IV.E. Endomorphisms

Recall from IV.B.21-IV.B.22 that for a free module \( M \) of rank \( n \) over a commutative ring \( R \), sending endomorphisms to their matrix (with respect to some base) yields a map

\[
\text{End}_R(M) \xrightarrow{\sim} M_n(R)
\]

which is in fact an isomorphism of rings and of \( R \)-modules. What happens if \( M \) is no longer free? In this section we will give an answer to this question in the case (henceforth assumed) that \( R \) is a PID. We begin with some easy

**IV.E.1. Examples.** (a) Suppose \( M = Rz \cong R/(d) \) is a cyclic \( R \)-module, and note that \( rz \) corresponds to \( \bar{r} \) under the isomorphism. The map

\[
\text{End}_R(M) \to R/(d) \\
\eta \mapsto \eta(\bar{1})
\]

(IV.E.2)

is an isomorphism of rings and \( R \)-modules. [Why? Clearly (IV.E.2) is an \( R \)-module homomorphism. It is injective because \( \eta \) is determined by where it sends a generator; and surjective because it sends

\[
\mu_r := \{\text{multiplication by } r\} \mapsto \bar{r}
\]

for any \( \bar{r} \in R/(d) \). So then \( \text{End}_R(M) \) consists entirely of \( \mu_r \)'s, and (IV.E.2) sends composition to multiplication.]

(b) If \( M \cong (R/(d))^\oplus n \), then writing \( \bar{e}_i \) for the “standard” generators \((\bar{e}_1 = (\bar{1},0,\ldots,0), \text{etc.})\), writing \( \eta(\bar{e}_j) = \sum_i r_{ij} \bar{e}_i \) defines a map

\[
\text{End}_R(M) \to M_n(R/(d)) \\
\eta \mapsto (\bar{r}_{ij})
\]

which one also shows is an isomorphism (of rings and \( R \)-modules), by combining the approach for free modules with that in (a).

(c) On the other hand, if \( M \cong \bigoplus R/(p_i) \) with \( p_i \) distinct primes of \( R \), then by Schur’s Lemma IV.B.32, \( \text{Hom}_R(R/(p_i), R/(p_j)) = \{0\} \) for
Combining this with (a) yields
\[ \text{End}_R(M) \cong \bigoplus_i \text{End}_R(R/(p_i)) \cong \bigoplus_i R/(p_i). \]

Alternatively, one can use the Chinese Remainder Theorem (see the proof of IV.C.25) to write \( M \cong R/(\prod p_i) \), apply (a), and use the CRT again on the RHS.

(d) Finally, if \( M \cong \bigoplus_i (R/(p_i))^\oplus n_i \), then combining Schur’s Lemma with (b) yields
\[ \text{End}_R(M) \cong \bigoplus_i M_{n_i}(R/(p_i)), \]
which is again an isomorphism as rings and as \( R \)-modules.

Now we turn to the general case: let
\[ M = Rz_1 \oplus \cdots \oplus Rz_s \cong \underbrace{Rz_1 \oplus \cdots \oplus Rz_\ell}_{\text{tor}(M)} \oplus R^t, \]
where \( \ell + t = s \), \( \text{Ann}(z_i) = (\delta_i), \delta_1, \ldots, \delta_\ell, \) and \( \delta_{\ell+1} = \cdots = \delta_s = 0 \).

We can present \( M \) in terms of generators and relations as
\[ M \cong R^s / K = \frac{R\langle e_1, \ldots, e_s \rangle}{\langle \delta_1 e_1, \ldots, \delta_\ell e_\ell \rangle}. \]

Our aim is to get a description of the endomorphism ring
\[ S := \text{End}_R(M) \]
in the spirit of the above examples, but in terms of the \( \{\delta_i\} \).

Recall the matrix description of endomorphisms of \( R^s \)
\[ \theta : \text{End}_R(R^s) \xrightarrow{\cong} M_s(R) \]
\[ \eta \mapsto e[\bar{\eta}] =: (n_{ij}) =: N, \]
where \( \bar{\eta}(e_j) = \sum_i n_{ij} e_i \). Given \( \bar{\eta} \in \text{End}_R(R^s) \), we can ask when it makes sense modulo \( K \), as an endomorphism of \( M (= R^s / K) \). Evidently,
- \( \bar{\eta} \) defines an element \( \eta \in S \iff \bar{\eta}(K) \subseteq K \); and
- \( \bar{\eta} \) defines the zero element in \( S \iff \bar{\eta}(R^s) \subseteq K. \)
For \( \bar{x} \in R^s \), we have

\[
\bar{x} \in K \iff \bar{x} = \sum_{i=1}^{\ell} d_i r_i e_i \quad \iff \quad e[\bar{x}] \in \left( \begin{array}{c} \delta_1 \\ \vdots \\ \delta_s \end{array} \right) R^s =: DR^s
\]

(thinking of \( R^s \) as column vectors on the RHS). Hence

\[
\bar{\eta}(K) \subseteq K \iff \bar{\eta}(\bar{x}) \in K \quad (\forall \bar{x} \in K)
\]

[apply \( \varepsilon[\cdot] \sim \)] \iff \( NDv \in DR^s \quad (\forall v \in R^s)\)

[apply to \( v = e_1, \ldots, e_s \sim \)] \iff \( ND \subset DM_s(R) \)

def. \( N \in M_s \),

and

\[
\bar{\eta}(R^s) \subseteq K \iff Nv \in DR^s \quad (\forall v \in R^s)
\]

\[
\iff N \in DM_s(R) =: J_S.
\]

Note that \( M_s \) is a subring of \( M_s(R) \): given \( N, N' \in M_s \), we can write

\[
(N'N)D = N'(ND) = N'(DM') = (N'D)M' = (DM)M' = DM''
\]

with \( M, M', M'' \in M_s(R) \); and so \( N'N \in M_s \). Furthermore, \( J_S \subset M_s \) is a (two-sided) ideal: given \( N \in M_s \),

\[
N J_S = NDM_s(R) \subset DM_s(R) = J_S
\]

and \( J_S N = DM_s(R)N \subset DM_s(R) = J_S \).

So \( M_s / J_S \) is a ring (and an \( R \)-module!); and we have the

**IV.E.3. Theorem.** \( \theta \) induces an isomorphism

\[
\bar{\theta} : S \xrightarrow{\cong} M_s / J_S
\]

of rings (and \( R \)-modules).

**Proof.** We just did it! To briefly recapitulate: applying \( \theta = e[\cdot] \) to the numerator and denominator of the RHS of

\[
S = \text{End}_R(M) = \text{End}_R(R^s/K) = \left\{ \bar{\eta} \in \text{End}_R(R^s) \mid \bar{\eta}(K) \subseteq K \right\}
\]

yields exactly \( M_s / J_S \). \( \square \)
IV.E.4. REMARK. Note that we can think of \( \overline{\theta} \) as “taking the matrix with respect to \( z_1, \ldots, z_s \)” even though this is not a base in the standard sense.

Now consider the conditions defining \( M_S \) if \( s = 2 \): keeping in mind that \( \delta_1 | \delta_2 \) (and denoting by \( r_{ij} \) arbitrary elements of \( R \)), we have

\[
N = \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix} \in M_S \iff \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix}
\]

\[
\iff \begin{pmatrix} n_{11} n_{12} \\ n_{21} n_{22} \end{pmatrix} = \begin{pmatrix} \delta_1 r_{11} & \delta_1 r_{12} \\ \delta_2 r_{21} & \delta_2 r_{22} \end{pmatrix}
\]

\[
\iff n_{21} \in (\delta_2);
\]

so \( n_{21} = n'_{21} \delta_2 \), with \( n'_{21} \) and the other \( n_{ij} \) arbitrary elements of \( R \).

(Note that if \( \delta_2 = 0 \neq \delta_1 \), this would make \( n_{21} = 0 \).) Furthermore, we have

\[
N \in J_S \iff \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix} = \begin{pmatrix} \delta_1 & \delta_2 \\ r_{21} & r_{22} \end{pmatrix} \begin{pmatrix} \delta_1 & \delta_2 \\ r_{11} & r_{12} \end{pmatrix}
\]

\[
\iff n_{11}, n_{12} \in (\delta_1) \text{ and } n_{21}, n_{22} \in (\delta_2).
\]

The upshot is that, for elements of \( \overline{\theta}(S) = M_S / J_S \), we need to consider \( n_{11} \) and \( n_{12} \) as elements of \( R / (\delta_1) \), \( n_{21} \) as an element of \( (\delta_2) / (\delta_2) \), and \( n_{22} \) in \( R / (\delta_2) \).

More generally, for any \( s \), this analysis leads to the following specifications for entries in the “regions” of the \( s \times s \) matrix \( N \) (corresponding via \( \overline{\theta} \) to elements of \( S \)) as shown:

\[
\begin{align*}
(I) & \quad i \leq j, \ell : \quad n_{ij} \in R / (\delta_i) \\
(II) & \quad j < i \leq \ell : \quad n_{ij} \in (\delta_j) / (\delta_i) \\
(III) & \quad i > \ell; j \leq \ell : \quad n_{ij} = 0 \\
(IV) & \quad i, j > \ell : \quad n_{ij} \in R
\end{align*}
\]

so we can write \( n_{ij} := n'_{ij} \delta_i \delta_j \) in (II) as above, with \( n'_{ij} \in R / (\delta_j) \). In the event that \( M \) is torsion, \( \ell = s \) and we don’t have regions (III) and (IV).

An immediate consequence is

IV.E.5. COROLLARY. The center of \( S = \text{End}_R(M) \) is \( R \).
PROOF. Let $\varepsilon_{is} \in S$ be the endomorphism with matrix given by\(^{22}\)

$\bar{\vartheta}(\varepsilon_{is}) = e_{is}$. (Note that this is possible because the $(i,s)^{th}$ entry lies in region (I) or (IV), never (II).) This endomorphism sends $z_s \mapsto z_i$ and kills all other $z_j$. So given $\eta \in C(S)$ (in the center), and writing $N = \bar{\vartheta}(\eta)$, we have

$$\eta(z_s) = \eta(\varepsilon_{ss}(z_s)) = \varepsilon_{ss}(\eta(z_s)) = \varepsilon_{ss}(\sum_i n_{is}z_i) = \sum_i n_{is}\varepsilon_{ss}z_i = n_{ss}z_s$$

and

$$\eta(z_i) = \eta(\varepsilon_{js}z_s) = \varepsilon_{js}(\eta(z_s)) = \varepsilon_{js}(\sum_i n_{is}z_i) = \sum_i n_{is}\varepsilon_{js}(z_i) = n_{ss}z_j,$$

so that $\eta$ is simply multiplication by $n_{ss}$ — which, being in region (I) or (IV), can be any element of $R$. □

Assume henceforth that $M$ is torsion. As $S$ is an $R$-module:

(a) if $R = \mathbb{Z}$, then $M = G$ is a finite abelian group, and $S = \text{End}_{\mathbb{Z}}(G)$ also has the structure of a finite abelian group, with a (finite) order; while

(b) if $R = F[\lambda]$, then $M = V$ is an $F$-vector space on which $\lambda$ acts by a linear transformation $T \in \text{End}_F(V)$, and $S = \text{End}_{F[\lambda]}(V)$ itself has the structure of an $F$-vector space, with a (finite) dimension.

So we can take the theory for a test-drive to see if we can compute the italicized numbers. For (a), we have the

IV.E.6. COROLLARY. Consider any finite abelian group, written in the form $G \cong \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_s}$ with $m_1 \mid \cdots \mid m_s$. Then the number of group homomorphisms from $G$ to itself is

$$|\text{End}_{\mathbb{Z}}(G)| = \prod_{j=1}^s m_j^{2s-2j+1}.$$

PROOF. With $S = \text{End}_{\mathbb{Z}}(G)$, one counts the possible choices for the $n_{ij}$ in a matrix $N \in \mathcal{M}_S/\mathcal{J}_S$. For (I) $i \leq j$, $n_{ij} \in \mathbb{Z}/(m_i) = \mathbb{Z}_{m_i}$; while for (II) $i > j$, $n_{ij} = n'_{ij} \frac{m_i}{m_j}$ with $n'_{ij} \in \mathbb{Z}/(m_j) = \mathbb{Z}_{m_j}$. So to compute $|S| = |\mathcal{M}_S/\mathcal{J}_S|$, we simply have to take the product of all

\(^{22}\)Recall that $e_{ij}$ is the matrix with $(i,j)^{th}$ entry 1 and all other entries 0.
entries of the matrix
\[
\begin{pmatrix}
m_1 & m_1 & m_1 & \cdots & m_1 \\
m_1 & m_2 & m_2 & \cdots & m_2 \\
m_1 & m_2 & m_3 & \cdots & m_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
m_1 & m_2 & m_3 \\
\end{pmatrix}
\]
which gives the result. \(\square\)

For (b), notice that
\[
S = \text{End}_{\mathbb{F}[\lambda]}(V) = \{\eta \in \text{End}_{\mathbb{F}}(V) \mid \eta T = T\eta\}
\]
is the centralizer of \(T\). Writing \(\bar{\chi}[T] = B\) and \(\bar{\chi}[\eta] = Z\) with respect to some basis of \(V\), \(S\) is identified with \(23\)
\[(S \cong) \text{End}_{\mathbb{F}[\lambda]}(\mathbb{F}^n) = \{Z \in M_n(\mathbb{F}) \mid ZB = BZ\},\]
the ring of matrices commuting with \(B\).

IV.E.7. COROLLARY. Let \(B \in M_n(\mathbb{F})\), with normal form

\[
nf(\lambda I_n - B) = \text{diag}(1, \ldots, 1, \delta_1(\lambda), \ldots, \delta_s(\lambda)).
\]
Then \(\dim_{\mathbb{F}}(S) = \sum_{j=1}^{s} (2s - 2j + 1) \deg(\delta_j(\lambda)).\)

PROOF. Once again, we use \(\bar{\theta}\) to identify \(S\) with \(s \times s\) matrices \(N\) with entries (I) \(n_{ij} \in \mathbb{F}[\lambda]/(\delta_i(\lambda))\) or (II) \(n_{ij} = n'_{ij}/\delta_j(\lambda)\) (and \(n'_{ij} \in \mathbb{F}[\lambda]/(\delta_j(\lambda))\)). So these \(n_{ij}\)'s each lie in a vector space of dimension (I) \(\deg(\delta_i)\) resp. (II) \(\deg(\delta_j)\), and we can record these degrees in a matrix exactly like that in the last proof. Only this time, to get the dimension of \(S\), we add these entries rather than multiplying them. \(\square\)

Call the transformation \(T\) cyclic if its action on \(V\) makes the latter into a cyclic \(\mathbb{F}[\lambda]\)-module (that is, \(s = 1\)).

IV.E.8. COROLLARY. A linear transformation \(T \in \text{End}_{\mathbb{F}}(V)\) is cyclic \iff the only transformations commuting with \(T\) are polynomials in \(T\).\(\text{\footnote{Here }\lambda\text{ acts on }\mathbb{F}^n\text{ via }B.}\)
PROOF. First let $T$ be an arbitrary transformation, and take $d = \deg(m_T) = \deg(\delta_s)$ to be the degree of the minimal polynomial. The polynomials in $T$ certainly commute with $T$, and so

(IV.E.9) \[ \mathbb{F}[\lambda]/(m_T) \cong \mathbb{F}[T] \hookrightarrow \text{End}_{\mathbb{F}[\lambda]}(V). \]

We have $\dim(\text{RHS}) = d + \sum_{j=1}^{s-1} (2s - 2j + 1) \deg(\delta_j)$ by IV.E.7, and $\dim(\text{LHS}) = d$. But then $V$ is cyclic $\iff s = 1 \iff \dim(\text{RHS})$ is $d \iff$ (IV.E.9) is an isomorphism $\iff$ the centralizer of $T$ consists of polynomials in $T$. \hfill \Box

IV.E.10. EXAMPLES. (i) The matrices commuting with a Jordan block are polynomials in the Jordan block.

(ii) Consider the matrix

\[ B = \begin{pmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix} \]

acting on $V = \mathbb{Q}^4$. This is in rational canonical form, hence the companion matrix for $\delta = \delta_1$ ($s = 1$), and we accordingly write

\[ V = \mathbb{Q}[\lambda]/(\delta(\lambda)), \quad \delta(\lambda) = \lambda^4 + \lambda^3 + \lambda^2 + \lambda + 1. \]

This is cyclic, and so IV.E.8 applies.

But we can also recognize $\delta$ as the 5th cyclotomic polynomial, and thus $V \cong \mathbb{Q}[\zeta_5]$ as the corresponding cyclotomic number field. So IV.E.8 tells us that $\text{End}_{\mathbb{Q}[\lambda]}(V) \cong \mathbb{Q}[\zeta_5]$ realizes the multiplicative action of the number field on itself via $4 \times 4$ rational matrices that are polynomials in $B$. In particular, $B$ corresponds to $\zeta_5$ itself.

If we replace $V$ by $V_C = \mathbb{C}^4$,

\[ V_C = \mathbb{C}[\lambda]/(\delta(\lambda)) = \bigoplus_{j=1}^4 \mathbb{C}[\lambda]/(\lambda - \zeta_5^j) \cong \mathbb{C}^4 \]

$\implies$ $\text{End}_{\mathbb{C}[\lambda]}(V_C) = \mathbb{C} \times \mathbb{C} \times \mathbb{C} \times \mathbb{C}$ is represented by diagonal matrices with respect to the (complex) eigenbasis for $B$. 


Notice that in going from $\mathbb{Q}$ to $\mathbb{C}$, the dimension as a vector space (over $\mathbb{Q}$ resp. $\mathbb{C}$) does not change, but the ring structure does dramatically — from a field to a non-domain!

(iii) Let $V = \mathbb{C}^3$. Recall from Example IV.D.12 that

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

is similar to its rational and Jordan forms

$$B' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 3 \end{pmatrix} \quad \text{and} \quad B'' = \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix}.$$

From $B'$, we see that $s = 2$, $\delta_1 = \lambda$ and $\delta_2 = \lambda^2 - 3\lambda$, from which IV.E.7 yields

$$\dim_{\mathbb{C}}(\text{End}_{\mathbb{C}[\lambda]}(V)) = 3 \deg(\delta_1) + 1 \deg(\delta_2) = 5.$$

But what the ring structure of $S = \text{End}_{\mathbb{C}[\lambda]}(V)$ is like, is much clearer from $B''$, which yields the decomposition into primary cyclic submodules $V \cong (\mathbb{C}[\lambda]/(\lambda))^\oplus 2 \oplus \mathbb{C}[\lambda]/(\lambda - 3)$. From there, we can use IV.E.3(d) to compute $S \cong M_2(\mathbb{C}) \times \mathbb{C}$ as a ring, since both $\mathbb{C}[\lambda]/(\lambda)$ and $\mathbb{C}[\lambda]/(\lambda - 3)$ are isomorphic to $\mathbb{C}$ as rings.