and the same minimal polynomial. Suppose also that for each i the solution spaces of $(A - c_iI)$ and $(B - c_iI)$ have the same dimension. If no d_i is greater than 6, then A and B are similar.

13. If N is a $k \times k$ elementary nilpotent matrix, i.e., $N^k = 0$ but $N^{k-1} \neq 0$, show that N^t is similar to N. Now use the Jordan form to prove that every complex $n \times n$ matrix is similar to its transpose.

14. What's wrong with the following proof? If A is a complex $n \times n$ matrix such that $A^t = -A$, then A is 0. (*Proof:* Let J be the Jordan form of A. Since $A^t = -A$, $J^t = -J$. But J is triangular so that $J^t = -J$ implies that every entry of J is zero. Since $J = 0$ and A is similar to J, we see that $A = 0$.) (Give an example of a non-zero A such that $A^t = -A$.)

15. If N is a nilpotent 3×3 matrix over C, prove that $A = I + \frac{1}{2}N - \frac{1}{8}N^2$ satisfies $A^2 = I + N$, i.e., A is a square root of $I + N$. Use the binomial series for $(1 + t)^{1/2}$ to obtain a similar formula for a square root of $I + N$, where N is any nilpotent $n \times n$ matrix over C.

16. Use the result of Exercise 15 to prove that if c is a non-zero complex number and N is a nilpotent complex matrix, then $(cI + N)$ has a square root. Now use the Jordan form to prove that every non-singular complex $n \times n$ matrix has a square root.

7.4. Computation of Invariant Factors

Suppose that A is an $n \times n$ matrix with entries in the field F. We wish to find a method for computing the invariant factors $p_1, \ldots, p_r$ which define the rational form for A. Let us begin with the very simple case in which A is the companion matrix (7.2) of a monic polynomial

$$p = x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0.$$

In Section 7.1 we saw that p is both the minimal and the characteristic polynomial for the companion matrix A. Now, we want to give a direct calculation which shows that p is the characteristic polynomial for A. In this case,

$$xI - A = \begin{bmatrix} x & 0 & 0 & \cdots & 0 & c_0 \\ -1 & x & 0 & \cdots & 0 & c_1 \\ 0 & -1 & x & \cdots & 0 & c_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & c_{n-2} \\ 0 & 0 & 0 & \cdots & -1 & x + c_{n-1} \end{bmatrix}.$$

Add x times row n to row $(n - 1)$. This will remove the x in the $(n - 1, n - 1)$ place and it will not change the determinant. Then, add x times the new row $(n - 1)$ to row $(n - 2)$. Continue successively until all of the x's on the main diagonal have been removed by that process. The result is the matrix

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & x^n + \cdots + c_1x + c_0 \\ -1 & 0 & 0 & \cdots & 0 & x^{n-1} + \cdots + c_2x + c_1 \\ 0 & -1 & 0 & \cdots & 0 & x^{n-2} + \cdots + c_3x + c_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & x^2 + c_{n-1}x + c_{n-2} \\ 0 & 0 & 0 & \cdots & -1 & x + c_{n-1} \end{bmatrix}$$

which has the same determinant as $xI - A$. The upper right-hand entry of this matrix is the polynomial p. We clean up the last column by adding to it appropriate multiples of the other columns:

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & p \\ -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 0 \end{bmatrix}$$

Multiply each of the first $(n - 1)$ columns by -1 and then perform $(n - 1)$ interchanges of adjacent columns to bring the present column n to the first position. The total effect of the $2n - 2$ sign changes is to leave the determinant unaltered. We obtain the matrix

$$\begin{bmatrix} p & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{7-28}$$

It is then clear that $p = \det (xI - A)$.

We are going to show that, for any $n \times n$ matrix A, there is a succession of row and column operations which will transform $xI - A$ into a matrix much like (7-28), in which the invariant factors of A appear down the main diagonal. Let us be completely clear about the operations we shall use.

We shall be concerned with $F[x]^{m\times n}$, the collection of $m \times n$ matrices with entries which are polynomials over the field F. If M is such a matrix, an **elementary row operation** on M is one of the following

1. multiplication of one row of M by a non-zero scalar in F;
2. replacement of the rth row of M by row r plus f times row s, where f is any polynomial over F and $r \neq s$;
3. interchange of two rows of M.

The inverse operation of an elementary row operation is an elementary row operation of the same type. Notice that we could not make such an assertion if we allowed non-scalar polynomials in (1). An $m \times m$ **ele-**

mentary matrix, that is, an elementary matrix in $F[x]^{m \times m}$, is one which can be obtained from the $m \times m$ identity matrix by means of a single elementary row operation. Clearly each elementary row operation on M can be effected by multiplying M on the left by a suitable $m \times m$ elementary matrix; in fact, if e is the operation, then

$$e(M) = e(I)M.$$

Let M, N be matrices in $F[x]^{m \times n}$. We say that N is **row-equivalent** to M if N can be obtained from M by a finite succession of elementary row operations:

$$M = M_0 \rightarrow M_1 \rightarrow \cdots \rightarrow M_k = N.$$

Evidently N is row-equivalent to M if and only if M is row-equivalent to N, so that we may use the terminology 'M and N are row-equivalent.' If N is row-equivalent to M, then

$$N = PM$$

where the $m \times m$ matrix P is a product of elementary matrices:

$$P = E_1 \cdots E_k.$$

In particular, P is an invertible matrix with inverse

$$P^{-1} = E_k^{-1} \cdots E_1^{-1}.$$

Of course, the inverse of E_j comes from the inverse elementary row operation.

All of this is just as it is in the case of matrices with entries in F. It parallels the elementary results in Chapter 1. Thus, the next problem which suggests itself is to introduce a row-reduced echelon form for polynomial matrices. Here, we meet a new obstacle. How do we row-reduce a matrix? The first step is to single out the leading non-zero entry of row 1 and to divide every entry of row 1 by that entry. We cannot (necessarily) do that when the matrix has polynomial entries. As we shall see in the next theorem, we can circumvent this difficulty in certain cases; however, there is not any entirely suitable row-reduced form for the general matrix in $F[x]^{m \times n}$. If we introduce column operations as well and study the type of equivalence which results from allowing the use of both types of operations, we can obtain a very useful standard form for each matrix. The basic tool is the following.

Lemma. *Let* M *be a matrix in* $F[x]^{m \times n}$ *which has some non-zero entry in its first column, and let* p *be the greatest common divisor of the entries in column* 1 *of* M. *Then* M *is row-equivalent to a matrix* N *which has*

$$\begin{bmatrix} p \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

as its first column.

Proof. We shall prove something more than we have stated. We shall show that there is an algorithm for finding N, i.e., a prescription which a machine could use to calculate N in a finite number of steps. First, we need some notation.

Let M be any $m \times n$ matrix with entries in $F[x]$ which has a non-zero first column

$$M_1 = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix}.$$

Define

$$\begin{aligned} l(M_1) &= \min_{f_i \neq 0} \deg f_i \\ p(M_1) &= \text{g.c.d. } (f_1, \ldots, f_m). \end{aligned} \tag{7-29}$$

Let j be some index such that $\deg f_j = l(M_1)$. To be specific, let j be the smallest index i for which $\deg f_i = l(M_1)$. Attempt to divide each f_i by f_j:

$$f_i = f_j g_i + r_i, \qquad r_i = 0 \quad \text{or} \quad \deg r_i < \deg f_j. \tag{7-30}$$

For each i different from j, replace row i of M by row i minus g_i times row j. Multiply row j by the reciprocal of the leading coefficient of f_j and then interchange rows j and 1. The result of all these operations is a matrix M' which has for its first column

$$M_1' = \begin{bmatrix} \tilde{f}_j \\ r_2 \\ \vdots \\ r_{j-1} \\ r_1 \\ r_{j+1} \\ \vdots \\ r_m \end{bmatrix}. \tag{7-31}$$

where $\tilde{f}_j$ is the monic polynomial obtained by normalizing f_j to have leading coefficient 1. We have given a well-defined procedure for associating with each M a matrix M' with these properties.

(a) M' is row-equivalent to M.
(b) $p(M_1') = p(M_1)$.
(c) Either $l(M_1') < l(M_1)$ or

$$M_1' = \begin{bmatrix} p(M_1) \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

It is easy to verify (b) and (c) from (7-30) and (7-31). Property (c) is just another way of stating that either there is some i such that $r_i \neq 0$ and $\deg r_i < \deg f_j$ or else $r_i = 0$ for all i and $\tilde{f}_j$ is (therefore) the greatest common divisor of $f_1, \ldots, f_m$.

The proof of the lemma is now quite simple. We start with the matrix M and apply the above procedure to obtain M'. Property (c) tells us that either M' will serve as the matrix N in the lemma or $l(M_1') < l(M_1)$. In the latter case, we apply the procedure to M' to obtain the matrix $M^{(2)} = (M')'$. If $M^{(2)}$ is not a suitable N, we form $M^{(3)} = (M^{(2)})'$, and so on. The point is that the strict inequalities

$$l(M_1) > l(M_1') > l(M_1^{(2)}) > \cdots$$

cannot continue for very long. After not more than $l(M_1)$ iterations of our procedure, we must arrive at a matrix $M^{(k)}$ which has the properties we seek. ∎

Theorem 6. *Let* P *be an* m × m *matrix with entries in the polynomial algebra* F[x]. *The following are equivalent.*

(i) P *is invertible.*
(ii) *The determinant of* P *is a non-zero scalar polynomial.*
(iii) P *is row-equivalent to the* m × m *identity matrix.*
(iv) P *is a product of elementary matrices.*

Proof. Certainly (i) implies (ii) because the determinant function is multiplicative and the only polynomials invertible in $F[x]$ are the non-zero scalar ones. As a matter of fact, in Chapter 5 we used the classical adjoint to show that (i) and (ii) are equivalent. Our argument here provides a different proof that (i) follows from (ii). We shall complete the merry-go-round

$$\begin{array}{ccc} \text{(i)} & \rightarrow & \text{(ii)} \\ \uparrow & & \downarrow \\ \text{(iv)} & \leftarrow & \text{(iii).} \end{array}$$

The only implication which is not obvious is that (iii) follows from (ii).

Assume (ii) and consider the first column of P. It contains certain polynomials $p_1, \ldots, p_m$, and

$$\text{g.c.d.}\ (p_1, \ldots, p_m) = 1$$

because any common divisor of $p_1, \ldots, p_m$ must divide (the scalar) $\det P$. Apply the previous lemma to P to obtain a matrix

$$Q = \begin{bmatrix} 1 & a_2 & \cdots & a_m \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{bmatrix} \tag{7-32}$$

which is row-equivalent to P. An elementary row operation changes the determinant of a matrix by (at most) a non-zero scalar factor. Thus $\det Q$

is a non-zero scalar polynomial. Evidently the $(m - 1) \times (m - 1)$ matrix B in (7-32) has the same determinant as does Q. Therefore, we may apply the last lemma to B. If we continue this way for m steps, we obtain an upper-triangular matrix

$$R = \begin{bmatrix} 1 & a_2 & \cdots & a_m \\ 0 & 1 & \cdots & b_m \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

which is row-equivalent to R. Obviously R is row-equivalent to the $m \times m$ identity matrix. ∎

Corollary. *Let* M *and* N *be* m $\times$ n *matrices with entries in the polynomial algebra* F[x]. *Then* N *is row-equivalent to* M *if and only if*

$$\mathrm{N} = \mathrm{PM}$$

where P *is an invertible* m $\times$ m *matrix with entries in* F[x].

We now define **elementary column operations** and **column-equivalence** in a manner analogous to row operations and row-equivalence. We do not need a new concept of elementary matrix because the class of matrices which can be obtained by performing one elementary column operation on the identity matrix is the same as the class obtained by using a single elementary row operation.

Definition. *The matrix* N *is* **equivalent** *to the matrix* M *if we can pass from* M *to* N *by means of a sequence of operations*

$$\mathrm{M} = \mathrm{M}_0 \rightarrow \mathrm{M}_1 \rightarrow \cdots \rightarrow \mathrm{M}_k = \mathrm{N}$$

each of which is an elementary row operation or an elementary column operation.

Theorem 7. *Let* M *and* N *be* m $\times$ n *matrices with entries in the polynomial algebra* F[x]. *Then* N *is equivalent to* M *if and only if*

$$\mathrm{N} = \mathrm{PMQ}$$

where P *is an invertible matrix in* $\mathrm{F}[\mathrm{x}]^{m \times m}$ *and* Q *is an invertible matrix in* $\mathrm{F}[\mathrm{x}]^{n \times n}$.

Theorem 8. *Let* A *be an* n $\times$ n *matrix with entries in the field* F, *and let* $\mathrm{p}_1, \ldots, \mathrm{p}_r$ *be the invariant factors for* A. *The matrix* $\mathrm{xI} - \mathrm{A}$ *is equivalent to the* n $\times$ n *diagonal matrix with diagonal entries* $\mathrm{p}_1, \ldots, \mathrm{p}_r, 1, 1, \ldots, 1$.

Proof. There exists an invertible $n \times n$ matrix P, with entries in F, such that PAP^{-1} is in rational form, that is, has the block form