begin{tikzcd}
 & G/[G,G] \arrow[dd, dotted, "\widetilde{f}"] \\
G \arrow[ur, "\pi"] \arrow[dr, "f"'] & \\
 & A
\end{tikzcd}$$

Definition 3.1. The *tensor product* $M \otimes_R N$ is an R-module equipped with a bilinear map $M \times N \xrightarrow{\otimes} M \otimes_R N$ such that for each bilinear map $M \times N \xrightarrow{B} P$ there is a unique linear map $M \otimes_R N \xrightarrow{L} P$ making the following diagram commute.

$$\begin{tikzcd}
 & M \otimes_R N \arrow[dd, dotted, "L"] \\
M \times N \arrow[ur, "\otimes"] \arrow[dr, "B"'] & \\
 & P
\end{tikzcd}$$

While the functions in the universal mapping property for $G/[G,G]$ are all group homomorphisms (out of G and $G/[G,G]$), functions in the universal mapping property for $M \otimes_R N$ are not all of the same type: those out of $M \times N$ are *bilinear* and those out of $M \otimes_R N$ are *linear*: bilinear maps out of $M \times N$ turn into linear maps out of $M \otimes_R N$.

The definition of the tensor product involves not just a new module $M \otimes_R N$, but also a special bilinear map to it, $\otimes\colon M \times N \longrightarrow M \otimes_R N$. This is similar to the universal mapping property for the abelianization $G/[G,G]$, which requires not just $G/[G,G]$ but also the homomorphism $\pi\colon G \longrightarrow G/[G,G]$ through which all homomorphisms from G to abelian groups factor. The universal mapping property requires fixing this extra information.

Before building a tensor product, let's show any two tensor products are essentially the same. Let R-modules T and T', and bilinear maps $M \times N \xrightarrow{b} T$ and $M \times N \xrightarrow{b'} T'$, satisfy the universal mapping property of the tensor product. From universality of $M \times N \xrightarrow{b} T$, the map $M \times N \xrightarrow{b'} T'$ factors uniquely through T: a unique linear map $f\colon T \to T'$ makes

(3.1)
$$\begin{tikzcd} & T \arrow[dd, dotted, "f"] \\ M \times N \arrow[ur, "b"] \arrow[dr, "b'"'] & \\ & T' \end{tikzcd}$$

commute. From universality of $M \times N \xrightarrow{b'} T'$, the map $M \times N \xrightarrow{b} T$ factors uniquely through T': a unique linear map $f'\colon T' \to T$ makes

(3.2)
$$\begin{tikzcd} & T' \arrow[dd, dotted, "f'"] \\ M \times N \arrow[ur, "b'"] \arrow[dr, "b"'] & \\ & T \end{tikzcd}$$

commute. We combine (3.1) and (3.2) into the commutative diagram

$$\begin{tikzcd} & T \arrow[d, "f"] \\ M \times N \arrow[ur, "b"] \arrow[r, "b'"] \arrow[dr, "b"'] & T' \arrow[d, "f'"] \\ & T \end{tikzcd}$$

Removing the middle, we have the commutative diagram

(3.3)
$$\begin{tikzcd} & T \arrow[dd, "f'\circ f"] \\ M \times N \arrow[ur, "b"] \arrow[dr, "b"'] & \\ & T \end{tikzcd}$$

From universality of (T, b), a unique linear map $T \to T$ fits in (3.3). The identity map works, so $f' \circ f = \mathrm{id}_T$. Similarly, $f \circ f' = \mathrm{id}_{T'}$ by stacking (3.1) and (3.2) together in the other order. Thus T and T' are isomorphic R-modules by f and also $f \circ b = b'$, which means f identifies b with b'. So two tensor products of M and N can be identified with each other in a unique way *compatible*[6] with the distinguished bilinear maps to them from $M \times N$.

Theorem 3.2. *A tensor product of M and N exists.*

Proof. Consider $M \times N$ simply as a set. We form the free R-module on this set:

$$F_R(M \times N) = \bigoplus_{(m,n)\in M\times N} R\delta_{(m,n)}.$$

(This is an *enormous* R-module. If $R = \mathbf{R}$ and $M = N = \mathbf{R}^3$ then $F_R(M \times N)$ is a direct sum of $\mathbf{R}^6$-many copies of $\mathbf{R}$. The direct sum runs over all pairs of vectors from $\mathbf{R}^3$, not just pairs coming from a basis of $\mathbf{R}^3$, and its components lie in $\mathbf{R}$. For most modules a basis doesn't even generally exist.) Let D be the submodule of $F_R(M \times N)$ spanned by all the elements

$$\delta_{(m+m',n)} - \delta_{(m,n)} - \delta_{(m',n)}, \quad \delta_{(m,n+n')} - \delta_{(m,n)} - \delta_{(m,n')}, \quad \delta_{(rm,n)} - \delta_{(m,rn)},$$
$$r\delta_{(m,n)} - \delta_{(rm,n)}, \quad r\delta_{(m,n)} - \delta_{(m,rn)}.$$

The quotient module by D will serve as the tensor product: set

$$M \otimes_R N := F_R(M \times N)/D.$$

We write the coset $\delta_{(m,n)} + D$ in $M \otimes_R N$ as $m \otimes n$.

From the definition of D, we get relations in $F_R(M \times N)/D$ like

$$\delta_{(m+m',n)} \equiv \delta_{(m,n)} + \delta_{(m',n)} \bmod D,$$

which is the same as

$$(m + m') \otimes n = m \otimes n + m' \otimes n$$

in $M \otimes_R N$. Similarly, $m \otimes (n + n') = m \otimes n + m \otimes n'$ and $r(m \otimes n) = rm \otimes n = m \otimes rn$ in $M \otimes_R N$. These relations are the reason D was defined the way it was, and they show that the function $M \times N \xrightarrow{\otimes} M \otimes_R N$ given by $(m, n) \mapsto m \otimes n$ is bilinear. (No other function $M \times N \to M \otimes_R N$ will be considered except this one.)

Now we will show all bilinear maps out of $M \times N$ factor uniquely through the bilinear map $M \times N \to M \otimes_R N$ that we just wrote down. Suppose P is an R-module and $M \times N \xrightarrow{B} P$ is a bilinear map. Treating $M \times N$ simply as a *set*, so B is just a function on this set (ignore its bilinearity), the universal mapping property of free modules extends B from a function

[6]The universal mapping property is not about modules T *per se*, but about pairs (T, b).

$M \times N \to P$ to a linear function $\ell \colon F_R(M \times N) \to P$ with $\ell(\delta_{(m,n)}) = B(m,n)$, so the diagram

$$\begin{tikzcd} & F_R(M \times N) \arrow[dd, "\ell"] \\ M \times N \arrow[ur, "(m,n)\mapsto\delta_{(m,n)}"] \arrow[dr, "B"'] & \\ & P \end{tikzcd}$$

commutes. We want to show ℓ makes sense as a function on $M \otimes_R N$, which means showing $\ker \ell$ contains D. From the bilinearity of B,

$$B(m+m',n) = B(m,n) + B(m,n'), \quad B(m,n+n') = B(m,n) + B(m,n'),$$
$$rB(m,n) = B(rm,n) = B(m,rn),$$

so

$$\ell(\delta_{(m+m',n)}) = \ell(\delta_{(m,n)}) + \ell(\delta_{(m',n)}), \quad \ell(\delta_{(m,n+n')}) = \ell(\delta_{(m,n)}) + \ell(\delta_{(m,n')}),$$
$$r\ell(\delta_{(m,n)}) = \ell(\delta_{(rm,n)}) = \ell(\delta_{(m,rn)}).$$

Since ℓ is linear, these conditions are the same as

$$\ell(\delta_{(m+m',n)}) = \ell(\delta_{(m,n)} + \delta_{(m',n)}), \quad \ell(\delta_{(m,n+n')}) = \ell(\delta_{(m,n)} + \delta_{(m,n')}),$$
$$\ell(r\delta_{(m,n)}) = \ell(\delta_{(rm,n)}) = \ell(\delta_{(m,rn)}).$$

Therefore the kernel of ℓ contains all the generators of the submodule D, so ℓ induces a linear map $L \colon F_R(M \times N)/D \to P$ where $L(\delta_{(m,n)} + D) = \ell(\delta_{(m,n)}) = B(m,n)$, which means the diagram

$$\begin{tikzcd} & F_R(M \times N)/D \arrow[dd, "L"] \\ M \times N \arrow[ur, "(m,n)\mapsto\delta_{(m,n)}+D"] \arrow[dr, "B"'] & \\ & P \end{tikzcd}$$

commutes. Since $F_R(M \times N)/D = M \otimes_R N$ and $\delta_{(m,n)} + D = m \otimes n$, the above diagram is

(3.4)
$$\begin{tikzcd} & M \otimes_R N \arrow[dd, "L"] \\ M \times N \arrow[ur, "\otimes"] \arrow[dr, "B"'] & \\ & P \end{tikzcd}$$

and that shows every bilinear map B out of $M \times N$ comes from a linear map L out of $M \otimes_R N$ such that $L(m \otimes n) = B(m,n)$ for all $m \in M$ and $n \in N$.

It remains to show the linear map $M \otimes_R N \xrightarrow{L} P$ in (3.4) is the only one that makes (3.4) commute. We go back to the definition of $M \otimes_R N$ as a quotient of the free module

$F_R(M \times N)$. From the construction of free modules, every element of $F_R(M \times N)$ is a finite sum

$$r_1\delta_{(m_1,n_1)} + \cdots + r_k\delta_{(m_k,n_k)}.$$

The reduction map $F_R(M \times N) \twoheadrightarrow F_R(M \times N)/D = M \otimes_R N$ is linear, so every element of $M \otimes_R N$ is a finite sum

$$r_1(m_1 \otimes n_1) + \cdots + r_k(m_k \otimes n_k). \tag{3.5}$$

This means the elements $m \otimes n$ in $M \otimes_R N$ span it as an R-module. Therefore linear maps out of $M \otimes_R N$ are completely determined by their values on all the elements $m \otimes n$, so there is at most one linear map $M \otimes_R N \to P$ with the effect $m \otimes n \mapsto B(m,n)$. Since we have created a linear map out of $M \otimes_R N$ with this very effect in (3.4), it is the only one. $\square$

Having shown a tensor product of M and N exists,[7] its essential uniqueness lets us call $M \otimes_R N$ "the" tensor product rather than "a" tensor product. Don't forget that the construction involves not only the module $M \otimes_R N$ but also the distinguished bilinear map $M \times N \xrightarrow{\otimes} M \otimes_R N$ given by $(m,n) \mapsto m \otimes n$, through which all bilinear maps out of $M \times N$ factor. We call this distinguished map the *canonical* bilinear map from $M \times N$ to the tensor product. Elements of $M \otimes_R N$ are called **tensors**, and will be denoted by the letter t. Tensors in $M \otimes_R N$ that have the form $m \otimes n$ are called *elementary tensors*. (Other names for elementary tensors are simple tensors, decomposable tensors, pure tensors, and monomial tensors.) Just as elements of the free R-module $F_R(A)$ on a set A are usually *not* of the form δ_a but are linear combinations of these, **elements of $M \otimes_R N$ are usually not elementary tensors**[8] but are linear combinations of elementary tensors. In fact each tensor is a *sum* of elementary tensors since $r(m \otimes n) = (rm) \otimes n$. This shows all elements of $M \otimes_R N$ have the form (1.4).

That every tensor is a sum of elementary tensors, but need not be an elementary tensor itself, is a feature that confuses people who are learning about tensor products. One source of the confusion is that in the direct sum $M \oplus N$ every element *is* a pair (m,n), so why shouldn't every element of $M \otimes_R N$ have the form $m \otimes n$? Here are two related ideas to keep in mind, so it seems less strange that not all tensors are elementary.

- The R-module $R[X,Y]$ is a tensor product of $R[X]$ and $R[Y]$ (see Example 4.12) and, as Eisenbud and Harris note in their book on schemes [4, p. 39], the study of polynomials in two variables is more than the study of polynomials of the form $f(X)g(Y)$. That is, most polynomials in $R[X,Y]$ are not $f(X)g(Y)$, but they are all a sum of such products (and in fact they are sums of monomials $a_{ij}X^iY^j$).

---

[7] What happens if R is a noncommutative ring? If M and N are left R-modules and B is bilinear on $M \times N$ then for all $m \in M$, $n \in N$, and r and s in R, $rsB(m,n) = rB(m,sn) = B(rm,sn) = sB(rm,n) = srB(m,n)$. Usually $rs \neq sr$, so asking that $rsB(m,n) = srB(m,n)$ for all m and n puts us in a delicate situation! The correct tensor product $M \otimes_R N$ for noncommutative R uses a *right* R-module M, a *left* R-module N, and a "middle-linear" map B where $B(mr,n) = B(m,rn)$. In fact $M \otimes_R N$ is not an R-module but just an abelian group! While we won't deal with tensor products over a noncommutative ring, they *are* important. They appear in the construction of induced representations of groups.

[8] An explicit example of a nonelementary tensor in $R^2 \otimes_R R^2$ will be provided in Example 4.11. We essentially already met one in Example 2.1 when we saw $e_1e_1^\top + e_2e_2^\top \neq \mathbf{v}\mathbf{w}^\top$ for all $\mathbf{v}$ and $\mathbf{w}$ in $\mathbf{R}^n$.

- The role of elementary tensors among all tensors is like that of separable solutions $f(x)g(y)$ to a 2-variable PDE among all solutions.[9] Solutions to a PDE are not generally separable, so one first aims to understand separable solutions and then tries to form the general solution as a sum (perhaps an *infinite* sum) of separable solutions.

From now on *forget* the explicit construction of $M \otimes_R N$ as the quotient of an enormous free module $F_R(M \times N)$. It will confuse you more than it's worth to try to think about $M \otimes_R N$ in terms of its construction. What is more important to remember is the universal mapping property of the tensor product, which we will start using systematically in the next section. To get used to the bilinearity of $\otimes$, let's prove two simple results.

Theorem 3.3. *Let M and N be R-modules with respective spanning sets $\{x_i\}_{i\in I}$ and $\{y_j\}_{j\in J}$. The tensor product $M \otimes_R N$ is spanned linearly by the elementary tensors $x_i \otimes y_j$.*

Proof. An elementary tensor in $M \otimes_R N$ has the form $m \otimes n$. Write $m = \sum_i a_i x_i$ and $n = \sum_j b_j y_j$, where the a_i's and b_j's are 0 for all but finitely many i and j. From the bilinearity of $\otimes$,

$$m \otimes n = \sum_i a_i x_i \otimes \sum_j b_j y_j = \sum_{i,j} a_i b_j x_i \otimes y_j$$

is a linear combination of the tensors $x_i \otimes y_j$. So every elementary tensor is a linear combination of the particular elementary tensors $x_i \otimes y_j$. Since every tensor is a sum of elementary tensors, the $x_i \otimes y_j$'s span $M \otimes_R N$ as an R-module. □

Example 3.4. Let $e_1, \ldots, e_k$ be the standard basis of R^k. The R-module $R^k \otimes_R R^k$ is linearly spanned by the k^2 elementary tensors $e_i \otimes e_j$. We will see later (Theorem 4.9) that these elementary tensors are a *basis* of $R^k \otimes_R R^k$, which for R a field is consistent with the physicist's "definition" of tensor products of vector spaces from Section 1 using bases.

Theorem 3.5. *In $M \otimes_R N$, $m \otimes 0 = 0$ and $0 \otimes n = 0$.*

Proof. This is just like the proof that $a \cdot 0 = 0$ in a ring: since $m \otimes n$ is additive in n with m fixed, $m \otimes 0 = m \otimes (0+0) = m \otimes 0 + m \otimes 0$. Subtracting $m \otimes 0$ from both sides, $m \otimes 0 = 0$. That $0 \otimes n = 0$ follows by a similar argument. □

Example 3.6. If A is a finite abelian group, $\mathbf{Q} \otimes_{\mathbf{Z}} A = 0$ since every elementary tensor is 0: for $a \in A$, let $na = 0$ for some positive integer n. Then in $\mathbf{Q} \otimes_{\mathbf{Z}} A$, $r \otimes a = n(r/n) \otimes a = r/n \otimes na = r/n \otimes 0 = 0$. Every tensor is a sum of elementary tensors, and every elementary tensor is 0, so all tensors are 0. (For instance, $(1/3) \otimes (5 \bmod 7) = 0$ in $\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Z}/7\mathbf{Z}$. Thus we can have $m \otimes n = 0$ without m or n being 0.)

To show $\mathbf{Q} \otimes_{\mathbf{Z}} A = 0$, we don't need A to be finite, but rather that each element of A has finite order. The group $\mathbf{Q}/\mathbf{Z}$ has that property, so $\mathbf{Q} \otimes_{\mathbf{Z}} (\mathbf{Q}/\mathbf{Z}) = 0$. By a similar argument, $\mathbf{Q}/\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Q}/\mathbf{Z} = 0$.

Since $M \otimes_R N$ is spanned additively by elementary tensors, each linear (or just additive) function out of $M \otimes_R N$ is determined on all tensors from its values on elementary tensors. This is why linear maps on tensor products are in practice described only by their values on elementary tensors. It is similar to describing a linear map between finite free modules

[9] In Brad Osgood's notes on the Fourier transform [13, pp. 343-344], he writes about functions of the form $f_1(x_1)f_2(x_2)\cdots f_n(x_n)$ "If you really want to impress your friends and confound your enemies, you can invoke *tensor products* in this context. [...] People run in terror from the $\otimes$ symbol. Cool."

using a matrix. The matrix directly tells you only the values of the map on a particular basis, but this information is enough to determine the linear map everywhere.

However, there is a key difference between basis vectors and elementary tensors: elementary tensors have lots of linear relations. A linear map out of $\mathbf{R}^2$ is determined by its values on $(1,0)$, $(2,3)$, $(8,4)$, and $(-1,5)$, but those values are not independent: they have to satisfy every linear relation the four vectors satisfy because a linear map preserves linear relations. Similarly, a random function on elementary tensors generally does not extend to a linear map on the tensor product: elementary tensors span the tensor product of two modules, but they are not linearly independent.

Functions of elementary tensors can't be created out of a random function of two variables. For instance, the "function" $f(m \otimes n) = m + n$ makes no sense since $m \otimes n = (-m) \otimes (-n)$ but $m + n$ is usually not $-m - n$. The *only* way to create linear maps out of $M \otimes_R N$ is with the universal mapping property of the tensor product (which creates linear maps out of bilinear maps), because all linear relations among elementary tensors – from the obvious to the obscure – are built into the universal mapping property of $M \otimes_R N$. There will be a lot of practice with this in Section 4. Understanding how the universal mapping property of the tensor product can be used to compute examples and to prove properties of the tensor product is the best way to get used to the tensor product; if you can't write down functions out of $M \otimes_R N$, you don't understand $M \otimes_R N$.

The tensor product can be extended to allow more than two factors. Given k modules $M_1, \ldots, M_k$, there is a module $M_1 \otimes_R \cdots \otimes_R M_k$ that is universal for k-multilinear maps: it admits a k-multilinear map $M_1 \times \cdots \times M_k \xrightarrow{\otimes} M_1 \otimes_R \cdots \otimes_R M_k$ and every k-multilinear map out of $M_1 \times \cdots \times M_k$ factors through this by composition with a unique linear map out of $M_1 \otimes_R \cdots \otimes_R M_k$:

$$\begin{tikzcd}
 & M_1 \otimes_R \cdots \otimes_R M_k \arrow[dd, dotted, "\exists \text{ unique linear}"] \\
M_1 \times \cdots \times M_k \arrow[ur, "\otimes"] \arrow[dr, "\text{multilinear}"'] & \\
 & P
\end{tikzcd}$$

The image of $(m_1, \ldots, m_k)$ in $M_1 \otimes_R \cdots \otimes_R M_k$ is written $m_1 \otimes \cdots \otimes m_k$. This k-fold tensor product can be constructed as a quotient of the free module $F_R(M_1 \times \cdots \times M_k)$. It can also be constructed using tensor products of modules two at a time:

$$(\cdots((M_1 \otimes_R M_2) \otimes_R M_3) \otimes_R \cdots) \otimes_R M_k.$$

The canonical k-multilinear map to this R-module from $M_1 \times \cdots \times M_k$ is $(m_1, \ldots, m_k) \mapsto (\cdots((m_1 \otimes m_2) \otimes m_3) \cdots) \otimes m_k$. This is not the same construction of the k-fold tensor product using $F_R(M_1 \times \cdots \times M_k)$, but it satisfies the same universal mapping property and thus can serve the same purpose (all constructions of a tensor product of $M_1, \ldots, M_k$ are isomorphic to each other in a unique way compatible with the distinguished k-multilinear maps to them from $M_1 \times \cdots \times M_k$).

The module $M_1 \otimes_R \cdots \otimes_R M_k$ is spanned additively by all $m_1 \otimes \cdots \otimes m_k$. Important examples of the k-fold tensor product are *tensor powers* $M^{\otimes k}$:

$$M^{\otimes 0} = R, \quad M^{\otimes 1} = M, \quad M^{\otimes 2} = M \otimes_R M, \quad M^{\otimes 3} = M \otimes_R M \otimes_R M,$$

and so on. (The formula $M^{\otimes 0} = R$ is a convention, like $a^0 = 1$.)

Let's address a few beginner questions about the tensor product:

Questions

(1) What is $m \otimes n$?
(2) What does it mean to say $m \otimes n = 0$?
(3) What does it mean to say $M \otimes_R N = 0$?
(4) What does it mean to say $m_1 \otimes n_1 + \cdots + m_k \otimes n_k = m_1' \otimes n_1' + \cdots + m_\ell' \otimes n_\ell'$?
(5) Where do tensor products arise outside of mathematics?
(6) Is there a way to picture the tensor product?

Answers

(1) Strictly speaking, $m \otimes n$ is the image of $(m, n) \in M \times N$ under the canonical bilinear map $M \times N \xrightarrow{\otimes} M \otimes_R N$ in the definition of the tensor product. Here's another answer, which is not a definition but more closely aligns with how $m \otimes n$ occurs in practice: $m \otimes n$ is that element of $M \otimes_R N$ at which the linear map $M \otimes_R N \to P$ corresponding to a bilinear map $M \times N \xrightarrow{B} P$ takes the value $B(m, n)$. Review the proof of Theorem 3.2 and check this property of $m \otimes n$ really holds.

(2) We have $m \otimes n = 0$ if and only if every bilinear map out of $M \times N$ vanishes at (m, n). Indeed, if $m \otimes n = 0$ then for each bilinear map $B \colon M \times N \to P$ we have a commutative diagram

$$\begin{tikzcd} & M \otimes_R N \arrow[dd, "L"] \\ M \times N \arrow[ur, "\otimes"] \arrow[dr, "B"'] & \\ & P \end{tikzcd}$$

for some linear map L, so $B(m, n) = L(m \otimes n) = L(0) = 0$. Conversely, if every bilinear map out of $M \times N$ sends (m, n) to 0 then the canonical bilinear map $M \times N \to M \otimes_R N$, which is a particular example, sends (m, n) to 0. Since this bilinear map actually sends (m, n) to $m \otimes n$, we obtain $m \otimes n = 0$.

A very important consequence is a tip about how to show a particular elementary tensor $m \otimes n$ is *not* 0: find a bilinear map B out of $M \times N$ such that $B(m, n) \neq 0$. Remember this idea! It will be used in Theorem 4.9.

That $m \otimes 0 = 0$ and $0 \otimes n = 0$ is related to $B(m, 0) = 0$ and $B(0, n) = 0$ for each bilinear map B on $M \times N$. This gives another proof of Theorem 3.5.

As an exercise, check from the universal mapping property that $m_1 \otimes \cdots \otimes m_k = 0$ in $M_1 \otimes_R \cdots \otimes_R M_k$ if and only if all k-multilinear maps out of $M_1 \times \cdots \times M_k$ vanish at $(m_1, \ldots, m_k)$.

(3) The tensor product $M \otimes_R N$ is 0 if and only if every bilinear map out of $M \times N$ (to *all modules*) is identically 0. First suppose $M \otimes_R N = 0$. Then all elementary tensors $m \otimes n$ are 0, so $B(m, n) = 0$ for all bilinear maps out of $M \times N$ by the answer to the second question. Thus B is identically 0. Next suppose every bilinear map out of $M \times N$ is identically 0. Then the canonical bilinear map $M \times N \xrightarrow{\otimes} M \otimes_R N$, which is a particular example, is identically 0. Since this function sends (m, n) to $m \otimes n$, we have $m \otimes n = 0$ for all m and n. Since $M \otimes_R N$ is additively spanned by all $m \otimes n$, the vanishing of all elementary tensors implies $M \otimes_R N = 0$.

Returning to Example 3.6, that $\mathbf{Q} \otimes_{\mathbf{Z}} A = 0$ if each element of A has finite order is another way of saying every $\mathbf{Z}$-bilinear map out of $\mathbf{Q} \times A$ is identically zero, which can be verified directly: if B is such a map (into an abelian group) and $na = 0$ with $n \geq 1$, then $B(r, a) = B(n(r/n), a) = B(r/n, na) = B(r/n, 0) = 0$.

Turning this idea around, to show some tensor product module $M \otimes_R N$ is *not* 0, find a bilinear map on $M \times N$ that is not identically 0.

(4) We have $\sum_{i=1}^k m_i \otimes n_i = \sum_{j=1}^{\ell} m'_j \otimes n'_j$ if and only if for all bilinear maps B out of $M \times N$, $\sum_{i=1}^k B(m_i, n_i) = \sum_{j=1}^{\ell} B(m'_j, n'_j)$. The justification is along the lines of the previous two answers and is left to the reader. For example, the condition $\sum_{i=1}^k m_i \otimes n_i = 0$ means $\sum_{i=1}^k B(m_i, n_i) = 0$ for all bilinear maps B on $M \times N$.

(5) Tensors are used in physics and engineering (stress, elasticity, electromagnetism, metrics, diffusion MRI), where they transform in a multilinear way under a change in coordinates. The treatment of tensors in physics is discussed in Section 7.

(6) There isn't a simple picture of a tensor (even an elementary tensor) analogous to how a vector is an arrow. Some physical manifestations of tensors are in the previous answer, but they won't help you understand tensor products of modules.

Nobody is comfortable with tensor products at first. Two quotes by Cathy O'Neil and Johan de Jong[10] nicely capture the phenomenon of learning about them:

- O'Neil: After a few months, though, I realized something. I hadn't gotten any better at understanding tensor products, but I was getting used to *not* understanding them. It was pretty amazing. I no longer felt anguished when tensor products came up; I was instead almost amused by their cunning ways.
- de Jong: It is the things you *can* prove that tell you how to think about tensor products. In other words, you let elementary lemmas and examples shape your intuition of the mathematical object in question. There's nothing else, no magical intuition will magically appear to help you "understand" it.

Remark 3.7. Hassler Whitney, who first defined tensor products beyond the setting of vector spaces, called abelian groups A and B a *group pair* relative to the abelian group C if there is a $\mathbf{Z}$-bilinear map $A \times B \to C$ and then wrote [18, p. 499] that "*any such group pair may be defined by choosing a homomorphism*" $A \otimes_{\mathbf{Z}} B \to C$. So the idea that $\otimes_{\mathbf{Z}}$ solves a universal mapping problem is essentially due to Whitney.

4. Examples of Tensor Products

Theorem 4.1. *For positive integers a and b with $d = (a, b)$, $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \cong \mathbf{Z}/d\mathbf{Z}$ as abelian groups. In particular, $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} = 0$ if and only if $(a, b) = 1$.*

Proof. Since 1 spans $\mathbf{Z}/a\mathbf{Z}$ and $\mathbf{Z}/b\mathbf{Z}$, $1 \otimes 1$ spans $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z}$ by Theorem 3.3. From

$$a(1 \otimes 1) = a \otimes 1 = 0 \otimes 1 = 0 \text{ and } b(1 \otimes 1) = 1 \otimes b = 1 \otimes 0 = 0,$$

the additive order of $1 \otimes 1$ divides a and b, and therefore also d, so $\#(\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z}) \leq d$.

To show $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z}$ has size at least d, we *create* a $\mathbf{Z}$-linear map from $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z}$ onto $\mathbf{Z}/d\mathbf{Z}$. Since $d|a$ and $d|b$, we can reduce $\mathbf{Z}/a\mathbf{Z} \to \mathbf{Z}/d\mathbf{Z}$ and $\mathbf{Z}/b\mathbf{Z} \to \mathbf{Z}/d\mathbf{Z}$ in the natural way. Consider the map $\mathbf{Z}/a\mathbf{Z} \times \mathbf{Z}/b\mathbf{Z} \xrightarrow{B} \mathbf{Z}/d\mathbf{Z}$ that is reduction mod d in each factor followed by multiplication: $B(x \bmod a, y \bmod b) = xy \bmod d$. This is $\mathbf{Z}$-bilinear, so

[10] See http://mathbabe.org/2011/07/20/what-tensor-products-taught-me-about-living-my-life/.

the universal mapping property of the tensor product says there is a (unique) $\mathbf{Z}$-linear map $f\colon \mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \to \mathbf{Z}/d\mathbf{Z}$ making the diagram

$$\begin{tikzcd} & \mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \arrow[dd, "f"] \\ \mathbf{Z}/a\mathbf{Z} \times \mathbf{Z}/b\mathbf{Z} \arrow[ur, "\otimes"] \arrow[dr, "B"'] & \\ & \mathbf{Z}/d\mathbf{Z} \end{tikzcd}$$

commute, so $f(x \otimes y) = xy$. In particular, $f(x \otimes 1) = x$, so f is onto. Therefore $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z}$ has size at least d, so the size is d and we're done. $\square$

Example 4.2. The abelian group $\mathbf{Z}/3\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/5\mathbf{Z}$ is 0. This type of collapsing in a tensor product often bothers people when they first see it, but it's saying something pretty concrete: each $\mathbf{Z}$-bilinear map $B\colon \mathbf{Z}/3\mathbf{Z} \times \mathbf{Z}/5\mathbf{Z} \to A$ to an abelian group A is identically 0, which is easy to show directly: $3B(a,b) = B(3a,b) = B(0,b) = 0$ and $5B(a,b) = B(a,5b) = B(a,0) = 0$, so $B(a,b)$ is killed by $3\mathbf{Z} + 5\mathbf{Z} = \mathbf{Z}$, hence $B(a,b)$ is killed by 1, which is another way of saying $B(a,b) = 0$.

In $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z}$ all tensors are elementary tensors: $x \otimes y = xy(1 \otimes 1)$ and a sum of multiples of $1 \otimes 1$ is again a multiple, so $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} = \mathbf{Z}(1 \otimes 1) = \{x \otimes 1 : x \in \mathbf{Z}\}$.

Notice in the proof of Theorem 4.1 how the map $f\colon \mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \to \mathbf{Z}/d\mathbf{Z}$ was created from the bilinear map $B\colon \mathbf{Z}/a\mathbf{Z} \times \mathbf{Z}/b\mathbf{Z} \to \mathbf{Z}/d\mathbf{Z}$ and the universal mapping property of tensor products. Quite generally, to define a linear map out of $M \otimes_R N$ that sends all elementary tensors $m \otimes n$ to particular places, *always* back up and start by defining a bilinear map out of $M \times N$ sending (m,n) to the place you want $m \otimes n$ to go. Make sure you show the map is bilinear! Then the universal mapping property of the tensor product gives you a linear map out of $M \otimes_R N$ sending $m \otimes n$ to the place where (m,n) goes, which gives you what you wanted: a (unique) linear map on the tensor product with specified values on the elementary tensors.

Theorem 4.3. *For ideals I and J in R, there is a unique R-module isomorphism*

$$R/I \otimes_R R/J \cong R/(I+J)$$

where $\overline{x} \otimes \overline{y} \mapsto \overline{xy}$. In particular, taking $I = J = 0$, $R \otimes_R R \cong R$ by $x \otimes y \mapsto xy$.

For $R = \mathbf{Z}$ and nonzero I and J, this is Theorem 4.1.

Proof. Start with the function $R/I \times R/J \to R/(I+J)$ given by $(x \bmod I, y \bmod J) \mapsto xy \bmod I+J$. This is well-defined and bilinear, so from the universal mapping property of the tensor product we get a linear map $f\colon R/I \otimes_R R/J \to R/(I+J)$ making the diagram

$$\begin{tikzcd} & R/I \otimes_R R/J \arrow[dd, "f"] \\ R/I \times R/J \arrow[ur, "\otimes"] \arrow[dr, "(x \bmod I,y \bmod J)\mapsto xy \bmod I+J"'] & \\ & R/(I+J) \end{tikzcd}$$

commute, so $f(x \bmod I \otimes y \bmod J) = xy \bmod I+J$. To write down the inverse map, let $R \to R/I \otimes_R R/J$ by $r \mapsto r(\overline{1} \otimes \overline{1})$. This is linear, and when $r \in I$ the value is $\overline{r} \otimes \overline{1} = \overline{0} \otimes \overline{1} = 0$. Similarly, when $r \in J$ the value is 0. Therefore $I+J$ is in the kernel, so we get a linear map $g\colon R/(I+J) \to R/I \otimes_R R/J$ by $g(r \bmod I+J) = r(\overline{1} \otimes \overline{1}) = \overline{r} \otimes \overline{1} = \overline{1} \otimes \overline{r}$.

To check f and g are inverses, a computation in one direction shows

$$f(g(r \bmod I+J)) = f(\overline{r} \otimes \overline{1}) = r \bmod I+J.$$

To show $g(f(t)) = t$ for all $t \in R/I \otimes_R R/J$, we show each tensor is a scalar multiple of $\overline{1} \otimes \overline{1}$. An elementary tensor is $\overline{x} \otimes \overline{y} = x\overline{1} \otimes y\overline{1} = xy(\overline{1} \otimes \overline{1})$, so sums of elementary tensors are multiples of $\overline{1} \otimes \overline{1}$ and thus all tensors are multiples of $\overline{1} \otimes \overline{1}$. We have

$$g(f(r(\overline{1} \otimes \overline{1}))) = rg(1 \bmod I+J) = r(\overline{1} \otimes \overline{1}).$$

□

Remark 4.4. For ideals I and J, a few operations produce new ideals: $I+J$, $I \cap J$, and IJ. The intersection $I \cap J$ is the kernel of the linear map $R \to R/I \oplus R/J$ where $r \mapsto (\overline{r}, \overline{r})$. Theorem 4.3 tells us $I+J$ is the kernel of the linear map $R \to R/I \otimes_R R/J$ where $r \mapsto r(\overline{1} \otimes \overline{1})$.

Theorem 4.5. *For an ideal I in R and R-module M, there is a unique R-module isomorphism*

$$(R/I) \otimes_R M \cong M/IM$$

such that $\overline{r} \otimes m \mapsto \overline{rm}$. In particular, taking $I = (0)$, $R \otimes_R M \cong M$ by $r \otimes m \mapsto rm$, so $R \otimes_R R \cong R$ as R-modules by $r \otimes r' \mapsto rr'$.

Proof. We start with the bilinear map $(R/I) \times M \to M/IM$ given by $(\overline{r}, m) \mapsto \overline{rm}$. From the universal mapping property of the tensor product, we get a linear map $f\colon (R/I) \otimes_R M \to M/IM$ where $f(\overline{r} \otimes m) = \overline{rm}$.

$$\begin{tikzcd} & (R/I) \otimes_R M \arrow[dd, "f"] \\ (R/I) \times M \arrow[ur, "\otimes"] \arrow[dr, "(\overline{r},m) \mapsto \overline{rm}"'] & \\ & M/IM \end{tikzcd}$$

To create an inverse map, start with the function $M \to (R/I) \otimes_R M$ given by $m \mapsto \overline{1} \otimes m$. This is linear in m (check!) and kills IM (generators for IM are products im for $i \in I$ and $m \in M$, and $\overline{1} \otimes rm = \overline{i} \otimes m = \overline{0} \otimes m = 0$), so it induces a linear map $g\colon M/IM \to (R/I) \otimes_R M$ given by $g(\overline{m}) = \overline{1} \otimes m$.

To check $f(g(\overline{m})) = \overline{m}$ and $g(f(t)) = t$ for all $\overline{m} \in M/IM$ and $t \in (R/I) \otimes_R M$, we do the first one by a direct computation:

$$f(g(\overline{m})) = f(\overline{1} \otimes m) = \overline{1 \cdot m} = \overline{m}.$$

To show $g(f(t)) = t$ for all $t \in M \otimes_R N$, we show all tensors in $R/I \otimes_R M$ are elementary. An elementary tensor looks like $\overline{r} \otimes m = \overline{1} \otimes rm$, and a sum of tensors $\overline{1} \otimes m_i$ is $\overline{1} \otimes \sum_i m_i$. Thus all tensors look like $\overline{1} \otimes m$. We have $g(f(\overline{1} \otimes m)) = g(\overline{m}) = \overline{1} \otimes m$. □

Example 4.6. For every abelian group A, $(\mathbf{Z}/n\mathbf{Z}) \otimes_{\mathbf{Z}} A \cong A/nA$ as abelian groups by $\overline{m} \otimes a \mapsto \overline{ma}$.

Remark 4.7. That $R \otimes_R M \cong M$ by $r \otimes m \mapsto rm$ says R-bilinear maps B out of $R \times M$ can be identified with R-linear maps out of M, since $B(r, m) = B(1, rm)$ by bilinearity and $B(1, -)$ is linear in the second component.

Remark 4.8. For an ideal I in R and R-module M there is an R-linear map $I \otimes_R M \to IM$ where $i \otimes m \mapsto im$, and it's surjective (IM is spanned by all im, which are in the image), but *not* necessarily injective! If $R = \mathbf{Z}$, $I = a\mathbf{Z}$, and $M = \mathbf{Z}/a\mathbf{Z}$ for $a \geq 2$ then $I \otimes_R M = (a\mathbf{Z}) \otimes_{\mathbf{Z}} (\mathbf{Z}/a\mathbf{Z}) \cong \mathbf{Z}/a\mathbf{Z}$ by $(ax) \otimes (y \bmod a) \mapsto xy \bmod a$, but $IM = \{0\}$.

In this section so far, M or N in $M \otimes_R N$ has been R or R/I. Such a module contains 1 or $\overline{1}$, making all tensors in $M \otimes_R N$ elementary. *Don't be misled.* Most tensors are not elementary, and don't think $m \otimes n = mn(1 \otimes 1)$; 1 is not in a general module, so $1 \otimes 1$ usually doesn't make sense.[11] The next theorem, which justifies the discussion in the introduction about bases for tensor products of free modules, will let us construct nonelementary tensors.

Theorem 4.9. *If F and F' are free R-modules, with respective bases $\{e_i\}_{i \in I}$ and $\{e'_j\}_{j \in J}$, then $F \otimes_R F'$ is a free R-module with basis $\{e_i \otimes e'_j\}_{(i,j) \in I \times J}$.*

Proof. The result is clear if F or F' is 0, so let them both be nonzero free modules (hence $R \neq 0$ and F and F' have bases). By Theorem 3.3, $\{e_i \otimes e'_j\}$ spans $F \otimes_R F'$ as an R-module.

To show this spanning set is linearly independent, suppose $\sum_{i,j} c_{ij} e_i \otimes e'_j = 0$, where all but finitely many c_{ij} are 0. We want to show every c_{ij} is 0. Pick two basis vectors e_{i_0} and e'_{j_0} in F and F'. To show the coefficient $c_{i_0 j_0}$ is 0, consider the *bilinear* function $F \times F' \to R$ by $(v, w) \mapsto v_{i_0} w_{j_0}$, where $v = \sum_i v_i e_i$ and $w = \sum_j w_j e'_j$. (Here v_i and w_j are coordinates in R.) By the universal mapping property of tensor products there is a linear map $f_0 \colon F \otimes_R F' \to R$ such that $f_0(v \otimes w) = v_{i_0} w_{j_0}$ on each elementary tensor $v \otimes w$.

$$\begin{tikzcd} & F \otimes_R F' \arrow[dd, "f_0"] \\ F \times F' \arrow[ur, "\otimes"] \arrow[dr, "(v,w)\mapsto a_{i_0} b_{j_0}"'] & \\ & R \end{tikzcd}$$

In particular, $f_0(e_{i_0} \otimes e'_{j_0}) = 1$ and $f_0(e_i \otimes e'_j) = 0$ for $(i, j) \neq (i_0, j_0)$. Applying f_0 to the equation $\sum_{i,j} c_{ij} e_i \otimes e'_j = 0$ in $F \otimes_R F'$ tells us $c_{i_0 j_0} = 0$ in R. Since i_0 and j_0 are arbitrary, all the coefficients are 0. $\square$

Theorem 4.9 can be interpreted in terms of bilinear maps out of $F \times F'$. It says that all bilinear maps out of $F \times F'$ are determined by their values on the pairs (e_i, e'_j), *and* that each assignment of values to these pairs extends in a unique way to a bilinear map out of $F \times F'$. (The uniqueness of the extension is connected to the linear independence of the elementary tensors $e_i \otimes e'_j$.) This is the bilinear analogue of the existence and uniqueness of a linear extension of a function from a basis of a free module to the whole module.

Example 4.10. Let K be a field and V and W be nonzero vector spaces over K with finite dimension. There are bases for V and W, say $\{e_1, \ldots, e_m\}$ for V and $\{f_1, \ldots, f_n\}$ for W. Every element of $V \otimes_K W$ can be written in the form $\sum_{i,j} c_{ij} e_i \otimes f_j$ for unique $c_{ij} \in K$.

[11] Each part of an elementary tensor in $M \otimes_R N$ belongs to M or to N.

In fact, this holds even for infinite-dimensional vector spaces, since Theorem 4.9 had no assumption that bases were finite. This justifies the basis-dependent description of tensor products of vector spaces used by physicists. on the first page.

Example 4.11. Let F be a finite free R-module of rank $n \geq 2$ with a basis $\{e_1, \ldots, e_n\}$. In $F \otimes_R F$, the tensor $e_1 \otimes e_1 + e_2 \otimes e_2$ is an **example of a tensor that is provably not an elementary tensor**. Any elementary tensor in $F \otimes_R F$ has the form

$$\sum_{i=1}^{n} a_i e_i \otimes \sum_{j=1}^{n} b_j e_j = \sum_{i,j=1}^{n} a_i b_j e_i \otimes e_j. \tag{4.1}$$

We know that the set of all $e_i \otimes e_j$ is a basis of $F \otimes_R F$, so if (4.1) equals $e_1 \otimes e_1 + e_2 \otimes e_2$ then comparing coefficients implies

$$a_1 b_1 = 1, \quad a_1 b_2 = 0, \quad a_2 b_1 = 0, \quad a_2 b_2 = 1.$$

Since $a_1 b_1 = 1$ and $a_2 b_2 = 1$, a_1 and b_2 are invertible, but that contradicts $a_1 b_2 = 0$. So $e_1 \otimes e_1 + e_2 \otimes e_2$ is not an elementary tensor.[12]

Example 4.12. As an R-module, $R[X] \otimes_R R[Y]$ is free with basis $\{X^i \otimes Y^j\}_{i,j \geq 0}$, so this tensor product is isomorphic to $R[X, Y]$ as R-modules by $\sum c_{ij}(X^i \otimes Y^j) \mapsto \sum c_{ij} X^i Y^j$. More generally, $R[X_1, \ldots, X_k] \cong R[X]^{\otimes k}$ as R-modules with X_i corresponding to the tensor $1 \otimes \cdots \otimes X \otimes \cdots \otimes 1$ where X is in the ith position. The difference between ordinary products and tensor products is like the difference between multiplying one-variable polynomials as $f(T)g(T)$ and as $f(X)g(Y)$.

Example 4.13. We return to Example 2.1. For $\mathbf{v}$ and $\mathbf{w}$ in $\mathbf{R}^n$, $B(\mathbf{v}, \mathbf{w}) = \mathbf{v}\mathbf{w}^\top \in \mathrm{M}_n(\mathbf{R})$. This is $\mathbf{R}$-bilinear, so there is an $\mathbf{R}$-linear map $L \colon \mathbf{R}^n \otimes_{\mathbf{R}} \mathbf{R}^n \to \mathrm{M}_n(\mathbf{R})$ where $L(\mathbf{v} \otimes \mathbf{w}) = \mathbf{v}\mathbf{w}^\top$ for all elementary tensors $\mathbf{v} \otimes \mathbf{w}$. In Example 2.1 we saw that, for $n \geq 2$, the image of B in $\mathrm{M}_n(\mathbf{R})$ is not closed under addition. In particular, $B(e_1, e_1) + B(e_2, e_2)$ is not of the form $B(\mathbf{v}, \mathbf{w})$. This is a typical "problem" with bilinear maps. However, using tensor products, $B(e_1, e_1) + B(e_2, e_2) = L(e_1 \otimes e_1) + L(e_2 \otimes e_2) = L(e_1 \otimes e_1 + e_2 \otimes e_2)$, which is a value of L.

In fact, L is an isomorphism. To prove this we use bases. By Theorem 4.9, $\mathbf{R}^n \otimes_{\mathbf{R}} \mathbf{R}^n$ has the basis $\{e_i \otimes e_j\}$. The value $L(e_i \otimes e_j) = e_i e_j^\top$ is the matrix with 1 in the (i, j) entry and 0 elsewhere, and these matrices are the standard basis of $\mathrm{M}_n(\mathbf{R})$. Therefore L sends a basis to a basis, so it is an isomorphism of $\mathbf{R}$-vector spaces.

Theorem 4.14. *Let F be a free R-module with basis $\{e_i\}_{i \in I}$. For $k \geq 1$, the kth tensor power $F^{\otimes k}$ is free with basis $\{e_{i_1} \otimes \cdots \otimes e_{i_k}\}_{(i_1, \ldots, i_k) \in I^k}$.*

Proof. This is similar to the proof of Theorem 4.9. □

Theorem 4.15. *If M is an R-module and F is a free R-module with basis $\{e_i\}_{i \in I}$, then every element of $M \otimes_R F$ has a unique representation in the form $\sum_{i \in I} m_i \otimes e_i$, where all but finitely many m_i equal 0.*

Proof. Using M as a spanning set of M and $\{e_i\}_{i \in I}$ as a spanning set for F as R-modules, by Theorem 3.3 every element of $M \otimes_R F$ is a linear combination of elementary tensors

[12] From (4.1), a necessary condition for $\sum_{i,j=1}^{n} c_{ij} e_i \otimes e_j$ to be elementary is that $c_{ii} c_{jj} = c_{ij} c_{ji}$ for all i and j. When $R = K$ is a field this condition is also sufficient, so the elementary tensors in $K^n \otimes_K K^n$ are characterized among all tensors by polynomial equations of degree 2. For more on this, see [7].

$m_i \otimes e_i$, where $m_i \in M$. Since $r(m_i \otimes e_i) = (rm_i) \otimes e_i$, we can write every tensor in $M \otimes_R F$ as a sum of elementary tensors of the form $m_i \otimes e_i$. So we have a surjective linear map $f\colon \bigoplus_{i\in I} M \mapsto M \otimes_R F$ given by $f((m_i)_{i\in I}) = \sum_{i\in I} m_i \otimes e_i$. (All but finitely many m_i are 0, so the sum makes sense.)

To create an inverse to f, consider the function $M \times F \to \bigoplus_{i\in I} M$ where $(m, \sum_i r_i e_i) \mapsto (r_i m)_{i\in I}$. This function is bilinear (check!), so there is a linear map $g\colon M \otimes_R F \to \bigoplus_{i\in I} M$ where $g(m \otimes \sum_i r_i e_i) = (r_i m)_{i\in I}$.

To check $f(g(t)) = t$ for all t in $M \otimes_R F$, we can't expect that all tensors in $M \otimes_R F$ are elementary (an idea used in the proofs of Theorems 4.3 and 4.5), but we only need to check $f(g(t)) = t$ when t is an elementary tensor since f and g are additive and the elementary tensors additively span $M \otimes_R F$. (We will use this kind of argument *a lot* to reduce the proof of an identity involving functions of all tensors to the case of elementary tensors even though most tensors are not themselves elementary. The point is all tensors are sums of elementary tensors and the formula we want to prove will involve additive functions.) Any elementary tensor looks like $m \otimes \sum_i r_i e_i$, and

$$\begin{aligned} f\left(g\left(m \otimes \sum_{i\in I} r_i e_i\right)\right) &= f((r_i m)_{i\in I}) \\ &= \sum_{i\in I} r_i m \otimes e_i \\ &= \sum_{i\in I} m \otimes r_i e_i \\ &= m \otimes \sum_{i\in I} r_i e_i. \end{aligned}$$

These sums have finitely many terms ($r_i = 0$ for all but finitely many i), from the definition of direct sums. Thus $f(g(t)) = t$ for all $t \in M \otimes_R F$.

For the composition in the other order,

$$g(f((m_i)_{i\in I})) = g\left(\sum_{i\in I} m_i \otimes e_i\right) = \sum_{i\in I} g(m_i \otimes e_i) = \sum_{i\in I} (\dots, 0, m_i, 0, \dots) = (m_i)_{i\in I}.$$

Now that we know $M \otimes_R F \cong \bigoplus_{i\in I} M$, with $\sum_{i\in I} m_i \otimes e_i$ corresponding to $(m_i)_{i\in I}$, the uniqueness of coordinates in the direct sum implies the sum representation $\sum_{i\in I} m_i \otimes e_i$ is unique. □

Example 4.16. For a ring $S \supset R$, elements of $S \otimes_R R[X]$ have unique expressions of the form $\sum_{n\geq 0} s_n \otimes X^n$, so $S \otimes_R R[X] \cong S[X]$ as R-modules by $\sum_{n\geq 0} s_n \otimes X^n \mapsto \sum_{n\geq 0} s_n X^n$.

Remark 4.17. When f and g are additive functions you can check $f(g(t)) = t$ for all tensors t by only checking it on elementary tensors, but it would be *wrong* to think you have proved injectivity of a linear map $f\colon M \otimes_R N \to P$ by only looking at elementary tensors.[13] That is, if $f(m \otimes n) = 0 \Rightarrow m \otimes n = 0$, there is no reason to believe $f(t) = 0 \Rightarrow t = 0$ for all $t \in M \otimes_R N$, since injectivity of a linear map is *not* an additive identity.[14] This is

[13]Unless every tensor in $M \otimes_R N$ is elementary, which is usually not the case.

[14]Here's an example. Let $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C} \to \mathbf{C}$ be the $\mathbf{R}$-linear map with the effect $z \otimes w \mapsto zw$ on elementary tensors. If $z \otimes w \mapsto 0$ then $z = 0$ or $w = 0$, so $z \otimes w = 0$, but the map on $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C}$ is not injective: $1 \otimes i - i \otimes 1 \mapsto 0$ but $1 \otimes i - i \otimes 1 \neq 0$ since $1 \otimes i$ and $i \otimes 1$ belong to a basis of $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C}$ by Theorem 4.9.

the main reason that proving a linear map out of a tensor product is injective can require technique. As a special case, if you want to prove a linear map out of a tensor product is an isomorphism, it might be easier to construct an inverse map and check the composite in both orders is the identity than to show the original map is injective and surjective.

Theorem 4.18. *If M is nonzero and finitely generated then $M^{\otimes k} \neq 0$ for all k.*

Proof. Write $M = Rx_1 + \cdots + Rx_d$, where $d \geq 1$ is minimal. Set $N = Rx_1 + \cdots + Rx_{d-1}$ ($N = 0$ if $d = 1$), so $M = N + Rx_d$ and $x_d \notin N$. Set $I = \{r \in R : rx_d \in N\}$, so I is an ideal in R and $1 \notin I$, so I is a proper ideal. When we write an element m of M in the form $n + rx$ with $n \in N$ and $r \in R$, n and r may not be well-defined from m but the value of $r \bmod I$ is well-defined: if $n + rx = n' + r'x$ then $(r - r')x = n' - n \in N$, so $r \equiv r' \bmod I$. Therefore the function $M^k \to R/I$ given by

$$(n_1 + r_1x_d, \ldots, n_k + r_kx_d) \mapsto r_1 \cdots r_d \bmod I$$

is well-defined and multilinear (check!), so there is an R-linear map $M^{\otimes k} \to R/I$ such that $\underbrace{x_d \otimes \cdots \otimes x_d}_{k \text{ terms}} \mapsto 1$. That shows $M^{\otimes k} \neq 0$. □

Example 4.19. By Theorem 4.1, $(\mathbf{Z}/a\mathbf{Z})^{\otimes 2} \cong \mathbf{Z}/a\mathbf{Z}$ as $\mathbf{Z}$-modules.

Example 4.20. Tensor powers of a non-finitely generated module could vanish: $(\mathbf{Q}/\mathbf{Z})^{\otimes 2} = 0$ as a $\mathbf{Z}$-module (Example 3.6). This example is interesting because $\mathbf{Q}/\mathbf{Z}$ is the union of cyclic subgroups $(1/a)\mathbf{Z}/\mathbf{Z}$ for all $a \geq 1$, and each $(1/a)\mathbf{Z}/\mathbf{Z}$ has a nonzero tensor square: $((1/a)\mathbf{Z}/\mathbf{Z})^{\otimes 2} \cong (1/a)\mathbf{Z}/\mathbf{Z}$ by an argument like the one used to prove Theorem 4.1. That $((1/a)\mathbf{Z}/\mathbf{Z})^{\otimes 2} \neq 0$ while $(\mathbf{Q}/\mathbf{Z})^{\otimes 2} = 0$ reflects something about bilinear maps: there are $\mathbf{Z}$-bilinear maps out of $(1/a)\mathbf{Z}/\mathbf{Z} \times (1/a)\mathbf{Z}/\mathbf{Z}$ that are not identically 0, but every $\mathbf{Z}$-bilinear map out of $\mathbf{Q}/\mathbf{Z} \times \mathbf{Q}/\mathbf{Z}$ is identically 0. For example, the $\mathbf{Z}$-bilinear map $(1/5)\mathbf{Z}/\mathbf{Z} \times (1/5)\mathbf{Z}/\mathbf{Z} \to \mathbf{Z}/5\mathbf{Z}$ given by $(x/5, y/5) \mapsto xy \bmod 5$ is nonzero at $(1/5, 1/5)$, but each bilinear map B out of $\mathbf{Q}/\mathbf{Z} \times \mathbf{Q}/\mathbf{Z}$ must vanish at $(1/5, 1/5)$ since $B(1/5, 1/5) = B(1/5, 5/25) = B(5/5, 1/25) = B(1, 1/25) = B(0, 1/25) = 0$. Thus $(1/5) \otimes (1/5) \neq 0$ in $((1/5)\mathbf{Z}/\mathbf{Z})^{\otimes 2}$ while $(1/5) \otimes (1/5) = 0$ in $(\mathbf{Q}/\mathbf{Z})^{\otimes 2}$. The lesson is that an elementary tensor requires context (which tensor product module is it in?).

The rest of this section is concerned with properties of tensor products over domains.

Theorem 4.21. *Let R be a domain with fraction field K and V be a K-vector space. There is an R-module isomorphism $K \otimes_R V \cong V$, where $x \otimes v \mapsto xv$.*

By Theorem 4.5, $K \otimes_K V \cong V$ by $x \otimes v \mapsto xv$, but Theorem 4.21 is different because the scalars in the tensor product are from R.

Proof. Multiplication is a function $K \times V \to V$. It is R-bilinear, so the universal mapping property of tensor products says there is an R-linear function $f \colon K \otimes_R V \to V$ where $f(x \otimes v) = xv$ on elementary tensors. That says the diagram

$$\begin{tikzcd}
 & K \otimes_R V \arrow[dd, "f"] \\
K \times V \arrow[ur, "\otimes"] \arrow[dr] & \\
 & V
\end{tikzcd}$$

commutes, where the lower diagonal map is scalar multiplication. Since $f(1 \otimes v) = v$, f is onto.

To show f is one-to-one, first we show every tensor in $K \otimes_R V$ is elementary with 1 in the first component. For an elementary tensor $x \otimes v$, write $x = a/b$ with a and b in R, and $b \neq 0$. Then

$$x \otimes v = \frac{a}{b} \otimes v = \frac{1}{b} \otimes av = \frac{1}{b} \otimes \frac{ab}{b}v = \frac{1}{b}b \otimes \frac{a}{b}v = 1 \otimes \frac{a}{b}v = 1 \otimes xv.$$

Notice how we moved $x \in K$ across $\otimes$ even though x need not be in R: we used K-scaling in V to *create* b and $1/b$ on the right side of $\otimes$ and bring b across $\otimes$ from right to left, which cancels $1/b$ on the left side of $\otimes$. This has the effect of moving $1/b$ from left to right.

Thus all elementary tensors in $K \otimes_R V$ have the form $1 \otimes v$ for some $v \in V$, so by adding, every tensor is $1 \otimes v$ for some v. Now we can show f has trivial kernel: if $f(t) = 0$ then, writing $t = 1 \otimes v$, we get $v = 0$, so $t = 1 \otimes 0 = 0$. □

Example 4.22. Taking $V = K$, we get $K \otimes_R K \cong K$ as R-modules by $x \otimes y \mapsto xy$ on elementary tensors. In particular, $\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Q} \cong \mathbf{Q}$ and $\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{R} \cong \mathbf{R}$ as $\mathbf{Z}$-modules.

Theorem 4.23. *Let R be a domain with fraction field K and V be a K-vector space. For each nonzero R-module M inside K, $M \otimes_R V \cong V$ as R-modules by $m \otimes v \mapsto mv$. In particular, $I \otimes_R K \cong K$ as R-modules for every nonzero ideal I in R.*

Proof. The proof is largely like that for the previous theorem.[15] Multiplication gives a function $M \times V \to V$ that is R-bilinear, so we get an R-linear map $f\colon M \otimes_R V \to V$ where $f(m \otimes v) = mv$. To show f is onto, we can't look at $f(1 \otimes v)$ as in the previous proof, since 1 is usually not in M. Instead we can just *pick* a nonzero $m \in M$. Then for all $v \in V$, $f(m \otimes (1/m)v) = v$.

To show f is injective, first we show all tensors in $M \otimes_R V$ are elementary. This sounds like our previous proof that all tensors in $K \otimes_R V$ are elementary, but M need not be K, so our manipulations need to be more careful than before. (We can't write $(a/b) \otimes v$ as $(1/b) \otimes av$, since $1/b$ usually won't be in M.) Given a finite set of nonzero elementary tensors $m_i \otimes v_i$, each m_i is nonzero. Write $m_i = a_i/b_i$ with nonzero a_i and b_i in R. Let $a \in R$ be the product of the a_i's and $c_i = a/a_i \in R$, so $a = a_i c_i = b_i c_i m_i \in M$. In V we can write $v_i = b_i c_i w_i$ for some $w_i \in V$, so

$$m_i \otimes v_i = m_i \otimes b_i c_i w_i = b_i c_i m_i \otimes w_i = a \otimes w_i.$$

The sum of these elementary tensors is $a \otimes \sum_i w_i$, which is elementary.

Now suppose $t \in M \otimes_R V$ is in the kernel of f. All tensors in $M \otimes_R V$ are elementary, so we can write $t = m \otimes v$. Then $f(t) = 0 \Rightarrow mv = 0$ in V, so $m = 0$ or $v = 0$, and thus $t = m \otimes v = 0$. □

Example 4.24. Let $R = \mathbf{Z}[\sqrt{10}]$ and $K = \mathbf{Q}(\sqrt{10})$. The ideal $I = (2, \sqrt{10})$ in R is not principal, so $I \not\cong R$ as R-modules. However, $I \otimes_R K \cong R \otimes_R K$ as R-modules since both are isomorphic to K.

Theorem 4.25. *Let R be a domain and F and F' be free R-modules. If x and x' are nonzero in F and F', then $x \otimes x' \neq 0$ in $F \otimes_R F'$.*

---

[15] Theorem 4.21 is just a special case of Theorem 4.23, but we worked it out separately first since the technicalities are simpler.

Proof. If we were working with vector spaces this would be trivial, since x and x' are each part of a basis of F and F', so $x \otimes x'$ is part of a basis of $F \otimes_R F'$ (Theorem 4.9). In a free module over a commutative ring, a nonzero element need not be part of a basis, so our proof needs to be a little more careful. We'll still uses bases, just not ones that necessarily include x or x'.

Pick a basis $\{e_i\}$ for F and $\{e'_j\}$ for F'. Write $x = \sum_i a_i e_i$ and $x' = \sum_j a'_j e'_j$. Then $x \otimes x' = \sum_{i,j} a_i a'_j e_i \otimes e'_j$ in $F \otimes_R F'$. Since x and x' are nonzero, they each have some nonzero coefficient, say a_{i_0} and a'_{j_0}. Then $a_{i_0} a'_{j_0} \neq 0$ since R is a domain, so $x \otimes x'$ has a nonzero coordinate in the basis $\{e_i \otimes e'_j\}$ of $F \otimes_R F'$. Thus $x \otimes x' \neq 0$. □

Remark 4.26. There is always a counterexample for Theorem 4.25 when R is not a domain. Let $F = F' = R$ and say $ab = 0$ with a and b nonzero in R. In $R \otimes_R R$ we have $a \otimes b = ab(1 \otimes 1) = 0$.

Theorem 4.27. *Let R be a domain with fraction field K.*

(1) *For each R-module M, $K \otimes_R M \cong K \otimes_R (M/M_{\rm tor})$ as R-modules, where $M_{\rm tor}$ is the torsion submodule of M.*
(2) *If M is a torsion R-module then $K \otimes_R M = 0$ and if M is not a torsion module then $K \otimes_R M \neq 0$.*
(3) *If N is a submodule of M such that M/N is a torsion module then $K \otimes_R N \cong K \otimes_R M$ as R-modules by $x \otimes n \mapsto x \otimes n$.*

Proof. (1) The map $K \times M \to K \otimes_R (M/M_{\rm tor})$ given by $(x, m) \mapsto x \otimes \overline{m}$ is R-bilinear, so there is a linear map $f \colon K \otimes_R M \to K \otimes_R (M/M_{\rm tor})$ where $f(x \otimes m) = x \otimes \overline{m}$.

To go the other way, the canonical bilinear map $K \times M \xrightarrow{\otimes} K \otimes_R M$ vanishes at (x, m) if $m \in M_{\rm tor}$: when $rm = 0$ for $r \neq 0$, $x \otimes m = r(x/r) \otimes m = x/r \otimes rm = x/r \otimes 0 = 0$. Therefore we get an induced bilinear map $K \times (M/M_{\rm tor}) \to K \otimes_R M$ given by $(x, \overline{m}) \mapsto x \otimes m$. (The point is that an elementary tensor $x \otimes m$ in $K \otimes_R M$ only depends on m through its coset mod $M_{\rm tor}$.) The universal mapping property of the tensor product now gives us a linear map $g \colon K \otimes_R (M/M_{\rm tor}) \to K \otimes_R M$ where $g(x \otimes \overline{m}) = x \otimes m$.

The composites $g \circ f$ and $f \circ g$ are both linear and fix elementary tensors, so they fix all tensors and thus f and g are inverse isomorphisms.

(2) It is immediate from (1) that $K \otimes_R M = 0$ if M is a torsion module, since $K \otimes_R M \cong K \otimes_R (M/M_{\rm tor}) = K \otimes_R 0 = 0$. We could also prove this in a direct way, by showing all elementary tensors in $K \otimes_R M$ are 0: for $x \in K$ and $m \in M$, let $rm = 0$ with $r \neq 0$, so $x \otimes m = r(x/r) \otimes m = x/r \otimes rm = x/r \otimes 0 = 0$.

To show $K \otimes_R M \neq 0$ when M is not a torsion module, from the isomorphism $K \otimes_R M \cong K \otimes_R (M/M_{\rm tor})$, we may replace M with $M/M_{\rm tor}$ and are reduced to the case when M *is torsion-free*. For torsion-free M we will create a nonzero R-module and a bilinear map onto it from $K \times M$. This will require a fair bit of work (as it usually does to prove a tensor product doesn't vanish if you don't have bases available).

We want to consider formal products xm with $x \in K$ and $m \in M$. To make this precise, we will use equivalence classes of ordered pairs in the same way that a fraction field is created out of a domain. On the product set $K \times M$, define an equivalence relation by

$$(a/b, m) \sim (c/d, n) \iff adm = bcn \text{ in } M.$$

Here a, b, c, and d are in R and b and d are not 0. The proof that this relation is well-defined (independent of the choice of numerators and denominators) and transitive requires M be torsion-free (check!). As an example, $(0, m) \sim (0, 0)$ for all $m \in M$.

Define $KM = (K \times M)/\sim$ and write the equivalence class of (x, m) as $x \cdot m$. Give KM the addition and K-scaling formulas

$$\frac{a}{b} \cdot m + \frac{c}{d} \cdot n = \frac{1}{bd} \cdot (adm + bcn), \quad x(y \cdot m) = (xy) \cdot m.$$

It is left to the reader to check these operations on KM are well-defined and make KM into a K-vector space (so in particular an R-module). The zero element of KM is $0 \cdot 0 = 0 \cdot m$. The function $M \to KM$ given by $m \mapsto 1 \cdot m$ is *injective*: if $1 \cdot m = 1 \cdot m'$ then $(1, m) \sim (1, m')$, so $m = m'$ in M. Thus $KM \neq 0$ since $M \neq 0$.

The function $K \times M \to KM$ given by $(x, m) \mapsto x \cdot m$ is R-bilinear and onto, so there is a linear map $K \otimes_R M \xrightarrow{f} KM$ such that $f(x \otimes m) = x \cdot m$, which is onto. Since $KM \neq 0$ we have $K \otimes_R M \neq 0$, and in fact $K \otimes_R M \cong KM$ by the map f (exercise).

(3) Since $N \subset M$, there is an obvious bilinear map $K \times N \to K \otimes_R M$, namely $(x, n) \mapsto x \otimes n$. So we get automatically a linear map $f \colon K \otimes_R N \to K \otimes_R M$ where $f(x \otimes n) = x \otimes n$. (This is not the identity: on the left $x \otimes n$ is in $K \otimes_R N$ and on the right $x \otimes n$ is in $K \otimes_R M$.)

To get a map inverse to f, we can't have $K \times M \to K \otimes_R N$ by $(x, m) \mapsto x \otimes m$, since m may not be in N. The trick to use is that some nonzero R-multiple of m is in N, since M/N is a torsion module: for some $r \in R - \{0\}$ we have $rm \in N$. Let $(x, m) \mapsto (1/r)x \otimes rm$ in $K \otimes_R N$. (Don't try to simplify $(1/r)x \otimes rm$ by moving r through $\otimes$ from right to left, since rm is in N but m usually is not.) We need to check that $(1/r)x \otimes rm$ is independent of the choice of r such that $rm \in N$. If also $r'm \in N$ with $r' \in R - \{0\}$, then

$$\frac{1}{r'}x \otimes r'm = \frac{r}{rr'}x \otimes r'm = \frac{1}{rr'}x \otimes rr'm = \frac{1}{rr'}x \otimes r'(rm) = \frac{r'}{rr'}x \otimes rm = \frac{1}{r}x \otimes rm.$$

So the function $K \times M \to K \otimes_R N$ where $(x, m) \mapsto (1/r)x \otimes rm$ is well-defined, and the reader can check it is bilinear. It leads to a linear map $g \colon K \otimes_R M \to K \otimes_R N$ where $g(x \otimes m) = (1/r)x \otimes rm$ when $rm \in N$, $r \neq 0$. Check $f(g(x \otimes m)) = x \otimes m$ and $g(f(x \otimes n)) = x \otimes n$, so $f \circ g$ and $g \circ f$ are both the identity by additivity. $\square$

Corollary 4.28. *Let R be a domain with fraction field K. In $K \otimes_R M$, $x \otimes m = 0$ if and only if $x = 0$ or $m \in M_{\mathrm{tor}}$. In particular, $M_{\mathrm{tor}} = \ker(M \to K \otimes_R M)$ where $m \mapsto 1 \otimes m$.*

Proof. If $x = 0$ then $x \otimes m = 0 \otimes m = 0$. If $m \in M_{\mathrm{tor}}$, with $rm = 0$ for some nonzero $r \in R$, then $x \otimes m = (x/r)r \otimes m = (x/r) \otimes rm = (x/r) \otimes 0 = 0$.

Conversely, suppose $x \otimes m = 0$. We want to show $x = 0$ or $m \in M_{\mathrm{tor}}$. Write $x = a/b$, so $(1/b) \otimes am = 0$. If $a = 0$ then $x = 0$, so we suppose $a \neq 0$ and will show $m \in M_{\mathrm{tor}}$. Multiply $(1/b) \otimes am$ by b to get $1 \otimes am = 0$. From the isomorphism $K \otimes_R M \cong K \otimes_R (M/M_{\mathrm{tor}})$, $1 \otimes \overline{am} = 0$ in $K \otimes_R (M/M_{\mathrm{tor}})$. Since M/M_{tor} is torsion-free, applying the R-linear map $K \otimes_R (M/M_{\mathrm{tor}}) \to K(M/M_{\mathrm{tor}})$ from the proof of Theorem 4.27 tells us that $1 \cdot \overline{am} = 0$ in $K(M/M_{\mathrm{tor}})$. The function $\overline{m} \mapsto 1 \cdot \overline{m}$ from M/M_{tor} to $K(M/M_{\mathrm{tor}})$ is injective, so $\overline{am} = 0$, so $am \in M_{\mathrm{tor}}$. Therefore there is nonzero $r \in R$ such that $0 = r(am) = (ra)m$. Since $ra \neq 0$, $m \in M_{\mathrm{tor}}$. $\square$

Example 4.29. The tensor product $\mathbf{Q} \otimes_{\mathbf{Z}} A$ is 0 when A is a torsion abelian group, so we recover Example 3.6.

5. General Properties of Tensor Products

There are canonical isomorphisms $M \oplus N \cong N \oplus M$ and $(M \oplus N) \oplus P \cong M \oplus (N \oplus P)$. We want to show similar isomorphisms for tensor products: $M \otimes_R N \cong N \otimes_R M$ and

$(M \otimes_R N) \otimes_R P \cong M \otimes_R (N \otimes_R P)$. Furthermore, there is a distributive property over direct sums: $M \otimes_R (N \oplus P) \cong (M \otimes_R N) \oplus (M \otimes_R P)$. *How* these modules are isomorphic is much more important than just that they are isomorphic.

Theorem 5.1. *There is a unique R-module isomorphism $M \otimes_R N \cong N \otimes_R M$ where $m \otimes n \mapsto n \otimes m$.*

Proof. We want to create a linear map $M \otimes_R N \to N \otimes_R M$ sending $m \otimes n$ to $n \otimes m$. To do this, we back up and start off with a map out of $M \times N$ to the desired target module $N \otimes_R M$. Define $M \times N \to N \otimes_R M$ by $(m, n) \mapsto n \otimes m$. This is a bilinear map since $n \otimes m$ is bilinear in m and n. Therefore by the universal mapping property of the tensor product, there is a unique linear map $f\colon M \otimes_R N \to N \otimes_R M$ such that $f(m \otimes n) = n \otimes m$ on elementary tensors: the diagram

$$\begin{tikzcd}
 & M \otimes_R N \arrow[dd, "f"] \\
M \times N \arrow[ur, "\otimes"] \arrow[dr, "(m,n)\mapsto n\otimes m"'] & \\
 & N \otimes_R M
\end{tikzcd}$$

commutes.

Running through the above argument with the roles of M and N interchanged, there is a unique linear map $g\colon N \otimes_R M \to M \otimes_R N$ where $g(n \otimes m) = m \otimes n$ on elementary tensors. We will show f and g are inverses of each other.

To show $f(g(t)) = t$ for all $t \in N \otimes_R M$, it suffices to check this when t is an elementary tensor, since both sides are R-linear (or even just additive) in t and $N \otimes_R M$ is spanned by its elementary tensors: $f(g(n \otimes m)) = f(m \otimes n) = n \otimes m$. Therefore $f(g(t)) = t$ for all $t \in N \otimes_R M$. The proof that $g(f(t)) = t$ for all $t \in M \otimes_R N$ is similar. We have shown f and g are inverses of each other, so f is an R-module isomorphism. □

Theorem 5.2. *There is a unique R-module isomorphism $(M \otimes_R N) \otimes_R P \cong M \otimes_R (N \otimes_R P)$ where $(m \otimes n) \otimes p \mapsto m \otimes (n \otimes p)$.*

Proof. By Theorem 3.3, $(M \otimes_R N) \otimes_R P$ is linearly spanned by all $(m \otimes n) \otimes p$ and $M \otimes_R (N \otimes_R P)$ is linearly spanned by all $m \otimes (n \otimes p)$. Therefore linear maps out of these two modules are determined by their values on these[16] elementary tensors. So there is at most one linear map $(M \otimes_R N) \otimes_R P \to M \otimes_R (N \otimes_R P)$ with the effect $(m \otimes n) \otimes p \mapsto m \otimes (n \otimes p)$, and likewise in the other direction.

To create such a linear map $(M \otimes_R N) \otimes_R P \to M \otimes_R (N \otimes_R P)$, consider the function $M \times N \times P \to M \otimes_R (N \otimes_R P)$ given by $(m, n, p) \mapsto m \otimes (n \otimes p)$. Since $m \otimes (n \otimes p)$ is *trilinear* in m, n, and p, for each p we get a *bilinear* map $b_p\colon M \times N \to M \otimes_R (N \otimes_R P)$ where $b_p(m, n) = m \otimes (n \otimes p)$, which induces a *linear* map $f_p\colon M \otimes_R N \to M \otimes_R (N \otimes_R P)$ such that $f_p(m \otimes n) = m \otimes (n \otimes p)$ on all elementary tensors $m \otimes n$ in $M \otimes_R N$.

Now we consider the function $(M \otimes_R N) \times P \to M \otimes_R (N \otimes_R P)$ given by $(t, p) \mapsto f_p(t)$. This is bilinear! First, it is linear in t with p fixed, since each f_p is a linear function. Next

[16] A general elementary tensor in $(M \otimes_R N) \otimes_R P$ is *not* $(m \otimes n) \otimes p$, but $t \otimes p$ where $t \in M \otimes_R N$ and t might not be elementary itself. Similarly, elementary tensors in $M \otimes_R (N \otimes_R P)$ are more general than $m \otimes (n \otimes p)$.

we show it is linear in p with t fixed:

$$f_{p+p'}(t) = f_p(t) + f_{p'}(t) \text{ and } f_{rp}(t) = rf_p(t)$$

for all p, p', and r. Both sides of these identities are additive in t, so to check them it suffices to check the case when $t = m \otimes n$:

$$\begin{aligned} f_{p+p'}(m \otimes n) &= (m \otimes n) \otimes (p + p') \\ &= (m \otimes n) \otimes p + (m \otimes n) \otimes p' \\ &= f_p(m \otimes n) + f_{p'}(m \otimes n) \\ &= (f_p + f_{p'})(m \otimes n). \end{aligned}$$

That $f_{rp}(m \otimes n) = rf_p(m \otimes n)$ is left to the reader. Since $f_p(t)$ is bilinear in p and t, the universal mapping property of the tensor product tells us there is a unique linear map $f\colon (M \otimes_R N) \otimes_R P \to M \otimes_R (N \otimes_R P)$ such that $f(t \otimes p) = f_p(t)$. Then $f((m \otimes n) \otimes p) = f_p(m \otimes n) = m \otimes (n \otimes p)$, so we have found a linear map with the desired values on the tensors $(m \otimes n) \otimes p$.

Similarly, there is a linear map $g\colon M \otimes_R (N \otimes_R P) \to (M \otimes_R N) \otimes_R P$ where $g(m \otimes (n \otimes p)) = (m \otimes n) \otimes p$. Easily $f(g(m \otimes (n \otimes p))) = m \otimes (n \otimes p)$ and $g(f((m \otimes n) \otimes p)) = (m \otimes n) \otimes p$. Since these particular tensors linearly span the two modules, these identities extend by linearity (f and g are *linear*) to show f and g are inverse functions. $\square$

Theorem 5.3. *There is a unique R-module isomorphism*

$$M \otimes_R (N \oplus P) \cong (M \otimes_R N) \oplus (M \otimes_R P)$$

where $m \otimes (n, p) \mapsto (m \otimes n, m \otimes p)$.

Proof. Instead of directly writing down an isomorphism, we will put to work the essential uniqueness of solutions to a universal mapping problem by showing $(M \otimes_R N) \oplus (M \otimes_R P)$ has the universal mapping property of the tensor product $M \otimes_R (N \oplus P)$. Therefore by abstract nonsense these modules must be isomorphic. That there is an isomorphism whose effect on elementary tensors in $M \otimes_R (N \oplus P)$ is as indicated in the statement of the theorem will fall out of our work.

For $(M \otimes_R N) \oplus (M \otimes_R P)$ to be a tensor product of M and $N \oplus P$, it needs a bilinear map to it from $M \times (N \oplus P)$. Let $b\colon M \times (N \oplus P) \to (M \otimes_R N) \oplus (M \otimes_R P)$ by $b(m, (n, p)) = (m \otimes n, m \otimes p)$. This function is bilinear. We verify the additivity of b in its second component, leaving the rest to the reader:

$$\begin{aligned} b(m, (n, p) + (n', p')) &= b(m, (n + n', p + p')) \\ &= (m \otimes (n + n'), m \otimes (p + p')) \\ &= (m \otimes n + m \otimes n', m \otimes p + m \otimes p') \\ &= (m \otimes n, m \otimes p) + (m \otimes n', m \otimes p') \\ &= b(m, (n, p)) + b(m, (n', p')). \end{aligned}$$

To show $(M\otimes_R N)\oplus(M\otimes_R P)$ and b have the universal mapping property of $M\otimes_R(N\oplus P)$ and $\otimes$, let $B\colon M\times(N\oplus P)\to Q$ be a bilinear map. We seek an R-linear map L making

(5.1)
$$\begin{tikzcd} & (M\otimes_R N)\oplus(M\otimes_R P) \arrow[d, "L"] \\ M\times(N\oplus P) \arrow[ur, "b"] \arrow[r, "B"'] & Q \end{tikzcd}$$

commute. Being linear, L would be determined by its values on the direct summands, and these values would be determined by the values of L on all pairs $(m\otimes n,0)$ and $(0,m\otimes p)$ by additivity. These values are forced by commutativity of (5.1) to be

$$L(m\otimes n,0)=L(b(m,(n,0)))=B(m,(n,0)) \text{ and } L(0,m\otimes p)=L(b(m,(0,p)))=B(m,(0,p)).$$

To construct L, the above formulas suggest the maps $M\times N\to Q$ and $M\times P\to Q$ given by $(m,n)\mapsto B(m,(n,0))$ and $(m,p)\mapsto B(m,(0,p))$. Both are bilinear, so there are R-linear maps $M\otimes_R N\xrightarrow{L_1}Q$ and $M\otimes_R P\xrightarrow{L_2}Q$ where

$$L_1(m\otimes n)=B(m,(n,0)) \text{ and } L_2(m\otimes p)=B(m,(0,p)).$$

Define L on $(M\otimes_R N)\oplus(M\otimes_R P)$ by $L(t_1,t_2)=L_1(t_1)+L_2(t_2)$. (Notice we are defining L not just on ordered pairs of elementary tensors, but on *all* pairs of tensors. We need L_1 and L_2 to be defined on the whole tensor product modules $M\otimes_R N$ and $M\otimes_R P$.) The map L is linear since L_1 and L_2 are linear, and (5.1) commutes:

$$\begin{aligned}
L(b(m,(n,p))) &= L(b(m,(n,0)+(0,p)))\\
&= L(b(m,(n,0))+b(m,(0,p)))\\
&= L((m\otimes n,0)+(0,m\otimes p)) \text{ by the definition of } b\\
&= L(m\otimes n,m\otimes p)\\
&= L_1(m\otimes n)+L_2(m\otimes p) \text{ by the definition of } L\\
&= B(m,(n,0))+B(m,(0,p))\\
&= B(m,(n,0)+(0,p))\\
&= B(m,(n,p)).
\end{aligned}$$

Now that we've shown $(M\otimes_R N)\oplus(M\otimes_R P)$ and the bilinear map b have the universal mapping property of $M\otimes_R(N\oplus P)$ and the canonical bilinear map $\otimes$, there is a unique linear map f making the diagram

$$\begin{tikzcd} & (M\otimes_R N)\oplus(M\otimes_R P) \arrow[d, "f"] \\ M\times(N\oplus P) \arrow[ur, "b"] \arrow[r, "\otimes"'] & M\otimes_R(N\oplus P) \end{tikzcd}$$

commute, and f is an isomorphism of R-modules because it transforms one solution of a universal mapping problem into another. Taking $(m,(n,p))$ around the diagram both ways,

$$f(b(m,(n,p))) = f(m\otimes n, m\otimes p) = m\otimes(n,p).$$

Therefore the inverse of f is an isomorphism $M\otimes_R(N\oplus P)\to(M\otimes_R N)\oplus(M\otimes_R P)$ with the effect $m\otimes(n,p)\mapsto(m\otimes n, m\otimes p)$. We look at the inverse because the theorem is saying something about an isomorphism out of $M\otimes_R(N\oplus P)$, which is the target of f. $\square$

Theorem 5.4. *There is a unique R-module isomorphism*

$$M\otimes_R\bigoplus_{i\in I}N_i\cong\bigoplus_{i\in I}(M\otimes_R N_i)$$

where $m\otimes(n_i)_{i\in I}\mapsto(m\otimes n_i)_{i\in I}$.

Proof. We extrapolate from the case $\#I=2$ in Theorem 5.3. The map $b\colon M\times(\bigoplus_{i\in I}N_i)\to\bigoplus_{i\in I}(M\otimes_R N_i)$ by $b((m,(n_i)_{i\in I}))=(m\otimes n_i)_{i\in I}$ is bilinear. We will show $\bigoplus_{i\in I}(M\otimes_R N_i)$ and b have the universal mapping property of $M\otimes_R\bigoplus_{i\in I}N_i$ and $\otimes$.

Let $B\colon M\times(\bigoplus_{i\in I}N_i)\to Q$ be bilinear. For each $i\in I$ the function $M\times N_i\to Q$ where $(m,n_i)\mapsto B(m,(\dots,0,n_i,0,\dots))$ is bilinear, so there is a linear map $L_i\colon M\otimes_R N_i\to Q$ where $L_i(m\otimes n_i)=B(m,(\dots,0,n_i,0,\dots))$. Define $L\colon\bigoplus_{i\in I}(M\otimes_R N_i)\to Q$ by $L((t_i)_{i\in I})=\sum_{i\in I}L_i(t_i)$. All but finitely many t_i equal 0, so the sum here makes sense, and L is linear. It is left to the reader to check the diagram

$$\begin{tikzcd} & \bigoplus_{i\in I}(M\otimes_R N_i) \arrow[dd, "L"] \\ M\times\bigoplus_{i\in I}N_i \arrow[ur, "b"] \arrow[dr, "B"'] & \\ & Q \end{tikzcd}$$

commutes. A map L making this diagram commute has its value on $(\dots,0,m\otimes n_i,0,\dots)=b(m,(\dots,0,n_i,0,\dots))$ determined by B, so L is unique. Thus $\bigoplus_{i\in I}(M\otimes_R N_i)$ and the bilinear map b to it have the universal mapping property of $M\otimes_R\bigoplus_{i\in I}N_i$ and the canonical map $\otimes$, so there is an R-module isomorphism f making the diagram

$$\begin{tikzcd} & \bigoplus_{i\in I}(M\otimes_R N_i) \arrow[dd, "f"] \\ M\times\bigoplus_{i\in I}N_i \arrow[ur, "b"] \arrow[dr, "\otimes"'] & \\ & M\otimes_R\bigoplus_{i\in I}N_i \end{tikzcd}$$

commute. Sending $(m,(n_i)_{i\in I})$ around the diagram both ways, $f((m\otimes n_i)_{i\in I})=m\otimes(n_i)_{i\in I}$, so the inverse of f is an isomorphism with the effect $m\otimes(n_i)_{i\in I}\mapsto(m\otimes n_i)_{i\in I}$. $\square$

Remark 5.5. The analogue of Theorem 5.4 for direct products is false. While there is a natural R-linear map

$$M \otimes_R \prod_{i\in I} N_i \to \prod_{i\in I}(M \otimes_R N_i) \tag{5.2}$$

where $m \otimes (n_i)_{i\in I} \mapsto (m \otimes n_i)_{i\in I}$, it may not be an isomorphism. Taking $R = \mathbf{Z}$, $M = \mathbf{Q}$, and $N_i = \mathbf{Z}/p^i\mathbf{Z}$ ($i \geq 1$), the right side of (5.2) is 0 since $\mathbf{Q} \otimes_{\mathbf{Z}} (\mathbf{Z}/p^i\mathbf{Z}) = 0$ for all $i \geq 1$ (Example 3.6). The left side of (5.2) is $\mathbf{Q} \otimes_{\mathbf{Z}} \prod_{i\geq 1} \mathbf{Z}/p^i\mathbf{Z}$, which is not 0 by Theorem 4.27 since $\prod_{i\geq 1} \mathbf{Z}/p^i\mathbf{Z}$, *unlike* $\bigoplus_{i\in I} \mathbf{Z}/p^i\mathbf{Z}$, is not a torsion abelian group.

In our proof of associativity of the tensor product, we started with a function on a direct product $M \times N \times P$ and collapsed this direct product to an iterated tensor product $(M \otimes_R N) \otimes_R P$ using bilinearity twice. It is useful to record a rather general result in that direction, as a technical lemma for future convenience.

Theorem 5.6. *Let $M_1, \dots, M_k, N$ be R-modules, with $k > 2$, and suppose*

$$M_1 \times \cdots \times M_{k-2} \times M_{k-1} \times M_k \xrightarrow{\varphi} N$$

is a function that is bilinear in M_{k-1} and M_k when other coordinates are fixed. There is a unique function

$$M_1 \times \cdots \times M_{k-2} \times (M_{k-1} \otimes_R M_k) \xrightarrow{\Phi} N$$

that is linear in $M_{k-1} \otimes_R M_k$ when the other coordinates are fixed and satisfies

$$\Phi(m_1, \dots, m_{k-2}, m_{k-1} \otimes m_k) = \varphi(m_1, \dots, m_{k-2}, m_{k-1}, m_k). \tag{5.3}$$

If φ is multilinear in $M_1, \dots, M_k$, then Φ is multilinear in $M_1, \dots, M_{k-2}$, $M_{k-1} \otimes_R M_k$.

Proof. Assuming a function Φ exists satisfying (5.3) and is linear in the last coordinate when other coordinates are fixed, its value everywhere is determined by additivity in the last coordinate: write each tensor $t \in M_{k-1} \otimes_R M_k$ in the form $t = \sum_{i=1}^p x_i \otimes y_i$, and then

$$\begin{aligned} \Phi(m_1, \dots, m_{k-2}, t) &= \Phi\left(m_1, \dots, m_{k-2}, \sum_{i=1}^p x_i \otimes y_i\right) \\ &= \sum_{i=1}^p \Phi(m_1, \dots, m_{k-2}, x_i \otimes y_i) \\ &= \sum_{i=1}^p \varphi(m_1, \dots, m_{k-2}, x_i, y_i). \end{aligned}$$

It remains to show Φ exists with the desired properties.

Fix $m_i \in M_i$ for $i = 1, \dots, k-2$. Define $\varphi_{m_1,\dots,m_{k-2}} \colon M_{k-1} \times M_k \to N$ by

$$\varphi_{m_1,\dots,m_{k-2}}(x, y) = \varphi(m_1, \dots, m_{k-2}, x, y).$$

By hypothesis $\varphi_{m_1,\dots,m_{k-2}}$ is bilinear in x and y, so from the universal mapping property of the tensor product there is a linear map $\Phi_{m_1,\dots,m_{k-2}} \colon M_{k-1} \otimes_R M_k \to N$ such that

$$\Phi_{m_1,\dots,m_{k-2}}(x \otimes y) = \varphi_{m_1,\dots,m_{k-2}}(x, y) = \varphi(m_1, \dots, m_{k-2}, x, y).$$

Define $\Phi \colon M_1 \times \cdots \times M_{k-2} \times (M_{k-1} \otimes_R M_k) \to N$ by

$$\Phi(m_1, \dots, m_{k-2}, t) = \Phi_{m_1,\dots,m_{k-2}}(t).$$

Since $\Phi_{m_1,\ldots,m_{k-2}}$ is a linear function on $M_{k-1} \otimes_R M_k$, $\Phi(m_1,\ldots,m_{k-2},t)$ is linear in t when $m_1,\ldots,m_{k-2}$ are fixed.

If φ is multilinear in $M_1,\ldots,M_k$ we want to show Φ is multilinear in $M_1,\ldots,M_{k-2}$, $M_{k-1} \otimes_R M_k$. We already know Φ is linear in $M_{k-1} \otimes_R M_k$ when the other coordinates are fixed. To show Φ is linear in each of the other coordinates (fixing the rest), we carry out the computation for M_1 (the argument is similar for other M_i's): is

$$\begin{aligned} \Phi(x+x', m_2,\ldots,m_{k-2},t) &\stackrel{?}{=} \Phi(x,m_2,\ldots,m_{k-2},t) + \Phi(x',m_2,\ldots,m_{k-2},t) \\ \Phi(rx,m_2,\ldots,m_{k-2},t) &\stackrel{?}{=} r\Phi(x,m_2,\ldots,m_{k-2},t) \end{aligned}$$

when $m_2,\ldots,m_{k-2},t$ are fixed in $M_2,\ldots,M_{k-2},M_{k-1}\otimes_R M_k$? In these two equations, both sides are additive in t so it suffices to verify these two equations when t is an elementary tensor $m_{k-1} \otimes m_k$. Then from (5.3), these two equations are true since we're assuming φ is linear in M_1 (fixing the other coordinates). $\square$

Theorem 5.6 is not specific to functions that are bilinear in the last two coordinates: any two coordinates can be used when the function is bilinear in those two coordinates. For instance, let's revisit the proof of associativity of the tensor product in Theorem 5.2 to see why the construction of the functions f_p in the proof of Theorem 5.3 is a special case of Theorem 5.6. Define

$$\varphi\colon M \times N \times P \to M \otimes_R (N \otimes_R P)$$

by $\varphi(m,n,p) = m \otimes (n \otimes p)$. This function is trilinear, so Theorem 5.6 says we can replace $M \times N$ with its tensor product: there is a bilinear function

$$\Phi\colon (M \otimes_R N) \times P \to M \otimes_R (N \otimes_R P)$$

such that $\Phi(m \otimes n, p) = m \otimes (n \otimes p)$. Since Φ is bilinear, there is a linear function

$$f\colon (M \otimes_R N) \otimes_R P \to M \otimes_R (N \otimes_R P)$$

such that $f(t \otimes p) = \Phi(t,p)$, so $f((m \otimes n) \otimes p) = \Phi(m \otimes n, p) = m \otimes (n \otimes p)$.

The remaining module properties we treat with the tensor product in this section involve its interaction with the Hom-module construction, so in particular the dual module construction ($M^\vee = \mathrm{Hom}_R(M,R)$).

Theorem 5.7. *For R-modules M, N, and P, there are R-module isomorphisms*

$$\mathrm{Bil}_R(M,N;P) \cong \mathrm{Hom}_R(M \otimes_R N, P) \cong \mathrm{Hom}_R(M, \mathrm{Hom}_R(N,P)).$$

Proof. The R-module isomorphism from $\mathrm{Bil}_R(M,N;P)$ to $\mathrm{Hom}_R(M \otimes_R N, P)$ comes from the universal mapping property of the tensor product: every bilinear map $B\colon M \times N \to P$ leads to a specific linear map $L_B\colon M \otimes_R N \to P$, and all linear maps $M \otimes_R N \to P$ arise in this way. The correspondence $B \mapsto L_B$ is an isomorphism from $\mathrm{Bil}_N(M,N;P)$ to $\mathrm{Hom}_R(M \otimes_R N, P)$.

Next we explain why $\mathrm{Bil}_R(M,N;P) \cong \mathrm{Hom}_R(M,\mathrm{Hom}_R(N,P))$, which amounts to thinking about a bilinear map $M \times N \to P$ as a family of linear maps $N \to P$ indexed by elements of M. If $B\colon M \times N \to P$ is bilinear, then for each $m \in M$ the function $B(m,-)$ is a linear map $N \to P$. Define $f_B\colon M \to \mathrm{Hom}_R(N,P)$ by $f_B(m) = B(m,-)$. (That is, $f_B(m)(n) = B(m,n)$.) Check f_B is linear, partly because B is linear in its first component when the second component is fixed.

Going in the other direction, if $L\colon M \to \mathrm{Hom}_R(N,P)$ is linear then for each $m \in M$ we have a linear function $L(m)\colon N \to P$. Define $B_L\colon M \times N \to P$ to be $B_L(m,n) = L(m)(n)$. Check B_L is bilinear.

It is left to the reader to check the correspondences $B \rightsquigarrow f_B$ and $L \rightsquigarrow B_L$ are each linear and are inverses of each other, so $\mathrm{Bil}_R(M,N;P) \cong \mathrm{Hom}_R(M, \mathrm{Hom}_R(N,P))$ as R-modules. $\square$

Here's a high-level way of interpreting the isomorphism between the second and third modules in Theorem 5.7. Write $\mathbf{F}_N(M) = M \otimes_R N$ and $\mathbf{G}_N(M) = \mathrm{Hom}_R(N,M)$, so $\mathbf{F}_N$ and $\mathbf{G}_N$ turn R-modules into new R-modules. Theorem 5.7 says

$$\mathrm{Hom}_R(\mathbf{F}_N(M), P) \cong \mathrm{Hom}_R(M, \mathbf{G}_N(P)).$$

This is analogous to the relation between a matrix A and its transpose $A^\top$ inside dot products:

$$Av \cdot w = v \cdot A^\top w$$

for all vectors v and w. So $\mathbf{F}_N$ and $\mathbf{G}_N$ are "transposes" of each other. Actually, $\mathbf{F}_N$ and $\mathbf{G}_N$ are called adjoints of each other because pairs of operators L and L' in linear algebra that satisfy the relation $Lv \cdot w = v \cdot L'w$ for all vectors v and w are called adjoints and the relation between $\mathbf{F}_N$ and $\mathbf{G}_N$ looks similar.

Corollary 5.8. *For R-modules M and N, there are R-module isomorphisms*

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

Proof. Using $P = R$ in Theorem 5.7, we get $\mathrm{Bil}_R(M,N;R) \cong (M \otimes_R N)^\vee \cong \mathrm{Hom}_R(M, N^\vee)$. From $M \otimes_R N \cong N \otimes_R M$ we get $(M \otimes_R N)^\vee \cong (N \otimes_R M)^\vee$, and the second dual module is isomorphic to $\mathrm{Hom}_R(N, M^\vee)$ by Theorem 5.7 with the roles of M and N there reversed and $P = R$. Thus we have obtained isomorphisms between the desired modules.

The isomorphism between $\mathrm{Hom}_R(M, N^\vee)$ and $\mathrm{Hom}_R(N, M^\vee)$ amounts to viewing a map in either Hom-module as as a bilinear map $B\colon M \times N \to R$. $\square$

The construction of $M \otimes_R N$ is "symmetric" in M and N in the sense that $M \otimes_R N \cong N \otimes_R M$ in a natural way, but Corollary 5.8 is *not* saying $\mathrm{Hom}_R(M,N) \cong \mathrm{Hom}_R(N,M)$ since those are not the Hom-modules in the corollary. For instance, if $R = M = \mathbf{Z}$ and $N = \mathbf{Z}/2\mathbf{Z}$ then $\mathrm{Hom}_R(M,N) \cong \mathbf{Z}/2\mathbf{Z}$ and $\mathrm{Hom}_R(N,M) = 0$.

Theorem 5.9. *For R-modules M and N, there is a linear map $M^\vee \otimes_R N \to \mathrm{Hom}_R(M,N)$ sending each elementary tensor $\varphi \otimes n$ in $M^\vee \otimes_R N$ to the linear map $M \to N$ defined by $(\varphi \otimes n)(m) = \varphi(m)n$. This is an isomorphism from $M^\vee \otimes_R N$ to $\mathrm{Hom}_R(M,N)$ if M and N are finite free. In particular, if F is finite free then $F^\vee \otimes_R F \cong \mathrm{End}_R(F)$ as R-modules.*

Proof. We need to make an element of $M^\vee$ and an element of N act together as a linear map $M \to N$. The function $M^\vee \times M \times N \to N$ given by $(\varphi, m, n) \mapsto \varphi(m)n$ is trilinear. Here the functional $\varphi \in M^\vee$ acts on m to give a scalar, which is then multiplied by n. By Theorem 5.6, this trilinear map induces a bilinear map $B\colon (M^\vee \otimes_R N) \times M \to N$ where $B(\varphi \otimes n, m) = \varphi(m)n$. For $t \in M^\vee \otimes_R N$, $B(t, -)$ is in $\mathrm{Hom}_R(M,N)$, so we have a linear map $f\colon M^\vee \otimes_R N \to \mathrm{Hom}_R(M,N)$ by $f(t) = B(t,-)$. (Explicitly, the elementary tensor $\varphi \otimes n$ acts as a linear map $M \to N$ by the rule $(\varphi \otimes n)(m) = \varphi(m)n$.)

Now let M and N be *finite free*. To show f is an isomorphism, we may suppose M and N are nonzero. Pick bases $\{e_i\}$ of M and $\{e_j'\}$ of N. Then f makes $e_i^\vee \otimes e_j'$ act on M by sending each e_k to $e_i^\vee(e_k)e_j' = \delta_{ik}e_j'$. So $f(e_i^\vee \otimes e_j') \in \mathrm{Hom}_R(M,N)$ sends e_i to e_j' and sends every other basis element e_k to 0. Writing elements of M and N as coordinate vectors using their bases, $\mathrm{Hom}_R(M,N)$ becomes matrices and $f(e_i^\vee \otimes e_j')$ becomes the matrix with a 1 in the (j,i) position and 0 elsewhere. Such matrices are a basis of all matrices, so the image of f contains a basis of $\mathrm{Hom}_R(M,N)$, so f is onto.

To show f is one-to-one, suppose $f(\sum_{i,j} c_{ij} e_i^\vee \otimes e_j') = O$ in $\mathrm{Hom}_R(M,N)$. Applying both sides to e_k, we get $\sum_{i,j} c_{ij}\delta_{ik} e_j' = 0$, which says $\sum_j c_{kj} e_j' = 0$, so $c_{kj} = 0$ for all j and all k. Thus every c_{ij} is 0. This concludes the proof that f is an isomorphism.

Let's work out the inverse map explicitly. For $L \in \mathrm{Hom}_R(M,N)$, write $L(e_i) = \sum_j a_{ji} e_j'$, so L has matrix representation (a_{ji}). (The matrix indices here look reversed from usual practice because we use i as the index for basis vectors in M and j as the index for basis vectors in N; review how linear maps become matrices when bases are chosen. If we had indexed bases of M and N with i and j in each other's places, then $L(e_j) = \sum a_{ij} e_i'$.) Suppose under the isomorphism f that $L = f(\sum_{i,j} c_{ij} e_i^\vee \otimes e_j')$, with the coefficients c_{ij} to be determined. Then

$$L(e_k) = \sum_{i,j} c_{ij}(e_i^\vee \otimes e_j')(e_k) = \sum_j c_{kj} e_j',$$

so $a_{jk} = c_{kj}$. Therefore $c_{ij} = a_{ji}$, so an $L \in \mathrm{Hom}_R(M,N)$ with a matrix representation (a_{ji}) corresponds to $\sum_{i,j} a_{ji} e_i^\vee \otimes e_j'$. That just says $e_i^\vee \otimes e_j$ corresponds to the "matrix unit" E_{ji}[17], which we already saw before when computing $f(e_i^\vee \otimes e_j')$. □

Example 5.10. For finite-dimensional K-vector spaces V and W, $V^\vee \otimes_K W \cong \mathrm{Hom}_K(V,W)$ by having $\varphi \otimes w$ act as the linear map $V \to W$ given by the rule $(\varphi \otimes w)(v) = \varphi(v)w$. This is one of the most basic ways tensor products occur in linear algebra. What is this isomorphism really saying? For each $\varphi \in V^\vee$ and $w \in W$, we get a linear map $V \to W$ by $v \mapsto \varphi(v)w$, whose image as v varies is the scalar multiples of w (unless $\varphi = 0$). Since the expression $\varphi(v)w$ is bilinear in φ and w, we can regard this linear map $V \to W$ as defining an effect of $\varphi \otimes w$ on V, with values in W, and the point is that all linear maps $V \to W$ are *sums* of such maps. This corresponds to the fact that every matrix is a sum of matrices that each have a single nonzero entry. By Theorem 5.1, $W \otimes_K V^\vee \cong \mathrm{Hom}_K(V,W)$ too.

When $V = W = K^2$, with standard basis e_1 and e_2, the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{M}_2(K) = \mathrm{Hom}_K(V,W)$ corresponds to the tensor $e_1^\vee \otimes \binom{a}{c} + e_2^\vee \otimes \binom{b}{d}$ in $V^\vee \otimes_K W$ since this tensor sends e_1 to $e_1^\vee(e_1)\binom{a}{c} + e_2^\vee(e_1)\binom{b}{d} = \binom{a}{c}$, and similarly this tensor sends e_2 to $\binom{b}{d}$, which is exactly how $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ acts on e_1 and e_2. In particular, $e_j^\vee \otimes e_i$ corresponds to the matrix in $\mathrm{M}_2(K)$ with 1 in the (i,j) position.

Remark 5.11. If M and N are not both finite free, the map $M^\vee \otimes_R N \to \mathrm{Hom}_R(M,N)$ in Theorem 5.9 may not be an isomorphism, or even injective or surjective. For example, let p be prime, $R = \mathbf{Z}/p^2\mathbf{Z}$, and $M = N = \mathbf{Z}/p\mathbf{Z}$ as R-modules. Check that $M^\vee \cong M$, $M \otimes_R M \cong M$, and $\mathrm{Hom}_R(M,M) \cong M$, but the map $M^\vee \otimes_R M \to \mathrm{Hom}_R(M,M)$ in Theorem 5.9 is identically 0 (it suffices to show each elementary tensor in $M^\vee \otimes_R M$ acts on M as 0). Notice $M^\vee \otimes_R M$ and $\mathrm{Hom}_R(M,M)$ are isomorphic, but the natural linear map between them happens to be identically 0.

When M and N are *finite free* R-modules, the isomorphisms in Corollary 5.8 and Theorem 5.9 lead to a basis-free description of $M \otimes_R N$ making **no mention of universal mapping**

[17]Notice the index switch: $e_i^\vee \otimes e_j$ goes to E_{ji} and not E_{ij}.

properties. Identify M with $M^{\vee\vee}$ by double duality, so Theorem 5.9 with $M^\vee$ in place of M assumes the form

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

where $m \otimes n$ acts as a linear map $M^\vee \to N$ by the rule $(m \otimes n)(\varphi) = \varphi(m)n$. Since $N \cong N^{\vee\vee}$ by double duality, $\mathrm{Hom}_R(M^\vee, N) \cong \mathrm{Hom}_R(M^\vee, (N^\vee)^\vee) \cong \mathrm{Bil}_R(M^\vee, N^\vee; R)$ by Corollary 5.8. Therefore

$$M \otimes_R N \cong \mathrm{Bil}_R(M^\vee, N^\vee; R), \tag{5.4}$$

where $m \otimes n$ acts as a bilinear map $M^\vee \times N^\vee \to R$ by the rule $(m \otimes n)(\varphi, \psi) = \varphi(m)\psi(n)$. Similarly, $M^{\otimes k}$ is isomorphic to the module of k-multilinear maps $(M^\vee)^k \to R$, with the elementary tensor $m_1 \otimes \cdots \otimes m_k$ defining the map sending $(\varphi_1, \ldots, \varphi_k)$ to $\varphi_1(m_1) \cdots \varphi_k(m_k)$.

The definition of the tensor product of finite-dimensional vector spaces in [1, p. 65] and [12, p. 35] is essentially (5.4).[18] It is a good exercise to check these interpretations of $m \otimes n$ as a member of $\mathrm{Hom}_R(M^\vee, N)$ and $\mathrm{Bil}_R(M^\vee, N^\vee; R)$ are identified with each other by Corollary 5.8 and double duality.

But watch out! The isomorphism (5.4) is false for general M and N (where double duality doesn't hold). While there is always a linear map $M \otimes_R N \to \mathrm{Bil}_R(M^\vee, N^\vee; R)$ given on elementary tensors by $m \otimes n \mapsto [(\varphi, \psi) \mapsto \varphi(m)\psi(n)]$, it is generally not an isomorphism.

Example 5.12. Let p be prime, $R = \mathbf{Z}/p^2\mathbf{Z}$, and $M = \mathbf{Z}/p\mathbf{Z}$ as an R-module. The R-modules $M \otimes_R M$ and $\mathrm{Bil}_R(M^\vee, M^\vee; R)$ both have size p and are isomorphic, but the natural map $M \otimes_R M \to \mathrm{Bil}_R(M^\vee, M^\vee; R)$ is identically 0.

Example 5.13. Let $R = \mathbf{Z}$ and $M = N = \mathbf{Q}$. Since $\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Q} \cong \mathbf{Q}$ as $\mathbf{Z}$-modules (Example 4.22) and $\mathbf{Q}^\vee = \mathrm{Hom}_{\mathbf{Z}}(\mathbf{Q}, \mathbf{Z}) = 0$, the left side of (5.4) is nonzero and the right side is 0.

6. Base Extension

In algebra, there are many times a module over one ring is replaced by a related module over another ring. For instance, in linear algebra it is useful to enlarge $\mathbf{R}^n$ to $\mathbf{C}^n$, creating in this way a complex vector space by letting the real coordinates be extended to complex coordinates. In ring theory, irreducibility tests in $\mathbf{Z}[X]$ involve viewing a polynomial in $\mathbf{Z}[X]$ as a polynomial in $\mathbf{Q}[X]$ or reducing the coefficients mod p to view it in $(\mathbf{Z}/p\mathbf{Z})[X]$. We will see that all these passages to modules with new coefficients ($\mathbf{R}^n \rightsquigarrow \mathbf{C}^n, \mathbf{Z}[X] \rightsquigarrow \mathbf{Q}[X], \mathbf{Z}[X] \rightsquigarrow (\mathbf{Z}/p\mathbf{Z})[X]$) can be described in a uniform way using tensor products.

Let $f\colon R \to S$ be a homomorphism of commutative rings. We use f to consider a S-module N as an R-module by $rn := f(r)n$. In particular, S itself is an R-module by $rs := f(r)s$. Passing from N as an S-module to N as an R-module in this way is called *restriction of scalars*.

Example 6.1. If $R \subset S$, f can be the inclusion map (*e.g.*, $\mathbf{R} \hookrightarrow \mathbf{C}$ or $\mathbf{Q} \hookrightarrow \mathbf{C}$). This is how a $\mathbf{C}$-vector space is thought of as an $\mathbf{R}$-vector space or a $\mathbf{Q}$-vector space.

Example 6.2. If $S = R/I$, f can be reduction modulo I: each R/I-module is also an R-module by letting r act in the way that $r \bmod I$ acts.

Here is a simple illustration of restriction of scalars.

[18] Using the first isomorphism in Corollary 5.8 and double duality, $M \otimes_R N \cong \mathrm{Bil}_R(M, N; R)^\vee$ for finite free M and N, where $m \otimes n$ in $M \otimes_R N$ corresponds to the function $B \mapsto B(m, n)$ in $\mathrm{Bil}_R(M, N; R)^\vee$. This is how tensor products of finite-dimensional vector spaces are defined in [8, p. 40], namely $V \otimes_K W$ is the dual space to $\mathrm{Bil}_K(V, W; K)$.

Theorem 6.3. *Let N and N' be S-modules. Any S-linear map $N \to N'$ is also an R-linear map when we treat N and N' as R-modules.*

Proof. Let $\varphi\colon N \to N'$ be S-linear, so $\varphi(sn) = s\varphi(n)$ for all $s \in S$ and $n \in N$. For $r \in R$,
$$\varphi(rn) = \varphi(f(r)n) = f(r)\varphi(n) = r\varphi(n),$$
so φ is R-linear. $\square$

As a notational convention, since we will be going back and forth between R-modules and S-modules a lot, we will write M (or M', and so on) for R-modules and N (or N', and so on) for S-modules. Since N is also an R-module by restriction of scalars, we can form the tensor product R-module $M \otimes_R N$, where
$$r(m \otimes n) = (rm) \otimes n = m \otimes rn,$$
with the third expression really being $m \otimes f(r)n$ since $rn := f(r)n$.

The idea of base extension is to *reverse* the process of restriction of scalars. For an R-module M we want to create an S-module of products sm that matches the old meaning of rm if $s = f(r)$. This new S-module is called an *extension of scalars* or *base extension*. It will be the R-module $S \otimes_R M$ equipped with a specific structure of an S-module.

Since S is a ring, not just an R-module, let's try making $S \otimes_R M$ into an S-module by
$$s'(s \otimes m) := s's \otimes m. \tag{6.1}$$
Is this S-scaling on elementary tensors well-defined and does it extend to S-scaling on all tensors?

Theorem 6.4. *The additive group $S \otimes_R M$ has a unique S-module structure satisfying (6.1), and this is compatible with the R-module structure in the sense that $rt = f(r)t$ for all $r \in R$ and $t \in S \otimes_R M$.*

Proof. Suppose the additive group $S \otimes_R M$ has an S-module structure satisfying (6.1). We will show the S-scaling on all tensors in $S \otimes_R M$ is determined by this. Any $t \in S \otimes_R M$ is a finite sum of elementary tensors, say
$$t = s_1 \otimes m_1 + \cdots + s_k \otimes m_k.$$
For $s \in S$,
$$\begin{aligned} st &= s(s_1 \otimes m_1 + \cdots + s_k \otimes m_k) \\ &= s(s_1 \otimes m_1) + \cdots + s(s_k \otimes m_k) \\ &= ss_1 \otimes m_1 + \cdots + ss_k \otimes m_k \qquad \text{by (6.1)}, \end{aligned}$$
so st is determined, although this formula for it is not obviously well-defined. (Does a different expression for t as a sum of elementary tensors change st?)

Now we show there really is an S-module structure on $S \otimes_R M$ satisfying (6.1). Describing the S-scaling on $S \otimes_R M$ means creating a function $S \times (S \otimes_R M) \to S \otimes_R M$ satisfying the relevant scaling axioms:
$$1 \cdot t = t, \quad s(t_1 + t_2) = st_1 + st_2, \quad (s_1 + s_2)t = s_1t + s_2t, \quad s_1(s_2t) = (s_1s_2)t. \tag{6.2}$$

For each $s' \in S$ we consider the function $S \times M \to S \otimes_R M$ given by $(s, m) \mapsto (s's) \otimes m$. This is R-bilinear, so by the universal mapping property of tensor products there is an R-linear map $\mu_{s'}\colon S \otimes_R M \to S \otimes_R M$ where $\mu_{s'}(s \otimes m) = (s's) \otimes m$ on elementary tensors. Define a multiplication $S \times (S \otimes_R M) \to S \otimes_R M$ by $s't := \mu_{s'}(t)$. This will be the scaling of S on $S \otimes_R M$. We check the conditions in (6.2):

(1) To show $1t = t$ means showing $\mu_1(t) = t$. On elementary tensors, $\mu_1(s \otimes m) = (1 \cdot s) \otimes m = s \otimes m$, so μ_1 fixes elementary tensors. Therefore μ_1 fixes all tensors by additivity.
(2) $s(t_1 + t_2) = st_1 + st_2$ since μ_s is additive.
(3) Showing $(s_1 + s_2)t = s_1 t + s_2 t$ means showing $\mu_{s_1+s_2} = \mu_{s_1} + \mu_{s_2}$ as functions on $S \otimes_R M$. Both sides are additive functions, so it suffices to check they agree on elementary tensors $s \otimes m$, where both sides have common value $(s_1 + s_2)s \otimes m$.
(4) To show $s_1(s_2 t) = (s_1 s_2)t$ means $\mu_{s_1} \circ \mu_{s_2} = \mu_{s_1 s_2}$ as functions on $S \otimes_R M$. Both sides are additive functions of t, so it suffices to check they agree on elementary tensors $s \otimes m$, where both sides have common value $(s_1 s_2 s) \otimes m$.

Let's check the S-module structure on $S \otimes_R M$ is compatible with its original R-module structure. For $r \in R$, if we treat r as $f(r) \in S$ then scaling by $f(r)$ on an elementary tensor $s \otimes m$ has the effect $f(r)(s \otimes m) = f(r)s \otimes m$. Since $f(r)s$ is the definition of rs (this is how we make S into an R-module), $f(r)s \otimes m = rs \otimes m = r(s \otimes m)$. Thus $f(r)(s \otimes m) = r(s \otimes m)$, so $f(r)t = rt$ for all t in $S \otimes_R M$ by additivity. □

By exactly the same kind of argument, $M \otimes_R S$ with S on the right has a unique S-module structure where $s'(m \otimes s) = m \otimes s's$. So whenever we meet $S \otimes_R M$ or $M \otimes_R S$, we know they are S-modules in a specific way. Moreover, these two S-modules are naturally isomorphic: by Theorem 5.1, there is an isomorphism $\varphi\colon S \otimes_R M \to M \otimes_R S$ of R-modules where $\varphi(s \otimes m) = m \otimes s$. To show φ is in fact an isomorphism of S-modules, all we need to do is check S-linearity since φ is known to be additive and a bijection. To show $\varphi(s't) = s'\varphi(t)$ for all s' and t, additivity of both sides in t means we may focus on the case $t = s \otimes m$:

$$\varphi(s'(s \otimes m)) = \varphi((s's) \otimes m) = m \otimes s's = s'(m \otimes s) = s'\varphi(s \otimes m).$$

This idea of creating an S-module isomorphism by using a known R-module isomorphism that is also S-linear will be used many more times, so watch for it.

Now we must be careful to refer to R-linear and S-linear maps, rather than just linear maps, so it is clear what our scalar ring is each time.

Example 6.5. In Example 4.5 we saw $(R/I) \otimes_R M \cong M/IM$ as R-modules by $\overline{r} \otimes m \mapsto \overline{rm}$. Since M/IM is naturally an R/I-module, and now we know $(R/I) \otimes_R M$ is an R/I-module, the R-module isomorphism $(R/I) \otimes_R M \cong M/IM$ turns out to be an R/I-module isomorphism too since it is R/I-linear (check!).

Theorem 6.6. *If F is a free R-module with basis $\{e_i\}_{i \in I}$ then $S \otimes_R F$ is a free S-module with basis $\{1 \otimes e_i\}_{i \in I}$.*

Proof. Since S is an R-module, we know from Theorem 4.15 that every element of $S \otimes_R F$ has a *unique* representation in the form $\sum_{i \in I} s_i \otimes e_i$, where all but finitely many s_i equal 0. Since $s_i \otimes e_i = s_i(1 \otimes e_i)$ in the S-module structure on $S \otimes_R F$, every element of $S \otimes_R F$ is a unique S-linear combination $\sum s_i(1 \otimes e_i)$, which says $\{1 \otimes e_i\}$ is an S-basis of $S \otimes_R F$. □

Example 6.7. As an S-module, $S \otimes_R R^n$ has S-basis $\{1 \otimes e_1, \ldots, 1 \otimes e_n\}$ where $\{e_1, \ldots, e_n\}$ is the standard basis of R^n, so $S^n \cong S \otimes_R R^n$ as S-modules by

$$(s_1, \ldots, s_n) \mapsto \sum_{i=1}^{n} s_i(1 \otimes e_i) = \sum_{i=1}^{n} s_i \otimes e_i$$

because this map is S-linear (check!) and sends an S-basis to an S-basis. In particular, $S \otimes_R R \cong S$ as S-modules by $s \otimes r \mapsto sr$.

For instance,

$$\mathbf{C} \otimes_{\mathbf{R}} \mathbf{R}^n \cong \mathbf{C}^n, \quad \mathbf{C} \otimes_{\mathbf{R}} \mathrm{M}_n(\mathbf{R}) \cong \mathrm{M}_n(\mathbf{C})$$

as $\mathbf{C}$-vector spaces, not just as $\mathbf{R}$-vector spaces. For an ideal I in R, $(R/I) \otimes_R R^n \cong (R/I)^n$, not just as R-modules. as R/I-modules.

Example 6.8. As an S-module, $S \otimes_R R[X]$ has S-basis $\{1 \otimes X^i\}_{i \geq 0}$, so $S \otimes_R R[X] \cong S[X]$ as S-modules[19] by $\sum_{i \geq 0} s_i \otimes X^i \mapsto \sum_{i \geq 0} s_i X^i$.

As particular examples, $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{R}[X] \cong \mathbf{C}[X]$ as $\mathbf{C}$-vector spaces, $\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Z}[X] \cong \mathbf{Q}[X]$ as $\mathbf{Q}$-vector spaces and $(\mathbf{Z}/p\mathbf{Z}) \otimes_{\mathbf{Z}} \mathbf{Z}[X] \cong (\mathbf{Z}/p\mathbf{Z})[X]$ as $\mathbf{Z}/p\mathbf{Z}$-vector spaces.

Example 6.9. If we treat $\mathbf{C}^n$ as a *real vector space*, then its base extension to $\mathbf{C}$ is the complex vector space $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C}^n$ where $c(z \otimes v) = cz \otimes v$ for c in $\mathbf{C}$. Since $\mathbf{C}^n \cong \mathbf{R}^{2n}$ as real vector spaces, we have a $\mathbf{C}$-vector space isomorphism

$$\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C}^n \cong \mathbf{C} \otimes_{\mathbf{R}} \mathbf{R}^{2n} \cong \mathbf{C}^{2n}.$$

That's interesting: restricting scalars on $\mathbf{C}^n$ to make it a real vector space and then extending scalars back up to $\mathbf{C}$ does *not* give us $\mathbf{C}^n$ back, but instead two copies of $\mathbf{C}^n$. The point is that when we restrict scalars, the real vector space $\mathbf{C}^n$ forgets it is a complex vector space. So the base extension of $\mathbf{C}^n$ from a real vector space to a complex vector space doesn't remember that it used to be a complex vector space.

Quite generally, if V is a finite-dimensional complex vector space and we view it as a real vector space, its base extension $\mathbf{C} \otimes_{\mathbf{R}} V$ to a complex vector space is not V but a direct sum of two copies of V. Let's do a dimension check. Set $n = \dim_{\mathbf{C}}(V)$, so $\dim_{\mathbf{R}}(V) = 2n$. Then $\dim_{\mathbf{R}}(\mathbf{C} \otimes_{\mathbf{R}} V) = \dim_{\mathbf{R}}(\mathbf{C}) \dim_{\mathbf{R}}(V) = 2(2n) = 4n$ and $\dim_{\mathbf{R}}(V \oplus V) = 2 \dim_{\mathbf{R}}(V) = 2(2n) = 4n$, so the two dimensions match. This match is of course not a proof that there is a natural isomorphism $\mathbf{C} \otimes_{\mathbf{R}} V \to V \oplus V$ of complex vector spaces. Work out such an isomorphism as an exercise. The proof had better use the fact that V is already a complex vector space to make sense of $V \oplus V$ as a complex vector space.

To get our bearing on this example, let's compare an S-module N with the S-module $S \otimes_R N$ (where $s'(s \otimes n) = s's \otimes n$). Since N is already an S-module, should $S \otimes_R N \cong N$? If you think so, reread Example 6.9 ($R = \mathbf{R}$, $S = \mathbf{C}$, $N = \mathbf{C}^n$). Scalar multiplication $S \times N \to N$ is R-bilinear, so there is an R-linear map $\varphi \colon S \otimes_R N \to N$ where $\varphi(s \otimes n) = sn$. This map is also S-linear: $\varphi(st) = s\varphi(t)$. To check this, since both sides are additive in t it suffices to check the case of elementary tensors, and

$$\varphi(s(s' \otimes n)) = \varphi((ss') \otimes n) = ss'n = s(s'n) = s\varphi(s' \otimes n).$$

In the other direction, the function $\psi \colon N \to S \otimes_R N$ where $\psi(n) = 1 \otimes n$ is R-linear but is generally not S-linear since $\psi(sn) = 1 \otimes sn$ has no reason to be $s\psi(n) = s \otimes n$ because we're using $\otimes_R$, not $\otimes_S$. We have created natural maps $\varphi \colon S \otimes_R N \to N$ and $\psi \colon N \to S \otimes_R N$; are they inverses? It's unlikely, since φ is S-linear and ψ is generally not. But let's work out the composites and see what happens. In one direction,

$$\varphi(\psi(n)) = \varphi(1 \otimes n) = 1 \cdot n = n.$$

In the other direction,

$$\psi(\varphi(s \otimes n)) = \psi(sn) = 1 \otimes sn \neq s \otimes n$$

[19] We saw $S \otimes_R R[X]$ and $S[X]$ are isomorphic as R-modules in Example 4.16 when $S \supset R$, and it holds now for all $R \xrightarrow{f} S$.

in general. So $\varphi\circ\psi$ is the identity but $\psi\circ\varphi$ is usually not the identity. Since $\varphi\circ\psi = \mathrm{id}_N$, ψ is a section to φ, so N is a direct summand of $S\otimes_R N$. Explicitly, $S\otimes_R N \cong \ker\varphi\oplus N$ by $s\otimes n\mapsto (s\otimes n - 1\otimes sn, sn)$ and its inverse map is $(k,n)\mapsto k+1\otimes n$. The phenomenon that $S\otimes_R N$ is typically larger than N when N is an S-module can be remembered by the example $\mathbf{C}\otimes_{\mathbf{R}}\mathbf{C}^n\cong\mathbf{C}^{2n}$.

Theorem 6.10. *For R-modules $\{M_i\}_{i\in I}$, there is an S-module isomorphism*

$$S\otimes_R \bigoplus_{i\in I} M_i \cong \bigoplus_{i\in I}(S\otimes_R M_i).$$

Proof. Since S is an R-module, by Theorem 5.4 there is an R-module isomorphism

$$\varphi\colon S\otimes_R \bigoplus_{i\in I} M_i \to \bigoplus_{i\in I}(S\otimes_R M_i)$$

where $\varphi(s\otimes(m_i)_{i\in I}) = (s\otimes m_i)_{i\in I}$. To show φ is an S-module isomorphism, we just have to check φ is S-linear, since we already know φ is additive and a bijection. It is obvious that $\varphi(st) = s\varphi(t)$ when t is an elementary tensor, and since both $\varphi(st)$ and $s\varphi(t)$ are additive in t the case of general tensors follows. $\square$

The analogue of Theorem 6.10 for direct products is false. There is a natural S-linear map $S\otimes_R\prod_{i\in I}M_i\to\prod_{i\in I}(S\otimes_R M_i)$, but it need not be an isomorphism. Here are two examples.

- $\mathbf{Q}\otimes_{\mathbf{Z}}\prod_{i\geq 1}\mathbf{Z}/p^i\mathbf{Z}$ is nonzero (Remark 5.5) but $\prod_{i\geq 1}(\mathbf{Q}\otimes_{\mathbf{Z}}\mathbf{Z}/p^i\mathbf{Z})$ is 0.
- $\mathbf{Q}\otimes_{\mathbf{Z}}\prod_{i\geq 1}\mathbf{Z}$ is isomorphic as a $\mathbf{Q}$-vector space not to $\prod_{i\geq 1}(\mathbf{Q}\otimes_{\mathbf{Z}}\mathbf{Z})\cong\prod_{i\geq 1}\mathbf{Q}$, but rather to the subgroup of $\prod_{i\geq 1}\mathbf{Q}$ consisting of rational sequences with a common denominator. Under the natural map $\mathbf{Q}\otimes_{\mathbf{Z}}\prod_{i\geq 1}\mathbf{Z}\to\prod_{i\geq 1}\mathbf{Q}$, the image of an elementary tensor has coordinates with a common denominator, and each tensor in $\mathbf{Q}\otimes_{\mathbf{Z}}\prod_{i\geq 1}\mathbf{Z}$ is a *finite* sum of elementary tenors, so its image in $\prod_{i\geq 1}\mathbf{Q}$ is a sequence with a common denominator.

We now put base extensions to work. Let M be a finitely generated R-module, say with n generators. That is the same as saying there is a linear surjection $R^n\twoheadrightarrow M$. To say M contains a subset of d linearly independent elements is the same as saying there is a linear injection $R^d\hookrightarrow M$. If both $R^n\twoheadrightarrow M$ and $R^d\hookrightarrow M$, it is natural to suspect $d\leq n$, *i.e.*, the size of a spanning set should always be an upper bound on the size of a linearly independent subset. Is it really true? If R is a field, so modules are vector spaces, we can use dimension inequalities on R^d, M, and R^n to see $d\leq n$. But if R is not a field, then what? We will settle the issue in the affirmative when R is a domain, by tensoring M with the fraction field of R to reduce to the case of vector spaces. We first tensored R-modules with the fraction field of R in Theorem 4.27, but not much use was made of the vector space structure of the tensor product with a field. Now we exploit it.

Theorem 6.11. *Let R be a domain with fraction field K. For a finitely generated R-module M, $K\otimes_R M$ is finite-dimensional as a K-vector space and $\dim_K(K\otimes_R M)$ is the maximal number of R-linearly independent elements in M and is a lower bound on the size of a spanning set for M. In particular, the size of each linearly independent subset of M is less than or equal to the size of each spanning set of M.*

Proof. If $x_1,\dots,x_n$ is a spanning set for M as an R-module then $1\otimes x_1,\dots,1\otimes x_n$ span $K\otimes_R M$ as a K-vector space, so $\dim_K(K\otimes_R M)\leq n$.

Let $y_1, \dots, y_d$ be R-linearly independent in M. We will show $\{1 \otimes y_i\}$ is K-linearly independent in $K \otimes_R M$, so $d \leq \dim_K(K \otimes_R M)$. Suppose $\sum_{i=1}^d c_i(1 \otimes y_i) = 0$ with $c_i \in K$. Write $c_i = a_i/b$ using a common denominator b in R. Then $0 = 1/b \otimes \sum_{i=1}^d a_i y_i$ in $K \otimes_R M$. By Corollary 4.28, this implies $\sum_{i=1}^d a_i y_i \in M_{\rm tor}$, so $\sum_{i=1}^d r a_i y_i = 0$ in M for some nonzero $r \in R$. By linear independence of the y_i's over R, every ra_i is 0, so every a_i is 0 (R is a domain). Thus every $c_i = a_i/b$ is 0.

It remains to prove M has a linearly independent subset of size $\dim_K(K \otimes_R M)$. Let $\{e_1, \dots, e_d\}$ be a linearly independent subset of M, where d is maximal. (Since $d \leq \dim_K(K \otimes_R M)$, there is a maximal d.) For every $m \in M$, $\{e_1, \dots, e_d, m\}$ has to be linearly dependent, so there is a nontrivial R-linear relation

$$a_1 e_1 + \cdots + a_d e_d + am = 0.$$

Necessarily $a \neq 0$, as otherwise all the a_i's are 0 by linear independence of the e_i's. In $K \otimes_R M$,

$$\sum_{i=1}^d a_i(1 \otimes e_i) + a(1 \otimes m) = 0$$

and from the K-vector space structure on $K \otimes_R M$ we can solve for $1 \otimes m$ as a K-linear combination of the $1 \otimes e_i$'s. Therefore $\{1 \otimes e_i\}$ spans $K \otimes_R M$ as a K-vector space. This set is also linearly independent over K by the previous paragraph, so it is a basis and therefore $d = \dim_K(K \otimes_R M)$. $\square$

While M has at most $\dim_K(K \otimes_R M)$ linearly independent elements and this upper bound is achieved, each spanning set has at least $\dim_K(K \otimes_R M)$ elements but this lower bound is not necessarily reached. For example, if R is not a field and M is a torsion module (*e.g.*, R/I for I a nonzero proper ideal) then $K \otimes_R M = 0$ and M certainly doesn't have a spanning set of size 0 if $M \neq 0$. It is also not true that finiteness of $\dim_K(K \otimes_R M)$ implies M is finitely generated as an R-module. Take $R = \mathbf{Z}$ and $M = \mathbf{Q}$, so $\mathbf{Q} \otimes_{\mathbf{Z}} M = \mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Q} \cong \mathbf{Q}$ (Example 4.22), which is finite-dimensional over $\mathbf{Q}$ but M is not finitely generated over $\mathbf{Z}$.

The maximal number of linearly independent elements in an R-module M, for R a domain, is called the *rank* of M.[20] This use of the word "rank" is consistent with its usage for finite free modules as the size of a basis: if M is free with an R-basis of size n then $K \otimes_R M$ has a K-basis of size n by Theorem 6.6.

Example 6.12. A nonzero ideal I in a domain R has rank 1. We can see this in two ways. First, any two nonzero elements in I are linearly dependent over R, so the maximal number of R-linearly independent elements in I is 1. Second, $K \otimes_R I \cong K$ as K-vector spaces (in Theorem 4.23 we showed they are isomorphic as R-modules, but that isomorphism is also K-linear; check!), so $\dim_K(K \otimes_R I) = 1$.

Example 6.13. A finitely generated R-module M has rank 0 if and only if it is a torsion module, since $K \otimes_R M = 0$ if and only if M is a torsion module.

Since $K \otimes_R M \cong K \otimes_R (M/M_{\rm tor})$ as K-vector spaces (the isomorphism between them as R-modules in Theorem 4.27 is easily checked to be K-linear – check!), M and $M/M_{\rm tor}$ have the same rank.

We return to general R, no longer a domain, and see how to make the tensor product of an R-module and S-module into an S-module.

[20]When R is not a domain, this concept of rank for R-modules is not quite the right one.

Theorem 6.14. *Let M be an R-module and N be an S-module.*

(1) *The additive group $M\otimes_R N$ has a unique structure of S-module such that $s(m\otimes n) = m\otimes sn$ for $s \in S$. This is compatible with the R-module structure in the sense that $rt = f(r)t$ for $r \in R$ and $t \in M \otimes_R N$.*
(2) *The S-module $M \otimes_R N$ is isomorphic to $(S \otimes_R M) \otimes_S N$ by sending $m \otimes_R n$ to $(1 \otimes_R m) \otimes_S n$.*

The point of part 2 is that it shows how the S-module structure on $M \otimes_R N$ can be described as an ordinary S-module tensor product by base extending M to an S-module $S \otimes_R M$. Part 2 has both R-module and S-module tensor products, and it is the first time that we must decorate the tensor product sign explicitly. Up to now it was actually unnecessary, as all the tensor products were over R.

Writing $S\otimes_R M$ as $M\otimes_R S$ makes the isomorphism in part 2 notationally obvious, since it becomes $(M \otimes_R S) \otimes_S N \cong M \otimes_R N$; this is similar to the "proof" of the chain rule in differential calculus, $dy/dx = (dy/du)(du/dx)$, by cancellation of du in the notation. This kind of notational trick will be proved in greater generality in Theorem 6.23(3).

Proof. (1) This is similar to the proof of Theorem 6.4 (which is the special case $N = S$). We just sketch the idea.

Since every tensor is a sum of elementary tensors, declaring how $s \in S$ scales elementary tensors in $M\otimes_R N$ determines its scaling on all tensors. To show the rule $s(m\otimes n) = m\otimes sn$ really corresponds to an S-module structure, for each $s \in S$ we consider the function $M \times N \to M \otimes_R N$ given by $(m, n) \mapsto m \otimes sn$. This is R-bilinear in m and n, so there is an R-linear map $\mu_s \colon M \otimes_R N \to M \otimes_R N$ such that $\mu_s(m \otimes n) = m \otimes sn$ on elementary tensors. Define a multiplication $S \times (M \otimes_R N) \to M \otimes_R N$ by $st := \mu_s(t)$. It is left to the reader to check that the maps μ_s on $M \otimes_R N$, as s varies, satisfy the scaling axioms that make $M \otimes_R N$ an S-module.

To check $rt = f(r)t$ for $r \in R$ and $t \in M \otimes_R N$, both sides are additive in t so it suffices to check equality when $t = m\otimes n$ is an elementary tensor. In that case $r(m\otimes n) = m\otimes rn = m \otimes f(r)n = f(r)(m \otimes n)$.

(2) Let $M \times N \to (S \otimes_R M) \otimes_S N$ by $(m, n) \mapsto (1 \otimes_R m) \otimes_S n$. We want to check this is R-bilinear. Biadditivity is clear. For R-scaling, we have

$$(1 \otimes_R rm) \otimes_S n = (r(1 \otimes_R m)) \otimes_S n = (f(r)(1 \otimes_R m)) \otimes_S n = f(r)((1 \otimes_R m) \otimes_S n)$$

and

$$(1 \otimes_R m) \otimes_S rn = (1 \otimes_R m) \otimes_S f(r)n = f(r)((1 \otimes_R m) \otimes_S n).$$

Now the universal mapping property of tensor products gives an R-linear map $\varphi\colon M\otimes_R N \to (S\otimes_R M)\otimes_S N$ where $\varphi(m\otimes_R n) = (1\otimes_R m)\otimes_S n$. This is exactly the map we were looking for, but we only know it is R-linear so far. It is also S-linear: $\varphi(st) = s\varphi(t)$. To check this, it suffices by additivity of φ to focus on the case of an elementary tensor:

$$\varphi(s(m \otimes_R n)) = \varphi(m \otimes_R sn) = (1 \otimes_R m) \otimes_S sn = s((1 \otimes_R m) \otimes_S n) = s\varphi(m \otimes_R n).$$

To show φ is an isomorphism, we create an inverse map $(S\otimes_R M)\otimes_S N \to M\otimes_R N$. The function $S \times M \times N \to M \otimes_R N$ given by $(s, m, n) \mapsto m \otimes sn$ is R-trilinear, so by Theorem 5.6 there is an R-bilinear map $B\colon (S \otimes_R M) \times N \to M \otimes_R N$ where $B(s \otimes m, n) = m \otimes sn$. This function is in fact S-bilinear:

$$B(st, n) = sB(t, n), \quad B(t, sn) = sB(t, n).$$

To check these equations, the additivity of both sides of the equations in t reduces us to case when t is an elementary tensor. Writing $t = s' \otimes m$,

$$B(s(s' \otimes m), n) = B(ss' \otimes m, n) = m \otimes ss'n = m \otimes s(s'n) = s(m \otimes s'n) = sB(s' \otimes m, n)$$

and

$$B(s' \otimes m, sn) = m \otimes s'(sn)m \otimes s(s'n) = s(m \otimes s'n) = sB(s' \otimes m, n).$$

Now the universal mapping property of the tensor product for S-modules tells us there is an S-linear map $\psi\colon (S \otimes_R M) \otimes_S N \to M \otimes_R N$ such that $\psi(t \otimes n) = B(t, n)$.

It is left to the reader to check $\varphi \circ \psi$ and $\psi \circ \varphi$ are identity functions, so φ is an S-module isomorphism. $\square$

In addition to $M \otimes_R N$ being an S-module because N is, the tensor product $N \otimes_R M$ in the other order has a unique S-module structure where $s(n \otimes m) = sn \otimes m$, and this is proved in a similar way.

Example 6.15. For an S-module N, let's show $R^k \otimes_R N \cong N^k$ as S-modules. By Theorem 5.4, $R^k \otimes_R N \cong (R \otimes_R N)^k \cong N^k$ as R-modules, an explicit isomorphism $\varphi\colon R^k \otimes_R N \to N^k$ being $\varphi((r_1, \ldots, r_k) \otimes n) = (r_1 n, \ldots, r_k n)$. Let's check φ is S-linear: $\varphi(st) = s\varphi(t)$. Both sides are additive in t, so we only need to check when t is an elementary tensor:

$$\varphi(s((r_1, \ldots, r_k) \otimes n)) = \varphi((r_1, \ldots, r_k) \otimes sn) = (r_1 sn, \ldots, r_k sn) = s\varphi((r_1, \ldots, r_k) \otimes n).$$

To reinforce the S-module isomorphism

$$M \otimes_R N \cong (S \otimes_R M) \otimes_S N \tag{6.3}$$

from Theorem 6.14(2), let's write out the isomorphism in both directions on appropriate tensors:

$$m \otimes_R n \mapsto (1 \otimes_R m) \otimes_S n, \quad (s \otimes_R m) \otimes_S n \mapsto m \otimes_R sn.$$

Corollary 6.16. *If M and M' are isomorphic R-modules, and N is an S-module, then $M \otimes_R N$ and $M' \otimes_R N$ are isomorphic S-modules, as are $N \otimes_R M$ and $N \otimes_R M'$.*

Proof. We will show $M \otimes_R N \cong M' \otimes_R N$ as S-modules. The other one is similar.

Let $\varphi\colon M \to M'$ be an R-module isomorphism. To write down an S-module isomorphism $M \otimes_R N \to M' \otimes_R N$, we will write down an R-module isomorphism that is also S-linear. Let $M \times N \to M' \otimes_R N$ by $(m, n) \mapsto \varphi(m) \otimes n$. This is R-bilinear (check!), so we get an R-linear map $\Phi\colon M \otimes_R N \to M' \otimes_R N$ such that $\Phi(m \otimes n) = \varphi(m) \otimes n$. This is also S-linear: $\Phi(st) = s\Phi(t)$. Since Φ is additive, it suffices to check this when $t = m \otimes n$:

$$\Phi(s(m \otimes n)) = \Phi(m \otimes sn) = \varphi(m) \otimes sn = s(\varphi(m) \otimes n) = s\Phi(m \otimes n).$$

Using the inverse map to φ we get an R-linear map $\Psi\colon M' \otimes_R N \to M \otimes_R N$ that is also S-linear, and a computation on elementary tensors shows Φ and Ψ are inverses of each other. $\square$

Example 6.17. We can use tensor products to prove the well-definedness of ranks of finite free R-modules when R is not the zero ring. Suppose $R^m \cong R^n$ as R-modules. Pick a maximal ideal $\mathfrak{m}$ in R (Zorn's lemma) and $R/\mathfrak{m} \otimes_R R^m \cong R/\mathfrak{m} \otimes_R R^n$ as $R/\mathfrak{m}$-vector spaces by Corollary 6.16. Therefore $(R/\mathfrak{m})^m \cong (R/\mathfrak{m})^n$ as $R/\mathfrak{m}$-vector spaces (Example 6.7), so taking dimensions of both sides over $R/\mathfrak{m}$ tells us $m = n$.

Here's a conundrum. If N and N' are both S-modules, then we can make $N \otimes_R N'$ into an S-module in two ways: $s(n \otimes_R n') = sn \otimes_R n'$ and $s(n \otimes_R n') = n \otimes_R sn'$. In the first S-module structure on $N \otimes_R N'$, N' only matters as an R-module. In the second S-module structure, N only matters as an R-module. These two S-module structures on $N \otimes_R N'$ are *not* generally the same because the tensor product is $\otimes_R$, not $\otimes_S$, so $sn \otimes_R n'$ need not equal $n \otimes_R sn'$. But are the two S-module structures on $N \otimes_R N'$ at least isomorphic to each other? In general, no.

Example 6.18. Let $R = \mathbf{Z}$ and $S = \mathbf{Z}[\sqrt{d}]$ where d is a nonsquare integer such that $d \equiv 1 \bmod 4$. Set I to be the ideal $(2, 1+\sqrt{d})$ in S, so as a $\mathbf{Z}$-module $I = \mathbf{Z}2 + \mathbf{Z}(1+\sqrt{d})$. We will look at the two S-module structures on $S \otimes_{\mathbf{Z}} I$, coming from scaling by S on the left and the right.

As $\mathbf{Z}$-modules, S and I are both free of rank 2. When $S \otimes_{\mathbf{Z}} I$ is an S-module by scaling by S on the left, I only matters as a $\mathbf{Z}$-module, so $S \otimes_{\mathbf{Z}} I \cong S \otimes_{\mathbf{Z}} \mathbf{Z}^2$ as S-modules by Corollary 6.16. By Example 6.7, $S \otimes_{\mathbf{Z}} \mathbf{Z}^2 \cong S^2$ as S-modules. Similarly, by making $S \otimes_{\mathbf{Z}} I$ into an S-module by scaling by S on the right, $S \otimes_{\mathbf{Z}} I \cong \mathbf{Z}^2 \otimes_{\mathbf{Z}} I \cong I \oplus I$ as S-modules. If we can show $I \oplus I \not\cong S^2$ as S-modules then $S \otimes_{\mathbf{Z}} I$ has different S-module structures from scaling by S on the left and the right.

The crucial property of I for us is that $I^2 = 2I$, which is left to the reader to check. Let's see how this implies that I is not a principal ideal: if $I = \alpha S$ then $I^2 = \alpha^2 S$, so $\alpha^2 S = 2\alpha S$, which implies $\alpha S = 2S$. However, $2S$ has index 4 in S while I has index 2 in S, so I is nonprincipal. Thus the ideal I is not a free S-module. Is it then obvious that $I \oplus I$ can't be a free S-module? No! A direct sum of two nonfree modules *can* be free. For instance, in $\mathbf{Z}[\sqrt{-5}]$ the ideals $I = (3, 1+\sqrt{-5})$ and $J = (3, 1-\sqrt{-5})$ are both nonprincipal but it can be shown that $I \oplus J \cong \mathbf{Z}[\sqrt{-5}] \oplus \mathbf{Z}[\sqrt{-5}]$ as $\mathbf{Z}[\sqrt{-5}]$-modules. The reason a direct sum of non-free modules sometimes is free is that there is more room to move around in a direct sum than just within the direct summands, and this extra room might contain a basis. So showing $I \oplus I$ is not a free S-module requires work.

Using more advanced tools in multilinear algebra (specifically, exterior powers), one can show that *if* $I \oplus I \cong S^2$ as S-modules then $I \otimes_S I \cong S$ as S-modules. Then since multiplication gives a surjective S-linear map $I \otimes_S I \twoheadrightarrow I^2$ (where $x \otimes y \mapsto xy$), there would be a surjective S-linear map $S \twoheadrightarrow I^2$, which means I^2 would be a principal ideal. However, $I^2 = 2I$ and I is not principal, so I^2 is not principal.

The lesson from this example is that if you want $N \otimes_R N'$ to be an S-module where N and N' are S-modules, you have to *specify* whether S scales on the left or the right. That two S-modules structures on $N \otimes_R N'$ are the same or at least isomorphic needs a proof. Here is one such example, where R is a domain and $S = K$ is its fraction field.

Theorem 6.19. *Let R be a domain, with fraction field K. For K-vector spaces V and W, the K-vector space structures on $V \otimes_R W$ using K-scaling on either the V or W factor are the same.*

Proof. The two K-vector space structures on $V \otimes_R W$ are based on either the formula $x(v \otimes_R w) = xv \otimes_R w$ on elementary tensors or the formula $x(v \otimes_R w) = v \otimes_R xw$ on elementary tensors, where $x \in K$ and we write $\otimes_R$ rather than $\otimes$ in the elementary tensors for emphasis.[21] Proving that the two K-vector space structures on $V \otimes_R W$ agree amounts

[21]These scaling formulas are *not* really definitions of scalar multiplication by K on $V \otimes_R W$, since tensors are not unique sums of elementary tensors. That such scaling formulas are well-defined operatios on $V \otimes_K W$ requires creating the scaling functions as in the proof of part 1 of Theorem 6.14.

to showing $xv \otimes_R w = v \otimes_R xw$ for all $x \in K$, $v \in V$, and $w \in W$. This is something we dealt with back in the proof of Theorem 4.21, and now we'll apply that argument again.

Write $x = a/b$ with $a, b \in R$ and $b \neq 0$. Then $(xv) \otimes_R w$ equals

$$\left(\frac{a}{b}v\right) \otimes_R w = a\left(\frac{1}{b}v\right) \otimes_R w = \frac{1}{b}v \otimes_R aw = \frac{1}{b}v \otimes_R b\left(\frac{a}{b}w\right) = b\left(\frac{1}{b}v\right) \otimes_R \frac{a}{b}w = v \otimes_R \frac{a}{b}w,$$

which is $v \otimes_R xw$. □

It would have been wrong to complete the proof by immediately writing $xv \otimes_R w = v \otimes_R xw$, because we are working with an R-module tensor product and the scalar x is not limited to R. We can say in one step that $xv \otimes_R w = v \otimes_R xw$ with $x \in R$, but to say this holds when $x \in K$ needed justification. (For comparison, in $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C}$ we have $1 \otimes_{\mathbf{R}} i \neq i \otimes_{\mathbf{R}} 1$ since these are different members of a basis. More generally, check using bases that $z \otimes_{\mathbf{R}} iw \neq iz \otimes_{\mathbf{R}} w$ in $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C}$ except when z or w is 0. This doesn't violate Theorem 6.19 since $\mathbf{C}$ is not the fraction field of $\mathbf{R}$.)

Corollary 6.20. *With notation as in Theorem 6.19, $V \otimes_R W \cong V \otimes_K W$ as K-vector spaces by $v \otimes_R w \mapsto v \otimes_K w$.*

There is no ambiguity about what we mean by $V \otimes_R W$ as a K-vector space, since the K-scaling via V or W is exactly the same.

Proof. We will give two proofs.

The mapping $V \times W \to V \otimes_K W$ given by $(v, w) \mapsto v \otimes_K w$ is R-bilinear, so there is a unique R-linear mapping $\varphi\colon V \otimes_R W \to V \otimes_K W$ given by $\varphi(v \otimes_R w) = v \otimes_K w$ on elementary tensors. To show φ is K-linear, it suffices to check $\varphi(xt) = x\varphi(t)$ when t is an elementary tensor, so we want to check $\varphi(x(v \otimes_R w)) = x(v \otimes_K w)$. By the *definition* of the K-vector space structure on $V \otimes_R W$, $\varphi(x(v \otimes_R w)) = \varphi((xv) \otimes_R w) = (xv) \otimes_K w$, which is $x(v \otimes_K w)$. (We could also say $\varphi(x(v \otimes_R w)) = \varphi(v \otimes_R xw) = v \otimes_K xw$, which is $x(v \otimes_K w)$.)

The mapping $V \times W \to V \otimes_R W$ given by $(v, w) \mapsto v \otimes_R w$ is not just R-bilinear, but K-bilinear. For example, $(xv, w) \mapsto (xv) \otimes_R w = x(v \otimes_R w)$. Thus we get a K-linear map $\psi\colon V \otimes_K W \mapsto V \otimes_R W$ by $\psi(v \otimes_K w) = v \otimes_R w$. Since $\varphi \circ \psi$ and $\psi \circ \varphi$ are both additive and fix all elementary tensors (in $V \otimes_K W$ and $V \otimes_R W$, respectively), they fix all tensors and thus φ and ψ are inverses. Therefore φ is an isomorphism of K-vector spaces.

Another reason $V \otimes_R W \cong V \otimes_K W$ as K-vector spaces is that $V \otimes_R W$ satisfies the universal mapping property of $V \otimes_K W$! Let's check this. The canonical R-bilinear map $V \times W \to V \otimes_R W$ is not just R-bilinear, but K-bilinear (why?). Then for each K-vector space U and K-bilinear map $B\colon V \times W \to U$, since B is R-bilinear the universal mapping property of $V \otimes_R W$ tells us that there is a unique R-linear map L making the diagram

$$\begin{tikzcd} & V \otimes_R W \arrow[dd, dotted, "L"] \\ V \times W \arrow[ur, "\otimes_R"] \arrow[dr, "B"'] & \\ & U \end{tikzcd}$$

commute, and on account of the fact that U and $V \otimes_R W$ are already K-vector spaces you can check that L is in fact K-linear (and is the only K-linear map that can fit into the above commutative diagram). Any two solutions of a universal mapping property are uniquely isomorphic, so $V \otimes_R W \cong V \otimes_K W$. More specifically, using for B the canonical K-bilinear map $V \times W \to V \otimes_K W$ *implies* that the diagram

$$\begin{tikzcd} & V \otimes_R W \arrow[dd, dotted, "v\otimes_R w\mapsto v\otimes_K w"] \\ V \times W \arrow[ur, "\otimes_R"] \arrow[dr, "\otimes_K"'] & \\ & V \otimes_K W \end{tikzcd}$$

commutes, and by universality the vertical map has to be an isomorphism. $\square$

Example 6.21. Since $\mathbf{R}$ is a $\mathbf{Q}$-vector space, $\mathbf{R} \otimes_{\mathbf{Z}} \mathbf{R} \cong \mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ as $\mathbf{Q}$-vector spaces by $x \otimes_{\mathbf{Z}} y \mapsto x \otimes_{\mathbf{Q}} y$. Is there a "formula" for $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$? Yes and no. Since $\mathbf{R}$ is uncountable, if $\{e_i\}$ is a $\mathbf{Q}$-basis of $\mathbf{R}$, then this basis has cardinality equal to the cardinality of $\mathbf{R}$, and $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ has $\mathbf{Q}$-basis $\{e_i \otimes e_j\}$, whose cardinality is also the same as $\mathbf{R}$, so we could say $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ is isomorphic to $\mathbf{R}$ as $\mathbf{Q}$-vector spaces. However, this isomorphism is completely nonconstructive.

Recalling that $\mathbf{R} \otimes_{\mathbf{R}} \mathbf{R} \cong \mathbf{R}$ as $\mathbf{R}$-vector spaces by $x \otimes_{\mathbf{R}} y \mapsto xy$, might the $\mathbf{Q}$-linear map $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R} \to \mathbf{R}$ given by $x \otimes_{\mathbf{Q}} y \mapsto xy$ on elementary tensors be an isomorphism? It is obviously surjective, but this map is far from being injective. For example, since π is transcendental its powers $1, \pi, \pi^2, \ldots$ are linearly independent over $\mathbf{Q}$, so the tensors $\pi^i \otimes_{\mathbf{Q}} \pi^j$ in $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ are linearly independent over $\mathbf{Q}$. (Zorn's lemma lets us enlarge the powers of π to a $\mathbf{Q}$-basis of $\mathbf{R}$, so the tensors $\pi^i \otimes_{\mathbf{Q}} \pi^j$ are part of a $\mathbf{Q}$-basis by Theorem 4.9 and thus are linearly independent.) Then for $n \geq 2$ the tensor $\pi^n \otimes 1 + \pi^{n-1} \otimes \pi + \cdots + \pi \otimes \pi^{n-1} - (n-1)(1 \otimes \pi^n)$ in $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ is nonzero and its image in $\mathbf{R}$ is $(n-1)\pi^n - (n-1)\pi^n = 0$.

Another failed attempt at trying to make $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ look like $\mathbf{R}$ concretely comes from treating $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ as a real vector space using scaling on the left (or on the right, but that is a *different* scaling structure since $\pi \otimes 1 \neq 1 \otimes \pi$). Its dimension over $\mathbf{R}$ is infinite by Theorem 6.6, which is pretty far from the behavior of $\mathbf{R}$ as a real vector space.

Remark 6.22. Theorem 6.19 and its corollary remain true, by the same proofs, with localizations in places of fraction fields. If R is a ring, D is a multiplicative subset of R, and N and N' are $D^{-1}R$-modules, then the two $D^{-1}R$-module structures on $N \otimes_R N'$, using $D^{-1}R$-scaling on either N or N', are the same: for $x \in D^{-1}R$, $xn \otimes_R n' = n \otimes_R xn'$. Moreover, the natural $D^{-1}R$-module mapping $N \otimes_R N' \to N \otimes_{D^{-1}R} N'$ determined by $n \otimes_R n' \mapsto n \otimes_{D^{-1}R} n'$ on elementary tensors is an isomorphism of $D^{-1}R$-modules.

The next theorem collects a number of earlier tensor product isomorphisms for R-modules and shows the same maps are S-module isomorphisms when one of the R-modules in the tensor product is an S-module.

Theorem 6.23. *Let M and M' be R-modules and N and N' be S-modules.*

(1) *There is a unique S-module isomorphism*

$$M \otimes_R N \to N \otimes_R M$$

where $m \otimes n \mapsto n \otimes m$. In particular, $S \otimes_R M$ and $M \otimes_R S$ are isomorphic S-modules.

(2) *There are unique S-module isomorphisms*

$$(M \otimes_R N) \otimes_R M' \to N \otimes_R (M \otimes_R M')$$

where $(m \otimes n) \otimes m' \mapsto n \otimes (m \otimes m')$ *and*

$$(M \otimes_R N) \otimes_R M' \cong M \otimes_R (N \otimes_R M')$$

where $(m \otimes n) \otimes m' \mapsto m \otimes (n \otimes m')$.

(3) *There is a unique* S*-module isomorphism*

$$(M \otimes_R N) \otimes_S N' \to M \otimes_R (N \otimes_S N')$$

where $(m \otimes n) \otimes n' \mapsto m \otimes (n \otimes n')$.

(4) *There is a unique* S*-module isomorphism*

$$N \otimes_R (M \oplus M') \to (N \otimes_R M) \oplus (N \otimes_R M')$$

where $n \otimes (m, m') \mapsto (n, m) \otimes (n, m')$.

In the first, second, and fourth parts, we are using R-module tensor products only and then endowing them with S-module structure from one of the factors being an S-module (Theorem 6.14). In the third part we have both $\otimes_R$ and $\otimes_S$.

Proof. There is a canonical R-module isomorphism $M \otimes_R N \to N \otimes_R M$ where $m \otimes n \mapsto n \otimes m$. This map is S-linear using the S-module structure on both sides (check!), so it is an S-module isomorphism. This settles part 1.

Part 2, like part 1, only uses R-module tensor products, so there is an R-module isomorphism $\varphi\colon (M \otimes_R N) \otimes_R M' \to N \otimes_R (M \otimes_R M')$ where $\varphi((m \otimes n) \otimes m') = n \otimes (m \otimes m')$. Using the S-module structure on $M \otimes_R N$, $(M \otimes_R N) \otimes_R M'$, and $N \otimes_R (M \otimes_R M')$, φ is S-linear (check!), so it is an S-module isomorphism. To derive $(M \otimes_R N) \otimes_R M' \cong M \otimes_R (N \otimes_R M')$ from $(M \otimes_R N) \otimes_R M' \cong N \otimes_R (M \otimes_R M')$, use a few commutativity isomorphisms.

Part 3 resembles associativity of tensor products. We will in fact derive part 3 from such associativity for $\otimes_S$:

$$\begin{aligned}(M \otimes_R N) \otimes_S N' &\cong ((S \otimes_R M) \otimes_S N) \otimes_S N' \text{ by (6.3)}\\ &\cong (S \otimes_R M) \otimes_S (N \otimes_S N') \text{ by associativity of } \otimes_S\\ &\cong M \otimes_R (N \otimes_S N') \text{ by (6.3)}.\end{aligned}$$

These successive S-module isomorphisms have the effect

$$\begin{aligned}(m \otimes n) \otimes n' &\mapsto ((1 \otimes m) \otimes n) \otimes n'\\ &\mapsto (1 \otimes m) \otimes (n \otimes n')\\ &\mapsto m \otimes (n \otimes n'),\end{aligned}$$

which is what we wanted.

For part 4, there is an R-module isomorphism $N \otimes_R (M \oplus M') \to (N \otimes_R M) \oplus (N \otimes_R M')$ by Theorem 5.4. Now it's just a matter of checking this map is S-linear using the S-module structure on both sides coming from N being an S-module, and this is left to the reader. As an alternate proof, we have a chain of S-module isomorphisms

$$\begin{aligned}N \otimes_R (M \oplus M') &\cong N \otimes_S \left(S \otimes_R (M \oplus M')\right) \text{ by part 1 and (6.3)}\\ &\cong N \otimes_S ((S \otimes_R M) \oplus (S \otimes_R M')) \text{ by Theorem 6.10}\\ &\cong (N \otimes_S (S \otimes_R M)) \oplus (N \otimes_S (S \otimes_R M')) \text{ by Theorem 5.4}\\ &\cong (N \otimes_R M) \oplus (N \otimes_R M') \text{ by part 1 and (6.3)}.\end{aligned}$$

Of course one needs to trace through these isomorphisms to check the overall result has the effect intended on elementary tensors, and it does (exercise). $\square$

The last part of Theorem 6.23 extends to arbitrary direct sums: the natural R-module isomorphism $N \otimes_R \bigoplus_{i \in I} M_i \cong \bigoplus_{i \in I}(N \otimes_R M_i)$ is also an S-module isomorphism.

As an application of Theorem 6.23, we can show the base extension of an R-module tensor product "is" the S-module tensor product of the base extensions:

Corollary 6.24. *For R-modules M and M', there is a unique S-module isomorphism*

$$S \otimes_R (M \otimes_R M') \to (S \otimes_R M) \otimes_S (S \otimes_R M')$$

where $s \otimes_R (m \otimes_R m') \mapsto s((1 \otimes_R m) \otimes_S (1 \otimes_R m'))$.

Proof. Since $M \otimes_R M'$ is additively spanned by all $m \otimes m'$, $S \otimes_R (M \otimes_R M')$ is additively spanned by all $s \otimes_R (m \otimes_R m')$. Therefore an S-linear (or even additive) map out of $S \otimes_R (M \otimes_R M')$ is determined by its values on the tensors $s \otimes_R (m \otimes_R m')$.

We have the S-module isomorphisms

$$\begin{aligned} S \otimes_R (M \otimes_R M') &\cong M \otimes_R (S \otimes_R M') \text{ by Theorem 6.23(2)} \\ &\cong (S \otimes_R M) \otimes_S (S \otimes_R M') \text{ by (6.3).} \end{aligned}$$

The effect of these isomorphisms on $s \otimes_R (m \otimes_R m')$ is

$$\begin{aligned} s \otimes_R (m \otimes_R m') &\mapsto m \otimes_R (s \otimes_R m') \\ &\mapsto (1 \otimes_R m) \otimes_S (s \otimes_R m') \\ &= (1 \otimes_R m) \otimes_S s(1 \otimes_R m') \\ &= s((1 \otimes_R m) \otimes_S (1 \otimes_R m')), \end{aligned}$$

as desired. The effect of the inverse isomorphism on $(s_1 \otimes_R m) \otimes_S (s_2 \otimes_R m')$ is

$$\begin{aligned} (s_1 \otimes_R m) \otimes_S (s_2 \otimes_R m') &\mapsto m \otimes_R s_1(s_2 \otimes_R m') \\ &= m \otimes_R ((s_1 s_2) \otimes_R m') \\ &\mapsto s_1 s_2 \otimes_R (m \otimes_R m'). \end{aligned}$$

□

Theorem 6.24 could also be proved by showing the S-module $S \otimes_R (M \otimes_R M')$ has the universal mapping property of $(S \otimes_R M) \otimes_S (S \otimes_R M')$ as a tensor product of S-modules. That is left as an exercise.

Corollary 6.25. *For R-modules $M_1, \ldots, M_k$,*

$$S \otimes_R (M_1 \otimes_R \cdots \otimes_R M_k) \cong (S \otimes_R M_1) \otimes_S \cdots \otimes_S (S \otimes M_k)$$

as S-modules, where $s \otimes_S (m_1 \otimes_R \cdots \otimes_R m_k) \mapsto s((1 \otimes_R m_1) \otimes_S \cdots \otimes_S (1 \otimes_R m_k))$. In particular, $S \otimes_R (M^{\otimes_R k}) \cong (S \otimes_R M)^{\otimes_S k}$ as S-modules.

Proof. Induct on k. □

Example 6.26. For a real vector space V, $\mathbf{C} \otimes_{\mathbf{R}} (V \otimes_{\mathbf{R}} V) \cong (\mathbf{C} \otimes_{\mathbf{R}} V) \otimes_{\mathbf{C}} (\mathbf{C} \otimes_{\mathbf{R}} V)$. The middle tensor product sign on the right is over $\mathbf{C}$, not $\mathbf{R}$. Note that $\mathbf{C} \otimes_{\mathbf{R}} (V \otimes_{\mathbf{R}} V) \not\cong (\mathbf{C} \otimes_{\mathbf{R}} V) \otimes_{\mathbf{R}} (\mathbf{C} \otimes_{\mathbf{R}} V)$ when $V \neq 0$, as the two sides have different dimensions over $\mathbf{R}$ (what are they?).

The base extension $M \rightsquigarrow S \otimes_R M$ turns R-modules into S-modules in a systematic way. So does $M \rightsquigarrow M \otimes_R S$, and this is essentially the same construction. This suggests there should be a universal mapping problem about R-modules and S-modules that is solved by base extension, and there is: it is the universal device for turning each R-linear map from M to an S-module into an S-linear map of S-modules.

Theorem 6.27. *Let M be an R-module. For every S-module N and R-linear map $\varphi\colon M \to N$, there is a unique S-linear map $\varphi^S\colon S \otimes_R M \to N$ such that the diagram*

$$\begin{tikzcd} M \arrow[rr, "m\mapsto 1\otimes m"] \arrow[dr, "\varphi"'] & & S \otimes_R M \arrow[dl, dotted, "\varphi^S"] \\ & N & \end{tikzcd}$$

commutes.

This says the single R-linear map $M \to S \otimes_R M$ from M to an S-module explains all other R-linear maps from M to S-modules using composition of it with S-linear maps from $S \otimes_R M$ to S-modules.

Proof. Assume there is such an S-linear map φ^S. We will derive a formula for it on elementary tensors:

$$\varphi^S(s \otimes m) = \varphi^S(s(1 \otimes m)) = s\varphi^S(1 \otimes m) = s\varphi(m).$$

This shows φ^S is unique if it exists.

To prove existence, consider the function $S \times M \to N$ by $(s, m) \mapsto s\varphi(m)$. This is R-bilinear (check!), so there is an R-linear map $\varphi^S\colon S \otimes_R M \to N$ such that $\varphi^S(s \otimes m) = s\varphi(m)$. Using the S-module structure on $S \otimes_R M$, φ^S is S-linear. $\square$

For φ in $\mathrm{Hom}_R(M, N)$, φ^S is in $\mathrm{Hom}_S(S \otimes_R M, N)$. Because $\varphi^S(1 \otimes m) = \varphi(m)$, we can recover φ from φ^S. But even more is true.

Theorem 6.28. *Let M be an R-module and N be an S-module. The function $\varphi \mapsto \varphi^S$ is an S-module isomorphism $\mathrm{Hom}_R(M, N) \to \mathrm{Hom}_S(S \otimes_R M, N)$.*

How is $\mathrm{Hom}_R(M, N)$ an S-module? Values of these functions are in N, which is an S-module, so S turns scales each function $M \to N$ to a new function $M \to N$ by just scaling the values.

Proof. For φ and φ' in $\mathrm{Hom}_R(M, N)$, $(\varphi + \varphi')^S = \varphi^S + \varphi'^S$ and $(s\varphi)^S = s\varphi^S$ by checking both sides are equal on all elementary tensors in $S \otimes_R M$. Therefore $\varphi \mapsto \varphi^S$ is S-linear. Its injectivity is discussed above (φ^S determines φ).

For surjectivity, let $h\colon S \otimes_R M \to N$ be S-linear. Set $\varphi\colon M \to N$ by $\varphi(m) = h(1 \otimes m)$. Then φ is R-linear and $\varphi^S(s \otimes m) = s\varphi(m) = sh(1 \otimes m) = h(s(1 \otimes m)) = h(s \otimes m)$, so $h = \varphi^S$ since both are additive and are equal at all elementary tensors. $\square$

The S-module isomorphism

$$\mathrm{Hom}_R(M, N) \cong \mathrm{Hom}_S(S \otimes_R M, N) \tag{6.4}$$

should be thought of as analogous to the R-module isomorphism

$$\mathrm{Hom}_R(M, \mathrm{Hom}_R(N, P)) \cong \mathrm{Hom}_R(M \otimes_R N, P) \tag{6.5}$$

from Theorem 5.7, where $- \otimes_R N$ is left adjoint to $\mathrm{Hom}_R(N, -)$. (In (6.5), N and P are R-modules, not S-modules! We're using the same notation as in Theorem 5.7.) If we look at (6.4), we see $S \otimes_R -$ is applied to M on the right but nothing special is applied to N on the left. Yet there *is* something different about N on the two sides of (6.4). It is an S-module on the right side of (6.4), but on the left side it is being treated as an R-module (restriction of scalars). That changes N, but we have introduced no notation to reflect this. We still just write it as N. Let's now write $\mathrm{Res}_{S/R}(N)$ to denote N as an R-module. It is the same underlying additive group as N, but the scalars are now taken from R with the rule $rn = f(r)n$. The appearance of (6.4) now looks like this:

$$\mathrm{Hom}_R(M, \mathrm{Res}_{S/R}(N)) \cong \mathrm{Hom}_S(S \otimes_R M, N). \tag{6.6}$$

So extension of scalars (from R-modules to S-modules) is left adjoint to restriction of scalars (from S-modules to R-modules) in a similar way that $-\otimes_R M$ is left adjoint to $\mathrm{Hom}_R(M, -)$.

Using this new notation for restriction of scalars, the important S-module isomorphism (6.3) can be written more explicitly as

$$M \otimes_R \mathrm{Res}_{S/R}(N) \cong (S \otimes_R M) \otimes_S N,$$

Theorem 6.29. *Let M be an R-module and N and P be S-modules. There is an S-module isomorphism*

$$\mathrm{Hom}_S(M \otimes_S N, P) \cong \mathrm{Hom}_R(M, \mathrm{Res}_{S/R}(\mathrm{Hom}_S(N, P))).$$

Example 6.30. Taking $N = S$,

$$\mathrm{Hom}_S(S \otimes_R M, P) \cong \mathrm{Hom}_R(M, \mathrm{Res}_{S/R}(P))$$

since $\mathrm{Hom}_S(S, P) \cong P$. We have recovered $S \otimes_R -$ being left adjoint to $\mathrm{Res}_{S/R}$.

Example 6.31. Taking $S = R$, so N and P are now R-modules,

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

We have recovered $- \otimes_R N$ being left adjoint to $\mathrm{Hom}_R(N, -)$ for R-modules N.

These two consequences of Theorem 6.29 are results we have already seen, and in fact we are going to use them in the proof, so they are together equivalent to Theorem 6.29.

Proof. Since $M \otimes_R N \cong (S \otimes_R M) \otimes_S N$ as S-modules,

$$\mathrm{Hom}_S(M \otimes_R N, P) \cong \mathrm{Hom}_S((S \otimes_R M) \otimes_S N, P).$$

Since $- \otimes_S N$ is left adjoint to $\mathrm{Hom}_S(N, -)$,

$$\mathrm{Hom}_S((S \otimes_R M) \otimes_S N, P) \cong \mathrm{Hom}_S(S \otimes_R M, \mathrm{Hom}_S(N, P)).$$

Since $S \otimes_R -$ is left adjoint to $\mathrm{Res}_{S/R}$,

$$\mathrm{Hom}_S(S \otimes_R M, \mathrm{Hom}_S(N, P)) \cong \mathrm{Hom}_R(M, \mathrm{Res}_{S/R}(\mathrm{Hom}_S(N, P))).$$

Combining these three isomorphisms,

$$\mathrm{Hom}_S(M \otimes_R N, P) \cong \mathrm{Hom}_R(M, \mathrm{Res}_{S/R}(\mathrm{Hom}_S(N, P))).$$

Here is an explicit (overall) isomorphism. If $\varphi\colon M \otimes_R N \to P$ is S-linear there is an R-linear map $L_\varphi\colon M \to \mathrm{Hom}_S(N, P)$ by $L_\varphi(m) = \varphi(m \otimes (-))$. If $\psi\colon M \to \mathrm{Hom}_S(N, P)$ is R-linear then $M \times N \to P$ by $(m, n) \mapsto \psi(m)(n)$ is R-bilinear and $\psi(m)(sn) = s\psi(m)(n)$, so the corresponding R-linear map $\widetilde{L}_\psi\colon M \otimes_R N \to P$ where $\widetilde{L}_\psi(m \otimes n) = \psi(m)(n)$ is S-linear. The functions $\varphi \mapsto L_\varphi$ and $\psi \mapsto \widetilde{L}_\psi$ are S-linear and are inverses. $\square$

7. Tensors in Physics

In physics and engineering, tensors are often defined not in terms of multilinearity, but by the way tensors look in different coordinate systems. Here is a definition of a tensor that can be found (more or less) in most physics textbooks. Let V be a vector space[22] with dimension n. A *tensor of rank* 0 on V is a scalar. For $k \geq 1$, a *contravariant tensor of rank* k (on V) is an object T with n^k components in every coordinate system of V such that if $\{T^{i_1,\dots,i_k}\}_{1\leq i_1,\dots,i_k\leq n}$ and $\{\widetilde{T}^{i_1,\dots,i_k}\}_{1\leq i_1,\dots,i_k\leq n}$ are the components of T in two coordinate systems of V then

$$\widetilde{T}^{i_1,\dots,i_k} = \sum_{1\leq j_1,\dots,j_k\leq n} T^{j_1,\dots,j_k} a_{i_1j_1}\cdots a_{i_kj_k}, \tag{7.1}$$

where (a_{ij}) is the matrix expressing the first coordinate system of V in terms of the second. In short, a contravariant tensor of rank k is a "quantity that transforms by the rule (7.1)."

What is being described here, with components, is just an element of $V^{\otimes k}$. To see this, note that a coordinate system means a choice of a basis of V. For each basis $\{e_1,\dots,e_n\}$ of V, in which T has components $\{T^{i_1,\dots,i_k}\}_{1\leq i_1,\dots,i_k\leq n}$, make these numbers into the coefficients of the basis $\{e_{i_1}\otimes\cdots\otimes e_{i_k}\}$ of $V^{\otimes k}$:

$$\sum_{1\leq i_1,\dots,i_k\leq n} T^{i_1,\dots,i_k} e_{i_1}\otimes\cdots\otimes e_{i_k}.$$

This belongs to $V^{\otimes k}$. Let's express this sum in terms of a second basis ("coordinate system") $\{f_1,\dots,f_n\}$ of V. Writing $e_j = \sum_{i=1}^n a_{ij} f_i$, the above sum equals, after a notational change,

$$\begin{aligned}
&\sum_{1\leq j_1,\dots,j_k\leq n} T^{j_1,\dots,j_k} e_{j_1}\otimes\cdots\otimes e_{j_k}\\
&= \sum_{1\leq j_1,\dots,j_k\leq n} T^{j_1,\dots,j_k}\left(\sum_{i_1=1}^n a_{i_1j_1} f_{i_1}\right)\otimes\cdots\otimes\left(\sum_{i_k=1}^n a_{i_kj_k} f_{i_k}\right)\\
&= \sum_{1\leq i_1,\dots,i_k\leq n}\left(\sum_{1\leq j_1,\dots,j_k\leq n} T^{j_1,\dots,j_k} a_{i_1j_1}\cdots a_{i_kj_k}\right) f_{i_1}\otimes\cdots\otimes f_{i_k}\\
&= \sum_{1\leq i_1,\dots,i_k\leq n} \widetilde{T}^{i_1,\dots,i_k} f_{i_1}\otimes\cdots\otimes f_{i_k} \quad \text{by (7.1)}.
\end{aligned}$$

So the physicist's contravariant rank k tensor T is just all the different coordinate representations of a single element of $V^{\otimes k}$.[23]

Switching from tensor powers of V to tensor powers of its dual space $V^\vee$, we now want to compare the representations of an element of $(V^\vee)^{\otimes \ell}$ in coordinate systems built from the two dual bases $e_1^\vee,\dots,e_n^\vee$ and $f_1^\vee,\dots,f_n^\vee$ of $V^\vee$. The formula we find will be similar to (7.1), but with a crucial change.

To align calculations with the way they're done in physics (and differential geometry), from now on write the dual bases of $e_1,\dots,e_n$ and $f_1,\dots,f_n$ as $e^1,\dots,e^n$ and $f^1,\dots,f^n$, not as $e_1^\vee,\dots,e_n^\vee$ and $f_1^\vee,\dots,f_n^\vee$. So $e^i(e_j) = f^i(f_j) = \delta_{ij}$ for all i and j. When a basis of

[22] The physicist is interested only in real or complex vector spaces.

[23] Strictly speaking this is false. Not all bases of $V^{\otimes k}$ are the k-fold elementary tensors built from a basis of V, so we don't actually see *all* coordinate representations of T in this way. Let's not dwell on that.

V is $e_1, \ldots, e_n$, and its dual basis is $e^1, \ldots, e^n$, general elements of V and $V^\vee$ are written as $\sum_{i=1}^n a^i e_i$ and $\sum_{i=1}^n b_i e^i$, respectively. A basis of V always has lower indices and its coefficients have upper indices, while a basis of $V^\vee$ always has upper indices and its coefficients have lower indices. This is consistent since the coefficients a^i of the vector $\sum_{i=1}^n a^i e_i$ in V are the values of $e^1, \ldots, e^n$ on this vector. Coordinate functions of a basis of V lie in $V^\vee$ and, by duality, coordinate functions of a basis of $V^\vee$ lie in $(V^\vee)^\vee \cong V$.

Pick a mathematician's tensor $\mathsf{T} \in (V^\vee)^{\otimes \ell}$ and write it in the basis $\{e^{i_1} \otimes \cdots \otimes e^{i_\ell}\}$ as

$$\mathsf{T} = \sum_{1 \leq i_1, \ldots, i_\ell \leq n} T_{i_1, \ldots, i_\ell} e^{i_1} \otimes \cdots \otimes e^{i_\ell}, \tag{7.2}$$

where lower indices on the coefficients, rather than upper indices, are consistent with the idea that this is a dual object (lies in a tensor power of $V^\vee$). To express T in terms of the second basis $\{f^{i_1} \otimes \cdots \otimes f^{i_\ell}\}$ of $(V^\vee)^{\otimes \ell}$, we want to express the e^j's in terms of the f^i's.

We already wrote $e_j = \sum_{i=1}^n a_{ij} f_i$ in V for all j, and it turns out that

$$e_j = \sum_{i=1}^n a_{ij} f_i \text{ for all } j \Longrightarrow f^j = \sum_{i=1}^n a_{ji} e^i \text{ for all } j. \tag{7.3}$$

Indeed, at each basis vector e_r for $r = 1, \ldots, n$, both f^j and $\sum_{i=1}^n a_{ji} e^i$ in $V^\vee$ have the same value a_{jr}, so they are equal. We meet *transposed* matrix entries (a_{ji}) on the right side of (7.3) in an essential way: j in (7.3) is the second index of a_{ij} and the first index of a_{ji}. It is a fact of life that passing to the dual space involves a transpose. Alas, this change of basis formula in $V^\vee$, from e^i's to f^j's, is not the direction we need (to transform (7.2) we want e^j in terms of f^i, not f^j in terms of e^i), so we will bring in an inverse matrix.

The inverse of the matrix (a_{ij}) is denoted (a^{ij}). Just as (a_{ij}) describes the e_j's in terms of the f_i's (that's the definition of (a_{ij})) and, transposed, the f^j's in terms of the e^i's as in (7.3), the inverse matrix (a^{ij}) describes the f_j's in terms of the e_i's and, transposed, the e^j's in terms of the f^i's:

$$e_j = \sum_{i=1}^n a_{ij} f_i \text{ for all } j \Longrightarrow f_j = \sum_{i=1}^n a^{ij} e_i, \text{ and } e^j = \sum_{i=1}^n a^{ji} f^i \text{ for all } j. \tag{7.4}$$

Example 7.1. Let $e_1 = f_1 + 2f_2$ and $e_2 = 3f_2$. By simple algebra, $f_1 = e_1 - \frac{2}{3}e_2$ and $f_2 = \frac{1}{3}e_2$. In $V^\vee$ we have $f^1 = e^1$ (check both sides are the same at f_1 and f_2) and $f^2 = 2e^1 + 3e^2$ (check both sides are the same at f_1 and f_2), and $e^1 = f^1$ and $e^2 = -\frac{2}{3}f^1 + \frac{1}{3}f^2$. This is consistent with (7.4) using

$$(a_{ij}) = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \text{ and } (a^{ij}) = \begin{pmatrix} 1 & 0 \\ -2/3 & 1/3 \end{pmatrix},$$

which are inverses.

The only change of basis we need for (7.2) is e^j in terms of f^i, so let's isolate that part of (7.4):

$$e_j = \sum_{i=1}^n a_{ij} f_i \text{ for all } j \Longrightarrow e^j = \sum_{i=1}^n a^{ji} f^i \text{ for all } j. \tag{7.5}$$

As a reminder, (a^{ji}) is the transpose of the inverse of the matrix (a_{ij}).

Returning to (7.2),

$$\begin{aligned}
\mathsf{T} &= \sum_{1\le j_1,\dots,j_\ell\le n} T_{j_1,\dots,j_\ell} e^{j_1}\otimes\cdots\otimes e^{j_\ell}\\
&= \sum_{1\le j_1,\dots,j_\ell\le n} T_{j_1,\dots,j_\ell}\left(\sum_{i_1=1}^n a^{j_1 i_1} f^{i_1}\right)\otimes\cdots\otimes\left(\sum_{i_\ell=1}^n a^{j_\ell i_\ell} f^{i_\ell}\right) \quad \text{by (7.2)}\\
&= \sum_{1\le i_1,\dots,i_\ell\le n}\left(\sum_{1\le j_1,\dots,j_\ell\le n} T_{j_1,\dots,j_\ell} a^{j_1 i_1}\cdots a^{j_\ell i_\ell}\right) f^{i_1}\otimes\cdots\otimes f^{i_\ell}.
\end{aligned}$$

The component-based approach to $(V^\vee)^{\otimes\ell}$ is based on the above calculation. Any "quantity that transforms by the rule"

$$(7.6)\qquad \widetilde{T}_{i_1,\dots,i_\ell} = \sum_{1\le j_1,\dots,j_\ell\le n} T_{j_1,\dots,j_\ell} a^{j_1 i_1}\cdots a^{j_\ell i_\ell}$$

when passing from the coordinate system $\{e^1,\dots,e^n\}$ to the coordinate system $\{f^1,\dots,f^n\}$ of $V^\vee$, is called a *covariant tensor of rank ℓ*. This is just an element of $(V^\vee)^{\otimes\ell}$, and (7.6) explains operationally how different coordinate representations of this tensor are related to one another.

The rules (7.1) for components in $V^{\otimes k}$ and (7.6) for components in $(V^\vee)^{\otimes\ell}$ are different, and not just on account of the convention about indices being upper on tensor components in (7.1) and lower on tensor components in (7.6). If we place (7.1) and (7.6) side by side and, to avoid being distracted by tensor index notational conventions, we *temporarily* make all tensor-component indices lower and give the tensor components the same number of indices ($\ell = k$, so we are in $V^{\otimes k}$ and $(V^\vee)^{\otimes k}$), we obtain this:

$$\widetilde{T}_{i_1,\dots,i_k} = \sum_{1\le j_1,\dots,j_k\le n} T_{j_1,\dots,j_k} a_{i_1 j_1}\cdots a_{i_k j_k}, \quad \widetilde{T}_{i_1,\dots,i_k} = \sum_{1\le j_1,\dots,j_k\le n} T_{j_1,\dots,j_k} a^{j_1 i_1}\cdots a^{j_k i_k}.$$

We did *not* lower the indices of a^{ji} in the second sum because its indices reflect something serious: (a_{ij}) is the matrix expressing a change of coordinates in V and (a^{ji}) is the matrix expressing the dual change of coordinates in $V^\vee$ in the same direction (see (7.5)). The use of a_{ij} or a^{ji} is the difference between the transformation rules in tensor powers of V and tensor powers of $V^\vee$. Both of the transformation rules involve a multilinear change of coordinates (as evidenced by the multiple products in the sums), but in the first rule the summation indices appear in the multipliers $a_{i_r j_r}$ as the second index, while in the second rule the summation indices appear in the multipliers $a^{j_r i_r}$ as the first index. This swap happens because physicists *always* start a change of basis in V, and passing to the effect in $V^\vee$ necessitates a transpose (and inverse). The reason for systematically using upper indices on tensor components satisfying (7.1) and lower indices on tensor components satisfying (7.6) is to know at a glance (with experience) what type of transformation rule the tensor components will satisfy under a change in coordinates.

Here is some terminology about tensors that is used by physicists.

- A contravariant tensor of rank k, which is an indexed quantity $T^{i_1\dots i_k}$ that transforms by (7.1), is also called a *tensor of rank k with upper indices* (easier to remember!).
- A covariant tensor of rank ℓ, which is an indexed quantity $T_{j_1\dots j_\ell}$ that transforms by (7.6), is also called a *tensor of rank ℓ with lower indices*.

- An indexed quantity $T^{i_1 \dots i_k}_{j_1 \dots j_\ell}$ that transforms by the rule

$$\widetilde{T}^{i_1 \dots i_k}_{j_1 \dots j_\ell} = \sum_{\substack{1 \leq p_1, \dots, p_k \leq n \\ 1 \leq q_1, \dots, q_\ell \leq n}} T^{p_1 \dots p_k}_{q_1 \dots q_\ell} a_{i_1 p_1} \cdots a_{i_k p_k} a^{q_1 j_1} \cdots a^{q_\ell j_\ell} \tag{7.7}$$

  is called a *tensor of type* (k, ℓ) *and rank* $k + \ell$. This "quantity" is just an element of $V^{\otimes k} \otimes (V^\vee)^{\otimes \ell}$ written in terms of elementary tensor product bases produced from two bases of V (check!). For instance, elements of $V \otimes V$, $V \otimes V^\vee$, and $V^\vee \otimes V^\vee$ are all rank 2 tensors. An element of $V^{\otimes 2} \otimes V^\vee$ is denoted $T^{i_1 i_2}_j$.

  If we permute the order of the spaces in the tensor product from the conventional "first every V, then every $V^\vee$," then the indexing rule on tensors needs to be adapted: $V \otimes V^\vee \otimes V$ is not the same space as $V \otimes V \otimes V^\vee$, so we shouldn't write its tensor components as $T^{i_1 i_2}_j$. Write them as $T^{i_1}{}_j{}^{i_2}$, so that as we read indices from left to right we see each index in the order its corresponding space appears in $V \otimes V^\vee \otimes V$: upper indices for V and lower indices for $V^\vee$.

Example 7.2. To compare transformation rules for rank 2 tensors in $V^{\otimes 2}$ (type (2,0)), $(V^\vee)^{\otimes 2}$ (type (0,2)), and $V \otimes V^\vee$ (type (1,1)), let bases $\{e_1, \dots, e_n\}$ and $\{f_1, \dots, f_n\}$ of V be related by numbers a_{ij} as in (7.3) and (7.4).

Case 1: (2,0)-tensors. By (7.3), in $V^{\otimes 2}$ we have $\sum_{j_1, j_2} T^{j_1, j_2} e_{j_1} \otimes e_{j_2} = \sum_{i_1, i_2} \widetilde{T}^{i_1, i_2} f_{i_1} \otimes f_{i_2}$ where

$$\widetilde{T}^{i_1 i_2} = \sum_{j_1, j_2} T^{j_1 j_2} a_{i_1 j_1} a_{i_2 j_2}. \tag{7.8}$$

Case 2: (0,2)-tensors. In $(V^\vee)^{\otimes 2}$, $\sum_{j_1, j_2} T_{j_1, j_2} e^{j_1} \otimes e^{j_2} = \sum_{i_1, i_2} \widetilde{T}_{i_1, i_2} f^{i_1} \otimes f^{i_2}$ where

$$\widetilde{T}_{i_1 i_2} = \sum_{j_1, j_2} T_{j_1 j_2} a^{j_1 i_1} a^{j_2 i_2}, \tag{7.9}$$

with the matrix (a^{ij}) being the inverse of (a_{ij}).

Case 3: (1,1)-tensors. In $V \otimes V^\vee$, $\sum_{j_1, j_2} T^{j_1}_{j_2} e_{j_1} \otimes e^{j_2} = \sum_{i_1, i_2} \widetilde{T}^{i_1}_{i_2} f_{i_1} \otimes f^{i_2}$ where

$$\widetilde{T}^{i_1}_{i_2} = \sum_{j_1, j_2} T^{j_1}_{j_2} a_{i_1 j_1} a^{j_2 i_2}, \tag{7.10}$$

The n^2 components of such tensors relative to the basis $\{e_1, \dots, e_n\}$ can be put into an $n \times n$ matrix (T^{ij}), (T_{ij}), or (T^i_j). We can rewrite (7.8), (7.9), and (7.10) so the sums on the right look like formulas from multiplying 3 matrices:

$$\widetilde{T}^{i_1 i_2} = \sum_{j_1, j_2} a_{i_1 j_1} T^{j_1 j_2} a_{i_2 j_2}, \quad \widetilde{T}_{i_1 i_2} = \sum_{j_1, j_2} a^{j_1 i_1} T_{j_1 j_2} a^{j_2 i_2}, \quad \widetilde{T}^{i_1}_{i_2} = \sum_{j_1, j_2} a_{i_1 j_1} T^{j_1}_{j_2} a^{j_2 i_2}.$$

By how indices in these sums *repeat*, the matrix of components of a tensor of rank 2 transform as indicated in the table below.

Type	Transformation Rule
(2,0)	$(\widetilde{T}^{ij}) = (a_{ij})(T^{ij})(a_{ij})^\top$
(0,2)	$(\widetilde{T}_{ij}) = (a^{ij})^\top (T_{ij})(a^{ij})$
(1,1)	$(\widetilde{T}^i_j) = (a_{ij})(T^i_j)(a_{ij})^{-1}$

The (0,2) case is how the coefficients in an n-variable quadratic form change under a linear change of variables and the (1,1) case is how the components of an $n \times n$ matrix change after a change of basis. This is why in physics a quadratic form on a vector space (such as a spacetime metric in relativity) may be called a (0,2)-tensor and a linear map of a vector space to itself may be called a (1,1)-tensor. We saw the interpretation of (1,1)-tensors as linear maps before, without coordinates: from Example 5.10, $V \otimes V^\vee \cong \mathrm{Hom}(V, V)$. **Warning**: the "rank" of a linear map $V \to V$ (dimension of its image) has nothing to do with its "rank" as a tensor, which is always 2.

While the vector spaces V and $V^\vee$ are not literally the same, they are isomorphic. If we fix an isomorphism between them and use it everywhere to replace $V^\vee$ with V then the different spaces of rank 2 tensors can all be made to look like $V^{\otimes 2}$, a process called "raising indices" since it turns T_{ij} and T^i_j into T^{ij}. This is done very often in geometry and physics since $\mathbf{R}^n$ is isomorphic to its dual space using the standard dot product to identify $(\mathbf{R}^n)^\vee$ with $\mathbf{R}^n$.

Let's compare how the mathematician and physicist think about a tensor:

- (Mathematician) Tensors belongs to a tensor space, which is a module defined by a multilinear universal mapping property.
- (Physicist) "Tensors are systems of components organized by one or more indices that transform according to specific rules under a set of transformations."[24]

In a tensor product of vector spaces, mathematicians and physicists can check two tensors t and t' are equal *in the same way*: check t and t' have the same components in one coordinate system. (Physicists don't deal with modules that aren't vector spaces, so they always have bases available.) The reason mathematicians and physicists consider this to be a sufficient test of equality is not the same. The mathematician thinks about the condition $t = t'$ in a coordinate-free way and knows that to check $t = t'$ it suffices to check t and t' have the same coordinates in one basis. The physicist considers the condition $t = t'$ to mean (by definition!) that the components of t and t' match in all coordinate systems, and the *multilinear* transformation rule (7.6), or (7.7), on tensors implies that if the components of t and t' are equal in one coordinate system then they are equal in every coordinate system. That's why the physicist is content to look in just one coordinate system.

An operation on tensors (like the flip $v \otimes w \mapsto w \otimes v$ on $V^{\otimes 2}$) is checked to be well-defined by the mathematician and physicist *in different ways*. The mathematician checks the operation respects the universal mapping property that defines tensor products, while the physicist checks the explicit formula for the operation on elementary tensors (such as $v \otimes w \mapsto w \otimes v$ on $V^{\otimes 2}$) changes in different coordinate systems by the tensor transformation rule (like (7.1)). The physicist would say an operation on tensors makes sense because it "transforms tensorially," which in more expansive terms means that the formulas for the operation in two different coordinate systems are *related by a multilinear change of variables*. However, textbooks on classical mechanics and quantum mechanics that treat tensors don't seem to use the word "multilinear," even though that word describes exactly what is going on. Instead, these textbooks nearly always say that a tensor's components transform by a "definite rule" or a "specific rule," which doesn't seem to have an actual meaning; isn't every computational rule a specific rule? Graduate textbooks on general relativity are an exception to this habit: [3], [10], and [17] all define tensors in terms of multilinearity.[25]

[24] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 6th ed., p. 136.

[25] I thank Don Marolf for bringing this point to my attention.

While mathematicians may shake their heads and wonder how physicists can work with tensors in terms of components, that viewpoint is crucial to understanding how tensors show up in physics (as well as being the way tensors were handled in mathematics until work of Murray and von Neumann [11] and Whitney [18] in the 1930s). The physical meaning of a vector is not just displacement, but linear displacement. For instance, forces at a point combine in the same way that vectors add (this is an experimental observation), so force is treated as a vector. The physical meaning of a tensor is *multilinear displacement.*[26]. That means each quantity (mathematical or physical) whose descriptions in two different coordinate systems are related to each other in the same way as coordinates of a tensor in two different coordinate systems is asking to be mathematically described as a tensor. Moreover, the transformation formula for that quantity in different coordinate systems tells the physicist what indexing to use on the tensor (*e.g.*, whether the tensor description of the quantity should have upper indices, lower indices, or both).

Example 7.3. The most basic example of a rank-2 tensor in mechanics is the stress tensor. When a force is applied to a body the stress it imparts at a point may not be in the direction of the force but in some other direction (compressing a piece of clay, say, can push it out orthogonally to the direction of the force), so stress is described by a linear transformation, and thus is a rank-2 tensor since $\mathrm{End}(V) \cong V^\vee \otimes V$ (Example 5.10). Since the stress from an applied force can act in different directions at different points, the stress tensor is not really a single tensor but rather is a varying family of tensors at different points: stress is a tensor *field*, which is a generalization of a vector field.

The end of that example is pervasive: tensors in mechanics, electromagnetism, and relativity are always part of a tensor field. A change of variables between coordinate systems $\mathsf{x} = \{x^i\}$ and $\mathsf{y} = \{y^i\}$ in a region of $\mathbf{R}^n$ involves partial derivatives $\frac{\partial y^i}{\partial x^j}$ or (in the reverse direction) $\frac{\partial x^j}{\partial y^i}$, and the tensor transformation rules occur with $\frac{\partial y^i}{\partial x^j}$ and $\frac{\partial x^j}{\partial y^i}$, which vary from point to point, in the role of a_{ij} and a^{ij}. For example, a tensor of rank 2 with upper indices is a doubly-indexed quantity $T^{ij}(\mathsf{x})$ in each coordinate system x, such that in a coordinate system y its components are

$$\widetilde{T}^{ij}(\mathsf{y}) = \sum_{r,s=1}^{n} T^{rs}(\mathsf{x}) \frac{\partial y_i}{\partial x_r} \frac{\partial y_j}{\partial x_s},$$

which should be compared to (7.1) with $k = 2$.

To see a physicist introduce tensors (really, tensor fields) as indexed quantities, watch Leonard Susskind's lectures on general relativity on YouTube from 2009, particularly lecture 3 (tensors first appear 42 minutes in, although some notation is introduced earlier) and lecture 4. In lecture 5 tensor calculus (covariant differentiation of tensor fields) is introduced.

Besides classical mechanics, electromagnetism, and relativity, tensors play an essential role in quantum mechanics, but for rather different reasons than we've seen already. In classical mechanics, the states of a system are modeled by the points on a finite-dimensional manifold, and when we combine two systems the corresponding manifold is the direct product of the manifolds for the original two systems. The states of a quantum system, on the other hand, are represented by the nonzero vectors (really, the 1-dimensional subspaces) in a complex Hilbert space, such as $L^2(\mathbf{R}^6)$. (A point in $\mathbf{R}^6$ has three position and three momentum coordinates, which is the classical description of a particle.) When we combine

[26]Tensors are "multilinear functions of several directions" in [15, p. 9].

two quantum systems, its corresponding Hilbert space is the tensor product of the original two Hilbert spaces, essentially because $L^2(\mathbf{R}^6 \times \mathbf{R}^6) = L^2(\mathbf{R}^6) \otimes_{\mathbf{C}} L^2(\mathbf{R}^6)$, which is the analytic[27] analogue of $R[X,Y] \cong R[X] \otimes_R R[Y]$. While in classical mechanics, electromagnetism, and relativity a physicist uses specific tensor fields (*e.g.*, the stress tensor, electromagnetic field tensor, or metric tensor), in quantum mechanics it is a whole tensor product space $H_1 \otimes_{\mathbf{C}} H_2$ that gets used. A video of a physicist introducing tensor products of Hilbert spaces on YouTube is Frederic Schuller's lecture 14 on quantum mechanics, where he writes an elementary tensor as $v \boxtimes w$ rather than $v \otimes w$ to avoid confusion with the use of $\otimes$ in the notation of the vector space $H_1 \otimes_{\mathbf{C}} H_2$.

The difference between a direct product of manifolds $M \times N$ and a tensor product of vector spaces $H_1 \otimes_{\mathbf{C}} H_2$ reflects mathematically some of the non-intuitive features of quantum mechanics. Every point in $M \times N$ is a pair (x, y) where $x \in M$ and $y \in N$, so we get a direct link from a point in $M \times N$ to something in M and something in N. On the other hand, most tensors in $H_1 \otimes_{\mathbf{C}} H_2$ are *not* elementary, and a non-elementary tensor in $H_1 \otimes_{\mathbf{C}} H_2$ has no simple-minded description in terms of a pair of elements of H_1 and H_2. Quantum states in $H_1 \otimes_{\mathbf{C}} H_2$ that correspond to non-elementary tensors are called *entangled states*, and they reflect the difficulty of trying to describe quantum phenomena for a combined system (*e.g.*, the two-slit experiment) in terms of quantum states of the two original systems individually. I've been told that physics students who get used to computing with tensors in relativity by learning to work with the "transform by a definite rule" description of tensors find the role of tensors in quantum mechanics to be difficult to learn, because the conceptual role of the tensors is so different.

We'll end this discussion of tensors in physics with a story. I was the math consultant for the 4th edition of the American Heritage Dictionary of the English Language (2000). The editors sent me all the words in the 3rd edition with mathematical definitions, and I had to find and correct the errors. Early on I came across a word I had never heard of before: *dyad*. It was defined in the 3rd edition as "an operator represented as a pair of vectors juxtaposed without multiplication." That's a ridiculous definition, as it conveys no meaning at all. I obviously had to fix this definition, but first I had to know what the word meant! In a physics book[28] a dyad is defined as "a pair of vectors, written in a definite order **ab**." This is just as useless, but the physics book also does something with dyads, which gives a clue about what they really are. The product of a dyad **ab** with a vector **c** is $\mathbf{a}(\mathbf{b} \cdot \mathbf{c})$, where $\mathbf{b} \cdot \mathbf{c}$ is the usual dot product ($\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$ are all vectors in $\mathbf{R}^n$). This reveals what a dyad is. Do you see it? Dotting with **b** is an element of the dual space $(\mathbf{R}^n)^\vee$, so the effect of **ab** on **c** is reminiscient of the way $V \otimes V^\vee$ acts on V by $(v \otimes \varphi)(w) = \varphi(w)v$. A dyad is the same thing as an elementary tensor $v \otimes \varphi$ in $\mathbf{R}^n \otimes (\mathbf{R}^n)^\vee$. In the 4th edition of the dictionary, I included two definitions for a dyad. For the general reader, a dyad is "a function that draws a correspondence[29] from any vector **u** to the vector $(\mathbf{v} \cdot \mathbf{u})\mathbf{w}$ and is denoted **vw**, where **v** and **w** are a fixed pair of vectors and $\mathbf{v} \cdot \mathbf{u}$ is the scalar product of **v** and **u**. For example, if $\mathbf{v} = (2, 3, 1)$, $\mathbf{w} = (0, -1, 4)$, and $\mathbf{u} = (a, b, c)$, then the dyad **vw** draws a correspondence from **u** to $(2a + 3b + c)\mathbf{w}$." The more concise second definition was: a dyad is "a tensor formed from a vector in a vector space and a linear functional on that vector space." Unfortunately, the definition of "tensor" in the dictionary is "A set of quantities that obey certain transformation laws relating the bases in one generalized

[27]This tensor product should be a completed tensor product, including infinite sums of products $f(\mathbf{x})g(\mathbf{y})$.

[28]H. Goldstein, *Classical Mechanics*, 2nd ed., p. 194

[29]Yes, this terminology sucks. Blame the unknown editor at the dictionary for that one.

coordinate system to those of another and involving partial derivative sums. Vectors are simple tensors." That is really the definition of a tensor field, and that sense of the word tensor is incompatible with my concise definition of a dyad in terms of tensors.

More general than a dyad is a *dyadic*, which is a sum of dyads: $\mathbf{ab}+\mathbf{cd}+\dots$. So a dyadic is a general tensor in $\mathbf{R}^n \otimes_{\mathbf{R}} (\mathbf{R}^n)^\vee \cong \mathrm{Hom}_{\mathbf{R}}(\mathbf{R}^n, \mathbf{R}^n)$. In other words, a dyadic is an $n \times n$ real matrix. The terminology of dyads and dyadics goes back to Gibbs [5, Chap. 3], who championed the development of linear and multilinear algebra, including his indeterminate product (that is, the tensor product), under the name "multiple algebra."

References

[1] D. V. Alekseevskij, V. V. Lychagin, A. M. Vinogradov, "Geometry I," Springer-Verlag, Berlin, 1991.

[2] N. Bourbaki, Livre II Algèbre Chapitre III (état 4) Algèbre Multilinéaire, http://archives-bourbaki.ahp-numerique.fr/archive/files/97a9fed708bdde4dc55547ab5a8ff943.pdf.

[3] S. Carroll, "Spacetime and Geometry: An Introduction to General Relativity," Benjamin Cummings, 2003.

[4] D. Eisenbud and J. Harris, "The Geometry of Schemes", Springer-Verlag, New York, 2000.

[5] J. W. Gibbs, "Elements of Vector Analysis Arranged for the Use of Students in Physics," Tuttle, Morehouse & Taylor, New Haven, 1884.

[6] J. W. Gibbs, *On Multiple Algebra*, Proceedings of the American Association for the Advancement of Science, **35** (1886). Available online at http://archive.org/details/onmultiplealgeb00gibbgoog.

[7] R. Grone, *Decomposable Tensors as a Quadratic Variety*, Proc. Amer. Math. Soc. **64** (1977), 227–230.

[8] P. Halmos, "Finite-Dimensional Vector Spaces," Springer-Verlag, New York, 1974.

[9] R. Hermann, "Ricci and Levi-Civita's Tensor Analysis Paper" Math Sci Press, Brookline, 1975.

[10] C. W. Misner, K. S. Thorne, and J. A. Wheeler, "Gravitation," W. H. Freeman and Co., San Francisco, 1973.

[11] F. J. Murray and J. von Neumann, *On Rings of Operators*, Annals Math. **37** (1936), 116–229.

[12] B. O'Neill, "Semi-Riemannian Geometry," Academic Press, New York, 1983.

[13] B. Osgood, Chapter 8 of The Fourier Transform and its Applications, http://see.stanford.edu/materials/lsoftaee261/chap8.pdf.

[14] G. Ricci, T. Levi-Civita, *Méthodes de Calcul Différentiel Absolu et Leurs Applications*, Math. Annalen **54** (1901), 125–201.

[15] F. J. Temple, "Cartesian Tensors," Wiley, New York, 1960.

[16] W. Voigt, *Die fundamentalen physikalischen Eigenschaften der Krystalle in elementarer Darstellung*, Verlag von Veit & Comp., Leipzig, 1898.

[17] R. Wald, "General Relativity," Univ. Chicago Press, Chicago, 1984.

[18] H. Whitney, *Tensor Products of Abelian Groups*, Duke Math. Journal **4** (1938), 495–528.