(2) If A is an $n \times n$ matrix with coefficients in F, an element λ is called an *eigenvalue* of A with corresponding eigenvector v if v is a nonzero $n \times 1$ column vector such that $Av = \lambda v$.
(3) If λ is an eigenvalue of the linear transformation T, the set $\{v \in V \mid T(v) = \lambda v\}$ is called the *eigenspace* of T corresponding to the eigenvalue λ. Similarly, if λ is an eigenvalue of the $n \times n$ matrix A, the set of $n \times 1$ matrices v with $Av = \lambda v$ is called the *eigenspace* of A corresponding to the eigenvalue λ.

Note that if we fix a basis $\mathcal{B}$ of V then any linear transformation T of V has an associated $n \times n$ matrix A. Conversely, if A is any $n \times n$ matrix then the map T defined by $T(v) = Av$ for $v \in V$, where the v on the right is the $n \times 1$ vector consisting of the coordinates of v with respect to the fixed basis $\mathcal{B}$ of V, is a linear transformation of V. Then v is an eigenvector of T with corresponding eigenvalue λ if and only if the coordinate vector of v with respect to $\mathcal{B}$ is an eigenvector of A with eigenvalue λ. In other words, the eigenvalues for the linear transformation T are the same as the eigenvalues for the matrix A of T with respect to any fixed basis for V.

Definition. The determinant of a linear transformation from V to V is the determinant of any matrix representing the linear transformation (note that this does not depend on the choice of the basis used).

Proposition 12. The following are equivalent:
(1) λ is an eigenvalue of T
(2) $\lambda I - T$ is a singular linear transformation of V
(3) $\det(\lambda I - T) = 0$.

Proof: Since λ is an eigenvalue of T with corresponding eigenvector v if and only if v is a nonzero vector in the kernel of $\lambda I - T$, it follows that (1) and (2) are equivalent. (2) and (3) are equivalent by our results on determinants.

Definition. Let x be an indeterminate over F. The polynomial $\det(xI - T)$ is called the *characteristic polynomial* of T and will be denoted $c_T(x)$. If A is an $n \times n$ matrix with coefficients in F, $\det(xI - A)$ is called the *characteristic polynomial* of A and will be denoted $c_A(x)$.

It is easy to see by expanding the determinant that the characteristic polynomial of either T or A is a monic polynomial of degree $n = \dim V$. Proposition 12 says that the set of eigenvalues of T (or A) is precisely the set of roots of the characteristic polynomial of T (of A, respectively). In particular, T has at most n distinct eigenvalues.

We have seen that V considered as a module over $F[x]$ via the linear transformation T is a torsion $F[x]$-module. Let $m(x) \in F[x]$ be the unique monic polynomial generating the annihilator of V in $F[x]$. Equivalently, $m(x)$ is the unique monic polynomial of minimal degree annihilating V (i.e., such that $m(T)$ is the 0 linear transformation), and if $f(x) \in F[x]$ is any polynomial annihilating V, $m(x)$ divides $f(x)$. Since the ring of all $n \times n$ matrices over F is isomorphic to the collection of all linear transformations of V to itself (an isomorphism is obtained by choosing a basis for V), it follows that for any $n \times n$ matrix A over F there is similarly a unique monic polynomial of minimal degree with $m(A)$ the zero matrix.

Definition. The unique monic polynomial which generates the ideal $\text{Ann}(V)$ in $F[x]$ is called the *minimal polynomial* of T and will be denoted $m_T(x)$. The unique monic polynomial of smallest degree which when evaluated at the matrix A is the zero matrix is called the *minimal polynomial* of A and will be denoted $m_A(x)$.

It is easy to see (cf. Exercise 5) that the degrees of these minimal polynomials are at most n^2 where n is the dimension of V. We shall shortly prove that the minimal polynomial for T is a divisor of the characteristic polynomial for T (this is the *Cayley–Hamilton Theorem*), and similarly for A, so in fact the degrees of these polynomials are at most n.

We now describe the *rational canonical form* of the linear transformation T (respectively, of the $n \times n$ matrix A). By Theorem 5 we have an isomorphism

$$V \cong F[x]/(a_1(x)) \oplus F[x]/(a_2(x)) \oplus \cdots \oplus F[x]/(a_m(x)) \tag{12.1}$$

of $F[x]$-modules where $a_1(x), a_2(x), \dots, a_m(x)$ are polynomials in $F[x]$ of degree at least one with the divisibility conditions

$$a_1(x) \mid a_2(x) \mid \cdots \mid a_m(x).$$

These invariant factors $a_i(x)$ are only determined up to a unit in $F[x]$ but since the units of $F[x]$ are precisely the nonzero elements of F (i.e., the nonzero constant polynomials), we may make these polynomials *unique* by stipulating that they be *monic*.

Since the annihilator of V is the ideal $(a_m(x))$ (part (3) of Theorem 5), we immediately obtain:

Proposition 13. The minimal polynomial $m_T(x)$ is the largest invariant factor of V. All the invariant factors of V divide $m_T(x)$.

We shall see below how to calculate not only the minimal polynomial for T but also the other invariant factors.

We now choose a basis for each of the direct summands for V in the decomposition (1) above for which the matrix for T is quite simple. Recall that the linear transformation T acting on the left side of (1) is the element x acting by multiplication on each of the factors on the right side of the isomorphism in (1).

We have seen in the example following Proposition 1 of Chapter 11 that the elements $1, \bar{x}, \bar{x}^2, \dots, \bar{x}^{k-1}$ give a basis for the vector space $F[x]/(a(x))$ where $a(x) = x^k + b_{k-1}x^{k-1} + \cdots + b_1 x + b_0$ is any monic polynomial in $F[x]$ and $\bar{x} = x \bmod (a(x))$. With respect to this basis the linear transformation of multiplication by x acts in a simple manner:

$$x : \quad \begin{array}{ll} 1 & \mapsto \bar{x} \\ \bar{x} & \mapsto \bar{x}^2 \\ \bar{x}^2 & \mapsto \bar{x}^3 \\ & \vdots \\ \bar{x}^{k-2} & \mapsto \bar{x}^{k-1} \\ \bar{x}^{k-1} & \mapsto \bar{x}^k = -b_0 - b_1\bar{x} - \cdots - b_{k-1}\bar{x}^{k-1} \end{array}$$

where the last equality is because $\bar{x}^k + b_{k-1}\bar{x}^{k-1} + \cdots + b_1\bar{x} + b_0 = 0$ since $a(\bar{x}) = 0$ in $F[x]/(a(x))$. With respect to this basis, the matrix for multiplication by x is therefore

$$\begin{pmatrix} 0 & 0 & \dots & \dots & \dots & -b_0 \\ 1 & 0 & \dots & \dots & \dots & -b_1 \\ 0 & 1 & \dots & \dots & \dots & -b_2 \\ 0 & 0 & \ddots & & & \vdots \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & \dots & 1 & -b_{k-1} \end{pmatrix}.$$

Such matrices are given a name:

Definition. Let $a(x) = x^k + b_{k-1}x^{k-1} + \cdots + b_1x + b_0$ be any monic polynomial in $F[x]$. The *companion matrix* of $a(x)$ is the $k \times k$ matrix with 1's down the first subdiagonal, $-b_0, -b_1, \dots, -b_{k-1}$ down the last column and zeros elsewhere. The companion matrix of $a(x)$ will be denoted by $\mathcal{C}_{a(x)}$.

We apply this to each of the cyclic modules on the right side of (1) above and let $\mathcal{B}_i$ be the elements of V corresponding to the basis chosen above for the cyclic factor $F[x]/(a_i(x))$ under the isomorphism in (1). Then by definition the linear transformation T acts on $\mathcal{B}_i$ by the companion matrix for $a_i(x)$ since we have seen that this is how multiplication by x acts. The union $\mathcal{B}$ of the $\mathcal{B}_i$'s gives a basis for V since the sum on the right of (1) is direct and with respect to this basis the linear transformation T has as matrix the *direct sum* of the companion matrices for the invariant factors, i.e.,

$$\begin{pmatrix} \mathcal{C}_{a_1(x)} & & & \\ & \mathcal{C}_{a_2(x)} & & \\ & & \ddots & \\ & & & \mathcal{C}_{a_m(x)} \end{pmatrix}. \tag{12.2}$$

Notice that this matrix is uniquely determined from the invariant factors of the $F[x]$-module V and, by Theorem 9, the list of invariant factors uniquely determines the module V up to isomorphism as an $F[x]$-module.

Definition.

(1) A matrix is said to be in *rational canonical form* if it is the direct sum of companion matrices for monic polynomials $a_1(x), \dots, a_m(x)$ of degree at least one with $a_1(x) \mid a_2(x) \mid \cdots \mid a_m(x)$. The polynomials $a_i(x)$ are called the *invariant factors* of the matrix. Such a matrix is also said to be a *block diagonal* matrix with blocks the companion matrices for the $a_i(x)$.

(2) A *rational canonical form* for a linear transformation T is a matrix representing T which is in rational canonical form.

We have seen that any linear transformation T has a rational canonical form. We now see that this rational canonical form is unique (hence is called *the* rational canonical form for T). To see this note that the process we used to determine the matrix of T