# TENSOR PRODUCTS II 

KEITH CONRAD

## 1. Introduction

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

## 2. Tensor Products of Linear Maps

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

Start with the map $M \times N \rightarrow M^{\prime} \otimes_{R} N^{\prime}$ 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 $M \otimes_{R} N \xrightarrow{\varphi \otimes \psi} M^{\prime} \otimes_{R} N^{\prime}$ where $(\varphi \otimes \psi)(m \otimes n)=\varphi(m) \otimes \psi(n)$, and more generally

$$
(\varphi \otimes \psi)\left(m_{1} \otimes n_{1}+\cdots+m_{k} \otimes n_{k}\right)=\varphi\left(m_{1}\right) \otimes \psi\left(n_{1}\right)+\cdots+\varphi\left(m_{k}\right) \otimes \psi\left(n_{k}\right) .
$$

We call $\varphi \otimes \psi$ the tensor product of $\varphi$ and $\psi$, 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 $M \xrightarrow{\varphi} M^{\prime}$ is linear, the linear maps $N \otimes_{R} M \xrightarrow{1 \otimes \varphi} N \otimes_{R} M^{\prime}$ or $M \otimes_{R} N \xrightarrow{\varphi \otimes 1}$ $M^{\prime} \otimes_{R} N$ 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 $(i \otimes 1)(a x \otimes y \bmod b)=a x \otimes y \bmod b$. 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 $a \mathbf{Z} \otimes \mathbf{Z} \mathbf{Z} / b \mathbf{Z} \rightarrow \mathbf{Z} / b \mathbf{Z}$ where $a x \otimes y \bmod b \mapsto a x y \bmod b$. Its image is $\{a z \bmod b: z \in \mathbf{Z} / b \mathbf{Z}\}$, which is $d \mathbf{Z} / b \mathbf{Z}$ where $d=(a, b)$; this is 0 if $b \mid a$ and is $\mathbf{Z} / b \mathbf{Z}$ if $(a, b)=1$.

Example 2.2. Let $A=\left(\begin{array}{ll}a & b \\ c & d\end{array}\right)$ and $A^{\prime}=\left(\begin{array}{cc}a^{\prime} & b^{\prime} \\ c^{\prime} & d^{\prime}\end{array}\right)$ in $\mathrm{M}_{2}(R)$. Then $A$ and $A^{\prime}$ are both linear maps $R^{2} \rightarrow R^{2}$, so $A \otimes A^{\prime}$ is a linear map from $\left(R^{2}\right)^{\otimes 2}=R^{2} \otimes_{R} R^{2}$ back to itself. Writing $e_{1}$ and $e_{2}$ for the standard basis vectors of $R^{2}$, let's compute the matrix for $A \otimes A^{\prime}$ on $\left(R^{2}\right)^{\otimes 2}$
with respect to the basis $\left\{e_{1} \otimes e_{1}, e_{1} \otimes e_{2}, e_{2} \otimes e_{1}, e_{2} \otimes e_{2}\right\}$. By definition,

$$
\begin{aligned}
\left(A \otimes A^{\prime}\right)\left(e_{1} \otimes e_{1}\right) & =A e_{1} \otimes A^{\prime} e_{1} \\
& =\left(a e_{1}+c e_{2}\right) \otimes\left(a^{\prime} e_{1}+c^{\prime} e_{2}\right) \\
& =a a^{\prime} e_{1} \otimes e_{1}+a c^{\prime} e_{1} \otimes e_{2}+c a^{\prime} e_{2} \otimes e_{1}+c c^{\prime} e_{2} \otimes e_{2} \\
\left(A \otimes A^{\prime}\right)\left(e_{1} \otimes e_{2}\right) & =A e_{1} \otimes A^{\prime} e_{2} \\
& =\left(a e_{1}+c e_{2}\right) \otimes\left(b^{\prime} e_{1}+d^{\prime} e_{2}\right) \\
& =c b^{\prime} e_{1} \otimes e_{1}+a d^{\prime} e_{1} \otimes e_{2}+c b^{\prime} e_{2} \otimes e_{2}+c d^{\prime} e_{2} \otimes e_{2}
\end{aligned}
$$

and similarly

$$
\begin{aligned}
& \left(A \otimes A^{\prime}\right)\left(e_{2} \otimes e_{1}\right)=b a^{\prime} e_{1} \otimes e_{1}+b c^{\prime} e_{1} \otimes e_{2}+d a^{\prime} e_{2} \otimes e_{1}+d c^{\prime} e_{2} \otimes e_{2}, \\
& \left(A \otimes A^{\prime}\right)\left(e_{2} \otimes e_{2}\right)=b b^{\prime} e_{1} \otimes e_{1}+b d^{\prime} e_{1} \otimes e_{2}+d b^{\prime} e_{2} \otimes e_{1}+d d^{\prime} e_{2} \otimes e_{2} .
\end{aligned}
$$

Therefore the matrix for $A \otimes A^{\prime}$ is

$$
\left(\begin{array}{cccc}
a a^{\prime} & a b^{\prime} & b a^{\prime} & b b^{\prime} \\
a c^{\prime} & a d^{\prime} & b c^{\prime} & b d^{\prime} \\
c a^{\prime} & c b^{\prime} & d a^{\prime} & d b^{\prime} \\
c c^{\prime} & c d^{\prime} & d c^{\prime} & d d^{\prime}
\end{array}\right)=\left(\begin{array}{l|l}
a A^{\prime} & b A^{\prime} \\
\hline c A^{\prime} & d A^{\prime}
\end{array}\right) .
$$

So $\operatorname{Tr}\left(A \otimes A^{\prime}\right)=a\left(a^{\prime}+d^{\prime}\right)+d\left(a^{\prime}+d^{\prime}\right)=(a+d)\left(a^{\prime}+d^{\prime}\right)=(\operatorname{Tr} A)\left(\operatorname{Tr} A^{\prime}\right)$, and $\operatorname{det}\left(A \otimes A^{\prime}\right)$ 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 $A^{\prime} \in \mathrm{M}_{n^{\prime}}(R)$ then the matrix for $A \otimes A^{\prime}$ with respect to the standard basis for $R^{n} \otimes_{R} R^{n^{\prime}}$ is the block matrix ( $a_{i j} A^{\prime}$ ) where $A=\left(a_{i j}\right)$. This $n n^{\prime} \times n n^{\prime}$ matrix is called the Kronecker product of $A$ and $A^{\prime}$, and is not symmetric in the roles of $A$ and $A^{\prime}$ in general (just as $A \otimes A^{\prime} \neq A^{\prime} \otimes A$ in general). In particular, $I_{n} \otimes A^{\prime}$ has block matrix representation $\left(\delta_{i j} A^{\prime}\right)$, whose determinant is $\left(\operatorname{det} A^{\prime}\right)^{n}$.

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 $\lambda$ and $\mu$ 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 $A v=\lambda v$ and $B w=\mu w$ for some $v \in K^{m}$ and $w \in K^{n}$. Then

$$
\begin{aligned}
\left(A \otimes I_{n}+I_{m} \otimes B\right)(v \otimes w) & =A v \otimes w+v \otimes B w \\
& =\lambda v \otimes w+v \otimes \mu w \\
& =(\lambda+\mu)(v \otimes w)
\end{aligned}
$$

and

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

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

$$
\begin{aligned}
A \otimes I_{2}+I_{2} \otimes B & =\left(\begin{array}{llll}
0 & 0 & 2 & 0 \\
0 & 0 & 0 & 2 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{array}\right)+\left(\begin{array}{llll}
0 & 3 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 3 \\
0 & 0 & 1 & 0
\end{array}\right) \\
& =\left(\begin{array}{llll}
0 & 3 & 2 & 0 \\
1 & 0 & 0 & 2 \\
1 & 0 & 0 & 3 \\
0 & 1 & 1 & 0
\end{array}\right),
\end{aligned}
$$

whose characteristic polynomial is $T^{4}-10 T^{2}+1$. 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, $\varphi$ and $\psi$ belong to the $R$-modules $\operatorname{Hom}_{R}\left(M, M^{\prime}\right)$ and $\operatorname{Hom}_{R}\left(N, N^{\prime}\right)$, so one could ask if the actual elementary tensor $\varphi \otimes \psi$ in $\operatorname{Hom}_{R}\left(M, M^{\prime}\right) \otimes_{R} \operatorname{Hom}_{R}\left(N, N^{\prime}\right)$ is related to the linear map $\varphi \otimes \psi: M \otimes_{R} N \rightarrow M^{\prime} \otimes_{R} N^{\prime}$.

Theorem 2.5. There is a linear map

$$
\operatorname{Hom}_{R}\left(M, M^{\prime}\right) \otimes_{R} \operatorname{Hom}_{R}\left(N, N^{\prime}\right) \rightarrow \operatorname{Hom}_{R}\left(M \otimes_{R} N, M^{\prime} \otimes_{R} N^{\prime}\right)
$$

that sends the elementary tensor $\varphi \otimes \psi$ to the linear map $\varphi \otimes \psi$. When $M, M^{\prime}, N$, and $N^{\prime}$ 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 $T(\varphi, \psi): M \otimes_{R} N \rightarrow M^{\prime} \otimes_{R} N^{\prime}$ is the linear map sending every $m \otimes n$ to $\varphi(m) \otimes \psi(n)$.

Define $\operatorname{Hom}_{R}\left(M, M^{\prime}\right) \times \operatorname{Hom}_{R}\left(N, N^{\prime}\right) \rightarrow \operatorname{Hom}_{R}\left(M \otimes_{R} N, M^{\prime} \otimes_{R} N^{\prime}\right)$ by $(\varphi, \psi) \mapsto T(\varphi, \psi)$. This is $R$-bilinear. For example, to show $T(r \varphi, \psi)=r T(\varphi, \psi)$, both sides are linear maps so to prove they are equal it suffices to check they are equal at the elementary tensors in $M \otimes_{R} N:$
$T(r \varphi, \psi)(m \otimes n)=(r \varphi)(m) \otimes \psi(n)=r \varphi(m) \otimes \psi(n)=r(\varphi(m) \otimes \psi(n))=r T(\varphi, \psi)(m \otimes n)$.
The other bilinearity conditions are left to the reader.
From the universal mapping property of tensor products, there is a unique $R$-linear map $\operatorname{Hom}_{R}\left(M, M^{\prime}\right) \otimes_{R} \operatorname{Hom}_{R}\left(N, N^{\prime}\right) \rightarrow \operatorname{Hom}_{R}\left(M \otimes_{R} N, M^{\prime} \otimes_{R} N^{\prime}\right)$ where $\varphi \otimes \psi \mapsto T(\varphi, \psi)$.

Suppose $M, M^{\prime}, N$, and $N^{\prime}$ are all finite free $R$-modules. Let them have respective bases $\left\{e_{i}\right\},\left\{e_{i^{\prime}}^{\prime}\right\},\left\{f_{j}\right\}$, and $\left\{f_{j^{\prime}}^{\prime}\right\}$. Then $\operatorname{Hom}_{R}\left(M, M^{\prime}\right)$ and $\operatorname{Hom}_{R}\left(N, N^{\prime}\right)$ are both free with bases $\left\{E_{i^{\prime} i}\right\}$ and $\left\{\widetilde{E}_{j^{\prime} j}\right\}$, where $E_{i^{\prime} i}: M \rightarrow M^{\prime}$ is the linear map sending $e_{i}$ to $e_{i^{\prime}}^{\prime}$ and is 0 at other basis vectors of $M$, and $\widetilde{E}_{j^{\prime} j}: N \rightarrow N^{\prime}$ is defined similarly. (The matrix representation of $E_{i^{\prime} i}$ with respect to the chosen bases of $M$ and $M^{\prime}$ has a 1 in the $\left(i^{\prime}, i\right)$ position and 0 elsewhere, thus justifying the notation.) A basis of $\operatorname{Hom}_{R}\left(M, M^{\prime}\right) \otimes_{R} \operatorname{Hom}_{R}\left(N, N^{\prime}\right)$ is
$\left\{E_{i^{\prime} i} \otimes \widetilde{E}_{j^{\prime} j}\right\}$ and $T\left(E_{i^{\prime} i} \otimes \widetilde{E}_{j^{\prime} j}\right): M \otimes_{R} N \rightarrow M^{\prime} \otimes_{R} N^{\prime}$ has the effect

$$
\begin{aligned}
T\left(E_{i^{\prime} i} \otimes \widetilde{E}_{j^{\prime} j}\right)\left(e_{\mu} \otimes f_{\nu}\right) & =E_{i^{\prime} i}\left(e_{\mu}\right) \otimes \widetilde{E}_{j^{\prime} j}\left(f_{\nu}\right) \\
& =\delta_{\mu i} e_{i^{\prime}}^{\prime} \otimes \delta_{\nu j} f_{j^{\prime}}^{\prime} \\
& = \begin{cases}e_{i^{\prime}}^{\prime} \otimes f_{j^{\prime}}^{\prime}, & \text { if } \mu=i \text { and } \nu=j, \\
0, & \text { otherwise },\end{cases}
\end{aligned}
$$

so $T\left(E_{i^{\prime} i} \otimes E_{j^{\prime} j}\right)$ sends $e_{i} \otimes f_{j}$ to $e_{i^{\prime}}^{\prime} \otimes f_{j^{\prime}}^{\prime}$ and sends other members of the basis of $M \otimes_{R} N$ to 0 . That means the linear map $\operatorname{Hom}_{R}\left(M, M^{\prime}\right) \otimes_{R} \operatorname{Hom}_{R}\left(N, N^{\prime}\right) \rightarrow \operatorname{Hom}_{R}\left(M \otimes_{R} N, M^{\prime} \otimes_{R} N^{\prime}\right)$ 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 $\operatorname{Hom}_{R}\left(M, M^{\prime}\right) \otimes_{R} \operatorname{Hom}_{R}\left(N, N^{\prime}\right)$ naturally acts as linear maps $M \otimes_{R} N \rightarrow M^{\prime} \otimes_{R} N^{\prime}$ 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 $\varphi_{1}: M_{1} \rightarrow N_{1}, \varphi_{2}: M_{2} \rightarrow$ $N_{2}$, and $\varphi_{3}: M_{3} \rightarrow N_{3}$ are linear maps, we have $\varphi_{1} \otimes\left(\varphi_{2} \oplus \varphi_{3}\right)=\left(\varphi_{1} \otimes \varphi_{2}\right) \oplus\left(\varphi_{1} \otimes \varphi_{3}\right)$ and $\left(\varphi_{1} \otimes \varphi_{2}\right) \otimes \varphi_{3}=\varphi_{1} \otimes\left(\varphi_{2} \otimes \varphi_{3}\right)$, in the sense that the diagrams

$$
\begin{gathered}
M_{1} \otimes_{R}\left(M_{2} \oplus M_{3}\right) \xrightarrow{\varphi_{1} \otimes\left(\varphi_{2} \oplus \varphi_{3}\right)} N_{1} \otimes_{R}\left(N_{2} \oplus N_{3}\right) \\
\left.\downarrow \downarrow M_{1} \otimes_{R} M_{2}\right) \oplus\left(M_{1} \otimes_{R} M_{3}\right) \xrightarrow{\left(\varphi_{1} \otimes \varphi_{2}\right) \oplus\left(\varphi_{1} \otimes \varphi_{3}\right)}\left(N_{1} \otimes_{R} N_{2}\right) \oplus\left(N_{1} \otimes_{R} N_{3}\right)
\end{gathered}
$$

and

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 $N, \operatorname{id}_{M} \otimes \operatorname{id}_{N}=\operatorname{id}_{M \otimes_{R} N}$. For linear maps $M \xrightarrow{\varphi}$ $M^{\prime}, M^{\prime} \xrightarrow{\varphi^{\prime}} M^{\prime \prime}, N \xrightarrow{\psi} N^{\prime}$, and $N^{\prime} \xrightarrow{\psi^{\prime}} N^{\prime \prime}$,

$$
\left(\varphi^{\prime} \otimes \psi^{\prime}\right) \circ(\varphi \otimes \psi)=\left(\varphi^{\prime} \circ \varphi\right) \otimes\left(\psi^{\prime} \circ \psi\right)
$$

as linear maps from $M \otimes_{R} N$ to $M^{\prime \prime} \otimes_{R} N^{\prime \prime}$.
Proof. The function $\operatorname{id}_{M} \otimes \operatorname{id}_{N}$ is a linear map from $M \otimes_{R} N$ to itself that fixes every elementary tensor, so it fixes all tensors.

Since $\left(\varphi^{\prime} \otimes \psi^{\prime}\right) \circ(\varphi \otimes \psi)$ and $\left(\varphi^{\prime} \circ \varphi\right) \otimes\left(\psi^{\prime} \circ \psi\right)$ 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 $\varphi^{\prime}(\varphi(m)) \otimes \psi^{\prime}(\psi(n))$.

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 $\ell$. For linear operators $M \xrightarrow{\varphi} M$ and $N \xrightarrow{\psi} N$, we will compute $\operatorname{det}(\varphi \otimes \psi)$ by breaking up $\varphi \otimes \psi$ into a composite of two maps $M \otimes_{R} N \rightarrow M \otimes_{R} N$ :

$$
\varphi \otimes \psi=\left(\varphi \otimes \mathrm{id}_{N}\right) \circ\left(\mathrm{id}_{M} \otimes \psi\right)
$$

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

$$
\operatorname{det}(\varphi \otimes \psi)=\operatorname{det}\left(\operatorname{id}_{N} \otimes \varphi\right) \operatorname{det}\left(\operatorname{id}_{M} \otimes \psi\right)
$$

What are the determinants on the right side? Pick bases $e_{1}, \ldots, e_{k}$ of $M$ and $e_{1}^{\prime}, \ldots, e_{\ell}^{\prime}$ of $N$. We will use the $k \ell$ elementary tensors $e_{i} \otimes e_{j}^{\prime}$ as a bases of $M \otimes_{R} N$. Let [ $\varphi$ ] be the matrix of $\varphi$ in the ordered basis $e_{1}, \ldots, e_{k}$. Since $\left(\varphi \otimes \operatorname{id}_{N}\right)\left(e_{i} \otimes e_{j}^{\prime}\right)=\varphi\left(e_{i}\right) \otimes e_{j}^{\prime}$, let's order the basis of $M \otimes_{R} N$ as

$$
e_{1} \otimes e_{1}^{\prime}, \ldots, e_{k} \otimes e_{1}^{\prime}, \ldots, e_{1} \otimes e_{\ell}^{\prime}, \ldots, e_{k} \otimes e_{\ell}^{\prime}
$$

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

$$
\left(\begin{array}{cccc}
{[\varphi]} & O & \cdots & O \\
O & {[\varphi]} & \cdots & O \\
\vdots & \vdots & \ddots & \vdots \\
O & O & \cdots & {[\varphi]}
\end{array}\right)
$$

whose determinant is $(\operatorname{det} \varphi)^{\ell}$.
Thus

$$
\begin{equation*}
\operatorname{det}(\varphi \otimes \psi)=(\operatorname{det} \varphi)^{\ell}(\operatorname{det} \psi)^{k} \tag{2.1}
\end{equation*}
$$

Note $\ell$ is the rank of the module on which $\psi$ is defined and $k$ is the rank of the module on which $\varphi$ is defined. In particular, in Example 2.2 we have $\operatorname{det}\left(A \otimes A^{\prime}\right)=(\operatorname{det} A)^{2}\left(\operatorname{det} A^{\prime}\right)^{2}$.

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, $\varphi \otimes \operatorname{id}_{N}$ becomes the $\ell$-fold direct sum $\varphi \oplus \cdots \oplus \varphi$, which has a block diagonal matrix representation in a suitable basis. So its determinant is $(\operatorname{det} \varphi)^{\ell}$.
Example 2.8. Taking $M=N$ and $\varphi=\psi$, the tensor square $\varphi^{\otimes 2}$ has determinant $(\operatorname{det} \varphi)^{2 k}$.
Corollary 2.9. Let $M$ be a free module of rank $k \geq 1$ and $\varphi: M \rightarrow M$ be a linear map. For every $i \geq 1, \operatorname{det}\left(\varphi^{\otimes i}\right)=(\operatorname{det} \varphi)^{i k^{i-1}}$.
Proof. Use induction and associativity of the tensor product of linear maps.
Remark 2.10. Equation (2.1) in the setting of vector spaces and matrices says $\operatorname{det}(A \otimes B)=$ $(\operatorname{det} A)^{\ell}(\operatorname{det} B)^{k}$, where $A$ is $k \times k, B$ is $\ell \times \ell$, and $A \otimes B=\left(a_{i j} B\right)$ is the matrix incarnation of a tensor product of linear maps, called the Kronecker product of $A$ and $B$ at the end of Example 2.2. While the label "Kronecker product" for the matrix $A \otimes B$ is completely standard, it is not historically accurate. It is based on Hensel's attribution of the formula for $\operatorname{det}(A \otimes B)$ to Kronecker, but the formula is due to Zehfuss. See [2].

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

Theorem 2.11. If $\varphi: M \rightarrow M^{\prime}$ and $\psi: N \rightarrow N^{\prime}$ 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 2.12. If $\varphi: M \rightarrow M^{\prime}$ and $\psi: N \rightarrow N^{\prime}$ 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 $M^{\prime} \otimes_{R} N^{\prime}$ is in the image. For such an elementary tensor $m^{\prime} \otimes n^{\prime}$, we can write $m^{\prime}=\varphi(m)$ and $n^{\prime}=\psi(n)$ since $\varphi$ and $\psi$ are onto. Therefore $m^{\prime} \otimes n^{\prime}=\varphi(m) \otimes \psi(n)=(\varphi \otimes \psi)(m \otimes n)$.

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

Example 2.13. Taking $R=\mathbf{Z}$, let $\alpha: \mathbf{Z} / p \mathbf{Z} \rightarrow \mathbf{Z} / p^{2} \mathbf{Z}$ be multiplication by $p: \alpha(x)=p x$. This is injective, and if we tensor with $\mathbf{Z} / p \mathbf{Z}$ we get the linear map $1 \otimes \alpha: \mathbf{Z} / p \mathbf{Z} \otimes \mathbf{z} \mathbf{Z} / p \mathbf{Z} \rightarrow$ $\mathbf{Z} / p \mathbf{Z} \otimes \mathbf{Z} \mathbf{Z} / p^{2} \mathbf{Z}$ with the effect $a \otimes x \mapsto a \otimes p x=p a \otimes x=0$, 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

$$
\operatorname{Hom}_{R}\left(M, M^{\prime}\right) \otimes_{R} \operatorname{Hom}_{R}\left(N, N^{\prime}\right) \rightarrow \operatorname{Hom}_{R}\left(M \otimes_{R} N, M^{\prime} \otimes_{R} N^{\prime}\right)
$$

in Theorem 2.5 is not an isomorphism; $R=\mathbf{Z}, M=M^{\prime}=N=\mathbf{Z} / p \mathbf{Z}$, and $N^{\prime}=\mathbf{Z} / p^{2} \mathbf{Z}$.
Because the tensor product of linear maps does not generally preserve injectivity, a tensor has to be understood in context: it is a tensor in a specific tensor product module $M \otimes_{R} N$. If $M \subset M^{\prime}$ and $N \subset N^{\prime}$, it is generally false that $M \otimes_{R} N$ can be thought of as a submodule of $M^{\prime} \otimes_{R} N^{\prime}$ since the natural map $M \otimes_{R} N \rightarrow M^{\prime} \otimes_{R} N^{\prime}$ may not be injective. We can say it this way: a tensor product of submodules need not be a submodule.

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

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

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

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

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

Darij Grinberg ${ }^{1}$ gave a nice clarification of this confusing situation: it is due to an abuse of notation when writing elementary tensors as $m \otimes n$ without indicating the ambient modules that $m$ and $n$ lie in. Including those modules in the notation, say by writing $m \otimes n$ as ( $m, M$ ) $\otimes_{R}(n, N)$, would avoid the confusion at the cost of very clumsy notation.
Example 2.16. Here is an example of a linear map $f: M \rightarrow N$ that is injective and its tensor square $f^{\otimes 2}: M^{\otimes 2} \rightarrow N^{\otimes 2}$ is not injective.

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

$$
\begin{equation*}
X \otimes Y=X Y(1 \otimes 1) \text { and } Y \otimes X=Y X(1 \otimes 1)=X Y(1 \otimes 1) \tag{2.2}
\end{equation*}
$$

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

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

Generalizing this, let $R=A\left[X_{1}, \ldots, X_{n}\right]$ where $n \geq 2$ and $I=\left(X_{1}, \ldots, X_{n}\right)$. The inclusion $i: I \hookrightarrow R$ is injective but the $n$th tensor power (as $R$-modules) $i^{\otimes n}: I^{\otimes n} \rightarrow R^{\otimes n}$ is not injective because the tensor

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

gets sent to $\sum_{\sigma \in S_{n}}(\operatorname{sign} \sigma) X_{1} \cdots X_{n}(1 \otimes \cdots \otimes 1)$ in $R^{\otimes n}$, which is 0 , but $t$ is not 0 in $I^{\otimes n}$ because there is an $R$-linear map $I^{\otimes n} \rightarrow A$ sending $t$ to 1 : use a product of partial derivatives at $(0,0, \ldots, 0)$, as in the $n=2$ case.
Remark 2.17. The ideal $I=(X, Y)$ in $R=A[X, Y]$ from Example 2.16 has another interesting feature when $A$ is a domain: it is a torsion-free $R$-module but $I^{\otimes 2}$ is not: $X(X \otimes Y)=X \otimes X Y=Y(X \otimes X)$ and $X(Y \otimes X)=X Y \otimes X=Y(X \otimes X)$, so in $I^{\otimes 2}$ we have $X(X \otimes Y-Y \otimes X)=0$, but $X \otimes Y-Y \otimes X \neq 0$. Similarly, $Y(X \otimes Y-Y \otimes X)=0$. Therefore a tensor product of torsion-free modules (even over a domain) need not be torsionfree.

While we have just seen a tensor power of an injective linear map need not be injective, here is a condition where injectivity holds.
Theorem 2.18. Let $\varphi: M \rightarrow N$ be injective and $\varphi(M)$ be a direct summand of $N$. For $k \geq 0, \varphi^{\otimes k}: M^{\otimes k} \rightarrow N^{\otimes k}$ is injective and the image is a direct summand of $N^{\otimes k}$.

[^0]Proof. Write $N=\varphi(M) \oplus P$. Let $\pi: N \rightarrow M$ by $\pi(\varphi(m)+p)=m$, so $\pi$ is linear and $\pi \circ \varphi=\operatorname{id}_{M}$. Then $\varphi^{\otimes k}: M^{\otimes k} \rightarrow N^{\otimes k}$ and $\pi^{\otimes k}: N^{\otimes k} \rightarrow M^{\otimes k}$ 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.

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 $W \rightarrow V$ are injective linear maps $W^{\otimes k} \rightarrow V^{\otimes k}$.

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

We will now compute the kernel of $M \otimes_{R} N \xrightarrow{\varphi \otimes \psi} M^{\prime} \otimes_{R} N^{\prime}$ in terms of the kernels of $\varphi$ and $\psi$, assuming $\varphi$ and $\psi$ are onto.

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

$$
\operatorname{ker}(\varphi \otimes \psi)=(\operatorname{ker} \varphi) \otimes_{R} N+M \otimes_{R}(\operatorname{ker} \psi),
$$

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

$$
\operatorname{ker}(\varphi \otimes \psi)=(i \otimes 1)\left((\operatorname{ker} \varphi) \otimes_{R} N\right)+(1 \otimes j)\left(M \otimes_{R}(\operatorname{ker} \psi)\right)
$$

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

Proof. Both $(i \otimes 1)\left((\operatorname{ker} \varphi) \otimes_{R} N\right)$ and $(1 \otimes j)\left(M \otimes_{R}(\operatorname{ker} \psi)\right)$ are killed by $\varphi \otimes \psi$ : if $m \in \operatorname{ker} \varphi$ and $n \in N$ then $(\varphi \otimes \psi)((i \otimes 1)(m \otimes n))=(\varphi \otimes \psi)(m \otimes n)=\varphi(m) \otimes \psi(n)=0$ since $^{2}$ $\varphi(m)=0$. Similarly $(1 \otimes j)(m \otimes n)$ is killed by $\varphi \otimes \psi$ if $m \in M$ and $n \in \operatorname{ker} \psi$. Set

$$
U=(i \otimes 1)\left((\operatorname{ker} \varphi) \otimes_{R} N\right)+(1 \otimes j)\left(M \otimes_{R}(\operatorname{ker} \psi)\right),
$$

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

$$
\Phi:\left(M \otimes_{R} N\right) / U \rightarrow M^{\prime} \otimes_{R} N^{\prime}
$$

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 $\Phi$ is injective, so the kernel of $\varphi \otimes \psi$ is $U$.

Because $\varphi$ and $\psi$ are assumed to be onto, every elementary tensor in $M^{\prime} \otimes_{R} N^{\prime}$ 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 $\operatorname{ker} \varphi$ and $\operatorname{ker} \psi$. For $m^{\prime} \in \operatorname{ker} \varphi$ and $n^{\prime} \in \operatorname{ker} \psi$,

$$
\left(m+m^{\prime}\right) \otimes\left(n+n^{\prime}\right)=m \otimes n+m^{\prime} \otimes n+m \otimes n^{\prime}+m^{\prime} \otimes n^{\prime} \in m \otimes n+U
$$

so the function $M^{\prime} \times N^{\prime} \rightarrow\left(M \otimes_{R} N\right) / U$ defined by $(\varphi(m), \psi(n)) \mapsto m \otimes n \bmod U$ is welldefined. It is $R$-bilinear, so we have an $R$-linear map $\Psi: M^{\prime} \otimes_{R} N^{\prime} \rightarrow\left(M \otimes_{R} N\right) / U$ 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.

[^1]Remark 2.20. If we remove the assumption that $\varphi$ and $\psi$ are onto, Theorem 2.19 does not correctly compute the kernel. For example, if $\varphi$ and $\psi$ are both injective then the formula for the kernel in Theorem 2.19 is 0 , and we know $\varphi \otimes \psi$ need not be injective.

Unlike the kernel computation in Theorem 2.19, it is not easy to describe the torsion submodule of a tensor product in terms of the torsion submodules of the original modules. While $\left(M \otimes_{R} N\right)_{\text {tor }}$ contains $(i \otimes 1)\left(M_{\text {tor }} \otimes_{R} N\right)+(1 \otimes j)\left(M \otimes_{R} N_{\text {tor }}\right)$, with $i: M_{\text {tor }} \rightarrow M$ and $j: N_{\text {tor }} \rightarrow N$ being the inclusions, it is not true that this is all of $\left(M \otimes_{R} N\right)_{\text {tor }}$, since $M \otimes_{R} N$ can have nonzero torsion when $M$ and $N$ are torsion-free (so $M_{\text {tor }}=0$ and $N_{\text {tor }}=0$ ). We saw this at the end of Example 2.16.
Corollary 2.21. If $M \xrightarrow{\varphi} M^{\prime}$ is an isomorphism of $R$-modules and $N \xrightarrow{\psi} N^{\prime}$ is surjective, then the linear map $M \otimes_{R} N \xrightarrow{\varphi \otimes \psi} M^{\prime} \otimes_{R} N^{\prime}$ has kernel $(1 \otimes j)\left(M \otimes_{R}(\operatorname{ker} \psi)\right)$, where $\operatorname{ker} \psi \xrightarrow{j} N$ is the inclusion.

Proof. This is immediate from Theorem 2.19 since $\operatorname{ker} \varphi=0$.
Corollary 2.22. Let $f: R \rightarrow S$ 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 $\operatorname{ker} \pi=M$, so $\operatorname{ker}(1 \otimes \pi)=$ $(1 \otimes i)\left(S \otimes_{R} M\right)$. Therefore $1 \otimes i$ is onto if and only if $\operatorname{ker}(1 \otimes \pi)=S \otimes_{R} N$ if and only if $1 \otimes \pi=0$, and since $1 \otimes \pi$ is onto we have $1 \otimes \pi=0$ if and only if $S \otimes_{R}(N / M)=0$.
Example 2.23. 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 2.22. Since $N$ is finitely generated, so is $N / M$, and we are reduced to showing $R / \mathfrak{m} \otimes_{R}(N / M)=0$ for all maximal ideals $\mathfrak{m}$ if and only if $N / M=0$. When $P$ is a finitely generated module, $P=0$ if and only if $P / \mathfrak{m} P=0$ for all maximal ideals ${ }^{3} \mathfrak{m}$ in $R$, so we can apply this to $P=N / M$ since $P / \mathfrak{m} P \cong R / \mathfrak{m} \otimes_{R} P$.

Corollary 2.24. Let $f: R \rightarrow S$ be a homomorphism of commutative rings and $I$ be an ideal in $R\left[X_{1}, \ldots, X_{n}\right]$. Write $I \cdot S\left[X_{1}, \ldots, X_{n}\right]$ for the ideal generated by the image of $I$ in $S\left[X_{1}, \ldots, X_{n}\right]$. Then

$$
S \otimes_{R} R\left[X_{1}, \ldots, X_{n}\right] / I \cong S\left[X_{1}, \ldots, X_{n}\right] /\left(I \cdot S\left[X_{1}, \ldots, X_{n}\right]\right) .
$$

as $S$-modules by $s \otimes h \bmod I \mapsto s h \bmod I \cdot S\left[X_{1}, \ldots, X_{n}\right]$.
Proof. The identity $S \rightarrow S$ and the natural reduction $R\left[X_{1}, \ldots, X_{n}\right] \rightarrow R\left[X_{1}, \ldots, X_{n}\right] / I$ are both onto, so the tensor product of these $R$-linear maps is an $R$-linear surjection

$$
\begin{equation*}
S \otimes_{R} R\left[X_{1}, \ldots, X_{n}\right] \rightarrow S \otimes_{R}\left(R\left[X_{1}, \ldots, X_{n}\right] / I\right) \tag{2.3}
\end{equation*}
$$

[^2]and the kernel is $(1 \otimes j)\left(S \otimes_{R} I\right)$ by Theorem 2.19, where $j: I \rightarrow R\left[X_{1}, \ldots, X_{n}\right]$ is the inclusion. Under the natural $R$-module isomorphism
\[

$$
\begin{equation*}
S \otimes_{R} R\left[X_{1}, \ldots, X_{n}\right] \cong S\left[X_{1}, \ldots, X_{n}\right] \tag{2.4}
\end{equation*}
$$

\]

$(1 \otimes j)\left(S \otimes_{R} I\right)$ on the left side corresponds to $I \cdot S\left[X_{1}, \ldots, X_{n}\right]$ on the right side, so (2.3) and (2.4) say

$$
S\left[X_{1}, \ldots, X_{n}\right] /\left(I \cdot S\left[X_{1}, \ldots, X_{n}\right]\right) \cong S \otimes_{R}\left(R\left[X_{1}, \ldots, X_{n}\right] / I\right) .
$$

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

## 3. Flat Modules

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

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

Theorem 3.2. Any free $R$-module $F$ is flat: if the linear map $\varphi: M \rightarrow M^{\prime}$ is injective, then $1 \otimes \varphi: F \otimes_{R} M \rightarrow F \otimes_{R} M^{\prime}$ is injective.

Proof. When $F=0$ it is clear, so take $F \neq 0$ with basis $\left\{e_{i}\right\}_{i \in I}$. 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 $F \otimes_{R} M^{\prime}$.

For $t \in \operatorname{ker}(1 \otimes \varphi)$, we can write $t=\sum_{i} e_{i} \otimes m_{i}$ with $m_{i} \in M$. Then

$$
0=(1 \otimes \varphi)(t)=\sum_{i} e_{i} \otimes \varphi\left(m_{i}\right),
$$

in $F \otimes_{R} M^{\prime}$, which forces each $\varphi\left(m_{i}\right)$ to be 0 . So every $m_{i}$ is 0 , since $\varphi$ is injective, and we get $t=\sum_{i} e_{i} \otimes 0=0$.

Note that in Theorem 3.2 we did not need to assume $F$ has a finite basis.
Theorem 3.3. Let $R$ be a domain and $K$ be its fraction field. $A s$ an $R$-module, $K$ is flat.
This is not a special case of the previous theorem: if $K$ were a free $R$-module then ${ }^{4}$ $K=R$, so whenever $R$ is a domain that is not a field (e.g., $R=\mathbf{Z}$ ) the fraction field of $R$ is a flat $R$-module that is not a free $R$-module.

[^3]Proof. Let $M \xrightarrow{\varphi} M^{\prime}$ 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 $\varphi(m) \in M_{\text {tor }}^{\prime}$. If $\varphi(m) \in M_{\text {tor }}^{\prime}$ then $r \varphi(m)=0$ for some nonzero $r \in R$, so $\varphi(r m)=0$, so $r m=0$ in $M$ ( $\varphi$ is injective), which means $m \in M_{\text {tor }}$. Thus $x=0$ or $m \in M_{\text {tor }}$, so $t=x \otimes m=0$.

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


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

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

Theorem 3.6. If $M$ is a flat $R$-module and $I$ is an ideal in $R$ then $I \otimes_{R} M \cong I M$ by $i \otimes m \mapsto i m$.

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

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

Example 3.7. For a nonzero torsion abelian group $A$, the natural map $\mathbf{Z} \hookrightarrow \mathbf{Q}$ is injective but if we apply $A \otimes_{\mathbf{Z}}$ we get the map $A \rightarrow 0$, which is not injective, so $A$ is not a flat $\mathbf{Z}$-module. This includes nonzero finite abelian groups and the infinite abelian group $\mathbf{Q} / \mathbf{Z}$.
Remark 3.8. Since $\mathbf{Q} / \mathbf{Z}$ is not flat as a $\mathbf{Z}$-module, for a homomorphism of abelian groups $G \xrightarrow{f} G^{\prime}$ the kernel of $\mathbf{Q} / \mathbf{Z} \otimes_{\mathbf{Z}} G \xrightarrow{1 \otimes f} \mathbf{Q} / \mathbf{Z} \otimes_{\mathbf{Z}} G^{\prime}$ need not be $\mathbf{Q} / \mathbf{Z} \otimes$ ker $f$ but could be larger. Therefore it is not easy to determine the kernel of a group homomorphism after base extension by $\mathbf{Q} / \mathbf{Z}$. Failure to take this into account created a gap in a proof of a widely used theorem in knot theory. See [1, p. 927].
Example 3.9. If $R$ is a domain with fraction field $K$, any nonzero torsion $R$-module $T$ (meaning every element of $T$ is killed by a nonzero element of $R$ ) is not a flat $R$-module
since tensoring the inclusion $R \hookrightarrow K$ with $T$ produces the $R$-linear map $T \rightarrow 0$, which is not injective. In particular, the quotient module $K / R$ is not a flat $R$-module. The previous example is the special case $R=\mathbf{Z}: \mathbf{Q} / \mathbf{Z}$ is not a flat $\mathbf{Z}$-module.

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

Proof. The diagram

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


The map $1 \otimes i$ is injective since $i$ is injective and $N$ is flat. Therefore the two maps out of $N \otimes_{R} M$ above have the same kernel. The kernel of $N \otimes_{R} M \rightarrow N \otimes_{R} \varphi(M)$ can be computed by Corollary 2.21 to be the natural image of $N \otimes_{R} \operatorname{ker} \varphi$ inside $N \otimes_{R} M$, and we can identify the image with $N \otimes_{R} \operatorname{ker} \varphi$ since $N$ is flat.

Theorem 3.11. A tensor product of two flat modules is flat.
Proof. Let $N$ and $N^{\prime}$ be flat. For any injective linear map $M \xrightarrow{\varphi} M^{\prime}$, we want to show the induced linear map $\left(N \otimes_{R} N^{\prime}\right) \otimes_{R} M \xrightarrow{1 \otimes \varphi}\left(N \otimes_{R} N^{\prime}\right) \otimes M^{\prime}$ is injective.

Since $N^{\prime}$ is flat, $N^{\prime} \otimes_{R} M \xrightarrow{1 \otimes \varphi} N^{\prime} \otimes_{R} M^{\prime}$ is injective. Tensoring now with $N, N \otimes_{R}$ $\left(N^{\prime} \otimes_{R} M\right) \xrightarrow{1 \otimes(1 \otimes \varphi)} N \otimes_{R}\left(N^{\prime} \otimes_{R} M^{\prime}\right)$ is injective since $N$ is flat. The diagram

commutes, where the vertical maps are the natural isomorphisms, so the bottom map is injective. Thus $N \otimes_{R} N^{\prime}$ is flat.

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

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


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.


Both $\varphi \otimes 1$ and $1 \otimes \psi$ are injective since $N^{\prime}$ 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

and the two diagonal maps are injective from flatness of $N$ and $M^{\prime}$, so $\varphi \otimes \psi$ is injective.
Corollary 3.13. Let $M_{1}, \ldots, M_{k}, N_{1}, \ldots, N_{k}$ be flat $R$-modules and $\varphi_{i}: M_{i} \rightarrow N_{i}$ be injective linear maps. Then the linear map

$$
\varphi_{1} \otimes \cdots \otimes \varphi_{k}: M_{1} \otimes_{R} \cdots \otimes_{R} M_{k} \rightarrow N_{1} \otimes \cdots \otimes N_{k}
$$

is injective. In particular, if $\varphi: M \rightarrow N$ is an injective linear map of flat modules then the tensor powers $\varphi^{\otimes k}: M^{\otimes k} \rightarrow N^{\otimes k}$ 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


The first diagonal map is injective because $M_{k}$ is flat, and the second diagonal map is injective because $N_{1} \otimes_{R} \cdots \otimes_{R} N_{k-1}$ is flat (Theorem 3.11 and induction).

Corollary 3.14. If $M$ and $N$ are free $R$-modules and $\varphi: M \rightarrow N$ is an injective linear map, any tensor power $\varphi^{\otimes k}: M^{\otimes k} \rightarrow N^{\otimes k}$ is injective.
Proof. Free modules are flat by Theorem 3.2.
Note the free modules in Corollary 3.14 are completely arbitrary. We make no assumptions about finite bases.

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

Corollary 3.15. If $M$ is a free module and $\left\{m_{1}, \ldots, m_{d}\right\}$ is a finite linearly independent subset then for any $k \leq d$ the $d^{k}$ elementary tensors

$$
\begin{equation*}
m_{i_{1}} \otimes \cdots \otimes m_{i_{k}} \text { where } i_{1}, \ldots, i_{k} \in\{1,2, \ldots, d\} \tag{3.1}
\end{equation*}
$$

are linearly independent in $M^{\otimes k}$.
Proof. There is an embedding $R^{d} \hookrightarrow M$ by $\sum_{i=1}^{d} r_{i} e_{i} \mapsto \sum_{i=1}^{d} r_{i} m_{i}$. Since $R^{d}$ and $M$ are free, the $k$ th tensor power $\left(R^{d}\right)^{\otimes k} \rightarrow M^{\otimes k}$ is injective. This map sends the basis

$$
e_{i_{1}} \otimes \cdots \otimes e_{i_{k}}
$$

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

Corollary 3.15 is not saying the elementary tensors in (3.1) can be extended to a basis of $M^{\otimes k}$, just as $m_{1}, \ldots, m_{d}$ usually can't be extended to a basis of $M$.

## 4. Tensor Products of Linear Maps and Base Extension

Fix a ring homomorphism $R \xrightarrow{f} S$. Every $S$-module becomes an $R$-module by restriction of scalars, and every $R$-module $M$ has a base extension $S \otimes_{R} M$, which is an $S$-module. In part I we saw $S \otimes_{R} M$ has a universal mapping property among all $S$-modules: an $R$-linear map from $M$ to any $S$-module "extends" uniquely to an $S$-linear map from $S \otimes_{R} M$ to the $S$-module. We discuss in this section an arguably more important role for base extension: it turns an $R$-linear map $M \xrightarrow{\varphi} M^{\prime}$ between two $R$-modules into an $S$-linear map between $S$-modules. Tensoring $M \xrightarrow{\varphi} M^{\prime}$ with $S$ gives us an $R$-linear map $S \otimes_{R} M \xrightarrow{\stackrel{\otimes \varphi}{\longrightarrow}} S \otimes_{R} M^{\prime}$ that is in fact $S$-linear: $(1 \otimes \varphi)(s t)=s(1 \otimes \varphi)(t)$ for all $s \in S$ and $t \in S \otimes_{R} M$. Since both sides are additive in $t$, to prove $1 \otimes \varphi$ is $S$-linear it suffices to consider the case when $t=s^{\prime} \otimes m$ is an elementary tensor. Then

$$
(1 \otimes \varphi)\left(s\left(s^{\prime} \otimes m\right)\right)=(1 \otimes \varphi)\left(s s^{\prime} \otimes m\right)=s s^{\prime} \otimes \varphi(m)=s\left(s^{\prime} \otimes \varphi(m)\right)=s(1 \otimes \varphi)\left(s^{\prime} \otimes m\right)
$$

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

$$
\varphi_{S}: S \otimes_{R} M \rightarrow S \otimes_{R} M^{\prime} \text { by } \varphi_{S}(s \otimes m)=s \otimes \varphi(m) .
$$

Since $1 \otimes \operatorname{id}_{M}=\mathrm{id}_{S \otimes_{R} M}$ and $(1 \otimes \varphi) \circ\left(1 \otimes \varphi^{\prime}\right)=1 \otimes\left(\varphi \circ \varphi^{\prime}\right)$, we have $\left(\mathrm{id}_{M}\right)_{S}=\mathrm{id}_{S \otimes_{R} M}$ and $\left(\varphi \circ \varphi^{\prime}\right)_{S}=\varphi_{S} \circ \varphi_{S}^{\prime}$. 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 $M \xrightarrow{\varphi} M^{\prime}$ is an isomorphism or is surjective then so is $S \otimes_{R} M \xrightarrow{\varphi_{S}}$ $S \otimes_{R} M^{\prime}$ (Theorems 2.11 and 2.12). But if $\varphi$ is injective then $\varphi_{S}$ need not be injective. (Examples 2.13, 2.14, and 2.15, which all have $S$ as a field).

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

Proof. Pick a maximal ideal $\mathfrak{m}$ in $R$. Tensoring $R$-linear maps $R^{m} \cong R^{n}$ or $R^{m} \rightarrow R^{n}$ with $R / \mathfrak{m}$ produces $R / \mathfrak{m}$-linear maps $(R / \mathfrak{m})^{m} \cong(R / \mathfrak{m})^{n}$ or $(R / \mathfrak{m})^{m} \rightarrow(R / \mathfrak{m})^{n}$. 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

$$
M \cong R^{d} \oplus R /\left(a_{1}\right) \oplus \cdots \oplus R /\left(a_{k}\right),
$$

where $a_{1}|\cdots| a_{k}$, the integer d equals $\operatorname{dim}_{K}\left(K \otimes_{R} M\right)$, 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 $K \otimes_{R}\left(R /\left(a_{i}\right)\right)=0$. Thus $d=\operatorname{dim}_{K}\left(K \otimes_{R} M\right)$.

Example 4.3. Let $A=\left(\begin{array}{cc}a & b \\ c & d\end{array}\right) \in \mathrm{M}_{2}(R)$. Regarding $A$ as a linear map $R^{2} \rightarrow R^{2}$, its base extension $A_{S}: S \otimes_{R} R^{2} \rightarrow S \otimes_{R} R^{2}$ is $S$-linear and $S \otimes_{R} R^{2} \cong S^{2}$ as $S$-modules.

Let $\left\{e_{1}, e_{2}\right\}$ be the standard basis of $R^{2}$. An $S$-basis for $S \otimes_{R} R^{2}$ is $\left\{1 \otimes e_{1}, 1 \otimes e_{2}\right\}$. Using this basis, we can compute a matrix for $A_{S}$ :

$$
A_{S}\left(1 \otimes e_{1}\right)=1 \otimes A\left(e_{1}\right)=1 \otimes\left(a e_{1}+c e_{2}\right)=a\left(1 \otimes e_{1}\right)+c\left(1 \otimes e_{2}\right)
$$

and

$$
A_{S}\left(1 \otimes e_{2}\right)=1 \otimes A\left(e_{2}\right)=1 \otimes\left(b e_{1}+d e_{2}\right)=b\left(1 \otimes e_{1}\right)+d\left(1 \otimes e_{2}\right) .
$$

Therefore the matrix for $A_{S}$ is $\left(\begin{array}{ll}a & b \\ c & d\end{array}\right) \in \mathrm{M}_{2}(S)$. (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 $M^{\prime}$ be nonzero finite-free $R$-modules and $M \xrightarrow{\varphi} M^{\prime}$ be an $R$-linear map. For any bases $\left\{e_{j}\right\}$ and $\left\{e_{i}^{\prime}\right\}$ of $M$ and $M^{\prime}$, the matrix for the $S$-linear map $S \otimes_{R} M \xrightarrow{\varphi_{S}} S \otimes_{R} M^{\prime}$ with respect to the bases $\left\{1 \otimes e_{j}\right\}$ and $\left\{1 \otimes e_{i}^{\prime}\right\}$ equals the matrix for $\varphi$ with respect to $\left\{e_{j}\right\}$ and $\left\{e_{i}^{\prime}\right\}$.
Proof. Say $\varphi\left(e_{j}\right)=\sum_{i} a_{i j} e_{i}^{\prime}$, so the matrix of $\varphi$ is $\left(a_{i j}\right)$. Then

$$
\varphi_{S}\left(1 \otimes e_{j}\right)=1 \otimes \varphi\left(e_{j}\right)=1 \otimes \sum_{i} a_{i j} e_{i}=\sum_{i} a_{i j}\left(1 \otimes e_{i}\right),
$$

so the matrix of $\varphi_{S}$ is also $\left(a_{i j}\right)$.
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 $M^{\prime}$ be $R$-modules. There is a unique $S$-linear map

$$
S \otimes_{R} \operatorname{Hom}_{R}\left(M, M^{\prime}\right) \rightarrow \operatorname{Hom}_{S}\left(S \otimes_{R} M, S \otimes_{R} M^{\prime}\right)
$$

sending $s \otimes \varphi$ to the function $s \varphi_{S}: t \mapsto s \varphi_{S}(t)$ and it is an isomorphism if $M$ and $M^{\prime}$ are finite free $R$-modules. In particular, using $M^{\prime}=M$, there is a unique $S$-linear map

$$
S \otimes_{R}\left(M^{\vee_{R}}\right) \rightarrow\left(S \otimes_{R} M\right)^{\vee_{S}}
$$

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

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

Proof. The function $S \times \operatorname{Hom}_{R}\left(M, M^{\prime}\right) \rightarrow \operatorname{Hom}_{S}\left(S \otimes_{R} M, S \otimes_{R} M^{\prime}\right)$ where $(s, \varphi) \mapsto s \varphi_{S}$ is $R$-bilinear (check!), so there is a unique $R$-linear map

$$
S \otimes_{R} \operatorname{Hom}_{R}\left(M, M^{\prime}\right) \xrightarrow{L} \operatorname{Hom}_{S}\left(S \otimes_{R} M, S \otimes_{R} M^{\prime}\right)
$$

such that $L(s \otimes \varphi)=s \varphi_{S}$. The map $L$ is $S$-linear (check!). If $M^{\prime}=R$ and we identify $S \otimes_{R} R$ with $S$ as $S$-modules by multiplication, then $L$ becomes an $S$-linear map $S \otimes_{R}\left(M^{\vee_{R}}\right) \rightarrow$ $\left(S \otimes_{R} M\right)^{\vee_{S}}$.

Now suppose $M$ and $M^{\prime}$ are both finite free. We want to show $L$ is an isomorphism. If $M$ or $M^{\prime}$ is 0 it is clear, so we may take them both to be nonzero with respective $R$-bases $\left\{e_{i}\right\}$ and $\left\{e_{j}^{\prime}\right\}$, say. Then $S$-bases of $S \otimes_{R} M$ and $S \otimes_{R} M^{\prime}$ are $\left\{1 \otimes e_{i}\right\}$ and $\left\{1 \otimes e_{j}^{\prime}\right\}$. An $R$-basis of $\operatorname{Hom}_{R}\left(M, M^{\prime}\right)$ is the functions $\varphi_{i j}$ sending $e_{i}$ to $e_{j}^{\prime}$ and other basis vectors $e_{k}$ of $M$ to 0 . An $S$-basis of $S \otimes_{R} \operatorname{Hom}_{R}\left(M, M^{\prime}\right)$ is the tensors $\left\{1 \otimes \varphi_{i j}\right\}$, and

$$
L\left(1 \otimes \varphi_{i j}\right)\left(1 \otimes e_{i}\right)=\left(\varphi_{i j}\right)_{S}\left(1 \otimes e_{i}\right)=\left(1 \otimes \varphi_{i j}\right)\left(1 \otimes e_{i}\right)=1 \otimes \varphi_{i j}\left(e_{i}\right)=1 \otimes e_{j}^{\prime}
$$

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

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

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

We have $S \otimes_{R} M=S \otimes_{R} S \cong S$ as $S$-modules. ${ }^{5}$ The natural linear map

$$
\begin{equation*}
S \otimes_{R}\left(M^{\vee_{R}}\right) \rightarrow\left(S \otimes_{R} M\right)^{\vee_{S}} \tag{4.1}
\end{equation*}
$$

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

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

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

Proof. We have a commutative diagram


[^4]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).

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

## 5. Vector Spaces

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

Let $K$ be a field. Tensor products of $K$-vector spaces involve no unexpected collapsing: if $V$ and $W$ are nonzero $K$-vector spaces then $V \otimes_{K} W$ is nonzero and in fact $\operatorname{dim}_{K}\left(V \otimes_{K} W\right)=$ $\operatorname{dim}_{K}(V) \operatorname{dim}_{K}(W)$ in the sense of cardinal numbers.

For any $K$-linear maps $V \xrightarrow{\varphi} V^{\prime}$ and $W \xrightarrow{\psi} W^{\prime}$, we have the tensor product linear map $V \otimes_{K} W \xrightarrow{\varphi \otimes \psi} V^{\prime} \otimes_{K} W^{\prime}$ that sends $v \otimes w$ to $\varphi(v) \otimes \psi(w)$. When $V \xrightarrow{\varphi} V^{\prime}$ and $W \xrightarrow{\psi} W^{\prime}$ are isomorphisms or surjective, so is $V \otimes_{K} W \xrightarrow{\varphi \otimes \psi} V^{\prime} \otimes_{K} W^{\prime}$ (Theorems 2.11 and 2.12). Moreover, because all $K$-vector spaces are free a tensor product of injective $K$-linear maps is injective (Theorem 3.2).
Example 5.1. If $V \xrightarrow{\varphi} W$ is an injective $K$-linear map and $U$ is any $K$-vector space, the $K$-linear map $U \otimes_{K} V \xrightarrow{1 \otimes \varphi} U \otimes_{K} W$ is injective.

Example 5.2. A tensor product of subspaces "is" a subspace: if $V \subset V^{\prime}$ and $W \subset W^{\prime}$ the natural linear map $V \otimes_{K} W \rightarrow V^{\prime} \otimes_{K} W^{\prime}$ 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 $V \xrightarrow{\varphi} V^{\prime}$ and $W \xrightarrow{\psi} W^{\prime}$ are linear then $\varphi(V) \subset V^{\prime}$ and $\psi(W) \subset W^{\prime}$, so we can regard $\varphi(V) \otimes_{K} \psi(W)$ as a subspace of $V^{\prime} \otimes_{K} W^{\prime}$, which we couldn't do with modules in general. The following result gives us some practice with this viewpoint.

Theorem 5.3. Let $V \subset V^{\prime}$ and $W \subset W^{\prime}$ where $V^{\prime}$ and $W^{\prime}$ are nonzero. Then $V \otimes_{K} W=$ $V^{\prime} \otimes_{K} W^{\prime}$ if and only if $V=V^{\prime}$ and $W=W^{\prime}$.

Proof. Since $V \otimes_{K} W$ is inside both $V \otimes_{K} W^{\prime}$ and $V^{\prime} \otimes_{K} W$, which are inside $V^{\prime} \otimes_{K} W^{\prime}$, by reasons of symmetry it suffices to assume $V \subsetneq V^{\prime}$ and show $V \otimes_{K} W^{\prime} \subsetneq V^{\prime} \otimes_{K} W^{\prime}$.

Since $V$ is a proper subspace of $V^{\prime}$, there is a linear functional $\varphi: V^{\prime} \rightarrow K$ that vanishes on $V$ and is not identically 0 on $V^{\prime}$, so $\varphi\left(v_{0}^{\prime}\right)=1$ for some $v_{0}^{\prime} \in V^{\prime}$. Pick nonzero $\psi \in W^{\prime \vee}$, and say $\psi\left(w_{0}^{\prime}\right)=1$. Then the linear function $V^{\prime} \otimes_{K} W^{\prime} \rightarrow K$ where $v^{\prime} \otimes w^{\prime} \mapsto \varphi\left(v^{\prime}\right) \psi\left(w^{\prime}\right)$ vanishes on all of $V \otimes_{K} W^{\prime}$ by checking on elementary tensors but its value on $v_{0}^{\prime} \otimes w_{0}^{\prime}$ is 1. Therefore $v_{0}^{\prime} \otimes w_{0}^{\prime} \notin V \otimes_{K} W^{\prime}$, so $V \otimes_{K} W^{\prime} \subsetneq V^{\prime} \otimes_{K} W^{\prime}$.

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

$$
\begin{equation*}
\operatorname{Hom}_{K}\left(V, V^{\prime}\right) \otimes_{K} \operatorname{Hom}_{K}\left(W, W^{\prime}\right) \rightarrow \operatorname{Hom}_{K}\left(V \otimes_{K} W, V^{\prime} \otimes_{K} W^{\prime}\right) \tag{5.1}
\end{equation*}
$$

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 $V^{\prime}=K$ and $W^{\prime}=K$ in (5.1) and identifying $K \otimes_{K} K$ with $K$ by multiplication, (5.1) says $V^{\vee} \otimes_{K} W^{\vee} \cong\left(V \otimes_{K} W\right)^{\vee}$ 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,

$$
V_{1}^{\vee} \otimes_{K} \cdots \otimes_{K} V_{k}^{\vee} \cong\left(V_{1} \otimes_{K} \cdots \otimes_{K} V_{k}\right)^{\vee}
$$

when the $V_{i}$ 's are finite-dimensional. Here an elementary tensor $\varphi_{1} \otimes \cdots \otimes \varphi_{k} \in \otimes_{i=1}^{k} V_{i}{ }^{\vee}$ acts on an elementary tensor $v_{1} \otimes \cdots \otimes v_{k} \in \bigotimes_{i=1}^{k} V_{i}$ with value $\varphi_{1}\left(v_{1}\right) \cdots \varphi_{k}\left(v_{k}\right) \in K$. In particular,

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

when $V$ is finite-dimensional.
Let's turn now to base extensions to larger fields. When $L / K$ is any field extension, ${ }^{6}$ base extension turns $K$-vector spaces into $L$-vector spaces $\left(V \rightsquigarrow L \otimes_{K} V\right)$ 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 $V \rightarrow W$ 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

$$
\begin{equation*}
L \otimes_{K} \operatorname{Hom}_{K}(V, W) \cong \operatorname{Hom}_{L}\left(L \otimes_{K} V, L \otimes_{K} W\right) \tag{5.2}
\end{equation*}
$$

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 $V \rightarrow W$ and its base extension by $L$ are the same (Theorem 4.4). Taking $W=K$, the natural $L$-linear map

$$
\begin{equation*}
L \otimes_{K} V^{\vee} \cong\left(L \otimes_{K} V\right)^{\vee} \tag{5.3}
\end{equation*}
$$

is an isomorphism for finite-dimensional $V$, using $K$-duals on the left and $L$-duals on the right. ${ }^{7}$

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

We will make no finite-dimensionality assumptions in the rest of this section.
The next theorem tells us the image and kernel of a tensor product of linear maps of vector spaces, with no surjectivity hypotheses as in Theorem 2.19.

[^5]Theorem 5.5. Let $V_{1} \xrightarrow{\varphi_{1}} W_{1}$ and $V_{2} \xrightarrow{\varphi_{2}} W_{2}$ be linear. Then
$\operatorname{ker}\left(\varphi_{1} \otimes \varphi_{2}\right)=\operatorname{ker} \varphi_{1} \otimes_{K} V_{2}+V_{1} \otimes_{K} \operatorname{ker} \varphi_{2}, \quad \operatorname{Im}\left(\varphi_{1} \otimes \varphi_{2}\right)=\varphi_{1}\left(V_{1}\right) \otimes_{K} \varphi_{2}\left(V_{2}\right)$.
In particular, if $V_{1}$ and $V_{2}$ are nonzero then $\varphi_{1} \otimes \varphi_{2}$ is injective if and only if $\varphi_{1}$ and $\varphi_{2}$ are injective, and if $W_{1}$ and $W_{2}$ are nonzero then $\varphi_{1} \otimes \varphi_{2}$ is surjective if and only if $\varphi_{1}$ and $\varphi_{2}$ 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: $\operatorname{ker} \varphi_{1} \otimes_{K} V_{2}$ can be identified with its image in $V_{1} \otimes_{K} V_{2}$ and $\varphi_{1}\left(V_{1}\right) \otimes_{K} \varphi_{2}\left(V_{2}\right)$ can be identified with its image in $W_{1} \otimes_{K} W_{2}$ under the natural maps. Theorem 2.19 for modules has weaker conclusions (e.g., injectivity of $\varphi_{1} \otimes \varphi_{2}$ doesn't imply injectivity of $\varphi_{1}$ and $\varphi_{2}$ ).

Proof. First we handle the image of $\varphi_{1} \otimes \varphi_{2}$. The diagrams

commute, with $i_{1}$ and $i_{2}$ being injections, so the composite diagram

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 $V_{1} \otimes_{K} V_{2} \rightarrow \varphi_{1}\left(V_{1}\right) \otimes_{K} \varphi_{2}\left(V_{2}\right)$ can be computed by Theorem 2.19 to be $\operatorname{ker} \varphi_{1} \otimes_{K} V_{2}+V_{1} \otimes_{K} \operatorname{ker} \varphi_{2}$, 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 $0=\operatorname{ker} \varphi_{1} \otimes_{K} V_{2}+V_{1} \otimes_{K} \operatorname{ker} \varphi_{2}$ from the kernel formula. Therefore the subspaces $\operatorname{ker} \varphi_{1} \otimes_{K} V_{2}$ and $V_{1} \otimes_{K} \operatorname{ker} \varphi_{2}$ both vanish, so $\operatorname{ker} \varphi_{1}$ and $\operatorname{ker} \varphi_{2}$ must vanish (because $V_{2}$ and $V_{1}$ are nonzero, respectively). Conversely, if $\varphi_{1}$ and $\varphi_{2}$ are injective then we already knew $\varphi_{1} \otimes \varphi_{2}$ is injective, but the formula for $\operatorname{ker}\left(\varphi_{1} \otimes \varphi_{2}\right)$ also shows us this kernel is 0 .

If $\varphi_{1} \otimes \varphi_{2}$ is surjective then the formula for its image shows $\varphi_{1}\left(V_{1}\right) \otimes_{K} \varphi_{2}\left(V_{2}\right)=W_{1} \otimes_{K} W_{2}$, so $\varphi_{1}\left(V_{1}\right)=W_{1}$ and $\varphi_{2}\left(V_{2}\right)=W_{2}$ by Theorem 5.3 (here we need $W_{1}$ and $W_{2}$ nonzero). Conversely, if $\varphi_{1}$ and $\varphi_{2}$ are surjective then so is $\varphi_{1} \otimes \varphi_{2}$ because that's true for all modules.

Corollary 5.6. Let $V \subset V^{\prime}$ and $W \subset W^{\prime}$. Then

$$
\left(V^{\prime} \otimes_{K} W^{\prime}\right) /\left(V \otimes_{K} W^{\prime}+V^{\prime} \otimes_{K} W\right) \cong\left(V^{\prime} / V\right) \otimes_{K}\left(W^{\prime} / W\right) .
$$

Proof. Tensor the natural projections $V^{\prime} \xrightarrow{\pi_{1}} V^{\prime} / V$ and $W^{\prime} \xrightarrow{\pi_{2}} W^{\prime} / W$ to get a linear map $V^{\prime} \otimes_{K} W^{\prime} \xrightarrow{\pi_{1} \otimes \pi_{2}}\left(V^{\prime} / V\right) \otimes_{K}\left(W^{\prime} / W\right)$ that is onto with $\operatorname{ker}\left(\pi_{1} \otimes_{2}\right)=V \otimes_{K} W^{\prime}+$ $V^{\prime} \otimes_{K} W$ by Theorem 5.5.

Remark 5.7. It is false that $\left(V^{\prime} \otimes_{K} W^{\prime}\right) /\left(V \otimes_{K} W\right) \cong\left(V^{\prime} / V\right) \otimes_{K}\left(W^{\prime} / W\right)$. The subspace $V \otimes_{K} W$ is generally too small ${ }^{8}$ to be the kernel. This is a distinction between tensor products and direct sums (where $\left(V^{\prime} \oplus W^{\prime}\right) /(V \oplus W) \cong\left(V^{\prime} / V\right) \oplus\left(W^{\prime} / W\right)$ ).
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

$$
\begin{equation*}
\operatorname{ker}(1 \otimes \varphi)=U \otimes_{K} \operatorname{ker} \varphi \quad \text { and } \quad \operatorname{Im}(1 \otimes \varphi)=U \otimes_{K} \varphi(V) \tag{5.4}
\end{equation*}
$$

In particular, for nonzero $U$ the map $\varphi$ 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 $U$.
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

$$
\operatorname{ker}\left(\varphi_{L}\right)=L \otimes_{K} \operatorname{ker} \varphi, \quad \operatorname{Im}\left(\varphi_{L}\right)=L \otimes_{K} \operatorname{Im}(\varphi)
$$

The map $\varphi$ is injective if and only if $\varphi_{L}$ is injective and $\varphi$ is surjective if and only if $\varphi_{L}$ is surjective.

Let's formulate this in the language of matrices. If $V$ and $W$ are finite-dimensional then $\varphi$ can be written as a matrix with entries in $K$ once we pick bases of $V$ and $W$. Then $\varphi_{L}$ 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, $\operatorname{dim}_{L}\left(L \otimes_{K} \operatorname{Im}(\varphi)\right)=\operatorname{dim}_{K} \operatorname{Im}(\varphi)$ and $\operatorname{dim}_{L}\left(L \otimes_{K} \operatorname{ker}(\varphi)\right)=\operatorname{dim}_{K} \operatorname{ker}(\varphi)$. That means $\varphi$ and $\varphi_{L}$ 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 $\left\{v_{i}\right\}_{i \in I}$ of $V$. Let $v=v_{i_{0}}$ in this indexing. Define $\varphi: V \rightarrow K$ by

$$
\varphi\left(\sum_{i} c_{i} v_{i}\right)=c_{i_{0}}
$$

Then $\varphi \in V^{\vee}$ and $\varphi(v)=\varphi\left(v_{i_{0}}\right)=1$.
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 $\varphi_{i}\left(v_{i}\right)=1$. Then $\varphi_{1} \otimes \cdots \otimes \varphi_{k}$ is a linear map $V_{1} \otimes_{K} \cdots \otimes_{K} V_{k} \rightarrow K$ having the effect

$$
v_{1} \otimes \cdots \otimes v_{k} \mapsto \varphi_{1}\left(v_{1}\right) \cdots \varphi_{k}\left(v_{k}\right)=1 \neq 0
$$

so $v_{1} \otimes \cdots \otimes v_{k} \neq 0$.
Corollary 5.12. Let $\varphi_{i}: V_{i} \rightarrow W_{i}$ be linear maps between $K$-vector spaces for $1 \leq i \leq k$. Then the linear map $\varphi_{1} \otimes \cdots \otimes \varphi_{k}: V_{1} \otimes_{K} \cdots \otimes_{K} V_{k} \rightarrow W_{1} \otimes_{K} \cdots \otimes_{K} W_{k}$ is $O$ if and only if some $\varphi_{i}$ is $O$.

[^6]Proof. For $(\Leftarrow)$, if some $\varphi_{i}$ is $O$ then $\left(\varphi_{1} \otimes \cdots \otimes \varphi_{k}\right)\left(v_{1} \otimes \cdots \otimes v_{k}\right)=\varphi_{1}\left(v_{1}\right) \otimes \cdots \otimes \varphi_{k}\left(v_{k}\right)=0$ since $\varphi_{i}\left(v_{i}\right)=0$. 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 $\varphi_{i}$ 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 $\varphi_{i}\left(v_{i}\right) \neq 0$ in $W_{i}$. Then $\varphi_{1} \otimes \cdots \otimes \varphi_{k}$ sends $v_{1} \otimes \cdots \otimes v_{k}$ to $\varphi_{1}\left(v_{1}\right) \otimes \cdots \otimes \varphi_{k}\left(v_{k}\right)$. Since each $\varphi_{i}\left(v_{i}\right)$ is nonzero in $W_{i}$, the elementary tensor $\varphi_{1}\left(v_{1}\right) \otimes \cdots \otimes \varphi_{k}\left(v_{k}\right)$ 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 $M \otimes_{R} N \rightarrow K \otimes_{R}\left(M \otimes_{R} N\right)$ sending $t$ to $1 \otimes t$. We want to show $1 \otimes(x \otimes y) \neq 0$.

The natural $K$-vector space isomorphism $K \otimes_{R}\left(M \otimes_{R} N\right) \cong\left(K \otimes_{R} M\right) \otimes_{K}\left(K \otimes_{R} N\right)$ 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 $\left(K \otimes_{R} M\right) \otimes_{K}\left(K \otimes_{R} N\right)$. Therefore $1 \otimes(x \otimes y) \neq 0$ in $K \otimes_{R}\left(M \otimes_{R} N\right)$, which is what we wanted to show.

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

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

Theorem 5.15. Let $V_{1}, \ldots, V_{k}$ be $K$-vector spaces. Pick pairs of nonzero vectors $v_{i}, v_{i}^{\prime}$ in $V_{i}$ for $i=1, \ldots, k$. Then $v_{1} \otimes \cdots \otimes v_{k}=v_{1}^{\prime} \otimes \cdots \otimes v_{k}^{\prime}$ in $V_{1} \otimes_{K} \cdots \otimes_{K} V_{k}$ if and only if there are nonzero constants $c_{1}, \ldots, c_{k}$ in $K$ such that $v_{i}=c_{i} v_{i}^{\prime}$ and $c_{1} \cdots c_{k}=1$.
Proof. If $v_{i}=c_{i} v_{i}^{\prime}$ for all $i$ and $c_{1} \cdots c_{k}=1$ then $v_{1} \otimes \cdots \otimes v_{k}=c_{1} v_{1}^{\prime} \otimes \cdots \otimes c_{k} v_{k}^{\prime}=$ $\left(c_{1} \cdots c_{k}\right) v_{1}^{\prime} \otimes \cdots \otimes v_{k}^{\prime}=v_{1}^{\prime} \otimes \cdots \otimes v_{k}^{\prime}$.

Now we want to go the other way. It is clear for $k=1$, so we may take $k \geq 2$.
By Theorem 5.11, $v_{1} \otimes \cdots \otimes v_{k}$ is not 0 since each $v_{i}$ is not 0 . Fix $\varphi_{i} \in V_{i}^{\vee}$ for $1 \leq i \leq k-1$ such that $\varphi_{i}\left(v_{i}\right)=1$. (Such $\varphi_{i}$ exist since $v_{1}, \ldots, v_{k-1}$ are nonzero vectors.) For arbitrary $\varphi \in V_{k}^{\vee}$, let $h_{\varphi}=\varphi_{1} \otimes \cdots \otimes \varphi_{k-1} \otimes \varphi$, so $h_{\varphi}\left(v_{1} \otimes \cdots \otimes v_{k-1} \otimes v_{k}\right)=\varphi\left(v_{k}\right)$. Also
$h_{\varphi}\left(v_{1} \otimes \cdots \otimes v_{k-1} \otimes v_{k}\right)=h_{\varphi}\left(v_{1}^{\prime} \otimes \cdots \otimes v_{k-1}^{\prime} \otimes v_{k}^{\prime}\right)=\varphi_{1}\left(v_{1}^{\prime}\right) \cdots \varphi_{k-1}\left(v_{k-1}^{\prime}\right) \varphi\left(v_{k}^{\prime}\right)=\varphi\left(c_{k} v_{k}^{\prime}\right)$, where $c_{k}:=\varphi_{1}\left(v_{1}^{\prime}\right) \cdots \varphi_{k-1}\left(v_{k-1}^{\prime}\right) \in K$. So we have

$$
\varphi\left(v_{k}\right)=\varphi\left(c_{k} v_{k}^{\prime}\right)
$$

for arbitrary $\varphi \in V_{k}^{\vee}$. Therefore $\varphi\left(v_{k}-c_{k} v_{k}^{\prime}\right)=0$ for all $\varphi \in V_{k}^{\vee}$, so $v_{k}-c_{k} v_{k}^{\prime}=0$, which says $v_{k}=c_{k} v_{k}^{\prime}$.

In the same way, for every $i=1,2, \ldots, k$ there is $c_{i}$ in $K$ such that $v_{i}=c_{i} v_{i}^{\prime}$. Then $v_{1} \otimes \cdots \otimes v_{k}=c_{1} v_{1}^{\prime} \otimes \cdots \otimes c_{k} v_{k}^{\prime}=\left(c_{1} \cdots c_{k}\right)\left(v_{1}^{\prime} \otimes \cdots \otimes v_{k}^{\prime}\right)$. Since $v_{1} \otimes \cdots \otimes v_{k}=v_{1}^{\prime} \otimes \cdots \otimes v_{k}^{\prime} \neq 0$, we get $c_{1} \cdots c_{k}=1$.

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

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

Here is the analogue of Theorem 5.15 for linear maps (compare to Corollary 5.12).
Theorem 5.16. Let $\varphi_{i}: V_{i} \rightarrow W_{i}$ and $\varphi_{i}^{\prime}: V_{i} \rightarrow W_{i}$ be nonzero linear maps between $K$ vector spaces for $1 \leq i \leq k$. Then $\varphi_{1} \otimes \cdots \otimes \varphi_{k}=\varphi_{1}^{\prime} \otimes \cdots \otimes \varphi_{k}^{\prime}$ as linear maps $V_{1} \otimes_{K}$ $\cdots \otimes_{K} V_{k} \rightarrow W_{1} \otimes_{K} \cdots \otimes_{K} W_{k}$ if and only if there are $c_{1}, \ldots, c_{k}$ in $K$ such that $\varphi_{i}=c_{i} \varphi_{i}^{\prime}$ and $c_{1} c_{2} \cdots c_{k}=1$.

Proof. Since each $\varphi_{i}: V_{i} \rightarrow W_{i}$ is not identically 0 , for $i=1, \ldots, k-1$ there is $v_{i} \in V_{i}$ such that $\varphi_{i}\left(v_{i}\right) \neq 0$ in $W_{i}$. Then there is $f_{i} \in W_{i}^{\bigvee}$ such that $f_{i}\left(\varphi_{i}\left(v_{i}\right)\right)=1$.

Pick any $v \in V_{k}$ and $f \in W_{k}^{\vee}$. Set $h_{f}=f_{1} \otimes \cdots \otimes f_{k-1} \otimes f \in\left(W_{1} \otimes_{K} \cdots \otimes_{K} W_{k}\right)^{\vee}$ where

$$
h_{f}\left(w_{1} \otimes \cdots \otimes w_{k-1} \otimes w_{k}\right)=f_{1}\left(w_{1}\right) \cdots f_{k-1}\left(w_{k-1}\right) f\left(w_{k}\right) .
$$

Then
$h_{f}\left(\left(\varphi_{1} \otimes \cdots \otimes \varphi_{k-1} \otimes \varphi_{k}\right)\left(v_{1} \otimes \cdots \otimes v_{k-1} \otimes v\right)\right)=f_{1}\left(\varphi_{1}\left(v_{1}\right)\right) \cdots f_{k-1}\left(\varphi_{k-1}\left(v_{k-1}\right)\right) f\left(\varphi_{k}(v)\right)$,

[^7]and since $f_{i}\left(\varphi_{i}\left(v_{i}\right)\right)=1$ for $i \neq k$, the value is $f\left(\varphi_{k}(v)\right)$. Also
$h_{f}\left(\left(\varphi_{1}^{\prime} \otimes \cdots \otimes \varphi_{k-1}^{\prime} \otimes \varphi_{k}^{\prime}\right)\left(v_{1} \otimes \cdots \otimes v_{k-1} \otimes v\right)\right)=f_{1}\left(\varphi_{1}^{\prime}\left(v_{1}\right)\right) \cdots f_{k-1}\left(\varphi_{k-1}^{\prime}\left(v_{k-1}\right)\right) f\left(\varphi_{k}^{\prime}(v)\right)$.
Set $c_{k}=f_{1}\left(\varphi_{1}^{\prime}\left(v_{1}\right)\right) \cdots f_{k-1}\left(\varphi_{k-1}^{\prime}\left(v_{k-1}\right)\right)$, so the value is $c_{k} f\left(\varphi_{k}^{\prime}(v)\right)=f\left(c_{k} \varphi_{k}^{\prime}(v)\right)$. Since $\varphi_{1} \otimes \cdots \otimes \varphi_{k-1} \otimes \varphi_{k}=\varphi_{1}^{\prime} \otimes \cdots \otimes \varphi_{k-1}^{\prime} \otimes \varphi_{k}^{\prime}$,
$$
f\left(\varphi_{k}(v)\right)=f\left(c_{k} \varphi_{k}^{\prime}(v)\right)
$$

This holds for all $f \in W_{k}^{\vee}$, so $\varphi_{k}(v)=c_{k} \varphi_{k}^{\prime}(v)$. This holds for all $v \in V_{k}$, so $\varphi_{k}=c_{k} \varphi_{k}^{\prime}$ as linear maps $V_{k} \rightarrow W_{k}$.

In a similar way, there is $c_{i} \in K$ such that $\varphi_{i}=c_{i} \varphi_{i}^{\prime}$ for all $i$, so

$$
\begin{aligned}
\varphi_{1} \otimes \cdots \otimes \varphi_{k} & =\left(c_{1} \varphi_{1}^{\prime}\right) \otimes \cdots \otimes\left(c_{k} \varphi_{k}^{\prime}\right) \\
& =\left(c_{1} \cdots c_{k}\right) \varphi_{1}^{\prime} \otimes \cdots \otimes \varphi_{k}^{\prime} \\
& =\left(c_{1} \cdots c_{k}\right) \varphi_{1} \otimes \cdots \otimes \varphi_{k},
\end{aligned}
$$

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 finitedimensional.

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:

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

The inverse on $\sigma$ is needed to get a genuine left group action (Check!). Here is a similar action of $S_{k}$ on the $k$ th tensor power.

Corollary 5.18. For $\sigma \in S_{k}$, there is a linear map $P_{\sigma}: V^{\otimes k} \rightarrow V^{\otimes k}$ such that

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

on elementary tensors. Then $P_{\sigma} \circ P_{\tau}=P_{\sigma \tau}$ for $\sigma$ and $\tau$ in $S_{k} .{ }^{10}$
When $\operatorname{dim}_{K}(V)>1, P_{\sigma}=P_{\tau}$ if and only if $\sigma=\tau$. In particular, $P_{\sigma}$ is the identity map if and only if $\sigma$ is the identity permutation.

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

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

is multilinear, so the universal mapping property of tensor products gives us a linear map $P_{\sigma}$ with the indicated effect on elementary tensors. It is clear that $P_{1}$ is the identity map. For any $\sigma$ and $\tau$ 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 $P_{\sigma}$ is the identity map on $V^{\otimes k}$ then $\sigma$ is the identity permutation.

We prove the contrapositive. Suppose $\sigma$ 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 $\operatorname{dim}_{K} V>1$.) If $P_{\sigma}\left(v_{1} \otimes \cdots \otimes v_{k}\right)=v_{1} \otimes \cdots \otimes v_{k}$ then $v_{j} \in K v_{i}$ by Theorem 5.15, which is not so.

[^8]The linear maps $P_{\sigma}$ 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 GL $(V)$ acts on $V^{\otimes k}$ via tensor powers of linear maps:

$$
g\left(v_{1} \otimes \cdots \otimes v_{k}\right):=g^{\otimes k}\left(v_{1} \otimes \cdots \otimes v_{k}\right)=g v_{1} \otimes \cdots \otimes g v_{k}
$$

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

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

$$
\sum_{\sigma \in S_{k}} P_{\sigma}\left(v_{1} \otimes v_{2} \otimes \cdots \otimes v_{k}\right)=\sum_{\sigma \in S_{k}} v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}=\sum_{\sigma \in S_{k}} v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(k)}
$$

An example of an anti-symmetric tensor in $V^{\otimes k}$ is $\sum_{\sigma \in S_{k}}(\operatorname{sign} \sigma) v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(k)}$. Both symmetric and anti-symmetric tensors occur in physics. See Table 1.

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

Table 1. Some tensors in physics.

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

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

(The map $P_{(12)}$ is the flip automorphism of $V^{\otimes 2}$ sending $v \otimes w$ to $w \otimes v$.) Concretely, if $\left\{e_{1}, \ldots, e_{n}\right\}$ is a basis of $V$ then each tensor in $V^{\otimes 2}$ has the form $\sum_{i, j} c_{i j} e_{i} \otimes e_{j}$, and it's symmetric if and only if $c_{i j}=c_{j i}$ for all $i$ and $j$, while it's anti-symmetric if and only if $c_{i j}=-c_{j i}$ for all $i$ and $j$ (in particular, $c_{i i}=0$ for all $i$ ).

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

## 6. TENSOR CONTRACTION

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

$$
\begin{equation*}
c(\mathbf{v} \otimes \widetilde{\mathbf{w}})=\sum_{i, j} v^{i} w_{j} e^{j}\left(e_{i}\right)=\sum_{i} v^{i} w_{i} \tag{6.1}
\end{equation*}
$$

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

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

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

on elementary tensors. This isomorphism converts contraction on $V \otimes V^{\vee}$ into a linear map $\operatorname{Hom}_{K}(V, V) \rightarrow K$ that turns out to be the trace. Why? A tensor in $V \otimes V^{\vee}$ can be written uniquely as a linear combination $\sum_{i, j} T_{j}^{i} e_{i} \otimes e^{j}$. The isomorphism $V \otimes V^{\vee} \cong \operatorname{Hom}_{K}(V, V)$ sends $\sum_{i, j} T_{j}^{i} e_{i} \otimes e^{j}$ to the linear operator $V \rightarrow V$ that maps each $e_{k}$ to $\sum_{i, j} T_{j}^{i} e^{j}\left(e_{k}\right) e_{i}=$ $\sum_{i, j} T_{j}^{i} \delta_{j k} e_{i}=\sum_{i} T_{k}^{i} e_{i}$, so this linear operator $V \rightarrow V$ has matrix representation $\left(T_{j}^{i}\right)$ with respect to the basis $\left\{e_{i}\right\}$. The contraction of $\sum_{i, j} T_{j}^{i} e_{i} \otimes e^{j}$ is $c\left(\sum_{i, j} T_{j}^{i} e_{i} \otimes e^{j}\right)=\sum_{i, j} T_{j}^{i} \delta_{i j}=$ $\sum_{i} T_{i}^{i}$, which is exactly the trace of the matrix $\left(T_{j}^{i}\right)$.

Contraction $V \otimes V^{\vee} \rightarrow K$ generalizes to maps $V^{\otimes k} \otimes\left(V^{\vee}\right)^{\otimes \ell} \rightarrow V^{\otimes(k-1)} \otimes\left(V^{\vee}\right)^{\otimes(\ell-1)}$ for $k \geq 1$ and $\ell \geq 1$ (recall $V^{\otimes 0}=K$ and $\left(V^{\vee}\right)^{\otimes 0}=K$ ) that are linear and also called contractions, as follows. For $r \in\{1, \ldots, k\}$ and $s \in\{1, \ldots, \ell\}$, the associated contraction $c_{r, s}: V^{\otimes k} \otimes\left(V^{\vee}\right)^{\otimes \ell} \rightarrow V^{\otimes(k-1)} \otimes\left(V^{\vee}\right)^{\otimes(\ell-1)}$ evaluates the $r$ th tensorand in $V^{\otimes k}$ in the $s$ th tensorand in $\left(V^{\vee}\right)^{\otimes \ell}$, leaving other tensorands untouched: on elementary tensors,

$$
\begin{aligned}
c_{r, s}\left(v_{1} \otimes \cdots \otimes v_{k} \otimes \varphi_{1} \otimes \cdots \otimes \varphi_{\ell}\right) & =c\left(v_{r} \otimes \varphi_{s}\right) \bigotimes_{i \neq r} v_{i} \otimes \bigotimes_{j \neq s} \varphi_{j} \\
& =\varphi_{s}\left(v_{r}\right) \bigotimes_{i \neq r} v_{i} \otimes \bigotimes_{j \neq s} \varphi_{j}
\end{aligned}
$$

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

$$
c_{r, s}\left(\sum_{\substack{i_{1}, \ldots, i_{k} \\ j_{1}, \ldots, j_{\ell}}} T_{j_{1} \cdots j_{\ell}}^{i_{1} \ldots i_{k}} e_{i_{1}} \otimes \cdots \otimes e_{i_{k}} \otimes e^{j_{1}} \otimes \cdots \otimes e^{j_{\ell}}\right)=\sum_{\substack{i_{1}, \ldots, j_{\ell} \\ i_{r}, j_{s} m i s s i n g}} \sum_{m=1}^{n} T_{j_{1} \cdots m \cdots j_{\ell}}^{i_{1} \cdots m \cdots i_{k}} \bigotimes_{a \neq r} e_{i_{a}} \otimes \bigotimes_{b \neq s} e^{j_{b}}
$$

[^9]where the $m$ appears in the $r$ th slot of the superscript and the $s$ th slot of the subscript of the coefficient on the right. The outer sum on the right initially runs over all sets of indices and the coefficient is multiplied by the numerical factor $e^{i_{s}}\left(e_{i_{r}}\right)$, which is nonzero only for $i_{s}=i_{r}$ (call this common value $m$, and its runs from 1 to $n$ ), when the factor is 1 .

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

$$
\begin{equation*}
c_{2,1}\left(\sum_{\substack{i_{1}, i_{2} \\ j_{1}}} T_{j_{1}}^{i_{1} i_{2}} e_{i_{1}} \otimes e_{i_{2}} \otimes e^{j_{1}}\right)=\sum_{\substack{i_{1}, i_{2} \\ j_{1}}} T_{j_{1}}^{i_{1} i_{2}} e^{j_{1}}\left(e_{i_{2}}\right) e_{i_{1}}=\sum_{i_{1}}\left(\sum_{m} T_{m}^{i_{1} m}\right) e_{i_{1}} \tag{6.2}
\end{equation*}
$$

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

$$
\begin{equation*}
c_{2,1}\left(\sum_{\substack{i_{1}, i_{2} \\ j_{1}, j_{2}}} T_{j_{1} j_{2}}^{i_{1} i_{2}} e_{i_{1}} \otimes e_{i_{2}} \otimes e^{j_{1}} \otimes e^{j_{2}}\right)=\sum_{i_{1}, j_{2}}\left(\sum_{m} T_{m j_{2}}^{i_{1} m}\right) e_{i_{1}} \otimes e^{j_{2}} \tag{6.3}
\end{equation*}
$$

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

$$
T_{j}^{i} \mapsto T_{i}^{i}
$$

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

$$
T_{j_{1}}^{i_{1} i_{2}} \mapsto T_{i_{2}}^{i_{1} i_{2}}, \quad T_{j_{1} j_{2}}^{i_{1} i_{2}} \mapsto T_{i_{2} j_{2}}^{i_{1} i_{2}}
$$

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

$$
\begin{equation*}
V \otimes V \xrightarrow{\text { id. } \otimes g} V \otimes V^{\vee} \xrightarrow{c} K \tag{6.4}
\end{equation*}
$$

This depends on the choice of $g$ but not on a choice of basis of $V$.
There should be nothing special about using $g$ in the second tensorand in (6.4). We want

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

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

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

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

$$
g(\mathbf{v})=g\left(\sum_{i} v^{i} e_{i}\right)=\sum_{i} v^{i} g\left(e_{i}\right)=\sum_{i} v^{i}\left(\sum_{j} g_{i j} e^{j}\right)=\sum_{j}\left(\sum_{i} g_{i j} v^{i}\right) e^{j} .
$$

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

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

$$
\begin{aligned}
\mathbf{x} \otimes \mathbf{y} & \mapsto \mathbf{x} \otimes g(\mathbf{y}) \\
& =\left(\sum_{i} x^{i} e_{i}\right) \otimes\left(\sum_{j} y_{j} e^{j}\right) \quad \text { where } y_{j}=\sum_{k} g_{k j} y^{k}=\sum_{k} g_{j k} y^{k} \\
& \stackrel{c}{\mapsto} \sum_{i} x^{i} y_{i} \text { by (6.1). }
\end{aligned}
$$

Unwrapping the definition of $y_{i}$ terms of $g_{i j}$ and $y^{j}$ (for $j=1, \ldots, n$ ),

$$
\sum_{i} x^{i} y_{i}=\sum_{i, j} g_{i j} x^{i} y^{j}
$$

This is usually not $\sum_{i} x^{i} y^{i}$ unless $g_{i j}=\delta_{i j}$ for all $i$ and $j$. More generally, the contraction of the tensor $\mathrm{T}=\sum_{i, j} T^{i j} e_{i} \otimes e_{j}$ in $V^{\otimes 2}$ using $g$ is the scalar $\sum_{i, j} g_{i j} T^{i j}$ in $K$. This contraction is independent of the basis $\left\{e_{1}, \ldots, e_{n}\right\}$ of $V$ but depends on the isomorphism $g: V \rightarrow V^{\vee}$.

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

$$
\sum_{i_{1}, \ldots, i_{k}} T^{i_{1} \ldots i_{k}} e_{i_{1}} \otimes \cdots \otimes g\left(e_{i_{s}}\right) \otimes \cdots \otimes e_{i_{k}}=\sum_{i_{1}, \ldots, i_{k}} T^{i_{1} \ldots i_{k}} e_{i_{1}} \otimes \cdots \otimes \underbrace{\left(\sum_{i} g_{i_{s}} e^{i}\right)}_{\text {sth tensorand }} \otimes \cdots \otimes e_{i_{k}}
$$

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

$$
\sum_{\substack{i_{1}, \ldots, i_{k} \\ i_{r}, i_{s} \text { missing }}}\left(\sum_{i_{r}, i_{s}} g_{i_{r} i_{s}} T^{i_{1} \ldots i_{k}}\right) \underbrace{e_{i_{1}} \otimes \cdots \otimes e_{i_{k}}}_{i_{r}, i_{s} \text { missing }} .
$$

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

Using a basis $\left\{e_{1}, \ldots, e_{n}\right\}$ for $V$ and dual basis $\left\{e^{1}, \ldots, e^{n}\right\}$ for $V^{\vee}$, in $V^{\otimes 2} \otimes V^{\vee}$ let

$$
\mathrm{T}=\sum_{\substack{i_{1}, i_{2} \\ j_{1}}} T_{j_{1}}^{i_{1} i_{2}} e_{i_{1}} \otimes e_{i_{2}} \otimes e^{j_{1}}
$$

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

$$
\mathbf{U}=(g \otimes 1 \otimes 1)(\mathbf{T})=\sum_{\substack{i_{2} \\ j_{1}, j_{2}}} U_{j_{2} j_{1}}^{i_{2}} e^{j_{2}} \otimes e_{i_{2}} \otimes e^{j_{1}}, \quad \text { where } \quad U_{j_{2} j_{1}}^{i_{2}}=\sum_{i_{1}} g_{i_{1} j_{2}} T_{j_{1}}^{i_{1} i_{2}}
$$

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

$$
\mathrm{U}^{\prime}=(g \otimes g \otimes 1)(\mathrm{T})=\sum_{j_{1}, j_{2}, j_{3}} U_{j_{2} j_{3} j_{1}} e^{j_{2}} \otimes e^{j_{3}} \otimes e^{j_{1}}, \quad \text { where } \quad U_{j_{2} j_{3} j_{1}}=\sum_{i_{1}, i_{2}} g_{i_{1} j_{2}} g_{i_{2} j_{3}} T_{j_{1}}^{i_{1} i_{2}} .
$$

Remark 6.4. Putting the Einstein summation convention (Remark 6.1) to work, the contraction $V \otimes V^{\vee} \rightarrow K$ sends $v^{i} e_{i} \otimes w_{j} e^{j}$ to $v^{i} w_{i}$, which is the same as $v^{j} w_{j}$ and is independent of the basis since it is simply evaluating $V^{\vee}$ on $V$. Using an isomorphism $g: V \rightarrow V^{\vee}$ that is symmetric (and nondegenerate) as a bilinear form $V \times V \rightarrow K$, any $\mathbf{v}=v^{i} e_{i}$ in $V$ turns into $g(\mathbf{v})=v_{i} e^{i}$ in $V^{\vee}$ where $v_{i}=g_{k i} v^{k}=g_{i k} v^{k}$ since $\left(g_{i j}\right)$ is symmetric, and the contraction $V \otimes V \rightarrow K$ that depends on $g$ has the effect $x^{i} e_{i} \otimes y^{j} e_{j} \mapsto x^{i} y_{i}=x_{j} y^{j}=g_{i j} x^{i} y^{j}$ (remember $g_{i j}=g_{j i}!$ ) and more generally $T^{i j} e_{i} \otimes e_{j} \mapsto g_{i j} T^{i j}$. The contraction on $V^{\otimes 2}$ depends on $g$ but not on the basis of $V$. Contraction on $V^{\otimes k}$, for $k \geq 2$, using the $r$ th and $s$ th tensorand is a linear map to $V^{\otimes(k-2)}$ with the effect $T^{i_{1} \cdots i_{k}} e_{i_{1}} \otimes \cdots \otimes e_{i_{k}} \mapsto g_{i_{r} i_{s}} T^{i_{1} \cdots i_{k}} \underbrace{e_{i_{1}} \otimes \cdots \otimes e_{i_{k}}}_{i_{r}, i_{s} \text { missing }}$. For $T_{j_{1}}^{i_{1} i_{2}}$ in $V \otimes V \otimes V^{\vee}$, if we lower the first upper index we get $g_{i_{1} j_{2}} T_{j_{1}}^{i_{1} i_{2}}$ in $V^{\vee} \otimes V \otimes V^{\vee}$ and if we lower both upper indices then we get $g_{i_{1} j_{2}} g_{i_{2} j_{3}} T_{j_{1}}^{i_{1} i_{2}}$ in $V^{\vee} \otimes V^{\vee} \otimes V^{\vee}$.

[^10]Contraction lets us map $V^{\otimes k} \otimes\left(V^{\vee}\right)^{\otimes \ell}$ to $V^{\otimes(k-1)} \otimes\left(V^{\vee}\right)^{\otimes(\ell-1)}$ by combining a choice of $V$ and $V^{\vee}$ in $V^{\otimes k} \otimes\left(V^{\vee}\right)^{\otimes \ell}$ using evaluation to get scalars. We can use an isomorphism $g: V \rightarrow V^{\vee}$ to change any $V$ in $V^{\otimes k} \otimes\left(V^{\vee}\right)^{\otimes \ell}$ into $V^{\vee}$ (in practice $g$ is symmetric when viewed as a bilinear form) and thus contract any two different $V$ tensorands (the result depends on $g$ ). To contract two different $V^{\vee}$ tensorands, use the inverse $g^{-1}: V^{\vee} \rightarrow V$, which as a bilinear form on $V^{\vee}$ is symmetric: writing $g^{-1}\left(e^{i}\right)=\sum_{j} g^{i j} e_{j},\left(g^{i j}\right)$ is the inverse matrix to $\left(g_{i j}\right)$ and is symmetric since $\left(g_{i j}\right)$ is. We have $g^{-1}\left(\sum_{i} v_{i} e^{i}\right)=\sum_{j} v^{j} e_{j}$, where $v^{j}=\sum_{i} g^{i j} v_{i}$. The passage from $\left\{v_{i}\right\}$ to $\left\{v^{j}\right\}$ is multiplying $v_{i}$ by $g^{i j}$ and summing over all $i$. It is called raising an index and depends on $g$ : this is the coordinate version (using a basis of $V$ and its dual basis in $V^{\vee}$ ) of going from elements of $V^{\vee}$ to elements of $V$ by $g^{-1} .{ }^{13}$ Using $g$ or $g^{-1}$ in enough places lets us turn the mixed tensors of $V^{\otimes k} \otimes\left(V^{\vee}\right)^{\otimes \ell}$ into pure tensors in $V^{\otimes(k+\ell)}$ or $\left(V^{\vee}\right)^{\otimes(k+\ell)}$.

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

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

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


## 7. Tensor Product of $R$-Algebras

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

[^11]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 $A \times A \rightarrow A$, called multiplication or product. Bilinearity of multiplication includes distributive laws for multiplication over addition as well as the extra rule

$$
\begin{equation*}
r(a b)=(r a) b=a(r b) \tag{7.1}
\end{equation*}
$$

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 $R\left[X_{1}, \ldots, X_{n}\right]$. We will assume, except in Example 7.5, that our $R$-algebras have associative multiplication and a multiplicative identity, so they are genuinely rings (perhaps not commutative) and being an $R$-algebra just means they have a little extra structure related to scaling by $R$. When an $R$-algebra contains $R,(7.1)$ is a special case of associative multiplication in the algebra.

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

Because of the universal mapping property of the tensor product, to give an $R$-bilinear multiplication $A \times A \rightarrow A$ in an $R$-algebra $A$ is the same thing as giving an $R$-linear map $A \otimes_{R} A \rightarrow A$. 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 $a b:=m(a \otimes b)$. Associativity of multiplication can be formulated in tensor language: the diagram

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

$$
\begin{equation*}
(a \otimes b)\left(a^{\prime} \otimes b^{\prime}\right)=a a^{\prime} \otimes b b^{\prime} \tag{7.2}
\end{equation*}
$$

for all elementary tensors. The multiplicative identity is $1 \otimes 1$.
Proof. If there is an $R$-algebra multiplication on $A \otimes_{R} B$ satisfying (7.2) then multiplication between any two tensors is determined:

$$
\sum_{i=1}^{k} a_{i} \otimes b_{i} \cdot \sum_{j=1}^{\ell} a_{j}^{\prime} \otimes b_{j}^{\prime}=\sum_{i, j}\left(a_{i} \otimes b_{i}\right)\left(a_{j}^{\prime} \otimes b_{j}^{\prime}\right)=\sum_{i, j} a_{i} a_{j}^{\prime} \otimes b_{i} b_{j}^{\prime} .
$$

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

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

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

$$
\begin{aligned}
\left(A \otimes_{R} B\right) \otimes_{R}\left(A \otimes_{R} B\right) & \cong\left(\left(A \otimes_{R} B\right) \otimes_{R} A\right) \otimes_{R} B \\
& \cong\left(A \otimes_{R}\left(B \otimes_{R} A\right)\right) \otimes_{R} B \\
& \cong\left(A \otimes_{R}\left(A \otimes_{R} B\right)\right) \otimes_{R} B \\
& \left.\cong\left(A \otimes_{R} A\right) \otimes_{R} B\right) \otimes_{R} B \\
& \cong\left(A \otimes_{R} A\right) \otimes_{R}\left(B \otimes_{R} B\right) .
\end{aligned}
$$

Tracking the effect of these maps on $(a \otimes b) \otimes\left(a^{\prime} \otimes b^{\prime}\right)$,

$$
\begin{aligned}
(a \otimes b) \otimes\left(a^{\prime} \otimes b^{\prime}\right) & \mapsto\left((a \otimes b) \otimes a^{\prime}\right) \otimes b^{\prime} \\
& \mapsto\left(a \otimes\left(b \otimes a^{\prime}\right)\right) \otimes b^{\prime} \\
& \mapsto\left(a \otimes\left(a^{\prime} \otimes b\right)\right) \otimes b^{\prime} \\
& \mapsto\left(\left(a \otimes a^{\prime}\right) \otimes b\right) \otimes b^{\prime} \\
& \mapsto\left(a \otimes a^{\prime}\right) \otimes\left(b \otimes b^{\prime}\right) .
\end{aligned}
$$

Composing these isomorphisms with $m_{A} \otimes m_{B}$ makes a map $\left(A \otimes_{R} B\right) \otimes_{R}\left(A \otimes_{R} B\right) \rightarrow A \otimes_{R} B$ that is $R$-linear and has the effect

$$
(a \otimes b) \otimes\left(a^{\prime} \otimes b^{\prime}\right) \mapsto\left(a \otimes a^{\prime}\right) \otimes\left(b \otimes b^{\prime}\right) \mapsto a a^{\prime} \otimes b b^{\prime},
$$

where $m_{A} \otimes m_{B}$ is used in the second step. This $R$-linear map $\left(A \otimes_{R} B\right) \otimes_{R}\left(A \otimes_{R} B\right) \rightarrow$ $A \otimes_{R} B$ pulls back to an $R$-bilinear map $\left(A \otimes_{R} B\right) \times\left(A \otimes_{R} B\right) \rightarrow A \otimes_{R} B$ with the effect $\left(a \otimes b, a^{\prime} \otimes b^{\prime}\right) \mapsto a a^{\prime} \otimes b b^{\prime}$ on pairs of elementary tensors, which is what we wanted for our multiplication on $A \otimes_{R} B$. This proves $A \otimes_{R} B$ has a multiplication satisfying (7.2).

To prove $1 \otimes 1$ is an identity and multiplication in $A \otimes_{R} B$ is associative, we want

$$
(1 \otimes 1) t=t, \quad t(1 \otimes 1)=t, \quad\left(t_{1} t_{2}\right) t_{3}=t_{1}\left(t_{2} t_{3}\right)
$$

for general tensors $t, t_{1}, t_{2}$, and $t_{3}$ in $A \otimes_{R} B$. These identities are additive in each tensor appearing on both sides, so verifying these equations reduces to the case that the tensors are all elementary, and this case is left to the reader.

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 $t t^{\prime}=t^{\prime} t$ for all $t$ and $t^{\prime}$ 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: $(a \otimes b) t^{\prime} \stackrel{?}{=} t^{\prime}(a \otimes b)$. Both sides of this are additive in $t^{\prime}$, so we are reduced further to the special case when $t^{\prime}=a^{\prime} \otimes b^{\prime}$ is also an elementary tensor: $(a \otimes b)\left(a^{\prime} \otimes b^{\prime}\right) \stackrel{?}{=}\left(a^{\prime} \otimes b^{\prime}\right)(a \otimes b)$. The validity of this is immediate from (7.2) since $A$ and $B$ are commutative.

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

$$
\begin{gathered}
(1 \otimes i)^{2}=1 \otimes(-1)=-(1 \otimes 1), \quad(1 \otimes i)(i \otimes 1)=i \otimes i, \quad(1 \otimes i)(i \otimes i)=i \otimes(-1)=-(i \otimes 1), \\
(i \otimes 1)^{2}=(-1) \otimes 1=-(1 \otimes 1), \quad(i \otimes 1)(i \otimes i)=(-1) \otimes i=-(1 \otimes i), \quad(i \otimes i)^{2}=1 \otimes 1 .
\end{gathered}
$$

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

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

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

with neither factor on the left side being 0 .
Example 7.5. Let $\mathbf{R}^{3}$ have the cross product $\times$ as multiplication: $\mathbf{v w}:=\mathbf{v} \times \mathbf{w}$. This is bilinear with $\mathbf{v v}=\mathbf{v} \times \mathbf{v}=\mathbf{0}$ for all $\mathbf{v}$ in $\mathbf{R}^{3}$ and it's assumed the reader has seen the products among the different basis vectors $\mathbf{i}, \mathbf{j}$, and $\mathbf{k}$, e.g., $\mathbf{i} \times \mathbf{j}=\mathbf{k}$ and $\mathbf{i} \times \mathbf{k}=-\mathbf{j}$.

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

$$
(\mathbf{v} \otimes \mathbf{w})\left(\mathbf{v}^{\prime} \otimes \mathbf{w}^{\prime}\right)=\mathbf{v}^{\prime} \otimes \mathbf{w w}^{\prime}=\left(\mathbf{v} \times \mathbf{v}^{\prime}\right) \otimes\left(\mathbf{w} \times \mathbf{w}^{\prime}\right)
$$

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

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

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

$$
\begin{aligned}
t\left(t^{\prime} t^{\prime \prime}\right)+t^{\prime}\left(t^{\prime \prime} t\right)+t^{\prime \prime}\left(t t^{\prime}\right) & =\mathbf{i} \otimes \mathbf{j}((\mathbf{i} \otimes \mathbf{k})(\mathbf{j} \otimes \mathbf{k}))+\mathbf{i} \otimes \mathbf{k}((\mathbf{j} \otimes \mathbf{k})(\mathbf{i} \otimes \mathbf{j}))+\mathbf{j} \otimes \mathbf{k}((\mathbf{i} \otimes \mathbf{j})(\mathbf{i} \otimes \mathbf{k})) \\
& =(\mathbf{i} \otimes \mathbf{j})(\mathbf{i} \mathbf{j} \otimes \mathbf{k} \mathbf{k})+(\mathbf{i} \otimes \mathbf{k})(\mathbf{j} \otimes \mathbf{k} \mathbf{j})+(\mathbf{j} \otimes \mathbf{k})(\mathbf{i} \otimes \mathbf{j} \mathbf{k}) \\
& =(\mathbf{i} \otimes \mathbf{j})(\mathbf{i} \mathbf{j} \otimes \mathbf{0})+(\mathbf{i} \otimes \mathbf{k})(-\mathbf{k} \otimes-\mathbf{i})+(\mathbf{j} \otimes \mathbf{k})(\mathbf{0} \otimes \mathbf{j} \mathbf{k}) \\
& =\mathbf{0}+(\mathbf{i} \otimes \mathbf{k})(\mathbf{k} \otimes \mathbf{i})+\mathbf{0} \\
& =\mathbf{i} \mathbf{k} \otimes \mathbf{k} \mathbf{i} \\
& =-\mathbf{j} \otimes \mathbf{j}
\end{aligned}
$$

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

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

For any $R$-algebras $A$ and $B$, there is an $R$-algebra homomorphism $A \rightarrow A \otimes_{R} B$ 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 $a \otimes(b \bmod 5)=a b \bmod 5)$, 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 $B \rightarrow A \otimes_{R} B$ by $b \mapsto 1 \otimes b$. Even when $A$ and $B$ are noncommutative, the images of $A$ and $B$ in $A \otimes_{R} B$ commute: $(a \otimes 1)(1 \otimes b)=a \otimes b=(1 \otimes b)(a \otimes 1)$. This is like groups $G$ and $H$ commuting in $G \times H$ even if $G$ and $H$ are nonabelian.

It is worth contrasting the direct product $A \times B$ (componentwise addition and multiplication, with $r(a, b)=(r a, r b))$ and the tensor product $A \otimes_{R} B$, which are both $R$-algebras. The direct product $A \times B$ is a ring structure on the $R$-module $A \oplus B$, which is usually quite different from $A \otimes_{R} B$ as an $R$-module. There are natural $R$-algebra homomorphisms $A \times B \xrightarrow{\pi_{1}} A$ and $A \times B \xrightarrow{\pi_{2}} B$ by projection, while there are natural $R$-algebra homomorphisms $A \rightarrow A \otimes_{R} B$ and $B \rightarrow A \otimes_{R} B$ in the other direction (out of $A$ and $B$ to the tensor product rather than to $A$ and $B$ from the direct product). The projections out of the direct product $A \times B$ to $A$ and $B$ are both surjective, but the maps to the tensor product $A \otimes_{R} B$ from $A$ and $B$ need not be injective, e.g., $\mathbf{Z} \rightarrow \mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z} / 5 \mathbf{Z}$. The maps $A \rightarrow A \otimes_{R} B$ and $B \rightarrow A \otimes_{R} B$ are ring homomorphisms and the images are subrings, but although there are natural functions $A \rightarrow A \times B$ and $B \rightarrow A \times B$ given by $a \mapsto(a, 0)$ and $b \mapsto(0, b)$, these are not ring homomorphisms and the images are ideals rather than subrings.

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

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

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

We will see in Example 7.18 that $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ has infinitely many square roots of 1.
The $R$-algebras $A \times B$ and $A \otimes_{R} B$ have dual universal mapping properties. For any $R$-algebra $C$ and $R$-algebra homomorphisms $C \xrightarrow{\varphi} A$ and $C \xrightarrow{\psi} B$, there is a unique
$R$-algebra homomorphism $C \rightarrow A \times B$ making the diagram

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 $A \otimes_{R} B \rightarrow C$ making the diagram

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

$$
\begin{aligned}
\varphi\left(\sum_{i} r_{i} a_{i} \cdot \sum_{j} r_{j}^{\prime} a_{j}\right) & =\varphi\left(\sum_{i, j} r_{i} r_{j}^{\prime} a_{i} a_{j}\right) \\
& =\sum_{i, j} r_{i} r_{j}^{\prime} \varphi\left(a_{i} a_{j}\right) \\
& =\sum_{i, j} r_{i} r_{j}^{\prime} \varphi\left(a_{i}\right) \varphi\left(a_{j}\right) \\
& =\sum_{i} r_{i} \varphi\left(a_{i}\right) \sum_{j} r_{j}^{\prime} \varphi\left(a_{j}\right) \\
& =\varphi\left(\sum_{i} r_{i} a_{i}\right) \varphi\left(\sum_{j} r_{j}^{\prime} a_{j}\right) .
\end{aligned}
$$

This will let us bootstrap a lot of known $R$-module isomorphisms between tensor products to $R$-algebra isomorphisms by checking the behavior only on products of elementary tensors (and checking the multiplicative identity is preserved, which is always easy). We give some concrete examples before stating some general theorems.
Example 7.7. For ideals $I$ and $J$ in $R$, there is an isomorphism $\varphi: R / I \otimes_{R} R / J \rightarrow$ $R /(I+J)$ of $R$-modules where $\varphi(\bar{x} \otimes \bar{y})=\overline{x y}$. Then $\varphi(\overline{1} \otimes \overline{1})=\overline{1}$ and

$$
\varphi\left((\bar{x} \otimes \bar{y})\left(\bar{x}^{\prime} \otimes \bar{y}^{\prime}\right)\right)=\varphi\left(\overline{x x^{\prime}} \otimes \overline{y y^{\prime}}\right)=\overline{x x^{\prime} y y^{\prime}}=\overline{x y} \overline{x^{\prime} y^{\prime}}=\varphi(\bar{x} \otimes \bar{y}) \varphi\left(\bar{x}^{\prime} \otimes \bar{y}^{\prime}\right) .
$$

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

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

$$
\begin{aligned}
\varphi\left(\left(f_{1}(X) \otimes g_{1}(Y)\right)\left(f_{2}(X) \otimes g_{2}(Y)\right)\right) & =\varphi\left(f_{1}(X) f_{2}(X) \otimes g_{1}(Y) g_{2}(Y)\right) \\
& =f_{1}(X) f_{2}(X) g_{1}(Y) g_{2}(Y) \\
& =f_{1}(X) g_{1}(Y) f_{2}(X) g_{2}(Y) \\
& =\varphi\left(f_{1}(X) \otimes g_{1}(Y)\right) \varphi\left(f_{2}(X) \otimes g_{2}(Y)\right),
\end{aligned}
$$

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 $\varphi$ is multiplicative on pairs of monomial tensors $X^{i} \otimes Y^{j}$.)

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

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

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

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

$$
t=\sum_{i, j} \frac{a_{i}(x)}{b(x)} \otimes_{F} \frac{c_{j}(y)}{d(y)}=\left(\frac{1}{b(x)} \otimes \frac{1}{d(y)}\right) \sum_{i, j} a_{i}(x) \otimes c_{j}(y) .
$$

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

$$
\frac{\sum_{i, j} a_{i}(x) c_{j}(y)}{b(x) d(y)}
$$

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

To show $F(x) \otimes_{F} F(y)$ is isomorphic as an $F$-algebra to the rational functions in $F(x, y)$ having a denominator in the factored form $b(x) d(y)$, show that the multiplication mapping $F(x) \otimes_{F} F(y) \rightarrow F(x, y)$ given by $f(x) \otimes g(y) \mapsto f(x) g(y)$ on elementary tensors is an embedding of $F$-algebras. That it is an $F$-algebra homomorphism follows by the same argument used in Example 7.8. It is left to the reader to show the kernel is 0 from the known fact that the multiplication mapping $F[x] \otimes_{F} F[y] \rightarrow F[x, y]$ is injective. (Hint: Justify the idea of clearing denominators.) Thus $F(x) \otimes_{F} F(y)$ is an integral domain that is not a field, since its image in $F(x, y)$ is not a field: the image contains $F[x, y]$ but is smaller
than $F(x, y)$. Concretely, the fact that $1-x y$ is the image of $1 \otimes 1-x \otimes y$ but $1 /(1-x y)$ is not in the image shows $1 \otimes 1-x \otimes y$ is not invertible in $F(x) \otimes_{F} F(y)$. (In terms of localizations, $F(x) \otimes_{F} F(y)$ is isomorphic as an $F$-algebra to the localization of $F[x, y]$ at the multiplicative set of all products $b(x) d(y)$.)
Example 7.11. For any $R$-module $M$, there is an $S$-linear map

$$
S \otimes_{R} \operatorname{End}_{R}(M) \longrightarrow \operatorname{End}_{S}\left(S \otimes_{R} M\right)
$$

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

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

Example 7.12. If $I$ is an ideal in $R$ and $A$ is an $R$-algebra, $R / I \otimes_{R} A \cong A / I A$ first as $R$ modules, then as $R$-algebras (the $R$-linear isomorphism is also multiplicative and preserves identities), and finally as $R / I$-algebras since the isomorphism is $R / I$-linear too.
Theorem 7.13. Let $A, B$, and $C$ be $R$-algebras. The standard $R$-module isomorphisms

$$
\begin{aligned}
A \otimes_{R} B & \cong B \otimes_{R} A \\
A \otimes_{R}(B \times C) & \cong\left(A \otimes_{R} B\right) \times\left(A \otimes_{R} C\right) \\
\left(A \otimes_{R} B\right) \otimes_{R} C & \cong A \otimes_{R}\left(B \otimes_{R} C\right) .
\end{aligned}
$$

are all $R$-algebra isomorphisms.
The distributivity of $\otimes$ over $\times$ suggests denoting the direct product of algebras as a direct sum with $\oplus$.

Proof. Exercise. Note the direct product of two $R$-algebras is the direct sum as $R$-modules with componentwise multiplication, so first just treat the direct product as a direct sum.
Corollary 7.14. For $R$-algebras $A$ and $B, A \otimes_{R} B^{n} \cong\left(A \otimes_{R} B\right)^{n}$ as $R$-algebras.
Proof. Induct on $n$. Note $B^{n}$ here means the $n$-fold product ring, not $B^{\otimes n}$.
We turn now to base extensions. Fix a homomorphism $f: R \rightarrow S$ 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 $r a=f(r) a$, and this makes $A$ into an $R$-algebra (restriction of scalars). More interesting is extension of scalars. For this we need a lemma.

Lemma 7.15. If $A, A^{\prime}, B$, and $B^{\prime}$ are all $R$-algebras and $A \xrightarrow{\varphi} A^{\prime}$ and $B \xrightarrow{\psi} B^{\prime}$ are $R$-algebra homomorphisms then the $R$-linear map $A \otimes_{R} B \xrightarrow{\varphi \otimes \psi} A^{\prime} \otimes_{R} B^{\prime}$ is an $R$-algebra homomorphism.
Proof. Exercise.
Theorem 7.16. Let $A$ be an $R$-algebra.
(1) The base extension $S \otimes_{R} A$, which is both an $R$-algebra and an $S$-module, is an $S$-algebra by its $S$-scaling.
(2) If $A \xrightarrow{\varphi} B$ is an $R$-algebra homomorphism then $S \otimes_{R} A \xrightarrow{1 \otimes \varphi} S \otimes_{R} B$ is an $S$-algebra homomorphism.

Proof. 1) We just need to check multiplication in $S \otimes_{R} A$ commutes with $S$-scaling (not just $R$-scaling): $s\left(t t^{\prime}\right)=(s t) t^{\prime}=t\left(s t^{\prime}\right)$. Since all three expressions are additive in $t$ and $t^{\prime}$, it suffices to check this when $t$ and $t^{\prime}$ are elementary tensors:

$$
s\left(\left(s_{1} \otimes a_{1}\right)\left(s_{2} \otimes a_{2}\right)\right) \stackrel{?}{=}\left(s\left(s_{1} \otimes a_{1}\right)\right)\left(s_{2} \otimes a_{2}\right) \stackrel{?}{=}\left(s_{1} \otimes a_{1}\right)\left(s\left(s_{2} \otimes a_{2}\right)\right) .
$$

From the way $S$-scaling on $S \otimes_{R} A$ is defined, all these products equal $s s_{1} s_{2} \otimes a_{1} a_{2}$.
2) For an $R$-algebra homomorphism $A \xrightarrow{\varphi} B$, the base extension $S \otimes_{R} A \xrightarrow{1 \otimes \varphi} S \otimes_{R} B$ is $S$-linear and it is an $R$-algebra homomorphism by Lemma 7.15. Therefore it is an $S$-algebra homomorphism.

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

Example 7.17. Let $I$ be an ideal in $R\left[X_{1}, \ldots, X_{n}\right]$. Check the $S$-module isomorphism $S \otimes_{R} R\left[X_{1}, \ldots, X_{n}\right] / I \cong S\left[X_{1}, \ldots, X_{n}\right] /\left(I \cdot S\left[X_{1}, \ldots, X_{n}\right]\right)$ from Corollary 2.24 is an $S$ algebra isomorphism.

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

$$
S \otimes_{R} R[X] /(h(X)) \cong S[X] /\left(h^{f}(X)\right),
$$

where $h^{f}(X)$ is the result of applying $f: R \rightarrow S$ to the coefficients of $h(X)$. (If $f: \mathbf{Z} \rightarrow$ $\mathbf{Z} / p \mathbf{Z}$ is reduction $\bmod p$, for instance, then $h^{f}(X)=h(X) \bmod p$.) This isomorphism is particularly convenient, as it lets us compute a lot of tensor products of fields.
Example 7.18. Writing $\mathbf{Q}(\sqrt{2})$ as $\mathbf{Q}[X] /\left(X^{2}-2\right)$ (as a $\mathbf{Q}$-algebra), we have

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

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

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

as $\mathbf{R}$-algebras since $X^{3}-2$ has irreducible factors in $\mathbf{R}[X]$ of degree 1 and 2 .
More generally, we have an $\mathbf{R}$-algebra isomorphism

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

The $\mathbf{Q}$-linear embedding $\mathbf{Q}(\sqrt[n]{2}) \hookrightarrow \mathbf{R}$ extends to an $\mathbf{R}$-linear embedding $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{Q}(\sqrt[n]{2}) \rightarrow$ $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ that is multiplicative (it suffices to check that on elementary tensors), so the ring $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ contains a subring isomorphic to $\mathbf{R}[X] /\left(X^{n}-2\right)$. When $n$ is odd, $X^{n}-2$ has one linear factor and $(n-1) / 2$ quadratic irreducible factors in $\mathbf{R}[X]$, so $\mathbf{R}[X] /\left(X^{n}-2\right) \cong$ $\mathbf{R} \times \mathbf{C}^{(n-1) / 2}$ as $\mathbf{R}$-algebras. Therefore $\mathbf{R}[X] /\left(X^{n}-2\right)$ contains $2^{1+(n-1) / 2}=2^{(n+1) / 2}$ square roots of 1 . Letting $n \rightarrow \infty$ shows $\mathbf{R} \otimes_{\mathbf{Q}} \mathbf{R}$ contains infinitely many square roots of 1 .
Example 7.19. We revisit Example 7.3, using Example 7.17:

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

as $\mathbf{R}$-algebras.

[^12]Let's make the $\mathbf{R}$-algebra isomorphism $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C} \cong \mathbf{C} \times \mathbf{C}$ explicit, as it is not $z \otimes w \mapsto$ $(z, w)$. Tracing the effect of the isomorphisms on elementary tensors,
$z \otimes(a+b i) \mapsto z \otimes(a+b X) \mapsto z a+z b X \mapsto(z a+z b i, z a+a b(-i))=(z(a+b i), z(a-b i))$, so $z \otimes w \mapsto(z w, z \bar{w})$. Thus $1 \otimes 1 \mapsto(1,1), z \otimes 1 \mapsto(z, z)$, and $1 \otimes w \mapsto(w, \bar{w})$.

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

Example 7.20. We have

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

which is a field because $X^{2}-3$ is irreducible in $\mathbf{Q}(\sqrt{2})[X]$.
Example 7.21. As an example of a tensor product involving a finite field and a ring,

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

since $X^{2}+1=(X-2)(X-3)$ in $(\mathbf{Z} / 5 \mathbf{Z})[X]$.
A general discussion of tensor products of fields is in [3, Sect. 8.18].
Theorem 7.22. If $A$ is $R$-algebra and $B$ is an $S$-algebra, then the $S$-module structure on the $R$-algebra $A \otimes_{R} B$ makes it an $S$-algebra, and

$$
A \otimes_{R} B \cong\left(A \otimes_{R} S\right) \otimes_{S} B
$$

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.

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


Then there is a ring isomorphism

$$
F \otimes_{K} L \cong\left(F \otimes_{K} E\right) \otimes_{E} L
$$

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

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

$$
S \otimes_{R}\left(A \otimes_{R} B\right) \rightarrow\left(S \otimes_{R} A\right) \otimes_{S}\left(S \otimes_{R} B\right)
$$

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

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 $S \otimes_{R}\left(A \otimes_{R} B\right)$ ), so it is an $S$-algebra isomorphism.

## 8. The tensor algebra of an $R$-module

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

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

To enlarge an $R$-module $M$ to an $R$-algebra, we want to multiply elements in $M$ without having any multiplication defined in advance. (As in Section $7, R$-algebras are associative.) The "most general" product $m_{1} m_{2}$ for $m_{1}$ and $m_{2}$ in $M$ should be bilinear in $m_{1}$ and $m_{2}$, so we want this product to be the elementary tensor $m_{1} \otimes m_{2}$, which lives not in $M$ but in $M^{\otimes 2}$. Similarly, an expression like $m_{1} m_{2}+m_{3} m_{4} m_{5}$ using five elements from $M$ should be $m_{1} \otimes m_{2}+m_{3} \otimes m_{4} \otimes m_{5}$ in $M^{\otimes 2} \oplus M^{\otimes 3}$. This suggests creating an $R$-algebra as

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

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

$$
\begin{equation*}
\sum_{k \geq 0} t_{k} \cdot \sum_{\ell \geq 0} t_{\ell}^{\prime}=\sum_{n \geq 0}\left(\sum_{k+\ell=n} t_{k} \otimes t_{\ell}^{\prime}\right), \tag{8.1}
\end{equation*}
$$

where $t_{k} \otimes t_{\ell}^{\prime} \in M^{\otimes n}$ if $k+\ell=n$. To show this multiplication makes the direct sum of all $M^{\otimes k}$ an $R$-algebra we use the following construction theorem.

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

1) (scaling by $R$ ) $\mu_{k, 0}: M_{k} \times R \rightarrow M_{k}$ and $\mu_{0, \ell}: R \times M_{\ell} \rightarrow M_{\ell}$ are both scaling by $R$ : $\mu_{k, 0}(x, r)=r x$ and $\mu_{0, \ell}(r, y)=r y$ for $r \in R, x \in M_{k}$, and $y \in M_{\ell}$,
2) (associativity) for $k, \ell, n \geq 0$ we have $\mu_{k, \ell+n}\left(x, \mu_{\ell, n}(y, z)\right)=\mu_{k+\ell, n}\left(\mu_{k, \ell}(x, y), z\right)$ in $M_{k+\ell+n}$ for all $x \in M_{k}, y \in M_{\ell}$, and $z \in M_{n}$.
The direct sum $\bigoplus_{k \geq 0} M_{k}$ is an associative $R$-algebra with identity using the multiplication rule $\sum_{k \geq 0} m_{k} \cdot \sum_{\ell \geq 0} m_{\ell}^{\prime}=\sum_{n \geq 0}\left(\sum_{k+\ell=n} \mu_{k, \ell}\left(m_{k}, m_{\ell}^{\prime}\right)\right)$.
Proof. The direct sum $\bigoplus_{k \geq 0} M_{k}$ is automatically an $R$-module. It remains to check (i) the multiplication defined on the direct sum is $R$-bilinear, (ii) $1 \in R=M_{0}$ is a multiplicative identity, and (iii) multiplication is associative. The $R$-bilinearity of multiplication is a
bookkeeping exercise left to the reader. In particular, this includes distributivity of multiplication over addition. To prove multiplication has 1 as an identity and is associative, it suffices by distributivity to consider only multiplication with factors from direct summands $M_{k}$, in which case we can use the two properties of the maps $\mu_{k, \ell}$ in the theorem.

Lemma 8.2. For $k, \ell \geq 0$, there is a unique bilinear map $\beta_{k, \ell}: M^{\otimes k} \times M^{\otimes \ell} \rightarrow M^{\otimes(k+\ell)}$ where $\beta_{k, 0}(t, r)=r t, \beta_{0, \ell}(r, t)=r t$, and for $k, \ell \geq 1, \beta_{k, \ell}\left(m_{1} \otimes \cdots \otimes m_{k}, m_{1}^{\prime} \otimes \cdots \otimes m_{\ell}^{\prime}\right)=$ $m_{1} \otimes \cdots \otimes m_{k} \otimes m_{1}^{\prime} \otimes \cdots \otimes m_{\ell}^{\prime}$ on pairs of elementary tensors.
Proof. The cases $k=0$ and $\ell=0$ are trivial, so let $k, \ell \geq 1$. It suffices to construct a bilinear $\beta_{k, \ell}$ with the indicated values on pairs of elementary tensors; uniqueness is then automatic since elementary tensors span $M^{\otimes k}$ and $M^{\otimes \ell}$.

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

$$
f\left(m_{1}, \ldots, m_{k}, m_{1}^{\prime}, \ldots, m_{\ell}^{\prime}\right)=m_{1} \otimes \cdots \otimes m_{k} \otimes m_{1}^{\prime} \otimes \cdots \otimes m_{\ell}^{\prime} .
$$

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

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

$$
\begin{aligned}
\beta_{k, \ell}\left(m_{1} \otimes \cdots \otimes m_{k}, m_{1}^{\prime} \otimes \cdots \otimes m_{\ell}^{\prime}\right) & =f\left(m_{1}, \cdots, m_{k}, m_{1}^{\prime}, \ldots, m_{\ell}\right) \\
& =m_{1} \otimes \cdots \otimes m_{k} \otimes m_{1}^{\prime} \otimes \cdots \otimes m_{\ell}^{\prime} .
\end{aligned}
$$

Theorem 8.3. The $M$-module $\bigoplus_{k \geq 0} M^{\otimes k}$ is an $R$-algebra using the multiplication in (8.1).
Proof. Use Theorem 8.1 with $M_{k}=M^{\otimes k}$ and $\mu_{k, \ell}=\beta_{k, \ell}$ from Lemma 8.2. The first property of the maps $\mu_{k, \ell}$ in Theorem 8.1 is automatic from the definition of $\beta_{k, 0}$ and $\beta_{0, \ell}$. To prove the second property of the maps $\mu_{k, \ell}$ in Theorem 8.1, namely $\mu_{k, \ell+n}\left(t_{1}, \mu_{\ell, n}\left(t_{2}, t_{3}\right)\right)=$ $\mu_{k+\ell, n}\left(\mu_{k, \ell}\left(t_{1}, t_{2}\right), t_{3}\right)$ in $M^{\otimes(k+\ell+n)}$ for all $t_{1} \in M^{\otimes k}, t_{2} \in M^{\otimes \ell}$, and $t_{3} \in M^{\otimes n}$, by multilinearity of each $\mu_{k, \ell}=\beta_{k, \ell}$ it suffices to consider the case when each $t_{i}$ is an elementary tensor, in which case the equality is a simple calculation.

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

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

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

Example 8.6. If $M$ is a finite free $R$-module with basis $e_{1}, \ldots, e_{n}$ then $T(M)$ is the polynomial ring over $R$ in $n$ noncommuting indeterminates $e_{1}, \ldots, e_{n}$ : in $M^{\otimes 2}, e_{i} \otimes e_{j} \neq e_{j} \otimes e_{i}$ when $i \neq j$, which says in $T(M)$ that $e_{i} e_{j} \neq e_{j} e_{i}$.

Remark 8.7. The tensor product construction of the polynomial ring over $R$ in $n$ noncommuting indeterminates is quite different from that of the tensor product construction of the commutative polynomial ring $R\left[X_{1}, \ldots, X_{n}\right]$ : the former is the tensor algebra $T\left(R^{n}\right)$ of a free $R$-module of rank $n$ while the latter is $R[X]^{\otimes n}$ (Example 7.8).

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

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

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

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

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

$$
F_{k}\left(m_{1} \otimes \cdots \otimes m_{k}\right)=f\left(m_{1}\right) \cdots f\left(m_{k}\right)
$$

Define $F_{0}: M^{\otimes 0} \rightarrow A$ by $F_{0}(r)=r \cdot 1_{A}$ for $r \in R$. Finally, define $F: T(M) \rightarrow A$ by

$$
F\left(\sum_{k \geq 0} t_{k}\right)=\sum_{k \geq 0} F_{k}\left(t_{k}\right)
$$

where $t_{k} \in M^{\otimes k}$ and $t_{k}=0$ for large $k$. Since each $F_{k}$ is $R$-linear and $F_{0}(1)=1_{A}, F$ is $R$-linear and $F(1)=1_{A}$. To prove $F$ is multiplicative, by linearity it suffices to check $F(x y)=F(x) F(y)$ where $x$ is an elementary tensor in some $M^{\otimes k}$ and $y$ is an elementary tensor in some $M^{\otimes \ell}$. The cases $k=0$ and $\ell=0$ are the linearity of $F$. If $k, \ell \geq 1$, write $x=m_{1} \otimes \cdots \otimes m_{k}$ and $y=m_{1}^{\prime} \otimes \cdots \otimes m_{\ell}^{\prime}$. Then $x y=m_{1} \otimes \cdots \otimes m_{k} \otimes m_{1}^{\prime} \otimes \cdots \otimes m_{\ell}^{\prime}$ in $T(M)$, so

$$
F(x y)=F_{k+\ell}(x y)=f\left(m_{1}\right) \cdots f\left(m_{k}\right) f\left(m_{1}^{\prime}\right) \cdots f\left(m_{\ell}^{\prime}\right)
$$

and

$$
\begin{aligned}
F(x) F(y) & =F_{k}(x) F_{\ell}(y) \\
& =\left(f\left(m_{1}\right) \cdots f\left(m_{k}\right)\right)\left(f\left(m_{1}^{\prime}\right) \cdots f\left(m_{\ell}^{\prime}\right)\right) \\
& =f\left(m_{1}\right) \cdots f\left(m_{k}\right) f\left(m_{1}^{\prime}\right) \cdots f\left(m_{\ell}^{\prime}\right) .
\end{aligned}
$$

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

## References

[1] S. Friedl, Eta invariants as sliceness obstructions and their relation to Casson-Gordon invariants, Algebr. Geom. Topol. 4 (2004), 893-934.
[2] H. V. Henderson and F. Pukelsheim, On the History of the Kronecker Product, Linear and Multilinear Algebra 14 (1983), 113-120.
[3] N. Jacobson, "Basic Algebra II," 2nd ed., W. H. Freeman \& Co., New York, 1989.
[4] E. Schrödinger, Discussion of Probability Relations between Separated Systems, Math. Proc. Camb. Philos. Society 31 (1935), 555-563. Online at https://www.cambridge.org/core/ journals/mathematical-proceedings-of-the-cambridge-philosophical-society/article/discussion-of-probability-relations-between-separated-systems/C1C71E1AA5BA56EBE6588AAACB9A222D.


[^0]:    ${ }^{1}$ See https://math.stackexchange.com/questions/3452281.

[^1]:    ${ }^{2}$ The first $m \otimes n$ is in $(\operatorname{ker} \varphi) \otimes_{R} N$, while the second $m \otimes n$ is in $M \otimes_{R} N$.

[^2]:    ${ }^{3} P / \mathfrak{m} P=0 \Rightarrow P=\mathfrak{m} P \Rightarrow P_{\mathfrak{m}}=\mathfrak{m} P_{\mathfrak{m}} \Rightarrow P_{\mathfrak{m}}=0$ by Nakayama's lemma. From $P_{\mathfrak{m}}=0$ for all maximal ideals $\mathfrak{m}, P=0$ : for all $x \in P, x=0$ in $P_{\mathfrak{m}}$ implies $a x=0$ in $P$ for some $a \in R-\mathfrak{m}$. Thus $\operatorname{Ann}_{R}(x)$ is not in any maximal ideal of $R$, so $\operatorname{Ann}_{R}(x)=R$ and thus $x=1 \cdot x=0$.

[^3]:    ${ }^{4}$ Any two nonzero elements of $K$ are $R$-linearly dependent, so if $K$ were a free $R$-module then it would have a basis of size $1: K=R x$ for some $x \in K$. Therefore $x^{2}=r x$ for some $r \in R$, so $x=r \in R$, which implies $K \subset R$, so $K=R$.

[^4]:    ${ }^{5}$ More generally, $R / I \otimes_{R} R / I \cong R / I$ as $R$-modules and then also as $R / I$-modules where $R / I$ scales elementary tensors $x \otimes y$ in $R / I \otimes_{R} R / I$ on the left $(a(x \otimes y)=a x \otimes y)$.

[^5]:    ${ }^{6}$ We allow infinite or even non-algebraic extensions, such as $\mathbf{R} / \mathbf{Q}$.
    ${ }^{7}$ If we drop finite-dimensionality assumptions, (5.1), (5.2), and (5.3) are all still injective but generally not surjective.

[^6]:    ${ }^{8}$ Exception: $V^{\prime}=V$ or $W=0$, and $W^{\prime}=W$ or $V=0$.

[^7]:    ${ }^{9}$ A nonzero vector in $H$, viewed as a quantum state, is often scaled to have length 1 , but this still allows ambiguity up to scaling by a complex number $e^{i \theta}$ of absolute value 1 , with $\theta$ called a phase-factor.

[^8]:    ${ }^{10}$ If we used $v_{\sigma(i)}$ instead of $v_{\sigma^{-1}(i)}$ in the definition of $P_{\sigma}$ then we'd have $P_{\sigma} \circ P_{\tau}=P_{\tau \sigma}$.

[^9]:    ${ }^{11}$ Here and elsewhere in this section, tensor products are always over $K: \otimes=\otimes_{K}$.

[^10]:    ${ }^{12}$ While $V^{\otimes(s-1)} \otimes\left(V^{\vee}\right) \otimes V^{\otimes(k-s)}$ is not in standard form, the meaning of contraction using the $r$ th tensorand $V$ and the $s$ th tensorand $V^{\vee}$ should be obvious to the reader.

[^11]:    ${ }^{13}$ Raising or lowering indices describes what happens to components in a basis. The basis undergoes the opposite change (lowering or raising its indices).

[^12]:    ${ }^{14}$ Not all ideals in $R[X]$ have to be principal, but this is just an example.

