Chapter 8: Matrix Representation of a Galois Field

8.1 A review of some matrix facts.

row index, column index, transpose, tensors, the antisymmetric tensor, determinant, cofactors, inverse, singular, trace, rank, nullity, degeneracy, characteristic polynomial, eigenvalues, eigenvectors, quantum mechanics, triangular, diagonal, similarity transformation, similar matrices, diagonalization, various theorems.

8.1 Polynomials of a Matrix

Fact: One can extend the meaning of any polynomial f(x) from (x = element of field F) to (x = square matrix whose matrix elements are elements of field F).

Big Fact: Every matrix is a root of its own characteristic polynomial. (Cayley-Hamilton)

8.3 The Remainder Theorem

(1) in the field of complex numbers; (2) in a Galois field; (3) in a matrix field

Fact: If A is an mxm matrix, then a polynomial f(A) are arbitrarily high degree is equal to another polynomial r(A) of degree < m, and we write f(A) = r(A). Moreoever, given f(A), and given A with m distinct eigenvalues, we have a method for computing the coefficients of r(A).

8.4 The Galois Field Connection

Definition: A *matrix representation* of a Galois Field.

Fact 1: If the characteristic polynomial d(x) of the mxm matrix A is irreducible over GF(p), then the set of p^m matrices $r(A) = r_0 I + r_1 A + r_2 A^2 + ... + r_{m-1} A^{m-1} = [r_0, r_1, r_2, ... r_{m-1}]$, forms a *matrix representation* of the Galois Field $GF(p^m)$.

Fact 2: If the characteristic polynomial d(x) of the mxm matrix A is a *primitive polynomial* of $GF(p^m)$, then matrix A is a generator of $\{GF(p^m) - 0, \bullet\}$, and we can enumerate all non-zero elements of $GF(p^m)$ as powers of A. Then all of $GF(p^m)$ can be enumerated as: $\{ 0, I, A, A^2, A^3, ... A^{q-2}\}$, where $A^{q-1} = I$.

Corollary 2: We can build a "table" for $GF(p^m)$ exactly as was done in Chapter 6. One replaces the symbol α with the matrix A everywhere.

Example for the Skeptical Reader: Consider the case $GF(2^9)$.

Fact: If a matrix A "works" as the primitive element of a matrix representation of $GF(p^m)$, then any *similar* matrix $A' = S^{-1} A S$ works just as well.

Fact 3: If matrices A and B have the same characteristic polynomial d(x), then if A is a primitive element of $GF(p^m)$, so is B. In this case, all Galois statements above concerning A also apply to B. We can regard either A or B as the matrix whose powers generate $GF(p^m)$.

Corollary 3 : If A and B are similar matrices , then the conclusions of Fact 3 follow.

8.5 The Companion Matrix

Definition: Companion matrix C, see picture.

Fact 1: The characteristic polynomial of matrix C is precisely f(x).

Fact 2: The characteristic polynomial d(x) = det(xI - A) remains unaltered if the matrix A is reflected in either or both diagonals.

Fact 3: Based on Fact 2, we can now write down three alternative forms of the companion matrix. All four forms, including C above, have the same characteristic polynomial:

Big Corollary: Let p(x) be a primitive polynomial of GF(p^m). Then the companion matrix C (or any of its alternative forms) with f(x) = p(x) can serve as the matrix A which represents a primitive element of a matrix representation of GF(p^m).

Example: Here we explicitly construct all matrices for GF(2^3),

Sample calculation: Compute A^{973}.

Chapter 8: Matrix Representation of a Galois Field

In this Chapter it will be shown that one can construct an explicit matrix representation of any Galois Field GF($q=p^m$) as a set of q mxm matrices, whose matrix elements are elements of GF(p). A certain matrix called the companion matrix is constructed using a primitive polynomial of GF(q), and this matrix can then be identified with a primitive element α of GF(q). Thus, all powers of the companion matrix become the non-zero field elements. The zeroth power is the identity matrix, and the zero element is the all-zero matrix.

The notion of a matrix representation of a *group* is extremely common in many scientific disciplines, for both finite and infinite groups, examples of the latter being the rotation group or the Lorentz group.

I have never before heard of the idea of a matrix representation of a *field*, so in that sense, the contents of this chapter may be regarded as "original work". It is my impression, possibly wrong, that the subject of Galois Fields has a somewhat limited literature, perhaps because there is a limited range of applications, one being error-correcting codes. This impression is likely due to the small amount of time I have been able to spend searching for information in this field. Surely the idea of a matrix representation of a Galois Field is not new, it's just that I don't know of any source that discusses this subject.

References: Motivation for this section was provided by Section 7.6 of a 1972 second edition of Peterson and Weldon, *Error-Correcting Codes*. This section discusses a matrix approach to state machines, but does not make the connection to Galois fields, although such fields are also discussed in the book. Since the first edition of this book is dated 1961, the references given for this section are all prior to that time, which is to say, they are over 30 years old. At Marriott library at the University of Utah, the Albert reference is so old that the book is in off-site storage. The MacDuffie 1943 reference (*Vectors and Matrices*) is in the Dewey section.

Most matrix facts below came from a book by Bronson, *Matrix methods...,* AP, 1969. Several such books can be found in the vicinity of QA267 at Marriott.

8.1 A review of some matrix facts.

This section is meant purely as a review of terminology. The reader must consult any book on matrices to find proofs of things unproven here. To add familiarity with the terms, some "peculiar facts" and "observations" are added ; they will not be used in following sections.

All the following facts deal with square matrices. We use an mxm matrix A as our prototype. It is of course assumed that the reader has a basic knowledge of matrices, such as how to multiply them.

It is assumed in this section that the elements of the matrix A are just "numbers", meaning complex numbers. More generally, the matrix elements are elements of some field F.

It is amazing how the theory of matrices seems to infiltrate every field of human interest in which there is some attempt to calculate something. There are probably several thousand "facts" about matrices which could be listed in an exhaustive treatise, which this small review is not.

Convention: For the definitions which follow, we shall assume that the "first" row or column or a matrix is labelled by 1. Sometimes in later sections we may use 0 instead. No facts are rendered untrue by such a change of labelling.

Definitions: If A is a matrix, and its elements are A_{ij} , the first index i is the *row index*, and the second is the *column index*. Thus, the first row of this matrix has elements A_{11}, A_{12}, A_{13}, ... The row index is constant across a row, and serves to define a row. Similarly for the column index.

Definition: The *transpose* of a matrix A is called A^T, and is a new matrix formed by interchanging the rows and columns of A. We have then $(A^T)_{ij} = A_{ji}$.

Definition: A *tensor of rank m* is a matrix with m subscripts instead of 2. For example, A_{ijk} is a rank-3 tensor. A matrix is a rank-2 tensor, a vector is a rank-1 tensor.

Definition: The *completely antisymmetric tensor* $\varepsilon_{abcde...}$ of rank m is defined as follows. If the subscripts are an even permutation of 123...m, it is (+1). If the subscripts are an odd permutation, it is (-1). If the subscripts are not a permutation, this means some subscripts are repeats, and in this case it is (0).

Example : The rank 6 totally antisymmetric tensor, some sample elements:

$$\varepsilon_{123456} = 1 \qquad \varepsilon_{132456} = -1 \qquad \varepsilon_{132465} = 1 \qquad \varepsilon_{123345} = 0$$

Definition: The *determinant* of an mxm matrix A is indicated by det(A) or |A| and is defined as:

$$\det(A) = \varepsilon_{abcde...} A_{1a}A_{2b}A_{3c}A_{4d} \ldots \quad \{ = \varepsilon_{abcde...} A_{a1}A_{b2}A_{c3}A_{d4} \ldots \}$$

where ε is the totally antisymmetric tensor of rank m defined above. In the above formula, each repeated index is implicitly summed from 1 to m. There are a total of m such indices, hence there are a total of m factors of A.

This form of the determinant is convenient for making some points below. It is assumed that the reader is familiar with the more normal "recursive" method of computing a determinant in terms of sub-determinants by starting with any row or column.

Example: Determinant of a 2x2 matrix:

$$\det(A) = \varepsilon_{ab}A_{1a}A_{2b} = \varepsilon_{12}A_{11}A_{22} + \varepsilon_{21}A_{12}A_{21} = A_{11}A_{22} - A_{12}A_{21}$$

Fact 1: Adding a multiple of one row to another row does not change the determinant of a matrix, although it certainly changes the matrix itself. The same applies for columns.

Fact 2: $\det(A) = \det(A^T)$

Fact 3: If A is an mxm matrix, and α is a constant, then $\det(\alpha A) = \alpha^m \det(A)$.

Definition: If A is a square matrix, then the *cofactor* of a matrix element A_{ij} is denoted by $[\text{cof}(A)]_{ij}$ and is equal to $(-1)^{i+j}$ times the determinant of the submatrix obtained by crossing out the row and column which contains A_{ij}. Thus, the cofactor is just some number. Since there is one cofactor for each element of a matrix, we can combine all these cofactors into a new matrix called the *cofactor matrix* cof(A).

Fact 4: The inverse of a square matrix A is given by

$$A^{-1} = [\text{cof}(A)]^T / \det(A)$$

Corollary 4: The inverse A^{-1} exists if and only if $\det(A) \neq 0$.

Definition: If $\det(A) = 0$ and this A^{-1} does not exist, A is said to be *singular*.

Definition: The matrix $[\text{cof}(A)]^T$ is sometimes called the *adjoint* matrix of A and denoted adj(A). That is , the adjoint is the transpose of the cofactor matrix of A. We will not use this adjoint terminology.

Definition: The *trace* of a square matrix A is denoted tr(A) and is defined as the sum of the diagonal elements.

Peculiar Fact 5: If A is a square matrix, then $\det(e^A) = e^{\text{tr}(A)}$.

Definitions: The *rank* of an mxm matrix A is denoted r(A) and is the number of linearly independent rows or columns. It is also the size of the largest non-zero sub-determinant. The size m of a square matrix is called its *order*. The quantity (order - rank) is called the *degeneracy* or the *nullity* of A, denoted by n(A). Thus, n(A) = m - r(A).

The rank of a matrix has nothing to do with the rank of a tensor mentioned above, it happens that the same word is used for both concepts.

Corollary 5: If an mxm matrix A is invertible, it has full rank m.

Proof: A invertible means $\det(A) \neq 0$, which means rank = m.

Peculiar Fact 6: Consider the equation AB=C for three square matrices. It turns out that $n(c) \geq n(a), n(b)$ but $n(c) \leq n(a) + n(b)$. This set of inequalities is called Sylvester's Law of Nullity.

Definition: The *characteristic polynomial* d(x) of a square matrix A is

$$d(x) + \det(xI - A).$$

Here, I is the identity matrix.

Fact 7: The characteristic polynomial of an mxm matrix A is of degree m. Two of the coefficients of d(x) are known at once: $d_0 = (-1)^m \det(A)$, and $d_m = 1$. From this last, d(x) is a monic polynomial.

Proof: From the earlier expression given for a determinant, it is clear that one term in the determinant is the product of the diagonal elements of the matrix with a (+1) coefficient. If the matrix is (xI-A), then this term is of the form $(x - A_{11}) (x-A_{22})....(x - A_{mm}) = x^m + ...$. This shows that the degree is m and the leading coefficient is 1. As for the constant term, $d(0) = d_0 = \det(0I-A) = \det(-A) = (-1)^m \det(A)$.

Definition: The m roots of the characteristic polynomial d(x) are called *eigenvalues* and are usually denoted as λ_i. Thus, $d(\lambda_i) = 0$ for any eigenvalue λ_i.

Definition: The equation d(x) = 0 is sometimes called the *secular equation*. Thus, the eigenvalues of the matrix A are the solutions of its secular equation.

Definition: "eigenvectors" : The significance of eigenvalues is that one often encounters problems in which one is solving for a vector v in the equation

$$Av = \lambda v$$

where λv means λIv . Writing this as $(\lambda I - A) v = 0$, one sees that if $(\lambda I - A)$ *has* an inverse, one has only the trivial solution v = 0, since then $v = (\lambda I - A)^{-1} 0 = 0$. However, if $\lambda = \lambda_i$, a root of $d(\lambda)$, then

$d(\lambda_i) = 0 = \det(\lambda I - A)$ $(\lambda I - A)$ has no inverse, see previous section

In this case, there might be (and usually is) some non-zero solution v_i . If this is the case, then v_i is called the *eigenvector* corresponding to the *eigenvalue* λ_i.

Example: In quantum mechanics, the matrix A is called the Hamiltonian H, the eigenvectors v are called stationary states ψ, the eigenvalues λ_i are the energies E_i of these states, and $Av = \lambda v$ is written $H\psi = E\psi$ and is called the Schrodinger equation.

Fact 8: The trace of a matrix tr(A) is equal to the sum of its eigenvalues.

Fact 9: The determinant of a matrix det(A) is equal to the product of its eigenvalues.

Definition: A *triangular* matrix is one which has nothing but zeros on one side of the diagonal.

Fact 10: The eigenvalues of a triangular matrix are its diagonal elements.

Proof: Easy to show this by the usual method of recursively computing a determinant.

Corollary 10: The eigenvalues of a diagonal matrix are its diagonal elements.

Example: Let A = I, the identity matrix. Then $d(x) = \det(xI - I) = \det[(x-1)I] = (x-1)^m \det(I) = (x-1)^m$. There are m roots of d(x) , they are all 1. This matrix has m eigenvalues equal to 1. These are its diagonal elements.

Fact 11: If any eigenvalue of an mxm matrix A is 0, then A is singular.

Proof: A zero eigenvalue means $0 = d(0) = \det(0I - A) = (-1)^m \det(A)$ $\det(A) = 0$ A singular.

Definition: A *similarity transformation* of a square matrix A is defined by

$$A' = S\,A\,S^{-1}$$

where S is any square matrix such that S^{-1} exists, ie, $\det(S) \neq 0$. One says that A and A' are *similar*.

Observation: Saying that A and A' are similar is like saying that a rotated set of axes (x',y',z') is similar to an unrotated set (x,y,z). Generally speaking, a similarity transformation does not change tha nature of what is going on, it just changes some specific internal details. It is a change of basis. The following fact should help impress upon the reader how little a similarity transformation changes things globally.

Fact 12: Similar matrices have the same determinant, the same trace, the same rank, the same characteristic polynomial, and the same eigenvalues. In addition, any equation involving matrices retains its same form after application of a similarity transformation.

Example: If AB=C, then apply S .. S^{-1} to get $(S\,A\,S^{-1})(S\,B\,S^{-1}) = (S\,C\,S^{-1})$ or $A'B' = C'$

Observation: There are certain interesting classes of matrices (Hermitian, real symmetric) which can be brought to diagonal form by a similarity transformation (unitary, real orthogonal). If one can find this similarity transformation, then one at once knows all the eigenvalues of the original matrix, since they are just the diagonal elements of the transformed matrix, and since eigenvalues are preserved under any similarity transformation. The process of bringing a matrix to diagonal form is called *diagonalization*.

8.1 Polynomials of a Matrix

Consider some polynomial f(x) of degree m with m+1 coefficients labelled in the standard manner,

$$f(x) = f_0 + f_1 x + f_2 x^2 \ldots + f_m x^m.$$

Normally, one assumes that the coefficients f_i and the variable x are elements of some field F.

Fact 1: One can extend the meaning of any polynomial f(x) from (x = element of F) to (x = square matrix whose matrix elements are elements of F).

Proof: Start with some f(x), a polynomial. Replace each occurrence of x^i with X^i where X is a square matrix. The coefficients are still elements (scalars) of the original field, but now each term is a scalar times a power of a matrix, which is of course just a matrix. The constant term f_0 is interpreted as the constant f_0 times the identity matrix. If you add up all the terms, the result is a square matrix. The results is:

$$f(X) = \text{some matrix } M$$

Comment: This is reminiscent of our Galois extension of a polynomial from a ground field GF(p) to an extension field GF(p^m). In the present case, we have so far elevated from some starting number field to a ring of matrices. We would like to work with some subset of matrices which in fact form a representation of the elements of the field GF(p^m), and then the analogy will be complete.

Big Fact 2: Let A be any square matrix. Let d(x) be the characteristic polynomial of A, as defined in our previous section. Then d(A) = 0. This is the Cayley-Hamilton Theorem. It says that every matrix is a root of its own characteristic polynomial. The result is non-obvious!

Comment on Proof: The claim is that you first compute d(x) from d(x) = det(xI - A). Then you replace x everywhere it appears in this polynomial d(x) with the matrix A. The amazing result is that you then get d(A) = 0. Note the following fact:

$$d(X) \neq \det(XI - A)$$

If this were an equality, then d(A) = 0 would be a trivial result. The above cannot be an equality because the thing on the right is a number (the determinant of the matrix X-A) while the thing on the left is a matrix. For a proof of the Cayley-Hamilton Theorem, see any respectable matrix text.

Corollary 2: For any mxm matrix A, there are coefficients d_i such that:

$$d_0 I + d_1 A + d_2 A^2 + \ldots + d_{m-1} A^{m-1} + A^m = 0$$

These coefficients are just the coefficients of the characteristic polynomial. We have simply written out d(A) = 0. Recall that $d_0 = (-1)^m \det(A)$.

8.3 The Remainder Theorem

We shall look at the remainder theorem in three different ways.

(1) Let d(x) be the characteristic polynomial of an mxm matrix A. Let f(x) be an arbitrary polynomial, expand it over d(x) according the polynomial remainder theorem of Chapter 3,

$$f(x) = q(x)\, d(x) + r(x).$$

As usual, the remainder polynomial r(x) has degree < m. If we evaluate the above equation at the eigenvalues of A, the λ_i , we know that $d(\lambda_i) = 0$, so we get

$$f(\lambda_i) = r(\lambda_i) \qquad i = 1,2....m$$

If we are given f(x) and we want to compute r(x), the above gives us m equations in m unknowns (ie, the coefficients) which let us compute the m coefficients of r(x). This assumes that all m eigenvalues are different, which will be the case in our application below. If some of the eigenvalues are the same, there is a method for still getting m equations in m unknowns. It involves derivatives of f(x), and we refer to the reader to Bronson, *Matrix Methods*.... .

In this subsection, we are assuming that the field over which things are defined is the complex numbers. In this case, d(x) of degree m is guaranteed to have m roots which are the m eigenvalues.

(2) Now consider the case where the coefficients of all polynomials are in the field GF(p) . We saw in Chapter 4 how one can extend a polynomial from the ground field GF(p) to the extension field $GF(p^m)$. We now rewrite the above remainder theorem from this point of view,

$$f(\beta) = q(\beta)\, d(\beta) + r(\beta).$$

where now β is some arbitrary element of $GF(p^m)$. The coefficients of all four polynomials still lie in the ground field GF(p).

Let us now assume that α_i is a root of d(x) lying in $GF(p^m)$. In this case we have

$$f(\alpha_i) = r(\alpha_i)$$

If there are m such roots, we again have m equations in m unknowns, and we can solve for the coefficients of the polynomial r(x). For a general root α_i , there may only be $k < m$ roots of d(x), if d(x) is a minimum polynomial. However, if d(x) happens to be a primitive polynomial, then from Chapter 5 we know there are m roots, and thus we are guaranteed in this case to find r(x).

(3) Finally, elevate all polynomials in the above remainder theorem to the matrix level,

$$f(X) = d(X)\, q(X) + r(X).$$

Select the matrix X = A. In this case, since d(A) = 0 by Cayley-Hamilton, we get

$$f(A) = r(A)$$

We have now proved the following

Fact : If A is an mxm matrix, then a polynomial f(A) of arbitrarily high degree is equal to another polynomial r(A) of degree < m, and we write f(A) = r(A). Moreoever, given f(A), and given A with m distinct eigenvalues, we have a straightforward method for computing the coefficients of r(A).

8.4 The Galois Field Connection

Definition: A *matrix representation* of a Galois Field is a set of matrices which satisfy all the requirements of being a field in general, and which also satisfy the Galois algebra, which means the matrices do the right thing under the • and + operations. In other words, these matrices behave in every respect just like the abstract elements of the Galois Field.

Let us now assume that the elements of the matrix A and all our polynomial coefficients lie in the Galois field GF(p). We have now identified a set of p^m objects that might serve as a representation of $GF(p^m)$. These are the set of all matrix remainder polynomials of the matrix A, the r(A) discussed above.

Fact 1: If the characteristic polynomial d(x) of the mxm matrix A is irreducible over GF(p), then the set of p^m matrices r(A),

$$r(A) = r_0 I + r_1 A + r_2 A^2 + ... + r_{m-1} A^{m-1} = [r_0, r_1, r_2, ...r_{m-1}],$$

forms a *matrix representation* of the Galois Field GF(p^m).

Proof: The basic idea here is to make a correspondence between A and {x} of Chapter 6. In Chapter 6, we made the identification:

$$\alpha = \{x\} = 01000... \qquad d(\{x\}) = 0$$

where we are replacing f(x) with d(x) in the construction GF(p^m) = Polys[x, GF(p)] / (d(x)). We argued in Chapter 6 that, if d(x) is a degree-m monic irreducible polynomial over GF(p), then it must be a minimum polynomial for some element α of GF(p^m) [See Chapter 5, Fact 14]. In this case, all roots of d(x) are known to be elements of GF(p^m). Thus, we argued, {x} must be an element of GF(p^m). Of course any {F(x)} is an element of GF(p^m), and the element is usually labeled as {r(x)} where r(x) = Rem[F(x)/d(x)]. The remainders r(x) had degree < m. Thus, we had $\{x^0 = 1\}$, {x}, $\{x^2\}$, $\{x^3\}$, ... $\{x^{m-1}\}$ as particular elements of GF(p^m), and we associated these powers with the m-tuple unit vectors. It was possible to express any element of GF(p^m) as a linear combination of these basic elements:

$$\{ r(x)\} = r_0 \{1\} + r_1\{x\} + r_2 \{x^2 \} + ... + r_{m-1} \{x^{m-1}\} = [r_0, r_1, r_2, ...r_{m-1}]$$

In exact analogy, we now wish to make the identification:

$$\alpha = A = 01000.. \qquad d(A) = 0$$

We now utter a set of words similar to the above paragraph:

If d(x) is a monic irreducible polynomial over GF(p), then it must be a minimum polynomial for some element α of GF(p^m). In this case, all roots of d(x) are known to be elements of GF(p^m). Thus, since Cayley-Hamilton tells us that d(A) = 0, A must be an element of GF(p^m). Of course any F(A) is an element of GF(p^m), and the element is usually labeled as r(A) where r(A) = F(A) as was shown in Section 8.3. The remainders r(A) had degree < m. Thus, we had $A^0 = 1$, A, A^2, A^3, ... A^{m-1} as particular elements of GF(p^m), and we associated these powers with the m-tuple unit vectors. It is possible to express any element of GF(p^m) as a linear combination of these basic elements:

$$r(A) = r_0 I + r_1 A + r_2 A^2 + ... + r_{m-1} A^{m-1} = [r_0, r_1, r_2, ...r_{m-1}]$$

That is, as you let the coefficients r_i assume all possible values in GF(p), you generate p^m matrices r(A) which represent the elements of GF(p^m). The condition that d(x) be irreducible is necessary so that the resulting structure ends up being a field and not just a ring. If d(x) were reducible, then some elements of r(A) would lack inverses.

One may ask why we deal with A instead of {A}. This is because A is what is a root of d(x), according to Caley-Hamilton. In the non-matrix world, we did not have d(x) = 0, but we did have d({x}) = 0, since this was another way to write { d(x) } which is zero, being a multiple of d(x). We can make this comparison as well:

non-matrix :	F(x) = q(x)d(x) + r(x)	r(x) = Rem[F(x)/d(x)]	{r(x) } = {F(x)}
matrix :	F(A) = q(A)d(A) + r(A)	r(A) = F(A)	

In the matrix world, we do not do a remainder. As noted earlier, any polynomial F(A) in the matrix A is identically equal to a polynomial r(A) of degree m-1 in A. There is no need for {} notation.

Corollary 1: If d(x) of mxm A is irreducible over GF(p), then all properties of the abstract element α of GF(q) which is a root of d(x) pass to the matrix A, where $q = p^m$. Many such properties are stated in Chapter 4. Here are two samples:

(a) There is some least integer $n \le q-1$ such that $A^n = I$. This is known as the order of A, and it divides q-1 evenly. [Chapter 4, Fact 3]

(b) Regardless of the order of A, it must be true that $A^{q-1} = I$. [Chapter 4, Big Theorem 2]

Fact 2: If the characteristic polynomial d(x) of the mxm matrix A is a *primitive polynomial* of $GF(p^m)$, then matrix A is a generator of $\{GF(p^m) - 0, \bullet\}$, and we can enumerate all non-zero elements of $GF(p^m)$ as powers of A. Then all of $GF(p^m)$ can be enumerated as:

$\{ 0, I, A, A^2, A^3, ... A^{q-2}\}$ where $A^{q-1} = I$

Proof: This follows from the definition of primitive polynomial, see Chapter 6.

Corollary 2: We can build a "table" for $GF(p^m)$ exactly as was done in Chapter 6. One replaces the symbol α with the matrix A everywhere.

Here, for example, is a particular table we built in Chapter 6. We have replaced α with A.

Example GF(8 = 2^3). From the primitive polynomial lookup table in Rhee we find that $p(x) = 1 + x + x^3$. Therefore, setting p(A) = 0 we get this reduction rule: $A^3 = 1 + A$. So we can build the entire table in 2 minutes:

0		000	0	
$A^0 = 1$		100	4	
$A^1 = A$	010	2		
A^2		001	1	
$A^3 = 1 + A$		110	6	
$A^4 = A + A^2$		011	3	$A^4 = A\ A^3 = A(1 + A) = A + A^2$
$A^5 = 1 + A + A^2$	111	7		$A^5 = A\ A^4 = A^2 + A^3 = A^2 + A + 1$
$A^6 = 1 + A^2$		101	5	$A^6 = (A^3)^2 = (1+A)^2 = 1 + A^2$

Notice that all 3-tuples are "hit", as claimed above. We have added some octal/hex numbers to make this fact more obvious. Also we know that $A^7 = I$, since the non-zero elements of GF(8) form a cyclic group under •. This fact was not needed to build that above table, but it is needed to construct the • table.

Note that the general "reduction rule" is that given in the previous section,

$$A^m = - d_0 I - d_1 A - d_2 A^2 - ... - d_{m-1} A^{m-1}$$

This then is our reduction rule. It is of course the same as the rule for α, since $d(x) = p(x)$, the primitive polynomial.

Example for the Skeptical Reader: Consider the case $GF(2^9)$. Imagine we have found some 9x9 matrix A which has as its characteristic polynomial $d(x) = 1 + x^4 + x^9$, which we know is a primitive polynomial of $GF(2^9)$. We have $q = 2^9 = 512$. Is it really true that $A^{511} = I$? You raise this 9x9 matrix A to the 511th power and you are magically supposed to get the identity 9x9 matrix? It does sound a bit unreasonable, yet it is really true.

According to the Remainder Theorem, we know right away that $A^{511} = r(A)$, where r(A) is a polynomial in matrix A of degree < 9. This is somewhat amazing in its own right. Now, we seek to find the 9 coefficients of r(A).

To this end, we use the method (1) suggested above which is to consider $\lambda_i^{511} = r(\lambda_i)$ evaluated at the 9 eivenvalues λ_i of matrix the A. This gives us 9 equations in 9 unknowns. We could really do this, thinking of the eigenvalues as lying in the complex number field. There really are 9 such eigenvalues, and they really do give 9 equations in 9 unknowns which can be solved for the 9 coefficients of r(x).

However, it is much easier to use method (2), to think of these eigenvalues as lying in GF(q). Recall that the 9 eigenvalues are the roots of d(x). One of these is element α. Thus, we write $\alpha^{511} = r(\alpha)$. But we know right away that $\alpha^{511} = 1$ from the abstract Galois theory, thus it must be that $r(\alpha) = 1$. Thus, the coefficient $r_0 = 1$ and the other 8 coefficients of r(x) are 0. We are done solving our m equations in m unknowns without going a bit further.

Finally, since $r_0 = 1$, r(A) = I, and thus $A^{511} = I$. Given the 9x9 matrix A, it would take a long time to verify this fact by brute force!

Fact 3: If matrices A and B have the same characteristic polynomial d(x), then if A is a primitive element of $GF(p^m)$, so is B. In this case, all Galois statements above concerning A also apply to B. We can regard either A or B as the matrix whose powers generate { $GF(p^m)$ -0, • }.

Proof: This is just a reminder that d(x) is the only property of A that was used in the above discussion.

Corollary 3 : If A and B are related by a similarity transformation, then the conclusions of Fact 3 follow.

Proof: This is because d(x) is preserved under a similarity transformation, Section 8.1, Fact 12

8.5 The Companion Matrix

In this section we shall explicitly construct forms of the matrix A which makes the above matrix representation of the Galois field GF(p^m) work out.

Let f(x) be some *monic* polynomial of degree m,

$$f(x) = f_0 + f_1 x + ... + f_{m-1} x^{m-1} + x^m.$$

Construct the following mxm matrix :

$$C = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & .. & 0 & -f_0 \\ 1 & 0 & 0 & 0 & 0 & .. & 0 & -f_1 \\ 0 & 1 & 0 & 0 & 0 & .. & 0 & -f_2 \\ 0 & 0 & 1 & 0 & 0 & .. & 0 & -f_3 \\ .. & & & & & & & \\ .. & & & & & & & \\ .. & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & .. & 1 & -f_{m-1} \end{vmatrix}$$

Note that this matrix is mxm, where m is the degree of f(x). There are 1's just off the main diagonal all the way through the matrix. The rightmost column contains the negatives of all but the leading coefficient of f(x), which is 1, since f(x) was assumed monic.

A compact way to write the elements of this matrix is as follows:

$$C_{ij} = \delta_{i,j+1} (1) + \delta_{j,m} (-f_i)$$

Definition: This matrix C is called the *companion matrix* of the polynomial f(x).

Fact 1: The characteristic polynomial of matrix C is precisely f(x).

Observations The m eigenvalues of the companion matrix C are the m roots of f(x). The sum of the eigenvalues is $-f_{m-1}$ since this is the trace. The product of the eigenvalues is is $f_0 (-1)^m$, since this is the determinant.

Proof of Fact 1: There are various well-known operations one can perform on rows or columns which do not change the determinant, such as adding some multiple of one row to another row. Here then is the basic idea. Write out the matrix xI - A. Number the rows from 1 to n, just so we can refer to them, where 1 is the top row. Now, perform this set of operations:

start with matrix $M_1 = xI - A$
do $row_{n-1} := row_{n-1} + x \bullet row_n$ to make matrix M_2
do $row_{n-2} := row_{n-2} + x \bullet row_{n-1}$ to make matrix M_3
....

do $row_1 := row_1 + x \bullet row_2$ to make matrix M_n

At each level, we add x times some row to the row above. What this does is cancel out the various x elements of the rows of M_1, and causes the right column to build up to our polynomial f(x) . What you end up with is something that you can easily take the determinant of. Here is an example for n=4:

$$M_1 = \begin{vmatrix} x & 0 & 0 & f_1 \\ -1 & x & 0 & f_2 \\ 0 & -1 & x & f_3 \\ 0 & 0 & -1 & f_4 + x \end{vmatrix}$$

$$M_2 = \begin{vmatrix} x & 0 & 0 & f_1 \\ -1 & x & 0 & f_2 \\ 0 & -1 & 0 & f_3 + f_4 x + x^2 \\ 0 & 0 & -1 & f_4 + x \end{vmatrix}$$

$$M_3 = \begin{vmatrix} x & 0 & 0 & f_1 \\ -1 & 0 & 0 & f_2 + f_3 x + f_4 x^2 + x^3 \\ 0 & -1 & 0 & f_3 + f_4 x + x^2 \\ 0 & 0 & -1 & f_4 + x \end{vmatrix}$$

$$M_4 = \begin{vmatrix} 0 & 0 & 0 & f_1 + f_2 x + f_3 x^2 + f_4 x^3 + x^4 \\ -1 & 0 & 0 & f_2 + f_3 x + f_4 x^2 + x^3 \\ 0 & -1 & 0 & f_3 + f_4 x + x^2 \\ 0 & 0 & -1 & f_4 + x \end{vmatrix}$$

At this point, the determinant is $(-1)^{4-1} \bullet [\, f_1 + f_2 x + f_3 x^2 + f_4 x^3 + x^4 \,] \bullet \det(-I_3)$. This last factor is just $(-1)^3 = -1$, so we get

$$\det(M) = \det(M_3) = (-1)^{4-1} f(x) (-1)^3 = f(x).$$

In the general case, the result is

$$d(x) = \det(xI - A) = \det(M_1) = \det(M_n) = (-1)^{n-1} f(x) \det(-I_{n-1}) = f(x) \qquad \text{Q.E.D.}$$

Fact 2: The characteristic polynomial d(x) = det(xI - A) remains unaltered if the matrix A is reflected in either or both diagonals.

Proof: (1) Consider det(B). Think of the numbers which comprise B as a square sheet of paper aligned with this page. If this piece of paper is rigidly transformed in any way which leaves it aligned with the page, the determinant is unchanged. These transformations include 180° rotation about the x or y axis, or about either diagonal, or various 90° rotations about an axis perpendicular to the page. (2) Not all of these rigid transformations leave the identity I unchanged. The following do: rotation about either diagonal, or rotation about one diagonal followed by rotation about the other. This last is the same as a perp-to-page rotation by 180°. These diagonal rotations are the same as reflections in the diagonals. The reflection in the main diagonal gives the transpose of the original matrix. The other two transformations give matrices which do not have standard names I am aware of. (3) Since the three transformations just noted leave I intact, we conclude that det(xI - A) = det(xI - A'), where A' indicates one of these transforms on A. Thus, these three transforms do not alter the characteristic polynomial. (4) No doubt these

transforms are similarity transforms, but I forget how to quickly find the matrix S for each such that that $A' = SAS^{-1}$.

Fact 3: Based on Fact 2, we can now write down three alternative forms of the companion matrix. All four forms, including C above, have the same characteristic polynomial:

To get C_1, reflect C in the / diagonal:

$$C_1 = \begin{vmatrix} -f_{m-1} & -f_{m-2} & -f_{m-3} & -f_{m-4} & -f_{m-5} & .. & -f_1 & -f_0 \\ 1 & 0 & 0 & 0 & 0 & .. & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & .. & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & .. & 0 & 0 \\ .. & & & & & & & \\ .. & & & & & & & \\ .. & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & .. & 1 & 0 \end{vmatrix}$$

To get C_2, reflect C in the \ diagonal:

$$C_2 = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 & .. & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & .. & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & .. & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & .. & 0 & 0 \\ .. & & & & & & & \\ .. & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & .. & 0 & 1 \\ -f_0 & -f_1 & -f_2 & -f_3 & -f_4 & .. & -f_{m-2} & -f_{m-1} \end{vmatrix}$$

To get C_3, reflect C_2 in the / diagonal:

$$C_3 = \begin{vmatrix} -f_{m-1} & 1 & 0 & 0 & 0 & .. & 0 & 0 \\ -f_{m-2} & 0 & 1 & 0 & 0 & .. & 0 & 0 \\ -f_{m-3} & 0 & 0 & 1 & 0 & .. & 0 & 0 \\ -f_{m-4} & 0 & 0 & 0 & 1 & .. & 0 & 0 \\ .. & & & & & & & \\ .. & & & & & & & \\ -f_1 & 0 & 0 & 0 & 0 & .. & 0 & 1 \\ -f_0 & 0 & 0 & 0 & 0 & .. & 0 & 0 \end{vmatrix}$$

Big Corollary: Let p(x) be a primitive polynomial of $GF(p^m)$. Then the companion matrix C (or any of its alternative forms) with f(x) = p(x) can serve as the matrix A which represents a primitive element of a matrix representation of $GF(p^m)$.

This corollary will have major application in the theory of scramblers and related circuits to come.

Example: Here we explicitly construct all matrices for GF(2^3),

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad A^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \qquad A^3 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

$$A^4 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \qquad A^5 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \qquad A^6 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

We constructed A from the companion C, then built up the others by hand taking powers of A. The following relationships were verified:

$$A^4 \bullet A^3 = I \qquad A^5 \bullet A^2 = I \qquad A^6 \bullet A = I \qquad A^3 = I + A$$

Here I is the 3x3 identify matrix, and we expect $A^7 = I$. Using the above matrices, one could verify every single entry in the GF(8) "table" quoted above, replacing abstract α with the matrix A.

Sample calculation: Compute A^{973}.

Solution: Rem(973/7) = 0, so $A^{973} = A^0 = I$.