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 and discuss tensor products of R-algebras.

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 $b|a$ 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 $\operatorname{Tr}(A \otimes A') = a(a' + d') + d(a' + d') = (a + d)(a' + d') = (\operatorname{Tr} A)(\operatorname{Tr} A')$, 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 $\mathrm{Hom}_R(M, M')$ and $\mathrm{Hom}_R(N, N')$, so one could ask if the actual elementary tensor $\varphi \otimes \psi$ in $\mathrm{Hom}_R(M, M') \otimes_R \mathrm{Hom}_R(N, N')$ 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*

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

which 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 $\mathrm{Hom}_R(M, M') \times \mathrm{Hom}_R(N, N') \to \mathrm{Hom}_R(M \otimes_R N, M' \otimes_R N')$ 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 $\mathrm{Hom}_R(M, M') \otimes_R \mathrm{Hom}_R(N, N') \to \mathrm{Hom}_R(M \otimes_R N, M' \otimes_R N')$ 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 $\mathrm{Hom}_R(M, M')$ and $\mathrm{Hom}_R(N, N')$ 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 $\mathrm{Hom}_R(M, M') \otimes_R \mathrm{Hom}_R(N, N')$ 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 $\mathrm{Hom}_R(M, M') \otimes_R \mathrm{Hom}_R(N, N') \to \mathrm{Hom}_R(M \otimes_R N, M' \otimes_R N')$ sends a basis to a basis, so it is an isomorphism when the modules are finite free. □

The upshot of Theorem 2.5 is that $\mathrm{Hom}_R(M, M') \otimes_R \mathrm{Hom}_R(N, N')$ 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 which 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 $\varphi'(\varphi(m)) \otimes \psi'(\psi(n))$. □

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 $e_1, \dots, e_k$ of M and $e'_1, \dots, e'_\ell$ 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 $e_1, \dots, e_k$. 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

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

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

$$\det(\varphi \otimes \psi) = (\det \varphi)^\ell (\det \psi)^k.$$

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. □

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

Theorem 2.10. *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. □

Theorem 2.11. *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 $m' \otimes n' = \varphi(m) \otimes \psi(n) = (\varphi \otimes \psi)(m \otimes n)$. □

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 2.12. 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 $M \otimes_R N$. 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 $M \otimes_R N \to M' \otimes_R N'$ might not be injective. We might say it this way: a tensor product of submodules need not be a submodule.

Example 2.13. Again with $R = \mathbf{Z}$, let $i\colon p\mathbf{Z} \hookrightarrow \mathbf{Z}$ be inclusion. Tensoring i with $\mathbf{Z}/p\mathbf{Z}$ gives the linear map $1 \otimes i\colon \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} p\mathbf{Z} \to \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}$ with the effect $a \otimes px \mapsto a \otimes px = pa \otimes x = 0$, so $1 \otimes i$ is identically 0. Since $p\mathbf{Z} \cong \mathbf{Z}$, $\mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} p\mathbf{Z} \cong \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z} \cong \mathbf{Z}/p\mathbf{Z}$ by Theorem 2.10, so $1 \otimes i$ is not injective.

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 2.14. Again with $R = \mathbf{Z}$, let $i\colon \mathbf{Z} \hookrightarrow \mathbf{Q}$ be the inclusion. Tensor this with $\mathbf{Z}/p\mathbf{Z}$ to get the $\mathbf{Z}$-linear map $1 \otimes i\colon \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z} \to \mathbf{Z}/p\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Q}$ with domain isomorphic to $\mathbf{Z}/p\mathbf{Z}$ (not 0) and target 0, so $1 \otimes i$ is not injective.

Example 2.15. Here is an example of a linear map $i\colon M \to N$ which is injective and its tensor square $i^{\otimes 2}\colon M^{\otimes 2} \to N^{\otimes 2}$ is not injective.

Let $R = A[X, Y]$ with A a nonzero commutative ring and $I = (X, Y)$. Let $i\colon I \hookrightarrow R$ be the inclusion map. We will show $i^{\otimes 2}\colon I^{\otimes 2} \to R^{\otimes 2}$ is not injective, which means the natural way to think of $I \otimes_R I$ "inside" $R \otimes_R R$ is not an embedding. In other words, 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 $R \otimes_R R$.

The tensor $X \otimes Y$ in $I^{\otimes 2}$ gets sent by $i^{\otimes 2}$ to the tensor $X \otimes Y$ in $R^{\otimes 2}$, and in $R^{\otimes 2}$ we can write $X \otimes Y = XY(1 \otimes 1)$. (We can't do that in $I^{\otimes 2}$, as 1 is not in I.) Similarly, $i^{\otimes 2}$ sends $Y \otimes X$ in $I^{\otimes 2}$ to $Y \otimes X = YX(1 \otimes 1)$ in $R^{\otimes 2}$. Since $XY = YX$, $i^{\otimes 2}(X \otimes Y) = i^{\otimes 2}(Y \otimes X)$. We will now show that in $I^{\otimes 2}$ the difference $X \otimes Y - Y \otimes X$ is not zero, so the kernel of $i^{\otimes 2}$ is not 0.

Letting the partial derivatives of a polynomial $f(X, Y)$ with respect to X and Y be denoted f_X and f_Y, the function $I \times I \to A$ given by $(f, g) \mapsto f_X(0,0)g_Y(0,0)$ is R-bilinear (where A is an R-module through multiplication by the constant term of polynomials in R, or just view A as R/I), so there is an R-linear map $I^{\otimes 2} \to A$ sending $f \otimes g$ 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 $I^{\otimes 2}$.

The tensor product $I^{\otimes 2}$ in Example 2.15 exhibits another interesting feature when A is a domain: I is torsion-free but $I^{\otimes 2}$ is not: $XY(X \otimes Y) = XY \otimes XY = XY(Y \otimes X)$, so $XY(X \otimes Y - Y \otimes X) = 0$ and $X \otimes Y - Y \otimes X \neq 0$. Therefore a tensor product of torsion-free modules (even over a domain) need not be torsion-free!

Generalizing Example 2.15, let $R = A[X_1, \ldots, X_n]$ where $n \geq 2$ and $I = (X_1, \ldots, X_n)$. The inclusion $i\colon I \hookrightarrow R$ is injective but the nth tensor power $i^{\otimes n}\colon I^{\otimes n} \to R^{\otimes n}$ is not injective because the tensor

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

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 the original sum is not 0 because we can construct a linear map $I^{\otimes n} \to A$ sending it to 1 (using a product of partial derivatives at $(0, 0, \ldots, 0)$, as in the $n = 2$ case above).

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 2.16. *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 $N^{\otimes k}$.*

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 3.9 and Theorem 4.8.

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 2.17. *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 $\psi(n) = 0$. 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 we don't write the kernel of $\varphi \otimes \psi$ as $(\ker \varphi) \otimes_R N + M \otimes_R (\ker \psi)$, even though that is the intuitive idea, is that strictly speaking these terms are their own tensor product modules and only the application of $i \otimes 1$ and $1 \otimes j$ – which might not be injective – puts them inside $M \otimes_R N$.

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

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

[1]This is $m \otimes n$ in $M \otimes_R N$, whereas the previous $m \otimes n$ is in $(\ker \varphi) \otimes_R N$.

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. $\square$

Remark 2.18. If we remove the assumption that φ and ψ are onto, Theorem 2.17 does not correctly compute the kernel. For example, if φ and ψ are both injective then the formula for the kernel in Theorem 2.17 is 0, and we know $\varphi \otimes \psi$ need not be injective.

Unlike the kernel computation in Theorem 2.17, 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 2.15.

Corollary 2.19. *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 $S \otimes_R (N/M) = 0$. $\square$

Example 2.20. 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 2.19. 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 $\mathfrak{m}$ in R, so we can apply this to $P = N/M$ since $P/\mathfrak{m}P \cong R/\mathfrak{m} \otimes_R P$.

Corollary 2.21. *Let $f\colon R \to S$ be a homomorphism of commutative rings and I be an ideal in $R[X_1, \ldots, X_n]$. Write $I \cdot S[X_1, \ldots, X_n]$ for the ideal generated by the image of I in $S[X_1, \ldots, X_n]$. Then*

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

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

$$S \otimes_R R[X_1, \dots, X_n] \twoheadrightarrow S \otimes_R (R[X_1, \dots, X_n]/I) \tag{2.1}$$

and the kernel is $(1 \otimes j)(S \otimes_R I)$ by Theorem 2.17, where $j: I \to R[X_1, \dots, X_n]$ is the inclusion. Under the natural R-module isomorphism

$$S \otimes_R R[X_1, \dots, X_n] \cong S[X_1, \dots, X_n], \tag{2.2}$$

$(1 \otimes j)(S \otimes_R I)$ on the left side corresponds to $I \cdot S[X_1, \dots, X_n]$ on the right side, so (2.1) and (2.2) 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. □

Example 2.22. 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 which 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.

Example 3.2. For a nonzero torsion abelian group A, the natural map $\mathbf{Z} \hookrightarrow \mathbf{Q}$ is injective but if we tensor with A we get the map $A \to 0$, which is not injective, so A is not a flat $\mathbf{Z}$-module.

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

Theorem 3.3. *Any free R-module F is flat: if the linear map $\varphi: M \to M'$ is injective, then $1 \otimes \varphi: F \otimes_R M' \to F \otimes_R M'$ 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 $t = \sum_i e_i \otimes 0 = 0$. □

Note that in Theorem 3.3 we did not need to assume F has a finite basis.

Theorem 3.4. *Let R be a domain and K be its fraction field. As an R-module, K is flat.*

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 $\varphi(m) \in M'_{\text{tor}}$. If $\varphi(m) \in M'_{\text{tor}}$ 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 $m \in M_{\text{tor}}$. Thus $x = 0$ or $m \in M_{\text{tor}}$, so $t = x \otimes m = 0$. $\square$

If M is a submodule of the R-module M' then Theorem 3.4 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 3.5. The natural inclusion $\mathbf{Z} \hookrightarrow \mathbf{Z}/3\mathbf{Z} \oplus \mathbf{Z}$ is $\mathbf{Z}$-linear and injective. Tensoring with $\mathbf{Q}$ and using properties of tensor products turns this into the identity map $\mathbf{Q} \to \mathbf{Q}$, which is also injective.

Remark 3.6. Theorem 3.4 generalizes: for any commutative ring R and multiplicative set D in R, the localization R_D is a flat R-module.

Theorem 3.7. *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 3.8. *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.*

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 3.9. *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 3.7 and induction). □

Corollary 3.10. *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 3.3. □

Note the free modules in Corollary 3.10 are completely arbitrary. We make no assumptions about finite bases.

Corollary 3.10 is not a special case of Theorem 2.16 because a free submodule of a free module need not be a direct summand (consider the ring $\mathbf{Z}$ and any proper subgroup of rank n in $\mathbf{Z}^n$ as a $\mathbf{Z}$-module inside of $\mathbf{Z}^n$).

Corollary 3.11. *If M is a free module and $\{m_1, \ldots, m_s\}$ is a finite linearly independent subset then for any $k \leq s$ the s^k elementary tensors*

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

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

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

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

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

Corollary 3.11 is *not* saying the elementary tensors in (3.1) can be extended to a basis of $\Lambda^k(M)$, any more than $m_1, \ldots, m_s$ can 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 $S \otimes_R M \xrightarrow{1 \otimes \varphi} S \otimes_R M'$ which 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 this 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 2.10 and 2.11). But if φ is injective then φ_S need not be injective. (Examples 2.12, 2.13, and 2.14, 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 $a_1 | \cdots | a_k$, 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. In particular, there is a unique S-linear map

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

where $s \otimes \varphi \mapsto s\varphi_S$, and it is an isomorphism if M is finite-free.

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. $\square$

Example 4.7. Let $R = \mathbf{Z}/p^2\mathbf{Z}$ and $S = \mathbf{Z}/p\mathbf{Z}$. Take $M = \mathbf{Z}/p\mathbf{Z}$ as an R-module. The linear map

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

has image 0, but the right side is isomorphic to $\mathrm{Hom}_{\mathbf{Z}/p\mathbf{Z}}(\mathbf{Z}/p\mathbf{Z}, \mathbf{Z}/p\mathbf{Z}) \cong \mathbf{Z}/p\mathbf{Z}$, so this map is not an isomorphism.

In Corollary 3.10 we saw any tensor power of an injective linear map between free modules is injective. Using base extension, we can drop the requirement that the target module be free provided we are working over a domain.

Theorem 4.8. *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 $k\geq 1$.*

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). $\square$

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 $V \otimes_K W \xrightarrow{\varphi\otimes\psi} V' \otimes_K W'$ which 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 2.10 and 2.11). Moreover, because all K-vector spaces are free a tensor product of injective K-linear maps is injective (Theorem 3.3).

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 $\varphi \colon V' \to K$ which 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 $V \otimes_K W' \subsetneq V' \otimes_K W'$. $\square$

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,[2] 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.[3]

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, we could define the characteristic polynomial of a linear operator $V \xrightarrow{\varphi} V$ 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 2.17.

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

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

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 2.17 for modules has weaker conclusions (*e.g.*, injectivity of $\varphi_1 \otimes \varphi_2$ doesn't imply injectivity of φ_1 and φ_2).

---

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

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

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 2.17 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. □

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 $V' \otimes_K W' \xrightarrow{\pi_1 \otimes \pi_2} (V'/V) \otimes_K (W'/W)$ which is onto with $\ker(\pi_1 \otimes \pi_2) = V \otimes_K W' + V' \otimes_K W$ by Theorem 5.5. □

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[4] 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*

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

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. □

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

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 $\varphi(v) = \varphi(v_{i_0}) = 1$. □

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.*

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 $i = 1, \ldots, k$ 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. $\square$

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 tensors with torsion that are not elementary, as we saw at the end of Example 2.15.

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 is 0. We now show two nonzero elementary tensors in vector spaces are equal only under the "obvious" circumstances.

Theorem 5.15. *Let $V_1, \ldots, V_k$ be K-vector spaces. Pick pairs of nonzero vectors v_i, v_i' in V_i for $i = 1, \ldots, k$. If $v_1 \otimes \cdots \otimes v_k = v_1' \otimes \cdots \otimes v_k'$ in $V_1 \otimes_K \cdots \otimes_K V_k$ then there are constants $c_1, \ldots, c_k$ 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. Pick $\varphi_i \in V_i^\vee$ for $1 \leq i \leq k-1$ such that $\varphi_i(v_i) = 1$. For *any* $\varphi \in V_k^\vee$, 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 all $\varphi \in V_k^\vee$. 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 $c_1 \cdots c_k = 1$. $\square$

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 $V_1 \otimes_K \cdots \otimes_K V_k \to W_1 \otimes_K \cdots \otimes_K W_k$ if and only if there are $c_1, \ldots, c_k$ in K such that $\varphi_i = c_i \varphi_i'$ and $c_1 c_2 \cdots c_k = 1$.*

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 $c_k f(\varphi'_k(v)) = f(c_k\varphi'_k(v))$. 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:

$$(v_1, \cdots, v_k) \mapsto (v_{\sigma(1)}, \cdots, v_{\sigma(k)}).$$

There 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*

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

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

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.

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

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

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 $v_j \in Kv_i$ by Theorem 5.15, which is not so. □

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, as a computation on elementary tensors shows.

Tensors $t \in V^{\otimes k}$ satisfying $\sigma(t) = t$ for all $\sigma \in S_k$ are called *symmetric* and tensors $t \in V^{\otimes k}$ satisfying $\sigma(t) = (\operatorname{sign}\sigma)t$ for all $\sigma \in S_k$ are called *skew-symmetric* or *anti-symmetric*. Both kinds of tensors occur in physics. We'll see some examples in Section 6. The symmetric tensors and skew-symmetric tensors each form subspaces of $V^{\otimes k}$. If K does not have characteristic 2, every tensor in $V^{\otimes 2}$ is a unique sum of a symmetric and skew-symmetric tensor:

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

where σ is the non-identity permutation in S_2. (It is the flip automorphism of $V^{\otimes 2}$ sending $v \otimes w$ to $w \otimes v$.) 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 $V^{\otimes k}$ which are needed to fill out the whole space, and these subspaces are 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. Tensors in Physics

The name tensor was introduced by the mathematical physicist Woldemar Voigt in the late 19th century in the context of measuring stress and strain. In the table below are examples of tensors in different areas of physics.

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

What is it about tensors that makes them show up in physics? Let's first understand why vectors show up in physics. The physical meaning of a vector is not just displacement, but linear displacement. For instance, forces at a point combine in the same way that vectors add (this is an experimental observation), so force is treated as a vector. While the physical meaning of vector is a linear displacement, the physical meaning of a tensor is *multilinear displacement*.

Here is a definition of a tensor that can be found (more or less) in physics textbooks. For $k \geq 1$, a *rank k tensor* on an n-dimensional vector space[5] V is a choice, for each basis

[5]The physicist may write $\mathbf{R}^3$ or $\mathbf{R}^n$ in place of V.

$\mathcal{B} = \{e_1, \ldots, e_n\}$, of a collection of n^k numbers $\{T_{\mathcal{B}}^{i_1,\ldots,i_k}\}_{1\leq i_1,\ldots,i_k\leq n}$ such that for two bases $\mathcal{B} = \{e_1, \ldots, e_n\}$ and $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ of V with change-of-basis matrix $A = (a_{ij})$ from $\mathcal{B}$ to $\mathcal{B}'$ (so $e_j = \sum_{i=1}^n a_{ij}e'_i$) the two systems of n^k numbers are related by the formula

$$T_{\mathcal{B}'}^{i_1,\ldots,i_k} = \sum_{1\leq j_1,\ldots,j_k\leq n} T_{\mathcal{B}}^{j_1,\ldots,j_k} a_{i_1j_1}\cdots a_{i_kj_k}. \tag{6.1}$$

Do you see what a rank k tensor on V means in mathematical language? What is being defined here is a member of the kth tensor power $V^{\otimes k}$. Place the array of numbers $\{T_{\mathcal{B}}^{i_1,\ldots,i_k}\}_{1\leq i_1,\ldots,i_k\leq n}$ associated to a particular basis $\mathcal{B} = \{e_1, \ldots, e_n\}$ of V as the coefficients for the associated basis $\{e_{i_1}\otimes\cdots\otimes e_{i_k}\}_{1\leq i_1,\ldots,i_k\leq n}$ of $V^{\otimes k}$:

$$t := \sum_{1\leq i_1,\ldots,i_k\leq n} T_{\mathcal{B}}^{i_1,\ldots,i_k} e_{i_1}\otimes\cdots\otimes e_{i_k}.$$

Let's write t in terms of the basis $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ for V: when $e_j = \sum_{i=1}^n a_{ij}e'_i$,

$$\begin{aligned} t &= \sum_{1\leq j_1,\ldots,j_k\leq n} T_{\mathcal{B}}^{j_1,\ldots,j_k} e_{j_1}\otimes\cdots\otimes e_{j_k} \\ &= \sum_{1\leq j_1,\ldots,j_k\leq n} T_{\mathcal{B}}^{j_1,\ldots,j_k} \left(\sum_{i_1=1}^n a_{i_1j_1}e'_{i_1}\right)\otimes\cdots\otimes\left(\sum_{i_k=1}^n a_{i_1kj_k}e'_{i_k}\right) \\ &= \sum_{1\leq i_1,\ldots,i_k\leq n}\left(\sum_{1\leq j_1,\ldots,j_k\leq n} T_{\mathcal{B}}^{j_1,\ldots,j_k} a_{i_1j_1}\cdots a_{i_kj_k}\right) e'_{i_1}\otimes\cdots\otimes e'_{i_k}. \end{aligned}$$

If we label the coefficients of t with respect to the basis $\{e'_{i_1}\otimes\cdots\otimes e'_{i_k}\}_{1\leq i_1,\ldots,i_k\leq n}$ as $T_{\mathcal{B}'}^{i_1,\ldots,i_k}$ then we recover the formula (6.1). So the physicist's rank k tensor is just all the different coordinate representations of a single element of $V^{\otimes k}$.

Physics textbooks say tensors are indexed quantities with "components that transform in a definite way"[6], but this is like defining a group to be a set equipped with a "definite law of composition" and then forgetting to add that the group law is supposed to satisfy some actual properties (not every set with a law of composition is a group). The key structure in the rule (6.1) is its *multilinearity* via the products $a_{i_1j_1}\cdots a_{i_kj_k}$. I have never seen a physics textbook whose discussion of tensors uses the word multilinear. The transformation rule (6.1) under a change of coordinates is not simply a "definite" rule but rather is a rule which *depends multilinearly on the change of coordinate data.* Tensors are "multilinear functions of several directions."[7]

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

- (Mathematician) A tensor belongs to a tensor space, which is a vector space defined by a multilinear universal mapping property. This is a coordinate-free point of view.
- (Physicist) A tensor is something defined by a multi-indexed system of components in all coordinate systems on V, and its components in two different coordinate systems are related by the transformation formula (6.1).

Mathematicians and physicists can check two rank k tensors t and t' are equal in the same way: check t and t' have the same components in one coordinate system. But the reason they consider this to be a verification of equality is not the same. The mathematician

[6] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 6th ed., p. 133

[7] G. F. J. Temple, *Cartesian Tensors*, 1960, p. 9

thinks about the condition $t = t'$ in a coordinate-free way but knows that to check $t = t'$ it suffices to check t and t' have the same coordinates in a basis. The physicist, who thinks about tensors in component form, considers the condition $t = t'$ to mean (by definition!) the components of t and t' match in all coordinate systems, and the rule (6.1) implies that if the components of t and t' are equal in one coordinate system then they are equal in any other coordinate system. That's why the physicist is content to look only in one coordinate system.

An operation on tensors (like the flip $v \otimes w \mapsto w \otimes v$ in $V^{\otimes 2}$) is checked to be well-defined by the mathematician and physicist in different ways. The mathematician checks the operation respects the universal mapping property that defines tensor products of vector spaces, while the physicist checks the explicit formula for the operation on individual tensors changes in different coordinate systems by the rule (6.1). The physicist would say an operation on tensors makes sense because it "transforms tensorially."

While the mathematicians may shake their heads and wonder how physicists can think about tensors in component-dependent ways, that more concrete viewpoint is crucial to understanding how tensors show up in physics. A physical quantity whose descriptions in two different coordinate systems are related to each other in the same way that the coordinates of a tensor in two different coordinate systems are related is asking to be mathematically described as a tensor.

A nice example of this is in electromagnetism. Electric fields and magnetic fields are first learned about separately as vector fields in $\mathbf{R}^3$. Experiments show that a changing electric field produces a magnetic field and a changing magnetic field produces an electric field, so electric and magnetic fields are not really well-defined objects independent of the coordinate system: an electric field that vanishes in one coordinate system is nonzero in a second coordinate system moving with respect to the first one. But the formulas describing how the components of the electric and magnetic fields interact (6 components total, 3 from each field) show that under a change of coordinates they transform together in the same way as an antisymmetric rank 2 tensor on a 4-dimensional space. This is why physicists describe an electromagnetic field as a rank 2 tensor (in the 4-dimensional Minkowski spacetime).

Physicists adopt the habit of identifying $\mathbf{R}^n$ with its dual space $(\mathbf{R}^n)^\vee$ via the dot product: $\mathbf{v}$ is associated to the linear map $\mathbf{w} \mapsto \mathbf{v} \cdot \mathbf{w}$. For example, a bilinear form $B \colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ is the same thing as a linear function $\mathbf{R}^n \otimes_{\mathbf{R}} \mathbf{R}^n \to \mathbf{R}$, and all such functions form the dual space $(\mathbf{R}^n \otimes_{\mathbf{R}} \mathbf{R}^n)^\vee \cong (\mathbf{R}^n)^\vee \otimes_{\mathbf{R}} (\mathbf{R}^n)^\vee$, which can be thought of as $\mathbf{R}^n \otimes_{\mathbf{R}} \mathbf{R}^n$, so bilinear forms on $\mathbf{R}^n$ are also rank 2 tensors on $\mathbf{R}^n$. Since $\mathrm{M}_n(\mathbf{R}) = \mathrm{Hom}_{\mathbf{R}}(\mathbf{R}^n, \mathbf{R}^n) \cong (\mathbf{R}^n)^\vee \otimes_{\mathbf{R}} \mathbf{R}^n$ can be thought of as $\mathbf{R}^n \otimes_{\mathbf{R}} \mathbf{R}^n$, linear maps from $\mathbf{R}^n$ to $\mathbf{R}^n$ are rank 2 tensors on $\mathbf{R}^n$. This accounts for the wide use of rank 2 tensors in physics, as they include both bilinear forms and linear transformations. General relativity is full of rank 2 tensors for this reason.

The most basic example of a tensor in mechanics is the stress tensor, which has rank 2. When a force is applied to a body the stress it imparts at a point may not be in the direction of the force but in some other direction (compressing a piece of clay, say, can push it out orthogonally to the direction of the force), so stress is described by a linear transformation, and thus is a rank 2 tensor. The link between stress and strain tensors is the elasticity tensor. The stress and strain tensors are rank 2, so they are elements of $V^{\otimes 2}$ ($V = \mathbf{R}^3$), and the elasticity tensor is a linear map in $\mathrm{Hom}_{\mathbf{R}}(V^{\otimes 2}, V^{\otimes 2}) \cong (V^{\otimes 2})^\vee \otimes_{\mathbf{R}} V^{\otimes 2} \cong V^{\otimes 4}$, so the elasticity tensor has rank 4. Since the stress from an applied force can act in different directions at different points, the stress tensor is not really a single tensor but a varying family of tensors at different points: stress is a tensor field, which is a generalization of a vector field. Tensors in physics usually occur as part of a tensor field, so "tensor" really means "tensor field" in physics. A change of variables between two coordinate systems $\{x_i\}$ and $\{x'_j\}$ in a region of $\mathbf{R}^n$ involves partial derivatives $\frac{\partial x_i}{\partial x'_j}$, and the tensor transformation formula (6.1) shows up in physics with the products $\frac{\partial x_{i_1}}{\partial x'_{j_1}} \cdots \frac{\partial x_{i_k}}{\partial x'_{j_k}}$, which vary from point to point, in the role of $a_{i_1 j_1} \cdots a_{i_k j_k}$.

Quantum mechanics is another part of physics where tensors play a role, but for rather different (non-computational) reasons than we've seen already. In classical mechanics, the states of a system are modeled by the points on a finite-dimensional manifold, and when we combine two systems the corresponding manifold is the direct product of the manifolds for the original two systems. The states of a quantum system, on the other hand, are represented by the nonzero vectors (really, the 1-dimensional subspaces) in an infinite-dimensional Hilbert space, such as $L^2(\mathbf{R}^6)$. (A point in $\mathbf{R}^6$ has three position and three momentum coordinates, which is the classical description of a particle.) When we combine two quantum systems the corresponding Hilbert space is the tensor product of the original two Hilbert spaces, essentially because $L^2(\mathbf{R}^6 \times \mathbf{R}^6) = L^2(\mathbf{R}^6) \otimes_{\mathbf{C}} L^2(\mathbf{R}^6)$, which is the analytic[8] analogue of $R[X,Y] \cong R[X] \otimes_R R[Y]$. While in electromagnetism and relativity the physicist uses specific tensors (*e.g.*, the electromagnetic or metric tensor), in quantum mechanics it is a whole tensor product space $H_1 \otimes_{\mathbf{C}} H_2$ that gets used.

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

We'll end this discussion of tensors in physics with a story. I was the mathematics consultant for the 4th edition of the American Heritage Dictionary of the English Language (2000). The editors sent me all the words in the 3rd edition with mathematical definitions, and I had to correct any errors. Early on I came across a word I had never heard of before: *dyad*. It was defined in the 3rd edition as "an operator represented as a pair of vectors juxtaposed without multiplication." That's a ridiculous definition, as it conveys no meaning at all. I obviously had to fix this definition, but first I had to know what the word meant! In a physics book[9] a dyad is defined as "a pair of vectors, written in a definite order $\mathbf{AB}$." This is just as useless, but the physics book also does something with dyads, which gives a clue about what they really are. The product of a dyad $\mathbf{AB}$ with a vector $\mathbf{C}$ is $\mathbf{A}(\mathbf{B} \cdot \mathbf{C})$, where $\mathbf{B} \cdot \mathbf{C}$ is the usual dot product ($\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$ are all vectors in $\mathbf{R}^n$). This reveals what a dyad is. Do you see it? Dotting with $\mathbf{B}$ is an element of the dual space $(\mathbf{R}^n)^\vee$, so the effect of $\mathbf{AB}$ on $\mathbf{C}$ is reminiscient of the way $V \otimes V^\vee$ acts on V by $(v \otimes \varphi)(w) = \varphi(w)v$. A dyad is really an elementary tensor $v \otimes \varphi$ in $\mathbf{R}^n \otimes (\mathbf{R}^n)^\vee$. In the 4th

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

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

edition of the dictionary, I included two definitions for a dyad. For the general reader, a dyad is "a function that draws a correspondence[10] from any vector $\mathbf{u}$ to the vector $(\mathbf{v}\cdot\mathbf{u})\mathbf{w}$ and is denoted $\mathbf{vw}$, where $\mathbf{v}$ and $\mathbf{w}$ are a fixed pair of vectors and $\mathbf{v}\cdot\mathbf{u}$ is the scalar product of $\mathbf{v}$ and $\mathbf{u}$. For example, if $\mathbf{v} = (2,3,1)$, $\mathbf{w} = (0,-1,4)$, and $\mathbf{u} = (a,b,c)$, then the dyad $\mathbf{vw}$ draws a correspondence from $\mathbf{u}$ to $(2a+3b+c)\mathbf{w}$." The more concise second definition was: a dyad is "a tensor formed from a vector in a vector space and a linear functional on that vector space."

More general than a dyad is a *dyadic*. The physicists define them[11] as a sum of dyads: $\mathbf{AB}+\mathbf{CD}+\dots$. So a dyadic is a general element of $\mathbf{R}^n \otimes_{\mathbf{R}} (\mathbf{R}^n)^\vee \cong \mathrm{Hom}_{\mathbf{R}}(\mathbf{R}^n, \mathbf{R}^n)$. The terminology of dyads and dyadics goes back to Gibbs (1884), who was aiming to develop a coordinate-free way of describing linear transformations. Elementary tensors in the tensor product of 3, 4, or more vector spaces are called triads, tetrads, and more generally polyads.

7. Tensor Product of R-Algebras

Many times our tensor product isomorphisms have involved rings, *e.g.*, $R[X] \otimes_R R[Y] \cong R[X,Y]$, but we only said the isomorphism holds at the level of R-modules since we didn't have a multiplication of tensors. 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 too.

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. The bilinearity of multiplication includes the distributive laws for multiplication over addition as well as the rule $r(ab) = (ra)b = a(rb)$ 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 $R[X_1,\dots,X_n]$. We will assume, as in all these examples, that our 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.[12]

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 $ab := m(a \otimes b)$.

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

[11]H. Goldstein, p. 194

[12]Lie algebras are an important class of nonassociative algebras; they are not rings.

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 R-algebra structure on $A \otimes_R B$ such that*

$$(a \otimes b)(a' \otimes b') = aa' \otimes bb' \tag{7.1}$$

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

Proof. If there is an R-algebra structure on $A \otimes_R B$ such that (7.1) holds 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 structure on $A \otimes_R B$ satisfying (7.1) is unique if it exists at all. Our task now is to write down an R-algebra structure on $A \otimes_R B$ satisfying (7.1).

One way to do this is to start from scratch and try to define what left multiplication by an elementary tensor $a \otimes b$ 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 advantange of various maps we already know between tensor products. Writing down a multiplication function on $A \otimes_R B$ which commutes with R-scaling means writing down an R-linear map $(A \otimes_R B) \otimes_R (A \otimes_R B) \to A \otimes_R B$, and that's what we're going to do. Using the commutativity and associativity isomorphisms on tensor products, there are natural isomorphisms

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

Let $A \otimes_R A \xrightarrow{m_A} A$ and $B \otimes_R B \xrightarrow{m_B} B$ be the linear maps corresponding to multiplication in A and B. Their tensor product is a linear map $(A \otimes_R A) \otimes_R (B \otimes_R B) \xrightarrow{m_A \otimes m_B} A \otimes_R B$, so composing $m_A \otimes m_B$ with the above isomorphism $(A \otimes_R B) \otimes_R (A \otimes_R B) \cong (A \otimes_R A) \otimes_R (B \otimes_R B)$ creates a linear map $(A \otimes_R B) \otimes_R (A \otimes_R B) \to A \otimes_R B$. Tracking the effect of these maps on $(a \otimes b) \otimes (a' \otimes b')$,

$$\begin{aligned} (a \otimes b) \otimes (a' \otimes b') &\mapsto ((a \otimes b) \otimes a') \otimes b' \\ &\mapsto (a \otimes (b \otimes a')) \otimes b' \\ &\mapsto (a \otimes (a' \otimes b)) \otimes b' \\ &\mapsto ((a \otimes a') \otimes b) \otimes b' \\ &\mapsto (a \otimes a') \otimes (b \otimes b') \\ &\mapsto aa' \otimes bb', \end{aligned}$$

where $m_A \otimes m_B$ is used in the last step. This R-linear map $(A\otimes_R B)\otimes_R (A\otimes_R B) \to A\otimes_R B$ can be pulled 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 $A \otimes_R B$.

To prove $1\otimes 1$ is an identity and that multiplication in $A\otimes_R B$ is associative, the equations

$$(1 \otimes 1)t' = t', \quad t(1 \otimes 1) = t, \quad (t_1t_2)t_3 = t_1(t_2t_3)$$

involving general tensors t', t_1, t_2, and t_3 are additive in each tensor appearing on both sides, so verifying these equations reduces to computations with elementary tensors, and those can be checked directly. $\square$

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 (7.1) since A and B are commutative. $\square$

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 tensor product $A \otimes_R B$ and the direct product $A \times B$ (componentwise addition and multiplication, with $r(a, b) = (ra, rb)$). Both are R-algebras. The most basic difference is that 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 $A \times B$ are surjective, but the maps to the tensor product $A \otimes_R B$ 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.

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

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

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 7.3. 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 7.4. There is an R-module isomorphism $\varphi\colon R[X] \otimes_R R[Y] \longrightarrow R[X,Y]$ where $\varphi(f(X) \otimes g(Y)) = f(X)g(Y)$. Then $\varphi(1 \otimes 1) = 1$ 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 using monomial tensors $X^i \otimes Y^j$.)

In a similar way, $R[X]^{\otimes k} \cong R[X_1, \dots, X_n]$ as R-algebras. 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.

Example 7.5. 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 R-algebra isomorphism.

Theorem 7.6. *Let A, B, and C be R-algebras. The standard R-module isomorphisms*

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

are all R-algebra isomorphisms.

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. □

Corollary 7.7. *For R-algebras A and B, $A \otimes_R B^n \cong (A \otimes_R B)^n$ as R-algebras.*

Proof. Induct on n. □

We turn now to base extensions. Fix a homomorphism $f\colon R \to S$ 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 7.8. *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. □

Theorem 7.9. *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 7.8. Therefore it is an S-algebra homomorphism. □

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

In one-variable, with $I = (h(X))$ a principal ideal in $R[X]$,[13] Example 7.10 gives us an S-algebra isomorphism

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

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*. For instance,

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

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 $X^3 - 2$ is a linear times a quadratic in $\mathbf{R}[X]$. And

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

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* $z \otimes w \mapsto (z, w)$; that's not additive or even a bijection, since $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C}$ contains non-elementary tensors. 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 sometimes it can be. For instance,

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

which is a field because $X^2 - 3$ is irreducible in $\mathbf{Q}(\sqrt{2})[X]$. 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 $(\mathbf{Z}/5\mathbf{Z})[X]$.

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

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.

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 it is R/I-linear too.

Theorem 7.13. *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 7.13 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.

$$\begin{tikzcd}
 & & & L \\
F & & E \arrow[ur, dash] & \\
 & K \arrow[ul, dash] \arrow[ur, dash] & &
\end{tikzcd}$$

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 7.14. *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. $\square$