CHAPTER III

RESIDUE MATRICES WITH POLYNOMIAL ELEMENTS

1. Introduction. We shall be concerned in the next chapter with the concept of "similarity modulo m" of residue matrices. Though these matrices are integral, various concepts related to the theory of similarity modulo m involve residue matrices whose elements are polynomials. It is for this reason that we shall now examine the theory of residue matrices with polynomial elements.

Since, in this paper, we shall be concerned primarily with the case in which m is a prime number, we shall assume from this point on that m is a prime and that F is the field of integers modulo m. We shall denote by $F[\lambda]$ the domain of integral polynomials in λ with coefficients in F. Our attention, then, will be focused on nxn matrices over $F[\lambda]$. We shall refer to these matrices as polynomial matrices modulo m or λ-matrices modulo m. A λ-matrix modulo m, $A(\lambda)$, can be represented as follows

$$A(\lambda) = [a_{ij}(\lambda)] = [a_{ij}^{(k)} + a_{ij}^{(k-1)}\lambda + \ldots + a_{ij}^{(1)}\lambda^{k-1} + a_{ij}^{(0)}\lambda^{k}],$$

where k is the largest of the degrees of the polynomials $a_{ij}(\lambda)$, or in the form of a matrix polynomial modulo m in λ with residue matrix coefficients,

$$A(\lambda) = \left|\,|A_k|_m + |A_{k-1}|_m\lambda + \ldots + |A_1|_m\lambda^{k-1} + |A_0|_m\lambda^{k}\right|_m ,$$

where

$$|A|_m = \left|[a_{ij}^{(\ell)}]\right|_m .$$

Most of the following definitions, theorems and proofs in this chapter and the next are analogous to those found in any book on matrix theory. See, for example, Gantmacher [1960] and Perlis [1952].

We assume from this point on that m is greater than the order of

all matrices and the degree of all polynomials considered.

2. Equivalence Modulo m over $F[\lambda]$. We introduce the following elementary operations on a λ-matrix modulo m:

(a) Interchange of any two rows (or two columns),

(b) Multiplication of any row (or column) by a nonzero constant, followed by reduction modulo m,

(c) Addition to the ith row of any other row, for example the jth, multiplied by any polynomial modulo m, and followed by reduction modulo m (and the analogous operation on columns).

An elementary λ-matrix modulo m, $E(\lambda)$, is a matrix obtained from I_n by performing one of the above elementary operations on I_n. An elementary operation can be performed on $A(\lambda)$ by multiplying by $E(\lambda)$ on the appropriate side, followed by reduction modulo m.

Clearly, every elementary λ-matrix modulo m has a nonzero determinant modulo m which does not depend on λ. Therefore, every elementary operation has an inverse. It can easily be shown that the matrix corresponding to this inverse is an elementary λ-matrix modulo m.

We have a theorem which is analogous to theorem (3.17).

(2.1) THEOREM. Multiplication of a matrix $A(\lambda)$ by an elementary λ-matrix modulo m does not change $r_m(A(\lambda))$.

Proof. The proof is analogous to that of theorem (3.17).

(2.2) DEFINITION. If $A(\lambda)$ and $B(\lambda)$ are two λ-matrices modulo m, then $B(\lambda)$ is said to be equivalent modulo m over $f[\lambda]$ to $A(\lambda)$ if and only if $B(\lambda)$ can be obtained from $A(\lambda)$ by a series of elementary operations, in other words,

$$|B(\lambda)|_m = \left||P(\lambda)|_m A(\lambda)|Q(\lambda)|_m\right|_m ,$$

where $P(\lambda)$ and $Q(\lambda)$ are products of elementary λ-matrices modulo m.

It is a simple matter to show that equivalence modulo m over $F[\lambda]$ is reflexive, symmetric, and transitive, and that rank modulo m is not affected by equivalence transformations modulo m over $F[\lambda]$. Thus, it is sufficient to say merely that $A(\lambda)$ and $B(\lambda)$ are equivalent modulo m over $F[\lambda]$.

EXAMPLE. Let m = 13,

$$A(\lambda) = \begin{bmatrix} 3\lambda & 1+\lambda & 2 \\ 2 & 4+\lambda^2 & -5 \\ 6 & -7\lambda & 15\lambda \end{bmatrix} ,$$

and

$$B(\lambda) = \begin{bmatrix} 3\lambda & 6+6\lambda & 2 \\ 2+3\lambda^2 & -2+6\lambda-\lambda^2 & -5+2\lambda \\ -2 & \lambda & -5\lambda \end{bmatrix} .$$

Then, $A(\lambda)$ and $B(\lambda)$ are equivalent modulo m over $F[\lambda]$, since

$$|B(\lambda)|_{13} = \left| \begin{bmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 3\lambda & 1+\lambda & 2 \\ 2 & 4+\lambda^2 & -5 \\ 6 & -7\lambda & 15\lambda \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right|_{13} .$$

Thus, in definition (2.2), we have

$$|P(\lambda)|_{13} = \begin{bmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

and

$$|Q(\lambda)|_{13} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A natural question is whether certain λ-matrices modulo m which are equivalent modulo m over $f[\lambda]$ can be strictly equivalent modulo m. That is, if

$$|B(\lambda)|_m = |P(\lambda)A(\lambda)Q(\lambda)|_m,$$

can we find matrices S and T over F such that

$$|B(\lambda)|_m = |S\ A(\lambda)\ T|_m\ ?$$

This is answered by the following lemma and theorem.

(2.3) LEMMA. (division algorithm for matrix polynomials modulo m). Let

(2.4)
$$A(\lambda) = \left||A_k|_m + |A_{k-1}|_m\lambda + \ldots + |A_1|_m\lambda^{k-1} + |A_0|_m\lambda^k\right|_m$$

$$B(\lambda) = \left||B_s|_m + |B_{s-1}|_m\lambda + \ldots + |B_1|_m\lambda^{s-1} + |B_0|_m\lambda^s\right|_m ,$$

where

(2.5)
$$\left|\det |B_0|_m\right|_m \neq 0.$$

Then there exist unique λ-matrices modulo m $Q_r(\lambda)$, $R_r(\lambda)$, $Q_\ell(\lambda)$, and $R_\ell(\lambda)$ such that

(2.6)
$$A(\lambda) = |Q_r(\lambda)B(\lambda) + R_r(\lambda)|_m$$

and

(2.7)
$$B(\lambda) = |B(\lambda)Q_\ell(\lambda) + R_\ell(\lambda)|_m ,$$

where either $R_r(\lambda) = \emptyset$ or the degree of $R_r(\lambda)$ is less than s, and either $R_\ell(\lambda) = \emptyset$ or the degree of $R_\ell(\lambda)$ is less than s.

Proof. [Perlis, 1952, pp. 134-135] [Gantmacher, 1960, pp. 78-79] If $A(\lambda)$ or $k < s$, then we can set $Q_r(\lambda) = Q_\ell(\lambda) = \emptyset$ and $R_r(\lambda) = R_\ell(\lambda) = A(\lambda)$. If $k \geq s$ and $A(\lambda) \neq \emptyset$, then we apply the usual scheme for division of polynomials modulo m. Let

$$C^{(1)}(\lambda) = \left|A(\lambda) - |A_0B_0^{-1}(m)\lambda^{k-s}|_m\ B(\lambda)\right|_m$$

$$= \left| |C^{(1)}_{k^{(1)}}|_m + |C^{(1)}_{k^{(1)}-1}|_m \lambda + \ldots + |C^{(1)}_1|_m \lambda^{k^{(1)}-1} + |C^{(1)}_0|_m \lambda^{k^{(1)}} \right|_m ,$$

where either $C^{(1)}(\lambda) = \emptyset$ or the degree of $C^{(1)}(\lambda)$, $k^{(1)}$, is less than k. If $C^{(1)}(\lambda) = \emptyset$ or if $k^{(1)} < s$, then (2.6) is satisfied by

$$R_r(\lambda) = C^{(1)}(\lambda)$$

and

$$Q_r(\lambda) = \left| |A_0|_m B_0^{-1}(m) \lambda^{k-s} \right|_m .$$

If $k^{(1)} > s$, then we repeat the process and compute

$$C^{(2)}(\lambda) = \left| C^{(1)}(\lambda) - |C^{(1)}_0 B_0^{-1}(m) \lambda^{k^{(1)}-s}|_m \; B(\lambda) \right|_m$$

$$= \left| |C^{(2)}_{k^{(2)}}|_m + |C^{(2)}_{k^{(2)}-1}|_m \lambda + \ldots + |C^{(2)}_0|_m \lambda^{k^{(2)}} \right|_m ,$$

where $k^{(2)} < k^{(1)}$.

Since the degrees of $A(\lambda)$, $C^{(1)}(\lambda)$, $C^{(2)}(\lambda)$, ... are decreasing, we must at some stage reach the situation in which $k^{(i)} < s$. Then (2.6) is satisfied by

$$R_r(\lambda) = C^{(i)}(\lambda)$$

and

$$Q_r(\lambda) = \left| A_0 B_0^{-1}(m) \lambda^{k-s} + C^{(1)}_0 B_0^{-1}(m) \lambda^{k^{(1)}-s} + \ldots + C^{(i-1)}_0 B_0^{-1}(m) \lambda^{k^{(i-1)}-s} \right|_m .$$

Thus, <u>right division</u> of matrix polynimials modulo m is possible if the conditions (2.4) and (2.5) exist. Similarly, we can show that <u>left division</u> is possible by computing the sequence $A(\lambda)$, $D^{(1)}(\lambda)$, $D^{(2)}(\lambda)$, ... , $D^{(\ell)}(\lambda)$, where

$$D^{(j)}(\lambda) = \left| D^{(j-1)}(\lambda) - B(\lambda) |B_0^{-1}(m) D_0^{(j-1)} \lambda^{m^{(j-1)}-s}|_m \right|_m$$

$$= \left| |D^{(j)}_{m^{(j)}}|_m + |D^{(j)}_{m^{(j)}-1}|_m \lambda + \ldots + |D^{(j)}_1|_m \lambda^{m^{(j)}-1} + |D^{(j)}_0|_m \lambda^{m^{(j)}} \right|_m$$

for $j=2,\ldots,\ell$ and $m^{(j)} < m^{(j-1)}$ is the degree of $D^{(j)}(\lambda)$, and where $m^{(\ell)} < s$ and

$$D^{(1)}(\lambda) = \left|A(\lambda) - B(\lambda)\left|B_0^{-1}(m)A_0\lambda^{k-s}\right|_m\right|_m$$

$$= \left|\left|D^{(1)}_{m^{(1)}}\right|_m + \left|D^{(1)}_{m^{(1)}-1}\right|_m\lambda + \ldots + \left|D_0^{(1)}\right|_m\lambda^{m^{(1)}}\right|_m .$$

In this case, (2.7) is satisfied by letting

$$R_\ell(\lambda) = D^{(\ell)}(\lambda)$$

and

$$Q_\ell(\lambda) = \left|B_0^{-1}(m)A_0\lambda^{k-s} + B_0^{-1}(m)D_0^{(1)}\lambda^{m^{(1)}-s} + \ldots + B_0^{-1}(m)D_0^{(\ell-1)}\lambda^{m^{(\ell-1)}-s}\right|_m .$$

We shall now prove uniqueness for $Q_r(\lambda)$ and $R_r(\lambda)$. The uniqueness of $Q_\ell(\lambda)$ and $R_\ell(\lambda)$ can be shown in a similar manner. Suppose we have

(2.8) $$A(\lambda) = \left|Q_r(\lambda)B(\lambda) + R_r(\lambda)\right|_m$$

and

(2.9) $$A(\lambda) = \left|Q_r^*(\lambda)B(\lambda) + R_r^*(\lambda)\right|_m ,$$

where the degrees of $R_r(\lambda)$ and $R_r^*(\lambda)$ are both less than s. Subtracting (2.8) from (2.9) modulo m, we obtain

$$\emptyset = \left|[Q_r^*(\lambda) - Q_r(\lambda)]B(\lambda) + [R_r^*(\lambda) - R_r(\lambda)]\right|_m$$

or

(2.10) $$\left|[Q_r(\lambda) - Q_r^*(\lambda)]B(\lambda)\right|_m = \left|R_r^*(\lambda) - R_r(\lambda)\right|_m .$$

The degree of the matrix polynomial modulo m on the right-hand side of (2.10) is less than s. If we had $\left|Q_r(\lambda) - Q_r^*(\lambda)\right|_m \neq \emptyset$, then the degree of the matrix polynomial on the left-hand side of (2.10) would be greater than s. Hence, we must have

$$|Q_r(\lambda) - Q_r^*(\lambda)|_m = \emptyset \ ,$$

and thus

$$|R_r^*(\lambda) - R_r(\lambda)|_m = \emptyset \ .$$

Therefore

$$|Q_r^*(\lambda)|_m = |Q_r(\lambda)|_m$$

and

$$|R_r^*(\lambda)|_m = |R_r(\lambda)|_m \ .$$

Since the matrix polynomials $Q_r^*(\lambda)$, $Q_r(\lambda)$, $R_r^*(\lambda)$, and $R_r(\lambda)$ are all polynomials modulo m, we must have

$$Q_r^*(\lambda) = Q_r(\lambda)$$

and

$$R_r^*(\lambda) = R_r(\lambda) \ . \qquad ///$$

(2.11) THEOREM. Let $A(\lambda)$ and $B(\lambda)$ be two λ-matrices modulo m of degree one:

$$A(\lambda) = |A_0\lambda + A_1|_m,$$

$$B(\lambda) = |B_0\lambda + B_1|_m,$$

and

$$|\det |B_0|_m|_m \neq 0.$$

Then $A(\lambda)$ and $B(\lambda)$ are equivalent modulo m over $F[\lambda]$ if and only if there exist residue matrices S and T over F which are nonsingular modulo m, such that

(2.12) $$B(\lambda) = |S\ A(\lambda)\ T|_m \ .$$

In other words, $A(\lambda)$ and $B(\lambda)$ are strictly equivalent modulo m.

Proof. [Gantmacher, 1960, pp. 146-147] If (2.12) holds for matrices S and T over F, then since matrices over F are also matrices over $F[\lambda]$, then $A(\lambda)$ and $B(\lambda)$ are equivalent modulo m over $F[\lambda]$.

Suppose $A(\lambda)$ and $B(\lambda)$ are equivalent modulo m over $F[\lambda]$, so that

$$B(\lambda) = |P(\lambda)A(\lambda)Q(\lambda)|_m , \tag{2.13}$$

where $P(\lambda)$ and $Q(\lambda)$ are products of elementary λ-matrices modulo m. Thus, $[P(\lambda)]^{-1}(m)$ exists, and from (2.13) we can write

$$|[P(\lambda)]^{-1}(m)\ B(\lambda)|_m = |A(\lambda)Q(\lambda)|_m . \tag{2.14}$$

Since $[P(\lambda)]^{-1}(m)$ is a λ-matrix modulo m, we can use lemma (2.2) to write

$$[P(\lambda)]^{-1}(m) = |A(\lambda)Q_\ell(\lambda) + R_\ell(\lambda)|_m \tag{2.15}$$

and

$$Q(\lambda) = |Q_r(\lambda)B(\lambda) + R_r(\lambda)|_m .$$

Since $A(\lambda)$ and $B(\lambda)$ are of degree one, then $R_\ell(\lambda)$ and $R_r(\lambda)$ must both be of degree zero. In other words, $R_\ell(\lambda)$ and $R_r(\lambda)$ are both matrices over F, and we can write

$$[P(\lambda)]^{-1}(m) = |A(\lambda)Q_\ell(\lambda) + R_\ell|_m \tag{2.16}$$

and

$$Q(\lambda) = |Q_r(\lambda)B(\lambda) + R_r|_m .$$

Substituting these expressions for $[P(\lambda)]^{-1}(m)$ and $Q(\lambda)$ into (2.14), we obtain

$$\left||A(\lambda)Q_\ell(\lambda) + R_\ell|_m\ B(\lambda)\right|_m = \left|A(\lambda)\ |Q_r(\lambda)B(\lambda) + R_r|_m\ \right|_m$$

which can be rewritten as

$$|A(\lambda)Q_\ell(\lambda)B(\lambda) + R_\ell B(\lambda)|_m = |A(\lambda)Q_r(\lambda)B(\lambda) + A(\lambda)R_r|_m$$

or

$$|A(\lambda)[Q_\ell(\lambda) - Q_r(\lambda)]\ B(\lambda)|_m = |A(\lambda)R_r - R_\ell B(\lambda)|_m .$$

Since the right-hand side is of degree $\leqq 1$, then the left-hand side must also

be of degree ≤ 1. This implies that the expression in brackets must be congruent to zero, for otherwise the left-hand side would be of degree ≥ 2. Hence, we have

$$|A(\lambda)\ R_r|_m = |R_\ell B(\lambda)|_m \ . \tag{2.17}$$

It remains to show that R_ℓ and R_r are nonsingular modulo m. In order to show that R_ℓ is nonsingular modulo m, we first write

$$P(\lambda) = |B(\lambda)C(\lambda) + D(\lambda)|_m$$

using lemma (2.2) Since $P(\lambda)$ is of degree one, $D(\lambda)$ must be of degree zero, and hence

$$P(\lambda) = |B(\lambda)C(\lambda) + D|_m \ . \tag{2.18}$$

From (2.14), (2.15), and (2.18), we have

$$\begin{aligned} I_n &= |[P(\lambda)]^{-1}(m)\ P(\lambda)|_m \\ &= |[P(\lambda)]^{-1}(m)\ [B(\lambda)C(\lambda) + D]|_m \\ &= \left|\,|[P(\lambda)]^{-1}(m)\ B(\lambda)|_m\ C(\lambda) + [P(\lambda)]^{-1}(m)\ D\right|_m \ . \end{aligned} \tag{2.19}$$

Substituting (2.14) and (2.16) into (2.19) we have

$$\begin{aligned} I_n &= \left|\,|A(\lambda)Q(\lambda)|_m C(\lambda) + |A(\lambda)Q_\ell(\lambda) + R_\ell|_m\ D\right|_m \\ &= |A(\lambda)Q(\lambda)C(\lambda) + A(\lambda)Q_\ell(\lambda)D + R_\ell D|_m \\ &= |A(\lambda)[Q(\lambda)C(\lambda) + Q_\ell(\lambda)\ D] + R_\ell D|_m . \end{aligned}$$

Because the above equation must be of degree zero (since it is equal to I_n), the expression in brackets must be congruent to zero, giving

$$I_n = |R_\ell\ D|_m$$

or

(2.20) $$R_{\ell}^{-1}(m) = D.$$

Hence R_{ℓ} is nonsingular modulo m. From (2.17) and(2.20) we have

$$|R_{\ell}^{-1}(m)A(\lambda)R_r|_m = |B(\lambda)|_m$$

$$= B(\lambda).$$

Thus, letting

$$S = R_{\ell}^{-1}(m)$$

and

$$T = R_r,$$

we obtain

$$B(\lambda) = |S\ A(\lambda)\ T|_m.$$

T is nonsingular modulo m, since

$$\left| |\det S|_m\ |\det A_o|_m\ |\det T|_m \right|_m = |\det B_o|_m$$

$$\neq 0.$$

Therefore, we have proved theorem (2.11). ///

3. The Canonical Form of a λ-Matrix Modulo m. We shall now examine a simple form for a λ-matrix modulo m which can be obtained by means of the elementary operations described in section 2.

(3.1) LEMMA. A nonzero λ-matrix modulo m $A(\lambda)$ is equivalent modulo m over $F[\lambda]$ to a matrix of the form

$$B_1(\lambda) = \left[\begin{array}{c|c} f_1(\lambda) & O \\ \hline O & A_2(\lambda) \end{array}\right],$$

where $f_1(\lambda)$ is a monic polynomial modulo m which divided

modulo m each of the nonzero elements of $A_2(\lambda)$.

Proof. [Perlis, 1952, p. 125], [Gantmacher, 1960, pp. 134-138] Of all the nonzero elements of $A(\lambda)$ we choose the one of least degree, and by performing a suitable combination of elementary operations of type (a), we can bring it into the $a_{11}(\lambda)$ position. By the division algorithm for polynomials modulo m (II-4.5), we can find polynomials modulo m, $q_{i1}(\lambda)$, $q_{1k}(\lambda)$, $r_{i1}(\lambda)$, and $r_{1k}(\lambda)$, such that

$$|a_{i1}(\lambda)|_m = |a_{11}(\lambda)q_{i1}(\lambda) + r_{i1}(\lambda)|_m$$

and

$$|a_{1k}(\lambda)|_m = |a_{11}(\lambda)q_{1k}(\lambda) + r_{1k}(\lambda)|_m , \quad (i,k=2,\ldots,n).$$

If any of the remainders $r_{i1}(\lambda)$ or $r_{1k}(\lambda)$ is not congruent to zero, for example $r_{i1}(\lambda)$, then we perform an elementary operation of type (c) and subtract $q_{i1}(\lambda)$ times the first row from the ith. This replaces the element $a_{i1}(\lambda)$ by $r_{i1}(\lambda)$, which is of smaller degree than $a_{i1}(\lambda)$. By carrying out an operation of type (a) again, we can reduce the degree of the element in the top left corner by bringing the new element $r_{i1}(\lambda)$ into the $a_{11}(\lambda)$ position. This process of replacing the elements $a_{i1}(\lambda)$ and $a_{1k}(\lambda)$ by $r_{i1}(\lambda)$ and $r_{1k}(\lambda)$, and of reducing the element in the top left corner must at some time reach the situation in which all of the $r_{i1}(\lambda)$ and $r_{1k}(\lambda)$ are zero. When this occurs, then by subtracting $q_{i1}(\lambda)$ times row one from row i and $q_{1k}(\lambda)$ times column one from column k, we can reduce the λ-matrix modulo m to a matrix in which $a_{i1}(\lambda) = a_{1k}(\lambda) \not\equiv 0$, $(i,k=2,\ldots,n)$.

If any of the nonzero elements $a_{ik}(\lambda)$ $(i,k=2,\ldots,n)$ is not divisible modulo m without remainder by $a_{11}(\lambda)$, then we can add the column containing this element to column one and again apply the procedure described above, thereby replacing $a_{11}(\lambda)$ again by an element of smaller degree. After

performing a finite number of elementary operations, we must obtain a matrix of the form stated in the lemma. ///

If $A_2(\lambda)$ is not identically equal to zero, then we can apply to this matrix the procedure just applied to $A(\lambda)$. During this process, clearly, all elements of $A_2(\lambda)$ remain multiples of $a_{11}(\lambda)$. Hence we obtain a matrix equivalent modulo m over $F[\lambda]$ to $A(\lambda)$ which has the form

$$B_2(\lambda) = \left[\begin{array}{cc|c} f_1(\lambda) & & \\ & f_2(\lambda) & \\ \hline & & A_3(\lambda) \end{array}\right]$$

in which $f_1(\lambda)$ divides $f_2(\lambda)$, and $f_2(\lambda)$ divides all the elements of $A_3(\lambda)$.

Continuing in this manner, we obtain a matrix of the form

$$B_r(\lambda) = \left[\begin{array}{ccc|c} f_1(\lambda) & & & \\ & f_2(\lambda) & & \\ & & \ddots & \\ & & f_r(\lambda) & \\ \hline & & & \end{array}\right], \tag{3.2}$$

where each $f_i(\lambda)$ divides $f_{i+1}(\lambda)$. The process terminates when $r = n$ or when $A_{r+1}(\lambda)$ is the null matrix modulo m. Since rank is invariant under equivalence transformations modulo m over $F[\lambda]$, we must have $r = r_m(A(\lambda))$. By multiplying the first r rows by suitable nonzero constants, followed by reduction modulo m, each $f_i(\lambda)$ can be made monic. (We know that such constants exist since m is a prime.) Thus, we have proved the following theorem:

(3.3) THEOREM. Each λ-matrix modulo m of rank r is equivalent modulo m over $F[\lambda]$ to a λ-matrix modulo m $B_r(\lambda)$ such that

(i) $B_r(\lambda)$ is diagonal,

(ii) the first r diagonal elements of $B_r(\lambda)$ are monic polynomials modulo m $f_1(\lambda),\ldots,f_r(\lambda)$,

(iii) the diagonal elements in row r+1 through row n are zero (if $r \neq n$),

(iv) $f_i(\lambda)$ divides $f_{i+1}(\lambda)$, i=1,...,r-1.

4. Invariant Factors Modulo m. Let $A(\lambda)$ be a λ-matrix modulo m of rank r which is in the form (3.2), and let $d_j(\lambda)$ denote the greatest common divisor modulo m of all the minors modulo m of order j (j=1,...,n) in $A(\lambda)$. If $j > r$, then all j-rowed minors modulo m are zero. For $j \leq r$, the only nonzero j-rowed minors modulo m are congruent to products of j of the diagonal elements $f_i(\lambda)$. Thus, since each $f_i(\lambda)$ is a factor of $f_{i+1}(\lambda)$ modulo m, we have

$$d_j(\lambda) = \left|f_1(\lambda) \ldots f_j(\lambda)\right|_m, \quad (j=1,\ldots,r). \tag{4.1}$$

If we let $d_o(\lambda) = 1$, then

$$f_j(\lambda) = \left|\frac{d_j(\lambda)}{d_{j-1}(\lambda)}\right|_m \quad (j=1,\ldots,r).$$

Before we state the main theorem of this section, we need to prove two lemmas.

(4.2) LEMMA. If $B(\lambda) = \left|P(\lambda)A(\lambda)\right|_m$, where $P(\lambda)$ is a product of elementary λ-matrices modulo m, then the jxj subdeterminants modulo m of $B(\lambda)$ are congruent to a linear combination, with coefficients over $F[\lambda]$, of the jxj subdeterminants modulo m of $A(\lambda)$.

Proof. [Perlis, 1952, p. 127] It is dufficient to let $P(\lambda)$ be an elementary λ-matrix modulo m, $E(\lambda)$, and to prove the theorem for $B(\lambda) = \left|E(\lambda)A(\lambda)\right|_m$.

We shall let $B_1(\lambda)$ be a jxj submatrix of $B(\lambda)$ and let $A_1(\lambda)$ be the submatrix of $A(\lambda)$ occupying the same rows and columns as $B_1(\lambda)$. If $E(\lambda)$ does not affect the rows of $A_1(\lambda)$, then $|\det B_1(\lambda)|_m = |\det A_1(\lambda)|_m$. If $E(\lambda)$ interchanges two rows of $A_1(\lambda)$, then $|\det B_1(\lambda)|_m = -|\det A_1(\lambda)|_m$, and if it interchanges a row of $A_1(\lambda)$ with a row of $A(\lambda)$ not in $A_1(\lambda)$, then $|\det B_1(\lambda)|_m$ is plus or minus some other jxj subdeterminant modulo m of $A(\lambda)$. If $E(\lambda)$ multiplies a row of $A_1(\lambda)$ by some nonzero constant c, then $|\det B_1(\lambda)|_m = |c \cdot \det A_1(\lambda)|_m$. When $E(\lambda)$ adds $f(\lambda)$ times row j to row i, and row j is in $A_1(\lambda)$, then $|\det B_1(\lambda)|_m = |\det A_1(\lambda)|_m$. If row i is in $A_1(\lambda)$ and row j is not, then $|\det B_1(\lambda)|_m = \big||\det A_1(\lambda)|_m + f(\lambda)\cdot|\det C(\lambda)|_m\big|_m$, where $|\det C(\lambda)|_m$ is some jxj subdeterminant modulo m of $A(\lambda)$. Thus, $|\det B(\lambda)|_m$ is congruent to a sum of terms of the form $\alpha_\lambda P_A(\lambda)$, where α_λ is a coefficient over $F[\lambda]$ and $P_A(\lambda)$ is a jxj subdeterminant modulo m of $A(\lambda)$.

///

(4.3) LEMMA. If $P(\lambda)$ and $Q(\lambda)$ are products of elementary λ-matrices modulo m, then the gcd of all jxj subdeterminants modulo m of $C(\lambda) = |P(\lambda)A(\lambda)Q(\lambda)|_m$ is the gcd of all jxj subdeterminants modulo m of $A(\lambda)$, a λ-matrix modulo m.

Proof. We shall use the notation $g_M(\lambda)$ to denote the gcd of all jxj subdeterminants modulo m of $M(\lambda)$, a λ-matrix modulo m. Let

$$B(\lambda) = |P(\lambda)A(\lambda)|_m .$$

Then

$$C(\lambda) = |B(\lambda)Q(\lambda)|_m$$

and

$$C^T(\lambda) = |Q^T(\lambda)B^T(\lambda)|_m .$$

From lemma (4.2), $g_A(\lambda)$ divides $g_B(\lambda)$. Also, $g_{B^T}(\lambda)$ divides $g_{C^T}(\lambda)$. But

$g_{B^T}(\lambda) = g_B(\lambda)$ and $g_{C^T}(\lambda) = g_C(\lambda)$. Thus, $g_B(\lambda)$ divides $g_C(\lambda)$, and hence $g_A(\lambda)$ divides $g_C(\lambda)$.

Since

$$A(\lambda) = \left| [P(\lambda)]^{-1}(m)\ C(\lambda)\ [Q(\lambda)]^{-1}(m) \right|_m ,$$

and $[P(\lambda)]^{-1}(m)$ and $[Q(\lambda)]^{-1}(m)$ are also λ-matrices modulo m, then an argument similar to the one above shows that $g_C(\lambda)$ divides $g_A(\lambda)$. Since both $g_A(\lambda)$ and $g_C(\lambda)$ are monic polynomials modulo m, we must have $g_A(\lambda) = g_C(\lambda)$. ///

We can now prove the uniqueness of the canonical form (3.2).

(4.4) THEOREM. [Perlis, 1952, p. 128] All λ-matrices which are equivalent modulo m over $F[\lambda]$ have the same canonical form (3.2).

Proof. Let $A(\lambda)$ be a nonzero λ-matrix modulo m and $B(\lambda)$ a λ-matrix which is equivalent modulo m over $F[\lambda]$ and which has the form stated in theorem (3.3). From (4.1) we see that the gcd of all jxj subdeterminants modulo m of $B(\lambda)$ is $d_j(\lambda) = \left| f_1(\lambda) \ \ldots\ f_j(\lambda) \right|_m$. By lemma (4.3), the gcd of all jxj subdeterminants modulo m of $A(\lambda)$ must also be $d_j(\lambda)$. It remains to show, then, that the $f_i(\lambda)$ are unique. Suppose

$$d_j(\lambda) = \left| f_1(\lambda) \ \ldots\ f_j(\lambda) \right|_m = \left| h_1(\lambda) \ \ldots\ h_j(\lambda) \right|_m ,$$

where the $h_i(\lambda)$ are monic polynomials modulo m. Then,

$$d_1(\lambda) = \left| f_1(\lambda) \right|_m = \left| h_1(\lambda) \right|_m$$

and

$$\left| \frac{d_j(\lambda)}{d_{j-1}(\lambda)} \right|_m = \left| \frac{\left| f_1(\lambda) \ \ldots\ f_j(\lambda) \right|_m}{\left| f_1(\lambda) \ \ldots\ f_{j-1}(\lambda) \right|_m} \right|_m = \left| \frac{\left| h_1(\lambda) \ \ldots\ h_j(\lambda) \right|_m}{\left| h_1(\lambda) \ \ldots\ h_{j-1}(\lambda) \right|_m} \right|_m$$

$$= \left| f_j(\lambda) \right|_m = \left| h_j(\lambda) \right|_m , \ (j=2,\ldots,r).$$

Thus, the canonical form (3.2) is unique.

///

(4.5) DEFINITION. The polynomials $f_i(\lambda)$ are called the invariant factors modulo m or the invariant polynomials modulo m of $A(\lambda)$, the square λ-matrix modulo m.

Hence, from theorem (4.4) and definition (4.5) we have the following theorem.

(4.6) THEOREM. Two λ-matrices modulo m are equivalent modulo m over $F[\lambda]$ if and only if they have the same invariant factors modulo m.

With this background, we are now prepared to discuss one of the most important λ-matrices modulo m.

5. The Characteristic Matrix Modulo m. An important role in the study of the residue matrix a over F is played by the nonzero vector x over F such that

(5.1)
$$|Ax|_m = |\lambda x|_m,$$

where

$$|x|_m = x$$

and

$$|\lambda|_m = \lambda,$$

for some scalar λ over F. The scalar λ is called an eigenvalue modulo m of A. The vector x is called the eigenvector modulo m of A which is associated with λ, the eigenvalue modulo m. Equation (5.1) may also be written in the form

(5.2)
$$|(A-\lambda I)x|_m = \emptyset$$

where the λ-matrix modulo m, $A-\lambda I$, is called the characteristic matrix

modulo m of A.

(5.3) THEOREM. There exists a nonzero (nontrivial) solution for the residue matrix equation (5.2) if and only if the matrix $A-\lambda I$ is singular modulo m for some value of λ. In other words, we have a nontrivial solution for (5.2) if and only if

$$|\det (A-\lambda I)|_m = 0. \tag{5.4}$$

Proof. Assume that for some value of λ the matrix $A-\lambda I$ is singular modulo m. If we let $r = r_m(A-\lambda I)$, where A is nxn, then $r < n$. Thus, we can solve for r of the unknowns, $x_{i_1}, \ldots, x_{i_r}$, in terms of the remaining n-r unknowns, $x_{i_{r+1}}, \ldots, x_{i_n}$, by using Gaussian elimination for residue arithmetic [Howell and Gregory, 1969a, pp. 217-220] and by assigning arbitrary values to the unknowns $x_{i_{r+1}}, \ldots, x_{i_n}$. Thus, if not all of the unknowns $x_{i_{r+1}}, \ldots, x_{i_n}$ are assigned the value zero, we have a nontrivial solution for (5.2).

Suppose (5.2) has a nontrivial solution. Then we cannot have $r = n$, for this would imply the existence of $(A-\lambda I)^{-1}(m)$, and hence we would have

$$\begin{aligned} x &= (A-\lambda I)^{-1}(m).\emptyset \\ &= \emptyset, \end{aligned}$$

the trivial solution. Therefore, $r < n$. ///

The values of λ for which (5.2) has a nontrivial solution are the roots of equation (5.4). The determinantal equation (5.4) can be expanded to give the polynomial equation modulo m of degree n

$$|(-1)^n(\lambda^n - b_1\lambda^{n-1} - \ldots - b_{n-1}\lambda - b_n)|_m = 0 \tag{5.5}$$

or

$$|a_o(-\lambda)^n + a_1(-\lambda)^{n-1} + \ldots + a_{n-1}(-\lambda) + a_n|_m = 0 \tag{5.6}$$

whose coefficient a_o is 1 and whose constant term $a_n = |\det A|_m$. In general, a_i is the residue modulo m of the sum of the principal minors modulo m of order i of the residue matrix A. Equation (5.4), (5.5), or (5.6) is called the characteristic equation modulo m of the residue matrix A, and the polynomial on the left-hand side of (5.4), (5.5), or (5.6) is called its characteristic polynomial modulo m. Clearly there are at most n eigenvalues modulo m for the matrix A (since $n < m$).

EXAMPLE. Let m = 13 and

$$A = \begin{bmatrix} 1 & -3 & -19 \\ -12 & 3 & 3 \\ 0 & 1 & 10 \end{bmatrix}.$$

Then the characteristic matrix modulo m for A is

$$|A-\lambda I|_{13} = \begin{bmatrix} 1-\lambda & -3 & -6 \\ 1 & 3-\lambda & 3 \\ 0 & 1 & -3-\lambda \end{bmatrix},$$

and the characteristic polynomial modulo m for A is

$$|\det (A-\lambda I)|_{13} = |(-1)^3(\lambda^3 - \lambda^2 + 4\lambda + 1)|_{13}.$$

It is easily verified that the eigenvalues modulo 13 and their associated eigenvectors modulo 13 for A are as follows:

$$\lambda_1 = 2,$$

$$\lambda_2 = -5,$$

$$\lambda_3 = 4,$$

$$x_1 = \begin{bmatrix} 5 \\ 5 \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 4 \\ -6 \\ 1 \end{bmatrix}.$$

We note that any multiple $k \not\equiv 0 \pmod{m}$ of the vectors x_i would also satisfy

$$|Ax|_m = |\lambda x|_m.$$

We have an explicit expression for the characteristic polynomial modulo m.

(5.7) THEOREM. The characteristic polynomial modulo m of A equals $|(-1)^n f_1(\lambda) \ldots f_n(\lambda)|_m$, where the $f_i(\lambda)$ are the invariant factors modulo m of $|A-\lambda I|_m$.

Proof. Since the characteristic polynomial modulo m is $|\det (A-\lambda I)|_m$, we shall examine $|A-\lambda I|_m$. By theorem (3.3) we can reduce $|A-\lambda I|_m$ to the diagonal canonical form (3.2), $B(\lambda)$, by means of equivalence transformations modulo m, given by

$$B(\lambda) = |P(\lambda)\ (A-\lambda I)\ Q(\lambda)|_m.$$

Since $|\det (A-\lambda I)|_m$ in its matrix polynomial form (5.5) has a leading coefficient of $(-1)^n$, then the rank modulo m of $|A-\lambda I|_m$ equals n. Furthermore since $|\det B(\lambda)|_m$ must be a monic polynomial modulo m, then we must have

$$|\det P(\lambda) \cdot \det Q(\lambda)|_m = |(-1)^n|_m.$$

Thus,

$$|\det B(\lambda)|_m = |\det P(\lambda) \cdot \det (A-\lambda I) \cdot \det Q(\lambda)|_m$$

$$= |(-1)^n \det (A-\lambda I)|_m,$$

and so

$$|\det (A-\lambda I)|_m = |(-1)^n \det B(\lambda)|_m$$

$$= |(-1)^n f_1(\lambda) \ \ldots \ f_n(\lambda)|_m. \qquad ///$$

EXAMPLE. If we let $A(\lambda)$ be the characteristic matrix modulo m in the previous example and compute the invariant factors modulo 13 of $A(\lambda)$ by the method described in the proof of lemma(3.1), then we find that

$$f_1(\lambda) = f_2(\lambda) = 1,$$

and

$$f_3(\lambda) = \lambda^3 - \lambda^2 + 4\lambda + 1.$$

Thus,

$$|\det A(\lambda)|_{13} = |(-1)^3 f_1(\lambda) f_2(\lambda) f_3(\lambda)|_{13}$$

$$= |(-1)^3(\lambda^3 - \lambda^4 + 4\lambda + 1)|_{13},$$

and hence

$$|\det (A-\lambda I)|_{13} = |(-1)^3(\lambda^3 - \lambda^2 + 4\lambda + 1)|_{13}$$

as stated in theorem (5.7).

We shall now exhibit a particular property of the characteristic equation modulo m. But first, we need to introduce a definition and a lemma.

(5.8) DEFINITION. Let $f(\lambda)$ be a polynomial modulo m of degree $n > 0$. If $f(A) = \emptyset$, then $f(\lambda)$ is said to be an annihilating polynomial modulo m for the matrix A.

(5.9) LEMMA. Let

$$C(\lambda) = |C_k\lambda^k + \ldots + C_1\lambda + C_o|_m$$

and

$$A(\lambda) = |A-\lambda I|_m.$$

If we divide $C(\lambda)$ b $A(\lambda)$ modulo m on the right or on the left as described in lemma (2.3), we obtain for $R_r(A)$ and $R_\ell(A)$

$$C_r(A) = |C_k A^k + \ldots + C_1 A + C_o|_m$$

or

$$C_\ell(A) = |A^k C_k + \ldots + A C_1 + C_0|_m \quad ,$$

respectively.

Proof. [Perlis, 1952, p. 135] By direct multiplication we see that

$$(5.10) \qquad \left| |\lambda^j I + \ldots + \lambda A^{j-2} + A^{j-1}|_m \; |A-\lambda I|_m \right|_m = |A^j - \lambda^j I|_m.$$

Multiplying both sides on the left by C_j, we have

$$(5.11) \qquad \left| |\lambda^j C_j + \ldots + \lambda C_j A^{j-2} + C_j A^{j-1}|_m \; |A-\lambda I|_m \right|_m = |C_j A^j - \lambda^j C_j|_m.$$

Then summing both sides over j, we obtain

$$\left| \sum_{j=1}^{k} |C_j A^j - \lambda^j C_j|_m \right|_m = \left| \sum_{j=1}^{k} |C_j A^j|_m - \sum_{j=1}^{k} |\lambda^j C_j|_m \right|_m$$

$$= \left| |C_o + \sum_{j=1}^{k} C_j A^j|_m - |C_o + \sum_{j=1}^{k} \lambda^j C_j|_m \right|_m$$

$$= |C_r(A) - C(\lambda)|_m$$

$$= \left| \sum_{j=1}^{k} |\lambda^{j-1} C_j + \ldots + \lambda C_j A^{j-2} + C_j A^{j-1}|_m \; |A-\lambda I|_m \right|_m$$

$$= \left| Q(\lambda) \; |A-\lambda I|_m \right|_m .$$

Thus,

$$|C_r(A) - C(\lambda)|_m = \left| Q(\lambda) \; |A-\lambda I|_m \right|_m$$

and so

$$|C(\lambda)|_m = \left| -Q(\lambda) \; |A-\lambda I|_m + C_r(A) \right|_m.$$

By uniqueness of the remainder in lemma (2.3), we have

$$C_r(A) = R_r(A).$$

Similarly, by reversing the two factors on the left in (5.10) and by multiplying on the right in (5.11) we can show that

$$C_\ell(A) = R_\ell(A). \qquad ///$$

(5.12) THEOREM. Let $f(\lambda) = |(-1)^n(\lambda^n - b_1\lambda^{n-1} - \ldots - b_{n-1}\lambda - b_n)|_m$ be the characteristic polynomial modulo m for A. Then $f(A) = \emptyset$.

Proof. [Eves, 1966, p. 201] From (II-3.12) we have

$$|(A-\lambda I)(A-\lambda I)^{adj}|_m = ||\det(A-\lambda I)|_m \cdot I\,|_m$$

$$= |(-1)^n(\lambda^n - b_1\lambda^{n-1} - \ldots - b_n)\cdot I\,|_m$$

$$= |(-1)^n(\lambda^n I - b_1\lambda^{n-1}I - \ldots - b_n I)|_m. \qquad (5.13)$$

Since the matrix polynomial on the right-hand side of (5.13) is divisible modulo m on the left by $|A-\lambda I|_m$, then by lemma (5.9), the remainder is

$$|R_\ell(A)|_m = |(-1)^n(A^n - b_1A^{n-1} - \ldots - b_n)|_m$$

$$= \emptyset.$$

Therefore

$$|f(A)|_m = \emptyset. \qquad ///$$

In other words, the characteristic polynomial modulo m is an annihilating polynomial modulo m.

EXAMPLE. If we let A be the same matrix as in the last two examples,

$$A = \begin{bmatrix} 1 & -3 & -19 \\ -12 & 3 & 3 \\ 0 & 1 & 10 \end{bmatrix} ,$$

and $m = 13$, then

$$|f(A)|_{13} = |(-1)^3(A^3 - A^2 + 4A + I)|_{13}$$

$$= \left| (-1)^3 \left(\begin{bmatrix} 6 & 7 & 1 \\ 0 & -4 & -5 \\ 1 & -4 & -3 \end{bmatrix} - \begin{bmatrix} -2 & -5 & 3 \\ 4 & -4 & -6 \\ 1 & 0 & -1 \end{bmatrix} + \begin{bmatrix} 4 & 1 & 2 \\ 4 & -1 & -1 \\ 0 & 4 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \right|_{13}$$

$$= |(-1)^3 \cdot \emptyset|_{13}$$

$$= \emptyset .$$

CHAPTER IV

SIMILARITY MODULO m OF MATRICES

1. Introduction. Transformations of the form

$$|B|_m = |C^{-1}(m)\ A\ C|_m \tag{1.1}$$

on the residue matrix A, where C is a residue matrix which is nonsingular modulo m, are of fundamental importance here and are known as similarity transformations modulo m. Then, B is said to be similar modulo m to the matrix A. The relation of similarity modulo m is symmetric, since from theorem (II-3.13) we get

$$\begin{aligned} |A|_m &= |C\cdot C^{-1}(m)\ A\ C\cdot C^{-1}(m)|_m \\ &= |C\ B\ C^{-1}(m)|_m \\ &= |(C^{-1}(m))^{-1}(m)\cdot B\cdot C^{-1}(m)|_m \ . \end{aligned} \tag{1.2}$$

Thus, it is sufficient to say merely that B and A are similar modulo m.

If we let C be any residue matrix which commutes with A, for example, the identity matrix, then we can show that similarity modulo m is reflexive, since

$$\begin{aligned} |C^{-1}(m)\ A\ C|_m &= \big||C^{-1}(m)C|_m\ A\big|_m \\ &= |A|_m \ . \end{aligned} \tag{1.3}$$

Transitivity also holds, since if

$$|D|_m = |F^{-1}(m)\ B\ F|_m$$

and

$$|B|_m = |C^{-1}(m)\ A\ C|_m,$$

then, using theorem (II-3.13) we get

$$|D|_m = \left|F^{-1}(m)\ |B|_m\ F\right|_m$$

$$= \left|F^{-1}(m)\ |C^{-1}(m)\ A\ C|_m\ F\right|_m$$

$$= \left|\,|F^{-1}(m)C^{-1}(m)|_m\ A\ |CF|_m\,\right|_m$$

$$= \left|(CF)^{-1}(m)\ A\ |CF|_m\ \right|_m . \tag{1.4}$$

Thus, we have proved the following

(1.5) THEOREM. Similarity modulo m is an equivalence relation.

EXAMPLE. Let A be the matrix used in the last example,

$$A = \begin{bmatrix} 1 & -3 & -19 \\ -12 & 3 & 3 \\ 0 & 1 & 10 \end{bmatrix},$$

and

$$B = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & -4 \\ 0 & 1 & 1 \end{bmatrix}.$$

Then A and B are similar modulo 13 since

$$|S^{-1}(m)\ A\ S|_{13} = \left|\begin{bmatrix} 1 & -1 & 6 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -3 & -19 \\ -12 & 3 & 3 \\ 0 & 1 & 10 \end{bmatrix}\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}\right|_{13}$$

$$= \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & -4 \\ 0 & 1 & 1 \end{bmatrix}$$

$$= B.$$

2. Basic Properties of Similarity Modulo m. We shall now exhibit the relationship between similarity modulo m of two matrices, A and B, and equivalence modulo m over $F[\lambda]$ of their corresponding λ-matrices modulo m, $A-\lambda I$ and $B-\lambda I$.

(2.1) THEOREM. Two residue matrices A and B are similar modulo m if and only if $A-\lambda I$ and $B-\lambda I$ have the same invariant factors modulo m, that is, if and only if $A-\lambda I$ and $B-\lambda I$ are equivalent modulo m over $F[\lambda]$.

Proof. [Perlis, 1952, pp. 143-144], [Gantmacher, 1960, pp. 147-148] Let A and B be similar modulo m. Then there exists a residue matrix C which is nonsingular modulo m and such that

$$B = |C^{-1}(m)\ A\ C|_m .$$

Then

$$\begin{aligned} |C^{-1}(m)(A-\lambda I)C|_m &= |C^{-1}(m)AC - C^{-1}(m)\ (\lambda I)\ C|_m \\ &= \left||C^{-1}(m)AC|_m - \lambda I\right|_m \\ &= |B-\lambda I|_m \\ &= B-\lambda I. \end{aligned}$$

Thus, $A-\lambda I$ and $B-\lambda I$ are strictly equivalent modulo m, and hence, by theorem (III-2.28) they are equivalent over $F[\lambda]$. Then, from theorem (III-4.14), $A-\lambda I$ and $B-\lambda I$ have the same invariant factors modulo m.

Suppose $A-\lambda I$ and $B-\lambda I$ have the same invariant factors modulo m. Then these λ-matrices modulo m are equivalent modulo m over $F[\lambda]$, and there exist λ-matrices modulo m $P(\lambda)$ and $Q(\lambda)$ such that

$$B-\lambda I = |P(\lambda)\ (A-\lambda I)\ Q(\lambda)|_m \ . \tag{2.2}$$

From theorem (III-2.28), we can write (2.2) as

$$B-\lambda I = |S\ (A-\lambda I)\ T|_m$$

$$= \big|\,|SAT|_m - |\lambda \cdot ST|_m\,\big|_m \ . \tag{2.3}$$

Equating coefficients on the powers of λ on both sides of (2.3), we obtain

$$B = |SAT|_m$$

and

$$I_n = |ST|_m .$$

Hence

$$S = T^{-1}(m)$$

and so

$$B = |T^{-1}(m)\ A\ T|_m \ .$$

Therefore, A and B are similar modulo m. ///

The following is an important result which concerns the characteristic polynomials modulo m of matrices which are similar modulo m.

(2.4) THEOREM. If $|B|_m = |C^{-1}(m)\ A\ C|_m$, then

$$\big|\det |B-\lambda I|_m\,\big|_m = \big|\det |A-\lambda I|_m\,\big|_m \ .$$

Proof.
$$\big|\det |B-\lambda I|_m\,\big|_m = \big|\det \big||B|_m - \lambda I\big|_m\,\big|_m$$

$$= \big|\det \big||C^{-1}(m)\ A\ C|_m - \lambda I\big|_m\,\big|_m$$

$$= \big|\det \big||C^{-1}(m)AC|_m - |C^{-1}(m)\ (\lambda I)\ C|_m\,\big|_m\,\big|_m$$

$$= \big|\det \big|C^{-1}(m)\cdot|A-\lambda I|_m\cdot C\big|_m\,\big|_m$$

$$= \left| \left| \det C^{-1}(m) \right|_m \cdot \left| \det \left| A-\lambda I \right|_m \right|_m \cdot \left| \det C \right|_m \right|_m$$

$$= \left| \det \left| A-\lambda I \right|_m \right|_m . \qquad ///$$

Thus, from the above theorem we see that matrices which are similar modulo m have the same eigenvalues modulo m.

Since the transforming matrix C is nonsingular modulo m, we have the following theorem.

(2.5) THEOREM. If $|B|_m = |C^{-1}(m)\ A\ C|_m$, then A and B have the same rank $(r_m(A) = r_m(B))$.

Proof. This follows from theorem (II-3.17).

Furthermore, if the matrix A is nonsingular modulo m, then we can prove the following theorem.

(2.6) THEOREM. If $|B|_m = |C^{-1}(m)\ A\ C|_m$, then A and B have the same determinant modulo m.

Proof. If A is singular modulo m then from theorem (2.5), both A and B have determinants modulo m which are 0. If A is nonsingular modulo m, then we must show that

$$|\det A|_m = |\det B|_m .$$

We have, from theorem (II-3.14) and (II-3.15)

$$|\det B|_m = |\det [C^{-1}(m)\ A\ C]|_m$$

$$= \left| \left| \det C^{-1}(m) \right|_m \cdot \left| \det A \right|_m \cdot \left| \det C \right|_m \right|_m$$

$$= \left| \left| \det C \right|_m^{-1}(m) \cdot \left| \det A \right|_m \cdot \left| \det C \right|_m \right|_m$$

$$= |\det A|_m . \qquad ///$$

EXAMPLE. We let

$$A = \begin{bmatrix} 1 & -3 & -19 \\ -12 & 3 & 3 \\ 0 & 1 & 10 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & -4 \\ 0 & 1 & 1 \end{bmatrix} .$$

In the example following theorem (1.5) we showed that A and B are similar modulo 13. It is easily verified that

$$|\det A|_{13} = -1$$

$$= |\det B|_{13} .$$

We note that if A is of the form

$$A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix}$$

then we can produce a matrix which is similar modulo m to A by permuting the blocks in A in an arbitrary manner along the diagonal.

(2.7) THEOREM. If

$$A = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix}$$

then A is similar modulo m to a matrix

$$B = \begin{bmatrix} A_{i_1} & & \\ & \ddots & \\ & & A_{i_k} \end{bmatrix}$$

where the subscripts $i_1,\ldots,i_k$ denote an arbitrary arrangement of the integers 1, ..., k.

Proof. [Perlis, 1952, pp. 150-151] It is sufficient to show that we can permute any two blocks, say the ith and jth, by means of similarity transformations modulo m. Suppose A is partitioned into k blocks each of order n_k. Then we partition the identity matrix similarly into k blocks of order n_k. Thus,

$$I_n = \begin{bmatrix} I_{n_1} & & & & & & \\ & \ddots & & & & & \\ & & I_{n_i} & & & & \\ & & & \ddots & & & \\ & & & & I_{n_j} & & \\ & & & & & \ddots & \\ & & & & & & I_{n_k} \end{bmatrix}.$$

Then we choose

$$P = \begin{bmatrix} I_{n_1} & & & & & & \\ & \ddots & & & & & \\ & & & & I_{n_i} & & \\ & & & \ddots & & & \\ & & I_{n_j} & & & & \\ & & & & & \ddots & \\ & & & & & & I_{n_k} \end{bmatrix}, \tag{2.8}$$

where the ith and jth columns of blocks have been interchanged. It is easy to see that postmultiplying A by P interchanges the ith and jth columns of blocks in A. Since $|P \cdot P^T|_m = I_n$, then $|P^T AP|_m = |P^{-1}(m)AP|_m$ is a similarity

transformation modulo m. Furthermore, premultiplying $|AP|_m$ by $P^{-1}(m)$ produces a matrix B with the ith and jth diagonal blocks interchanged, thus proving the theorem.

///

(2.9) COROLLARY. If $A(\lambda)$ is a λ-matrix modulo m in block-diagonal form, then $A(\lambda)$ is equivalent modulo m over $F[\lambda]$ to a matrix $B(\lambda)$ which is a block-diagonal λ-matrix modulo m, where the blocks of $B(\lambda)$ are the same as those in $A(\lambda)$, but in permuted arrangement.

Proof. The proof is similar to the above proof. We choose P as in (2.8) and form $|P^T A(\lambda)P|_m = B(\lambda)$. Since P is a residue matrix over F which is nonsingular modulo m, then it is also a matrix over $F[\lambda]$. Thus, $|P^T A(\lambda)P|_m = B(\lambda)$ is an equivalence transformation over $F[\lambda]$.

///

3. The Minimum Polynomial Modulo m of a Matrix. Before applying our knowledge of the properties of matrices which are similar modulo m, we need to introduce one further concept regarding residue matrices. This is the concept of the minimum polynomial modulo m. We have shown in chapter III that every residue matrix A satisfies its characteristic equation modulo m. A natural question is whether there are other polynomial equations modulo m which are satisfied by the matrix A. In particular we would like to know when there exist polynomial equations of lower degree than the characteristic equation modulo m which are satisfied by A. We shall answer these questions in this section.

(3.1) DEFINITION. A monic polynomial modulo m of positive degree, $p(\lambda)$, for which $|p(A)|_m = \emptyset$ and which has minimal degree among all annihilating polynomials modulo m is called a minimum polynomial modulo m of A.

Among all monic annihilating polynomials modulo m for a matrix A, clearly there is one of minimal degree. The relationship between this minimum polynomial modulo m and all other annihilating polynomials modulo m for A is stated specifically below.

(3.2) THEOREM. Every annihilating polynomial modulo m for a matrix A is divisible modulo m by any minimum polynomial modulo m for A.

Proof. [Perlis, 1952, pp. 145-146], [Gantmacher, 1960, pp. 89-90] Let $p(\lambda)$ be a minimum polynomial modulo m for A and $f(\lambda)$ an arbitrary annihilating polynomial modulo m for A. Since m is a prime, we know from theorem (II-4.6) that we can find polynomials $q(\lambda)$ and $r(\lambda)$ such that

$$|f(\lambda)|_m = |p(\lambda)q(\lambda) + r(\lambda)|_m$$

where either $r(\lambda) \equiv \emptyset \pmod{m}$ or the degree of $r(\lambda)$ is less than that of $p(\lambda)$. In order to show that $f(\lambda)$ is divisible modulo m by $p(\lambda)$, we must show that $r(\lambda) \equiv \emptyset \pmod{m}$. We know that

$$|f(A)|_m = |p(A)q(A) + r(A)|_m .$$

Since $p(\lambda)$ and $f(\lambda)$ are both annihilating polynomials modulo m, we have that

$$|p(A)|_m = \emptyset$$

and

$$|f(A)|_m = \emptyset .$$

Hence, we have

$$|r(A)|_m = \emptyset,$$

in which case, $r(\lambda)$ is also an annihilating polynomial modulo m. But, because of the minimality of $p(\lambda)$, $r(\lambda)$ cannot be of degree less than that

of $p(\lambda)$. Therefore, $r(\lambda) \equiv \emptyset \pmod{m}$, and $p(\lambda)$ divides $f(\lambda)$ modulo m. ///

From the above theorem, we see that, in particular, a minimum polynomial modulo m divides the characteristic polynomial modulo m.

We shall now prove the uniqueness of a minimum polynomial modulo m for a matrix A.

(3.3) THEOREM. A matrix A has a unique minimum polynomial modulo m.

Proof. [Hohn, 1964, pp. 286-287] Suppose $p_1(\lambda)$ and $p_2(\lambda)$ are both minimum polynomials modulo m for A, and that both are of degree k. The, A satisfies both

$$p_1(\lambda) = |\lambda^k + \alpha_o\lambda^{k-1} + \ldots + \alpha_{k-1}\lambda + \alpha_k|_m$$

and

$$p_2(\lambda) = |\lambda^k + \beta_o\lambda^{k-1} + \ldots + \beta_{k-1}\lambda + \beta_k|_m .$$

Thus, A also satisfies

$$|p_1(\lambda) - p_2(\lambda)|_m = |(\alpha_o-\beta_o)\lambda^{k-1} + \ldots + (\alpha_{k-1}-\beta_{k-1})\lambda + (\alpha_k-\beta_k)|_m .$$

If $p_1(\lambda)$ and $p_2(\lambda)$ are different, then the polynomial $|p_1(\lambda) - p_2(\lambda)|_m$ must be of degree less than k. But, by the minimality of $p_1(\lambda)$ and $p_2(\lambda)$, A can satisfy no polynomial of degree less than k. Hence, $|p_1(\lambda) - p_2(\lambda)|_m$ must be the zero polynomial modulo m. Therefore, $p_1(\lambda) = p_2(\lambda)$ and the minimum polynomial modulo m for A is unique. ///

We shall now show that the invariant factor modulo m for $|A-\lambda I|_m$ of highest order is the minimum polynomial modulo m for A.

(3.4) LEMMA. The minimum polynomial modulo m for A, $p(\lambda)$, divides the highest order invariant factor modulo m for $|A-\lambda I|_m$.

Proof. [Perlis, 1952, p. 146], [Gantmacher, 1960, p. 90] In chapter III, section 4, we saw that the highest order invariant factor modulo m for the matrix $|A-\lambda I|_m$ is given by

$$|f_n(\lambda)|_m = \left|\frac{d_n(\lambda)}{d_{n-1}(\lambda)}\right|_m , \tag{3.5}$$

where $d_k(\lambda)$ is the greatest common divisor modulo m of all minors modulo m of order k for $|A-\lambda I|_m$. Then we can write

$$|f_n(\lambda)\cdot d_{n-1}(\lambda)|_m = |d_n(\lambda)|_m . \tag{3.6}$$

If we look at the adjoint modulo m for $|A-\lambda I|_m$, we see that the elements of $|(A-\lambda I)^{adj}|_m$ are all multiples modulo m of $d_{n-1}(\lambda)$. Furthermore, $d_{n-1}(\lambda)$ is the greatest common divisor modulo m of the elements of $|(A-\lambda I)^{adj}|_m$. Hence, we can write

$$|(A-\lambda I)^{adj}|_m = |d_{n-1}(\lambda)Q(\lambda)|_m , \tag{3.7}$$

where the greatest common divisor modulo m of the elements of the matrix $Q(\lambda)$ is 1. From theorem (II-3.12) we can write

$$|(A-\lambda I)\cdot(A-\lambda I)^{adj}|_m = |f(\lambda)\cdot I|_m ,$$

where $f(\lambda)$ is the characteristic polynomial modulo m of A. Since $|f(\lambda)|_m = |d_n(\lambda)|_m$, this becomes

$$|(A-\lambda I)\cdot(A-\lambda I)^{adj}|_m = |d_n(\lambda)\cdot I|_m . \tag{3.8}$$

Combining (3.6), (3.7), and (3.8), we have

$$|(A-\lambda I)\cdot d_{n-1}(\lambda)\cdot Q(\lambda)|_m = |f_n(\lambda)\cdot d_{n-1}(\lambda)\cdot I|_m ,$$

and hence

$$|(A-\lambda I)\cdot Q(\lambda)|_m = |f_n(\lambda)\cdot I|_m . \tag{3.9}$$

Therefore,

$$|f_n(A)|_m = \emptyset \ ,$$

and by theorem (3.2) we have that $p(\lambda)$ divides $f_n(\lambda)$. ///

(3.10) THEOREM. The minimum polynomial modulo m for A, $p(\lambda)$, is equal to $f_n(\lambda)$.

Proof. [Perlis, 1952, pp. 146-147], [Gantmacher, 1960, pp. 90-91] By lemma (III-5.9), we can divide the matrix polynomial modulo m $|p(\lambda)\cdot I|_m$, by the characteristic matrix modulo m of A, obtaining

(3.11) $$|p(\lambda)\cdot I|_m = |B(\lambda)\cdot(A-\lambda I) + p_r(A)|_m \ ,$$

where $p_r(A)$ is identically equal to zero. From lemma (3.4) we have that

(3.12) $$|f_n(\lambda)|_m = |h(\lambda)\cdot p(\lambda)|_m \ ,$$

and hence

(3.13) $$|f_n(\lambda)I|_m = |h(\lambda)\cdot p(\lambda)I|_m .$$

Substituting (3.9) and (3.11) into (3.13), we obtain

$$|(A-\lambda I)Q(\lambda)|_m = |g(\lambda)B(\lambda)(A-\lambda I)|_m$$

and hence

(3.14) $$|Q(\lambda)|_m = |h(\lambda)B(\lambda)|_m \ .$$

This implies that the polynomial $h(\lambda)$ divides all of the elements of $|Q(\lambda)|_m$ modulo m. From (3.7), we note that the greatest common divisor modulo m of the elements of $Q(\lambda)$ is one, so that $h(\lambda)$ must be a constant, k. Then

$$|f_n(\lambda)|_m = |k\cdot p(\lambda)|_m \ .$$

But, both $f_n(\lambda)$ and $p(\lambda)$ are monic polynomials modulo m. Thus,

$$f_n(\lambda) = p(\lambda).$$

///

From the above lemma and theorem, we have an expression for the minimum polynomial modulo m for a matrix A. Since $d_n(\lambda) = \left|(-1)^n \left|\det(A-\lambda I)\right|_m\right|_m$, then from (3.5) we obtain

$$(3.15) \qquad \left|p(\lambda)\right|_m = \left|\frac{(-1)^n \left|\det\,(A-\lambda I)\right|_m}{d_{n-1}(\lambda)}\right|_m$$

In case $\left|d_{n-1}(\lambda)\right|_m = 1$, then we see that the minimum polynomial modulo m for A is congruent to $(-1)^n$ times the characteristic polynomial modulo m for A.

(3.16) DEFINITION. A matrix A whose minimum polynomial modulo m is congruent to $(-1)^n$ times its characteristic polynomial modulo m is called nonderogatory modulo m. Otherwise, A is called derogatory modulo m.

We see that a matrix which is nonderogatory modulo m has only one invariant factor modulo m which is not congruent to one.

EXAMPLE. Let m = 13 and

$$A = \begin{bmatrix} 1 & -3 & -19 \\ -12 & 3 & 3 \\ 0 & 1 & 10 \end{bmatrix}.$$

The minimum polynomial modulo m for A is

$$p(\lambda) = \lambda^3 - \lambda^2 + 4\lambda + 1$$

$$= \left|(-1)^3 \det\,(A-\lambda I)\right|_{13}.$$

Thus, the matrix A is nonderogatory modulo 13.

EXAMPLE. Let m = 13 and

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 1 & 6 \end{bmatrix} .$$

The minimum polynomial modulo m for A is

$$p(\lambda) = \lambda^2 - 6\lambda - 5 ,$$

and the characteristic polynomial modulo m is

$$f(\lambda) = \left|(-1)^3(\lambda^3 + 3\lambda^2 + 5\lambda - 2)\right|_{13}.$$

Thus, A is derogatory modulo 13.

In the next section we examine a particular matrix which is nonderogatory modulo m, and we apply the results of section 2.

4. The Companion Matrix Modulo m. If we let $|g(\lambda)|_m$ be a monic polynomial modulo m,

$$(4.1) \qquad \left|(-1)^n(\lambda^n - b_1\lambda^{n-1} - \ldots - b_{n-1}\lambda - b_n)\right|_m = |g(\lambda)|_m ,$$

then associated with $|g(\lambda)|_m$ is an nxn residue matrix, $|C(g)|_m$, where

$$(4.2) \qquad |C(g)|_m = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & b_n \\ 1 & 0 & 0 & \ldots & 0 & b_{n-1} \\ 0 & 1 & 0 & \ldots & 0 & b_{n-2} \\ . & . & . & \ldots & . & . \\ 0 & 0 & 0 & \ldots & 0 & b_2 \\ 0 & 0 & 0 & \ldots & 1 & b_1 \end{bmatrix} ,$$

which we call the companion matrix modulo m for $|g(\lambda)|_m$. It is not difficult to show that $|g(\lambda)|_m$ is the characteristic polynomial modulo m for $|C(g)|_m$.

Furthermore, if we examine

$$(4.3) \qquad |C(g) - \lambda I|_m = \begin{bmatrix} -\lambda & 0 & 0 & \dots & 0 & b_n \\ 1 & -\lambda & 0 & \dots & 0 & b_{n-1} \\ 0 & 1 & -\lambda & \dots & 0 & b_{n-2} \\ . & . & . & \dots & . & . \\ 0 & 0 & 0 & \dots & -\lambda & b_2 \\ 0 & 0 & 0 & \dots & 1 & b_1-\lambda \end{bmatrix},$$

we see that the minor modulo m for the element b_n is $(-1)^{n-1}$. Then, the greatest common divisor modulo m for all n-1 x n-1 subdeterminants modulo m of $|C(g) - yI|_m$ is 1. Hence, we have $|d_{n-1}(\lambda)|_m = 1$. From (3.15), it follows that the minimum polynomial modulo m for $|C(g)|_m$ equals $(-1)^n$ times the characteristic polynomial modulo m for $|C(g)|_m$. Therefore, by definition (3.16) we have proved the following:

(4.4) THEOREM. The companion matrix modulo m, $|C(g)|_m$, is nonderogatory modulo m.

It follows from the above that $|C(g) - \lambda I|_m$ has only one nontrivial invariant factor modulo m.

(4.5) THEOREM. If $|B|_m = |S^{-1}(m)\ A\ S|_m$ and $|B|_m$ is the companion matrix modulo m for $|\det (A-\lambda I)|_m$ then $|A|_m$ is nonderogatory modulo m.

Proof. From theorem (2.1), $|A-\lambda I|_m$ and $|B-\lambda I|_m$ have the same invariant factors modulo m. Since $|B|_m$ is in the form (4.2), it is nonderogatory modulo m and has only one nontrivial invariant factor modulo m. Thus, $|A|_m$ also has only one nontrivial invariant factor modulo m and is nonderogatory modulo m. ///

(4.6) COROLLARY. If A is derogatory modulo m, then A cannot be similar modulo m to the companion matrix modulo m of $|\det (A-\lambda I)|_m$.

Proof. The proof is by contradiction. Suppose A is similar modulo m to the companion matrix modulo m of $|\det (A-\lambda I)|_m$. Then A is nonderogatory modulo m. Hence, we have a contradiction to the assumption that A is derogatory modulo m. ///

In the above discussion we have shown that A is similar modulo m to $|C(f)|_m$, where $f(\lambda) = |\det (A-\lambda I)|_m$ if and only if A is nonderogatory modulo m. We shall now show that when A is similar modulo m to

$$B = \begin{bmatrix} C(g_1) & \\ & C(g_2) \end{bmatrix} , \tag{4.7}$$

then A is nonderogatory modulo m if and only if $g_1(\lambda)$ and $g_2(\lambda)$ are relatively prime modulo m. In order to prove this, we need to exhibit a canonical form for residue matrices and prove two lemmas.

(4.8) THEOREM. Let $|A-\lambda I|_m$ have invariant factors modulo m $f_1(\lambda)$, ... , $f_n(\lambda)$. Then A is similar modulo m to

$$B = \begin{bmatrix} C(f_i) & & & \\ & C(f_{i+1}) & & \\ & & \ddots & \\ & & & C(f_n) \end{bmatrix} ,$$

where $f_1(\lambda) = \ldots = f_{i-1}(\lambda) = 1$ and $f_j(\lambda)$ divides $f_{j+1}(\lambda)$ modulo m $(j = 1,\ldots,n)$.

Proof. [Gantmacher, 1960, p. 142], [Perlis, 1952, p. 153] Let the matrices

$C(f_j)$, $j = i,\ldots,n$, have order n_j, so that the polynomial modulo m $f_j(\lambda)$ has degree n_j. Then the characteristic polynomial modulo m and the minimum polynomial modulo m of $C(f_j)$ are both equal to $f_j(\lambda)$. Since $|C(f_j) - \lambda I_{n_j}|_m$ has only one nontrivial invariant factor modulo m, it is equivalent modulo m to a matrix $|K_j(\lambda)|_m$ of order n_j, where

$$|K_j(\lambda)|_m = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & f_j(\lambda) \end{bmatrix} .$$

Then $|B-\lambda I|_m$ is equivalent modulo m to

$$|E(\lambda)|_m = \begin{bmatrix} K_i(\lambda) & & & \\ & K_{i+1}(\lambda) & & \\ & & \ddots & \\ & & & K_n(\lambda) \end{bmatrix} .$$

By corollary (2.9) we can reorder the elements in the blocks and obtain a matrix with all of the ones in the upper left-hand corner

$$|F(\lambda)|_m = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & f_i(\lambda) & \\ & & & & \ddots \\ & & & & & f_n(\lambda) \end{bmatrix} .$$

But $|F(\lambda)|_m$ is equivalent modulo m to $|A-\lambda I|_m$. Therefore, $|B-\lambda I|_m$ is equivalent modulo m to $|A-\lambda I|_m$, and by theorem (2.1), A and B are similar modulo m. ///

(4.9) LEMMA. If we let

$$|B|_m = \begin{bmatrix} D & \\ & G \end{bmatrix}$$

where D and G have minimum polynomials modulo m $d(\lambda)$ and $g(\lambda)$, respectively, then the minimum polynomial modulo m for $|B|_m$ is the least common multiple modulo m of $d(\lambda)$ and $g(\lambda)$.

Proof. [Perlis, 1952, p. 151] Let $|p(\lambda)|_m$ by the minimum polynomial modulo m for $|B|_m$. Then, by definition (3.1)

$$\emptyset = |p(B)|_m = \begin{bmatrix} p(D) & \\ & p(G) \end{bmatrix} .$$

But $|p(B)|_m = \emptyset$ if and only if both $|p(D)|_m = \emptyset$ and $|p(G)|_m = \emptyset$. By theorem (3.2), this implies that both $d(\lambda)$ and $g(\lambda)$ divide $p(\lambda)$ modulo m. From the definition of $|p(\lambda)|_m$, it is the monic polynomial modulo m of least degree such that $|p(B)|_m = \emptyset$, and hence it is also the polynomial of least degree such that $d(\lambda)$ and $g(\lambda)$ both divide $p(\lambda)$ modulo m. In other words, $|p(\lambda)|_m$ is the monic polynomial modulo m of smallest degree such that $|p(\lambda)|_m$ is a common multiple modulo m of $d(\lambda)$ and $g(\lambda)$. By definition (II-4.14), we have that $|p(\lambda)|_m$ is the least common multiple modulo m of $d(\lambda)$ and $g(\lambda)$.

///

EXAMPLE. Let m = 13 and

$$B = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 1 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} D & \\ & G \end{bmatrix} .$$

The minimum polynomial modulo m for D is

$$d(\lambda) = \lambda - 4$$

and the minimum polynomial modulo m for G is

$$g(\lambda) = \lambda^2 - 6\lambda - 5$$
$$= (\lambda - 4)(\lambda - 2) \pmod{13}.$$

Thus, the minimum polynomial modulo m for B is

$$p(\lambda) = |(\lambda - 4)(\lambda - 2)|_{13}$$

$$= \lambda^2 - 6\lambda - 5 \quad ,$$

which verifies the result obtained in the last example.

From the above lemma we see that if $d(\lambda)$ and $g(\lambda)$ are relatively prime modulo m, then

$$|p(\lambda)|_m = |d(\lambda)g(\lambda)|_m \quad . \tag{4.10}$$

(4.11) LEMMA. If the characteristic polynomial modulo m of $|A|_m$ is $|d(\lambda)g(\lambda)|_m$, where $d(\lambda)$ and $g(\lambda)$ are relatively prime modulo m, then $|A|_m$ is similar modulo m to a matrix B, where

$$B = \left| \begin{bmatrix} D & \\ & G \end{bmatrix} \right|_m ,$$

and where D and G have characteristic polynomials modulo m $d(\lambda)$ and $g(\lambda)$, respectively.

Proof. [Perlis, 1952, pp. 154-155] Let $|A-\lambda I|_m$ have the following invariant factors modulo m: $f_1(\lambda),\ldots,f_n(\lambda)$. Then, from theorem (4.8), A is similar modulo m to a matrix F, where

$$F = \begin{bmatrix} C(f_i) & & \\ & \ddots & \\ & & C(f_n) \end{bmatrix} ,$$

and where $f_1(\lambda) = \ldots = f_{i-1}(\lambda) = 1$. Then, by theorem (III-5.7),

$$|\det (A-\lambda I)|_m = \left| (-1)^n \prod_{j=i}^{n} f_j(\lambda) \right|_m \quad .$$

By hypothesis, then,

$$|d(\lambda)\cdot g(\lambda)|_m = |(-1)^n f_i(\lambda) \ \ldots \ f_n(\lambda)|_m \ .$$

From theorem (II-4.18), we can find unique monic polynomials modulo m, $d'(\lambda)$ and $g'(\lambda)$, such that

$$|f_j(\lambda)|_m = |d'_j(\lambda)g'_j(\lambda)|_m \ ,$$

where $d'_j(\lambda)$ divides $d(\lambda)$ modulo m and $g'_j(\lambda)$ divides $g(\lambda)$ modulo m. Since $d(\lambda)$ and $g(\lambda)$ are relatively prime modulo m, then $d'(\lambda)$ and $g'(\lambda)$ are relatively prime modulo m. Thus, from lemma (4.9), the the matrix

$$|G_j|_m = \begin{bmatrix} C(d'_j) & \\ & C(g'_j) \end{bmatrix}$$

has minimum polynomial modulo m $|d'_j(\lambda)g'_j(\lambda)|_m$. But this is also the characteristic polynomial modulo m. Hence, $|G_j|_m$ is nonderogatory modulo m and is similar modulo m to $|C(d'_j g'_j)|_m = C(f_j)$. This implies that F is similar modulo m to

$$\begin{bmatrix} G_i & & \\ & \ddots & \\ & & G_n \end{bmatrix} = \begin{bmatrix} C(d'_i) & & & & \\ & C(g'_i) & & & \\ & & \ddots & & \\ & & & C(d'_n) & \\ & & & & C(g'_n) \end{bmatrix} .$$

By theorem (2.7) we can shuffle the blocks and obtain a matrix similar modulo m to F,

$$\left[\begin{array}{ccc|ccc} C(d'_i) & & & & & \\ & \ddots & & & \bigcirc & \\ & & C(d'_n) & & & \\ \hline & & & C(g'_i) & & \\ & \bigcirc & & & \ddots & \\ & & & & & C(g'_n) \end{array}\right] = \begin{bmatrix} D & \\ & G \end{bmatrix} ,$$

where the characteristic polynomials modulo m for D and G are $d_i'(\lambda) \ldots d_n'(\lambda) = d(\lambda)$ and $g_i'(\lambda) \ldots g_n'(\lambda) = g(\lambda)$, respectively.

///

(4.12) THEOREM. Let A be similar modulo m to B, where

$$|B|_m = \begin{bmatrix} D & \\ & G \end{bmatrix},$$

and where D and G are companion matrices modulo m to $d(\lambda)$ and $g(\lambda)$. Then A is nonderogatory modulo m if and only if $d(\lambda)$ and $g(\lambda)$ are relatively prime modulo m.

Proof. First let $d(\lambda)$ and $g(\lambda)$ be relatively prime modulo m. By lemma (4.9), the minimum polynomial modulo m for $|B|_m$ is $|d(\lambda)g(\lambda)|_m$. But, since D and G are companion matrices modulo m for $d(\lambda)$ and $g(\lambda)$, then the characteristic polynomial modulo m for $|B|_m$ is $|d(\lambda)g(\lambda)|_m$. Thus, $|B|_m$ is nonderogatory modulo m, and hence, so is A.

Now let A be nonderogatory modulo m. Then, the minimum polynomial modulo m of A equals the characteristic polynomial modulo m of A, which is $|d(\lambda)g(\lambda)|_m$. But, from lemma (4.9), the minimum polynomial modulo m of B, and hence of A, is the least common multiple modulo m of $d(\lambda)$ and $g(\lambda)$. By uniqueness of the minimum polynomial modulo m, $d(\lambda)$ and $g(\lambda)$ are relatively prime modulo m.

///

In chapters II, III, and IV we surveyed the main theorems of residue arithmetic. We are now prepared to discuss the modified Danilewski method, which uses residue arithmetic.