GALOIS FIELDS

6.441 Supplementary Notes 7, 4/30/92

R. G. Gallager

A field F Is a set of two or more elements, F={α,β,...) closed under two operations, + (addition) and * (multiplication) with the following properties:

a) F is an Abelian group under addition; the additive identity is called 0 and the additive inverse of α is called -α.

b) The set of non-zero elements is closed under multiplication (i.e., $\alpha\neq 0,\beta\neq 0$ implies $\alpha*\beta\neq 0$); there is a multiplicative identity called 1 (i.e., $1*\alpha=\alpha*1=\alpha$ for all α; and multiplication satisfies the associative and commutative laws.

c) There is a multiplicative inverse, denoted α^{-1} for each non-zero element α.

d) The distributive law is satisfied: $(\alpha+\beta)*\gamma=(\alpha*\gamma)+(\beta*\gamma)$

A Galois field is a field with a finite number of elements. If a Galois field has q elements, we denote it as GF(q). An integral domain is a structure that satisfies conditions a), b), and d) above. Conditions b) and c) above are equivalent to stating that the set of non-zero elements is an Abelian group under multiplication. We separate these conditions here for several reasons. First, lemma 1 below will show that condition c) is implied by the other conditions if the set F is finite (i.e., that all finite integral domains are Galois fields). Second, integral domains are important in their own right. They include the set of integers (under ordinary addition and multiplication), and, as discussed later, the set of polynomials over a field. Third, condition c) asserts that division is defined (i.e., for any $\alpha\in F$ and any non-zero $\beta\in F$, α/β is defined as $\alpha*(\beta^{-1})$. When c) is not satisfied, one typically gets both a quotient and a remainder when one tries to divide.

EXAMPLES:

1) The set of real numbers (or rational numbers, or complex numbers) using ordinary addition and multiplication.

2) The set F={0,1} using mod 2 addition and multiplication. This is called the binary field or GF(2) and satisfies the above conditions by inspection.

3) For any prime number p, the set F={0,1,...p-1} using mod p addition and multiplication. This is called the p-ary field or GF(p) and we demonstrate shortly that it is a field.

The latter two examples are finite fields, or Galois fields; Galois fields are of interest for algebraic coding, but the real field provides a guide in dealing with Galois fields, since the usual rules for addition, subtraction, multiplication, and division apply for all fields. In particular, the following rules apply to all integral domains, and thus also to fields.

1) $\alpha*0=0*\alpha=0$ for all $\alpha \in F$
2) $-(\alpha*\beta)=(-\alpha)*\beta=\alpha*(-\beta)$
3) For any $\beta \neq 0$, $\alpha*\beta=\gamma*\beta \Rightarrow \alpha=\gamma$ (Cancellation law)

To verify rule 1, note that for any α and β, $\alpha*\beta = \alpha*(\beta+0) = \alpha*\beta+\alpha*0$. Thus $\alpha*0$ must be the additive identity, which is 0. By commutativity, $0*\alpha = 0$. For rule 2, we have $0 = 0*\beta = (\alpha+(-\alpha))*\beta = \alpha*\beta+(-\alpha)*\beta$. Thus $(-\alpha)*\beta$ is the additive inverse of $\alpha*\beta$, which means it is equal to $-(\alpha*\beta)$. For rule 3, assume $\beta \neq 0$. Then $\alpha*\beta=\gamma*\beta \Rightarrow \alpha*\beta-(\gamma*\beta)=0 \Rightarrow \alpha*\beta+(-\gamma)*\beta=0 \Rightarrow (\alpha-\gamma)*\beta=0 \Rightarrow \alpha-\gamma=0 \Rightarrow \alpha=\gamma$.

Now consider F={0,1,...p-1} under mod p addition and multiplication. If p is not prime, then for some positive integers α, β in F, $\alpha\beta=p$ under ordinary multiplication, and thus $\alpha*\beta=0$ under mod p multiplication. Thus, condition b) is violated and we see that F is not a field (under mod p addition and multiplication) if p is not prime. If p is prime, conditions a), b), and d) are easily verified. Since p is finite, the following lemma demonstrates the

existence of a multiplicative inverse for all non-zero elements. Thus the set of a prime number p of elements with mod p addition and multiplication is a field, denoted GF(p).

LEMMA 1: If an integral domain has a finite number of elements, it is a field.

Proof: We must establish the existence of a multiplicative inverse for an arbitrary element $\alpha \neq 0$. In particular, consider $\{\alpha^1, \alpha^2, ...\}$ where $\alpha^i = \alpha * \alpha * ... * \alpha$ with m terms. We must have $\alpha^i = \alpha^{i+j}$ for some positive i,j, since the set is finite. By rule 3 above, this implies that $1 = \alpha^j$. Thus $1 = \alpha * \alpha^{j-1}$, so that α^{j-1} is the multiplicative inverse of α.

VECTOR SPACES

We have already used the idea of representing binary code words as vectors of elements from GF(2). Such vectors can be added (by adding the individual elements component-wise) and multiplied by field elements (i.e., multiplying $\mathbf{v}$ by 1 yields $\mathbf{v}$, and by 0 yields the all zero vector $\mathbf{0}$). Formalizing and generalizing this, we have the definition:

A vector space V over a given field F is a set of elements (called vectors) closed under an operation + called vector addition. There is also an operation * called scalar multiplication, which operates on an element of F (called a scalar) and an element of V to produce an element in V. The following properties are satisfied:

a) V is an Abelian group under vector addition. Let $\mathbf{0}$ denote the additive identity.

b) For every $\mathbf{v}, \mathbf{w} \in V$ and every $\alpha, \beta \in F$, we have

$(\alpha * \beta) * \mathbf{v} = \alpha * (\beta * \mathbf{v})$,

$\alpha * (\mathbf{v} + \mathbf{w}) = \alpha * \mathbf{v} + \alpha * \mathbf{w}$

$(\alpha + \beta) * \mathbf{v} = \alpha * \mathbf{v} + \beta * \mathbf{v}$

$1 * \mathbf{v} = \mathbf{v}$ where 1 is the multiplicative identity of F.

Note that the operators + and * do double duty as the operations in the field and vector space; this causes no confusion since it is always clear what is being operated on. We will often leave out the * (i.e., replacing $\alpha * \mathbf{v}$ with $\alpha\mathbf{v}$) when there is no danger of confusion. There is nothing subtle hidden in the definition above; it is the same as the vector spaces you are used to, with an arbitrary field F in place of the real field. In most cases, V can be represented by n-tuples of field elements for some given n; as we shall soon see, this is always true when V is a finite set of vectors.

A set of vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ is said to be linearly independent if $\alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+...+\alpha_k\mathbf{v}_k \neq \mathbf{0}$ for all choices of scalars $\alpha_1, \alpha_2, ..., \alpha_k$ in which at least one scalar is non-zero. A set of vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ is said to be a basis of the vector space if the set $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ is linearly independent and if every vector $\mathbf{w}$ in V can be represented as $\alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+...+\alpha_k\mathbf{v}_k$ for some choice of scalars $\alpha_1, \alpha_2, ..., \alpha_k$. The space is said to be k dimensional if there are k vectors that form a basis. Note that, given a basis $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$, we can represent each vector $\mathbf{w} = \alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+...+\alpha_k\mathbf{v}_k$ by the k tuple $(\alpha_1, \alpha_2, ..., \alpha_k)$; vector addition is then just component wise addition on these k tuples and scalar multiplication is just multiplication of the scalar by each element of the k tuple.

LEMMA 2: Let V be a vector space over a Galois field with q elements. If V has a finite number of elements, then the number of elements is q^n for some integer $n>0$.

Proof: Choose a basis by selecting arbitrary non-zero vectors one at a time such that, after selecting $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$, the next vector selected is linearly independent of $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$. If there is no such vector, then all vectors can be represented as $\mathbf{w} = \alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+...+\alpha_k\mathbf{v}_k$ for some choice of $\alpha_1, \alpha_2, ..., \alpha_k$ and $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ is a basis. Each distinct choice of scalars $\alpha_1, \alpha_2, ..., \alpha_k$ corresponds to a distinct vector, since if $\alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+...+\alpha_k\mathbf{v}_k = \beta_1\mathbf{v}_1+\beta_2\mathbf{v}_2+...+\beta_k\mathbf{v}_k$, then we would have $(\alpha_1-\beta_1)\mathbf{v}_1+(\alpha_2-\beta_2)\mathbf{v}_2+...+(\alpha_k-\beta_k)\mathbf{v}_k = 0$, which

is a contradiction of the independence of $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$. Finally, there are q^k choices for the set of scalars $\alpha_1, \alpha_2, ..., \alpha_k$, completing the proof.

POLYNOMIALS

An expression of the form $f_nD^n+f_{n-1}D^{n-1}+...+f_1D+f_0$ (usually denoted f(D)) is called a polynomial of degree n over F, where $f_n, f_{n-1}, ..., f_0$ are elements of a field F and $f_n \neq 0$. In the special case n=0, we regard the field elements as polynomials of degree 0. In particular, the zero field element is regarded as the zero polynomial. Two polynomials f(D) and g(D) are said to be equal iff they have the same degree n and $f_i=g_i$ for $0 \leq i \leq n$. The element D in a polynomial is referred to as an indeterminate. It should not be thought of as a variable within the field F and f(D) should not be thought of as a function mapping elements of F into F. To see the issue here, consider the polynomials D^2 and D over GF(2) (we consistently use the shorthand of writing 1 D^k as just D^k). We see that D^2 and D, considered as functions of D, both take value 1 for D=1 and value 0 for D=0. Thus they are equal as functions from GF(2) into GF(2), but not equal as polynomials.

The sum of two polynomials f(D)+g(D) is defined to be

$$f(D)+g(D) = \Sigma_i (f_i+g_i)D^i$$

Similarly, the product of two polynomials is given by convolution:

$$f(D)g(D) = \sum_{i \geq 0} (\sum_{j=0}^{i} f_j g_{i-j})D^i$$

As an example, over GF(2),

$$(1 + D + D^2)(1 + D) = 1 + (1+1)D + (1+1)D^2 + D^3 = 1 + D^3$$

Note that the degree of the product of two non-zero polynomials is always the sum of the degrees of the individual polynomials.

The set of polynomials over any given field can be seen to be an Abelian group under polynomial addition. Conditions b) and d) also hold but c) does not. This gives another

example where lemma 1 fails when the condition of a finite number of elements is violated (note that the set of polynomials, even over GF(2), is infinite since the degree is unbounded; if we consider the set of polynomials of degree less than n, then the set is finite, but is not closed under multiplication).

The set of polynomials over a field F can be regarded as a vector space. The vector addition is, of course, the polynomial addition defined above, and scalar multiplication is defined by $\alpha\, f(D) = \sum_i \alpha f_i D^i$ for $\alpha \in F$. Here the set of polynomials of degree less than n is an n dimensional vector space, and the polynomials 1, D, D^2, ...,D^{n-1} form a basis. For this vector space, the powers of D serve as little more than "place holders."

Polynomials can be divided by one another if one is willing to tolerate a remainder term. The algorithm for dividing polynomials over an arbitrary field is the same as the long division algorithm one learns in high school (except, of course, that the division of field elements uses the division rules of the given field). Example (for GF(2)):

$$
\begin{array}{r|l}
D^3 + D^2 + 1 & D + 1 \\
\cline{2-2}
D^4 + D^2 + 1 & \\
D^4 + D^3 + D & \\
\cline{1-1}
D^3 + D^2 + D + 1 & \\
D^3 + D^2 + 1 & \\
\cline{1-1}
D &
\end{array}
$$

The remainder (D in the above example) is always of lower degree than the divisor (if not, the algorithm continues until it is). If we divide f(D) by g(D) using the algorithm above, we can represent f(D) in terms of the quotient h(D) and the remainder r(D) by f(D) = g(D)h(D) + r(D). If g(D) has degree greater than 0, then there is a unique h(D) and a unique r(D) of degree less than g(D) that satisfies this equation. If r(D) = 0 (i.e., if f(D) = g(D)h(D)) then g(D) is said to be a factor or divisor of f(D). If g(D) and h(D) are both of degree greater than 0, then f(D) is said to be reducible; if f(D) has no such divisors, f(D) is

irreducible. A polynomial f(D) is said to be monic if its leading coefficient is 1 (i.e., the multiplicative identity of the field). The following theorem is familiar from elementary algebra. Its proof, however, is surprisingly tricky and is given in the text (Theorem 6.4.3)

THEOREM 1: (Unique factorization) A polynomial f(D) over a given field has a unique factorization into a field element times a product of monic irreducible polynomials over the field, each of degree greater than 0.

An element α of a field is defined to be a root of a polynomial f(D) over that field if $f(\alpha)=0$, i.e., if $\sum_i f_i * \alpha^i = 0$. We shall see that understanding Galois fields is very closely coupled with understanding the roots of various polynomials.

THEOREM 2: An element α of a field F is a root of a non-zero polynomial f(D) over F iff $(D-\alpha)$ is a factor of f(D). If f(D) has degree n, then at most n field elements are roots of f(D).

Proof: Dividing $D-\alpha$ into f(D), we have $f(D) = (D-\alpha)h(D)+r(D)$. Since r(D) has degree lower than $D-\alpha$, it has degree 0; thus $f(D) = (D-\alpha)h(D)+r_0$ where r_0 is a field element. It follows that $f(\alpha)=r_0$, so α is a root of f(D) iff $r_0=0$, i.e., iff $D-\alpha$ is a factor of f(D). Since each root corresponds to a factor of degree 1, the unique factorization theorem assures us that there are at most n factors and thus at most n roots.

We can now construct another example of a Galois field. We will show later that all Galois fields can be represented in this way. Suppose f(D) is a given monic irreducible polynomial of degree n over the field GF(p) for some prime p. Consider the set of polynomials of degree less than n. We define two operations on these polynomials, one of which is the polynomial addition already defined. The other is a new kind of multiplication, namely polynomial multiplication modulo f(D); we denote this by the symbol $*$, so that $g(D)*h(D)$ = g(D)h(D) modulo f(D), where g(D)h(D) modulo f(D) means the remainder when

g(D)h(D) is divided by f(D). Since this remainder is of degree less than n, we see that the set of polynomials of degree less than n is closed under multiplication modulo f(D).

THEOREM 3: Assume f(D) is an irreducible polynomial over GF(p) of degree n. Then the set of polynomials over GF(p) of degree less than n, under polynomial addition and multiplication modulo f(D) forms a Galois field of p^n elements.

Proof: We have already seen that the given set of polynomials forms an Abelian group under polynomial addition, and it is straightforward to verify the distributive law, and the associative and commutative laws for multiplication modulo f(D). By lemma 1, then, it is enough to verify that h(D)*g(D) is non-zero for all h(D)≠0 and g(D)≠0. Thus we must verify that there is no polynomial q(D) such that h(D)g(D) = f(D)q(D). Applying the unique factorization theorem to h(D) and g(D), we see that h(D)g(D) uniquely factors into polynomials of degree at most that of h(D) and g(D). Since f(D) is irreducible and of degree greater than h(D) and g(D), there can be no q(D) such that h(D)g(D)=f(D)q(D).

Example: Consider the irreducible polynomial D^2+D+1 over GF(2) (to verify that D^2+D+1 is irreducible, show that neither D nor D+1 are factors; these are the only degree 1 polynomials over GF(2)). The resulting field of polynomials of degree less than 2 over GF(2) modulo D^2+D+1 contains 4 elements, namely 0, 1, D, D+1. These elements can be represented (in the vector form) by 00, 01, 10, 11 respectively. The addition and * multiplication tables are then given below. Note that 01 is the multiplicative identity. In understanding the multiplication table, note that, for example, $(D+1)(D+1) \bmod (D^2+D+1) = (D^2+1) \bmod (D^2+D+1) = D$. Thus the lower right hand corner of the table indicates (11)*(11) = (10).

+	00	01	10	11
00	00	01	10	11
01	01	00	11	10
10	10	11	00	01
11	11	10	01	00

*	00	01	10	11
00	00	00	00	00
01	00	01	10	11
10	00	10	11	01
11	00	11	01	10

THE STRUCTURE OF GALOIS FIELDS

A subfield is a field whose elements are a subset of the original field and whose multiplication and addition operations are the same as those of the original field.

LEMMA 3: If the elements of a subset F' of a Galois field F are closed under the addition and multiplication operations of F, then F' is a subfield.

Proof: For any α in F', repeated addition of α to itself comes back to the additive inverse $-\alpha$, and with one more addition of α, to the additive identity element 0. Thus F', with the addition operation of F, is an Abelian group. Similarly, repeated multiplication shows that α^{-1} and 1 are in the subset, so that the non-zero elements of F' forms an Abelian group under multiplication. Since we have closure under the addition and multiplication of F, and the distributive law is satisfied, F' then satisfies all the axioms of a field.

Note that this lemma is not true for fields in general; for example in the real field, the set of numbers greater than or equal to 1 is closed under addition and multiplication, but it is certainly not a subfield.

THEOREM 4: Every Galois field has a subfield with a prime number of elements.

Proof: Consider the subset {0, 1, 1+1, 1+1+1, ...}. This is a cyclic subgroup of the additive group of the field and has some number p of elements. Denoting the sum of i 1's as the element i for i<p, we note that addition on this subgroup is just addition modulo p.

Using the distributive law to discover the multiplication law on these elements, we see that it has to be multiplication modulo p. Thus this subset is closed under multiplication. Finally, if p were not prime, and had two factors, say i and j, then i*j, in the * operation of the field, would be the sum of p 1's, which is 0. This is a contradiction since F is a field, and thus p is prime.

It is clear from the construction above that the prime subfield above is unique and that any other subfield must contain all these elements (since it contains 0, 1, and sums of 1's). Thus p is called the characteristic of the finite field and the elements of the subfield are called the integers of the field. Since any field of p elements has these same integers with the same mod p operations, this is the only field of p elements, justifying calling it GF(p). What we really mean when we say that there is only one field of p elements is that any field of p elements is isomorphic to the field above in the sense that the elements can be mapped into these integers with preservation of the + and * rules. We show later that all fields with any given number q of elements are isomorphic, thus justifying the notation GF(q).

Given a Galois field, it is useful to consider the elements of the Galois field as vectors in a vector space in which the scalars are the integers of the field. The elements of the field can be added, subtracted, and can be multiplied by field elements within the vector space structure. It follows then from lemma 2 that the field must have p^n elements for some positive integral n; we state this as a theorem.

THEOREM 5: A Galois field of characteristic p has p^n elements for some integer $n>0$.

Carrying the linear vector space view a little further, let $\alpha \neq 0$ be an arbitrary element in a field $GF(p^n)$, and let m be the largest integer such that $1, \alpha, \alpha^2, \ldots, \alpha^{m-1}$ (considered as vectors over GF(p) are linearly independent. This means that the set of elements

$$S(\alpha) = \{ \sum_{i=0}^{m-1} k_i \alpha^i \ : \ k_i \in GF(p),\ 0 \leq i \leq m-1 \} \qquad (1)$$

contains p^m distinct elements (note that m might be smaller than n since $S(\alpha)$ need not contain all elements of the field; for example, $S(\alpha)$ contains only p elements if α itself is an integer of the field). Since α^m is a linear combination of $1, \alpha, \ldots \alpha^{m-1}$, we see that $\alpha^m + f_{m-1}\alpha^{m-1} + \ldots + f_1\alpha + f_0 = 0$ for some choice of field integers $f_0, \ldots f_{m-1}$. This means that α is a root of the polynomial $f(D) = D^m + f_{m-1}D^{m-1} + \ldots + f_1 D + f_0$.

The minimal polynomial $f_\alpha(D)$ of an element α of a field $GF(p^n)$ is defined as the monic polynomial of lowest degree over GF(p) for which α is a root. Thus f(D) above is the minimal polynomial of α and we see that the minimal polynomial of each element has degree less than or equal to n. Recall that the only example we have seen of a Galois field with p^n, $n>1$, elements is that of the set of polynomials modulo an irreducible polynomial. The polynomial representation of such an element is not at all the same as the minimal polynomial of the element, and to avoid confusion between the two, it is often preferable to represent elements of the field either abstractly (e.g., an element α) or as a vector.

Now suppose that m, the degree of the above polynomial f(D), is strictly less than n. Thus $S(\alpha)$ is strictly a subset of F. On the other hand, $\alpha^m = -f_{m-1}\alpha^{m-1} - \ldots - f_1\alpha - f_0$ is contained in $S(\alpha)$. It follows that $\alpha^{m+1} = -\alpha(f_{m-1}\alpha^{m-1} + \ldots + f_1\alpha + f_0)$ is also contained in $S(\alpha)$, since the first term in the above expression is $f_{m-1}\alpha^m$, which, as we have just seen, is in $S(\alpha)$. Extending this argument, α^i is in $S(\alpha)$ for all i, and it follows easily that $S(\alpha)$ is closed under multiplication and addition, and thus $S(\alpha)$ is a subfield of F. Since all the elements in $S(\alpha)$ are generated by additions and multiplications between α and the integers of the field, α cannot be in any smaller subfield than $S(\alpha)$. Finally, we can consider $GF(p^n)$ to be a linear vector space over the subfield $S(\alpha)$ (instead of over GF(p)). Lemma 2 then shows that the number of elements in the field is $(p^m)^j$ for some integer j. This means that m must be a factor of n. We have proved the following theorem:

THEOREM 6: In a field of p^n elements, each element $\alpha \neq 0$ has a minimal polynomial with a degree $m(\alpha)$ that is either equal to n or divides n. The element α is contained in a subfield of F with $p^{m(\alpha)}$ elements (if $m(\alpha)=n$, this subfield is F itself) and α is contained in no subfield with fewer than $p^{m(\alpha)}$ elements.

LEMMA 4: Let P(D) be a polynomial over GF(p) and let $\alpha \in GF(p^n)$ have minimal polynomial $f_\alpha(D)$. Then α is a root of P(D) if and only if $f_\alpha(D)$ is a factor of P(D); also $f_\alpha(D)$ is irreducible over GF(p).

Proof: If $f_\alpha(D)$ is reducible, then α is a root of one of its factors, which contradicts the definition of a minimal polynomial. If we divide $f_\alpha(D)$ into P(D), we get

$$P(D) = f_\alpha(D)q(D) + r(D)$$

where r(D) is of degree smaller than $f_\alpha(D)$. Thus we have $P(\alpha)=r(\alpha)$. If α is a root of P(D), then $r(\alpha)$ is zero. If r(D) is not the zero polynomial, then it can be multiplied by some element of GF(p) to make it monic; α is a root of this monic polynomial and this polynomial has degree less than $f_\alpha(D)$; this is a contradiction, so r(D) is zero and $f_\alpha(D)$ divides P(D). If α is not a root of P(D), then $P(\alpha)\neq 0$, $r(\alpha)\neq 0$, $r(D)\neq 0$, and $f_\alpha(D)$ does not divide P(D).

We have seen that if an irreducible polynomial of degree n over GF(p) exists, then a Galois field of $q=p^n$ elements exists. We also know from Lagrange's theorem that the multiplicative order of each non-zero element divides q-1 and thus that each non-zero element is a root of $D^{q-1}-1$.

THEOREM 7: Assume that GF(q) exists for $q = p^n$. Let $f_1(D), f_2(D),...f_L(D)$ be the distinct minimal polynomials (over GF(p)) of GF(q). Then

$$D^{q-1} - 1 = \prod_{i=1}^{L} f_i(D) \qquad (2)$$

Proof: Each non-zero element of GF(q) is a root of the polynomial on the right, which therefore has degree at least q-1. Each minimal polynomial is an irreducible factor of $D^{q-1}-1$, however, so the right hand side divides the left ; since the right side is monic and its degree is not less than the degree of the left side, the two must be equal.

The left side of (2) can be factored in another way - for any k that divides q-1, we have

$$D^{q-1} - 1 = (D^k-1)(D^{q-1-k}+D^{q-1-2k}+\ldots+D^k+1) \qquad (3)$$

Some of the minimal polynomials on the right side of (2) are factors of D^k-1 and some are factors of $D^{q-1-k}+D^{q-1-2k}+\ldots+D^k+1$. The elements of GF(q) that have a multiplicative order equal to k or a divisor of k are roots of D^k-1 and there must be k such elements (since every non-zero element of GF(q) is a root of one of the polynomials on the right side of (3).

THEOREM 8: The multiplicative group of every Galois field $GF(p^n)$ is cyclic. That is, there is some element α (called a primitive element) whose powers include all non-zero elements of the field; the degree of $f_\alpha(D)$ is n.

Proof: For any m dividing q-1, the number of elements whose multiplicative order is m or a divisor of m is m; this is the same as in a cyclic group of q-1 elements (i.e., if km=q-1 in a cyclic group $\{\alpha, \alpha^2, \ldots, \alpha^{q-1}=1\}$, then $\{\alpha^k, \alpha^{2k}, \ldots, \alpha^{mk}=1\}$ is the set of elements of order m or a divisor of m).. It follows that the number of elements whose order is exactly q-1 is also the same as in a cyclic group. This is at least one since a cyclic group of q-1 elements by definition has an element of order q-1. Taking α as such a primitive element, let $m(\alpha)$ be the degree of $f_\alpha(D)$. From theorem 6, if $m(\alpha)<n$, then α is an element of a subfield with $p^{m(\alpha)}$ elements, and the multiplicative order of α is at most $p^{m(\alpha)}-1$; this is a contradiction, so $m(\alpha)=n$.

EXAMPLE: Consider $GF(2^4)$. Using (3), we have $D^{15}-1 = (D^5-1)(D^{10}+D^5+1)$.

$$D^5-1 = (D-1)(D^4+D^3+D^2+D+1) \qquad (4)$$

$$D^{10}+D^5+1 = (D^2+D+1)(D^4+D+1)(D^4+D^3+1) \qquad (5)$$

The minimal polynomial (D-1) on the right side of (4) is the minimal polynomial of the identity element 1. $D^4+D^3+D^2+D+1$ is the minimal polynomial of the other four roots of D^5-1, which are the four elements of multiplicative order 5. In terms of a primitive element α, these elements are α^3, α^6, α^9, and α^{12} (to see this for α^3, for example, note that $(\alpha^3)^5 = \alpha^{15} = 1$). The minimal polynomial D^2+D+1 is a factor of D^3-1; it is the minimal polynomial of the elements (other than 0, 1) in the subfield GF(4); in terms of α, these elements are α^5 and α^{10}. For α a root of D^4+D+1, we see that $\alpha^5=\alpha+\alpha^2$ and $\alpha^{10}=1+\alpha+\alpha^2$, from which the fact that these elements, along with 0 and 1, form a subfield is not too surprising. The final two minimal polynomials in (5) are primitive polynomials. Primitive polynomials over GF(2) always come in symmetric pairs (where the symmetry is to interchange coefficients f_i and f_{n-i} for each $i \le n/2$). If one is the minimal polynomial for α, the other is the minimal polynomial for α^{-1}.

THEOREM 9: All Galois fields with the same number of elements are isomorphic.

Proof: Theorem 7 showed that Eq. (2) determines the set of minimal polynomials of an arbitrary field with $q=p^n$ elements, so all fields with q elements have the same minimal polynomials. Theorem 5 shows that the set of minimal polynomials includes at least one of degree n. Letting α be an arbitrary root of a given polynomial of degree n, $S(\alpha)$ in (1) contains all the elements of the field and gives the addition operation in terms of the vector addition over GF(p). The multiplication operation is uniquely defined since α^n is determined as an element of $S(\alpha)$ by $f_\alpha(\alpha)=0$. Since all fields with p^n elements have this same addition and multiplication rule, they are isomorphic.

We next want to show that $GF(p^n)$ exists for all primes p and all integers n. In order to do this, we first develop a result about the irreducible factors of D^{p^n-1}, and then we demonstrate the existence of irreducible polynomials for all p and n.

THEOREM 10: Let f(D) be an irreducible monic polynomial over GF(p) of degree m; then f(D) is a factor of $D^{p^n-1} - 1$ iff m divides n.

Note that this theorem is quite similar to theorem 6, except that here we do not assume the existence of $GF(p^n)$. Before proving the theorem, we need two lemmas.

LEMMA 5: Assume $GF(p^n)$ exists for some prime p and integer n. For α, β in $GF(p^n)$ and any integer m>0,

$$(\alpha+\beta)^{p^m} = \alpha^{p^m} + \beta^{p^m}. \tag{6}$$

Proof: We can use the binomial theorem to expand $(\alpha+\beta)^p$,

$$(\alpha+\beta)^p = \alpha^p + p\alpha^{p-1}\beta + \ldots + \binom{p}{i}\alpha^i\beta^{p-i} + \ldots + \beta^p \tag{7}$$

$\binom{p}{i}\alpha^i\beta^{p-i}$ is to be interpreted as the sum of $\binom{p}{i}$ terms, each $\alpha^i\beta^{p-i}$. We have

$$\binom{p}{i} = \frac{p!}{i!(p-i)!} = \frac{p(p-1)!}{i!(p-i)!}$$

This is an integer and p is prime. For 0<i<p, the denominator does not contain the factor p, so the denominator must divide (p-1)!. Thus $\binom{p}{i}$ is a multiple of p, which in a field of characteristic p is 0; thus only the outer terms in (7) remain, completing the proof for m=1.

Applying the result $(\alpha+\beta)^p = \alpha^p + \beta^p$ twice, once to α and β, and then to α^p and β^p,

$$(\alpha+\beta)^{p^2} = ((\alpha+\beta)^p)^p = (\alpha^p + \beta^p)^p = \alpha^{p^2} + \beta^{p^2} \tag{8}$$

Repeating this argument m-1 times, we get Eq. (6).

LEMMA 6: Suppose $GF(p^m)$ exists. For any positive integer n, the set of elements of $GF(p^m)$ that are roots of D^{p^n}-D form a subfield of $GF(p^m)$ (where the subfield might be $GF(p^m)$ itself).

Proof: Note that 0 is a root of D^{p^n}-D and the other roots are the elements whose multiplicative order divides p^n-1. Let T be the set of elements of GF(p^m) that are roots of D^{p^n}-D. From lemma 3, it suffices to show that T is closed under addition and multiplication. Suppose $\alpha,\beta \in T$. From (6),

$$(\alpha+\beta)^{p^n} = \alpha^{p^n} + \beta^{p^n} = \alpha + \beta$$

The final equality above results from $\alpha,\beta \in T$, and this shows that $\alpha+\beta \in T$. Also

$$(\alpha*\beta)^{p^n} = \alpha^{p^n} * \beta^{p^n} = \alpha * \beta$$

completing the demonstration.

Proof of theorem 10: Let GF(p^m) be the field formed by the polynomials over GF(p) with multiplication modulo f(D) and let $\alpha \in$ GF(p^m) be a root of f(D). Assume that f(D) is a factor of D^{p^n-1} - 1. Lemma 6 implies that the roots of D^{p^n-1} - 1 constitute all the non-zero elements of a subfield of GF(p^m). Since α is in this subfield, and since theorem 6 shows that α is not in any subfield smaller than GF(p^m), the subfield must be GF(p^m) itself. Thus the primitive elements of GF(p^m) must be roots of D^{p^n-1} - 1, so that p^m-1 must divide p^n-1. Carrying out the division, $p^n-1 = (p^m-1)(p^{m-n} + p^{m-2n} + ...)$ which shows that m must divide n. Next assume that m divides n. Then, from the division above, p^m-1 divides p^n-1 and thus D^{p^m-1}-1 is a factor of D^{p^n-1}-1. Thus f(D) is also a factor of D^{p^n-1}-1.

THEOREM 11: GF(p^n) exists for all primes p and all positive integers n.

Proof: We already know that GF(p) exists for all prime p, so we need consider only $n \geq 2$. We also know that GF(p^n) exists if an irreducible polynomial over GF(p) of degree n exists. We complete the proof by considering the factorization of D^{p^n-1}-1 over GF(p). The irreducible monic factors of this polynomial, from theorem 10, all have degrees of n or divisors of n. Thus each irreducible factor has degree n or a degree at most n/2. We merely

need to show that there are not enough monic irreducible polynomials of degree n/2 or less to yield a product of degree p^n-1. Let $m \le n/2$. The sum of the degrees of all monic irreducible polynomials of degree m is at most p^m-1 (from theorem 7). Upper bounding this by $p^{n/2}-1$ and summing over $m \le n/2$, we see that there are at most $(n/2)p^{n/2}-1$ irreducible monic polynomials of degree n/2 or less. Showing that $p^n - (n/2)p^{n/2} > 0$ is thus sufficient to show that monic irreducible polynomials of degree n exist. We note by inspection that this inequality is satisfied for p=2, n=2, and the left side is an increasing function of n and p for all larger n and p.

We complete this section with the following theorem, which is sometimes useful in hardware manipulations with Galois fields.

THEOREM 12: For any non-zero α in $GF(p^n)$, let f(D) be the minimal polynomial of α. Then the roots of f(D) are $\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{m-1}}$ where m is the degree of f(D).

Proof: Let $f(D) = D^m + f_{m-1}D^{m-1} + \ldots + f_1 D + f_0$ be the minimal polynomial of α; thus $f(\alpha)=0$. We first show that $f(\alpha^p)=0$, demonstrating that α and α^p have the same minimal polynomial. Using (6), we have

$$[f(\alpha)]^p = \left(\alpha^m + \sum_{i=0}^{m-1} f_i \alpha^i\right)^p = \alpha^{pm} + \left(\sum_{i=0}^{m-1} f_i \alpha^i\right)^p \qquad (9)$$

Using (6) again on the final term in (9),

$$\left(\sum_{i=0}^{m-1} f_i \alpha^i\right)^p = \left(f_{m-1}\alpha^{m-1} + \sum_{i=0}^{m-2} f_i \alpha^i\right)^p = (f_{m-1})^p \alpha^{p(m-1)} + \left(\sum_{i=0}^{m-2} f_i \alpha^i\right)^p$$

Since f_{m-1} is an integer of the field, $(f_{m-1})^p = f_{m-1}$, and we have

$$[f(\alpha)]^p = \alpha^{pm} + f_{m-1}\alpha^{p(m-1)} + \left(\sum_{i=0}^{m-2} f_i \alpha^i\right)^p$$

Continuing with the remaining terms in the same way,

$$[f(\alpha)]^p = \alpha^{pm} + f_{m-1}\alpha^{p(m-1)} + \ldots + f_i\alpha^{pi} + \ldots + f_1\alpha^p + f_0 = f(\alpha^p) \quad (10)$$

Since $f(\alpha)=0$, the left side of (10) is 0 and thus $f(\alpha^p) = 0$. Since α is arbitrary, we can now substitute α^p for α and assert that α^p and $(\alpha^p)^p$ have the same minimal polynomial. In the same way α^{p^i} has the same minimal polynomial for all i.

Next observe that $\alpha = \alpha^{p^i}$ iff α is a root of $D^{p^i-1} - 1$. From theorem 10, this is true iff m divides i. Thus $\alpha \neq \alpha^{p^i}$ for $1 \leq i \leq m-1$. Similarly, if $\alpha^{p^i} = \alpha^{p^j}$ for $j>i$, we can define $\beta = \alpha^{p^i}$ to get $\beta = \beta^{p^{j-i}}$. Since β is a root of f(D) just like α, it follows from the previous argument that $\beta \neq \beta^{p^{j-i}}$ for $1 \leq j-i \leq m-1$. Thus $\alpha^{p^i} \neq \alpha^{p^j}$ for $0 \leq i < j \leq m-1$. Since f(D) has m roots, they must be $\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{m-1}}$.

For $GF(2^4)$ in the preceding example, if α is a root of D^4+D+1, then α^2, α^4, and α^8 are the other roots . The roots of D^4+D^3+1 are then $\alpha^{-1} = \alpha^{14}$, $\alpha^{-2} = \alpha^{13}$, $\alpha^{-4} = \alpha^{11}$, and $\alpha^{-8} = \alpha^7$.