(d) In particular, prove that $\Phi_\ell(x)$ is irreducible modulo p if and only if $l-1$ is the smallest power of p which is congruent to 1 modulo ℓ, i.e., p is a primitive root modulo ℓ.

21. Prove that the first two elementary row and column operations described before Theorem 21 do not change the determinant of the matrix and the third elementary operation multiplies the determinant by a unit. Conclude from Theorem 21 that the characteristic polynomial of A differs by a unit from the product of the invariant factors of A. Since both these polynomials are monic by definition, conclude that they are equal (this gives an alternate proof of Proposition 20).

The following exercises outline the proof of Theorem 21. They carry out explicitly the construction described in Exercises 16 to 19 of the previous section for the Euclidean Domain $F[x]$. Let V be an n-dimensional vector space with basis $v_1, v_2, \ldots, v_n$ and let T be the linear transformation of V defined by the matrix A and this choice of basis, i.e., T is the linear transformation with

$$T(v_j) = \sum_{i=1}^{n} a_{ij} v_i, \qquad j = 1, 2, \ldots, n$$

where $A = (a_{ij})$. Let $F[x]^n$ be the free module of rank n over $F[x]$ and let $\xi_1, \xi_2, \ldots, \xi_n$ denote a basis. Then we have a natural surjective $F[x]$-module homomorphism

$$\varphi : F[x]^n \to V$$

defined by mapping ξ_i to v_i, $i = 1, 2, \ldots, n$. As indicated in the exercises of the previous section the invariant factors for the $F[x]$-module V can be determined once we have determined a set of generators and the corresponding relations matrix for $\ker \varphi$. Since by definition x acts on V by the linear transformation T, we have

$$x(v_j) = \sum_{i=1}^{n} a_{ij} v_i, \quad j = 1, 2, \ldots, n.$$

22. Show that the elements

$$v_j = -a_{1j}\xi_1 - \cdots - a_{j-1\,j}\xi_{j-1} + (x - a_{jj})\xi_j - a_{j+1\,j}\xi_{j+1} - \cdots - a_{nj}\xi_n$$

for $j = 1, 2, \ldots, n$ are elements of the kernel of φ.

23. **(a)** Show that $x\xi_j = v_j + f_j$ where $f_j \in F\xi_1 + \cdots + F\xi_n$ is an element in the F-vector space spanned by $\xi_1, \ldots, \xi_n$.

(b) Show that

$$F[x]\xi_1 + \cdots + F[x]\xi_n = (F[x]v_1 + \cdots + F[x]v_n) + (F\xi_1 + \cdots + F\xi_n).$$

24. Show that $v_1, v_2, \ldots, v_n$ generate the kernel of φ. [Use the previous result to show that any element of $\ker \varphi$ is the sum of an element in the module generated by $v_1, v_2, \ldots, v_n$ and an element of the form $b_1\xi_1 + \cdots + b_n\xi_n$ where the b_i are elements of F. Then show that such an element is in $\ker \varphi$ if and only if all the b_i are 0 since $v_1, \ldots, v_n$ are a basis for V over F.]

25. Show that the generators $v_1, v_2, \ldots, v_n$ of $\ker \varphi$ have corresponding relations matrix

$$\begin{pmatrix} x - a_{11} & -a_{21} & \ldots & -a_{n1} \\ -a_{12} & x - a_{22} & \ldots & -a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \ldots & x - a_{nn} \end{pmatrix} = xI - A^t,$$

where A^t is the transpose of A. Conclude that Theorem 21 and the algorithm for determining the invariant factors of A follows by Exercises 16 to 19 in the previous section (note that the row and column operations necessary to diagonalize this relations matrix are the column and row operations necessary to diagonalize the matrix in Theorem 21, which explains why the invariant factor algorithm keeps track of the *row* operations used).

12.3 THE JORDAN CANONICAL FORM

We continue with the notation in the previous section: F is a field, $F[x]$ is the ring of polynomials in x with coefficients in F, V is a finite dimensional vector space over F of dimension n, T is a fixed linear transformation of V by which we make V into an $F[x]$-module, and A is an $n \times n$ matrix with coefficients in F. Recall that once a basis for V has been fixed any linear transformation T defines a matrix A and conversely any matrix A defines a linear transformation T.

In the previous section we used the invariant factor form of the Fundamental Theorem for finitely generated modules over the Principal Ideal Domain $F[x]$ to obtain the rational canonical form for such a linear transformation T and the rational canonical form for such an $n \times n$ matrix A. In this section we use the elementary divisor form of the Fundamental Theorem to obtain the *Jordan canonical form*. We shall see that matrices in this canonical form are as close to being diagonal matrices as possible, so the matrices are simpler than in the rational canonical form (but we lose some of the "rationality" results).

The elementary divisors of a module are the prime power divisors of its invariant factors (this was Corollary 10). For the $F[x]$-module V the invariant factors were monic polynomials $a_1(x), a_2(x), \ldots, a_m(x)$ of degree at least one (with $a_1(x) \mid a_2(x) \mid \cdots \mid a_m(x)$), so the associated elementary divisors are the powers of the irreducible polynomial factors of these polynomials. These polynomials are only defined up to multiplication by a unit and, as in the case of the invariant factors, we can specify them uniquely by requiring that they be monic.

To obtain the simplest possible elementary divisors we shall assume that the polynomials $a_1(x), a_2(x), \ldots, a_m(x)$ factor completely into linear factors, i.e., that the elementary divisors of V are powers $(x - \lambda)^k$ of linear polynomials. Since the product of the elementary divisors is the characteristic polynomial, this is equivalent to the assumption that the field F contains all the eigenvalues of the linear transformation T (equivalently, of the matrix A representing the linear transformation T).

Under this assumption on F, it follows immediately from Theorem 6 that V is the direct sum of finitely many cyclic $F[x]$-modules of the form $F[x]/(x - \lambda)^k$ where $\lambda \in F$ is one of the eigenvalues of T, corresponding to the elementary divisors of V.

We now choose a vector space basis for each of the direct summands corresponding to the elementary divisors of V for which the corresponding matrix for T is particularly simple. Recall that by definition of the $F[x]$-module structure the linear transformation T acting on V is the element x acting by multiplication on each of the direct summands $F[x]/(x - \lambda)^k$.

Consider the elements

$$(\bar{x} - \lambda)^{k-1},\ (\bar{x} - \lambda)^{k-2}, \ldots,\ \bar{x} - \lambda,\ 1,$$

in the quotient $F[x]/(x-\lambda)^k$. Expanding each of these polynomials in $\bar{x}$ we see that the matrix relating these elements to the F-basis $\bar{x}^{k-1}, \bar{x}^{k-2}, \ldots, \bar{x}, 1$ of $F[x]/(x-\lambda)^k$ is upper triangular with 1's along the diagonal. Since this is an invertible matrix (having determinant 1), it follows that the elements above are an F-basis for $F[x]/(x-\lambda)^k$. With respect to this basis the linear transformation of multiplication by x acts in a particularly simple manner (note that $x = \lambda + (x-\lambda)$ and that $(\bar{x}-\lambda)^k = 0$ in the quotient):

$$
\begin{array}{lll}
 & (\bar{x}-\lambda)^{k-1} & \mapsto \quad \lambda\cdot(\bar{x}-\lambda)^{k-1} + (\bar{x}-\lambda)^k = \lambda\cdot(\bar{x}-\lambda)^{k-1} \\
 & (\bar{x}-\lambda)^{k-2} & \mapsto \quad \lambda\cdot(\bar{x}-\lambda)^{k-2} + (\bar{x}-\lambda)^{k-1} \\
x: & \vdots & \\
 & \bar{x}-\lambda & \mapsto \quad \lambda\cdot(\bar{x}-\lambda) + (\bar{x}-\lambda)^2 \\
 & 1 & \mapsto \quad \lambda\cdot 1 + (\bar{x}-\lambda).
\end{array}
$$

With respect to this basis, the matrix for multiplication by x is therefore

$$
\begin{pmatrix}
\lambda & 1 & & & \\
 & \lambda & \ddots & & \\
 & & \ddots & 1 & \\
 & & & \lambda & 1 \\
 & & & & \lambda
\end{pmatrix}
$$

where the blank entries are all zero. Such matrices are given a name:

Definition. The $k \times k$ matrix with λ along the main diagonal and 1 along the first superdiagonal depicted above is called the *$k \times k$ elementary Jordan matrix with eigenvalue λ* or the *Jordan block of size k with eigenvalue λ*.

Applying this to each of the cyclic factors of V in its elementary divisor decomposition we obtain a vector space basis for V with respect to which the linear transformation T has as matrix the direct sum of the Jordan blocks corresponding to the elementary divisors of V, i.e., is block diagonal with Jordan blocks along the diagonal:

$$
\begin{pmatrix}
J_1 & & & \\
 & J_2 & & \\
 & & \ddots & \\
 & & & J_t
\end{pmatrix}.
$$

Notice that this matrix is uniquely determined up to permutation of the blocks along the diagonal by the elementary divisors of the $F[x]$-module V and conversely, by Theorem 9, the list of elementary divisors uniquely determines the module V up to $F[x]$-module isomorphism.

Definition.

(1) A matrix is said to be in *Jordan canonical form* if it is a block diagonal matrix with Jordan blocks along the diagonal.

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

We have proved that any linear transformation T has a Jordan canonical form. As in the case of the rational canonical form, it follows from the uniqueness of the elementary divisors that the Jordan canonical form is unique up to a permutation of the Jordan blocks along the diagonal (hence is called *the* Jordan canonical form for T). We summarize this in the following theorem.

Theorem 22. *(Jordan Canonical Form for Linear Transformations)* Let V be a finite dimensional vector space over the field F and let T be a linear transformation of V. Assume F contains all the eigenvalues of T.

(1) There is a basis for V with respect to which the matrix for T is in Jordan canonical form, i.e., is a block diagonal matrix whose diagonal blocks are the Jordan blocks for the elementary divisors of V.

(2) The Jordan canonical form for T is unique up to a permutation of the Jordan blocks along the diagonal.

As for the rational canonical form, the following theorem gives the corresponding statement for $n \times n$ matrices over F.

Theorem 23. *(Jordan Canonical Form for Matrices)* Let A be an $n \times n$ matrix over the field F and assume F contains all the eigenvalues of A.

(1) The matrix A is similar to a matrix in Jordan canonical form, i.e., there is an invertible $n \times n$ matrix P over F such that $P^{-1}AP$ is a block diagonal matrix whose diagonal blocks are the Jordan blocks for the elementary divisors of A.

(2) The Jordan canonical form for A is unique up to a permutation of the Jordan blocks along the diagonal.

The Jordan canonical form differs from a diagonal matrix only by the possible presence of some 1's along the first superdiagonal (and then only if there are Jordan blocks of size greater than one), hence is close to being a diagonal matrix. The following result shows in particular that the Jordan canonical form for a matrix A is as close to being a diagonal matrix as possible.

Corollary 24.

(1) If a matrix A is similar to a diagonal matrix D, then D is the Jordan canonical form of A.

(2) Two diagonal matrices are similar if and only if their diagonal entries are the same up to a permutation.

Proof: The first assertion is immediate from the uniqueness of Jordan canonical forms because a diagonal matrix is itself in Jordan form (with Jordan blocks of size 1). The uniqueness of the Jordan canonical form gives (2).

The next corollary gives a criterion to determine when a matrix A can be diagonalized.