Section 1: Fields

Let us begin with the definition of a field.

A field F is a set with two operation, $+$, $\cdot$, with the usual properties as follows. Note that, like everyone else, we often abbreviate $a \cdot b$ as ab.

1) For all $a, b, c \in F$, $(a + (b + c)) = ((a + b) + c)$ and $(a(bc)) = ((ab)c)$. (Associativity of both operations)

2) For all $a, b \in F$, $a + b = b + a$ and $ab = ba$. (Commutativity of both operations)

3) F has an element $0 \in F$ such that $0 + a = a$ for all $a \in F$ (Zero element).

4) For all $a \in F$, there is an element $-a \in F$ such that $a + -a = 0$. (Additive inverses)

5) There is an element $1 \in F$ such that $a \cdot 1 = a$ for all $a \in F$. (Unit element)

6) For all nonzero $a \in F$, there is an element $1/a \in F$ such that $a(1/a) = 1$. (Multiplicative inverses)

7) For all $a, b, c \in F$, $a(b + c) = ab + ac$. (Distributive Law)

8) $0 \neq 1$

Let me make some remarks about this definition. First of all, it has all the usual consequences. For example $0 \cdot a = 0$ for all $a \in F$, and $-(ab) = (-a)b$, and many others.

Our particular interest is FINITE fields. That is, fields where F is a finite set. There is one immediate example, namely $F = \{0, 1\}$ where $1 + 1 = 0$ and all other operations are forced by the field properties. As another example, let p be a prime in the integers $\mathbb{Z}$. We consider $\mathbb{Z}/p\mathbb{Z}$ which we define as follows. As a set $\mathbb{Z}/p\mathbb{Z} = \{\bar{0}, \bar{1}, \ldots \bar{p-1}\}$. We define $\bar{i} + \bar{j} = \bar{k}$ and $\bar{i} \cdot \bar{j} = \bar{k}$ if k is the remainder when $i + j$ or ij is divided by p.

Theorem 1.1. *$\mathbb{Z}/p\mathbb{Z}$ is a field.*

Proof. Properties 1), 2), 3), 5), 7) and 8) are clear, once you take $\bar{0}$ as the zero element and $\bar{1}$ as the multiplicative unit. Property 4) is clear once you take $-\bar{i} = \bar{p-i}$. (Note that none of these relied on the fact that p is prime). As for 6), if $1 \leq a \leq p - 1$, then the fact that p is prime implies that a, p are relatively prime. From Euclid's algorithm it follows that there are integers x, y such that $xa + yp = 1$. Let r be the remainder when you divide x by p. That is, write $x = np + r$ where $0 \leq r < p$. Then $ra + (y + an)p = 1$. It follows that $\bar{r} \cdot \bar{a} = \bar{1}$, proving 6). ∎

In particular, the field $\{0, 1\}$ mentioned above is just $\mathbb{Z}/2\mathbb{Z}$.

Section 2: Linear algebra

For us, the point about defining fields is that is exactly what you need for the "scalars" in order to make some standard facts from linear algebra true more generally. Perhaps, when you studied linear algebra, you only encountered the real or rational field (or the complexes). We will VERY BRIEFLY point out the any field will do.

Recall that a vector space was a set V with an addition operation, and scalar multiplication. Our vector spaces are just the same, except the "scalars" can come from an arbitrary but fixed field. To be precise, let F be a field. An F vector space V is a set with two "operations". We call the elements of V the "vectors" and the elements of F the "scalars". The first is **vector addition** and has the form $v, w \to v + w$. The second is **scalar multiplication** and has the form $\alpha, v \to \alpha v$. Note that if all you know is that V is a vector space, you CANNOT MULTIPLY VECTORS. You also CANNOT ADD A VECTOR and a SCALAR. To be a vector space these two operations must satisfy:

1) For all $u, v, w \in V$, $(u + (v + w)) = ((u + v) + w)$ and $u + v = v + u$. (Associativity and commutativity of vector addition)

2) There is a $0 \in V$ such that $0 + v = v$ for all $v \in V$.

3) For all $v \in V$, there is a $-v \in V$ such that $v + -v = 0$.

4) For all $\alpha, \beta \in F$, $v \in V$, $(\alpha\beta)v = \alpha(\beta v)$. (Associativity of scalar multiplication)

5) $1 \cdot v = v$ all $v \in V$.

6) For all $\alpha, \beta \in F$, $v \in V$, $(\alpha + \beta)v = \alpha v + \beta v$. (Distributivity I)

7) For all $\alpha \in F$, $v, w \in V$, $\alpha(v + w) = \alpha v + \alpha w$.

Let me again make a few comments. First of all, there are symbols with two meanings in this set up. For example, there are TWO zeroes, one in F and one in V. This goes along with the fact that there are two "additions", one in F and one in V. That is why there are two distributivity laws. On the other hand, only F has a "1".

There is one vector space that is enormously important to us. Let F be a field, and F^n the set of "n-tuples" $(\alpha_1, \ldots, \alpha_n)$ where the α_i are allowed to be any elements of F. Then vector addition is "component-wise". That is, $(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$. Note that this definition makes sense because on the right the only additions are in the field, which we presumably already know about. Scalar multiplication is defined by setting $\alpha(\alpha_1, \ldots, \alpha_n) = (\alpha\alpha_1, \ldots, \alpha\alpha_n)$. That the above seven rules hold is pretty easy. Note that $(0, \ldots, 0)$ is the zero element of F^n and $-(\alpha_1, \ldots, \alpha_n) = (-\alpha_1, \ldots, -\alpha_n)$.

Let us be even more specific. If $F = \{0, 1\}$ is the field of 2 elements, then F^n is a "bit stream" of length n. Since real life data transmission can be modeled as a bit stream, vector spaces over F are all of a sudden important in communication.

Most undergraduates taking this course have had linear algebra, and I am going to assume this. However, of these linear algebra courses assumed the scalars were the real, and perhaps rational and perhaps complex fields. The point of what I am about to say is that "it all works over any field", but "all" here does not include any linear algebra associated with "length" or "positive". For example, in linear algebra you define a **subspace** $W \subset V$ of a vector space V to be a subset closed under addition and scalar multiplication. We do the same here for any field of scalars. If $W \subset V$ is such a subspace, recall that a **basis** of W is a set $v_1, \ldots, v_m \in W$ such that every element of W can be expressed as a sum $\alpha_1 v_1 \oplus \ldots \oplus \alpha_m v_m$ and in such

an expression the α_i's are unique. The key fact about bases we need, and which you proved in linear algebra, is that:

Theorem 2.1. *Let $V = F^n$ and let $W \subset V$ be a subspace. Then W has a basis and any two bases of W have the same number of elements.*

We call the number of elements in a basis the **dimension** of W. The proof of the above theorem, which you saw in linear algebra, ONLY USED FIELD properties. For example, recall that the **span** of a set of vectors $w_1, \ldots, w_s$ is the set of elements of the form $\alpha_1 w_1 + \ldots + \alpha_s w_s$. You can check that the span of a set of vectors form a subspace. Recall that a set of vectors $w_1, \ldots, w_s$ is called **linearly independent** if and only if the only time $\alpha_1 w_1 + \ldots \alpha_s w_s = 0$ is when all the $\alpha_i = 0$. Without too much trouble one shows that basis of W is precisely a linearly independent set that spans W. A key lemma in the proof of the above theorem, whose proof we give as an example, is the following.

Lemma 2.2. *Suppose $w_1, \ldots, w_s$ are linearly independent and w is not in the span of the w_i. Then $w_1, \ldots, w_s, w$ are linearly independent.*

Proof. Suppose not. Then there are $\beta, \beta_i \in F$, not all zero, such that $\beta_1 w_1 + \ldots \beta_s w_s + \beta w = 0$. Since the w_i are linearly independent, β cannot be 0. Thus $w = \beta^{-1}(-\beta_1)w_1 + \ldots + \beta^{-1}(-\beta_s)w_s$, contradicting the assumption. ∎

In the above argument, note the need to invert β, which is precisely why vector spaces are assumed to have fields for the set of scalars. We assert that the full proof of the Theorem also only uses the field properties.

Another bit of linear algebra we need to consider over any field is the subject of linear maps. If V, V' are vector spaces over the field F, then a **linear** map $T : V \to V'$ is a map such that for all $\alpha \in F$, $v, w \in V$, $T(\alpha v) = \alpha T(v)$ and $T(v + w) = T(v) + T(w)$. A linear map defines two important subspaces. We call the **kernel** or **null space** of T the set $\{v \in V | T(v) = 0\}$. We call the **image** or **range** of T the subspace of V' of elements of the form $T(w)$ for $w \in V$. It is easy to verify that both these subsets are subspaces of V, V' respectively. The key result from linear algebra we need follows. Again, the proof only needs the field properties.

Theorem 2.3. *Suppose V, V' are vector spaces over F and $T : V \to V'$ is a linear map. Then the dimension of the kernel of T, plus the dimension of the range of T, equals the dimension of V.*

Finally, we recall something about matrices. An $n \times m$ matrix over the field F is a rectangular array of elements of F with n rows and m columns. We use the notation that if $A = (a_{ij})$ is an $n \times m$ matrix, then this means i has range 1 to n, j has range 1 to m, and a_{ij} is the entry of the matrix in the i row and j column.

The set of all $n \times m$ matrices forms an F vector space under the operations $(a_{ij}) + (b_{ij}) = ((a_{ij} + b_{ij})_{ij})$ and $\alpha(a_{ij}) = ((\alpha a_{ij})_{ij})$. In particular, F^n can be viewed as the space of $n \times 1$ matrices or synonymously the space of row vectors of length n. We set e_i to be the row vector with 1 in the i place and zeroes everywhere else. It is easy to see that the e_i are a basis for F^n.

If T is a linear map $F^n \to F^m$, then defines the matrix A_T associated to T to be the matrix where the i row is the row vector $T(e_i)$. Since the e_i are a basis, A_T completely determines T, and any $n \times m$ matrix is the matrix of a linear map. If $T : F^n \to F^m$ and $S : V^m \to V^r$ are linear maps, it is easy to see that the composition $S \circ T : V^n \to V^r$ is linear. One can also compute that the matrix of $S \circ T$ is just A_{TS} where TS is the so called product matrix. One can compute that the i, k entry of TS is $a_{i1}b_{1k} + \ldots + a_{im}b_{mk}$. Using the association with linear maps, one can prove $(ST)U = S(TU)$ for all matrices S, T, U where the product is defined. Finally, if v is a row vector, and A_T is the matrix for T, one can compute that $T(v) = vA_T$.

The point of all of the above is not to compress a semester of linear algebra into a few pages. The point is to emphasize that all of this linear algebra only needed the scalars to form a field. In particular, all of your linear algebra above applies to vector spaces over finite fields.

Section 3: Polynomials

From high school algebra, you should be familiar with adding, multiplying, and generally manipulating polynomials. One again, as in the last section, the point of this section is that it also all works for any field, in particular any finite field. If F is a field, a polynomial over f is a formal expression of the form $f_0 + f_1x + \ldots f_nx^n$ where the $f_i \in F$. The f_i are called the coefficients. Sometimes, when $f_i = 0$, we drop the f_ix^i in our expression of a polynomial. Note that a polynomial is a formal expression, NOT a function. Two polynomials are equal if and only if all the coefficients are equal. If F is the field $\{0, 1\}$, the FUNCTIONS $x \to x$ and $x \to x^2$ are the same function on F, but the POLYNOMIALS x and x^2 are different. We write $F[x]$ for the set of all polynomials with coefficients in F.

One defines the addition of two polynomials in the obvious way, namely, that $a_0 + a_1x + \ldots + a_nx^n + b_0 + \ldots + b_nx^n = (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_n + b_n)x^n$. Note a convention we often use, namely, by adding zero coefficients, a finite set of polynomials can always be written in the form $a_0 + a_1x + \ldots + a_nx^n$ for one n. The degree of the polynomial $a(x) = a_0 + a_1x + \ldots + a_nx^n$ is the largest i such that $a_i \neq 0$. As a matter of convention, we do not define the degree of the 0 polynomial. Let d be the degree of $a(x)$ above. The corresponding term a_dx^d is called the leading term of $a(x)$ and a_d is the leading coefficient. Thus a polynomial $a(x)$ can always be written as $a_0 + \ldots a_dx^d$ where d is the degree. We say $a(x)$ is monic if the leading coefficient is $1 \in F$.

One can also multiply $F[x]$ polynomials in the usual way, namely, that $(a_0 + a_1x + \ldots + a_nx^n)(b_0 + \ldots + b_mx^m) = a_0b_0 + (a_1b_0 + a_0b_1)x + \ldots + (a_0b_i + a_1b_{i-1} + \ldots + a_ib_0)x^i + \ldots + (a_nb_m)x^{n+m}$. There are a series of facts one can check for these $F[x]$ polynomials that are extensions of the facts about polynomials you know from high school. For example, the addition and multiplication operations are associative and commutative as in 1) and 2) of section one. There is a zero (the polynomial where all coefficients are 0), and additive inverses (3) and 4)). The polynomial "$1 \neq 0$" is the multiplicative unit as in 5), 8), and the distributive law holds (7)).

What fails to hold is 6). In fact, it is easy to see that the leading term of $a(x)b(x)$ is just $a_n b_m x^{n+m}$ where $a_n x^n$, $b_m x^m$ are the leading terms of $a(x)$ and $b(x)$ respectively. In particular, the degree of $a(x)b(x)$ is the sum of the degrees of $a(x)$ and $b(x)$. Note also that the product of two monic polynomials is again monic. Finally, it is clear that if $a(x)b(x) = 1$ then $a(x)$ must be a nonzero polynomial of degree 0, i.e. a nonzero "constant". On the other hand the same discussion shows that if $a(x), b(x)$ are nonzero, then $a(x)b(x) \neq 0$. In particular, we have cancellation. If $a(x)b(x) = a(x)c(x)$ and $a(x) \neq 0$, then $b(x) = c(x)$. While on the subject, note that the degree of a sum $a(x)+b(x)$ is less than or equal to the maximum of degrees. In fact, if you think about it, the degree IS the maximum if $a(x), b(x)$ have different degrees, but if they have the same degree there can be cancellation and the degree can be strictly smaller that maximum (in this case common) degree.

Even though polynomials are not functions, they can be used to define functions in the usual way. That is, if $a_n x^n + \ldots + a_1 x + a_0 \in f(x) \in F[x]$, and $a \in F$, we define $f(a) = a_n a^n + \ldots + a_1 a + a_0 \in F$. As usual, we say a is a root of $f(x)$ if $f(a) = 0$. Note that if $h(x) = f(x)g(x)$ and $k(x) = f(x) + g(x)$, then $h(a) = f(a)g(a)$ and $k(a) = f(a) + g(a)$.

The main theme of this section is that $F[x]$ has properties that make it a lot like the integers. The very key and very important reason for this is that one can divide polynomials and get quotients and remainders. Stated as a (overly) formal theorem the fact is:

Theorem 3.1. *Let $f(x), g(x) \in F[x]$ with $g(x) \neq 0$. Then there are unique polynomials $q(x), r(x)$ such that $f(x) = q(x)g(x) + r(x)$ where $r(x) = 0$ or the degree of $r(x)$ is less than the degree of $g(x)$.*

Of course, in the above, $q(x)$ is called the quotient and $r(x)$ is called the remainder.

Proof. We begin by showing $q(x), r(x)$ exist and then prove existence. Let $r'(x)$ be a polynomial of smallest degree which can be written in the form $f(x) - q'(x)g(x)$. Assume $r'(x)$ is nonzero of degree strictly larger than that of $g(x)$. In fact, let $g_n x^n$ be the leading term of $g(x)$ and $r_{n+d} x^{n+d}$ the leading term of $r'(x)$. Note that in $f(x) - q'(x)g(x) - g_n^{-1} r_{n+d} x^d g(x)$ the x^{n+d} term is canceled and so this is a polynomial of smaller degree, a contradiction. This proves the existence part.

To prove uniqueness, suppose $f(x) = q_1(x)g(x) + r_1(x) = q_2(x)g(x) + r_2(x)$ where both $r_i(x)$ are as in the theorem. Then $(q_1(x) - q_2(x))g(x) = (r_2(x) - r_1(x))$. Now if right side is nonzero, it has degree less than that of $g(x)$. But then the left side has degree greater or equal to that of $g(x)$, a contradiction. Thus $r_1(x) = r_2(x)$ and so $q_1(x) = q_2(x)$. ∎

Let me make two remarks about the above argument. First, the "existence" part of the proof is really the algorithm for dividing polynomials. That is, the terms of $q(x)$ are successively determined by the need to cancel terms in $f(x) - q'(x)g(x)$, starting with $f(x)$ itself. The remainder is what remains when the process has to stop. Secondly, note that, once again, the field property that g_n^{-1} exists is crucial in the proof.

There is a special case of this division algorithm that is worth pointing out.

Corollary 3.2. *Suppose $a \in F$. Then $f(x) = q(x)(x-a) = f(a)$.*

Proof. Using the division algorithm we have $f(x) = q(x)(x-a)+r(x)$ where $r(x)$ has degree less than 1, i.e. is zero or has degree 0. Either way, $f(x) = q(x)(x-a)+r$. If we substitute a for x, we have $f(a) = q(a)0 + r$. ∎

One can repeat the division algorithm to perform what is called Euclid's algorithm. To be precise, let $f(x), g(x)$ be nonzero polynomials. Use the above theorem to write $f(x) = q(x)g(x) + r_1(x)$. For the next step, we "slide" the polynomials and write $g(x) = q_1(x)r_1(x) + r_2(x)$. The general step of this algorithm is that if $r_{i-2} = q_{i-1}(x)r_{i-1}(x) + r_i(x)$ is the previous step, then we use theorem and form $r_{i-1}(x) = q_i(x)r_i(x) + r_{i+1}(x)$. Note that the degrees of the r_i are decreasing, so the process must stop. In fact, the process must stop when some remainder, say $r_n(x)$, is zero (because then we CANNOT form the next quotient and remainder.

Theorem 3.3. *Suppose $f(x), g(x), r_i(x)$ are as above and n is the first integer where $r_n(x) = 0$. Then $r_{n-1}(x)$ is the greatest common divisor of $f(x)$ and $g(x)$. That is, $r_{n-1}(x)|f(x)$, $r_{n-1}(x)|g(x)$ and if $d(x)|f(x)$, $d(x)|g(x)$ then $d(x)|r_{n-1}(x)$.*

Proof. The last equation is $r_{n-2} = q_{n-1}(x)r_{n-1}(x)$ because $r_n(x) = 0$. Thus $r_{n-1}(x)|r_{n-2}(x)$. From the equation $r_{n-3}(x) = q_{n-2}(x)r_{n-2}(x) + r_{n-1}(x)$ we conclude that $r_{n-1}(x)|r_{n-3}(x)$. Arguing up the list of equations we conclude $r_{n-1}(x)|g(x)$ and then $r_{n-1}(x)|f(x)$.

If $d(x)|f(x)$ and $d(x)|g(x)$ then using the first equation we conclude $d(x)|r_1(x)$. Arguing down the list, we conclude $d(x)|r_{n-1}(x)$. ∎

Note that by its defining property, the gcd is unique up to multiplication by a nonzero constant (i.e. degree 0 polynomial). This is because if $d(x)$ and $e(x)$ are both gcd's of $f(x), g(x)$, then $e(x)|d(x)$ and $d(x)|e(x)$. Thus $e(x), d(x)$ have equal degrees. Thus if $e(x) = q(x)d(x)$, then $q(x)$ must be nonzero and of degree 0. The above method of constructing the gcd has the following key consequence.

Corollary 3.4. *Suppose $d(x)$ is the gcd of $f(x)$, $g(x)$, both nonzero polynomials. Then there are polynomials $a(x), b(x)$ such that $d(x) = a(x)f(x) + b(x)g(x)$.*

Proof. If $d(x) = a(x)f(x) + b(x)g(x)$ for some $a(x), b(x)$ we say $d(x)$ is a linear combination of $f(x), g(x)$. To prove the corollary, may as well assume $d(x)$ is the $r_{n-1}(x)$ from the above theorem. Then $r_{n-1}(x) = r_{n-3}(x) - q_{n-2}(x)r_{n-2}(x)$. Substituting $r_{n-2}(x) = r_{n-4}(x) - q_{n-3}(x)r_{n-3}(x)$ into the first equation, we get $r_{n-1}(x)$ is a linear combination of $r_{n-4}(x)$ and $r_{n-3}(x)$. Again, repeating this argument up the equation list we get the result we want. ∎

The above two results are of enormous importance, implying that polynomials behave a lot like integers. In particular, we next observe that there is unique factorization of polynomials into primes, but it is traditional to use the term "irreducible" instead of prime. To be formal, a polynomial, $p(x)$, of positive degree is **irreducible** if and only if its only divisors have degree 0 or degree equal to that of $p(x)$.

Theorem 3.5. *Every positive degree polynomial, $f(x)$, can be factor into a product of irreducible polynomials. If $f(x) = p_1(x)\dots p_n(x) = q_1(x)\dots q_m(x)$ are two factorizations into irreducibles, then $n = m$ and one can reorder the $q_i(x)$'s such that $p_i(x) = c_i q_i(x)$ where c_i is a nonzero constant.*

Proof. We first prove the factorization exists by induction on the degree of $f(x)$. If this degree is one, $f(x)$ is automatically irreducible and we are done. For general $f(x)$, we are done it it is irreducible, so write $f(x) = g_1(x)g_2(x)$ where both $g_i(x)$ have positive degrees smaller than that of $f(x)$. By induction both the $g_i(x)$ can be written as a product of irreducibles, and combining these we get such a product for $f(x)$.

To prove uniqueness, we first note the following.

Lemma 3.6. *Suppose $p(x)$ is an irreducible polynomial and $p(x)|a(x)b(x)$. Then $p(x)|a(x)$ or $p(x)|b(x)$.*

Proof. If $p(x)$ does not divide $a(x)$, then they must be relatively prime (because $p(x)$ has so few divisors). Thus there are $c(x), d(x)$ such that $c(x)p(x) + d(x)a(x) = 1$. Multiplying by $b(x)$, we have $b(x) = b(x)c(x)p(x) + d(x)a(x)b(x)$. Since $p(x)|a(x)b(x)$, we have $p(x)|b(x)$. ∎

Returning to the proof of the theorem, suppose $f(x) = p_1(x)\dots p_n(x) = q_1(x)\dots q_m(x)$ are two such factorizations. By the lemma, $p_1(x)|q_i(x)$ for some i. Since $q_i(x)$ is irreducible, $p_1(x) = cq_i(x)$ for a constant c. Reordering, we can assume $i = 1$. Canceling, we can finish the proof by induction. ∎

Let $g(x) \in F[x]$ be a polynomial of positive degree and let $F[x]/(g(x))$ be the set of polynomials of degree strictly less than that of $g(x)$. This set has an obvious addition, and define the product of $f_1(x)f_2(x)$ as the remainder of the usual product $f_1(x)f_2(x)$ upon division by $g(x)$. We note (but not prove) that $F[x]/(g(x))$ satisfies the field properties 1) through 8) excluding 6). However, we also note:

Theorem 3.7. *Suppose $g(x) \in F[x]$ is irreducible. Then $F[x]/(g(x))$ is a field.*

Proof. Suppose $f(x) \in F[x]/(g(x))$ is nonzero. Since $g(x)$ is irreducible, and $f(x)$ has smaller degree, then can have no common divisors of degree bigger than 0. Thus $a(x)f(x)+b(x)g(x) = 1$ for some $a(x), b(x)$. If $a(x) = q(x)g(x)+a_1(x)$ where $a_1(x)$ has degree smaller than that of $g(x)$, then $a_1(x)f(x) - 1 = (a(x) - q(x)g(x))f(x) - 1 = b(x)g(x) - q(x)f(x)g(x)$. Thus $a_1(x)f(x)$ has remainder 1 when divided by $g(x)$, proving 6) for $F[x]/(g(x))$. ∎

Let us observe that F can be identified with the set of constant polynomials (i.e. 0 plus degree 0 polynomials), and so we can view $F \subset F[x]/(g(x))$. As a subset of $F[x]/(g(x))$, F is a so called **subfield**. That is, F is a subset, contains 0,1, is closed under addition, multiplication, and with respect to 0,1, and these operations forms a field.

Let me explain some of this notation a bit. The study of $F[x]$ often involves subsets calls ideals. We define an ideal $I \subset F[x]$ to be a subset closed under addition and the property that if $f(x) \in F[x]$, $g(x) \in F[x]$, then $f(x)g(x) \in F[x]$.

Another way of viewing 3.4 above is that all ideals of $F[x]$ are determined by a single element. More precisely, we say $g(x)$ generates an ideal I if I is precisely the set of multiples of $g(x)$. We write $I = (g(x))$ in this case.

Corollary 3.8. *Every ideal of $F[x]$ has the form $(g(x))$.*

Proof. Let I be an ideal, and choose $g(x)$ an element of the smallest positive degree in I. Clearly, from the definition of an ideal, $(g(x)) \subset I$. We claim $I = (g(x))$, proving the corollary. To this end, let $f(x) \in I$. Write $f(x) = q(x)g(x) + r(x)$ as in the division algorithm. If $r(x)$ is nonzero, then $r(x)$ has degree smaller than that of $g(x)$, and $r(x) = f(x) - q(x)g(x)$ implies $r(x) \in I$. This contradiction proves $r(x) = 0$ and so $f(x) \in (g(x))$. ∎

Theorem 3.1 says that every element of $F[x]$ can be written in a unique set of the form $h(x) + ((g(x))$ where $h(x)$ is zero or has degree less than that of $g(x)$. Then $F[x]/(g(x))$ can be viewed as the set of these subsets.
Add to previous notes:

The part of the above theorem 3.5 that talks about "differing by a nonzero constant" is clumsy to say if clear in meaning. There is a way of avoiding this, by concentrating on monic polynomials. Note that if $p(x)$, $q(x)$ and monic polynomials, and $p(x) = cq(x)$, then $c = 1$ by looking at leading terms. Furthermore, if $f(x)$ is any polynomial, then $f(x) = cg(x)$ where $g(x)$ is monic. Thus we can rewrite 3.6 as:

Corollary 3.6.3. *Every positive degree monic polynomial, $f(x)$, can be factored into a product of monic irreducible polynomials. If $f(x) = p_1(x) \dots p_n(x) = q_1(x) \dots q_m(x)$ are two factorizations into monic irreducibles, then $n = m$ and one can reorder the $q_i(x)$'s such that $p_i(x) = q_i(x)$.*

We can use 3.5 to prove a result that is well known to everyone, though perhaps it is less clear that it is also true over a finite field.

Corollary 3.6.5. *Suppose $f(x) \in F[x]$ is a polynomial of degree n. Then $f(x)$ has, at most, n roots in $F[x]$.*

Proof. We can assume $f(x)$ is monic for convenience. Suppose $a_1, \dots, a_{n+1}$ are all roots of $f(x)$ and all distinct. Then all the $x - a_i$ are distinct prime divisors of $f(x)$. Corollary 3.6.3 implies that $(x - a_1) \dots (x - a_{n+1})$ appears in the decomposition of $f(x)$ into irreducible factors. Since this product has degree $n + 1$, this is impossible. ∎

Section 4: Finite fields

The key concept of this whole course is that of a finite field, that is, a field which as a set has finitely many elements. We have already seen examples of this structures, namely, the fields $\mathbb{Z}/p\mathbb{Z}$ for a prime p. The purpose of this section is to study the other finite fields. Of course, 3.7 above will be a major tool in constructing these objects.

To begin with, let F be a finite field. In F, there is a **prime** subfield that has the minimum number of elements. That is, we start with $0, 1 \in F$ and only add what is needed to make a field. To be precise, we consider the subset of F consisting of $0, 1, 1+1, 1+1+1, \ldots$. Note that this set is closed under addition (by definition) and multiplication (by the distributive law). We write $n \cdot 1$ to mean the sum $1 + 1 + \ldots + 1$ where "1" appears n times. Later on, we will abbreviate $n \cdot 1$ by n, but note that this convention hides an important distinction. There is the integer n, and the element $n \in F$. These are different objects.

Since F is finite, the list $0, 1, 1+1, \ldots$ must have duplications. Let $n \cdot 1 = (n+r) \cdot 1$ be the first duplication. By cancellation, we have $0 = r \cdot 1$ and so the first duplication is in fact the second appearance of 0 in the list. We define the **characteristic** of F to be this least positive r such that $r \cdot 1 = 0$.

Lemma 4.1. *The characteristic of F is a prime.*

Proof. Suppose r is the characteristic of F, and $r = ab$ for $a, b \in \mathbb{Z}$ and both bigger than $1 \in \mathbb{Z}$. Then the distributive law easily shows $0 = r \cdot 1 = (a \cdot 1)(b \cdot 1)$. But F has no zero divisors, so $a \cdot 1 = 0$ or $b \cdot 1 = 0$, which contradicts the minimality of r. ∎

All fields have a characteristic, but many fields (like the rationals) have NO positive r such that $r \cdot 1 = 0$. For perverse reasons, such fields are said to have characteristic 0. Let F have characteristic p and set $F_p \subset F$ to be the subset of elements of the form $n \cdot 1$. It is clear that F_p consists of the p distinct elements $r \cdot 1$ where $0 \leq r < p$. We observed above that F_p is closed under addition and multiplication. Since $-(r \cdot 1) = (p - r) \cdot 1$, F_p is also closed under additive inverses. If $0 < a < p$, then a, p have no common divisor and so there are $m, n \in \mathbb{Z}$ such that $ma + pn = 1$. It follows that $(m \cdot 1)(a \cdot 1) = (1 - pn) \cdot 1 = 1$. Hence F_p is a field.

As a matter of fact, F_p is a field we have seen before. Recall we defined the field $\mathbb{Z}/p\mathbb{Z}$ thought of as the set of integers $0 \leq r < p$. Define $\phi : \mathbb{Z}/p\mathbb{Z} \to F_p$ by setting $\phi(r) = r \cdot 1$. It is immediate that ϕ is a bijection, and that $\phi(a+b) = \phi(a) + \phi(b), \phi(ab) = \phi(a)\phi(b), \phi(0) = 0, \phi(1) = 1$. Such a ϕ is called an **isomorphism** of fields. One major goal of this section is that, in fact, all fields of the same finite order are isomorphic.

Now the fact that $F_p \subset F$ puts a severe restriction on the possible orders for F.

Lemma 4.2. *Suppose F is a finite field of characteristic p, where p is of course a prime. Then there is an integer n such that F has p^n elements.*

Proof. This argument is an example of a frequent proof strategy in mathematics, namely, forgetting structure. F is a field, but more weakly F is a vector space over F_p. That is, we use the addition in F to define vector addition, and we limit the multiplication in F to products of the form αf, where $\alpha \in F_p$ and $f \in F$. This limited product obeys the rules of scalar multiplication in a vector space, where F_p are the scalars.

In this way F is a vector space over F_p. In particular, F has a basis $v_1, \ldots v_n$ over F_p. Since any element of F can be uniquely written in the form $\alpha_1 v_1 + \ldots \alpha_n v_n$, and there p^n choices of the α_i, we are done. ∎

A huge chunk of the rest of this section can be summarized as proving that for ANY prime power order $q = p^n$, there is a finite field of order q, and all such fields are isomorphic.

Section 5: Finite fields continued

Before we build finite fields, let's discuss them more, and in particular discuss their multiplicative "group". For any field F, the set $F^* = F - \{0\}$ is a **abelian group**. under the operation of multiplication. That is, F^* has an associative commutative operation (multiplication), a unit ("1"), and inverses for every element.

Let us begin with:

Lemma 5.1. *Suppose F is a finite field of order $q = p^r$ and $a \in F^*$. Then $a^{q-1} = 1$.*

Proof. Let $a_1, \ldots, a_{q-1}$ be the list of elements of F^*. Consider the list $aa_1, \ldots, aa_{q-1}$. Note that no two elements on the second list are equal, because $aa_i = aa_j$ implies $a_i = a_j$. Thus the second list is also all the elements of F^*. Since multiplication is commutative and associative, $a_1 \ldots a_{q-1} = aa_1 aa_2 \ldots aa_{q-1} = a^{q-1} a_1 \ldots a_{q-1}$. Thus, by canceling, $a^{q-1} = 1$. ∎

In fact, we can make 5.1 much more precise and say something about the smallest n for which $a^n = 1$ for all a in F. To this end, define the **order** of $a \in F^*$ to be the least positive n such that $a^n = 1$. Note that if a has order n and $a^m = 1$, then we can write $m = qn + r$, $0 \leq r < n$, and so $a^r = 1$ which is a contradiction unless $r = 0$. That is, n must divide m.

We can say more about orders. Suppose $a, b \in F^*$ have orders m, n respectively and $(m, n) = 1$. Consider $c = ab$. Suppose $c^r = 1$. Then $1 = c^{rn} = a^{rn} b^{rn} = a^{rn}$. Thus $m | rn$. Since $(m, n) = 1$, $m | r$. Similarly, $n | r$ and so $mn | r$. That is, mn is the order of c.

One might hope that a more general result is possible, but it is not exactly what one might think.

Lemma 5.2. *Suppose $a, b \in F^*$ have orders m, n respectively. Then there is an element $c \in F^*$ with order the least common multiple $[m, n]$.*

Proof. Write the prime factorizations $m = p_1^{r_1} \ldots p_s^{r_s}$ and $n = p_1^{t_1} \ldots p_s^{t_s}$ where we can assume the same primes appear by allowing 0 exponents. Let m' the the product of those p^{r_i} where $r_i \geq t_i$ and n' the product of the $p_i^{t_i}$ where $t_i > r_i$. Since the have no prime factors in common, $(m', n') = 1$. Also, $m'd = m$ and $n'e = n$ for some d, e. It is quite easy to check that a^d, b^e have orders m', n' respectively, and $m'n' = [m, n]$. Thus $c = a^d b^e$ has order $[m, n]$. ∎

By invoking induction, 5.2 shows that there is an integer n and an element $a \in F^*$ such that a has order n and every other order of an element of F^* divides n. In particular, $b^n = 1$ for all $b \in F^*$. In other words, F^* consists of roots of the polynomial $x^n - 1$. This almost shows:

Theorem 5.3. *Let F be a finite field of order q. Then F^* has an element, say a, of order $q-1$. Every element of F^* can be written as a power of a.*

Proof. We observed above there was an element $a \in F^*$ of order n such that every element of F^* is a root of $x^n - 1 = 0$. By 3.6.5 $q-1 \geq n$. Since $a^{q-1} = 1$, $n|q-1$ and $n \leq q-1$. Thus $n = q-1$. As for the second statement, the list $1, a, a^2, \ldots, a^{q-2}$ has no duplicates and thus must consist of all the elements of F^*. ∎

It is sometimes more convenient to add the root 0 to the equation $x^{q-1} - 1$. That is, we observe (with no proof needed):

Lemma 5.4. *Suppose F is a field of order q. Then F consists of roots of the polynomial $x^q - x$.*

Simple as it is, 5.4 will tell us something about subfields of a finite field.

Theorem 5.5. *Suppose K is a finite field of order $q = p^m$ and p is prime. Any subfield $F \subset K$ has order p^r where $r|m$. If $r|m$, there is a UNIQUE subfield of K of order r.*

Proof. Write $q = p^r$. Since K is a vector space over F, K must have order q^t where t is the dimension of K over F. Thus $p^n = q^t = p^{rt}$, implying $n = rt$. Note that F consists precisely of the roots of $x^q - x$, since F has q elements and this polynomial can have no more than q roots. Thus if F exists, it is unique.

Thus to finish the proof of the theorem we must show that if $r|n$, K has a subfield of order p^r. The uniqueness argument above shows how we will proceed, namely, we will choose F to be the roots of the polynomial $x^{p^r} - x$. Once again, write $q = p^r$ and also write $q' = p^n$ and $t = n/r$. Consider the integers $q'-1$ and $q-1$. Since $q' = q^t$, we have $q' - 1 = (q-1)(1 + q + q^2 + \ldots + q^{t-1})$. Write $m = (q'-1)/(q-1)$. Let $b \in K^*$ have order $q'-1$. Consider the $q-1$ elements $b^m, b^{2m}, \ldots, b^{(q-1)m} = 1$. Then these elements, along with 0, are all roots of $x^q - x$ and, once again, there is no room for K to have any other roots of this polynomial. Define F to be the set of roots of $x^q - x$.

It suffices to show F is a subfield. It clearly contains 0,1. If $a^q = a$, and $b^q = b$, then it is obvious that $(ab)^q = a^q b^q = ab$. That is F is closed under multiplication. In almost the same way, F^* is closed under inverses. It can be somewhat surprising that F is closed under addition, but this will follow from what I call the "stupid" binomial theorem.

Lemma 5.6. *Suppose L is a field of characteristic p and $x, y \in L$. Then $(x+y)^p = x^p + y^p$. If q is a power of p, then $(x+y)^q = x^q + y^q$.*

Proof. The second statement follows by an easy induction from the first. As for the first, the usual proof (in any characteristic) shows that $(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^r y^{n-r}$ where $\binom{n}{r} = n!/r!(n-r)!$ is an integer. If $n = p$ is prime and $0 < r < p$, then the numerator of $\binom{n}{r}$ is divisible by p but the denominator is not. That is, as an integer, $\binom{p}{r}$ is divisible by p when $0 < r < p$. In a field of characteristic p, it follows that $\binom{p}{r} = 0$ for the same r, which proves the Lemma. ∎

5.5 shops how to build finite subfields of a finite field, but we need the reverse. That is, we need to build our fields from below. The most direct way would be to show that if F is any finite field, and $n > 1$ is an integer, then $F[x]$ has an irreducible polynomial of degree n. This would be convenient because we could choose $F = \mathbb{Z}/p\mathbb{Z}$, and $f(x)$ irreducible of degree n, and construct $K = F[x]/(f(x))$, which has order p^n.

The above result is true, but we do not have a direct route to proving it. Instead, we will construct K first, and THEN show there is an irreducible polynomial using K. The whole thing will be woefully nonpractical, in the sense that the actual irreducible polynomials whose existence we prove will not be explicitly displayed.

If, however, n is prime and F is a finite field, it is easier to construct irreducible polynomials of degree n.

Lemma 5.7. *Let F be a finite field of order q and let ℓ be a prime (NOT necessarily equal to the characteristic).*

a) In $F[x]$, $x^q - x = \prod_{\theta \in F}(x - \theta)$.

b) $F[x]$ contains an irreducible polynomial of degree ℓ.

Proof. All the elements of F are roots of $x^q - x$ so $x^q - x$ is divisible by $g(x) = \prod_{\theta \in F}(x - \theta)$ divides $f(x)$. But $g(x)$ is also monic, and also has degree q, and we have $g(x) = x^q - x$.

Next consider the polynomial $f(x) = x^{q^\ell} - x$. Note that $q^\ell - 1 = (q-1)(1 + q + q^2 + \ldots + q^{\ell-1}$ which we write as $(q-1)r$. Note that $x^{q^l - 1} - 1 = (x^{q-1} - 1)(1 + x^{q-1} + \ldots x^{(r-1)(q-1)})$. which we write as $(x^{q-1} - 1)f_1(x)$. Multiplying both sides by x we have $f(x) = (x^q - x)f_1(x)$.

We claim that $f_1(x)$ has no element of F as a root. Suppose $\theta \in F$ were a root of $f_1(x)$. Then $1 + \theta^{q-1} + \theta^{(r-1)(q-1)} = 0$. Since this implies $\theta \neq 0$, we know $\theta^{q-1} = 1$ and so we have $0 = 1 + 1 + \ldots + 1$ where there are r "1"'s on the right side. That is, $0 = r \cdot 1$ in F. But $r - 1$ is divisible by q so this means $0 = 1$, a contradiction. We conclude $f_1(x)$ does not have a root in F.

Suppose $p(x)$ is an irreducible divisor of $f_1(x)$. We will show $p(x)$ has degree ℓ, and thus finish the lemma. We know from the above that $p(x)$ has degree bigger than 1. Write $K = F[x]/(p(x))$ and let $\mu \in K$ be the element with remainder x. Since $p(\mu) = 0$ in K, and $p(x)$ divides $f(x)$, we know that μ is a root of $f(x) = x^{q^\ell} - x$, as is all of F. Thus all the elements gotten from F and μ by addition and multiplication are also roots of $f(x)$, implying that all the elements of K are roots of $f(x)$. It follows that the order of K, say q^s, satisfies $q < q^s \leq q^\ell$ and s is the degree of $p(x)$. Also note that $\mu^{q^s} = \mu$.

If $s < \ell$, there are integers a, b such that $as + b\ell = 1$. But then $\mu^q = \mu^{q^{as+b\ell}} = \mu^{q^{as}q^{b\ell}} = (\mu^{q^{as}})^{q^{b\ell}}$. Since $\mu^{q^s} = \mu$, $\mu^{q^{as}} = (\ldots((\mu^{q^s})^{q^s})\ldots)^{q^s} = \mu$, and similarly $\mu^{q^{b\ell}} = \mu$. Thus, $\mu^q = \mu$, implying $\mu \in F$, which is a contradiction. Thus $p(x)$ has degree precisely ℓ. ∎

We have done all the hard work to show:

Theorem 5.8. *Let p be a prime and $n > 0$ an integer. Then there is a finite field of order p^n. More generally, if F is a finite field of order q, there is a field $K \supset F$ such that K has order q^n.*

Proof. The second statement proves the first if we set $F = \mathbb{Z}/p\mathbb{Z}$. We prove the second statement by induction on n. If $n = 1$, $K = F$ is the needed field. Suppose $n = \ell a$ where ℓ is prime. By induction there is a field F of order p^a. By 5.7 there is an irreducible polynomial $f(x) \in F[x]$ of degree ℓ. Set $K = F[x]/(f(x))$. Then K has order $(p^a)^\ell = p^n$. ∎

If the method of 5.8 were the only way to construct finite fields, they would have very complicated structures as being built up step by step. In fact, there is a single polynomial that will define the structure of a K as in 5.8. It's only that we needed the existence of K to prove the existence of such a polynomial. But to make this all clear, it is useful to know about isomorphisms and minimum polynomials, two subjects to which we now turn.

The first subject we wish to deal with is the isomorphisms of fields. We begin with the general definition. We say a map $\phi : K \to L$ between two fields is an **isomorphism** if and only ϕ is one to one, onto, $\phi(0) = 0$, $\phi(1) = 1$, $\phi(a + b) = \phi(a) + \phi(b)$, and $\phi(ab) = \phi(a)\phi(b)$ for all $a, b \in K$. Note that if ϕ is an isomorphism from K to L, then ϕ^{-1} is an isomorphism from L to K. We say the fields K, L are isomorphic if there is such a isomorphism ϕ. The intuitive idea is that isomorphic fields are the "same", since the only differ by the (arbitrary) names of their elements.

The most important tool for constructing isomorphisms is 5.10 to follow. A key part of this result is the so called minimum polynomial. To explain this, let $F \subset K$ be fields (of course this means F is a subfield of K). If $a \in K$, consider a monic polynomial $f(x) \in F[x]$ with minimum degree such that $f(a) = 0$. We call $f(x)$ a minimum polynomial of a over K. Note that such an $f(x)$ may not exist. That is, there may be no polynomial in $F[x]$ with a as a root, in which case we call a transcendental over F. If there are polynomials in $F[x]$ with a as a root, we say a is **algebraic** over F. The next lemma says that $f(x)$ is unique and irreducible, which in particular justifies use of the phrase "the minimum polynomial".

Lemma 5.9. *Let $a \in K \supset F$ be as above and is algebraic over F. Let $f(x)$ be the minimum polynomial of a over F. If $g(x) \in F[x]$ is any polynomial with a as a root, then $f(x)|g(x)$. Also, $f(x)$ is irreducible. If $g(x)$ is monic and irreducible, and has a as a root, then $g(x) = f(x)$. In particular, the minimum polynomial of a is unique.*

Proof. We begin with the first statement. Suppose $g(x)$ is as given. Then $g(x) = q(x)f(x) + r(x)$ where $r(x) = 0$ or $r(x)$ has degree less than that of $f(x)$. But $r(a) = g(a) - q(a)f(a) = 0$, and $f(x)$ has minimum degree, so $r(x) = 0$. Note that the set of $g(x)$ with $g(a) = 0$ form an ideal, and so this part of 5.10 follows from 3.8. We chose to repeat the argument to reinforce the reader's understanding.

As for the second statement, suppose $f(x) = g(x)h(x)$. Then $f(a) = g(a)h(a) = 0$. Thus we may assume a is a root of $g(x)$. But $f(x)$ has minimum degree, so $g(x)$

must have degree equal to that of $f(x)$, and $h(x)$ must have degree 0. This says $f(x)$ is irreducible.

If $g(x)$ is as given. another then $f(x)|g(x)$. Since $g(x)$ is irreducible, $g(x) = cf(x)$ where $c \in F^*$, that is, c is a polynomial of degree 0. Since both $g(x)$ and $f(x)$ are monic, examining leading terms shows that $c = 1$ or $g(x) = f(x)$. The final statement is now clear. ∎

Now we can turn to the isomorphism theorem we want.

Theorem 5.10. *Suppose $F \subset K$ are fields and $a \in K$ is algebraic over F. Let $f(x)$ be the minimum polynomial of a. Then there is an isomorphism $\phi : F[x]/(f(x)) \cong L$ where L is a field, and $F \subset L \subset K$. Furthermore, L is the smallest subfield of K containing both F and a.*

Proof. Define the map $\phi : F[x]/(f(x)) \to K$ by setting $\phi(g(x)) = g(a)$. Here we view $F[x]/(f(x))$ as the set of polynomials of degree less than that of $f(x)$. It is immediate that $\phi(0) = 0$, $\phi(1) = 1$, and $\phi(g(x) + h(x)) = \phi(g(x)) + \phi(h(x))$. To consider $\phi(g(x)h(x))$, write $g(x)h(x) = q(x)f(x) + r(x)$, so $r(x)$ is the product of $g(x), h(x)$ in $F[x]/(f(x))$. Then $g(a)h(a) = q(a)f(a) + r(a) = r(a)$, implying $\phi(g(x)h(x)) = \phi(g(x))\phi(h(x))$. Finally, 5.10 implies that ϕ is one to one.

Since $F[x]/(f(x))$ is a field, it is immediate that the image, L, is a field. Since $F[x]/(f(x))$ contains F as the polynomials of degree 0, it is immediate that $F \subset L$. If $L' \subset K$ is any other field containing F and a, then L' must (by closure) contain all expressions of the form $g_0 + g_1 a + \ldots + g_r a^r$ where all $g_i \in F$. In other words, L' must contain all $g(a)$ for $g(x) \in F[x]$ and, in particular, must contain L. ∎

We left the issue of constructing irreducible polynomials hanging while we dealt with 5.9 and 5.10, but now we are ready to return to it.

Proposition 5.11. *Let F be a finite field and $n \geq 1$ an integer. Then there is an irreducible polynomial in $F[x]$ of degree n.*

Proof. Let F have order q. Use 5.8 to construct $K \supset F$ of order q^n. Let $a \in K^*$ be an element of order $q^n - 1$. Let $f(x)$ be the minimum polynomial of a and $\phi : F[x]/(f(x)) \cong L \subset K$ the isomorphism from 5.10. Since all the nonzero elements of K are powers of a, it is clear that $K = L$. If m is the degree of $f(x)$, then $F[x]/(f(x))$ has order q^m. Since ϕ is one to one and onto, $q^m = q^n$ implying $f(x)$ has degree n. Finally, by 5.9, $f(x)$ is irreducible. ∎

Note that the $f(x)$ constructed above was "too" special. Namely, the method we used did not need that a generated K but only that $L = K$, that is, that a was not in any subfield of K containing a. But we know all the subfields, so we can say what this means:

Corollary 5.12. *Suppose $F \subset K$ are finite fields, F has order q and K has order q^n. Suppose $a \in K^*$ has order n and n does NOT divide $q^s - 1$ where $s < n$ divides n. Then the minimum polynomial of a over $F[x]$ is irreducible of degree n.*

Proof. By the proof of 5.11, it suffices to show that a is not a member of any proper subfield of K containing F. Suppose L is such a subfield. Then L has order q^s where $s < n$ divides n. Any element of L^* has order dividing $q^s - 1$, so $a \notin L^*$. ∎

It is a remarkable fact that 5.12 is the ONLY way to produce irreducible polynomials of given degree. This is closely related to the fact that there is only one choice of the field K, at least up to isomorphism. In fact, we will prove both facts in the same theorem.

Theorem 5.13. *Let F be a finite field of order q. Let $f(x)$ be an irreducible monic polynomial of degree n. Then $f(x)$ is a divisor of $x^{q^n} - x$. Let $K \supset F$ be a field of order q^n. Then $f(x)$ is a product of the form $\prod(x - a_i)$ in $K[x]$. Finally, if $K' \supset F$ is another field of order q^n, then there is an isomorphism $\phi : K \cong K'$ which is the identity on F.*

Proof. Let $f(x)$ be as given. Set $K' = F[x]/f(x))$. Then K' is a field of order q^n. Let $\mu \in K'$ correspond to x. Then $f(\mu) = 0$, so $f(x)$ (being irreducible) must be the minimum polynomial of μ. Since μ is a root of $x^{q^n} - x$, we have that $f(x)$ is a divisor of $x^{q^n} - x$.

Let K be as given. Then $x^{q^n} - x = \prod_{a \in K}(x - a)$. Since $f(x)$ is a divisor of $x^{q^n} - x$ in $F[x]$, this is also true in $K[x]$. But unique factorization shows that $f(x)$ is a product $\prod_i (x - a_i)$ where the a_i form a subset of K.

In particular, $f(x)$ has a root in K. By 5.10, there is an isomorphism $\phi : F[x]/(f(x)) \cong L \subset K$. Since both L and K have order q^n, $L = K$. Note that as defined ϕ is the identity on F. If K_1 is as given, $K_1 \cong F[x]/(f(x)) \cong K$. ∎

There is one more remark we need to make now about finite fields. It concerns another way of writing the irreducible polynomials, and computing their degrees. Alternatively, it is a way of finding the "conjugates" of an element, that is to say, a way of, given one root, to find all the other roots of an irreducible polynomial. We begin with a comment which a repeat of what we have already said. Namely, suppose $K \supset F$ are finite fields, and F has order q. Then F is precisely the set of elements of K which satisfy $\alpha^q = \alpha$.

Lemma 5.14. *Suppose K/F are as above and $f(x) \in F[x]$ is a polynomial. Let $\beta \in K$ be a root of $f(x)$. Then β^q is a root of $F[x]$.*

Proof. Write $f(x) = f_0 x^n + f_1 x^{n-1} + \ldots + f_n$, so $f_0\beta^n + \ldots + f_n = 0$. Take this last equation and raise both sides to the q power. Using the characteristic p binomial theorem repeatedly, we have $f_0^q \beta^{qn} + f_1^q \beta q(n-1) + \ldots + f_n^q = 0$. But being elements of F, we have $f_i^q = f_i$ for all i. Thus the last equation can be read as saying $f(\beta^q) = 0$, as needed. ∎

Of course the process discussed in 5.14 can be repeated. That is, if β is a root of $f(x)$, then so is β^q, and hence $(\beta^q)^q = \beta^{q^2}$, β^{q^3} etc. Of course the above is not a infinite list, since the field K is finite. Thus we are interested when we get the first repeat. Note that the map $x \to x^q$ is injective on K. If $x^q = y^q$ then $(x - y)^q = 0$ and so $x - y = 0$. Thus if $\beta^{q^r} = \beta^{q^s}$, and $r < s$, then $\beta = \beta^{q^{s-r}}$. That is, the first

repeat appears when β reappears. Note that, if n is the order of β, then $\beta = \beta^{q^r}$ if and only if n divides $q^r - 1$. Thus the first repeat appears at the first r such that n divides $q^r - 1$. We call this r the order of q modulo n.

We now claim:

Theorem 5.15. *Suppose $K \supset F$ are as above, F has order q, and $\beta \in K^*$ has order n. Let $f(x) \in F[x]$ be the minimum polynomial of β over F, and r the order of q modulo n. Then $f(x) = \prod_{i=0}^{r-1}(x - \beta^{q^i})$. In particular, r is the degree of $f(x)$.*

Proof. The second statement follows from the first. As for the first, set $g(x) = \prod_{i=0}^{r-1}(x - \beta^{q^i})$. and write this polynomial out as

$$\prod_{i=0}^{r-1}(x - \beta^{q^i}) = x^r + g_1 x^{r-1} + \ldots g_r.$$

Take this equation and raise both sides to the q power. One gets

$$\prod_{i=0}^{r-1}(x^q - \beta^{q^{i+1}} = x^{qr} + g_1^q x^{q(r-1)} + \ldots g_r^q$$

and the left side can be written $\prod_{i=1}^{r}(x^q - \beta^{q^i})$. Now $\beta^{q^r} = \beta = \beta^{q^0}$ so the left side is really $\prod_{i=0}^{r-1}(x^q - \beta^{q^i})$ which is really $g(x^q) = x^{qr} + g_1 x^{q(r-1)} + \ldots + g_r$. Equating this with the right side above we have $g_i^q = g_i$ for all i, or that $g(x) \in F[x]$.

Since $f(x)$ is the minimum polynomial, and β is clearly a root of $g(x)$, we have $f(x)$ divides $g(x)$. On the other hand, by 5.14, $x - \beta^{q^i}$ divides $f(x)$ for all i. Since $(x - \beta^{q^i})$ are all distinct for $i = 0, \ldots, r-1$, we have $g(x)$ divides $f(x)$. Since both are monic polynomials, $g(x) = f(x)$. ∎