TENSOR PRODUCTS II

KEITH CONRAD

1. Introduction

Continuing our study of tensor products, we will see how to combine two linear maps $M \longrightarrow M'$ and $N \longrightarrow N'$ into a linear map $M \otimes_R N \to M' \otimes_R N'$. This leads to flat modules and linear maps between base extensions. Then we will look at special features of tensor products of vector spaces (including contraction), the tensor products of R-algebras, and finally the tensor algebra of an R-module.

2. Tensor Products of Linear Maps

If $M \xrightarrow{\varphi} M'$ and $N \xrightarrow{\psi} N'$ are linear, then we get a linear map between the direct sums, $M \oplus N \xrightarrow{\varphi\oplus\psi} M' \oplus N'$, defined by $(\varphi \oplus \psi)(m, n) = (\varphi(m), \psi(n))$. We want to define a linear map $M \otimes_R N \longrightarrow M' \otimes_R N'$ such that $m \otimes n \mapsto \varphi(m) \otimes \psi(n)$.

Start with the map $M \times N \longrightarrow M' \otimes_R N'$ where $(m, n) \mapsto \varphi(m) \otimes \psi(n)$. This is R-bilinear, so the universal mapping property of the tensor product gives us an R-linear map $M \otimes_R N \xrightarrow{\varphi\otimes\psi} M' \otimes_R N'$ where $(\varphi \otimes \psi)(m \otimes n) = \varphi(m) \otimes \psi(n)$, and more generally

$$(\varphi \otimes \psi)(m_1 \otimes n_1 + \cdots + m_k \otimes n_k) = \varphi(m_1) \otimes \psi(n_1) + \cdots + \varphi(m_k) \otimes \psi(n_k).$$

We call $\varphi \otimes \psi$ the *tensor product* of φ and ψ, but be careful to appreciate that $\varphi \otimes \psi$ is *not* denoting an elementary tensor. This is just notation for a new linear map on $M \otimes_R N$.

When $M \xrightarrow{\varphi} M'$ is linear, the linear maps $N \otimes_R M \xrightarrow{1\otimes\varphi} N \otimes_R M'$ or $M \otimes_R N \xrightarrow{\varphi\otimes 1} M' \otimes_R N$ are called *tensoring with* N. The map on N is the identity, so $(1 \otimes \varphi)(n \otimes m) = n \otimes \varphi(m)$ and $(\varphi \otimes 1)(m \otimes n) = \varphi(m) \otimes n$. This construction will be particularly important for base extensions in Section 4.

Example 2.1. Tensoring inclusion $a\mathbf{Z} \xrightarrow{i} \mathbf{Z}$ with $\mathbf{Z}/b\mathbf{Z}$ is $a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \xrightarrow{i\otimes 1} \mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z}$, where $(i \otimes 1)(ax \otimes y \bmod b) = ax \otimes y \bmod b$. Since $\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \cong \mathbf{Z}/b\mathbf{Z}$ by multiplication, we can regard $i \otimes 1$ as a function $a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \to \mathbf{Z}/b\mathbf{Z}$ where $ax \otimes y \bmod b \mapsto axy \bmod b$. Its image is $\{az \bmod b : z \in \mathbf{Z}/b\mathbf{Z}\}$, which is $d\mathbf{Z}/b\mathbf{Z}$ where $d = (a, b)$; this is 0 if $b \mid a$ and is $\mathbf{Z}/b\mathbf{Z}$ if $(a, b) = 1$.

Example 2.2. Let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and $A' = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$ in $\mathrm{M}_2(R)$. Then A and A' are both linear maps $R^2 \to R^2$, so $A \otimes A'$ is a linear map from $(R^2)^{\otimes 2} = R^2 \otimes_R R^2$ back to itself. Writing e_1 and e_2 for the standard basis vectors of R^2, let's compute the matrix for $A \otimes A'$ on $(R^2)^{\otimes 2}$

with respect to the basis $\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\}$. By definition,

$$\begin{aligned}
(A \otimes A')(e_1 \otimes e_1) &= Ae_1 \otimes A'e_1 \\
&= (ae_1 + ce_2) \otimes (a'e_1 + c'e_2) \\
&= aa'e_1 \otimes e_1 + ac'e_1 \otimes e_2 + ca'e_2 \otimes e_1 + cc'e_2 \otimes e_2, \\
(A \otimes A')(e_1 \otimes e_2) &= Ae_1 \otimes A'e_2 \\
&= (ae_1 + ce_2) \otimes (b'e_1 + d'e_2) \\
&= cb'e_1 \otimes e_1 + ad'e_1 \otimes e_2 + cb'e_2 \otimes e_2 + cd'e_2 \otimes e_2,
\end{aligned}$$

and similarly

$$\begin{aligned}
(A \otimes A')(e_2 \otimes e_1) &= ba'e_1 \otimes e_1 + bc'e_1 \otimes e_2 + da'e_2 \otimes e_1 + dc'e_2 \otimes e_2, \\
(A \otimes A')(e_2 \otimes e_2) &= bb'e_1 \otimes e_1 + bd'e_1 \otimes e_2 + db'e_2 \otimes e_1 + dd'e_2 \otimes e_2.
\end{aligned}$$

Therefore the matrix for $A \otimes A'$ is

$$\begin{pmatrix} aa' & ab' & ba' & bb' \\ ac' & ad' & bc' & bd' \\ ca' & cb' & da' & db' \\ cc' & cd' & dc' & dd' \end{pmatrix} = \left(\begin{array}{c|c} aA' & bA' \\ \hline cA' & dA' \end{array}\right).$$

So $\mathrm{Tr}(A \otimes A') = a(a' + d') + d(a' + d') = (a + d)(a' + d') = (\mathrm{Tr}\, A)(\mathrm{Tr}\, A')$, and $\det(A \otimes A')$ looks painful to compute from the matrix. We'll do this later, in Example 2.7, in an almost painless way.

If, more generally, $A \in \mathrm{M}_n(R)$ and $A' \in \mathrm{M}_{n'}(R)$ then the matrix for $A \otimes A'$ with respect to the standard basis for $R^n \otimes_R R^{n'}$ is the block matrix $(a_{ij}A')$ where $A = (a_{ij})$. This $nn' \times nn'$ matrix is called the *Kronecker product* of A and A', and is not symmetric in the roles of A and A' in general (just as $A \otimes A' \neq A' \otimes A$ in general). In particular, $I_n \otimes A'$ has block matrix representation $(\delta_{ij}A')$, whose determinant is $(\det A')^n$.

The construction of tensor products (Kronecker products) of matrices has the following application to finding polynomials with particular roots.

Theorem 2.3. *Let K be a field and suppose $A \in \mathrm{M}_m(K)$ and $B \in \mathrm{M}_n(K)$ have eigenvalues λ and μ in K. Then $A \otimes I_n + I_m \otimes B$ has eigenvalue $\lambda + \mu$ and $A \otimes B$ has eigenvalue $\lambda\mu$.*

Proof. We have $Av = \lambda v$ and $Bw = \mu w$ for some $v \in K^m$ and $w \in K^n$. Then

$$\begin{aligned}
(A \otimes I_n + I_m \otimes B)(v \otimes w) &= Av \otimes w + v \otimes Bw \\
&= \lambda v \otimes w + v \otimes \mu w \\
&= (\lambda + \mu)(v \otimes w)
\end{aligned}$$

and

$$(A \otimes B)(v \otimes w) = Av \otimes Bw = \lambda v \otimes \mu w = \lambda\mu(v \otimes w),$$

□

Example 2.4. The numbers $\sqrt{2}$ and $\sqrt{3}$ are eigenvalues of $A = \left(\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & 3 \\ 1 & 0 \end{smallmatrix}\right)$. A matrix with eigenvalue $\sqrt{2} + \sqrt{3}$ is

$$\begin{aligned} A \otimes I_2 + I_2 \otimes B &= \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 3 & 2 & 0 \\ 1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 3 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \end{aligned}$$

whose characteristic polynomial is $T^4 - 10T^2 + 1$. So this is a polynomial with $\sqrt{2} + \sqrt{3}$ as a root.

Although we stressed that $\varphi \otimes \psi$ is not an elementary tensor, but rather is the notation for a linear map, φ and ψ belong to the R-modules $\operatorname{Hom}_R(M, M')$ and $\operatorname{Hom}_R(N, N')$, so one could ask if the actual elementary tensor $\varphi \otimes \psi$ in $\operatorname{Hom}_R(M, M') \otimes_R \operatorname{Hom}_R(N, N')$ is related to the linear map $\varphi \otimes \psi \colon M \otimes_R N \to M' \otimes_R N'$.

Theorem 2.5. *There is a linear map*

$$\operatorname{Hom}_R(M, M') \otimes_R \operatorname{Hom}_R(N, N') \to \operatorname{Hom}_R(M \otimes_R N, M' \otimes_R N')$$

that sends the elementary tensor $\varphi \otimes \psi$ to the linear map $\varphi \otimes \psi$. When $M, M', N,$ and N' are finite free, this is an isomorphism.

Proof. We adopt the temporary notation $T(\varphi, \psi)$ for the linear map we have previously written as $\varphi \otimes \psi$, so we can use $\varphi \otimes \psi$ to mean an elementary tensor in the tensor product of Hom-modules. So $T(\varphi, \psi) \colon M \otimes_R N \to M' \otimes_R N'$ is the linear map sending every $m \otimes n$ to $\varphi(m) \otimes \psi(n)$.

Define $\operatorname{Hom}_R(M, M') \times \operatorname{Hom}_R(N, N') \to \operatorname{Hom}_R(M \otimes_R N, M' \otimes_R N')$ by $(\varphi, \psi) \mapsto T(\varphi, \psi)$. This is R-bilinear. For example, to show $T(r\varphi, \psi) = rT(\varphi, \psi)$, both sides are linear maps so to prove they are equal it suffices to check they are equal at the elementary tensors in $M \otimes_R N$:

$$T(r\varphi, \psi)(m \otimes n) = (r\varphi)(m) \otimes \psi(n) = r\varphi(m) \otimes \psi(n) = r(\varphi(m) \otimes \psi(n)) = rT(\varphi, \psi)(m \otimes n).$$

The other bilinearity conditions are left to the reader.

From the universal mapping property of tensor products, there is a unique R-linear map $\operatorname{Hom}_R(M, M') \otimes_R \operatorname{Hom}_R(N, N') \to \operatorname{Hom}_R(M \otimes_R N, M' \otimes_R N')$ where $\varphi \otimes \psi \mapsto T(\varphi, \psi)$.

Suppose M, M', N, and N' are all finite free R-modules. Let them have respective bases $\{e_i\}$, $\{e'_{i'}\}$, $\{f_j\}$, and $\{f'_{j'}\}$. Then $\operatorname{Hom}_R(M, M')$ and $\operatorname{Hom}_R(N, N')$ are both free with bases $\{E_{i'i}\}$ and $\{\widetilde{E}_{j'j}\}$, where $E_{i'i} \colon M \to M'$ is the linear map sending e_i to $e'_{i'}$ and is 0 at other basis vectors of M, and $\widetilde{E}_{j'j} \colon N \to N'$ is defined similarly. (The matrix representation of $E_{i'i}$ with respect to the chosen bases of M and M' has a 1 in the (i', i) position and 0 elsewhere, thus justifying the notation.) A basis of $\operatorname{Hom}_R(M, M') \otimes_R \operatorname{Hom}_R(N, N')$ is

$\{E_{i'i} \otimes \widetilde{E}_{j'j}\}$ and $T(E_{i'i} \otimes \widetilde{E}_{j'j})\colon M \otimes_R N \to M' \otimes_R N'$ has the effect

$$\begin{aligned} T(E_{i'i} \otimes \widetilde{E}_{j'j})(e_\mu \otimes f_\nu) &= E_{i'i}(e_\mu) \otimes \widetilde{E}_{j'j}(f_\nu) \\ &= \delta_{\mu i} e'_{i'} \otimes \delta_{\nu j} f'_{j'} \\ &= \begin{cases} e'_{i'} \otimes f'_{j'}, & \text{if } \mu = i \text{ and } \nu = j, \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

so $T(E_{i'i} \otimes E_{j'j})$ sends $e_i \otimes f_j$ to $e'_{i'} \otimes f'_{j'}$ and sends other members of the basis of $M \otimes_R N$ to 0. That means the linear map $\operatorname{Hom}_R(M, M') \otimes_R \operatorname{Hom}_R(N, N') \to \operatorname{Hom}_R(M \otimes_R N, M' \otimes_R N')$ sends a basis to a basis, so it is an isomorphism when the modules are finite free. $\square$

The upshot of Theorem 2.5 is that $\operatorname{Hom}_R(M, M') \otimes_R \operatorname{Hom}_R(N, N')$ naturally acts as linear maps $M \otimes_R N \to M' \otimes_R N'$ and it turns the elementary tensor $\varphi \otimes \psi$ into the linear map we've been writing as $\varphi \otimes \psi$. This justifies our use of the notation $\varphi \otimes \psi$ for the linear map, but it should be kept in mind that we will continue to write $\varphi \otimes \psi$ for the linear map itself (on $M \otimes_R N$) and not for an elementary tensor in a tensor product of Hom-modules.

Properties of tensor products of modules carry over to properties of tensor products of linear maps, by checking equality on all tensors. For example, if $\varphi_1\colon M_1 \to N_1$, $\varphi_2\colon M_2 \to N_2$, and $\varphi_3\colon M_3 \to N_3$ are linear maps, we have $\varphi_1 \otimes (\varphi_2 \oplus \varphi_3) = (\varphi_1 \otimes \varphi_2) \oplus (\varphi_1 \otimes \varphi_3)$ and $(\varphi_1 \otimes \varphi_2) \otimes \varphi_3 = \varphi_1 \otimes (\varphi_2 \otimes \varphi_3)$, in the sense that the diagrams

$$\begin{tikzcd} M_1 \otimes_R (M_2 \oplus M_3) \arrow[r, "\varphi_1 \otimes (\varphi_2 \oplus \varphi_3)"] \arrow[d] & N_1 \otimes_R (N_2 \oplus N_3) \arrow[d] \\ (M_1 \otimes_R M_2) \oplus (M_1 \otimes_R M_3) \arrow[r, "(\varphi_1 \otimes \varphi_2) \oplus (\varphi_1 \otimes \varphi_3)"] & (N_1 \otimes_R N_2) \oplus (N_1 \otimes_R N_3) \end{tikzcd}$$

and

$$\begin{tikzcd} M_1 \otimes_R (M_2 \otimes_R M_3) \arrow[r, "\varphi_1 \otimes (\varphi_2 \otimes \varphi_3)"] \arrow[d] & N_1 \otimes_R (N_2 \otimes_R N_3) \arrow[d] \\ (M_1 \otimes_R M_2) \otimes_R M_3 \arrow[r, "(\varphi_1 \otimes \varphi_2) \otimes \varphi_3"] & (N_1 \otimes_R N_2) \otimes_R N_3 \end{tikzcd}$$

commute, with the vertical maps being the canonical isomorphisms.

The properties of the next theorem are called the *functoriality* of the tensor product of linear maps.

Theorem 2.6. *For R-modules M and N, $\mathrm{id}_M \otimes \mathrm{id}_N = \mathrm{id}_{M \otimes_R N}$. For linear maps $M \xrightarrow{\varphi} M'$, $M' \xrightarrow{\varphi'} M''$, $N \xrightarrow{\psi} N'$, and $N' \xrightarrow{\psi'} N''$,*

$$(\varphi' \otimes \psi') \circ (\varphi \otimes \psi) = (\varphi' \circ \varphi) \otimes (\psi' \circ \psi)$$

as linear maps from $M \otimes_R N$ to $M'' \otimes_R N''$.

Proof. The function $\mathrm{id}_M \otimes \mathrm{id}_N$ is a linear map from $M \otimes_R N$ to itself that fixes every elementary tensor, so it fixes all tensors.

Since $(\varphi' \otimes \psi') \circ (\varphi \otimes \psi)$ and $(\varphi' \circ \varphi) \otimes (\psi' \circ \psi)$ are linear maps, to prove their equality it suffices to check they have the same value at any elementary tensor $m \otimes n$, at which they both have the value $\varphi'(\varphi(m)) \otimes \psi'(\psi(n))$. $\square$

Example 2.7. The composition rule for tensor products of linear maps helps us compute determinants of tensor products of linear operators. Let M and N be *finite free* R-modules of respective ranks k and ℓ. For linear operators $M \xrightarrow{\varphi} M$ and $N \xrightarrow{\psi} N$, we will compute $\det(\varphi \otimes \psi)$ by breaking up $\varphi \otimes \psi$ into a composite of two maps $M \otimes_R N \to M \otimes_R N$:

$$\varphi \otimes \psi = (\varphi \otimes \mathrm{id}_N) \circ (\mathrm{id}_M \otimes \psi),$$

so the multiplicativity of the determinant implies $\det(\varphi \otimes \psi) = \det(\varphi \otimes \mathrm{id}_N)\det(\mathrm{id}_M \otimes \psi)$ and we are reduced to the case when one of the "factors" is an identity map. Moreover, the isomorphism $M \otimes_R N \to N \otimes_R M$ where $m \otimes n \mapsto n \otimes m$ converts $\varphi \otimes \mathrm{id}_N$ into $\mathrm{id}_N \otimes \varphi$, so $\det(\varphi \otimes \mathrm{id}_N) = \det(\mathrm{id}_N \otimes \varphi)$, so

$$\det(\varphi \otimes \psi) = \det(\mathrm{id}_N \otimes \varphi)\det(\mathrm{id}_M \otimes \psi).$$

What are the determinants on the right side? Pick bases $e_1, \ldots, e_k$ of M and $e'_1, \ldots, e'_\ell$ of N. We will use the $k\ell$ elementary tensors $e_i \otimes e'_j$ as a bases of $M \otimes_R N$. Let $[\varphi]$ be the matrix of φ in the ordered basis $e_1, \ldots, e_k$. Since $(\varphi \otimes \mathrm{id}_N)(e_i \otimes e'_j) = \varphi(e_i) \otimes e'_j$, let's order the basis of $M \otimes_R N$ as

$$e_1 \otimes e'_1, \ldots, e_k \otimes e'_1, \ldots, e_1 \otimes e'_\ell, \ldots, e_k \otimes e'_\ell.$$

The $k\ell \times k\ell$ matrix for $\varphi \otimes \mathrm{id}_N$ in this ordered basis is the block diagonal matrix

$$\begin{pmatrix} [\varphi] & O & \cdots & O \\ O & [\varphi] & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & [\varphi] \end{pmatrix},$$

whose determinant is $(\det \varphi)^\ell$.

Thus

$$\det(\varphi \otimes \psi) = (\det \varphi)^\ell (\det \psi)^k. \tag{2.1}$$

Note ℓ is the rank of the module on which ψ is defined and k is the rank of the module on which φ is defined. In particular, in Example 2.2 we have $\det(A \otimes A') = (\det A)^2(\det A')^2$.

Let's review the idea in this proof. Since $N \cong R^\ell$, $M \otimes_R N \cong M \otimes_R R^\ell \cong M^{\oplus \ell}$. Under such an isomorphism, $\varphi \otimes \mathrm{id}_N$ becomes the ℓ-fold direct sum $\varphi \oplus \cdots \oplus \varphi$, which has a block diagonal matrix representation in a suitable basis. So its determinant is $(\det \varphi)^\ell$.

Example 2.8. Taking $M = N$ and $\varphi = \psi$, the tensor square $\varphi^{\otimes 2}$ has determinant $(\det \varphi)^{2k}$.

Corollary 2.9. *Let M be a free module of rank $k \geq 1$ and $\varphi \colon M \to M$ be a linear map. For every $i \geq 1$, $\det(\varphi^{\otimes i}) = (\det \varphi)^{ik^{i-1}}$.*

Proof. Use induction and associativity of the tensor product of linear maps. □

Remark 2.10. Equation (2.1) in the setting of vector spaces and matrices says $\det(A \otimes B) = (\det A)^\ell (\det B)^k$, where A is $k \times k$, B is $\ell \times \ell$, and $A \otimes B = (a_{ij}B)$ is the matrix incarnation of a tensor product of linear maps, called the Kronecker product of A and B at the end of Example 2.2. While the label "Kronecker product" for the matrix $A \otimes B$ is completely standard, it is not historically accurate. It is based on Hensel's attribution of the formula for $\det(A \otimes B)$ to Kronecker, but the formula is due to Zehfuss. See [2].

Let's see how the tensor product of linear maps behaves for isomorphisms, surjections, and injections.

Theorem 2.11. *If $\varphi\colon M \to M'$ and $\psi\colon N \to N'$ are isomorphisms then $\varphi \otimes \psi$ is an isomorphism.*

Proof. The composite of $\varphi \otimes \psi$ with $\varphi^{-1} \otimes \psi^{-1}$ in both orders is the identity. $\square$

Theorem 2.12. *If $\varphi\colon M \to M'$ and $\psi\colon N \to N'$ are surjective then $\varphi \otimes \psi$ is surjective.*

Proof. Since $\varphi \otimes \psi$ is linear, to show it is onto it suffices to show every elementary tensor in $M' \otimes_R N'$ is in the image. For such an elementary tensor $m' \otimes n'$, we can write $m' = \varphi(m)$ and $n' = \psi(n)$ since φ and ψ are onto. Therefore $m' \otimes n' = \varphi(m) \otimes \psi(n) = (\varphi \otimes \psi)(m \otimes n)$. $\square$

It is a fundamental feature of tensor products that if φ and ψ are both injective then $\varphi \otimes \psi$ *might not* be injective. This can occur even if one of φ or ψ is the identity function.

Example 2.13. Taking $R = \mathbf{Z}$, let $\alpha\colon \mathbf{Z}/p\mathbf{Z} \to \mathbf{Z}/p^2\mathbf{Z}$ be multiplication by p: $\alpha(x) = px$. This is injective, and if we tensor with $\mathbf{Z}/p\mathbf{Z}$ we get the linear map $1 \otimes \alpha\colon \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/p\mathbf{Z} \to \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/p^2\mathbf{Z}$ with the effect $a \otimes x \mapsto a \otimes px = pa \otimes x = 0$, so $1 \otimes \alpha$ is identically 0 and its domain is $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/p\mathbf{Z} \cong \mathbf{Z}/p\mathbf{Z} \neq 0$, so $1 \otimes \alpha$ is not injective.

This provides an example where the natural linear map

$$\mathrm{Hom}_R(M, M') \otimes_R \mathrm{Hom}_R(N, N') \to \mathrm{Hom}_R(M \otimes_R N, M' \otimes_R N')$$

in Theorem 2.5 is not an isomorphism; $R = \mathbf{Z}$, $M = M' = N = \mathbf{Z}/p\mathbf{Z}$, and $N' = \mathbf{Z}/p^2\mathbf{Z}$.

Because the tensor product of linear maps does not generally preserve injectivity, a tensor has to be understood in *context*: it is a tensor in a specific tensor product module $M \otimes_R N$. If $M \subset M'$ and $N \subset N'$, it is generally false that $M \otimes_R N$ can be thought of as a submodule of $M' \otimes_R N'$ since the natural map $M \otimes_R N \to M' \otimes_R N'$ may not be injective. We can say it this way: a tensor product of submodules need not be a submodule.

Example 2.14. Since $p\mathbf{Z} \cong \mathbf{Z}$ as abelian groups, by $pn \mapsto n$, we have $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} p\mathbf{Z} \cong \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z} \cong \mathbf{Z}/p\mathbf{Z}$ as abelian groups by $a \otimes pn \mapsto a \otimes n \mapsto na \bmod p$. Therefore $1 \otimes p$ in $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} p\mathbf{Z}$ is nonzero, since the isomorphism identifies it with 1 in $\mathbf{Z}/p\mathbf{Z}$. However, $1 \otimes p$ in $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}$ is 0, since $1 \otimes p = p \otimes 1 = 0 \otimes 1 = 0$. (This calculation with $1 \otimes p$ doesn't work in $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} p\mathbf{Z}$ since we can't bring p to the left side of $\otimes$ and leave 1 behind, as $1 \notin p\mathbf{Z}$.)

It might seem weird that $1 \otimes p$ is nonzero in $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} p\mathbf{Z}$ while $1 \otimes p$ is zero in the "larger" abelian group $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}$! The reason there isn't a contradiction is that $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} p\mathbf{Z}$ is not really a subgroup of $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}$ even though $p\mathbf{Z}$ is a subgroup of $\mathbf{Z}$. The inclusion mapping $i\colon p\mathbf{Z} \to \mathbf{Z}$ gives us a natural mapping $1 \otimes i\colon \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} p\mathbf{Z} \to \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}$, with the effect $a \otimes pn \mapsto a \otimes pn$, but this is not an embedding. In fact its image is 0: in $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}$, $a \otimes pn = pa \otimes n = 0 \otimes n = 0$. The moral is that an elementary tensor $a \otimes pn$ means something different in $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} p\mathbf{Z}$ and in $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}$.

This example also shows the image of $M \otimes_R N \xrightarrow{\varphi \otimes \psi} M' \otimes_R N'$ need not be isomorphic to $\varphi(M) \otimes_R \psi(N)$, since $1 \otimes i$ has image 0 and $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} i(p\mathbf{Z}) \cong \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z} \cong \mathbf{Z}/p\mathbf{Z}$.

Example 2.15. While $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z} \cong \mathbf{Z}/p\mathbf{Z}$, if we enlarge the second tensor factor $\mathbf{Z}$ to $\mathbf{Q}$ we get a huge collapse: $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Q} = 0$ since $a \otimes r = a \otimes p(r/p) = pa \otimes r/p = 0 \otimes r/p = 0$. In particular, $1 \otimes 1$ is nonzero in $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}$ but $1 \otimes 1 = 0$ in $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Q}$.

In terms of tensor products of linear mappings, this example says that tensoring the inclusion $i\colon \mathbf{Z} \hookrightarrow \mathbf{Q}$ with $\mathbf{Z}/p\mathbf{Z}$ gives us a $\mathbf{Z}$-linear map $1 \otimes i\colon \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z} \to \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Q}$ that is not injective: the domain is isomorphic to $\mathbf{Z}/p\mathbf{Z}$ and the target is 0.

Darij Grinberg[1] gave a nice clarification of this confusing situation: it is due to an abuse of notation when writing elementary tensors as $m \otimes n$ without indicating the ambient modules that m and n lie in. Including those modules in the notation, say by writing $m \otimes n$ as $(m, M) \otimes_R (n, N)$, would avoid the confusion at the cost of very clumsy notation.

Example 2.16. Here is an example of a linear map $f\colon M \to N$ that is injective and its tensor square $f^{\otimes 2}\colon M^{\otimes 2} \to N^{\otimes 2}$ is not injective.

Let $R = A[X, Y]$ with A a nonzero commutative ring and $I = (X, Y)$. In $R^{\otimes 2}$, we have

$$X \otimes Y = XY(1 \otimes 1) \text{ and } Y \otimes X = YX(1 \otimes 1) = XY(1 \otimes 1), \tag{2.2}$$

so $X \otimes Y = Y \otimes X$. We will now show that in $I^{\otimes 2}$, $X \otimes Y \neq Y \otimes X$. (The calculation in (2.2) makes no sense in $I^{\otimes 2}$ since 1 is not an element of I.) To show two tensors are not equal, the best approach is to construct a linear map from the tensor product space that has different values at the two tensors. The function $I \times I \to A$ given by $(f, g) \mapsto f_X(0, 0)g_Y(0, 0)$, where f_X and g_Y are partial derivatives of f and g with respect to X and Y, is R-bilinear. (Treat the target A as an R-module through multiplication by the constant term of polynomials in R, or just view A as R/I with ordinary multiplication by R.) Thus there is an R-linear map $I^{\otimes 2} \to A$ sending any elementary tensor $f \otimes g$ to $f_X(0, 0)g_Y(0, 0)$. In particular, $X \otimes Y \mapsto 1$ and $Y \otimes X \mapsto 0$, so $X \otimes Y \neq Y \otimes X$ in $I^{\otimes 2}$.

It might seem weird that $X \otimes Y$ and $Y \otimes X$ are equal in $R^{\otimes 2}$ but are not equal in $I^{\otimes 2}$, even though X and Y are elements of I and $I \subset R$. We must be careful when we think about a tensor $t \in I \otimes_R I$ as a tensor in $R \otimes_R R$. Letting $i\colon I \hookrightarrow R$ be the inclusion map, thinking about t in $R^{\otimes 2}$ means looking at $i^{\otimes 2}(t)$, where $i^{\otimes 2}\colon I^{\otimes 2} \to R^{\otimes 2}$. For the tensor $t = X \otimes Y - Y \otimes X$ in $I^{\otimes 2}$, we computed above that $t \neq 0$ but $i^{\otimes 2}(t) = 0$, so $i^{\otimes 2}$ is not injective even though i is injective. In other words, the natural way to think of $I \otimes_R I$ "inside" $R \otimes_R R$ is actually not an embedding. For polynomials f and g in I, you have to distinguish between the tensor $f \otimes g$ in $I \otimes_R I$ and the tensor $f \otimes g$ in $R \otimes_R R$.

Generalizing this, let $R = A[X_1, \dots, X_n]$ where $n \geq 2$ and $I = (X_1, \dots, X_n)$. The inclusion $i\colon I \hookrightarrow R$ is injective but the nth tensor power (as R-modules) $i^{\otimes n}\colon I^{\otimes n} \to R^{\otimes n}$ is not injective because the tensor

$$t := \sum_{\sigma \in S_n} (\operatorname{sign} \sigma) X_{\sigma(1)} \otimes \cdots \otimes X_{\sigma(n)} \in I^{\otimes n}$$

gets sent to $\sum_{\sigma \in S_n} (\operatorname{sign} \sigma) X_1 \cdots X_n (1 \otimes \cdots \otimes 1)$ in $R^{\otimes n}$, which is 0, but t is not 0 in $I^{\otimes n}$ because there is an R-linear map $I^{\otimes n} \to A$ sending t to 1: use a product of partial derivatives at $(0, 0, \dots, 0)$, as in the $n = 2$ case.

Remark 2.17. The ideal $I = (X, Y)$ in $R = A[X, Y]$ from Example 2.16 has another interesting feature when A is a domain: it is a torsion-free R-module but $I^{\otimes 2}$ is not: $X(X \otimes Y) = X \otimes XY = Y(X \otimes X)$ and $X(Y \otimes X) = XY \otimes X = Y(X \otimes X)$, so in $I^{\otimes 2}$ we have $X(X \otimes Y - Y \otimes X) = 0$, but $X \otimes Y - Y \otimes X \neq 0$. Similarly, $Y(X \otimes Y - Y \otimes X) = 0$. Therefore a tensor product of torsion-free modules (even over a domain) *need not* be torsion-free.

While we have just seen a tensor power of an injective linear map need not be injective, here is a condition where injectivity holds.

Theorem 2.18. *Let $\varphi\colon M \to N$ be injective and $\varphi(M)$ be a direct summand of N. For $k \geq 0$, $\varphi^{\otimes k}\colon M^{\otimes k} \to N^{\otimes k}$ is injective and the image is a direct summand of $N^{\otimes k}$.*

[1] See https://math.stackexchange.com/questions/3452281.

Proof. Write $N = \varphi(M) \oplus P$. Let $\pi \colon N \twoheadrightarrow M$ by $\pi(\varphi(m) + p) = m$, so π is linear and $\pi \circ \varphi = \mathrm{id}_M$. Then $\varphi^{\otimes k} \colon M^{\otimes k} \to N^{\otimes k}$ and $\pi^{\otimes k} \colon N^{\otimes k} \twoheadrightarrow M^{\otimes k}$ are linear maps and

$$\pi^{\otimes k} \circ \varphi^{\otimes k} = (\pi \circ \varphi)^{\otimes k} = \mathrm{id}_M^{\otimes k} = \mathrm{id}_{M^{\otimes k}},$$

so $\varphi^{\otimes k}$ has a left inverse. That implies $\varphi^{\otimes k}$ is injective and $M^{\otimes k}$ is isomorphic to a direct summand of $N^{\otimes k}$ by criteria for when a short exact sequence of modules splits. $\square$

We can apply this to vector spaces: if V is a vector space and W is a subspace, there is a direct sum decomposition $V = W \oplus U$ (U is non-canonical), so tensor powers of the inclusion $W \to V$ are injective linear maps $W^{\otimes k} \to V^{\otimes k}$.

Other criteria for a tensor power of an injective linear map to be injective will be met in Corollary 3.13 and Theorem 4.9.

We will now compute the kernel of $M \otimes_R N \xrightarrow{\varphi \otimes \psi} M' \otimes_R N'$ in terms of the kernels of φ and ψ, assuming φ and ψ are *onto*.

Theorem 2.19. *Let $M \xrightarrow{\varphi} M'$ and $N \xrightarrow{\psi} N'$ be R-linear and surjective. The kernel of $M \otimes_R N \xrightarrow{\varphi \otimes \psi} M' \otimes_R N'$ is the submodule of $M \otimes_R N$ spanned by all $m \otimes n$ where $\varphi(m) = 0$ or $\psi(n) = 0$. That is, intuitively*

$$\ker(\varphi \otimes \psi) = (\ker \varphi) \otimes_R N + M \otimes_R (\ker \psi),$$

while rigorously in terms of the inclusion maps $\ker \varphi \xrightarrow{i} M$ and $\ker \psi \xrightarrow{j} N$,

$$\ker(\varphi \otimes \psi) = (i \otimes 1)((\ker \varphi) \otimes_R N) + (1 \otimes j)(M \otimes_R (\ker \psi)).$$

The reason $(\ker \varphi) \otimes_R N + M \otimes_R (\ker \psi)$ is only an *intuitive* formula for the kernel of $\varphi \otimes \psi$ is that, strictly speaking, these tensor product modules are not submodules of $M \otimes_R N$. Only after applying $i \otimes 1$ and $1 \otimes j$ to them – and these might not be injective – do those modules become submodules of $M \otimes_R N$.

Proof. Both $(i \otimes 1)((\ker \varphi) \otimes_R N)$ and $(1 \otimes j)(M \otimes_R (\ker \psi))$ are killed by $\varphi \otimes \psi$: if $m \in \ker \varphi$ and $n \in N$ then $(\varphi \otimes \psi)((i \otimes 1)(m \otimes n)) = (\varphi \otimes \psi)(m \otimes n) = \varphi(m) \otimes \psi(n) = 0$ since[2] $\varphi(m) = 0$. Similarly $(1 \otimes j)(m \otimes n)$ is killed by $\varphi \otimes \psi$ if $m \in M$ and $n \in \ker \psi$. Set

$$U = (i \otimes 1)((\ker \varphi) \otimes_R N) + (1 \otimes j)(M \otimes_R (\ker \psi)),$$

so $U \subset \ker(\varphi \otimes \psi)$, which means $\varphi \otimes \psi$ induces a linear map

$$\Phi \colon (M \otimes_R N)/U \to M' \otimes_R N'$$

where $\Phi(m \otimes n \bmod U) = (\varphi \otimes \psi)(m \otimes n) = \varphi(m) \otimes \psi(n)$. We will now write down an inverse map, which proves Φ is injective, so the kernel of $\varphi \otimes \psi$ is U.

Because φ and ψ are assumed to be onto, every elementary tensor in $M' \otimes_R N'$ has the form $\varphi(m) \otimes \psi(n)$. Knowing $\varphi(m)$ and $\psi(n)$ only determines m and n up to addition by elements of $\ker \varphi$ and $\ker \psi$. For $m' \in \ker \varphi$ and $n' \in \ker \psi$,

$$(m + m') \otimes (n + n') = m \otimes n + m' \otimes n + m \otimes n' + m' \otimes n' \in m \otimes n + U,$$

so the function $M' \times N' \to (M \otimes_R N)/U$ defined by $(\varphi(m), \psi(n)) \mapsto m \otimes n \bmod U$ is well-defined. It is R-bilinear, so we have an R-linear map $\Psi \colon M' \otimes_R N' \to (M \otimes_R N)/U$ where $\Psi(\varphi(m) \otimes \psi(n)) = m \otimes n \bmod U$ on elementary tensors.

Easily the linear maps $\Phi \circ \Psi$ and $\Psi \circ \Phi$ fix spanning sets, so they are both the identity. $\square$

[2]The first $m \otimes n$ is in $(\ker \varphi) \otimes_R N$, while the second $m \otimes n$ is in $M \otimes_R N$.

Remark 2.20. If we remove the assumption that φ and ψ are onto, Theorem 2.19 does not correctly compute the kernel. For example, if φ and ψ are both injective then the formula for the kernel in Theorem 2.19 is 0, and we know $\varphi \otimes \psi$ need not be injective.

Example 2.21. When I is an ideal in R and M and N are R-modules, let $\varphi : M \to M/IM$ and $\psi : N \to N/IN$ be the natural maps. Theorem 2.19 tells us the map $\varphi\otimes\psi : M \otimes_R N \to (M/IM) \otimes_R (N/IN)$ has kernel spanned as an R-module by all $m \otimes n$ where $m \in IM$ or $n \in IN$. The span of such elementary tensors is $I(M \otimes_R N)$ (why?), so

$$(M \otimes_R N)/I(M \otimes_R N) \cong (M/IM) \otimes_R (N/IN).$$

Unlike the kernel computation in Theorem 2.19, it is not easy to describe the torsion submodule of a tensor product in terms of the torsion submodules of the original modules. While $(M\otimes_R N)_{\rm tor}$ contains $(i\otimes 1)(M_{\rm tor}\otimes_R N)+(1\otimes j)(M\otimes_R N_{\rm tor})$, with $i\colon M_{\rm tor} \to M$ and $j\colon N_{\rm tor} \to N$ being the inclusions, it is not true that this is all of $(M\otimes_R N)_{\rm tor}$, since $M\otimes_R N$ can have nonzero torsion when M and N are torsion-free (so $M_{\rm tor} = 0$ and $N_{\rm tor} = 0$). We saw this at the end of Example 2.16.

Corollary 2.22. *If $M \xrightarrow{\varphi} M'$ is an isomorphism of R-modules and $N \xrightarrow{\psi} N'$ is surjective, then the linear map $M\otimes_R N \xrightarrow{\varphi\otimes\psi} M'\otimes_R N'$ has kernel $(1\otimes j)(M\otimes_R (\ker\psi))$, where $\ker\psi \xrightarrow{j} N$ is the inclusion.*

Proof. This is immediate from Theorem 2.19 since $\ker\varphi = 0$. □

Corollary 2.23. *Let $f\colon R \to S$ be a homomorphism of commutative rings and $M \subset N$ as R-modules, with $M \xrightarrow{i} N$ the inclusion map. The following are equivalent:*

(1) $S \otimes_R M \xrightarrow{1\otimes i} S \otimes_R N$ *is onto.*
(2) $S \otimes_R (N/M) = 0$.

Proof. Let $N \xrightarrow{\pi} N/M$ be the reduction map, so we have the sequence $S \otimes_R M \xrightarrow{1\otimes i} S \otimes_R N \xrightarrow{1\otimes\pi} S \otimes_R (N/M)$. The map $1 \otimes \pi$ is onto, and $\ker\pi = M$, so $\ker(1 \otimes \pi) = (1 \otimes i)(S \otimes_R M)$. Therefore $1 \otimes i$ is onto if and only if $\ker(1 \otimes \pi) = S \otimes_R N$ if and only if $1 \otimes \pi = 0$, and since $1 \otimes \pi$ is onto we have $1 \otimes \pi = 0$ if and only if $S \otimes_R (N/M) = 0$. □

Example 2.24. If $M \subset N$ and N is finitely generated, we show $M = N$ if and only if the natural map $R/\mathfrak{m} \otimes_R M \xrightarrow{1\otimes i} R/\mathfrak{m} \otimes_R N$ is onto for all maximal ideals $\mathfrak{m}$ in R, where $M \xrightarrow{i} N$ is the inclusion map. The "only if" direction is clear. In the other direction, if $R/\mathfrak{m}\otimes_R M \xrightarrow{1\otimes i} R/\mathfrak{m}\otimes_R N$ is onto then $R/\mathfrak{m}\otimes_R (N/M) = 0$ by Corollary 2.23. Since N is finitely generated, so is N/M, and we are reduced to showing $R/\mathfrak{m}\otimes_R (N/M) = 0$ for all maximal ideals $\mathfrak{m}$ if and only if $N/M = 0$. When P is a finitely generated module, $P = 0$ if and only if $P/\mathfrak{m}P = 0$ for all maximal ideals[3] $\mathfrak{m}$ in R, so we can apply this to $P = N/M$ since $P/\mathfrak{m}P \cong R/\mathfrak{m}\otimes_R P$.

Corollary 2.25. *Let $f\colon R \to S$ be a homomorphism of commutative rings and I be an ideal in $R[X_1,\dots,X_n]$. Write $I \cdot S[X_1,\dots,X_n]$ for the ideal generated by the image of I in $S[X_1,\dots,X_n]$. Then*

$$S \otimes_R R[X_1,\dots,X_n]/I \cong S[X_1,\dots,X_n]/(I \cdot S[X_1,\dots,X_n]).$$

[3] $P/\mathfrak{m}P = 0 \Rightarrow P = \mathfrak{m}P \Rightarrow P_\mathfrak{m} = \mathfrak{m}P_\mathfrak{m} \Rightarrow P_\mathfrak{m} = 0$ by Nakayama's lemma. From $P_\mathfrak{m} = 0$ for all maximal ideals $\mathfrak{m}$, $P = 0$: for all $x \in P$, $x = 0$ in $P_\mathfrak{m}$ implies $ax = 0$ in P for some $a \in R - \mathfrak{m}$. Thus $\mathrm{Ann}_R(x)$ is not in any maximal ideal of R, so $\mathrm{Ann}_R(x) = R$ and thus $x = 1 \cdot x = 0$.

as S-modules by $s \otimes h \bmod I \mapsto sh \bmod I \cdot S[X_1, \dots, X_n]$.

Proof. The identity $S \to S$ and the natural reduction $R[X_1, \dots, X_n] \twoheadrightarrow R[X_1, \dots, X_n]/I$ are both onto, so the tensor product of these R-linear maps is an R-linear surjection

$$S \otimes_R R[X_1, \dots, X_n] \twoheadrightarrow S \otimes_R (R[X_1, \dots, X_n]/I) \tag{2.3}$$

and the kernel is $(1 \otimes j)(S \otimes_R I)$ by Theorem 2.19, where $j \colon I \to R[X_1, \dots, X_n]$ is the inclusion. Under the natural R-module isomorphism

$$S \otimes_R R[X_1, \dots, X_n] \cong S[X_1, \dots, X_n], \tag{2.4}$$

$(1 \otimes j)(S \otimes_R I)$ on the left side corresponds to $I \cdot S[X_1, \dots, X_n]$ on the right side, so (2.3) and (2.4) say

$$S[X_1, \dots, X_n]/(I \cdot S[X_1, \dots, X_n]) \cong S \otimes_R (R[X_1, \dots, X_n]/I).$$

as R-modules. The left side is naturally an S-module and the right side is too using extension of scalars. It is left to the reader to check the isomorphism is S-linear. $\square$

Example 2.26. For $h(X) \in \mathbf{Z}[X]$, $\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Z}[X]/(h(X)) \cong \mathbf{Q}[X]/(h(X))$ as $\mathbf{Q}$-vector spaces and $\mathbf{Z}/m\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}[X]/(h(X)) = (\mathbf{Z}/m\mathbf{Z})[X]/(\overline{h(X)})$ as $\mathbf{Z}/m\mathbf{Z}$-modules where $m > 1$.

3. Flat Modules

Because a tensor product of injective linear maps might not be injective, it is important to give a name to those R-modules N that always preserve injectivity, in the sense that $M \xrightarrow{\varphi} M'$ being injective implies $N \otimes_R M \xrightarrow{1 \otimes \varphi} N \otimes_R M'$ is injective. (Notice the map on N is the identity.)

Definition 3.1. An R-module N is called *flat* if for all injective linear maps $M \xrightarrow{\varphi} M'$ the linear map $N \otimes_R M \xrightarrow{1 \otimes \varphi} N \otimes_R M'$ is injective.

The concept of a flat module is pointless unless one has some good examples. The next two theorems provide some.

Theorem 3.2. *Any free R-module F is flat: if the linear map $\varphi \colon M \to M'$ is injective, then $1 \otimes \varphi \colon F \otimes_R M \to F \otimes_R M'$ is injective.*

Proof. When $F = 0$ it is clear, so take $F \neq 0$ with basis $\{e_i\}_{i \in I}$. From our previous development of the tensor product, every element of $F \otimes_R M$ can be written as $\sum_i e_i \otimes m_i$ for a unique choice of $m_i \in M$, and similarly for $F \otimes_R M'$.

For $t \in \ker(1 \otimes \varphi)$, we can write $t = \sum_i e_i \otimes m_i$ with $m_i \in M$. Then

$$0 = (1 \otimes \varphi)(t) = \sum_i e_i \otimes \varphi(m_i),$$

in $F \otimes_R M'$, which forces each $\varphi(m_i)$ to be 0. So every m_i is 0, since φ is injective, and we get $t = \sum_i e_i \otimes 0 = 0$. $\square$

Note that in Theorem 3.2 we did not need to assume F has a finite basis.

Theorem 3.3. *Let R be a domain and K be its fraction field. As an R-module, K is flat.*

This is not a special case of the previous theorem: if K were a free R-module then[4] $K = R$, so whenever R is a domain that is not a field (*e.g.*, $R = \mathbf{Z}$) the fraction field of R is a flat R-module that is not a free R-module.

Proof. Let $M \xrightarrow{\varphi} M'$ be an injective linear map of R-modules. Every tensor in $K \otimes_R M$ is elementary (use common denominators in K) and an elementary tensor in $K \otimes_R M$ is 0 if and only if its first factor is 0 or its second factor is torsion. (Here we are using properties of $K \otimes_R M$ proved in part I.)

Supposing $(1 \otimes \varphi)(t) = 0$, we may write $t = x \otimes m$, so $0 = (1 \otimes \varphi)(t) = x \otimes \varphi(m)$. Therefore $x = 0$ in K or $\varphi(m) \in M'_{\rm tor}$. If $\varphi(m) \in M'_{\rm tor}$ then $r\varphi(m) = 0$ for some nonzero $r \in R$, so $\varphi(rm) = 0$, so $rm = 0$ in M (φ is injective), which means $m \in M_{\rm tor}$. Thus $x = 0$ or $m \in M_{\rm tor}$, so $t = x \otimes m = 0$. □

If M is a submodule of the R-module M' then Theorem 3.3 says we can consider $K \otimes_R M$ as a subspace of $K \otimes_R M'$ since the natural map $K \otimes_R M \to K \otimes_R M'$ is injective. (See diagram below.) Notice this works even if M or M' has torsion; although the natural maps $M \to K \otimes_R M$ and $M' \to_R K \otimes_R M'$ might not be injective, the map $K \otimes_R M \to K \otimes_R M'$ is injective.

$$\begin{tikzcd} M \arrow[r, hook, "\varphi"] \arrow[d] & M' \arrow[d] \\ K \otimes_R M \arrow[r, hook, "1\otimes\varphi"'] & K \otimes_R M' \end{tikzcd}$$

Example 3.4. The natural inclusion $\mathbf{Z} \hookrightarrow \mathbf{Z}/3\mathbf{Z} \oplus \mathbf{Z}$ is $\mathbf{Z}$-linear and injective. Applying $\mathbf{Q}\otimes_{\mathbf{Z}}$ to both sides and using properties of tensor products turns this into the identity map $\mathbf{Q} \to \mathbf{Q}$, which is also injective.

Remark 3.5. Theorem 3.3 generalizes: for any commutative ring R and multiplicative set D in R, the localization R_D is a flat R-module.

Theorem 3.6. *If M is a flat R-module and I is an ideal in R then $I \otimes_R M \cong IM$ by $i \otimes m \mapsto im$.*

Proof. The inclusion $I \to R$ is injective. Applying $\otimes_R M$ to this makes an *injective* R-linear map $I \otimes_R M \to R \otimes_R M$ since M is flat, and composing with the isomorphism $R \otimes_R M \cong M$ where $r \otimes m \mapsto rm$ makes the injective map $I \otimes_R M \to M$ where $i \otimes m \mapsto im$. The image is IM, so $I \otimes_R M \cong IM$ as R-modules with the desired effect on elementary tensors. □

To say an R-module N is *not* flat means there is some example of an injective linear map $M \xrightarrow{\varphi} M'$ whose induced linear map $N \otimes_R M \xrightarrow{1\otimes\varphi} N \otimes_R M'$ is not injective.

Example 3.7. For a nonzero torsion abelian group A, the natural map $\mathbf{Z} \hookrightarrow \mathbf{Q}$ is injective but if we apply $A\otimes_{\mathbf{Z}}$ we get the map $A \to 0$, which is not injective, so A is not a flat $\mathbf{Z}$-module. This includes nonzero finite abelian groups and the infinite abelian group $\mathbf{Q}/\mathbf{Z}$.

[4]Any two nonzero elements of K are R-linearly dependent, so if K were a free R-module then it would have a basis of size 1: $K = Rx$ for some $x \in K$. Therefore $x^2 = rx$ for some $r \in R$, so $x = r \in R$, which implies $K \subset R$, so $K = R$.

Remark 3.8. Since $\mathbf{Q}/\mathbf{Z}$ is not flat as a $\mathbf{Z}$-module, for a homomorphism of abelian groups $G \xrightarrow{f} G'$ the kernel of $\mathbf{Q}/\mathbf{Z} \otimes_{\mathbf{Z}} G \xrightarrow{1 \otimes f} \mathbf{Q}/\mathbf{Z} \otimes_{\mathbf{Z}} G'$ need not be $\mathbf{Q}/\mathbf{Z} \otimes \ker f$ but could be *larger*. Therefore it is not easy to determine the kernel of a group homomorphism after base extension by $\mathbf{Q}/\mathbf{Z}$. Failure to take this into account created a gap in a proof of a widely used theorem in knot theory. See [1, p. 927].

Example 3.9. If R is a domain with fraction field K, any nonzero torsion R-module T (meaning every element of T is killed by a nonzero element of R) is not a flat R-module since tensoring the inclusion $R \hookrightarrow K$ with T produces the R-linear map $T \to 0$, which is not injective. In particular, the quotient module K/R is not a flat R-module. The previous example is the special case $R = \mathbf{Z}$: $\mathbf{Q}/\mathbf{Z}$ is not a flat $\mathbf{Z}$-module.

Theorem 3.10. *If N is a flat R-module and $M \xrightarrow{\varphi} M'$ is R-linear then the kernel of $N \otimes_R M \xrightarrow{1\otimes\varphi} N \otimes_R M'$ is $N \otimes_R \ker\varphi$, viewed as a submodule of $N \otimes_R M$ in a natural way.*

Proof. The diagram

$$\begin{tikzcd} & \varphi(M) \arrow[dr, "i"] & \\ M \arrow[ur, "m \mapsto \varphi(m)"] \arrow[rr, "\varphi"'] & & M' \end{tikzcd}$$

commutes, where i is the inclusion. Tensoring with N produces a commutative diagram

$$\begin{tikzcd} & N \otimes_R \varphi(M) \arrow[dr, "1\otimes i"] & \\ N \otimes_R M \arrow[ur, "m\otimes n \mapsto m \otimes \varphi(n)"] \arrow[rr, "1\otimes\varphi"'] & & N \otimes_R M'. \end{tikzcd}$$

The map $1 \otimes i$ is injective since i is injective and N is flat. Therefore the two maps out of $N \otimes_R M$ above have the same kernel. The kernel of $N \otimes_R M \longrightarrow N \otimes_R \varphi(M)$ can be computed by Corollary 2.22 to be the natural image of $N \otimes_R \ker\varphi$ inside $N \otimes_R M$, and we can identify the image with $N \otimes_R \ker\varphi$ since N is flat. $\square$

Theorem 3.11. *A tensor product of two flat modules is flat.*

Proof. Let N and N' be flat. For any injective linear map $M \xrightarrow{\varphi} M'$, we want to show the induced linear map $(N \otimes_R N') \otimes_R M \xrightarrow{1\otimes\varphi} (N \otimes_R N') \otimes M'$ is injective.

Since N' is flat, $N' \otimes_R M \xrightarrow{1\otimes\varphi} N' \otimes_R M'$ is injective. Tensoring now with N, $N \otimes_R (N' \otimes_R M) \xrightarrow{1\otimes(1\otimes\varphi)} N \otimes_R (N' \otimes_R M')$ is injective since N is flat. The diagram

$$\begin{tikzcd} N \otimes_R (N' \otimes_R M) \arrow[r, hook, "1\otimes(1\otimes\varphi)"] \arrow[d] & N \otimes_R (N' \otimes M') \arrow[d] \\ (N \otimes_R N') \otimes_R M \arrow[r, "1\otimes\varphi"] & (N \otimes_R N') \otimes M' \end{tikzcd}$$

commutes, where the vertical maps are the natural isomorphisms, so the bottom map is injective. Thus $N \otimes_R N'$ is flat. $\square$

Theorem 3.12. *Let $M \xrightarrow{\varphi} M'$ and $N \xrightarrow{\psi} N'$ be injective linear maps. If the four modules are all flat then $M \otimes_R N \xrightarrow{\varphi\otimes\psi} M' \otimes_R N'$ is injective. More precisely, if M and N' are flat, or M' and N are flat, then $\varphi \otimes \psi$ is injective.*

The precise hypotheses (M and N' flat, or M' and N flat) can be remembered using dotted lines in the diagram below; if both modules connected by one of the dotted lines are flat, then $\varphi \otimes \psi$ is injective.

$$\begin{tikzcd} M \arrow[r, "\varphi"] \arrow[dr, dash, dashed] & M' \arrow[dl, dash, dashed] \\ N \arrow[r, "\psi"'] & N' \end{tikzcd}$$

Proof. The trick is to break up the linear map $\varphi \otimes \psi$ into a composite of linear maps $\varphi \otimes 1$ and $1 \otimes \psi$ in the following commutative diagram.

$$\begin{tikzcd} & M \otimes_R N' \arrow[dr, "\varphi\otimes 1"] & \\ M \otimes_R N \arrow[ur, "1\otimes\psi"] \arrow[rr, "\varphi\otimes\psi"'] & & M' \otimes_R N' \end{tikzcd}$$

Both $\varphi \otimes 1$ and $1 \otimes \psi$ are injective since N' and M are flat, so their composite $\varphi \otimes \psi$ is injective. Alternatively, we can write $\varphi \otimes \psi$ as a composite fitting in the commutative diagram

$$\begin{tikzcd} & M' \otimes_R N \arrow[dr, "1\otimes\psi"] & \\ M \otimes_R N \arrow[ur, "\varphi\otimes 1"] \arrow[rr, "\varphi\otimes\psi"'] & & M' \otimes_R N' \end{tikzcd}$$

and the two diagonal maps are injective from flatness of N and M', so $\varphi \otimes \psi$ is injective. □

Corollary 3.13. *Let $M_1, \ldots, M_k, N_1, \ldots, N_k$ be flat R-modules and $\varphi_i \colon M_i \to N_i$ be injective linear maps. Then the linear map*

$$\varphi_1 \otimes \cdots \otimes \varphi_k \colon M_1 \otimes_R \cdots \otimes_R M_k \to N_1 \otimes \cdots \otimes N_k$$

is injective. In particular, if $\varphi \colon M \to N$ is an injective linear map of flat modules then the tensor powers $\varphi^{\otimes k} \colon M^{\otimes k} \to N^{\otimes k}$ are injective for all $k \geq 1$.

Proof. We argue by induction on k. For $k = 1$ there is nothing to show. Suppose $k \geq 2$ and $\varphi_1 \otimes \cdots \otimes \varphi_{k-1}$ is injective. Then break up $\varphi_1 \otimes \cdots \otimes \varphi_k$ into the composite

$$\begin{tikzcd} & (N_1 \otimes_R \cdots \otimes_R N_{k-1}) \otimes_R M_k \arrow[dr, "1\otimes\varphi_k"] & \\ M_1 \otimes_R \cdots \otimes_R M_{k-1} \otimes_R M_k \arrow[ur, "(\varphi_1\otimes\cdots\otimes\varphi_{k-1})\otimes 1"] \arrow[rr, "\varphi_1\otimes\cdots\otimes\varphi_k"'] & & N_1 \otimes_R \cdots \otimes_R N_k. \end{tikzcd}$$

The first diagonal map is injective because M_k is flat, and the second diagonal map is injective because $N_1 \otimes_R \cdots \otimes_R N_{k-1}$ is flat (Theorem 3.11 and induction). □

Corollary 3.14. *If M and N are free R-modules and $\varphi \colon M \to N$ is an injective linear map, any tensor power $\varphi^{\otimes k} \colon M^{\otimes k} \to N^{\otimes k}$ is injective.*

Proof. Free modules are flat by Theorem 3.2. □

Note the free modules in Corollary 3.14 are completely arbitrary. We make no assumptions about finite bases.

Corollary 3.14 is not a special case of Theorem 2.18 because a free submodule of a free module need not be a direct summand (*e.g.*, $2\mathbf{Z}$ is not a direct summand of $\mathbf{Z}$).

Corollary 3.15. *If M is a free module and $\{m_1, \ldots, m_d\}$ is a finite linearly independent subset then for any $k \leq d$ the d^k elementary tensors*

$$m_{i_1} \otimes \cdots \otimes m_{i_k} \text{ where } i_1, \ldots, i_k \in \{1, 2, \ldots, d\} \tag{3.1}$$

are linearly independent in $M^{\otimes k}$.

Proof. There is an embedding $R^d \hookrightarrow M$ by $\sum_{i=1}^d r_i e_i \mapsto \sum_{i=1}^d r_i m_i$. Since R^d and M are free, the kth tensor power $(R^d)^{\otimes k} \to M^{\otimes k}$ is injective. This map sends the basis

$$e_{i_1} \otimes \cdots \otimes e_{i_k}$$

of $(R^d)^{\otimes k}$, where $i_1, \ldots, i_k \in \{1, 2, \ldots, d\}$, to the elementary tensors in (3.1), so they are linearly independent in $M^{\otimes k}$ □

Corollary 3.15 is *not* saying the elementary tensors in (3.1) can be extended to a basis of $M^{\otimes k}$, just as $m_1, \ldots, m_d$ usually can't be extended to a basis of M.

4. Tensor Products of Linear Maps and Base Extension

Fix a ring homomorphism $R \xrightarrow{f} S$. Every S-module becomes an R-module by restriction of scalars, and every R-module M has a base extension $S \otimes_R M$, which is an S-module. In part I we saw $S \otimes_R M$ has a universal mapping property among all S-modules: an R-linear map from M to any S-module "extends" uniquely to an S-linear map from $S \otimes_R M$ to the S-module. We discuss in this section an arguably more important role for base extension: it turns an R-linear map $M \xrightarrow{\varphi} M'$ between two R-modules into an S-linear map between S-modules. Tensoring $M \xrightarrow{\varphi} M'$ with S gives us an R-linear map $S \otimes_R M \xrightarrow{1 \otimes \varphi} S \otimes_R M'$ that is in fact S-linear: $(1 \otimes \varphi)(st) = s(1 \otimes \varphi)(t)$ for all $s \in S$ and $t \in S \otimes_R M$. Since both sides are additive in t, to prove $1 \otimes \varphi$ is S-linear it suffices to consider the case when $t = s' \otimes m$ is an elementary tensor. Then

$$(1 \otimes \varphi)(s(s' \otimes m)) = (1 \otimes \varphi)(ss' \otimes m) = ss' \otimes \varphi(m) = s(s' \otimes \varphi(m)) = s(1 \otimes \varphi)(s' \otimes m).$$

We will write the base extended linear map $1 \otimes \varphi$ as φ_S to make the S-dependence clearer, so

$$\varphi_S \colon S \otimes_R M \to S \otimes_R M' \text{ by } \varphi_S(s \otimes m) = s \otimes \varphi(m).$$

Since $1 \otimes \mathrm{id}_M = \mathrm{id}_{S \otimes_R M}$ and $(1 \otimes \varphi) \circ (1 \otimes \varphi') = 1 \otimes (\varphi \circ \varphi')$, we have $(\mathrm{id}_M)_S = \mathrm{id}_{S \otimes_R M}$ and $(\varphi \circ \varphi')_S = \varphi_S \circ \varphi'_S$. That means the process of creating S-modules and S-linear maps out of R-modules and R-linear maps is *functorial*.

If an R-linear map $M \xrightarrow{\varphi} M'$ is an isomorphism or is surjective then so is $S \otimes_R M \xrightarrow{\varphi_S} S \otimes_R M'$ (Theorems 2.11 and 2.12). But if φ is injective then φ_S need not be injective. (Examples 2.13, 2.14, and 2.15, which all have S as a field).

Theorem 4.1. *Let R be a nonzero commutative ring. If $R^m \cong R^n$ as R-modules then $m = n$. If there is a linear surjection $R^m \twoheadrightarrow R^n$ then $m \geq n$.*

Proof. Pick a maximal ideal $\mathfrak{m}$ in R. Tensoring R-linear maps $R^m \cong R^n$ or $R^m \twoheadrightarrow R^n$ with $R/\mathfrak{m}$ produces $R/\mathfrak{m}$-linear maps $(R/\mathfrak{m})^m \cong (R/\mathfrak{m})^n$ or $(R/\mathfrak{m})^m \twoheadrightarrow (R/\mathfrak{m})^n$. Taking dimensions over the field $R/\mathfrak{m}$ implies $m = n$ or $m \geq n$, respectively. □

We can't extend this method of proof to show a linear injection $R^m \hookrightarrow R^n$ forces $m \leq n$ because injectivity is not generally preserved under base extension. We will return to this later when we meet exterior powers.

Theorem 4.2. *Let R be a* PID *and M be a finitely generated R-module. Writing*

$$M \cong R^d \oplus R/(a_1) \oplus \cdots \oplus R/(a_k),$$

where $a_1 \mid \cdots \mid a_k$, the integer d equals $\dim_K(K \otimes_R M)$, where K is the fraction field of R. Therefore d is uniquely determined by M.

Proof. Tensoring the displayed R-module isomorphism by K gives a K-vector space isomorphism $K \otimes_R M \cong K^d$ since $K \otimes_R (R/(a_i)) = 0$. Thus $d = \dim_K(K \otimes_R M)$. □

Example 4.3. Let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{M}_2(R)$. Regarding A as a linear map $R^2 \to R^2$, its base extension $A_S \colon S \otimes_R R^2 \to S \otimes_R R^2$ is S-linear and $S \otimes_R R^2 \cong S^2$ as S-modules.

Let $\{e_1, e_2\}$ be the standard basis of R^2. An S-basis for $S \otimes_R R^2$ is $\{1 \otimes e_1, 1 \otimes e_2\}$. Using this basis, we can compute a matrix for A_S:

$$A_S(1 \otimes e_1) = 1 \otimes A(e_1) = 1 \otimes (ae_1 + ce_2) = a(1 \otimes e_1) + c(1 \otimes e_2)$$

and

$$A_S(1 \otimes e_2) = 1 \otimes A(e_2) = 1 \otimes (be_1 + de_2) = b(1 \otimes e_1) + d(1 \otimes e_2).$$

Therefore the matrix for A_S is $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{M}_2(S)$. (Strictly speaking, we should have entries $f(a)$, $f(b)$, and so on.)

The next theorem says base extension doesn't change matrix representations, as in the previous example.

Theorem 4.4. *Let M and M' be nonzero finite-free R-modules and $M \xrightarrow{\varphi} M'$ be an R-linear map. For any bases $\{e_j\}$ and $\{e'_i\}$ of M and M', the matrix for the S-linear map $S \otimes_R M \xrightarrow{\varphi_S} S \otimes_R M'$ with respect to the bases $\{1 \otimes e_j\}$ and $\{1 \otimes e'_i\}$ equals the matrix for φ with respect to $\{e_j\}$ and $\{e'_i\}$.*

Proof. Say $\varphi(e_j) = \sum_i a_{ij} e'_i$, so the matrix of φ is (a_{ij}). Then

$$\varphi_S(1 \otimes e_j) = 1 \otimes \varphi(e_j) = 1 \otimes \sum_i a_{ij} e_i = \sum_i a_{ij}(1 \otimes e_i),$$

so the matrix of φ_S is also (a_{ij}). □

Example 4.5. Any $n \times n$ real matrix acts on $\mathbf{R}^n$, and its base extension to $\mathbf{C}$ acts on $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{R}^n \cong \mathbf{C}^n$ as the same matrix. An $n \times n$ integral matrix acts on $\mathbf{Z}^n$ and its base extension to $\mathbf{Z}/m\mathbf{Z}$ acts on $\mathbf{Z}/m\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}^n \cong (\mathbf{Z}/m\mathbf{Z})^n$ as the same matrix reduced mod m.

Theorem 4.6. *Let M and M' be R-modules. There is a unique S-linear map*

$$S \otimes_R \mathrm{Hom}_R(M, M') \to \mathrm{Hom}_S(S \otimes_R M, S \otimes_R M')$$

sending $s \otimes \varphi$ to the function $s\varphi_S \colon t \mapsto s\varphi_S(t)$ and it is an isomorphism if M and M' are finite free R-modules. In particular, using $M' = M$, there is a unique S-linear map

$$S \otimes_R (M^{\vee_R}) \to (S \otimes_R M)^{\vee_S}$$

where $s \otimes \varphi \mapsto s\varphi_S$ on elementary tensors, and it is an isomorphism if M is a finite-free R-module.

The point of this theorem in the finite-free case is that it says base extension on linear maps accounts (through S-linear combinations) for all S-linear maps between base extended R-modules. This doesn't mean every S-linear map *is* a base extension, which would be like saying every tensor is an elementary tensor rather than just a sum of them.

Proof. The function $S \times \mathrm{Hom}_R(M, M') \to \mathrm{Hom}_S(S \otimes_R M, S \otimes_R M')$ where $(s, \varphi) \mapsto s\varphi_S$ is R-bilinear (check!), so there is a unique R-linear map

$$S \otimes_R \mathrm{Hom}_R(M, M') \xrightarrow{L} \mathrm{Hom}_S(S \otimes_R M, S \otimes_R M')$$

such that $L(s \otimes \varphi) = s\varphi_S$. The map L is S-linear (check!). If $M' = R$ and we identify $S \otimes_R R$ with S as S-modules by multiplication, then L becomes an S-linear map $S \otimes_R (M^{\vee_R}) \to (S \otimes_R M)^{\vee_S}$.

Now suppose M and M' are both finite free. We want to show L is an isomorphism. If M or M' is 0 it is clear, so we may take them both to be nonzero with respective R-bases $\{e_i\}$ and $\{e'_j\}$, say. Then S-bases of $S \otimes_R M$ and $S \otimes_R M'$ are $\{1 \otimes e_i\}$ and $\{1 \otimes e'_j\}$. An R-basis of $\mathrm{Hom}_R(M, M')$ is the functions φ_{ij} sending e_i to e'_j and other basis vectors e_k of M to 0. An S-basis of $S \otimes_R \mathrm{Hom}_R(M, M')$ is the tensors $\{1 \otimes \varphi_{ij}\}$, and

$$L(1 \otimes \varphi_{ij})(1 \otimes e_i) = (\varphi_{ij})_S(1 \otimes e_i) = (1 \otimes \varphi_{ij})(1 \otimes e_i) = 1 \otimes \varphi_{ij}(e_i) = 1 \otimes e'_j$$

while $L(1 \otimes \varphi_{ij})(1 \otimes e_k) = 0$ for $k \neq i$. That means L sends a basis to a basis, so L is an isomorphism. □

Example 4.7. Taking $R = \mathbf{R}$, $S = \mathbf{C}$, $M = \mathbf{R}^n$, and $M' = \mathbf{R}^m$, the theorem says $\mathbf{C} \otimes_{\mathbf{R}} \mathrm{M}_{m,n}(\mathbf{R}) \cong \mathrm{M}_{m,n}(\mathbf{C})$ as complex vector spaces by sending the elementary tensor $z \otimes A$ for $z \in \mathbf{C}$ and $A \in \mathrm{M}_{m,n}(\mathbf{R})$ to the matrix zA. In particular, $\mathbf{C} \otimes_{\mathbf{R}} \mathrm{M}_n(\mathbf{R}) \cong \mathrm{M}_n(\mathbf{C})$.

Example 4.8. Let $R = \mathbf{Z}/p^2\mathbf{Z}$ and $S = R/pR$ as rings. Then S is an R-module by using the natural reduction map $R \to S$. Let M be the R-module S (an additive group of order p). Note S is *not* a free R-module: $pS = 0$ and $p \neq 0$ in R.

We have $S \otimes_R M = S \otimes_R S \cong S$ as S-modules.[5] The natural linear map

$$S \otimes_R (M^{\vee_R}) \to (S \otimes_R M)^{\vee_S} \tag{4.1}$$

has image 0: check that each φ in $M^{\vee_R} = \mathrm{Hom}_R(R/pR, R)$ has values in pR, so the image of $1 \otimes \varphi$ on the right side of (4.1) has values in pS, which is 0. The right side of (4.1) is not 0 since $\mathrm{Hom}_S(S, S) \cong S$, so (4.1) isn't an isomorphism.

In Corollary 3.14 we saw the tensor powers of an injective linear map between free modules over *any commutative ring* are all injective. Using base extension, we can drop the requirement that the target module be free provided we are working over a domain.

Theorem 4.9. *Let R be a domain and $\varphi\colon M \hookrightarrow N$ be an injective linear map where M is free. Then $\varphi^{\otimes k}\colon M^{\otimes k} \to N^{\otimes k}$ is injective for any $k \geq 1$.*

[5]More generally, $R/I \otimes_R R/I \cong R/I$ as R-modules and then also as R/I-modules where R/I scales elementary tensors $x \otimes y$ in $R/I \otimes_R R/I$ on the left ($a(x \otimes y) = ax \otimes y$).

Proof. We have a commutative diagram

$$\begin{tikzcd} M^{\otimes k} \arrow[r, "\varphi^{\otimes k}"] \arrow[d] & N^{\otimes k} \arrow[d] \\ K \otimes_R M^{\otimes k} \arrow[r, "1\otimes(\varphi^{\otimes k})"] \arrow[d] & K \otimes_R N^{\otimes k} \arrow[d] \\ (K \otimes_R M)^{\otimes k} \arrow[r, "(1\otimes\varphi)^{\otimes k}"] & (K \otimes_R N)^{\otimes k} \end{tikzcd}$$

where the top vertical maps are the natural ones ($t \mapsto 1 \otimes t$) and the bottom vertical maps are the base extension isomorphisms. (Tensor powers along the bottom are over K while those on the first and second rows are over R.) From commutativity, to show $\varphi^{\otimes k}$ along the top is injective it suffices to show the composite map along the left side and the bottom is injective. The K-linear map map $K \otimes_R M \xrightarrow{1\otimes\varphi} K \otimes_R N$ is injective since K is a flat R-module, and therefore the map along the bottom is injective (tensor products of injective linear maps of vector spaces are injective). The bottom vertical map on the left is an isomorphism. The top vertical map on the left is injective since $M^{\otimes k}$ is free and thus torsion-free (R is a domain). $\square$

This theorem may not be true if M isn't free. Look at Example 2.16.

5. Vector Spaces

Because all (nonzero) vector spaces have bases, the results we have discussed for modules assume a simpler form when we are working with vector spaces. We will review what we have done in the setting of vector spaces and then discuss some further special properties of this case.

Let K be a field. Tensor products of K-vector spaces involve no unexpected collapsing: if V and W are nonzero K-vector spaces then $V \otimes_K W$ is nonzero and in fact $\dim_K(V \otimes_K W) = \dim_K(V)\dim_K(W)$ in the sense of cardinal numbers.

For any K-linear maps $V \xrightarrow{\varphi} V'$ and $W \xrightarrow{\psi} W'$, we have the tensor product linear map $V \otimes_K W \xrightarrow{\varphi\otimes\psi} V' \otimes_K W'$ that sends $v \otimes w$ to $\varphi(v) \otimes \psi(w)$. When $V \xrightarrow{\varphi} V'$ and $W \xrightarrow{\psi} W'$ are isomorphisms or surjective, so is $V \otimes_K W \xrightarrow{\varphi\otimes\psi} V' \otimes_K W'$ (Theorems 2.11 and 2.12). Moreover, because all K-vector spaces are free a tensor product of injective K-linear maps is injective (Theorem 3.2).

Example 5.1. If $V \xrightarrow{\varphi} W$ is an injective K-linear map and U is any K-vector space, the K-linear map $U \otimes_K V \xrightarrow{1\otimes\varphi} U \otimes_K W$ is injective.

Example 5.2. A tensor product of subspaces "is" a subspace: if $V \subset V'$ and $W \subset W'$ the natural linear map $V \otimes_K W \to V' \otimes_K W'$ is injective.

Because of this last example, we can treat a tensor product of subspaces as a subspace of the tensor product. For example, if $V \xrightarrow{\varphi} V'$ and $W \xrightarrow{\psi} W'$ are linear then $\varphi(V) \subset V'$ and $\psi(W) \subset W'$, so we can regard $\varphi(V) \otimes_K \psi(W)$ as a subspace of $V' \otimes_K W'$, which we couldn't do with modules in general. The following result gives us some practice with this viewpoint.

Theorem 5.3. *Let $V \subset V'$ and $W \subset W'$ where V' and W' are nonzero. Then $V \otimes_K W = V' \otimes_K W'$ if and only if $V = V'$ and $W = W'$.*

Proof. Since $V \otimes_K W$ is inside both $V \otimes_K W'$ and $V' \otimes_K W$, which are inside $V' \otimes_K W'$, by reasons of symmetry it suffices to assume $V \subsetneq V'$ and show $V \otimes_K W' \subsetneq V' \otimes_K W'$.

Since V is a proper subspace of V', there is a linear functional $\varphi \colon V' \to K$ that vanishes on V and is not identically 0 on V', so $\varphi(v_0') = 1$ for some $v_0' \in V'$. Pick nonzero $\psi \in W'^\vee$, and say $\psi(w_0') = 1$. Then the linear function $V' \otimes_K W' \to K$ where $v' \otimes w' \mapsto \varphi(v')\psi(w')$ vanishes on all of $V \otimes_K W'$ by checking on elementary tensors but its value on $v_0' \otimes w_0'$ is 1. Therefore $v_0' \otimes w_0' \notin V \otimes_K W'$, so $V \otimes_K W' \subsetneq V' \otimes_K W'$. □

When V and W are finite-dimensional, the K-linear map

$$\operatorname{Hom}_K(V, V') \otimes_K \operatorname{Hom}_K(W, W') \to \operatorname{Hom}_K(V \otimes_K W, V' \otimes_K W') \tag{5.1}$$

sending the elementary tensor $\varphi \otimes \psi$ to the linear map denoted $\varphi \otimes \psi$ is an isomorphism (Theorem 2.5). So the two possible meanings of $\varphi\otimes\psi$ (elementary tensor in a tensor product of Hom-spaces or linear map on a tensor product of vector spaces) really match up. Taking $V' = K$ and $W' = K$ in (5.1) and identifying $K \otimes_K K$ with K by multiplication, (5.1) says $V^\vee \otimes_K W^\vee \cong (V \otimes_K W)^\vee$ using the obvious way of making a tensor $\varphi \otimes \psi$ in $V^\vee \otimes_K W^\vee$ act on $V \otimes_K W$, namely through multiplication of the values: $(\varphi \otimes \psi)(v \otimes w) = \varphi(v)\psi(w)$. By induction on the numbers of terms,

$$V_1^\vee \otimes_K \cdots \otimes_K V_k^\vee \cong (V_1 \otimes_K \cdots \otimes_K V_k)^\vee$$

when the V_i's are finite-dimensional. Here an elementary tensor $\varphi_1 \otimes \cdots \otimes \varphi_k \in \bigotimes_{i=1}^k V_i^\vee$ acts on an elementary tensor $v_1 \otimes \cdots \otimes v_k \in \bigotimes_{i=1}^k V_i$ with value $\varphi_1(v_1) \cdots \varphi_k(v_k) \in K$. In particular,

$$(V^\vee)^{\otimes k} \cong (V^{\otimes k})^\vee$$

when V is finite-dimensional.

Let's turn now to base extensions to larger fields. When L/K is any field extension,[6] base extension turns K-vector spaces into L-vector spaces ($V \rightsquigarrow L \otimes_K V$) and K-linear maps into L-linear maps ($\varphi \rightsquigarrow \varphi_L := 1 \otimes \varphi$). Provided V and W are finite-dimensional over K, base extension of linear maps $V \longrightarrow W$ accounts for all the linear maps between $L \otimes_K V$ and $L \otimes_K W$ using L-linear combinations, in the sense that the natural L-linear map

$$L \otimes_K \operatorname{Hom}_K(V, W) \cong \operatorname{Hom}_L(L \otimes_K V, L \otimes_K W) \tag{5.2}$$

is an isomorphism (Theorem 4.6). When we choose K-bases for V and W and use the corresponding L-bases for $L \otimes_K V$ and $L \otimes_K W$, the matrix representations of a K-linear map $V \to W$ and its base extension by L are the same (Theorem 4.4). Taking $W = K$, the natural L-linear map

$$L \otimes_K V^\vee \cong (L \otimes_K V)^\vee \tag{5.3}$$

is an isomorphism for finite-dimensional V, using K-duals on the left and L-duals on the right.[7]

[6]We allow infinite or even non-algebraic extensions, such as $\mathbf{R}/\mathbf{Q}$.

[7]If we drop finite-dimensionality assumptions, (5.1), (5.2), and (5.3) are all still injective but generally not surjective.

Remark 5.4. We don't really need L to be a field; K-vector spaces are free and therefore their base extensions to modules over any commutative ring containing K will be free as modules over the larger ring. For example, the characteristic polynomial of a linear operator $V \xrightarrow{\varphi} V$ could be defined in a coordinate-free way using base extension of V from K to $K[T]$: the characteristic polynomial of φ is the determinant of the linear operator $T \otimes \mathrm{id}_V - \varphi_{K[T]} \colon K[T] \otimes_K V \longrightarrow K[T] \otimes_K V$ since $\det(T \otimes \mathrm{id}_V - \varphi_{K[T]}) = \det(TI_n - A)$, where A is a matrix representation of φ.

We will make no finite-dimensionality assumptions in the rest of this section.

The next theorem tells us the image and kernel of a tensor product of linear maps of vector spaces, with no surjectivity hypotheses as in Theorem 2.19.

Theorem 5.5. *Let $V_1 \xrightarrow{\varphi_1} W_1$ and $V_2 \xrightarrow{\varphi_2} W_2$ be linear. Then*

$$\ker(\varphi_1 \otimes \varphi_2) = \ker \varphi_1 \otimes_K V_2 + V_1 \otimes_K \ker \varphi_2, \quad \mathrm{Im}(\varphi_1 \otimes \varphi_2) = \varphi_1(V_1) \otimes_K \varphi_2(V_2).$$

In particular, if V_1 and V_2 are nonzero then $\varphi_1 \otimes \varphi_2$ is injective if and only if φ_1 and φ_2 are injective, and if W_1 and W_2 are nonzero then $\varphi_1 \otimes \varphi_2$ is surjective if and only if φ_1 and φ_2 are surjective.

Here we are taking advantage of the fact that in vector spaces a tensor product of subspaces is naturally a subspace of the tensor product: $\ker \varphi_1 \otimes_K V_2$ can be identified with its image in $V_1 \otimes_K V_2$ and $\varphi_1(V_1) \otimes_K \varphi_2(V_2)$ can be identified with its image in $W_1 \otimes_K W_2$ under the natural maps. Theorem 2.19 for modules has weaker conclusions (*e.g.*, injectivity of $\varphi_1 \otimes \varphi_2$ doesn't imply injectivity of φ_1 and φ_2).

Proof. First we handle the image of $\varphi_1 \otimes \varphi_2$. The diagrams

$$\begin{tikzcd} & \varphi_1(V_1) \arrow[dr, "i_1"] & \\ V_1 \arrow[ur, "v_1 \mapsto \varphi_1(v_1)"] \arrow[rr, "\varphi_1"'] & & W_1 \end{tikzcd} \qquad \begin{tikzcd} & \varphi_2(V_2) \arrow[dr, "i_2"] & \\ V_2 \arrow[ur, "v_2 \mapsto \varphi_2(v_2)"] \arrow[rr, "\varphi_2"'] & & W_2 \end{tikzcd}$$

commute, with i_1 and i_2 being injections, so the composite diagram

$$\begin{tikzcd} & \varphi_1(V_1) \otimes_K \varphi_2(V_2) \arrow[dr, "i_1 \otimes i_2"] & \\ V_1 \otimes_K V_2 \arrow[ur, "v_1 \otimes v_2 \mapsto \varphi_1(v_1) \otimes \varphi_2(v_2)"] \arrow[rr, "\varphi_1 \otimes \varphi_2"'] & & W_1 \otimes_K W_2 \end{tikzcd}$$

commutes. As $i_1 \otimes i_2$ is injective, both maps out of $V_1 \otimes_K V_2$ have the same kernel. The kernel of the map $V_1 \otimes_K V_2 \longrightarrow \varphi_1(V_1) \otimes_K \varphi_2(V_2)$ can be computed by Theorem 2.19 to be $\ker \varphi_1 \otimes_K V_2 + V_1 \otimes_K \ker \varphi_2$, where we identify tensor products of subspaces with a subspace of the tensor product.

If $\varphi_1 \otimes \varphi_2$ is injective then its kernel is 0, so $0 = \ker \varphi_1 \otimes_K V_2 + V_1 \otimes_K \ker \varphi_2$ from the kernel formula. Therefore the subspaces $\ker \varphi_1 \otimes_K V_2$ and $V_1 \otimes_K \ker \varphi_2$ both vanish, so $\ker \varphi_1$ and $\ker \varphi_2$ must vanish (because V_2 and V_1 are nonzero, respectively). Conversely, if φ_1 and φ_2 are injective then we already knew $\varphi_1 \otimes \varphi_2$ is injective, but the formula for $\ker(\varphi_1 \otimes \varphi_2)$ also shows us this kernel is 0.

If $\varphi_1 \otimes \varphi_2$ is surjective then the formula for its image shows $\varphi_1(V_1) \otimes_K \varphi_2(V_2) = W_1 \otimes_K W_2$, so $\varphi_1(V_1) = W_1$ and $\varphi_2(V_2) = W_2$ by Theorem 5.3 (here we need W_1 and W_2 nonzero).

Conversely, if φ_1 and φ_2 are surjective then so is $\varphi_1 \otimes \varphi_2$ because that's true for all modules. $\square$

Corollary 5.6. *Let $V \subset V'$ and $W \subset W'$. Then*

$$(V' \otimes_K W')/(V \otimes_K W' + V' \otimes_K W) \cong (V'/V) \otimes_K (W'/W).$$

Proof. Tensor the natural projections $V' \xrightarrow{\pi_1} V'/V$ and $W' \xrightarrow{\pi_2} W'/W$ to get a linear map $V' \otimes_K W' \xrightarrow{\pi_1 \otimes \pi_2} (V'/V) \otimes_K (W'/W)$ that is onto with $\ker(\pi_1 \otimes \pi_2) = V \otimes_K W' + V' \otimes_K W$ by Theorem 5.5. $\square$

Remark 5.7. It is *false* that $(V' \otimes_K W')/(V \otimes_K W) \cong (V'/V) \otimes_K (W'/W)$. The subspace $V \otimes_K W$ is generally too small[8] to be the kernel. This is a distinction between tensor products and direct sums (where $(V' \oplus W')/(V \oplus W) \cong (V'/V) \oplus (W'/W)$).

Corollary 5.8. *Let $V \xrightarrow{\varphi} W$ be a linear map and U be a K-vector space. The linear map $U \otimes_K V \xrightarrow{1\otimes\varphi} U \otimes_K W$ has kernel and image*

$$\ker(1 \otimes \varphi) = U \otimes_K \ker\varphi \quad \textit{and} \quad \operatorname{Im}(1 \otimes \varphi) = U \otimes_K \varphi(V). \tag{5.4}$$

In particular, for nonzero U the map φ is injective or surjective if and only if $1 \otimes \varphi$ has that property.

Proof. This is immediate from Theorem 5.5 since we're using the identity map on U. $\square$

Example 5.9. Let $V \xrightarrow{\varphi} W$ be a linear map and L/K be a field extension. The base extension $L \otimes_K V \xrightarrow{\varphi_L} L \otimes_K W$ has kernel and image

$$\ker(\varphi_L) = L \otimes_K \ker\varphi, \quad \operatorname{Im}(\varphi_L) = L \otimes_K \operatorname{Im}(\varphi).$$

The map φ is injective if and only if φ_L is injective and φ is surjective if and only if φ_L is surjective.

Let's formulate this in the language of matrices. If V and W are finite-dimensional then φ can be written as a matrix with entries in K once we pick bases of V and W. Then φ_L has the same matrix representation relative to the corresponding bases of $L \otimes_K V$ and $L \otimes_K W$. Since the base extension of a free module to another ring doesn't change the size of a basis, $\dim_L(L \otimes_K \operatorname{Im}(\varphi)) = \dim_K \operatorname{Im}(\varphi)$ and $\dim_L(L \otimes_K \ker(\varphi)) = \dim_K \ker(\varphi)$. That means φ and φ_L have the same rank and the same nullity: the rank and nullity of a matrix in $\mathrm{M}_{m\times n}(K)$ do not change when it is viewed in $\mathrm{M}_{m\times n}(L)$ for any field extension L/K.

In the rest of this section we will look at tensor products of many vector spaces at once.

Lemma 5.10. *For $v \in V$ with $v \neq 0$, there is $\varphi \in V^\vee$ such that $\varphi(v) = 1$.*

Proof. The set $\{v\}$ is linearly independent, so it extends to a basis $\{v_i\}_{i\in I}$ of V. Let $v = v_{i_0}$ in this indexing. Define $\varphi\colon V \to K$ by

$$\varphi\left(\sum_i c_i v_i\right) = c_{i_0}.$$

Then $\varphi \in V^\vee$ and $\varphi(v) = \varphi(v_{i_0}) = 1$. $\square$

Theorem 5.11. *Let $V_1, \ldots, V_k$ be K-vector spaces and $v_i \in V_i$. Then $v_1 \otimes \cdots \otimes v_k = 0$ in $V_1 \otimes_K \cdots \otimes_K V_k$ if and only if some v_i is 0.*

[8] Exception: $V' = V$ or $W = 0$, and $W' = W$ or $V = 0$.

Proof. The direction ($\Leftarrow$) is clear. To prove ($\Rightarrow$), we show the contrapositive: if every v_i is nonzero then $v_1 \otimes \cdots \otimes v_k \neq 0$. By Lemma 5.10, for $i = 1, \dots, k$ there is $\varphi_i \in V_i^\vee$ with $\varphi_i(v_i) = 1$. Then $\varphi_1 \otimes \cdots \otimes \varphi_k$ is a linear map $V_1 \otimes_K \cdots \otimes_K V_k \to K$ having the effect

$$v_1 \otimes \cdots \otimes v_k \mapsto \varphi_1(v_1) \cdots \varphi_k(v_k) = 1 \neq 0,$$

so $v_1 \otimes \cdots \otimes v_k \neq 0$. □

Corollary 5.12. *Let $\varphi_i \colon V_i \to W_i$ be linear maps between K-vector spaces for $1 \leq i \leq k$. Then the linear map $\varphi_1 \otimes \cdots \otimes \varphi_k \colon V_1 \otimes_K \cdots \otimes_K V_k \to W_1 \otimes_K \cdots \otimes_K W_k$ is O if and only if some φ_i is O.*

Proof. For ($\Leftarrow$), if some φ_i is O then $(\varphi_1 \otimes \cdots \otimes \varphi_k)(v_1 \otimes \cdots \otimes v_k) = \varphi_1(v_1) \otimes \cdots \otimes \varphi_k(v_k) = 0$ since $\varphi_i(v_i) = 0$. Therefore $\varphi_1 \otimes \cdots \otimes \varphi_k$ vanishes on all elementary tensors, so it vanishes on $V_1 \otimes_K \cdots \otimes_K V_k$, so $\varphi_1 \otimes \cdots \otimes \varphi_k = O$.

To prove ($\Rightarrow$), we show the contrapositive: if every φ_i is nonzero then $\varphi_1 \otimes \cdots \otimes \varphi_k \neq O$. Since $\varphi_i \neq O$, we can find some v_i in V_i with $\varphi_i(v_i) \neq 0$ in W_i. Then $\varphi_1 \otimes \cdots \otimes \varphi_k$ sends $v_1 \otimes \cdots \otimes v_k$ to $\varphi_1(v_1) \otimes \cdots \otimes \varphi_k(v_k)$. Since each $\varphi_i(v_i)$ is nonzero in W_i, the elementary tensor $\varphi_1(v_1) \otimes \cdots \otimes \varphi_k(v_k)$ is nonzero in $W_1 \otimes_K \cdots \otimes_K W_k$ by Theorem 5.11. Thus $\varphi_1 \otimes \cdots \otimes \varphi_k$ takes a nonzero value, so it is not the zero map. □

Corollary 5.13. *If R is a domain and M and N are R-modules, for non-torsion x in M and y in N, $x \otimes y$ is non-torsion in $M \otimes_R N$.*

Proof. Let K be the fraction field of R. The torsion elements of $M \otimes_R N$ are precisely the elements that go to 0 under the map $M \otimes_R N \to K \otimes_R (M \otimes_R N)$ sending t to $1 \otimes t$. We want to show $1 \otimes (x \otimes y) \neq 0$.

The natural K-vector space isomorphism $K \otimes_R (M \otimes_R N) \cong (K \otimes_R M) \otimes_K (K \otimes_R N)$ identifies $1 \otimes (x \otimes y)$ with $(1 \otimes x) \otimes (1 \otimes y)$. Since x and y are non-torsion in M and N, $1 \otimes x \neq 0$ in $K \otimes_R M$ and $1 \otimes y \neq 0$ in $K \otimes_R N$. An elementary tensor of nonzero vectors in two K-vector spaces is nonzero (Theorem 5.11), so $(1 \otimes x) \otimes (1 \otimes y) \neq 0$ in $(K \otimes_R M) \otimes_K (K \otimes_R N)$. Therefore $1 \otimes (x \otimes y) \neq 0$ in $K \otimes_R (M \otimes_R N)$, which is what we wanted to show. □

Remark 5.14. If M and N are torsion-free, Corollary 5.13 is not saying $M \otimes_R N$ is torsion-free. It only says all (nonzero) elementary tensors have no torsion. There could be non-elementary tensors with torsion, as we saw at the end of Example 2.16.

In Theorem 5.11 we saw an elementary tensor in a tensor product of vector spaces is 0 only under the obvious condition that one of the vectors appearing in the tensor is 0. We now show two nonzero elementary tensors in vector spaces are equal only under the "obvious" circumstances.

Theorem 5.15. *Let $V_1, \dots, V_k$ be K-vector spaces. Pick pairs of nonzero vectors v_i, v_i' in V_i for $i = 1, \dots, k$. Then $v_1 \otimes \cdots \otimes v_k = v_1' \otimes \cdots \otimes v_k'$ in $V_1 \otimes_K \cdots \otimes_K V_k$ if and only if there are nonzero constants $c_1, \dots, c_k$ in K such that $v_i = c_i v_i'$ and $c_1 \cdots c_k = 1$.*

Proof. If $v_i = c_i v_i'$ for all i and $c_1 \cdots c_k = 1$ then $v_1 \otimes \cdots \otimes v_k = c_1 v_1' \otimes \cdots \otimes c_k v_k' = (c_1 \cdots c_k) v_1' \otimes \cdots \otimes v_k' = v_1' \otimes \cdots \otimes v_k'$.

Now we want to go the other way. It is clear for $k = 1$, so we may take $k \geq 2$.

By Theorem 5.11, $v_1 \otimes \cdots \otimes v_k$ is not 0 since each v_i is not 0. Fix $\varphi_i \in V_i^\vee$ for $1 \leq i \leq k-1$ such that $\varphi_i(v_i) = 1$. (Such φ_i exist since $v_1, \dots, v_{k-1}$ are nonzero vectors.) For *arbitrary*

$\varphi \in V_k^\vee$, let $h_\varphi = \varphi_1 \otimes \cdots \otimes \varphi_{k-1} \otimes \varphi$, so $h_\varphi(v_1 \otimes \cdots \otimes v_{k-1} \otimes v_k) = \varphi(v_k)$. Also

$$h_\varphi(v_1 \otimes \cdots \otimes v_{k-1} \otimes v_k) = h_\varphi(v_1' \otimes \cdots \otimes v_{k-1}' \otimes v_k') = \varphi_1(v_1') \cdots \varphi_{k-1}(v_{k-1}')\varphi(v_k') = \varphi(c_k v_k'),$$

where $c_k := \varphi_1(v_1') \cdots \varphi_{k-1}(v_{k-1}') \in K$. So we have

$$\varphi(v_k) = \varphi(c_k v_k')$$

for arbitrary $\varphi \in V_k^\vee$. Therefore $\varphi(v_k - c_k v_k') = 0$ for all $\varphi \in V_k^\vee$, so $v_k - c_k v_k' = 0$, which says $v_k = c_k v_k'$.

In the same way, for every $i = 1, 2, \ldots, k$ there is c_i in K such that $v_i = c_i v_i'$. Then $v_1 \otimes \cdots \otimes v_k = c_1 v_1' \otimes \cdots \otimes c_k v_k' = (c_1 \cdots c_k)(v_1' \otimes \cdots \otimes v_k')$. Since $v_1 \otimes \cdots \otimes v_k = v_1' \otimes \cdots \otimes v_k' \neq 0$, we get $c_1 \cdots c_k = 1$. $\square$

Theorem 5.15 has an interesting interpretation in terms of subspaces. An elementary tensor $v_1 \otimes \cdots \otimes v_k$ does *not* determine all the individual vectors v_i, since they can each be scaled by a nonzero element c_i of K without changing $v_1 \otimes \cdots \otimes v_k$ as long as the product of the c_i's is 1. That's the easier direction of Theorem 5.15. What the harder direction of Theorem 5.15 tells us is that such scaling is the *only* way we can change the v_i's while keeping the elementary tensor $v_1 \otimes \cdots \otimes v_k$ unchanged. In other words, $v_1 \otimes \cdots \otimes v_k$ does not determine the v_i's but it does determine the 1-dimensional subspaces Kv_i in V_i. And since scaling $v_1 \otimes \cdots \otimes v_k$ is the same as scaling one of the v_i's (any of them), the *subspace* $K(v_1 \otimes \cdots \otimes v_k)$ determines the *subspaces* Kv_i (the converse is much easier).

In quantum mechanics, the quantum states of a system are described by the nonzero vectors in a complex Hilbert space H where a (nonzero) scalar multiple of a vector in H determines the same quantum state as the original vector (this condition is motivated by physics), so the states of a quantum system can be described by the 1-dimensional subspaces of H instead of by the individual (nonzero) elements of H.[9] When two quantum systems with corresponding Hilbert spaces H_1 and H_2 are combined, the Hilbert space for the combined system is the (completed) tensor product $H_1 \otimes_{\mathbf{C}} H_2$. In $H_1 \otimes_{\mathbf{C}} H_2$, a 1-dimensional subspace $\mathbf{C}(v_1 \otimes v_2)$ spanned by an elementary tensor determines the individual subspaces $\mathbf{C}v_1$ and $\mathbf{C}v_2$ of H_1 and H_2 by Theorem 5.15, but most 1-dimensional subspaces of $H_1 \otimes_{\mathbf{C}} H_2$ are *not* spanned by an elementary tensor and thus do not "come from" particular 1-dimensional subspaces of H_1 and H_2. The non-elementary tensors in $H_1 \otimes_{\mathbf{C}} H_2$ describe states that are called "entangled" and Schrödinger [4, p. 555] called this phenomenon "*the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought [...] the best possible knowledge of a *whole* does not necessarily include the best possible knowledge of all its *parts*." He did not use the terminology of tensor products, but you can see it at work in what he did say [4, p. 556], where his Hilbert spaces are L^2-spaces of functions on some $\mathbf{R}^n$: "Let x and y stand for all the coordinates of the first and second systems respectively and $\Psi(x, y)$ [stand for the] state of the composed system [...]. What constitutes the entanglement is that Ψ is not a product of a function of x and a function of y." That is analogous to most elements of $\mathbf{R}[x, y]$ not being of the form $f(x)g(y)$, *e.g.*, $x^3y - xy^2 + 7$ is an "entangled" polynomial.

Here is the analogue of Theorem 5.15 for linear maps (compare to Corollary 5.12).

Theorem 5.16. *Let $\varphi_i \colon V_i \to W_i$ and $\varphi_i' \colon V_i \to W_i$ be nonzero linear maps between K-vector spaces for $1 \leq i \leq k$. Then $\varphi_1 \otimes \cdots \otimes \varphi_k = \varphi_1' \otimes \cdots \otimes \varphi_k'$ as linear maps $V_1 \otimes_K$*

[9] A nonzero vector in H, viewed as a quantum state, is often scaled to have length 1, but this still allows ambiguity up to scaling by a complex number $e^{i\theta}$ of absolute value 1, with θ called a phase-factor.

$\cdots \otimes_K V_k \to W_1 \otimes_K \cdots \otimes_K W_k$ *if and only if there are* $c_1, \ldots, c_k$ *in* K *such that* $\varphi_i = c_i\varphi_i'$ *and* $c_1c_2\cdots c_k = 1$.

Proof. Since each $\varphi_i \colon V_i \to W_i$ is not identically 0, for $i = 1, \ldots, k-1$ there is $v_i \in V_i$ such that $\varphi_i(v_i) \neq 0$ in W_i. Then there is $f_i \in W_i^\vee$ such that $f_i(\varphi_i(v_i)) = 1$.

Pick any $v \in V_k$ and $f \in W_k^\vee$. Set $h_f = f_1 \otimes \cdots \otimes f_{k-1} \otimes f \in (W_1 \otimes_K \cdots \otimes_K W_k)^\vee$ where

$$h_f(w_1 \otimes \cdots \otimes w_{k-1} \otimes w_k) = f_1(w_1)\cdots f_{k-1}(w_{k-1})f(w_k).$$

Then

$$h_f((\varphi_1 \otimes \cdots \otimes \varphi_{k-1} \otimes \varphi_k)(v_1 \otimes \cdots \otimes v_{k-1} \otimes v)) = f_1(\varphi_1(v_1))\cdots f_{k-1}(\varphi_{k-1}(v_{k-1}))f(\varphi_k(v)),$$

and since $f_i(\varphi_i(v_i)) = 1$ for $i \neq k$, the value is $f(\varphi_k(v))$. Also

$$h_f((\varphi_1' \otimes \cdots \otimes \varphi_{k-1}' \otimes \varphi_k')(v_1 \otimes \cdots \otimes v_{k-1} \otimes v)) = f_1(\varphi_1'(v_1))\cdots f_{k-1}(\varphi_{k-1}'(v_{k-1}))f(\varphi_k'(v)).$$

Set $c_k = f_1(\varphi_1'(v_1))\cdots f_{k-1}(\varphi_{k-1}'(v_{k-1}))$, so the value is $c_kf(\varphi_k'(v)) = f(c_k\varphi_k'(v))$. Since $\varphi_1 \otimes \cdots \otimes \varphi_{k-1} \otimes \varphi_k = \varphi_1' \otimes \cdots \otimes \varphi_{k-1}' \otimes \varphi_k'$,

$$f(\varphi_k(v)) = f(c_k\varphi_k'(v)).$$

This holds for all $f \in W_k^\vee$, so $\varphi_k(v) = c_k\varphi_k'(v)$. This holds for all $v \in V_k$, so $\varphi_k = c_k\varphi_k'$ as linear maps $V_k \to W_k$.

In a similar way, there is $c_i \in K$ such that $\varphi_i = c_i\varphi_i'$ for all i, so

$$\begin{aligned}\varphi_1 \otimes \cdots \otimes \varphi_k &= (c_1\varphi_1') \otimes \cdots \otimes (c_k\varphi_k')\\ &= (c_1\cdots c_k)\varphi_1' \otimes \cdots \otimes \varphi_k'\\ &= (c_1\cdots c_k)\varphi_1 \otimes \cdots \otimes \varphi_k,\end{aligned}$$

so $c_1\cdots c_k = 1$ since $\varphi_1 \otimes \cdots \otimes \varphi_k \neq O$. □

Remark 5.17. When the V_i's and W_i's are finite-dimensional, the tensor product of linear maps between them can be identified with elementary tensors in the tensor product of the vector spaces of linear maps (Theorem 2.5), so in this special case Theorem 5.16 is a special case of Theorem 5.15. Theorem 5.16 does not assume the vector spaces are finite-dimensional.

When we have k copies of a vector space V, any permutation $\sigma \in S_k$ acts on the direct sum $V^{\oplus k}$ by permuting the coordinates:

$$(v_1, \ldots, v_k) \mapsto (v_{\sigma^{-1}(1)}, \ldots, v_{\sigma^{-1}(k)}).$$

The inverse on σ is needed to get a genuine left group action (Check!). Here is a similar action of S_k on the kth tensor power.

Corollary 5.18. *For* $\sigma \in S_k$, *there is a linear map* $P_\sigma \colon V^{\otimes k} \to V^{\otimes k}$ *such that*

$$v_1 \otimes \cdots \otimes v_k \mapsto v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}$$

on elementary tensors. Then $P_\sigma \circ P_\tau = P_{\sigma\tau}$ *for* σ *and* τ *in* S_k.[10]

When $\dim_K(V) > 1$, $P_\sigma = P_\tau$ *if and only if* $\sigma = \tau$. *In particular,* P_σ *is the identity map if and only if* σ *is the identity permutation.*

[10] If we used $v_{\sigma(i)}$ instead of $v_{\sigma^{-1}(i)}$ in the definition of P_σ then we'd have $P_\sigma \circ P_\tau = P_{\tau\sigma}$.

Proof. The function $V \times \cdots \times V \to V^{\otimes k}$ given by

$$(v_1, \ldots, v_k) \mapsto v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}$$

is multilinear, so the universal mapping property of tensor products gives us a linear map P_σ with the indicated effect on elementary tensors. It is clear that P_1 is the identity map. For any σ and τ in S_k, $P_\sigma \circ P_\tau = P_{\sigma\tau}$ by checking equality of both sides at all elementary tensors in $V^{\otimes k}$. Therefore the injectivity of $\sigma \mapsto P_\sigma$ is reduced to showing if P_σ is the identity map on $V^{\otimes k}$ then σ is the identity permutation.

We prove the contrapositive. Suppose σ is not the identity permutation, so $\sigma(i) = j \neq i$ for some i and j. Choose $v_1, \ldots, v_k \in V$ all nonzero such that v_i and v_j are not on the same line. (Here we use $\dim_K V > 1$.) If $P_\sigma(v_1 \otimes \cdots \otimes v_k) = v_1 \otimes \cdots \otimes v_k$ then $v_j \in K v_i$ by Theorem 5.15, which is not so. $\square$

The linear maps P_σ provide an action of S_k on $V^{\otimes k}$ by linear transformations. We usually write $\sigma(t)$ for $P_\sigma(t)$. Not only does S_k acts on $V^{\otimes k}$ but also the group $\mathrm{GL}(V)$ acts on $V^{\otimes k}$ via tensor powers of linear maps:

$$g(v_1 \otimes \cdots \otimes v_k) := g^{\otimes k}(v_1 \otimes \cdots \otimes v_k) = gv_1 \otimes \cdots \otimes gv_k$$

on elementary tensors. These actions of the groups S_k and $\mathrm{GL}(V)$ on $V^{\otimes k}$ commute with each other: $P_\sigma(g(t)) = g(P_\sigma(t))$ for all $t \in V^{\otimes k}$. To verify that, since both sides are additive in t, it suffices to check it on elementary tensors, which is left to the reader. This commuting action of S_k and $\mathrm{GL}(V)$ on $V^{\otimes k}$ leads to Schur–Weyl duality in representation theory.

A tensor $t \in V^{\otimes k}$ satisfying $P_\sigma(t) = t$ for all $\sigma \in S_k$ are called *symmetric*, and if $P_\sigma(t) = (\operatorname{sign}\sigma)t$ for all $\sigma \in S_k$ we call t *anti-symmetric* or *skew-symmetric*. An example of a symmetric tensor in $V^{\otimes k}$ is $\sum_{\sigma \in S_k} P_\sigma(t)$ for any $t \in V^{\otimes k}$. For elementary t, this sum is

$$\sum_{\sigma \in S_k} P_\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = \sum_{\sigma \in S_k} v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(k)} = \sum_{\sigma \in S_k} v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(k)}.$$

An example of an anti-symmetric tensor in $V^{\otimes k}$ is $\sum_{\sigma \in S_k} (\operatorname{sign}\sigma) v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(k)}$.

Both symmetric and anti-symmetric tensors occur in physics. See Table 1.

Area	Name of Tensor	Symmetry
Mechanics	Stress	Symmetric
	Strain	Symmetric
	Elasticity	Symmetric
	Moment of Inertia	Symmetric
Electromagnetism	Electromagnetic	Anti-symmetric
	Polarization	Symmetric
Relativity	Metric	Symmetric
	Stress-Energy	Symmetric

TABLE 1. Some tensors in physics.

The set of all symmetric tensors and the set of all anti-symmetric tensors in $V^{\otimes k}$ each form subspaces. If K does not have characteristic 2, every tensor in $V^{\otimes 2}$ is a unique sum of a symmetric and anti-symmetric tensor:

$$t = \frac{t + P_{(12)}(t)}{2} + \frac{t - P_{(12)}(t)}{2}.$$

(The map $P_{(12)}$ is the flip automorphism of $V^{\otimes 2}$ sending $v \otimes w$ to $w \otimes v$.) Concretely, if $\{e_1, \dots, e_n\}$ is a basis of V then each tensor in $V^{\otimes 2}$ has the form $\sum_{i,j} c_{ij} e_i \otimes e_j$, and it's symmetric if and only if $c_{ij} = c_{ji}$ for all i and j, while it's anti-symmetric if and only if $c_{ij} = -c_{ji}$ for all i and j (in particular, $c_{ii} = 0$ for all i).

For $k > 2$, the symmetric and anti-symmetric tensors in $V^{\otimes k}$ do not span the whole space. There are *additional* subspaces of tensors in $V^{\otimes k}$ connected to the representation theory of the group $\mathrm{GL}(V)$. The appearance of representations of $\mathrm{GL}(V)$ inside tensor powers of V is an important role for tensor powers in algebra.

6. Tensor Contraction

Continuing the theme of the previous section, let V be a finite-dimensional vector space over a field K. The evaluation pairing $V \times V^\vee \to K$, where $(v, \varphi) \mapsto \varphi(v)$, is K-bilinear and thus induces a K-linear map $c \colon V \otimes V^\vee \to K$ called a *contraction*[11], where $c(v \otimes \varphi) = \varphi(v)$. This map is independent of the choice of a basis, but it's worth seeing how this looks in a basis. We adopt the notation of physicists and geometers for bases and coefficients in V and $V^\vee$, as in the section about tensors in physics in the first handout about tensor products: for a basis $\{e_1, \dots, e_n\}$ of V, let $\{e^1, \dots, e^n\}$ be its dual basis in $V^\vee$, with elements of V written as $\sum_i v^i e_i$ and elements of $V^\vee$ written as $\sum_i v_i e^i$. (Here $v^i, v_i \in K$.) The space $V \otimes V^\vee$ has basis $\{e_i \otimes e^j\}_{i,j}$ and $c(e_i \otimes e^j) = e^j(e_i) = \delta_{ij}$. For $\mathbf{v} \in V$ and $\widetilde{\mathbf{w}} \in V^\vee$, write $\mathbf{v} = \sum_i v^i e_i$ and $\widetilde{\mathbf{w}} = \sum_j w_j e^j$, where $v^i, w_j \in K$. Then

$$c(\mathbf{v} \otimes \widetilde{\mathbf{w}}) = \sum_{i,j} v^i w_j e^j(e_i) = \sum_i v^i w_i. \tag{6.1}$$

The final sum is independent of the basis used on V, since contraction comes from evaluating $V^\vee$ on V and that does not depend on a choice of basis.

From the first tensor product handout, $V \otimes V^\vee \cong \mathrm{Hom}_K(V, V)$ by

$$v \otimes \varphi \mapsto [w \mapsto \varphi(w)v]$$

on elementary tensors. This isomorphism converts contraction on $V \otimes V^\vee$ into a linear map $\mathrm{Hom}_K(V, V) \to K$ that turns out to be the *trace*. Why? A tensor in $V \otimes V^\vee$ can be written uniquely as a linear combination $\sum_{i,j} T^i_j e_i \otimes e^j$. The isomorphism $V \otimes V^\vee \cong \mathrm{Hom}_K(V, V)$ sends $\sum_{i,j} T^i_j e_i \otimes e^j$ to the linear operator $V \to V$ that maps each e_k to $\sum_{i,j} T^i_j e^j(e_k) e_i = \sum_{i,j} T^i_j \delta_{jk} e_i = \sum_i T^i_k e_i$, so this linear operator $V \to V$ has matrix representation (T^i_j) with respect to the basis $\{e_i\}$. The contraction of $\sum_{i,j} T^i_j e_i \otimes e^j$ is $c(\sum_{i,j} T^i_j e_i \otimes e^j) = \sum_{i,j} T^i_j \delta_{ij} = \sum_i T^i_i$, which is exactly the trace of the matrix (T^i_j).

Contraction $V \otimes V^\vee \to K$ generalizes to maps $V^{\otimes k} \otimes (V^\vee)^{\otimes \ell} \to V^{\otimes (k-1)} \otimes (V^\vee)^{\otimes(\ell-1)}$ for $k \geq 1$ and $\ell \geq 1$ (recall $V^{\otimes 0} = K$ and $(V^\vee)^{\otimes 0} = K$) that are linear and also called contractions, as follows. For $r \in \{1, \dots, k\}$ and $s \in \{1, \dots, \ell\}$, the associated contraction $c_{r,s} \colon V^{\otimes k} \otimes (V^\vee)^{\otimes \ell} \to V^{\otimes(k-1)} \otimes (V^\vee)^{\otimes(\ell-1)}$ evaluates the rth tensorand in $V^{\otimes k}$ in the sth

[11]Here and elsewhere in this section, tensor products are always over K: $\otimes = \otimes_K$.

tensorand in $(V^\vee)^{\otimes \ell}$, leaving other tensorands untouched: on elementary tensors,

$$\begin{aligned} c_{r,s}(v_1 \otimes \cdots \otimes v_k \otimes \varphi_1 \otimes \cdots \otimes \varphi_\ell) &= c(v_r \otimes \varphi_s) \bigotimes_{i \neq r} v_i \otimes \bigotimes_{j \neq s} \varphi_j \\ &= \varphi_s(v_r) \bigotimes_{i \neq r} v_i \otimes \bigotimes_{j \neq s} \varphi_j. \end{aligned}$$

How does $c_{r,s}$ look in a basis? For a basis $\{e_1, \ldots, e_n\}$ of V and dual basis $\{e^1, \ldots, e^n\}$ of $V^\vee$, a basis of $V^{\otimes k} \otimes (V^\vee)^{\otimes \ell}$ is the $n^{k+\ell}$ elementary tensors $e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes e^{j_1} \otimes \cdots \otimes e^{j_\ell}$, and

$$c_{r,s}\left(\sum_{\substack{i_1,\ldots,i_k \\ j_1,\ldots,j_\ell}} T^{i_1 \ldots i_k}_{j_1 \ldots j_\ell} e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes e^{j_1} \otimes \cdots \otimes e^{j_\ell}\right) = \sum_{\substack{i_1,\ldots,j_\ell \\ i_r, j_s \text{missing}}} \sum_{m=1}^{n} T^{i_1 \cdots m \cdots i_k}_{j_1 \cdots m \cdots j_\ell} \bigotimes_{a \neq r} e_{i_a} \otimes \bigotimes_{b \neq s} e^{j_b},$$

where the m appears in the rth slot of the superscript and the sth slot of the subscript of the coefficient on the right. The outer sum on the right initially runs over all sets of indices and the coefficient is multiplied by the numerical factor $e^{i_s}(e_{i_r})$, which is nonzero only for $i_s = i_r$ (call this common value m, and its runs from 1 to n), when the factor is 1.

The contraction $c_{r,s}$ depends on its domain $V^{\otimes k} \otimes (V^\vee)^{\otimes \ell}$, not just r and s (which are at most k and ℓ, respectively). For example, on $V^{\otimes 2} \otimes V^\vee$ and $V^{\otimes 2} \otimes (V^\vee)^{\otimes 2}$, the contraction $c_{2,1}$ on elementary tensors is $v \otimes w \otimes \varphi \mapsto \varphi(w)v$ and $v \otimes w \otimes \varphi \otimes \psi \mapsto \varphi(w)v \otimes \psi$. Using bases, the contraction $c_{2,1}$ on $V^{\otimes 2} \otimes V^\vee$ is

$$c_{2,1}\left(\sum_{\substack{i_1,i_2 \\ j_1}} T^{i_1 i_2}_{j_1} e_{i_1} \otimes e_{i_2} \otimes e^{j_1}\right) = \sum_{\substack{i_1,i_2 \\ j_1}} T^{i_1 i_2}_{j_1} e^{j_1}(e_{i_2})\, e_{i_1} = \sum_{i_1} \left(\sum_m T^{i_1 m}_{m}\right) e_{i_1} \tag{6.2}$$

and the contraction $c_{2,1}$ on $V^{\otimes 2} \otimes (V^\vee)^{\otimes 2}$ is

$$c_{2,1}\left(\sum_{\substack{i_1,i_2 \\ j_1,j_2}} T^{i_1 i_2}_{j_1 j_2} e_{i_1} \otimes e_{i_2} \otimes e^{j_1} \otimes e^{j_2}\right) = \sum_{i_1,j_2} \left(\sum_m T^{i_1 m}_{m j_2}\right) e_{i_1} \otimes e^{j_2}. \tag{6.3}$$

Remark 6.1. In physics and engineering, tensors are often described just by components (coordinates), with summing over the basis being understood: $\sum_i T^i e_i$ is written as T^i and $\sum_{i,j} T_{ij} e^i \otimes e^j$ is written as T_{ij}. That components have indices in the opposite position (up vs. down) to basis vectors needs to be known to reconstruct the tensor from how its components are written. Another convention, which is widely used in geometry as well, is that summation signs within a component (due to contraction or to applying the metric – see below) are omitted and an implicit summation running from 1 to the dimension of V is intended for each index repeated as both a superscript and subscript. This is called the *Einstein summation convention*. For example, the most basic contraction $V \otimes V^\vee \to K$, where $\sum_{i,j} T^i_j e_i \otimes e^j \mapsto \sum_i T^i_i$, is written in this convention as

$$T^i_j \mapsto T^i_i.$$

The contractions $c_{2,1}$ in (6.2) and (6.3), on $V^{\otimes 2} \otimes V^\vee$ and $V^{\otimes 2} \otimes (V^\vee)^{\otimes 2}$ respectively, are denoted in the Einstein summation convention as

$$T^{i_1 i_2}_{j_1} \mapsto T^{i_1 i_2}_{i_2}, \qquad T^{i_1 i_2}_{j_1 j_2} \mapsto T^{i_1 i_2}_{i_2 j_2}.$$

Contraction only makes sense when we combine V and $V^\vee$. We can't contract V and V – it doesn't make sense to evaluate elements of V on V. However, if we have a *preferred* isomorphism $g\colon V \to V^\vee$ of vector spaces then we can use g to turn V into $V^\vee$. We have $V \otimes V \cong V \otimes V^\vee$ using $\mathrm{id.} \otimes g$ and we can contract the right side. So we *can* contract on $V \otimes V$ after all, provided we use an isomorphism g from V to $V^\vee$ and *remember what g is.* Contraction on $V \otimes V$ (using g) means the composite map

$$V \otimes V \xrightarrow{\mathrm{id.}\otimes g} V \otimes V^\vee \xrightarrow{c} K. \tag{6.4}$$

This depends on the choice of g but not on a choice of basis of V.

There should be nothing special about using g in the *second* tensorand in (6.4). We want

$$V \otimes V \xrightarrow{g\otimes \mathrm{id.}} V^\vee \otimes V \longrightarrow K \tag{6.5}$$

to be the same overall mapping as (6.4), where the second mapping in (6.5) is induced by evaluation of $V^\vee$ on V. The agreement of (6.4) and (6.5) means $g(\mathbf{v})(\mathbf{w}) = g(\mathbf{w})(\mathbf{v})$ for all $\mathbf{v}$ and $\mathbf{w}$ in V. This means that if we regard $g\colon V \to V^\vee$ as a bilinear map $V \times V \to K$ by $(\mathbf{v}, \mathbf{w}) \mapsto g(\mathbf{v})(\mathbf{w})$ or $(\mathbf{v}, \mathbf{w}) \mapsto g(\mathbf{w})(\mathbf{v})$ we want the same result: we want g to be *symmetric* as a bilinear form on V. To check a given g is symmetric, it enough to check this on a basis $\{e_1, \ldots, e_n\}$ of V: $g(e_i)(e_j) = g(e_j)(e_i)$ in K for all i and j. Set $g_{ij} = g(e_i)(e_j)$, so $\boxed{g(e_i) = \sum_j g_{ij} e^j}$ for all i and $g_{ij} = g_{ji}$ for all i, j from 1 to n: the matrix (g_{ij}) is symmetric.

We can think of the isomorphism $g\colon V \to V^\vee$ as a bilinear map $V \times V \to K$ in two ways: either as $(\mathbf{v}, \mathbf{w}) \mapsto g(\mathbf{v})(\mathbf{w})$ or as $(\mathbf{v}, \mathbf{w}) \mapsto g(\mathbf{w})(\mathbf{v})$. The fact that g is an isomorphism is equivalent to g as a bilinear form on V being nondegenerate and is also equivalent to the matrix (g_{ij}) coming from any basis of V ($g_{ij} = g(e_i)(e_j)$) being invertible. In practice we always assume g is symmetric and nondegenerate as a bilinear form on V.

Let's write g and then the contraction on $V \otimes V$ from (6.4) or (6.5) in coordinates. Pick a basis $\{e_1, \ldots, e_n\}$ of V and dual basis $\{e^1, \ldots, e^n\}$ of $V^\vee$. For $\mathbf{v} = \sum_i v^i e_i$ in V, how does $g(\mathbf{v})$ look in the dual basis of $V^\vee$? That $g_{ij} = g(e_i)(e_j)$ means $g(e_i) = \sum_j g_{ij} e^j$, so

$$g(\mathbf{v}) = g\left(\sum_i v^i e_i\right) = \sum_i v^i g(e_i) = \sum_i v^i \left(\sum_j g_{ij} e^j\right) = \sum_j \left(\sum_i g_{ij} v^i\right) e^j.$$

Thus $g(\sum_i v^i e_i) = \sum_j v_j e^j$, where $\boxed{v_j = \sum_i g_{ij} v^i = \sum_i g_{ji} v^i}$ for all j (recall $g_{ij} = g_{ji}$). The passage from the numbers $\{v^i\}$ (components of $\mathbf{v}$) to the numbers $\{v_j\}$ (components of $g(\mathbf{v})$) by multiplying each v^i by g_{ij} and summing over all i is called *lowering an index*. It is the coordinate version (using a basis of V and its dual basis in $V^\vee$) of going from $\mathbf{v}$ in V to $g(\mathbf{v})$ in $V^\vee$, which depends on g but not on the choice of basis of V.

For $\mathbf{x} = \sum_i x^i e_i$ and $\mathbf{y} = \sum_i y^i e_i$ in $V^{\otimes 2}$, (6.4) has the following effect on the elementary tensor $\mathbf{x} \otimes \mathbf{y}$:

$$\begin{aligned}\mathbf{x} \otimes \mathbf{y} &\mapsto \mathbf{x} \otimes g(\mathbf{y}) \\ &= \left(\sum_i x^i e_i\right) \otimes \left(\sum_j y_j e^j\right) \quad \text{where } y_j = \sum_k g_{kj} y^k = \sum_k g_{jk} y^k \\ &\overset{c}{\mapsto} \sum_i x^i y_i \text{ by (6.1).}\end{aligned}$$

Unwrapping the definition of y_i terms of g_{ij} and y^j (for $j = 1, \dots, n$),

$$\sum_i x^i y_i = \sum_{i,j} g_{ij} x^i y^j.$$

This is usually *not* $\sum_i x^i y^i$ unless $g_{ij} = \delta_{ij}$ for all i and j. More generally, the contraction of the tensor $\mathsf{T} = \sum_{i,j} T^{ij} e_i \otimes e_j$ in $V^{\otimes 2}$ using g is the scalar $\sum_{i,j} g_{ij} T^{ij}$ in K. This contraction is independent of the basis $\{e_1, \dots, e_n\}$ of V but depends on the isomorphism $g\colon V \to V^\vee$.

Example 6.2. Let g be an isomorphism $V \to V^\vee$ that is symmetric (and nondegenerate) as a bilinear form $V \times V \to K$. Applying g to the sth tensorand of $V^{\otimes k}$, where $k \geq 2$ and $1 \leq s \leq k$, turns the tensor $\sum_{i_1,\dots,i_k} T^{i_1 \dots i_k} e_{i_1} \otimes \cdots \otimes e_{i_k}$ into

$$\sum_{i_1,\dots,i_k} T^{i_1\dots i_k} e_{i_1} \otimes \cdots \otimes g(e_{i_s}) \otimes \cdots \otimes e_{i_k} = \sum_{i_1,\dots,i_k} T^{i_1\dots i_k} e_{i_1} \otimes \cdots \otimes \underbrace{\left(\sum_i g_{i i_s} e^i\right)}_{s\text{th tensorand}} \otimes \cdots \otimes e_{i_k}$$

and contracting the rth and sth tensorands[12] where $r \neq s$ gives us the following tensor in $V^{\otimes(k-2)}$:

$$\sum_{\substack{i_1,\dots,i_k \\ i_r, i_s \text{ missing}}} \left(\sum_{i_r, i_s} g_{i_r i_s} T^{i_1 \dots i_k}\right) \underbrace{e_{i_1} \otimes \cdots \otimes e_{i_k}}_{i_r, i_s \text{ missing}}.$$

Example 6.3. Let g be an isomorphism $V \to V^\vee$ that is symmetric (and nondegenerate) as a bilinear form $V \times V \to K$. For a rank 3 tensor T in $V^{\otimes 2} \otimes V^\vee$, we have the rank 3 tensors $\mathsf{U} = (g \otimes 1 \otimes 1)(\mathsf{T}) \in V^\vee \otimes V \otimes V^\vee$ (not standard form with all V's first) and $\mathsf{U}' = (g \otimes g \otimes 1)(\mathsf{T}) \in (V^\vee)^{\otimes 3}$. What do U and U' look like in terms of how T looks?

Using a basis $\{e_1, \dots, e_n\}$ for V and dual basis $\{e^1, \dots, e^n\}$ for $V^\vee$, in $V^{\otimes 2} \otimes V^\vee$ let

$$\mathsf{T} = \sum_{\substack{i_1, i_2 \\ j_1}} T^{i_1 i_2}_{j_1} e_{i_1} \otimes e_{i_2} \otimes e^{j_1}.$$

Passing from $V^{\otimes 2} \otimes V^\vee = V \otimes V \otimes V^\vee$ to $V^\vee \otimes V \otimes V^\vee$ by $g \otimes 1 \otimes 1$ only affects the first tensorand:

$$\mathsf{U} = (g \otimes 1 \otimes 1)(\mathsf{T}) = \sum_{\substack{i_2 \\ j_1, j_2}} U^{i_2}_{j_2 j_1} e^{j_2} \otimes e_{i_2} \otimes e^{j_1}, \quad \text{where } U^{i_2}_{j_2 j_1} = \sum_{i_1} g_{i_1 j_2} T^{i_1 i_2}_{j_1}.$$

[12]While $V^{\otimes(s-1)} \otimes (V^\vee) \otimes V^{\otimes(k-s)}$ is not in standard form, the meaning of contraction using the rth tensorand V and the sth tensorand $V^\vee$ should be obvious to the reader.

Passing from $V^{\otimes 2}\otimes V^\vee = V\otimes V\otimes V^\vee$ to $(V^\vee)^{\otimes 3}$ by $g\otimes g\otimes 1$ affects the first two tensorands:

$$\mathsf{U}' = (g\otimes g\otimes 1)(\mathsf{T}) = \sum_{j_1,j_2,j_3} U_{j_2j_3j_1}e^{j_2}\otimes e^{j_3}\otimes e^{j_1}, \quad \text{where } U_{j_2j_3j_1} = \sum_{i_1,i_2} g_{i_1j_2}g_{i_2j_3}T^{i_1i_2}_{j_1}.$$

Remark 6.4. Putting the *Einstein summation convention* (Remark 6.1) to work, the contraction $V\otimes V^\vee \to K$ sends $v^ie_i\otimes w_je^j$ to v^iw_i, which is the same as v^jw_j and is independent of the basis since it is simply evaluating $V^\vee$ on V. Using an isomorphism $g\colon V\to V^\vee$ that is symmetric (and nondegenerate) as a bilinear form $V\times V\to K$, any $\mathbf{v} = v^ie_i$ in V turns into $g(\mathbf{v}) = v_ie^i$ in $V^\vee$ where $v_i = g_{ki}v^k = g_{ik}v^k$ since (g_{ij}) is symmetric, and the contraction $V\otimes V\to K$ that depends on g has the effect $x^ie_i\otimes y^je_j\mapsto x^iy_i = x_jy^j = g_{ij}x^iy^j$ (remember $g_{ij} = g_{ji}$!) and more generally $T^{ij}e_i\otimes e_j\mapsto g_{ij}T^{ij}$. The contraction on $V^{\otimes 2}$ depends on g but not on the basis of V. Contraction on $V^{\otimes k}$, for $k\geq 2$, using the rth and sth tensorand is a linear map to $V^{\otimes(k-2)}$ with the effect $T^{i_1\dots i_k}e_{i_1}\otimes\cdots\otimes e_{i_k}\mapsto g_{i_ri_s}T^{i_1\cdots i_k}\underbrace{e_{i_1}\otimes\cdots\otimes e_{i_k}}_{i_r,i_s\text{ missing}}$.
For $T^{i_1i_2}_{j_1}$ in $V\otimes V\otimes V^\vee$, if we lower the first upper index we get $g_{i_1j_2}T^{i_1i_2}_{j_1}$ in $V^\vee\otimes V\otimes V^\vee$ and if we lower both upper indices then we get $g_{i_1j_2}g_{i_2j_3}T^{i_1i_2}_{j_1}$ in $V^\vee\otimes V^\vee\otimes V^\vee$.

Contraction lets us map $V^{\otimes k}\otimes(V^\vee)^{\otimes\ell}$ to $V^{\otimes(k-1)}\otimes(V^\vee)^{\otimes(\ell-1)}$ by combining a choice of V and $V^\vee$ in $V^{\otimes k}\otimes(V^\vee)^{\otimes\ell}$ using evaluation to get scalars. We can use an isomorphism $g\colon V\to V^\vee$ to change any V in $V^{\otimes k}\otimes(V^\vee)^{\otimes\ell}$ into $V^\vee$ (in practice g is symmetric when viewed as a bilinear form) and thus contract any two different V tensorands (the result depends on g). To contract two different $V^\vee$ tensorands, use the inverse $g^{-1}\colon V^\vee\to V$, which as a bilinear form on $V^\vee$ is symmetric: writing $\boxed{g^{-1}(e^i) = \sum_j g^{ij}e_j}$, (g^{ij}) is the inverse matrix to (g_{ij}) and is symmetric since (g_{ij}) is. We have $g^{-1}(\sum_i v_ie^i) = \sum_j v^je_j$, where $\boxed{v^j = \sum_i g^{ij}v_i}$. The passage from $\{v_i\}$ to $\{v^j\}$ is multiplying v_i by g^{ij} and summing over all i. It is called *raising an index* and depends on g: this is the coordinate version (using a basis of V and its dual basis in $V^\vee$) of going from elements of $V^\vee$ to elements of V by g^{-1}.[13] Using g or g^{-1} in enough places lets us turn the mixed tensors of $V^{\otimes k}\otimes(V^\vee)^{\otimes\ell}$ into pure tensors in $V^{\otimes(k+\ell)}$ or $(V^\vee)^{\otimes(k+\ell)}$.

The operations of raising/lowering an index and contraction do not depend on a choice of basis of V. That is because raising/lowering an index is just applying a choice of isomorphism from V to $V^\vee$ or *vice versa* and contraction is based on the natural bilnear evaluation map $V^\vee\times V\to K$ thought of as a linear map $V^\vee\otimes V\to K$ or $V\otimes V^\vee\to K$, and *none of these depend on bases*. For physics or engineering students who don't know about dual spaces, the independence of basis for these operations is verified by tedious calculations (if they care at all) and they check raising/lowering indices and contractions send tensors to tensors by checking the output of such operations satisfies tensor transformation rules, so it is a tensor. At least that it is how books for such students handle the tasks and it might look like a miracle to those who can't think about concepts without bases: what do raising/lowering an index and multiplying by g_{ij} mean if you don't know what $V^\vee$ is?

[13]Raising or lowering indices describes what happens to *components* in a basis. The basis undergoes the opposite change (lowering or raising its indices).

When $K = \mathbf{R}$, a nondegenerate symmetric bilinear form g on V is called a *metric* on V. This generalizes the dot product on $\mathbf{R}^n$, so intuitively it is more like squared distance than distance itself. (A metric in the sense of metric spaces is always nonnnegative, but g as a bilinear form might be positive and negative.) There is an integer p from 0 to n and a basis $\{e_1, \ldots, e_n\}$ of V in which $g(\sum_i x^i e_i, \sum_j y^j e_j) = \sum_{i=1}^p x^i y^i - \sum_{i=p+1}^n x^i y^i$. For $p = 0$ or n one sum on the right is empty and that's treated as 0. Here are two important special cases.

- The case $p = n$ means $g(\mathbf{v}, \mathbf{v}) \geq 0$ with equality if and only if $\mathbf{v} = \mathbf{0}$. Such g are called positive-definite or inner products. They are used in Riemannian geometry.
- The cases $p = 1$ or $p = n-1$ are important in Lorentzian geometry, and in relativity when $n = 4$.

7. Tensor Product of R-Algebras

Our tensor product isomorphisms of modules often involve rings, *e.g.*, $\mathbf{C} \otimes_{\mathbf{R}} \mathrm{M}_n(\mathbf{R}) \cong \mathrm{M}_n(\mathbf{C})$ as $\mathbf{C}$-vector spaces (Example 4.7). Now we will show how to turn the tensor product of two rings into a ring. Then we will revisit a number of previous module isomorphisms where the modules are also rings and find that the isomorphism holds at the level of rings.

Because we want to be able to say $\mathbf{C} \otimes_{\mathbf{R}} \mathrm{M}_n(\mathbf{R}) \cong \mathrm{M}_n(\mathbf{C})$ as rings, not just as vector spaces (over $\mathbf{R}$ or $\mathbf{C}$), and matrix rings are noncommutative, we are going to allow our R-modules to be possibly noncommutative rings. But R itself remains commutative!

Our rings will all be R-algebras. An *R-algebra* is an R-module A equipped with an R-bilinear map $A \times A \to A$, called multiplication or product. Bilinearity of multiplication includes distributive laws for multiplication over addition as well as the extra rule

$$r(ab) = (ra)b = a(rb) \tag{7.1}$$

for $r \in R$ and a and b in B, which says R-scaling commutes with multiplication in the R-algebra. We also want $1 \cdot a = a$ for $a \in A$, where 1 is the identity element of R.

Examples of R-algebras include the matrix ring $\mathrm{M}_n(R)$, a quotient ring R/I, and the polynomial ring $R[X_1, \ldots, X_n]$. We will assume, except in Example 7.5, that our R-algebras have associative multiplication and a multiplicative identity, so they are genuinely rings (perhaps not commutative) and being an R-algebra just means they have a little extra structure related to scaling by R. When an R-algebra contains R, (7.1) is a special case of associative multiplication in the algebra.

The difference between an R-algebra and a ring is exactly like that between an R-module and an abelian group. An R-algebra is a ring on which we have a scaling operation by R that behaves nicely with respect to the addition and multiplication in the R-algebra, in the same way that an R-module is an abelian group on which we have a scaling operation by R that behaves nicely with respect to the addition in the R-module. While $\mathbf{Z}$-modules are nothing other than abelian groups, $\mathbf{Z}$-algebras in our lexicon are nothing other than rings (possibly noncommutative).

Because of the universal mapping property of the tensor product, to give an R-bilinear multiplication $A \times A \to A$ in an R-algebra A is the *same thing* as giving an R-linear map $A \otimes_R A \to A$. So we could define an R-algebra as an R-module A equipped with an R-linear map $A \otimes_R A \xrightarrow{m} A$, and declare the product of a and b in A to be $ab := m(a \otimes b)$.

Associativity of multiplication can be formulated in tensor language: the diagram

$$\begin{tikzcd} A \otimes_R A \otimes_R A \arrow[r, "1\otimes m"] \arrow[d, "m\otimes 1"'] & A \otimes_R A \arrow[d, "m"] \\ A \otimes_R A \arrow[r, "m"] & A \end{tikzcd}$$

commutes.

Theorem 7.1. *Let A and B be R-algebras. There is a unique multiplication on $A \otimes_R B$ making it an R-algebra such that*

$$(a \otimes b)(a' \otimes b') = aa' \otimes bb' \tag{7.2}$$

for all elementary tensors. The multiplicative identity is $1 \otimes 1$.

Proof. If there is an R-algebra multiplication on $A \otimes_R B$ satisfying (7.2) then multiplication between any two tensors is determined:

$$\sum_{i=1}^{k} a_i \otimes b_i \cdot \sum_{j=1}^{\ell} a'_j \otimes b'_j = \sum_{i,j} (a_i \otimes b_i)(a'_j \otimes b'_j) = \sum_{i,j} a_i a'_j \otimes b_i b'_j.$$

So the R-algebra multiplication on $A \otimes_R B$ satisfying (7.2) is unique if it exists at all. Our task now is to write down a multiplication on $A \otimes_R B$ satisfying (7.2).

One way to do this is to define what left multiplication by each elementary tensor $a \otimes b$ on $A \otimes_R B$ should be, by introducing a suitable bilinear map and making it into a linear map. But rather than proceed by this route, we'll take advantage of various maps we already know between tensor products. Writing down an associative R-bilinear multiplication on $A \otimes_R B$ with identity $1 \otimes 1$ means writing down an R-linear map $(A \otimes_R B) \otimes_R (A \otimes_R B) \to A \otimes_R B$ satisfying certain conditions, and that's what we're going to do.

Let $A \otimes_R A \xrightarrow{m_A} A$ and $B \otimes_R B \xrightarrow{m_B} B$ be the R-linear maps corresponding to multiplication on A and on B. Their tensor product $m_A \otimes m_B$ is an R-linear map from $(A \otimes_R A) \otimes_R (B \otimes_R B)$ to $A \otimes_R B$. Using the commutativity and associativity isomorphisms on tensor products, there are natural isomorphisms

$$\begin{aligned} (A \otimes_R B) \otimes_R (A \otimes_R B) &\cong ((A \otimes_R B) \otimes_R A) \otimes_R B \\ &\cong (A \otimes_R (B \otimes_R A)) \otimes_R B \\ &\cong (A \otimes_R (A \otimes_R B)) \otimes_R B \\ &\cong ((A \otimes_R A) \otimes_R B) \otimes_R B \\ &\cong (A \otimes_R A) \otimes_R (B \otimes_R B). \end{aligned}$$

Tracking the effect of these maps on $(a \otimes b) \otimes (a' \otimes b')$,

$$\begin{aligned} (a \otimes b) \otimes (a' \otimes b') &\mapsto ((a \otimes b) \otimes a') \otimes b' \\ &\mapsto (a \otimes (b \otimes a')) \otimes b' \\ &\mapsto (a \otimes (a' \otimes b)) \otimes b' \\ &\mapsto ((a \otimes a') \otimes b) \otimes b' \\ &\mapsto (a \otimes a') \otimes (b \otimes b'). \end{aligned}$$

Composing these isomorphisms with $m_A \otimes m_B$ makes a map $(A \otimes_R B) \otimes_R (A \otimes_R B) \to A \otimes_R B$ that is R-linear and has the effect

$$(a \otimes b) \otimes (a' \otimes b') \mapsto (a \otimes a') \otimes (b \otimes b') \mapsto aa' \otimes bb',$$

where $m_A \otimes m_B$ is used in the second step. This R-linear map $(A \otimes_R B) \otimes_R (A \otimes_R B) \to A \otimes_R B$ pulls back to an R-bilinear map $(A \otimes_R B) \times (A \otimes_R B) \to A \otimes_R B$ with the effect $(a \otimes b, a' \otimes b') \mapsto aa' \otimes bb'$ on pairs of elementary tensors, which is what we wanted for our multiplication on $A \otimes_R B$. This proves $A \otimes_R B$ has a multiplication satisfying (7.2).

To prove $1 \otimes 1$ is an identity and multiplication in $A \otimes_R B$ is associative, we want

$$(1 \otimes 1)t = t, \quad t(1 \otimes 1) = t, \quad (t_1 t_2) t_3 = t_1 (t_2 t_3)$$

for general tensors t, t_1, t_2, and t_3 in $A \otimes_R B$. These identities are additive in each tensor appearing on both sides, so verifying these equations reduces to the case that the tensors are all elementary, and this case is left to the reader. $\square$

Corollary 7.2. *If A and B are commutative R-algebras then $A \otimes_R B$ is a commutative R-algebra.*

Proof. We want to check $tt' = t't$ for all t and t' in $A \otimes_R B$. Both sides are additive in t, so it suffices to check the equation when $t = a \otimes b$ is an elementary tensor: $(a \otimes b)t' \stackrel{?}{=} t'(a \otimes b)$. Both sides of this are additive in t', so we are reduced further to the special case when $t' = a' \otimes b'$ is also an elementary tensor: $(a \otimes b)(a' \otimes b') \stackrel{?}{=} (a' \otimes b')(a \otimes b)$. The validity of this is immediate from (7.2) since A and B are commutative. $\square$

Example 7.3. Let's look at the ring $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C}$. It is 4-dimensional as a real vector space, and its multiplication is determined from linearity by the products of its standard basis $1 \otimes 1$, $1 \otimes i$, $i \otimes 1$, and $i \otimes i$. The tensor $1 \otimes 1$ is the multiplicative identity, so we'll look at the products of the three other basis elements, and since multiplication is commutative we only need one product per basis pair:

$$(1 \otimes i)^2 = 1 \otimes (-1) = -(1 \otimes 1), \quad (1 \otimes i)(i \otimes 1) = i \otimes i, \quad (1 \otimes i)(i \otimes i) = i \otimes (-1) = -(i \otimes 1),$$

$$(i \otimes 1)^2 = (-1) \otimes 1 = -(1 \otimes 1), \quad (i \otimes 1)(i \otimes i) = (-1) \otimes i = -(1 \otimes i), \quad (i \otimes i)^2 = 1 \otimes 1.$$

Setting $x = i \otimes 1$ and $y = 1 \otimes i$, we have $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C} = \mathbf{R} + \mathbf{R}x + \mathbf{R}y + \mathbf{R}xy$, where $(\pm x)^2 = -1$ and $(\pm y)^2 = -1$. This commutative ring is not a field (-1 can't have more than two square roots in a field), and in fact it is "clearly" the product ring $(\mathbf{R} + \mathbf{R}x) \times (\mathbf{R} + \mathbf{R}y) \cong \mathbf{C} \times \mathbf{C}$ with componentwise operations. (Warning: an isomorphism $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C} \to \mathbf{C} \times \mathbf{C}$ is not obtained by $z \otimes w \mapsto (z, w)$ since that is not well-defined: $(-z) \otimes (-w) = z \otimes w$ but $(-z, -w) \neq (z, w)$ in general. We'll see an explicit isomorphism in Example 7.19.)

Example 7.4. The tensor product $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ is isomorphic to $\mathbf{R}$ as a $\mathbf{Q}$-vector space for the nonconstructive reason that they have the same (infinite) dimension over $\mathbf{Q}$. When we compare $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ to $\mathbf{R}$ as commutative rings, they look very different: the tensor product ring has zero divisors. The elementary tensors $\sqrt{2} \otimes 1$ and $1 \otimes \sqrt{2}$ are linearly independent over $\mathbf{Q}$ and square to $2(1 \otimes 1)$, so

$$(\sqrt{2} \otimes 1 + 1 \otimes \sqrt{2})(\sqrt{2} \otimes 1 - 1 \otimes \sqrt{2}) = 2 \otimes 1 - 1 \otimes 2 = 0$$

with neither factor on the left side being 0.

Example 7.5. Let $\mathbf{R}^3$ have the cross product $\times$ as multiplication: $\mathbf{vw} := \mathbf{v} \times \mathbf{w}$. This is bilinear with $\mathbf{vv} = \mathbf{v} \times \mathbf{v} = \mathbf{0}$ for all $\mathbf{v}$ in $\mathbf{R}^3$ and it's assumed the reader has seen the products among the different basis vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, *e.g.*, $\mathbf{i} \times \mathbf{j} = \mathbf{k}$ and $\mathbf{i} \times \mathbf{k} = -\mathbf{j}$.

Since the cross product has no identity and is not associative, the proof of Theorem 7.1 except for its last paragraph (which is about multiplicative identities and associativity) works on $\mathbf{R}^3 \otimes_{\mathbf{R}} \mathbf{R}^3$: it admits a unique multiplication such that on elementary tensors

$$(\mathbf{v} \otimes \mathbf{w})(\mathbf{v}' \otimes \mathbf{w}') = \mathbf{vv}' \otimes \mathbf{ww}' = (\mathbf{v} \times \mathbf{v}') \otimes (\mathbf{w} \times \mathbf{w}').$$

The cross product on $\mathbf{R}^3$ satisfies the Jacobi identity: for all $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in $\mathbf{R}^3$,

$$\mathbf{u}(\mathbf{vw}) + \mathbf{v}(\mathbf{wu}) + \mathbf{w}(\mathbf{uv}) = \mathbf{u} \times (\mathbf{v} \times \mathbf{w}) + \mathbf{v} \times (\mathbf{w} \times \mathbf{u}) + \mathbf{w} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{0}.$$

Such an identity does *not* hold for the multiplication on $\mathbf{R}^3 \otimes_{\mathbf{R}} \mathbf{R}^3$: when $t = \mathbf{i} \otimes \mathbf{j}$, $t' = \mathbf{i} \otimes \mathbf{k}$, and $t'' = \mathbf{j} \otimes \mathbf{k}$,

$$\begin{aligned} t(t't'') + t'(t''t) + t''(tt') &= \mathbf{i} \otimes \mathbf{j}((\mathbf{i} \otimes \mathbf{k})(\mathbf{j} \otimes \mathbf{k})) + \mathbf{i} \otimes \mathbf{k}((\mathbf{j} \otimes \mathbf{k})(\mathbf{i} \otimes \mathbf{j})) + \mathbf{j} \otimes \mathbf{k}((\mathbf{i} \otimes \mathbf{j})(\mathbf{i} \otimes \mathbf{k})) \\ &= (\mathbf{i} \otimes \mathbf{j})(\mathbf{ij} \otimes \mathbf{kk}) + (\mathbf{i} \otimes \mathbf{k})(\mathbf{ji} \otimes \mathbf{kj}) + (\mathbf{j} \otimes \mathbf{k})(\mathbf{ii} \otimes \mathbf{jk}) \\ &= (\mathbf{i} \otimes \mathbf{j})(\mathbf{ij} \otimes \mathbf{0}) + (\mathbf{i} \otimes \mathbf{k})(-\mathbf{k} \otimes -\mathbf{i}) + (\mathbf{j} \otimes \mathbf{k})(\mathbf{0} \otimes \mathbf{jk}) \\ &= \mathbf{0} + (\mathbf{i} \otimes \mathbf{k})(\mathbf{k} \otimes \mathbf{i}) + \mathbf{0} \\ &= \mathbf{ik} \otimes \mathbf{ki} \\ &= -\mathbf{j} \otimes \mathbf{j}, \end{aligned}$$

which is not 0. This shows the tensor product $\mathfrak{g} \otimes \mathfrak{g}$ of a Lie algebra $\mathfrak{g}$ with itself may not be a Lie algebra using the multiplication naturally inherited from the Lie bracket on $\mathfrak{g}$.

A *homomorphism* of R-algebras is a function between R-algebras that is both R-linear and a ring homomorphism. An *isomorphism* of R-algebras is a bijective R-algebra homomorphism. That is, an R-algebra isomorphism is simultaneously an R-module isomorphism and a ring isomorphism. For example, the reduction map $R[X] \to R[X]/(X^2 + X + 1)$ is an R-algebra homomorphism (it is R-linear and a ring homomorphism) and $\mathbf{R}[X]/(X^2+1) \cong \mathbf{C}$ as $\mathbf{R}$-algebras by $a + bX \mapsto a + bi$: this function is not just a ring isomorphism, but also $\mathbf{R}$-linear.

For any R-algebras A and B, there is an R-algebra homomorphism $A \to A \otimes_R B$ by $a \mapsto a \otimes 1$ (check!). The image of A in $A \otimes_R B$ might not be isomorphic to A. For instance, in $\mathbf{Z} \otimes_{\mathbf{Z}} (\mathbf{Z}/5\mathbf{Z})$ (which is isomorphic to $\mathbf{Z}/5\mathbf{Z}$ by $a \otimes (b \bmod 5) = ab \bmod 5$), the image of $\mathbf{Z}$ by $a \mapsto a \otimes 1$ is isomorphic to $\mathbf{Z}/5\mathbf{Z}$. There is also an R-algebra homomorphism $B \to A \otimes_R B$ by $b \mapsto 1 \otimes b$. Even when A and B are noncommutative, the images of A and B in $A \otimes_R B$ commute: $(a \otimes 1)(1 \otimes b) = a \otimes b = (1 \otimes b)(a \otimes 1)$. This is like groups G and H commuting in $G \times H$ even if G and H are nonabelian.

It is worth contrasting the direct product $A \times B$ (componentwise addition and multiplication, with $r(a, b) = (ra, rb)$) and the tensor product $A \otimes_R B$, which are both R-algebras. The direct product $A \times B$ is a ring structure on the R-module $A \oplus B$, which is usually quite different from $A \otimes_R B$ as an R-module. There are natural R-algebra homomorphisms $A \times B \xrightarrow{\pi_1} A$ and $A \times B \xrightarrow{\pi_2} B$ by projection, while there are natural R-algebra homomorphisms $A \to A \otimes_R B$ and $B \longrightarrow A \otimes_R B$ in the other direction (out of A and B to the tensor product rather than to A and B from the direct product). The projections out of the direct product $A \times B$ to A and B are both surjective, but the maps to the tensor product $A \otimes_R B$ from A and B need not be injective, *e.g.*, $\mathbf{Z} \to \mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/5\mathbf{Z}$. The maps $A \to A \otimes_R B$ and $B \to A \otimes_R B$ are ring homomorphisms and the images are subrings, but although there are natural functions $A \to A \times B$ and $B \to A \times B$ given by $a \mapsto (a, 0)$ and $b \mapsto (0, b)$, these are *not* ring homomorphisms and the images are ideals rather than subrings.

Example 7.6. We saw in Example 7.3 that $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C} \cong \mathbf{C} \times \mathbf{C}$ as $\mathbf{R}$-algebras. How do $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ and $\mathbf{R} \times \mathbf{R}$ compare? They are not isomorphic as real vector spaces since $\dim_{\mathbf{R}}(\mathbf{R} \times \mathbf{R}) = 2$ while $\dim_{\mathbf{R}}(\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}) = \dim_{\mathbf{Q}}(\mathbf{R}) = \infty$. An $\mathbf{R}$-algebra isomorphism would in particular be an $\mathbf{R}$-vector space isomorphism, so $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R} \not\cong \mathbf{R} \times \mathbf{R}$ as $\mathbf{R}$-algebras. To show $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ is not isomorphic to $\mathbf{R} \times \mathbf{R}$ just as rings, we'll count square roots of 1. In $\mathbf{R} \times \mathbf{R}$ there are four square roots of 1, namely $(\pm 1, \pm 1)$ using independent choices of signs, but in $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ there are at least six square roots of 1:

$$\pm(1 \otimes 1), \quad \pm\frac{1}{2}(\sqrt{2} \otimes \sqrt{2}) \quad \text{and} \quad \pm\frac{1}{3}(\sqrt{3} \otimes \sqrt{3}).$$

These six tensors look different from each other, but how do we know they really are different? The numbers $1, \sqrt{2}$, and $\sqrt{3}$ are linearly independent over $\mathbf{Q}$ (why?), so any elementary tensors formed from these numbers in $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ are linearly independent over $\mathbf{Q}$. This proves the six elementary tensors above are distinct.

We will see in Example 7.18 that $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ has infinitely many square roots of 1.

The R-algebras $A \times B$ and $A \otimes_R B$ have dual universal mapping properties. For any R-algebra C and R-algebra homomorphisms $C \xrightarrow{\varphi} A$ and $C \xrightarrow{\psi} B$, there is a unique R-algebra homomorphism $C \longrightarrow A \times B$ making the diagram

$$\begin{tikzcd}
 & C \arrow[d, dotted] \arrow[ddl, bend right, "\varphi"'] \arrow[ddr, bend left, "\psi"] & \\
 & A \times B \arrow[dl, "\pi_1"] \arrow[dr, "\pi_2"'] & \\
A & & B
\end{tikzcd}$$

commute. For any R-algebra C and R-algebra homomorphisms $A \xrightarrow{\varphi} C$ and $B \xrightarrow{\psi} C$ such that *the images of A and B in C commute* ($\varphi(a)\psi(b) = \psi(b)\varphi(a)$), there is a unique R-algebra homomorphism $A \otimes_R B \longrightarrow C$ making the diagram

$$\begin{tikzcd}
A \arrow[dr, "a \mapsto a \otimes 1"] \arrow[ddr, bend right, "\varphi"'] & & B \arrow[dl, "b \mapsto 1 \otimes b"'] \arrow[ddl, bend left, "\psi"] \\
 & A \otimes_R B \arrow[d, dotted] & \\
 & C &
\end{tikzcd}$$

commute.

A practical criterion for showing an R-linear map of R-algebras is an R-algebra homomorphism is as follows. If $\varphi\colon A \to B$ is an R-linear map of R-algebras and $\{a_i\}$ is a spanning set for A as an R-module (that is, $A = \sum_i Ra_i$), then φ is multiplicative as long as it is so on these module generators: $\varphi(a_i a_j) = \varphi(a_i)\varphi(a_j)$ for all i and j. Indeed, if this equation holds then

$$\begin{aligned}
\varphi\left(\sum_i r_i a_i \cdot \sum_j r'_j a_j\right) &= \varphi\left(\sum_{i,j} r_i r'_j a_i a_j\right) \\
&= \sum_{i,j} r_i r'_j \varphi(a_i a_j) \\
&= \sum_{i,j} r_i r'_j \varphi(a_i)\varphi(a_j) \\
&= \sum_i r_i \varphi(a_i) \sum_j r'_j \varphi(a_j) \\
&= \varphi\left(\sum_i r_i a_i\right) \varphi\left(\sum_j r'_j a_j\right).
\end{aligned}$$

This will let us bootstrap a lot of known R-module isomorphisms between tensor products to R-algebra isomorphisms by checking the behavior only on products of elementary tensors (and checking the multiplicative identity is preserved, which is always easy). We give some concrete examples before stating some general theorems.

Example 7.7. For ideals I and J in R, there is an isomorphism $\varphi\colon R/I \otimes_R R/J \longrightarrow R/(I+J)$ of R-modules where $\varphi(\overline{x} \otimes \overline{y}) = \overline{xy}$. Then $\varphi(\overline{1} \otimes \overline{1}) = \overline{1}$ and

$$\varphi((\overline{x} \otimes \overline{y})(\overline{x}' \otimes \overline{y}')) = \varphi(\overline{xx'} \otimes \overline{yy'}) = \overline{xx'yy'} = \overline{xy}\ \overline{x'y'} = \varphi(\overline{x} \otimes \overline{y})\varphi(\overline{x}' \otimes \overline{y}').$$

So $R/I \otimes_R R/J \cong R/(I+J)$ as R-algebras, not just as R-modules. In particular, the additive isomorphism $\mathbf{Z}/a\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/b\mathbf{Z} \cong \mathbf{Z}/(a,b)\mathbf{Z}$ is in fact an isomorphism of rings.

Example 7.8. There is an R-module isomorphism $\varphi\colon R[X] \otimes_R R[Y] \longrightarrow R[X,Y]$ where $\varphi(f(X) \otimes g(Y)) = f(X)g(Y)$. Let's show it's an R-algebra isomorphism: $\varphi(1 \otimes 1) = 1$ and

$$\begin{aligned}
\varphi((f_1(X) \otimes g_1(Y))(f_2(X) \otimes g_2(Y))) &= \varphi(f_1(X)f_2(X) \otimes g_1(Y)g_2(Y)) \\
&= f_1(X)f_2(X)g_1(Y)g_2(Y) \\
&= f_1(X)g_1(Y)f_2(X)g_2(Y) \\
&= \varphi(f_1(X) \otimes g_1(Y))\varphi(f_2(X) \otimes g_2(Y)),
\end{aligned}$$

so $R[X] \otimes_R R[Y] \cong R[X,Y]$ as R-algebras, not just as R-modules. (It would have sufficed to check φ is multiplicative on pairs of monomial tensors $X^i \otimes Y^j$.)

In a similar way, the natural R-module isomorphism $R[X]^{\otimes k} \cong R[X_1, \ldots, X_n]$, where the indeterminate X_i on the right corresponds on the left to the tensor $1 \otimes \cdots \otimes X \otimes \cdots \otimes 1$ with X in the ith position, is an isomorphism of R-algebras.

Example 7.9. When R is a domain with fraction field K, $K \otimes_R K \cong K$ as R-modules by $x \otimes y \mapsto xy$. This sends $1 \otimes 1$ to 1 and preserves multiplication on elementary tensors, so it is an isomorphism of R-algebras.

Example 7.10. Let F be a field. When x and y are independent indeterminates over F, $F[x] \otimes_F F[y] \cong F[x,y]$ as F-algebras by Example 7.8. It is natural to think that we should also have $F(x) \otimes_F F(y) \cong F(x,y)$ as F-algebras, but this is always false! Why should it be false, and is there a concrete way to think about $F(x) \otimes_F F(y)$?

Every tensor t in $F(x) \otimes_F F(y)$ is a finite sum $\sum_{i,j} f_i(x) \otimes g_j(y)$. We can give all the $f_i(x)$'s a common denominator and all the $g_j(y)$'s a common denominator, say $f_i(x) = a_i(x)/b(x)$ and $g_j(y) = c_j(y)/d(y)$ where $a_i(x) \in F[x]$ and $c_j(y) \in F[y]$. Then

$$t = \sum_{i,j} \frac{a_i(x)}{b(x)} \otimes_F \frac{c_j(y)}{d(y)} = \left(\frac{1}{b(x)} \otimes \frac{1}{d(y)}\right) \sum_{i,j} a_i(x) \otimes c_j(y).$$

Since $F[x] \otimes_F F[y] \cong F[x, y]$ as F-algebras by multiplication, this suggests comparing t with the rational function we get by multiplying terms in each elementary tensor, which leads to

$$\frac{\sum_{i,j} a_i(x)c_j(y)}{b(x)d(y)}.$$

The numerator is a polynomial in x and y, and every polynomial in $F[x, y]$ has that form (all polynomials in x and y are sums of polynomials in x times polynomials in y). The denominator, however, is quite special: it is a single polynomial in x times a single polynomial in y. Most rational functions in $F(x, y)$ don't have such a denominator. For example, $1/(1 - xy)$ can't be written to have a denominator of the form $b(x)d(y)$ (proof?).

To show $F(x) \otimes_F F(y)$ is isomorphic as an F-algebra to the rational functions in $F(x, y)$ having a denominator in the factored form $b(x)d(y)$, show that the multiplication mapping $F(x) \otimes_F F(y) \to F(x, y)$ given by $f(x) \otimes g(y) \mapsto f(x)g(y)$ on elementary tensors is an embedding of F-algebras. That it is an F-algebra homomorphism follows by the same argument used in Example 7.8. It is left to the reader to show the kernel is 0 from the known fact that the multiplication mapping $F[x] \otimes_F F[y] \to F[x, y]$ is injective. (Hint: Justify the idea of clearing denominators.) Thus $F(x) \otimes_F F(y)$ is an integral domain that is not a field, since its image in $F(x, y)$ is not a field: the image contains $F[x, y]$ but is smaller than $F(x, y)$. Concretely, the fact that $1 - xy$ is the image of $1 \otimes 1 - x \otimes y$ but $1/(1 - xy)$ is not in the image shows $1 \otimes 1 - x \otimes y$ is not invertible in $F(x) \otimes_F F(y)$. (In terms of localizations, $F(x) \otimes_F F(y)$ is isomorphic as an F-algebra to the localization of $F[x, y]$ at the multiplicative set of all products $b(x)d(y)$.)

Example 7.11. For any R-module M, there is an S-linear map

$$S \otimes_R \mathrm{End}_R(M) \longrightarrow \mathrm{End}_S(S \otimes_R M)$$

where $s \otimes \varphi \mapsto s\varphi_S = s(1 \otimes \varphi)$. Both sides are S-algebras. Check this S-linear map is an S-algebra map. When M is finite free this map is a bijection (chase bases), so it is an S-algebra isomorphism. For other M it might not be an isomorphism.

As a concrete instance of this, when $M = R^n$ we get $S \otimes_R \mathrm{M}_n(R) \cong \mathrm{M}_n(S)$ as S-algebras, not just as S-modules. In particular, $\mathbf{C} \otimes_{\mathbf{R}} M_n(\mathbf{R}) \cong M_n(\mathbf{C})$ as $\mathbf{C}$-algebras.

Example 7.12. If I is an ideal in R and A is an R-algebra, $R/I \otimes_R A \cong A/IA$ first as R-modules, then as R-algebras (the R-linear isomorphism is also multiplicative and preserves identities), and finally as R/I-algebras since the isomorphism is R/I-linear too.

Theorem 7.13. *Let A, B, and C be R-algebras. The standard R-module isomorphisms*

$$A \otimes_R B \cong B \otimes_R A,$$
$$A \otimes_R (B \times C) \cong (A \otimes_R B) \times (A \otimes_R C)$$
$$(A \otimes_R B) \otimes_R C \cong A \otimes_R (B \otimes_R C).$$

are all R-algebra isomorphisms.

The distributivity of $\otimes$ over $\times$ suggests denoting the direct product of algebras as a direct sum with $\oplus$.

Proof. Exercise. Note the direct product of two R-algebras is the direct sum as R-modules with componentwise multiplication, so first just treat the direct product as a direct sum. $\square$

Corollary 7.14. *For R-algebras A and B, $A \otimes_R B^n \cong (A \otimes_R B)^n$ as R-algebras.*

Proof. Induct on n. Note B^n here means the n-fold product ring, not $B^{\otimes n}$. $\square$

We turn now to base extensions. Fix a homomorphism $f: R \to S$ of commutative rings. We can restrict scalars from S-modules to R-modules and extend scalars from R-modules to S-modules. What about between R-algebras and S-algebras? An example is the formation of $\mathbf{C} \otimes_{\mathbf{R}} \mathrm{M}_n(\mathbf{R})$, which ought to look like $\mathrm{M}_n(\mathbf{C})$ as rings (really, as $\mathbf{C}$-algebras) and not just as complex vector spaces.

If A is an S-algebra, then we make A into an R-module in the usual way by $ra = f(r)a$, and this makes A into an R-algebra (restriction of scalars). More interesting is extension of scalars. For this we need a lemma.

Lemma 7.15. *If A, A', B, and B' are all R-algebras and $A \xrightarrow{\varphi} A'$ and $B \xrightarrow{\psi} B'$ are R-algebra homomorphisms then the R-linear map $A \otimes_R B \xrightarrow{\varphi \otimes \psi} A' \otimes_R B'$ is an R-algebra homomorphism.*

Proof. Exercise. $\square$

Theorem 7.16. *Let A be an R-algebra.*

(1) *The base extension $S \otimes_R A$, which is both an R-algebra and an S-module, is an S-algebra by its S-scaling.*
(2) *If $A \xrightarrow{\varphi} B$ is an R-algebra homomorphism then $S \otimes_R A \xrightarrow{1\otimes\varphi} S \otimes_R B$ is an S-algebra homomorphism.*

Proof. 1) We just need to check multiplication in $S \otimes_R A$ commutes with S-scaling (not just R-scaling): $s(tt') = (st)t' = t(st')$. Since all three expressions are additive in t and t', it suffices to check this when t and t' are elementary tensors:

$$s((s_1 \otimes a_1)(s_2 \otimes a_2)) \stackrel{?}{=} (s(s_1 \otimes a_1))(s_2 \otimes a_2) \stackrel{?}{=} (s_1 \otimes a_1)(s(s_2 \otimes a_2)).$$

From the way S-scaling on $S \otimes_R A$ is defined, all these products equal $ss_1s_2 \otimes a_1a_2$.

2) For an R-algebra homomorphism $A \xrightarrow{\varphi} B$, the base extension $S \otimes_R A \xrightarrow{1\otimes\varphi} S \otimes_R B$ is S-linear and it is an R-algebra homomorphism by Lemma 7.15. Therefore it is an S-algebra homomorphism. $\square$

We can also give $A \otimes_R S$ an S-algebra structure by S-scaling and the natural S-module isomorphism $S \otimes_R A \cong A \otimes_R S$ is an S-algebra isomorphism.

Example 7.17. Let I be an ideal in $R[X_1, \ldots, X_n]$. Check the S-module isomorphism $S \otimes_R R[X_1, \ldots, X_n]/I \cong S[X_1, \ldots, X_n]/(I \cdot S[X_1, \ldots, X_n])$ from Corollary 2.25 is an S-algebra isomorphism.

In one-variable, with $I = (h(X))$ a principal ideal in $R[X]$,[14] Example 7.17 gives us an S-algebra isomorphism

$$S \otimes_R R[X]/(h(X)) \cong S[X]/(h^f(X)),$$

[14]Not all ideals in $R[X]$ have to be principal, but this is just an example.

where $h^f(X)$ is the result of applying $f\colon R \to S$ to the coefficients of $h(X)$. (If $f\colon \mathbf{Z} \to \mathbf{Z}/p\mathbf{Z}$ is reduction mod p, for instance, then $h^f(X) = h(X) \bmod p$.) This isomorphism is particularly convenient, as it lets us compute a lot of tensor products of *fields*.

Example 7.18. Writing $\mathbf{Q}(\sqrt{2})$ as $\mathbf{Q}[X]/(X^2-2)$ (as a $\mathbf{Q}$-algebra), we have

$$\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{Q}(\sqrt{2}) \cong \mathbf{R} \otimes_{\mathbf{Q}} \mathbf{Q}[X]/(X^2-2) \cong \mathbf{R}[X]/(X^2-2) \cong \mathbf{R} \times \mathbf{R}$$

as $\mathbf{R}$-algebras since X^2-2 factors into distinct linear polynomials in $\mathbf{R}[X]$, and

$$\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{Q}(\sqrt[3]{2}) \cong \mathbf{R} \otimes_{\mathbf{Q}} \mathbf{Q}[X]/(X^3-2) \cong \mathbf{R}[X]/(X^3-2) \cong \mathbf{R} \times \mathbf{C}$$

as $\mathbf{R}$-algebras since X^3-2 has irreducible factors in $\mathbf{R}[X]$ of degree 1 and 2.

More generally, we have an $\mathbf{R}$-algebra isomorphism

$$\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{Q}(\sqrt[n]{2}) \cong \mathbf{R}[X]/(X^n-2).$$

The $\mathbf{Q}$-linear embedding $\mathbf{Q}(\sqrt[n]{2}) \hookrightarrow \mathbf{R}$ extends to an $\mathbf{R}$-linear embedding $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{Q}(\sqrt[n]{2}) \to \mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ that is multiplicative (it suffices to check that on elementary tensors), so the ring $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ contains a subring isomorphic to $\mathbf{R}[X]/(X^n-2)$. When n is odd, X^n-2 has one linear factor and $(n-1)/2$ quadratic irreducible factors in $\mathbf{R}[X]$, so $\mathbf{R}[X]/(X^n-2) \cong \mathbf{R} \times \mathbf{C}^{(n-1)/2}$ as $\mathbf{R}$-algebras. Therefore $\mathbf{R}[X]/(X^n-2)$ contains $2^{1+(n-1)/2} = 2^{(n+1)/2}$ square roots of 1. Letting $n \to \infty$ shows $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ contains infinitely many square roots of 1.

Example 7.19. We revisit Example 7.3, using Example 7.17:

$$\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C} \cong \mathbf{C} \otimes_{\mathbf{R}} (\mathbf{R}[X]/(X^2+1)) \cong \mathbf{C}[X]/(X^2+1) = \mathbf{C}[X]/(X-i)(X+i) \cong \mathbf{C} \times \mathbf{C}$$

as $\mathbf{R}$-algebras.

Let's make the $\mathbf{R}$-algebra isomorphism $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C} \cong \mathbf{C} \times \mathbf{C}$ explicit, as it is *not* $z \otimes w \mapsto (z, w)$. Tracing the effect of the isomorphisms on elementary tensors,

$$z \otimes (a+bi) \mapsto z \otimes (a+bX) \mapsto za + zbX \mapsto (za + zbi, za + ab(-i)) = (z(a+bi), z(a-bi)),$$

so $z \otimes w \mapsto (zw, z\overline{w})$. Thus $1 \otimes 1 \mapsto (1,1)$, $z \otimes 1 \mapsto (z,z)$, and $1 \otimes w \mapsto (w, \overline{w})$.

In these examples, a tensor product of fields is not a field. But the tensor product of fields *can* be a field (besides the trivial case $K \otimes_K K \cong K$). Here is an example.

Example 7.20. We have

$$\mathbf{Q}(\sqrt{2}) \otimes_{\mathbf{Q}} \mathbf{Q}(\sqrt{3}) \cong \mathbf{Q}(\sqrt{2}) \otimes_{\mathbf{Q}} \mathbf{Q}[X]/(X^2-3) \cong \mathbf{Q}(\sqrt{2})[X]/(X^2-3),$$

which is a field because X^2-3 is irreducible in $\mathbf{Q}(\sqrt{2})[X]$.

Example 7.21. As an example of a tensor product involving a finite field and a ring,

$$\mathbf{Z}/5\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}[i] \cong \mathbf{Z}/5\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}[X]/(X^2+1) \cong (\mathbf{Z}/5\mathbf{Z})[X]/(X^2+1) \cong \mathbf{Z}/5\mathbf{Z} \times \mathbf{Z}/5\mathbf{Z}$$

since $X^2+1 = (X-2)(X-3)$ in $(\mathbf{Z}/5\mathbf{Z})[X]$.

A general discussion of tensor products of fields is in [3, Sect. 8.18].

Theorem 7.22. *If A is R-algebra and B is an S-algebra, then the S-module structure on the R-algebra $A \otimes_R B$ makes it an S-algebra, and*

$$A \otimes_R B \cong (A \otimes_R S) \otimes_S B$$

as S-algebras sending $a \otimes b$ to $(a \otimes 1) \otimes b$.

Proof. It is left as an exercise to check the S-module and R-algebra structure on $A \otimes_R B$ make it an S-algebra. As for the isomorphism, from part I we know there is an S-module isomorphism with the indicated effect on elementary tensors. This function sends $1 \otimes 1$ to $(1 \otimes 1) \otimes 1$, which are the multiplicative identities. It is left to the reader to check this function is multiplicative on products of elementary tensors too. $\square$

Theorem 7.22 is particularly useful in field theory. Consider two field extensions L/K and F/K with an intermediate field $K \subset E \subset L$, as in the following diagram.

$$\begin{tikzcd}[every arrow/.append style={dash}] & & & L \\ F & & E \arrow[ur] & \\ & K \arrow[ul] \arrow[ur] & & \end{tikzcd}$$

Then there is a ring isomorphism

$$F \otimes_K L \cong (F \otimes_K E) \otimes_E L$$

which is also an isomorphism as E-algebras, F-algebras (from the left factor) and L-algebras (from the right factor).

Theorem 7.23. *Let A and B be R-algebras. There is an S-algebra isomorphism*

$$S \otimes_R (A \otimes_R B) \to (S \otimes_R A) \otimes_S (S \otimes_R B)$$

by $s \otimes (a \otimes b) \mapsto s((1 \otimes a) \otimes (1 \otimes b))$.

Proof. By part I, there is an S-module isomorphism with the indicated effect on tensors of the form $s \otimes (a \otimes b)$. This function preserves multiplicative identities and is multiplicative on such tensors (which span $S \otimes_R (A \otimes_R B)$), so it is an S-algebra isomorphism. $\square$

8. The Tensor Algebra of an R-module

Modules don't usually have a multiplication operation. That is, R-modules are not usually R-algebras. However, there is a construction that turns an R-module M into the generating set of an R-algebra in the "minimal" way possible. This R-algebra is the tensor algebra of M, which we'll construct in this section.

To start off, let's go over the difference between a generating set of an R-module and a generating set of an R-algebra. When we say an R-module M is generated by $m_1, \ldots, m_n$, we mean every element of M is an *R-linear combination* of $m_1, \ldots, m_n$. When we say an R-algebra A is generated by $a_1, \ldots, a_n$ we mean every element of A is a *polynomial* in $a_1, \ldots, a_n$ with coefficients in R, *i.e.*, is an R-linear combination of products of the a_i's. For example, the ring $R[X]$ is both an R-module and an R-algebra, but as an R-module it is generated by $\{1, X, X^2, \ldots\}$ while as an R-algebra it is generated by X alone. A generating set of an R-module is also called a spanning set, but the generating set of an R-algebra is not called a spanning set (the term "span" is used for linear things).

To enlarge an R-module M to an R-algebra, we want to multiply elements in M without having any multiplication defined in advance. (As in Section 7, R-algebras are associative.) The "most general" product $m_1 m_2$ for m_1 and m_2 in M should be bilinear in m_1 and m_2, so we want this product to be the elementary tensor $m_1 \otimes m_2$, which lives not in M but in

$M^{\otimes 2}$. Similarly, an expression like $m_1m_2 + m_3m_4m_5$ using five elements from M should be $m_1 \otimes m_2 + m_3 \otimes m_4 \otimes m_5$ in $M^{\otimes 2} \oplus M^{\otimes 3}$. This suggests creating an R-algebra as

$$\bigoplus_{k\geq 0} M^{\otimes k} = R \oplus M \oplus M^{\otimes 2} \oplus M^{\otimes 3} \oplus \cdots,$$

whose elements are formal sums $\sum t_k$ with $t_k \in M^{\otimes k}$ and $t_k = 0$ for all large k. We want to multiply by the intuitive rule

$$\sum_{k\geq 0} t_k \cdot \sum_{\ell\geq 0} t'_\ell = \sum_{n\geq 0}\left(\sum_{k+\ell=n} t_k \otimes t'_\ell\right), \tag{8.1}$$

where $t_k \otimes t'_\ell \in M^{\otimes n}$ if $k + \ell = n$. To show this multiplication makes the direct sum of all $M^{\otimes k}$ an R-algebra we use the following construction theorem.

Theorem 8.1. *Let $\{M_k\}_{k\geq 0}$ be a sequence of R-modules with $M_0 = R$ and let there be R-bilinear mappings ("multiplications") $\mu_{k,\ell}\colon M_k \times M_\ell \to M_{k+\ell}$ for all k and ℓ such that*

1) *(scaling by R) $\mu_{k,0}\colon M_k \times R \to M_k$ and $\mu_{0,\ell}\colon R \times M_\ell \to M_\ell$ are both scaling by R: $\mu_{k,0}(x, r) = rx$ and $\mu_{0,\ell}(r, y) = ry$ for $r \in R$, $x \in M_k$, and $y \in M_\ell$,*
2) *(associativity) for $k, \ell, n \geq 0$ we have $\mu_{k,\ell+n}(x, \mu_{\ell,n}(y, z)) = \mu_{k+\ell,n}(\mu_{k,\ell}(x, y), z)$ in $M_{k+\ell+n}$ for all $x \in M_k$, $y \in M_\ell$, and $z \in M_n$.*

The direct sum $\bigoplus_{k\geq 0} M_k$ is an associative R-algebra with identity using the multiplication rule $\sum_{k\geq 0} m_k \cdot \sum_{\ell\geq 0} m'_\ell = \sum_{n\geq 0}\left(\sum_{k+\ell=n} \mu_{k,\ell}(m_k, m'_\ell)\right)$.

Proof. The direct sum $\bigoplus_{k\geq 0} M_k$ is automatically an R-module. It remains to check (i) the multiplication defined on the direct sum is R-bilinear, (ii) $1 \in R = M_0$ is a multiplicative identity, and (iii) multiplication is associative. The R-bilinearity of multiplication is a bookkeeping exercise left to the reader. In particular, this includes distributivity of multiplication over addition. To prove multiplication has 1 as an identity and is associative, it suffices by distributivity to consider only multiplication with factors from direct summands M_k, in which case we can use the two properties of the maps $\mu_{k,\ell}$ in the theorem. □

Lemma 8.2. *For $k, \ell \geq 0$, there is a unique bilinear map $\beta_{k,\ell}\colon M^{\otimes k} \times M^{\otimes \ell} \to M^{\otimes(k+\ell)}$ where $\beta_{k,0}(t, r) = rt$, $\beta_{0,\ell}(r, t) = rt$, and for $k, \ell \geq 1$, $\beta_{k,\ell}(m_1 \otimes \cdots \otimes m_k, m'_1 \otimes \cdots \otimes m'_\ell) = m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell$ on pairs of elementary tensors.*

Proof. The cases $k = 0$ and $\ell = 0$ are trivial, so let $k, \ell \geq 1$. It suffices to construct a bilinear $\beta_{k,\ell}$ with the indicated values on pairs of elementary tensors; uniqueness is then automatic since elementary tensors span $M^{\otimes k}$ and $M^{\otimes \ell}$.

Define $f\colon \underbrace{M \times \cdots \times M}_{k \text{ times}} \times \underbrace{M \times \cdots \times M}_{\ell \text{ times}} \to M^{\otimes(k+\ell)}$ by

$$f(m_1, \ldots, m_k, m'_1, \ldots, m'_\ell) = m_1 \otimes \cdots \otimes m_k \otimes m'_1 \otimes \cdots \otimes m'_\ell.$$

This is $(k + \ell)$-multilinear. In particular, f is multilinear in the first k factors when the last ℓ factors are fixed and it is multilinear in the last ℓ factors when the first k factors are fixed, so we can collapse the first k factors and the last ℓ factors into tensor powers to get a bilinear mapping $\beta_{k,\ell}\colon M^{\otimes k} \times M^{\otimes \ell} \to M^{\otimes(k+\ell)}$ making the diagram

$$\begin{tikzcd}
 & M^{\otimes k} \times M^{\otimes \ell} \arrow[d, "\beta_{k,\ell}"] \\
(M \times \cdots \times M) \times (M \times \cdots \times M) \arrow[ur, "(\otimes,\otimes)"] \arrow[r, "f"'] & M^{\otimes(k+\ell)}
\end{tikzcd}$$

commute. (This collapsing is analogous to the proof of associativity of the tensor product in part I.) By commutativity of the diagram, on a pair of elementary tensors we have

$$\begin{aligned}
\beta_{k,\ell}(m_1 \otimes \cdots \otimes m_k, m_1' \otimes \cdots \otimes m_\ell') &= f(m_1, \dots, m_k, m_1', \dots, m_\ell) \\
&= m_1 \otimes \cdots \otimes m_k \otimes m_1' \otimes \cdots \otimes m_\ell'.
\end{aligned}$$

□

Theorem 8.3. *The M-module $\bigoplus_{k\geq 0} M^{\otimes k}$ is an R-algebra using the multiplication in* (8.1).

Proof. Use Theorem 8.1 with $M_k = M^{\otimes k}$ and $\mu_{k,\ell} = \beta_{k,\ell}$ from Lemma 8.2. The first property of the maps $\mu_{k,\ell}$ in Theorem 8.1 is automatic from the definition of $\beta_{k,0}$ and $\beta_{0,\ell}$. To prove the second property of the maps $\mu_{k,\ell}$ in Theorem 8.1, namely $\mu_{k,\ell+n}(t_1, \mu_{\ell,n}(t_2, t_3)) = \mu_{k+\ell,n}(\mu_{k,\ell}(t_1, t_2), t_3)$ in $M^{\otimes(k+\ell+n)}$ for all $t_1 \in M^{\otimes k}$, $t_2 \in M^{\otimes \ell}$, and $t_3 \in M^{\otimes n}$, by multilinearity of each $\mu_{k,\ell} = \beta_{k,\ell}$ it suffices to consider the case when each t_i is an elementary tensor, in which case the equality is a simple calculation. □

Definition 8.4. For an R-module M, its *tensor algebra* is $T(M) := \bigoplus_{k\geq 0} M^{\otimes k}$ with multiplication defined by (8.1).

Since multiplication in $T(M)$ is the tensor product, a generating set of M as an R-module is a generating set of $T(M)$ as an R-algebra.

Example 8.5. If $M = R$ then $M^{\otimes k} \cong R$ as an R-module and $T(M) \cong R[X]$ with X^k corresponding to the k-fold tensor $1 \otimes \cdots \otimes 1$ in $R^{\otimes k}$.

Example 8.6. If M is a finite free R-module with basis $e_1, \dots, e_n$ then $T(M)$ is the polynomial ring over R in n *noncommuting* indeterminates $e_1, \dots, e_n$: in $M^{\otimes 2}$, $e_i \otimes e_j \neq e_j \otimes e_i$ when $i \neq j$, which says in $T(M)$ that $e_i e_j \neq e_j e_i$.

Remark 8.7. The tensor product construction of the polynomial ring over R in n noncommuting indeterminates is quite different from that of the tensor product construction of the commutative polynomial ring $R[X_1, \dots, X_n]$: the former is the tensor algebra $T(R^n)$ of a free R-module of rank n while the latter is $R[X]^{\otimes n}$ (Example 7.8).

The mapping $i\colon M \to T(M)$ that identifies M with the $k = 1$ component of $T(M)$ is R-linear and injective, so we can view M as a submodule of $T(M)$ (the "degree 1" terms) using i. Just as the bilinear map $M \times N \to M \otimes_R N$ is universal for R-bilinear maps from $M \times N$ to all R-modules, the mapping $i\colon M \to T(M)$ is universal for R-linear maps from M to all R-algebras.

Theorem 8.8. *Let M be an R-module. For each R-algebra A and R-linear map $f \colon M \to A$, there is a unique R-algebra map $F \colon T(M) \to A$ such that the diagram*

$$\begin{tikzcd} & T(M) \arrow[dd, "F"] \\ M \arrow[ur, "i"] \arrow[dr, "f"'] & \\ & A \end{tikzcd}$$

commutes.

This says R-*linear* maps from M to an R-algebra A turn into R-*algebra homomorphisms* from $T(M)$ to A. It works by extending an R-linear map from $M \to A$ to an R-algebra map $T(M) \to A$ by forcing multiplicativity, and there are no relations to worry about keeping track of because $T(M)$ is an R-algebra formed from M in the most general way possible.

Proof. First suppose there is an R-algebra map F that makes the indicated diagram commute. For $r \in R$, $F(r) = r$ since F is an R-algebra homomorphism. Since $T(M)$ is generated as an R-algebra by M, an R-algebra homomorphism out of $T(M)$ is determined by its values on M, which really means its values on $i(M)$. For $m \in M$, we have $F(i(m)) = f(m)$, and thus there is at most one R-algebra homomorphism F that fits into the above commutative diagram.

To construct F, we will define it first on each $M^{\otimes k}$ in $T(M)$ and then extend by additivity. For $k \geq 1$, the R-linear map $f \colon M \to A$ leads to an R-linear map $f^{\otimes k} \colon M^{\otimes k} \to A^{\otimes k}$ that is $m_1 \otimes \cdots \otimes m_k \mapsto f(m_1) \otimes \cdots \otimes f(m_k)$ on elementary tensors. Multiplication on A gives us an R-linear map $A^{\otimes k} \to A$ that is $a_1 \otimes \cdots \otimes a_k \mapsto a_1 \cdots a_k$ on elementary tensors. Composing this with $f^{\otimes k}$ gives us an R-linear map $F_k \colon M^{\otimes k} \to A$ whose value on elementary tensors is

$$F_k(m_1 \otimes \cdots \otimes m_k) = f(m_1) \cdots f(m_k).$$

Define $F_0 \colon M^{\otimes 0} \to A$ by $F_0(r) = r \cdot 1_A$ for $r \in R$. Finally, define $F \colon T(M) \to A$ by

$$F\left(\sum_{k \geq 0} t_k\right) = \sum_{k \geq 0} F_k(t_k)$$

where $t_k \in M^{\otimes k}$ and $t_k = 0$ for large k. Since each F_k is R-linear and $F_0(1) = 1_A$, F is R-linear and $F(1) = 1_A$. To prove F is multiplicative, by linearity it suffices to check $F(xy) = F(x)F(y)$ where x is an elementary tensor in some $M^{\otimes k}$ and y is an elementary tensor in some $M^{\otimes \ell}$. The cases $k = 0$ and $\ell = 0$ are the linearity of F. If $k, \ell \geq 1$, write $x = m_1 \otimes \cdots \otimes m_k$ and $y = m_1' \otimes \cdots \otimes m_\ell'$. Then $xy = m_1 \otimes \cdots \otimes m_k \otimes m_1' \otimes \cdots \otimes m_\ell'$ in $T(M)$, so

$$F(xy) = F_{k+\ell}(xy) = f(m_1) \cdots f(m_k) f(m_1') \cdots f(m_\ell')$$

and

$$\begin{aligned} F(x)F(y) &= F_k(x)F_\ell(y) \\ &= (f(m_1) \cdots f(m_k))(f(m_1') \cdots f(m_\ell')) \\ &= f(m_1) \cdots f(m_k) f(m_1') \cdots f(m_\ell'). \end{aligned}$$

$\square$

Tensor algebras are useful preliminary constructions for other structures that can be defined as a quotient of them, such as the exterior algebra of a module, the Clifford algebra of a quadratic form, and the universal enveloping algebra of a Lie algebra.

References

[1] S. Friedl, *Eta invariants as sliceness obstructions and their relation to Casson–Gordon invariants*, Algebr. Geom. Topol. **4** (2004), 893–934.

[2] H. V. Henderson and F. Pukelsheim, *On the History of the Kronecker Product*, Linear and Multilinear Algebra **14** (1983), 113–120.

[3] N. Jacobson,"Basic Algebra II," 2nd ed., W. H. Freeman & Co., New York, 1989.

[4] E. Schrödinger, *Discussion of Probability Relations between Separated Systems*, Math. Proc. Camb. Philos. Society **31** (1935), 555–563. Online at https://www.cambridge.org/core/journals/mathematical-proceedings-of-the-cambridge-philosophical-society/article/discussion-of-probability-relations-between-separated-systems/C1C71E1AA5BA56EBE6588AAACB9A222D.