ON THE REDUCTION OF A MATRIX TO FROBENIUS FORM USING RESIDUE ARITHMETIC

by

JO ANN SHAW HOWELL, B.A., M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN
August 1971

ON THE REDUCTION OF A MATRIX TO FROBENIUS FORM USING RESIDUE ARITHMETIC

APPROVED BY SUPERVISORY COMMITTEE:

Robert Todd Gregory

Roger Osborn

David M. Young Jr

G. W. Stewart

To Benton

PREFACE

This thesis is concerned with an algorithm for computing exactly the characteristic polynomial of an integral matrix A (or in some cases, a factorization of it over the integers). The algorithm described here, which uses residue arithmetic, is analogous to the Danilewski method for reducing a matrix to Frobenius form. The algorithm can be performed using either single-modulus or multiple-modulus residue arithmetic, and examples are given for both cases.

We take advantage of the fact that the integers modulo a prime form a finite field, F. Thus, all the theorems relative to matrices and polynomials over a field can be utilized in describing the algorithm. Using the residue arithmetic algorithm, we reduce a matrix to Frobenius form. The Frobenius form obtained using residue arithmetic is the residue modulo m of the Frobenius form obtained using rational arithmetic. Then, if the modulus is sufficiently large, the two Frobenius forms are the same. Thus, from the blocks along the diagonal of the Frobenius form we obtain the exact multiple-precision coefficients of the characteristic polynomial or of its factors over the integers.

The thesis begins in chapter I with a description of the reduction to Frobenius form using rational arithmetic. Chapters II, III, and IV survey the theory of residue arithmetic for integers, matrices, and polynomials. In chapter V the single-modulus algorithm for reducing a matrix to Frobenius form is described , and examples are given. The multiple-modulus algorithm is described in chapter VI. It is shown that different moduli may yield different factorizations. In section 3 a theorem is proved which gives an algorithm for determining which factorizations are incorrect. All of the material related to the multiple-modulus algorithm is thought to be new.

Bounds are given in section 5 for the number of moduli required to guarantee that the coefficients can be reconstructed using the Chinese Remainder Theorem. Examples are given which illustrate the algorithm in section 6. Chapter VII gives numerical results from a computer program.

The author wishes to express her appreciation to the Computation Center of The University of Texas at Austin for the use of their facilities in preparing computer programs based on the material in this thesis.

I wish to express my gratitude to my supervising professor Dr. Robert T. Gregory for his assistance during my final years of graduate study, and to both Dr. Gregory and Dr. David M. Young for giving me the opportunity to work for the Center for Numerical Analysis. I also with to thank Dr. Roger C. Osborn and Dr. G. W. Stewart for their suggestions. Special thanks go to my husband Robert Benton Howell for his patience and encouragement throughout my graduate study.

August 1971

ABSTRACT

This thesis is concerned with the reduction of an integral matrix to Frobenius form exactly using residue arithmetic. Thus, exact integral factors of the characteristic polynomial are obtained. The algorithm is based on a modification of the Danilewski method. This algorithm can be performed using either single-modulus or multiple-modulus residue arithmetic, and examples are given for both cases.

Included in this thesis is a description of the Danilewski method. The theory of residue arithmetic for integers, matrices, and polynomials is surveyed in order to provide an adequate background for describing the modified Danilewski method. The selection of the moduli is discussed, and numerical results from a computer program are given.

TABLE OF CONTENTS

CHAPTER I

THE DANILEWSKI METHOD

1. Introduction. It is well known that the Danilewski method [Danilewski, 1937] for reducing a matrix A to Frobenius form,

$$F = \begin{bmatrix} F_1 & & \ast & \\ & F_2 & & \\ \bigcirc & & \ddots & \\ & & & F_\ell \end{bmatrix}, \tag{1.1}$$

where each diagonal block has the form

$$F_i = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & x^{(i)}_{r_i-1} \\ 1 & 0 & 0 & \cdots & 0 & x^{(i)}_{r_i-2} \\ 0 & 1 & 0 & \cdots & 0 & x^{(i)}_{r_i-3} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & x^{(i)}_{1} \end{bmatrix}, \tag{1.2}$$

is numerically unstable [Frank, 1958]. Several attempts have been made to reduce the inaccuracies by using multiple-precision arithmetic and pivoting for size [Chartres, 1964], [Hansen, 1963]. These variations yield a Frobenius form much more accurately than previously reported. However, it has been shown that because of the ill-condition of the Frobenius form of a matrix, the Danilewski method and its variations usually prove unsatisfactory for determining eigenvalues [Wilkinson, 1965, pp. 405-411]. Even small errors in the diagonal blocks, the F_i, may lead to catastrophic errors in the eigenvalues.

Owing to the fact that a Frobenius form* of a matrix does give us a factorization of the characteristic polynomial and some information on the derogatory nature of the matrix, this condensed form is still of some interest to us. For matrices arising from damped mechanical or electrical systems it is common for the Frobenius form to be well-conditioned [Wilkinson, 1965, p. 482].

It is for these reasons that we describe here a modification of the Danilewski method with which we can reduce a matrix to Frobenius form, that is compute the F_i exactly, without the use of multiple-precision arithmetic or pivoting for size [Slotnick, 1963, pp. 4-42 - 4-46]. Since this modification uses residue (or modular) arithmetic, it is applicable only to integral matrices. This restriction is not serious, however, since fixed-word-length computers store only rational numbers which can be scaled to integer form. We observe that if the Frobenius form of the matrix A is

$$F_A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & x_n \\ 1 & 0 & 0 & \cdots & 0 & x_{n-1} \\ 0 & 1 & 0 & \cdots & 0 & x_{n-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & x_1 \end{bmatrix},$$

then the Frobenius form of the scaled matrix k·A is given by

$$F_{kA} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & k^n x_n \\ 1 & 0 & 0 & \cdots & 0 & k^{n-1} x_{n-1} \\ 0 & 1 & 0 & \cdots & 0 & k^{n-2} x_{n-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & kx_1 \end{bmatrix}.$$

* The nonuniqueness of this form (and hence of the factors) is discussed below.

Hence, given F_{kA} and k we can compute F_A.

In the remainder of this chapter we briefly describe the Danilewski method in order to provide an adequate background for describing the modified method. Other descriptions of the method are given in Wilkinson [1965, pp. 405-407], Householder and Bauer [1959], Householder[1964, pp. 156-158], and Wayland [1945].

2. The Algorithm. Using the Danilewski method we transform a matrix A, by means of similarity transformations, into a matrix F, which is in the form (1.1). The elements in the last column of the F_i are coefficients of the characteristic polynomial for F_i,

$$p_i(\lambda) = (-1)^{r_i}[\lambda^{r_i} - x_1^{(i)}\lambda^{r_i-1} - \dots - x_{r_i-2}^{(i)}\lambda^2 - x_{r_i-1}^{(i)}\lambda - x_{r_i}^{(i)}].$$

Thus, the characteristic polynomial for F is

$$p(\lambda) = p_1(\lambda) \dots p_\ell(\lambda).$$

Since A and F are similar, then $p(\lambda)$ is also the characteristic polynomial for A. Thus, by using the Danilewski method, we can compute the characteristic polynomial (or a factorization of it) for the matrix A. We observe also that the matrix F is a special case of a Hessenberg form for A.

The matrix F is obtained after a finite number of similarity transformations of the form

$$A_{k+1} = S_k^{-1}A_kS_k \qquad (k = 0,1,2,\dots M)$$

where $A_o = A$. It is recommended by both Hansen[1963] and Wilkinson [1965, p. 409] that the computation be broken into two stages. During the first stage, the matrix A is reduced to Hessenberg form. Wilkinson has shown that for this step of the algorithm, single-precision arithmetic is usually sufficient. In the second stage, the Hessenberg matrix if further reduced to

Frobenius form. It is during this stage that we generally need to work in higher precision arithmetic. (See, for example, [Chartres, 1964].)

We shall assume here that step one and step two are carried out using elementary similarity transformations, since analogous transformations will be used in carrying out the modified algorithm. These transformations consist of one of the following three types of operations:

(a.) Interchange of rows i and j (or columns i and j),

(b.) Multiplication of row i (or column i) by a nonzero constant,K,

(c.) Addition to the ith row of an arbitrary multiple, K, of row j (and the analogous operation on columns).

Thus, the S_k can be assumed to be the elementary matrices which are obtained from the identity matrix by performing one of the above operations on it. We denote these by I_{ij} , $E_i(K)$, and $E_{ij}(K)$, respectively.

We now consider the problem of reducing a matrix A to Hessenberg form. This is accomplished in n-2 major steps of the form

$$A_{k+1} = J_k^{-1} A_k J_k \ ,$$

where $A_o = A$ and J_k is a product of elementary matrices. Each transformation changes a matrix A_k into a matrix, A_{k+1} , in which there is a row of zeros (where there was not one before) below the subdiagonal. After the first transformation we have

$$A_1 = J_o^{-1} A_o J_o$$

$$= \left[\begin{array}{c|cc} & x & x \\ & x & x \\ * & \vdots & \vdots \\ & x & x \\ \hline O & x & x \end{array}\right] ,$$

where

$$J_o = \left[\begin{array}{cccc|c} \multicolumn{4}{c|}{I_{n-2}} & \bigcirc \\ \hline \mu_{n,1} & \mu_{n,2} & \cdots & \mu_{n,n-2} & \\ 0 & 0 & \cdots & 0 & I_2 \end{array}\right]$$

and

$$\mu_{n,i} = \frac{-a^{(0)}_{n,i}}{a^{(0)}_{n,n-1}} .$$

The case in which a pivotal element $a^{(0)}_{i,i-1} = 0$ is treated below.

The $(j + 1)$st transformation produces

$$\begin{aligned} A_{j+1} &= J_j^{-1} A_j J_j \\ &= (J_j^{-1} J_{j-1}^{-1} \cdots J_o^{-1}) A_o (J_o \cdots J_{j-1} J_j) \\ &= \left[\begin{array}{c|c} * & * \\ \hline x \quad x \ \ldots \ x & \\ \hline \bigcirc & H \end{array}\right] , \end{aligned}$$

where H is a Hessenberg matrix of order j+2,

$$J_j = \left[\begin{array}{cccc|c|ccc} \multicolumn{4}{c|}{I_{n-j-2}} & \begin{matrix}0\\ \vdots \\ 0\end{matrix} & \multicolumn{3}{c}{\bigcirc} \\ \hline \mu_{n-j,1} & \mu_{n-j,2} & \cdots & \mu_{n-j,n-j-2} & 1 & 0 & \cdots & 0 \\ \hline \multicolumn{4}{c|}{\bigcirc} & \begin{matrix}0\\ \vdots \\ 0\end{matrix} & \multicolumn{3}{c}{I_{j+1}} \end{array}\right] ,$$

and

$$\mu_{n-j,i} = \frac{-a^{(j)}_{n-j,i}}{a^{(j)}_{n-j,n-j-1}} .$$

Finally at the (n-2)nd step, if no pivots are zero, we have a matrix in Hessenberg form,

$$
\begin{aligned}
A_{n-2} &= J^{-1}_{n-3}\, A_{n-3}\, J_{n-3} \\
&= (J^{-1}_{n-3}\, J^{-1}_{n-4} \cdots J^{-1}_{o})\, A_o\, (J_o \cdots J_{n-4}\, J_{n-3}) \\
&= \begin{bmatrix}
x & x & x & \ldots & x & x & x \\
x & x & x & \ldots & x & x & x \\
 & x & x & \ldots & x & x & x \\
 & & \cdot & \ldots & \cdot & \cdot & \cdot \\
 & & & & x & x & x \\
 & & & & & x & x
\end{bmatrix} ,
\end{aligned}
$$

where

$$
J_{n-3} = \left[\begin{array}{cc|c}
1 & 0 & \\
\mu_{31} & 1 & \bigcirc \\
\hline
\bigcirc & & I_{n-2}
\end{array}\right]
$$

and

$$
\mu_{31} = \frac{-a^{(n-3)}_{31}}{a^{(n-3)}_{32}} .
$$

Clearly if a pivotal element $a^{(n-i)}_{i,i-1}$ is small in magnitude with respect to other elements in the row, then we can expect excessive roundoff errors to occur. Thus, by searching through the elements $a^{(n-i)}_{ij}$ $(j=1,\ldots,i-1)$ for the largest element, and interchanging columns and corresponding rows, we can pivot a relatively large element into the $(i,i-1)$ position before annihilating row i.

In case we have $a^{(n-i)}_{ij} = 0$ $(j=1,\ldots,i-1)$, then we simply partition the matrix as follows and apply the algorithm to the principal submatrix S in rows one through i-1:

$$
\left[\begin{array}{c|c}
S & \ast \\
\hline
\bigcirc & H
\end{array}\right] ,
$$

where H is a Hessenberg matrix of order n-i+1. If the pivotal element is not zero, but less than some threshold value ϵ , say $2^{-t}\| A \|_E$, where t is the number of bits in the mantissa of the floating-point computer word, then we can replace ϵ by zero and partition the matrix as described above. What we have, then, at the end of the first stage of computation is a matrix of the form

$$\begin{bmatrix} B_1 & & * & \\ & B_2 & & \\ 0 & & \ddots & \\ & & & B_\ell \end{bmatrix}$$

which is similar to A, and where each B_i is in Hessenberg form with nonzero subdiagonal elements.

In stage two of the reduction of A we must annihilate the elements on and above the main diagonal of each of the diagonal blocks B_i except for the elements in the last column, and we must normalize the subdiagonal elements to unity. Hansen[1963] suggests deferring the normalization until after producing the zeros in order to save arithmetic and to increace accuracy.

It is desirable to produce the zero elements in an order that will save some arithmetic [Hansen, 1963]. We have at least two choices of orderings which possess this property. One choice is to annihilate the elements by rows from the top using row operations. In a given row we must annihilate the diagonal element first in order not to destroy zeros already produced in that row. Here, the inverse operations involve only one multiplication and one addition for every non-zero element in the row under consideration.

Another choice is to annihilate the elements by columns, beginning with column one. In this case, the diagonal element of a given column should be the last element annihilated in order to save arithmetic. After these operations, the matrix A is reduced to block triangular form

$$\begin{bmatrix} D_1 & & * \\ & D_2 & \\ O & & \ddots & \\ & & & D_\ell \end{bmatrix}$$

where each diagonal block D_i is of the form

$$D_i = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & y_{r_i}^{(i)} \\ d_{21}^{(i)} & 0 & 0 & \dots & 0 & y_{r_i-1}^{(i)} \\ 0 & d_{32}^{(i)} & 0 & \dots & 0 & y_{r_i-2}^{(i)} \\ . & . & . & \dots & . & . \\ 0 & 0 & 0 & \dots & d_{r_i,r_i-1} & y_1^{(i)} \end{bmatrix}$$

In both choices above it is usually necessary to utilize double-precision accumulation of inner products in order to save accuracy.

The coefficients, $x_j^{(i)}$, of the characteristic polynomial of D_i ,

$$p_i(\lambda) = (-1)^{r_i}[\lambda^{r_i} - x_1^{(i)}\lambda^{r_i-1} - \dots - x_{r_i-2}^{(i)}\lambda^2 - x_{r_i-1}^{(i)}\lambda - x_{r_i}^{(i)}],$$

are thus given by

$$x_1^{(i)} = y_1^{(i)}$$

and

$$x_j^{(i)} = y_j^{(i)} \prod_{k=r_i-j+1}^{r_i-1} d_{k+1,k} \qquad (j=2,\dots,r_i).$$

(These products usually must be computed in double precision to reduce error.) Hence, replacing the $y_j^{(i)}$ by the $x_j^{(i)}$ and replacing the $d_{j+1,j}^{(i)}$ by unity completes the computation of the F_i of the form (1.1).

We should point out that the form (1.1) is not unique for a given matrix. In fact, the form obtained depends upon the order in which the elements below the first subdiagonal are annihilated. For example, if the elements are annihilated by columns, beginning with column one, we might product a Frobenius form with the F_i permuted or with a completely different set of F_i than we would obtain by annihilating the elements by rows, beginning with row n.

Before describing a modification of the Danilewski method which uses residue arithmetic, it is necessary to review the theorems of residue arithmetic. Chapters II, III, and IV summarize these theorems.

CHAPTER II

RESIDUE ARITHMETIC FOR INTEGERS, MATRICES, AND POLYNOMIALS

1. Introduction. In this chapter we review some of the main definitions and theorems on residue arithmetic for integers, matrices, and polynomials. The results on integers can be found in Szabó and Tanaka [1967], and the results on polynomials in most elementary number theory books. For example see Griffin [1954]. The results on matrices can be found in Howell and Gregory [1969a]. The proofs of these theorems can be found in Howell [1969] and Howell and Gregory [1969c]. For this reason, some of the theorems will be stated here without proof.

2. Integers. Since the moduli m and -m generate the same residue classes, and since the case m=1 is not of interest to us, we shall assume in the following discussion that m is an integer greater than one.

(2.1) DEFINITION. Given integers a and b, if $-m/2 < a \leq m/2$ and if $a \equiv b \pmod{m}$, then we write

$$a = |b|_m$$

and say a is a residue of b modulo m.

Notice that Szabó and Tanaka define $|b|_m$ to be the unique number a in the interval [0,m-1]. We find it more convenient, however, to use the "symmetric" residue, a, where $a \in (-m/2, m/2]$. We can show that this residue is unique. Hence we may say that a is the residue of b modulo m. From the definitions of congruence and residue, we can prove the following.

(2.2) THEOREM. [Szabó and Tanaka, 1967, pp. 16-24] Let a and k be integers. Then

(a) $|km|_m = 0$.

(b) $k|a|_m = |ka|_{km}$.

(c) $|a|_m = a$, if and only if $-m/2 < a \leq m/2$.

(d) $|a \pm km|_m = |a|_m$.

(e) $|-a|_m = |m-a|_m$.

(f) If m is a prime, then $|a^m|_m = |a|_m$.

(g) If $m|M$, then $||a|_M|_m = |a|_m$.

Multiplication, addition, and subtraction have the following properties.

(2.3) THEOREM. [Szabó and Tanaka, 1967, pp. 18,19] If a and b are integers, then

(a) $|a \pm b|_m = ||a|_m \pm b|_m = |a \pm |b|_m|_m = ||a|_m \pm |b|_m|_m$

and

(b) $|ab|_m = ||a|_m b|_m = |a|b|_m|_m = ||a|_m |b|_m|_m = |ba|_m$.

Division, on the other hand, must be treated with more care. The basic rules concerning division modulo m are established by the following theorems. We begin by introducing the concept of multiplicative inverse modulo m.

(2.4) DEFINITION. [Szabó and Tanaka, 1967, p. 21] If a and b are integers, and if

(i) $-m/2 < b \leq m/2$

(ii) $|ab|_m = |ba|_m = 1$

then we write

$$b = a^{-1}(m)$$

and say b is a multiplicative inverse of a modulo m.

(2.5) <u>THEOREM</u>. [Szabó and Tanaka, 1967, p. 22] If a is an integer, then $a^{-1}(m)$ exists if and only if

(i) $|a|_m \neq 0$

(ii) $(a,m) = 1$.

(2.6) <u>THEOREM</u>. [Szabó and Tanaka, 1967, p. 22] If $a^{-1}(m)$ exists, it is unique.

The following properties can be proved using the above theorems and definitions.

(2.7) <u>THEOREM</u>. [Szabó and Tanaka, 1967, p. 26] If a and b are integers, then

(a) $(a^{-1}(m))^{-1}(m) = |a|_m$,

(b) $(ab)^{-1}(m) = |a^{-1}(m) \cdot b^{-1}(m)|_m = |b^{-1}(m) \cdot a^{-1}(m)|_m$,

(c) If $m|M$, then $|a^{-1}(M)|_m = a^{-1}(m)$.

If m is a prime then we have an explicit expression for the multiplicative inverse.

(2.8) <u>THEOREM</u>. [Szabó and Tanaka, 1967, p. 24] If a is an integer and m a prime, then if $a^{-1}(m)$ exists, we have

$$a^{-1}(m) = |a^{m-2}|_m.$$

Using the above theorems, we can now perform division to a limited extent.

(2.9) <u>THEOREM</u>. [Szabó and Tanaka, 1967, p. 38] If a and b are integers, and

(i) $a|b$

(ii) $a^{-1}(m)$ exists,

then

$$\left|\frac{b}{a}\right| = |b \cdot a^{-1}(m)|_m.$$

If, additionally, m is a prime, then

$$\left|\frac{b}{a}\right| = |b \cdot a^{m-2}|_m.$$

Furthermore, there is a <u>cancellation law for integers modulo m</u>.

(2.10) <u>THEOREM</u>. [Szabó and Tanaka, 1967, p. 21] If a and b and $k \neq 0$ are integers, and

(i) $|ka|_m = |kb|_m$

(ii) $k^{-1}(m)$ exists

then

$$|a|_m = |b|_m.$$

This leads to the solution of a single linear equation in one unknown.

(2.11) <u>THEOREM</u>. [Szabó and Tanaka, 1967, p. 24] If a and b are integers and $a^{-1}(m)$ exists, then $|ax|_m = |b|_m$ has a solution x which lies in a unique residue class which is given by

$$|x|_m = |b \cdot a^{-1}(m)|_m.$$

If, additionally, m is a prime, then

$$|x|_m = |b \cdot a^{m-2}|_m.$$

The results in the above theorems relating to multiplicative inverses modulo m show that if m is properly chosen, then division by an integer a in ordinary arithmetic is analogous to multiplication by $a^{-1}(m)$ in residue arithmetic.

EXAMPLE. Let

$$\begin{cases} m = 13 \\ a = 7 \\ b = 11. \end{cases}$$

If

$$|ax|_m = |b|_m$$

then

$$\begin{aligned} |x|_m &= |a^{-1}(m)\cdot b|_m \\ &= |7^{-1}(13)\cdot 11|_{13} \\ &= |2 \cdot 11|_{13} \\ &= |22|_{13} \\ &= -4. \end{aligned}$$

3. Matrices. Most of the theorems on integers have an analogue in matrix theory for residue arithmetic. Where there are differences, they will be pointed out. As before we shall assume $m > 1$. All matrices are assumed to have dimensions which are conformable for the operations indicated. We also assume that m is always greater than the order of the matrices.

(3.1) THEOREM. [Howell and Gregory, 1969a, p. 210] Given pxq integral matrices A and B, if

$$a_{ij} = |b_{ij}|_m$$

for all i and j, then we write

$$A = |B|_m,$$

and say A is a residue of B modulo m.

As with integers, this residue modulo m of a matrix A is unique. Thus, we can say that A is the residue of B modulo m.

From the definition of residue modulo m for matrices we can establish the following basic rules for doing matrix arithmetic modulo m.

(3.2) THEOREM. [Howell and Gregory, 1969c, p. 24] Let A and B be integral matrices and k an integer. Then

(a) $|mA|_m = \emptyset$ (the null matrix).

(b) $k|A|_m = |kA|_{km}$.

(c) $|A|_m = A$, if and only if $-m/2 < a_{ij} \leq m/2$ for all i and j.

(Here, we call A a residue matrix or a matrix modulo m.)

(d) $|A \pm mB|_m = |A|_m$.

(e) $|-A|_m = |mI-A|_m$.

(f) If $m|M$, then $||A|_M|_m = |A|_m$.

Addition, subtraction, and multiplication of matrices modulo m have the following properties.

(3.3) THEOREM. [Howell and Gregory, 1969a, p. 211] If A and B are integral matrices, then

(a) $|A \pm B|_m = ||A|_m \pm B|_m = |A \pm |B|_m|_m = ||A|_m \pm |B|_m|_m$

and

(b) $|AB|_m = ||A|_m B|_m = |A\ |B|_m|_m = ||A|_m|B|_m|_m$.

Notice, however, that multiplication modulo m of matrices modulo m is not commutative.

With matrices, we have a cancellation law for scalar multiplication.

(3.4) THEOREM. [Howell and Gragory, 1969a, p. 212] If A and B are pxq integral matrices, and $k \neq 0$ is an integer, and if

(i) $|kA|_m = |kB|_m$

(ii) $k^{-1}(m)$ exists,

then

$$|A|_m = |B|_m .$$

For the following discussion on matrix inverses modulo m, we must restrict ourselves to square matrices. As before, we assume that all matrices are conformable for the operations indicated.

(3.5) DEFINITION. [Howell and Gregory, 1969a, p. 212] If A and B are nxn integral matrices, and if

(i) $|AB|_m = |BA|_m = I_n$

(ii) $|B|_m = B$,

then we write

$$B = A^{-1}(m)$$

and call B a multiplicative inverse modulo m of A.

The question of the uniqueness of the multiplicative inverse modulo m is answered by the following theorem.

(3.6) THEOREM. [Howell and Gregory, 1969a, p. 212] If A is an nxn integral matrix, and if $A^{-1}(m)$ exists, then it is unique.

In order to discuss existence, we first need to introduce the concepts of nonsingularity modulo m and adjoint matrix modulo m.

(3.7) DEFINITION. [Howell and Gregory, 1969a, p. 212] If A is an nxn integral matrix, then A is said to be nonsingular modulo m if and only if both

(i) $|\det A|_m \neq 0$

(ii) $(\det A, m) = 1$.

Otherwise A is called singular modulo m.

We frequently refer to $|\det A|_m$ as the determinant modulo m of A.

(3.8) THEOREM. [Howell and Gregory, 1969a, p. 212] If A is an nxn integral matrix, then

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

EXAMPLE. Let m = 13 and

$$A = \begin{bmatrix} 7 & -15 \\ -1 & 4 \end{bmatrix}.$$

Then A is singular modulo 13 since

$$|\det |A|_{13}|_{13} = 0,$$

even though A is nonsingular over the real number field. By changing m to 11, we have

$$|\det |A|_{11}|_{11} = 2$$

and

$$(2,11) = 1.$$

Thus, A is nonsingular modulo 11.

(3.9) DEFINITION. (i) The determinant modulo m of an rxr submatrix of A is called a minor modulo m of A of order r. (ii) When the submatrix is located symmetrically with respect to the main diagonal of A, we call the corresponding minor modulo m a principal minor modulo m. (iii) If the submatrix is formed by deleting row i and column j, then we define the cofactor modulo m of the element a_{ij} to be

$$A_{ij} = |(-1)^{i+j}M_{ij}|_m,$$

where M_{ij} is the determinant modulo m of that submatrix.

(3.10) DEFINITION. We define the adjoint modulo m of a matrix A to be

$$|A^{adj}|_m = |(A_{ji})|_m,$$

where A_{ij} is the cofactor modulo m of the element a_{ij}.

(3.11) DEFINITION. An integral matrix A is said to have rank $r_m(A)$ if and only if it has at least one nonzero minor modulo m of order r, but has no nonzero minor modulo m of order greater than r.

EXAMPLE. Let

$$A = \begin{bmatrix} 5 & 3 \\ 5 & 10 \end{bmatrix}.$$

For the modulus m = 19, we have

$$|\det A|_{19} = -3$$

and

$$r_{19}(A) = 2.$$

For the modulus m = 7, on the other hand, we have

$$|\det A\ |_7 = 0$$

and

$$r_7(A) = 1.$$

We are now prepared to discuss the existence of the multiplicative inverse modulo m of a matrix A.

(3.12) THEOREM. [Howell and Gregory, 1969a, p. 213] $A^{-1}(m)$ exists if and only if A is nonsingular modulo m. In this case

$$A^{-1}(m) = \left| d^{-1}(m) \; |A^{adj}|_m \right|_m ,$$

where d = det A.

In other words, $r_m(A)$ must be n.

EXAMPLE. Let m = 11 and

$$A = \begin{bmatrix} 7 & -15 & 3 \\ -1 & 4 & -6 \\ 0 & 1 & -2 \end{bmatrix} .$$

Then

$$|d|_{11} = |\det A|_{11}$$

$$=2$$

and

$$|A^{adj}|_{11} = \begin{bmatrix} -2 & -5 & 1 \\ -2 & -3 & -5 \\ -1 & 4 & 2 \end{bmatrix} .$$

Thus, from theorem (3.12),

$$|A^{-1}(m)|_m = \left| 2^{-1}(11) \cdot \begin{bmatrix} -2 & -5 & 1 \\ -2 & -3 & -5 \\ -1 & 4 & 2 \end{bmatrix} \right|_{11}$$

$$= \left| -5 \cdot \begin{bmatrix} -2 & -5 & 1 \\ -2 & -3 & -5 \\ -1 & 4 & 2 \end{bmatrix} \right|_{11}$$

$$= \begin{bmatrix} -1 & 3 & -5 \\ -1 & 4 & 3 \\ 5 & 2 & 1 \end{bmatrix} .$$

As a computational check, observe that

$$|A \cdot A^{-1}(11)|_{11} = \left| \begin{bmatrix} 7 & -15 & 3 \\ -1 & 4 & -6 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} -1 & 3 & -5 \\ -1 & 4 & 3 \\ 5 & 2 & 1 \end{bmatrix} \right|_{11}$$

$$= \left| \begin{bmatrix} 23 & -33 & -77 \\ -33 & 1 & 11 \\ -11 & 0 & 1 \end{bmatrix} \right|_{11}$$

$$= I_3 .$$

(3.13) <u>THEOREM</u>. If A and B are nxn integral matrices, then

(a) $(A^{-1}(m))^{-1}(m) = |A|_m$,

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

(c) If $m|M$, then $|A^{-1}(M)|_m = A^{-1}(m)$.

We have two theorems regarding the determinant modulo m of a matrix.

(3.14) <u>THEOREM</u>. $\left|\det \prod_{i=1}^{k} A_i \right|_m = \left| \prod_{i=1}^{k} |\det A_i|_m \right|_m .$

<u>Proof</u>.
$$\left|\det \prod_{i=1}^{k} A_i \right|_m = \left| \prod_{i=1}^{k} \det A_i \right|_m$$

$$= \left| \prod_{i=1}^{k} |\det A_i|_m \right|_m . \qquad ///$$

(3.15) <u>THEOREM</u>. $|\det A^{-1}(m)|_m = [|\det A|_m]^{-1}(m)$.

Proof.

$$\left|\,|\det A^{-1}(m)|_m \cdot |\det A|_m\right|_m = |\det [A^{-1}(m) \cdot A]|_m$$
$$= 1$$
$$= |\det [A \cdot A^{-1}(m)]|_m$$
$$= \left|\,|\det A|_m \cdot |\det A^{-1}(m)|_m\right|_m .$$

By the uniqueness of the multiplicative inverse, we have

$$|\det A^{-1}(m)|_m = [|\det A|_m]^{-1}(m). \qquad ///$$

The following theorem gives us an expression for the solution of a residue system of equations,

$$|AX|_m = |B|_m ,$$

where A, B, and X are integral matrices.

(3.16) THEOREM. If A is an nxn integral matrix which is nonsingular modulo m, B is an nxp integral matrix, and $|AX|_m = |B|_m$, then

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

where X is an nxp integral matrix.

A discussion with examples of the solution of $|Ax|_m = |b|_m$, where x and b are nx1 vectors, can be found in Howell and Gregory [1969a, pp.214, 217-224]. The method of solution is based on the following three types of operations, which are called elementary operations or elementary transformations modulo m:

(a) Interchange of rows i and j (or columns i and j),

(b) Multiplication of row i (or column i) by a nonzero constant, k, where (k,m) = 1, and followed by reduction modulo m,

(c) Addition to the ith row of an arbitrary multiple, k, of row j, followed by reduction modulo m (and the analogous operation on columns).

An elementary matrix modulo m is a matrix obtained from the identity matrix by performing one of the above operations on it. We shall denote matrices of these kinds by $|I_{ij}|_m$, $|E_i(k)|_m$, and $|E_{ij}(k)|_m$, respectively. An elementary operation modulo m can be performed on a general matrix A by multiplying A by an elementary matrix on the appropriate side followed by reduction modulo m.

Clearly, since (k,m) = 1, elementary matrices modulo m are nonsingular modulo m, and their inverses modulo m are as follows:

$$|I_{ij}|_m^{-1}(m) = |I_{ij}|_m ,$$

$$|E_i(k)|_m^{-1}(m) = |E_i(k^{-1}(m))|_m ,$$

and

$$|E_{ij}(k)|_m^{-1}(m) = |E_{ij}(-k)|_m .$$

Thus, their inverses modulo m are also elementary matrices modulo m. If m is a prime, and (k,m) = 1, then we can guarantee that inverses modulo m exist.

(3.17) THEOREM. Performing an elementary operation modulo m on a matrix A does not change the rank $r_m(A)$, provided (k,m)=1.

Proof. [Hohn, 1964, p. 116] We shall assume that the elementary operation modulo m is a row operation, since analogous arguments hold for column operations. If the rank $r_m(A)$ is n, then the theorem follows from theorem (3.14). Thus, we shall let $r_m(A)$ be $r<n$, so that the determinants modulo m of all submatrices of order r+1 are zero. We shall first show that applying these row operations cannot increase $r_m(A)$.

Let a transformation of type (a) be applied to the matrix A, forming a new residue matrix A'. If we consider any submatrix of A' of order r+1, we see that either it is identical to the corresponding submatrix of A, or it is the corresponding submatrix with two rows interchanged, or it is equal to some other submatrix of A of order r+1. In all three cases, the determinant modulo m of this submatrix of A' is zero.

If a transformation of type (b) is performed on A, then any submatrix of order r+1 of A' is either identical to the corresponding submatrix of A, or it is the corresponding submatrix of A with one row multiplied by k, or it is some other submatrix of A with one row multiplied by k. In the first case, the determinant modulo m of the submatrix is zero. In the second and third cases the determinant modulo m is of the form $d = |k \cdot d_1|_m$, where d_1 is the determinant modulo m of some r+1 order submatrix of A. Since $d_1 = 0$, then we have $d = 0$.

For transformations of type (c), we see that any submatrix of order r+1 of A' either is identical to the corresponding submatrix of A, or it has a row which is a sum of two rows of the matrix A. In the first case, its determinant modulo m is zero. In the second case, its determinant modulo m, d, may be written in the form $|d_1 + kd_2|_m$, where d_1 is the determinant modulo m of an r+1 order submatrix of A, and d_2 is either the determinant modulo m of another r+1 order submatrix of A, or it has two equal rows. Thus, d_1 and d_2 are both zero, and hence d is zero.

From the above result, we see that performing an elementary operation on A does not increase $r_m(A)$. Thus, $r_m(A') \leq r$. If it were less than r, then performing the inverse transformation would raise the rank, contradicting what we have just shown above. Therefore, $r_m(A') = r$.

///

4. Polynomials. We shall assume here that all polynomials have integral coefficients, that $m > 1$, and that m is greater than the degree of the

polynomials.

(4.1) <u>DEFINITION</u>. Given integral polynomials

$$p_1(x) = a_o + a_1x + a_2x^2 + \ldots + a_{n-1}x^{n-1} + a_nx^n$$

and

$$p_2(x) = b_o + b_1x + b_2x^2 + \ldots + b_{k-1}x^{k-1} + b_kx^k$$

where $|b_k|_m \neq 0$ and $n \geqq k$, such that

(i) $b_i = |a_i|_m$ $(i = 0, \ldots, k)$

(ii) $|a_i|_m = 0$ $(i = k+1, \ldots, n)$

then we write

$$p_2(x) = |p_1(x)|_m$$

and say $p_2(x)$ is a residue of $p_1(x)$ molulo m.

(4.1a) <u>DEFINITION</u>. A polynomial modulo m

$$p(x) = a_o + a_1x + \ldots + a_{n-1}x^{n-1} + a_nx^n$$

is said to be of <u>degree k</u> if $|a_k|_m \neq 0$ and $|a_{k+1}|_m = \ldots = |a_n|_m = 0$. We refer to a_k as the <u>leading nonzero coefficient modulo m</u> of p(x).

<u>EXAMPLE</u>. Let

$$p(x) = 22x^4 + 15x^3 + 8x^2 - 17x - 14.$$

Then

$$|p(x)|_{11} = 4x^3 - 3x^2 + 5x - 3,$$

and p(x) is of degree 3 modulo 11.

This residue can be shown to be unique. We say that a polynomial p(x) is a monic polynomial modulo m if a_k , its leading nonzero coefficient modulo m, is congruent to 1 modulo m [Griffin, 1954, p. 187], and the polynomial whose coefficients are all congruent to zero is called a null (or zero) polynomial modulo m.

The following theorem establishes some of the basic rules for doing arithmetic with polynomials modulo m.

(4.2) THEOREM. Let $p_1(x)$ and $p_2(x)$ be integral polynomials and K an integer. Then

(a) $|m \cdot p_1(x)|_m = 0$ (the null polynomial modulo m).

(b) $K \cdot |p_1(x)|_m = |K \cdot p_1(x)|_{Km}$.

(c) $|p_1(x)|_m = p_1(x)$ if and only if $-m/2 < a_i \leqq m/2$ $(i=1,\ldots,n)$. (Here we call $p_1(x)$ a residue polynomial or a polynomial modulo m.)

(d) $|p_1(x) \pm m \cdot p_2(x)|_m = |p_1(x)|_m$.

Addition, subtraction, and multiplication of polynomials modulo m have the following properties.

(4.3) THEOREM. If $p_1(x)$ and $p_2(x)$ are integral polynomials, then

(a) $$|p_1(x) \pm p_2(x)|_m = \left||p_1(x)|_m \pm p_2(x)\right|_m = \left|p_1(x) \pm |p_2(x)|_m\right|_m$$
$$= \left||p_1(x)|_m \pm |p_2(x)|_m\right|_m$$

and

(b) $$|p_1(x)\ p_2(x)|_m = \left||p_1(x)|_m\ p_2(x)\right|_m = \left|p_1(x)\ |p_2(x)|_m\right|_m$$
$$= \left||p_1(x)|_m\ |p_2(x)|_m\right|_m = |p_2(x)\ p_1(x)|_m .$$

Thus, addition and multiplication of polynomial modulo m are commutative.

Division of polynomials modulo m is an operation which causes more difficulty. We have first the cancellation law for scalar multiplication of polynomials modulo m.

(4.4) THEOREM. If $p_1(x)$ and $p_2(x)$ are integral polynomials, and $K \neq 0$ is an integer, and if

(i) $|K \cdot p_1(x)|_m = |K \cdot p_2(x)|_m$

(ii) $K^{-1}(m)$ exists

then

$$|p_1(x)|_m = |p_2(x)|_m .$$

(4.5) DEFINITION. [Griffin, 1954, p. 72] If $p_1(x)$ and $p_2(x)$ are integral polynomials, and if it is possible to find integral polynomials $q(x)$ and $r(x)$ such that

$$|p_1(x)|_m = |p_2(x) \cdot q(x) + r(x)|_m ,$$

where either the degree of $r(x)$ is less than the degree of $p_2(x)$ or $|r(x)|_m = 0$, then we call $q(x)$ the quotient modulo m and $r(x)$ the remainder modulo m in the division modulo m of $p_1(x)$ by $p_2(x)$.

When $|r(x)|_m = 0$, we say that $p_2(x)$ is a divisor modulo m or a factor modulo m of $p_1(x)$, and that $p_1(x)$ is a multiple modulo m of $p_2(x)$ [Griffin, 1954, p. 72]. Then we write

$$|q(x)|_m = \left| \frac{p_1(x)}{p_2(x)} \right|_m .$$

In order to determine whether or not such a $q(x)$ and $r(x)$ can be found, we need to examine the coefficients of $p_1(x)$ and $p_2(x)$. If

$$p_1(x) = a_o + a_1x + \ldots + a_{n-1}x^{n-1} + a_nx^n$$

and

$$p_2(x) = b_o + b_1x + \ldots + b_{k-1}x^{k-1} + b_kx^k \quad ,$$

where $|b_k|_m \neq 0$, and where $n \geqq k$, then a sufficient condition for the existence of $q(x)$ and $r(x)$ is that $b_k^{-1}(m)$ exists.

(4.6) THEOREM. [Griffin, 1954, p. 186] If m is a prime, then we can compute

$$|q(x)|_m = \left| \frac{p_1(x) - r(x)}{\overline{\overline{p_2(x)}}} \right|_m$$

provided $p_2(x)$ is not the zero polynomial modulo m.

EXAMPLE. Let

$$p_1(x) = 4x^4 - 6x^3 + 5x^2 - 2x - 6$$

and

$$p_2(x) = 3x^2 - 6x + 5$$

and

$$m = 13.$$

We shall compute q(x) and r(x) using long division, reducing all results at intermediate steps modulo m:

$$
\begin{array}{rrrrrr}
 & & & -3x^2 & +5x & -5 \\ \cline{2-6}
3x^2 - 6x + 5 \,\big) & 4x^4 & -6x^3 & +5x^2 & -2x & -6 \\
 & 4x^4 & +5x^3 & -2x^2 & & \\ \cline{2-4}
 & & 2x^3 & -6x^2 & -2x & \\
 & & 2x^3 & -4x^2 & -x & \\ \cline{3-5}
 & & & -2x^2 & -x & -6 \\
 & & & -2x^2 & +4x & +1 \\ \cline{4-6}
 & & & & -5x & +6
\end{array}
$$

Thus,

$$q(x) = -3x^2 + 5x - 5$$

and

$$r(x) = -5x + 6.$$

We note that the leading coefficient of q(x) is obtained by solving the residue equation

$$|3 \cdot y|_m = 4.$$

It is easily verified that

$$|p_1(x)|_m = |p_2(x) \cdot q(x) + r(x)|_m .$$

(4.7) DEFINITION. [Griffin, 1954, pp. 179, 187] A residue polynomial d(x) which divides two or more polynomials modulo m, not all of which are congruent to zero modulo m, is called a common divisor modulo m. If d(x) is divisible modulo m by every other common divisor modulo m, and if it is monic, then it is called a greatest common divisor modulo m.

(4.8) THEOREM. The greatest common divisor modulo m of two or more polynomials is unique, provided m is a prime.

Proof. [Griffin, 1954, p. 181] Let d(x) and d'(x) be two polynomials which are both greatest common divisors modulo m for f(x) and g(x). Then d(x) divides d'(x) modulo m, and so

$$|d'(x)|_m = |d(x)q(x)|_m .$$

Also, d'(x) divides d(x) modulo m. Hence,

$$|d(x)|_m = |d'(x)p(x)|_m.$$

Thus,

$$|d(x)q(x)p(x)|_m = |d'(x)p(x)|_m$$

$$= |d(x)|_m .$$

Then, $|q(x)p(x)|_m = 1$, and both $q(x)$ and $p(x)$ are integers. Since $d(x)$ and $d'(x)$ are both monic polynomials modulo m, then $q(x)$ and $p(x) = 1$ and

$$|d'(x)|_m = |d(x)|_m .$$

Since $d'(x)$ and $d(x)$ are residue polynomials, we have

$$d'(x) = d(x).$$

///

Polynomials which have no common divisors modulo m other than the trivial ones, the integers ± 1 , are of special interest to us.

(4.9) DEFINITION. [Eames, 1967, p. 115] Two or more polynomials are said to be relatively prime modulo m when their greatest common divisor modulo m is one.

EXAMPLE. Let m = 13,

$$p_1(x) = x^2 + 20x - 18 ,$$

and

$$p_2(x) = x^2 - 6x + 6 .$$

Since we can write

$$|p_1(x)|_{13} = |(x-2)(x-4)|_{13}$$

and

$$|p_2(x)|_{13} = |(x+1)(x+6)|_{13} ,$$

then $p_1(x)$ and $p_2(x)$ have no common divisors modulo m other than ± 1. Hence $p_1(x)$ and $p_2(x)$ are relatively prime modulo m.

If we change $p_1(x)$ to

$$p_1(x) = x^2 + 16x - 11$$

then we can write

$$|p_1(x)|_{13} = |(x+1)(x+2)|_{13} .$$

Thus $p_1(x)$ and $p_2(x)$ have $(x+1)$ as a common divisor modulo 13, and hence $p_1(x)$ and $p_2(x)$ are not relatively prime modulo 13.

The next theorem gives us an expression for the greatest common divisor modulo m of two polynomials.

(4.10) THEOREM. If $p_1(x)$ and $p_2(x)$ are polynomials modulo m neither of which is the zero polynomial modulo m, and if m is a prime, then d(x), the greatest common divisor modulo m of $p_1(x)$ and $p_2(x)$ exists. Furthermore, we can find polynomials r(x) and s(x) such that

$$|d(x)|_m = |r(x)p_1(x) + s(x)p_2(x)|_m .$$

Proof. See [Griffin, 1954, pp. 187-188].

An algorithm for computing d(x), r(x), and s(x) is given in Griffin [1954, p. 187-188].

EXAMPLE. Let m = 13 and

$$p_1(x) = x^3 - 6x^2 - 6x - 2$$

and

$$p_2(x) = x^2 + 14x + 11.$$

Then simple computation shows that the greatest common divisor modulo m of $p_1(x)$ and $p_2(x)$ is x-1, and that

$$x\text{-}1 = |-4(x^3 - 6x^2 - 6x - 2) + (4x - 2)(x^2 + 14x + 11)|_{13}$$

$$= |-4\cdot p_1(x) + (4x - 2)\cdot p_2(x)|_{13} .$$

Thus, in theorem (4.10)

$$d(x) = x\text{-}1,$$

$$r(x) = -4$$

and

$$s(x) = 4x\text{-}2.$$

(4.11) COROLLARY. [Eames, 1968, p. 116] If m is a prime, two polynomials $p_1(x)$ and $p_2(x)$ are relatively prime modulo m if and only if there exist polynomials r(x) and s(x) such that

$$|r(x)p_1(x) + s(x)p_2(x)|_m = 1.$$

(4.12) THEOREM. If m is a prime and if two polynomials $p_1(x)$ and $p_2(x)$ are relatively prime modulo m and $q_i(x)$ divides $p_i(x)$ modulo m (i=1,2), then $q_1(x)$ and $q_2(x)$ are relatively prime modulo m.

Proof. From corollary (4.11) we can find polynomials r(x) and s(x) such that

$$|r(x)p_1(x) + s(x)p_2(x)|_m = 1$$

Since $q_i(x)$ divides $p_i(x)$ modulo m (i=1,2), then

$$|p_i(x)|_m = |q_i(x)h_i(x)|_m , \qquad (i=1,2).$$

Thus,

$$1 = |r(x)q_1(x)h_1(x) + s(x)q_2(x)h_2(x)|_m$$

$$= |f_1(x)q_1(x) + f_2(x)q_2(x)|_m ,$$

and we have that $q_1(x)$ and $q_2(x)$ are relatively prime modulo m.

///

In the discussion following definition (4.5) we introduced the concept of multiple modulo m of a polynomial. Any two nonzero polynomials have a common multiple modulo m and a least common multiple modulo m. These two concepts play an important role in chapter IV.

(4.13) DEFINITION. If $p_1(x)$ divides $p(x)$ modulo m and $p_2(x)$ divides $p(x)$ modulo m then $p(x)$ is a common multiple modulo m of $p_1(x)$ and $p_2(x)$.

(4.14) DEFINITION. A least common multiple modulo m of two or more polynomials is a monic common multiple modulo m that is a divisor modulo m of every common multiple modulo m of the given polynomials.

EXAMPLE. Let $m = 13$,

$$p_1(x) = x^2 + 16x - 11,$$

and

$$p_2(x) = x^2 - 6x + 6.$$

Then

$$|p_1(x)|_{13} = |(x + 1)(x + 2)|_{13}$$

and

$$|p_2(x)|_{13} = |(x + 1)(x + 6)|_{13} .$$

Thus, a common multiple modulo m of $p_1(x)$ and $p_2(x)$ is

$$|(x+1)(x+2)(x+1)(x+6)|_{13} = |x^4 - 3x^3 + 3x^2 + 6x - 1|_{13} .$$

The least common multiple modulo m is

$$|(x+1)(x+2)(x+6)|_{13} = |x^3 - 4x^2 - 6x - 1|_{13} .$$

The following definition and two lemmas will be used in the proof of theorem (4.18).

(4.15) DEFINITION. [Niven and Zuckermen, 1966, p. 194] A polynomial $p(x)$ which is not the zero polynomial modulo m is irreducible modulo m if there is no factoring $p(x) = |g(x)h(x)|_m$ of $p(x)$ into two polynomials $g(x)$ and $h(x)$ of positive degree.

(4.16) LEMMA. If m is a prime, $p_1(x)$ and $p_2(x)$ are polynomials which are relatively prime modulo m, and if $p_1(x)$ divides $|p_2(x)p_3(x)|_m$ modulo m, then $p_1(x)$ divides $p_3(x)$ modulo m.

Proof. See [Nagell, 1964, p. 96].

(4.17) LEMMA. Every polynomial $p(x)$ of degree n can be written in the form

$$|p(x)|_m = |c \cdot p_1(x)^{e_1} p_2(x)^{e_2} \dots p_r(x)^{e_r}|_m$$

where m is a prime, the $p_i(x)$ are distinct, monic, irreducible polynomials modulo m, c is an integer, $e_1+\dots+e_r = n$, and the form is unique apart from the order of the $p_i(x)$.

Proof. See [Nagell, 1964, p. 97].

EXAMPLE. Let $p(x) = 3x^4 + 12x^3 + 4x^2 - 19x - 6$ and $m = 13$. Then we can write

$$|p(x)|_{13} = |3x^4 + 12x^3 + 4x^2 - 19x - 6|_{13}$$
$$= |3(x^2 + 6)(x + 2)^2|_{13} .$$

Thus, in lemma (4.17)

$$\begin{cases} c = 3, & e_1 = 1, \\ p_1(x) = x^2 + 6, & e_2 = 2, \\ p_2(x) = x + 2, & r = 2. \end{cases}$$

(4.18) THEOREM. [Perlis, 1952, p. 118] If m is a prime, c(x) and d(x) are relatively prime modulo m, and g(x) is a monic polynomial which divides $|c(x)\ d(x)|_m$ modulo m, then there exist unique monic polynomials modulo m, c'(x) and d'(x), such that

$$|g(x)|_m = |c'(x)d'(x)|_m$$

where c'(x) divides c(x) and d'(x) divides d(x) modulo m.

Proof. Let the degrees of c(x) and d(x) be n_1 and n_2 , respectively. Then, since g(x) divides c(x)d(x) modulo m, we have

$$\begin{aligned}|g(x)q(x)|_m &= |c(x)d(x)|_m \\ &= |p_1c_1(x)^{e_1} \ldots c_r(x)^{e_r}p_2d_1(x)^{f_1} \ldots d_s(x)^{f_s}|_m ,\end{aligned}$$

where the $c_i(x)$ and $d_i(x)$ are the unique factors of c(x) and d(x) which are monic and irreducible modulo m, $e_1 + \ldots + e_r = n_1$, and $f_1 + \ldots + f_s = n_2$. Thus, because of the uniqueness of these factors, g(x) must have the same irreducible factors, though possibly to different powers. Hence,

$$|g(x)|_m = |c_1(x)^{j_1} \ldots c_r(x)^{j_r}d_1(x)^{i_1} \ldots d_s(x)^{i_s}|_m ,$$

where $j_k \leq e_k$ and $i_k \leq f_k$, for all k. Then we can write

$$|g(x)|_m = |c'(x)d'(x)|_m ,$$

where

$$|c'(x)|_m = |c_1(x)^{j_1} \ldots c_r(x)^{j_r}|_m$$

and

$$|d'(x)|_m = |d_1(x)^{i_1} \ldots d_s(x)^{i_s}|_m$$

and c'(x) divides c(x) and d'(x) divides d(x) modulo m.

We shall now prove the uniqueness of c'(x) and d'(x). Suppose

$$|c_1(x)^{j_1} \ldots c_r(x)^{j_r} d_1(x)^{i_1} \ldots d_s(x)^{i_s}|_m = |g(x)|_m$$
$$= |c_1(x)^{k_1} \ldots c_r(x)^{k_r} d_1(x)^{\ell_1} \ldots d_s(x)^{\ell_s}|_m .$$

Then $c_1(x)^{k_1}$ divides $|c_1(x)^{j_1} \ldots c_r(x)^{j_r} d_1(x)^{i_1} \ldots d_s(x)^{i_s}|_m$ modulo m. By lemma (4.16) $c_1(x)^{k_1}$ divides $c_1(x)^{j_1}$, and so

$$|c_1(x)^{k_1} c_1(x)^h|_m = |c_1(x)^{j_1}|_m .$$

By a similar argument

$$|c_1(x)^{j_1} c_1(x)^t|_m = |c_1(x)^{k_1}|_m .$$

Thus,

$$|c_1(x)^{k_1} c_1(x)^h c_1(x)^t|_m = |c_1(x)^{j_1} c_1(x)^t|_m$$
$$= |c_1(x)^{k_1}|_m .$$

This implies that

$$|c_1(x)^h c_1(x)^t|_m = 1.$$

Hence, $c_1(x)^h$ and $c_1(x)^t$ must be constants. But since the $c_i(x)$ are monic, we must have

$$|c_1(x)^h|_m = |c_1(x)^t|_m = 1.$$

Therefore,

$$|c_1(x)^{k_1}|_m = |c_1(x)^{j_1}|_m ,$$

and

$$|c_2(x)^{j_2} \ldots c_r(x)^{j_r} d_1(x)^{i_1} \ldots d_s(x)^{i_s}|_m = |c_2(x)^{k_2} \ldots c_r(x)^{k_r} d_1(x)^{\ell_1} \ldots d_s(x)^{\ell_s}|_m .$$

Similar arguments show that

$$|c_i(x)^{k_i}|_m = |c_i(x)^{j_i}|_m \ , \quad (i = 2,\ldots,r)$$

and

$$|d_j(x)^{\ell_j}|_m = |d_j(x)^{i_j}|_m \ , \quad (j = 1,\ldots,s).$$

Uniqueness of c'(x) and d'(x) is thus established. ///