A connection between number theory and linear algebra

Mark Steinberger

Contents

1. Some basics 1
2. Rational canonical form 2
3. Prime factorization in $\mathbb{F}[x]$ 4
4. Units and order 5
5. Finite fields 8
6. Matrices whose order is a power of the characteristic 10
7. Cyclotomic polynomials 11
8. Cyclotomic polynomials over $\mathbb{Z}_p$ 13

This is meant to be used in a variety of courses with different backgrounds so we will summarize some background material that can be learned elsewhere.

1. Some basics

Throughout this document $\mathbb{F}$ will denote a field. We write $\mathbb{Q}$ for the rational numbers, $\mathbb{R}$ for the real numbers, $\mathbb{C}$ for the complex numbers, $\mathbb{Z}$ for the integers, and $\mathbb{Z}_n$ for the integers modulo $n$. In particular, $\mathbb{Z}_p$ is a field important in number theory.

Write $\mathbb{F}[x]$ for the polynomials with coefficients in $\mathbb{F}$ and write $M_n(\mathbb{F})$ for the $n \times n$ matrices with coefficients in $\mathbb{F}$. Let

$$f = a_k x^k + \cdots + a_0 \in \mathbb{F}[x],$$

i.e., $a_0, \ldots, a_k \in \mathbb{F}$. If $a_k \neq 0$, we say $f$ has degree $k$ and (written $\deg f = k$) that $a_k$ is its leading coefficient. (If all the coefficients of $f$ are 0, we say $f = 0$ and its degree is $-\infty$.) We say $f$ is monic if its leading coefficient is 1.

We say that $A \in M_n(\mathbb{F})$ is a matrix root of the polynomial $f \in \mathbb{F}[x]$ if $f(A) = 0$. Here, $f(A) = a_k A^k + \cdots + a_0 I_n$, where $A^i$ is the $i$th power of $A$ under matrix multiplication and $I_n$ is the identity matrix in $M_n(\mathbb{F})$. The minimal polynomial, $\min_A(x)$ (or simply $\min_A$) is the monic polynomial of
lowest degree having $A$ as a matrix root. By the Cayley–Hamilton theorem below, $A$ is a matrix root of some monic polynomial. Therefore, our definition of the minimal polynomial is well-defined by the following basic result from linear algebra.

**Lemma 1.1.** Let $f \in \mathbb{F}[x]$. Then $f(A) = 0$ if and only if $\text{min}_A(x)$ divides $f$.

Here, $g$ divides $f$ if $f = gh$ for $h \in \mathbb{F}[x]$. We write $g | f$ for “$g$ divides $f$”.

The characteristic polynomial $\text{ch}_A(x)$ (or simply $\text{ch}_A$) is defined by

$$\text{ch}_A(x) = \det(xI_n - A),$$

where $xI_n$ is the matrix with coefficients in $\mathbb{F}[x]$ whose diagonal entries are all $x$ and whose off-diagonal entries are all 0. $\text{ch}_A(x)$ is a monic polynomial in $\mathbb{F}[x]$ of degree $n$.

**Theorem 1.2** (Cayley–Hamilton theorem). $A$ is a matrix root of $\text{ch}(x)$, and hence $\text{min}_A(x)$ divides $\text{ch}_A(x)$.

The matrices $A, B \in M_n(\mathbb{F})$ are said to be similar (written $A \sim B$) if $B = PAP^{-1}$ for some invertible matrix $P \in M_n(\mathbb{F})$. This means they differ by a linear change of variables, and hence have the same basic properties. It is easy to see that similar matrices have the same minimal polynomial. It is also true that similar matrices have the same characteristic polynomial.

If $A$ and $B$ are invertible, then the notion of similarity coincides with the notion of conjugacy in group theory. More on this below.

An important concept in linear algebra is the block sum of matrices. If $A \in M_n(\mathbb{F})$ and $B \in M_m(\mathbb{F})$, then the block sum $A \oplus B$ is the $(n+m) \times (n+m)$ matrix

$$A \oplus B = \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix}.$$  

Here, the 0’s are the appropriate sized 0-matrices.

### 2. Rational canonical form

A connection between polynomials and matrices is given by the following construction.

**Definition 2.1.** Let $f = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be a monic polynomial of degree $n$ in $\mathbb{F}[x]$. Then the companion matrix $C(f) \in M_n(\mathbb{F})$ is given by

$$C(f) = \begin{bmatrix}
0 & \cdots & 0 & -a_0 \\
1 & \cdots & 0 & -a_1 \\
\vdots & & \ddots & \vdots \\
0 & \cdots & 1 & -a_{n-1}
\end{bmatrix} = \begin{bmatrix}
0 & -a_0 \\
1 & -a_1 \\
\vdots & \ddots & \ddots \\
0 & \cdots & 1 & -a_{n-1}
\end{bmatrix}, \quad \text{with} \quad v = \begin{bmatrix}
-a_1 \\
\vdots \\
-a_{n-1}
\end{bmatrix}.$$  

An important property of companion matrices is the following.
Proposition 2.2. Let \( f \) be a monic polynomial of degree \( n \) in \( \mathbb{F}[x] \). Then the minimal and characteristic polynomials of its companion matrices are both equal to \( f \):
\[
\min_{C(f)}(x) = \text{ch}_{C(f)}(x) = f.
\]

There and back again, as it were. One of the main theorems in linear algebra is the following.

Theorem 2.3 (Rational canonical form). Let \( \mathbb{F} \) be a field and let \( A \in M_n(\mathbb{F}) \). Then there is a unique \( k \geq 1 \) and a unique sequence \( f_1, \ldots, f_k \in \mathbb{F}[x] \) of monic polynomials such that
\[
(2.1) \quad f_i \text{ divides } f_{i-1} \quad \text{for } i = 2, \ldots, k
\]
and
\[
(2.2) \quad A \sim C(f_1) \oplus \cdots \oplus C(f_k).
\]
Moreover, \( \min_A(x) = f_1 \) and \( \text{ch}_A(x) = f_1 \cdots f_k \).

If \( f_1, \ldots, f_k \) satisfy (2.1) and \( A \) satisfies (2.2) we say \( f_1, \ldots, f_k \) are the invariant factors of \( A \) and that \( C(f_1) \oplus \cdots \oplus C(f_k) \) is the rational canonical form of \( A \). The theorem then implies that two matrices are similar if and only if they have the same rational canonical form.

Note that rational canonical form requires (2.1). Thus, since neither of \( x^2 \) and \( x^2 + 1 \) divides the other, \( C(x^2) \oplus C(x^2 + 1) \) is not in rational canonical form. In fact, the rational canonical form of \( C(x^2) \oplus C(x^2 + 1) \) is
\[
C(x^4 + x^2) = \begin{bmatrix}
0 & 0 & 0 & 0 \\
1 & 0 & 0 & -1 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}.
\]
This may be shown by a Chinese remainder theorem for companion matrices. See Theorem 3.3, below.

Regardless, if there is only one polynomial \( f \) in our list, then (2.1) is satisfied automatically, and hence \( C(f) \) is in rational canonical form for every monic polynomial \( f \).

Remark 2.4. One important application of rational canonical form is to the understanding of the natural inclusion \( M_n(\mathbb{F}) \subset M_n(\mathbb{E}) \) induced by an inclusion \( \mathbb{F} \subset \mathbb{E} \) of fields. We also have an induced inclusion \( \mathbb{F}[x] \subset \mathbb{E}[x] \), which preserves division of polynomials and preserves the monic property. So a rational canonical form \( C(f_1) \oplus \cdots \oplus C(f_k) \) in \( M_n(\mathbb{F}) \) is still in rational canonical form in \( M_n(\mathbb{E}) \). So the rational canonical form of \( A \in M_n(\mathbb{F}) \) remains unchanged when we regard \( A \) as an element of \( M_n(\mathbb{E}) \). In particular, the rational canonical form of \( A \) actually has coefficients in the smallest subfield of \( \mathbb{F} \) containing all the coefficients of \( A \) (or sometimes even a smaller field), and remains unchanged over any larger field containing \( \mathbb{F} \). The theorem then shows that if \( \mathbb{F} \subset \mathbb{E} \) and if \( A, B \in M_n(\mathbb{F}) \), then \( A \) is similar to \( B \) in \( M_n(\mathbb{F}) \) if and only if \( A \) is similar to \( B \) in \( M_n(\mathbb{E}) \).
3. Prime factorization in $\mathbb{F}[x]$

Finding the sequences satisfying (2.1) requires understanding the prime factorization of polynomials. We review that here.

**Definition 3.1.** A polynomial $f \in \mathbb{F}[x]$ is irreducible (in $\mathbb{F}[x]$) if:

1. $\deg f > 0$.
2. If $f = gh$ with $g, h \in \mathbb{F}[x]$, then either $\deg g = 0$ or $\deg h = 0$.

This is a nearly precise analogue of prime numbers in $\mathbb{Z}$: a polynomial $f$ has a multiplicative inverse in $\mathbb{F}[x]$ if and only if $\deg f = 0$. So $\deg f > 0$ means that $f$ is nonzero and not invertible. In the usual definition of prime number, (1) is replaced by saying $p > 1$, whereas the strict analogue of (1) would be that the absolute value of $p$ be greater than 1, which would allow $-p$ to also be prime. In particular, if $0 \neq a \in \mathbb{F}$, then $f$ is irreducible if and only if $af$ is irreducible, but $f$ divides all the same polynomials as $af$, so that “prime decompositions” by irreducibles are not unique. The easiest way to fix this is by using monic irreducibles. So the following gives a clean statement.

**Theorem 3.2** (Prime decomposition for polynomials). Let $f \in \mathbb{F}[x]$ with $\deg f > 0$. Then $f$ may be written uniquely in the form

$$f = ap_1^{r_1} \cdots p_k^{r_k},$$

where $0 \neq a \in \mathbb{F}$, $k \geq 1$, $p_1, \ldots, p_k$ are distinct monic irreducibles in $\mathbb{F}[x]$, and $r_i > 0$ for $i = 1, \ldots, k$. Moreover, for $f$ as in (3.1), $g$ divides $f$ in $\mathbb{F}[x]$ if and only if

$$g = bp_1^{s_1} \cdots p_k^{s_k}$$

with $0 \neq b \in \mathbb{F}$ and $0 \leq s_i \leq r_i$ for all $i$.

Note that polynomials of degree 1 are always irreducible, because they cannot be factored as products of polynomials of positive degree, as

$$\deg(gh) = \deg g + \deg h$$

for all $g, h \in \mathbb{F}[x]$. (That is why we defined the degree of the 0-polynomial to be $-\infty$.)

Unlike the rational canonical form, prime decompositions can change as you go from $\mathbb{F}[x]$ to $\mathbb{E}[x]$ when $\mathbb{F} \subset \mathbb{E}$. For instance, if $f = x^2 - 2$, then the prime decomposition of $f$ in $\mathbb{R}[x]$ is

$$f = (x - \sqrt{2})(x + \sqrt{2}).$$

If $f$ had any degree 1 monic factors in $\mathbb{Q}[x]$ they would have to persist as factors in $\mathbb{R}[x]$. But neither $(x - \sqrt{2})$ nor $(x + \sqrt{2})$ lies in $\mathbb{Q}[x]$, so by uniqueness of prime decomposition $f$ must be irreducible in $\mathbb{Q}[x]$.

Similarly, $f = x^2 + 1$ is irreducible in $\mathbb{R}[x]$ (and hence also $\mathbb{Q}[x]$) as it factors as $(x - i)(x + i)$ in $\mathbb{C}[x]$.

Now that we have prime decomposition for polynomials we can state the following.
Theorem 3.3 (Chinese remainder theorem for companion matrices). Let \( f \in \mathbb{F}[x] \) be monic, with prime decomposition

\[
f = p_1^{r_1} \cdots p_k^{r_k}
\]

in \( \mathbb{F}[x] \), i.e., \( p_1, \ldots, p_k \) are distinct monic irreducibles in \( \mathbb{F}[x] \) and \( r_i > 0 \) for \( i = 1, \ldots, k \). Then

\[
C(f) \sim C(p_1^{r_1}) \oplus \cdots \oplus C(p_k^{r_k})
\]

in \( \mathbb{F}[x] \).

Indeed, this may be proven using a Chinese remainder theorem for quotients of the polynomial ring \( \mathbb{F}[x] \), or by studying the theory of minimal polynomials in greater depth.

Applying (3.2) to each summand in (2.2), we see that every matrix in \( M_n(\mathbb{F}) \) is similar to a block sum of matrices of the form \( C(p^r) \) with \( p \) monic and irreducible in \( \mathbb{F}[x] \). It is then not hard to show that such a decomposition is unique in the sense that for each \( p \) and \( r \), the number of summands of the form \( C(p^r) \) is unique, and is determined by the dimensions of certain nullspaces. Here \( p \) has to be an irreducible factor of the characteristic polynomial \( \chi_A \). One may then calculate these nullspaces and then apply (3.2) in reverse to put the matrix in rational canonical form. Details are given in a standard course in advanced linear algebra. The key is obtaining the prime decomposition of the characteristic polynomial, a highly nontrivial task. The rest is easy. Note that unlike the rational canonical form, a decomposition in terms of block sums of matrices \( C(p^r) \) with \( p \) irreducible is not invariant under enlarging the field \( \mathbb{F} \), as the irreducibles may then factor.

Exercises.

1. Investigate prime decompositions and primality testing in \( \mathbb{Z}_p[x] \) for a particular choice of \( p \).
2. Show that \( x^p - x + a \) is irreducible in \( \mathbb{Z}_p[x] \) for all \( 0 \neq a \in \mathbb{Z}_p \).
3. Deduce from the discussion of finite fields below that these are the only monic irreducibles of degree \( p \) in \( \mathbb{Z}_p[x] \).
4. How do you find irreducibles of degree \( p^2 \) in \( \mathbb{Z}_p[x] \)?

4. Units and order

A unit in a ring \( R \) is an element of \( R \) with a multiplicative inverse, i.e., there exists \( b \in R \) with \( ab = ba = 1 \). The collection of all units in \( R \) forms a group under multiplication. We write \( R^\times \) for the group of units of \( R \).

As an example, consider \( M_n(\mathbb{F}) \), a ring under standard addition and multiplication of matrices. The usual definition for a matrix \( A \) to be invertible is precisely that there is a matrix \( B \) with \( AB = BA = I_n \), where the identity matrix \( I_n \) is the multiplicative identity of \( M_n(\mathbb{F}) \). Thus, the group of units of \( M_n(\mathbb{F}) \) is precisely the group of invertible \( n \times n \) matrices over \( \mathbb{F} \). We call it the \( n \)th general linear group, \( \text{GL}_n(\mathbb{F}) \) of \( \mathbb{F} \):

\[
M_n(\mathbb{F})^\times = \text{GL}_n(\mathbb{F}).
\]
Another important example is the unit group of \(\mathbb{Z}_n\). We write \(\overline{a} \in \mathbb{Z}_n\) for the congruence class of the integer \(a\). The following is standard.

**Proposition 4.1.** An element \(\overline{a} \in \mathbb{Z}_n\) is a unit if and only if \((a,n) = 1\). The number of units in \(\mathbb{Z}_n\) is given by the Euler \(\phi\)-function \(\phi(n)\), which may be computed as follows: if \(n = p_1^{r_1} \cdots p_k^{r_k}\) where \(p_1 < \cdots < p_k\) are prime and \(r_i > 0\) for \(i = 1, \ldots, k\), then

\[
\phi(n) = \prod_{i=1}^{k} p_i^{r_i-1}(p_i - 1).
\]

Of course, a field \(\mathbb{F}\) is precisely a commutative ring in which every nonzero element is a unit. So \(\mathbb{F}^* = \mathbb{F} - \{0\}\), where the minus sign means set theoretic difference. Since the prime numbers are precisely those numbers \(n\) for which \(a \not\equiv 0 \text{ mod } n\) if and only if \((a,n) = 1\), \(\mathbb{Z}_n\) is a field if and only if \(n\) is prime.

An important issue in group theory is the concept of order. It has important applications to unit groups.

**Definition 4.2.** Let \(G\) be a group and let \(g \in G\). We say \(g\) has finite order if \(g^k = 1\) for some \(k > 0\). An integer \(k\) with \(g^k = 1\) is called an exponent of \(g\). If \(g\) has finite order, then the order of \(g\), written \(|g|\), is the smallest positive exponent of \(g\). If \(g\) does not have finite order, we write \(|g| = \infty\).

We write \(\langle g \rangle\) for the set of all integer powers of \(g\):

\[
\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}.
\]

It is a subgroup of \(G\), and hence is the smallest subgroup of \(G\) containing \(g\). We call it the cyclic subgroup of \(G\) containing \(g\). We say \(G\) itself is cyclic if \(G = \langle g \rangle\) for some \(g \in G\). In this case we say \(g\) generates \(G\).

If \(G\) is finite, we write \(|G|\) for the number of elements of \(G\), and call it the order of \(G\). More generally, if \(X\) is a finite set, we write \(|X|\) for the number of elements of \(X\).

The notions of order and of cyclic subgroups tie together nicely.

**Lemma 4.3.** Let \(g\) be an element of order \(n < \infty\) in the group \(G\). Then:

1. \(g^k = 1\) if and only if \(n \mid k\).
2. \(g^k = g^\ell\) if and only if \(k \equiv \ell \text{ mod } n\).
3. \(\langle g \rangle\) has exactly \(n\) distinct elements: \(\{g^k \mid 0 \leq k < n\}\). Thus, the order of \(\langle g \rangle\) is equal to the order of \(g\).

Another important relationship between group order and the orders of elements is the following.

**Proposition 4.4** (Lagrange’s theorem). Let \(G\) be a finite group and let \(g \in G\). Then the order of \(g\) divides the order of \(G\).

The following allows us to determine the orders of the elements in \(\langle g \rangle\) via number theory.
**Proposition 4.5.** Let \( g \) be an element of finite order in the group \( G \) and let \( k \in \mathbb{Z} \). Then

\[
|g^k| = \frac{|g|}{(|g|, k)}.
\]

**Corollary 4.6.** Let \( G \) be a cyclic group of order \( n \). Then there are exactly \( \phi(n) \) generators for \( G \). Specifically, if \( g \) is one generator for \( G \), then the complete set of generators is \( \{g^k \mid 0 \leq k < n, (k, n) = 1 \} \). These exponents \( k \) are precisely integers representing the distinct units in \( \mathbb{Z}_n \), so there are \( \phi(n) \) of them. More generally, if \( d \) divides \( n \) there are exactly \( \phi(d) \) elements of degree \( d \) in \( G \). Thus,

\[
n = \sum_{d|n} \phi(d).
\]

**Proof.** \( g^k \) generates \( G \) if and only if \( |g^k| = |G| = n \). In particular, \( |g| = n \). Thus, by Proposition 4.5 \( |g^k| = n \) if and only if \( (n, k) = 1 \).

More generally, if \( d|n \), let \( n = dm \). Then \( |g^k| = d \) if and only if \( (n, k) = m \). If \( (n, k) = m \), then \( k = mr \), and

\[
(d, r) = \left( \frac{n}{m}, \frac{k}{m} \right) = 1,
\]

as \( m = (n, k) \). Note that \( g^m \) has order \( d \) and that \( g^k = (g^m)^r \) is a generator of \( \langle g^m \rangle \), as \( (r, d) = 1 \). Since \( |\langle g^m \rangle| = d \), there are \( \phi(d) \) such generators, and the result follows. \( \square \)

Determining the orders of elements in the groups \( \text{GL}_n(F) \) is important in the theory of \( F \)-representations of a finite group. Also, if \( F \) is a finite field, \( \text{GL}_n(F) \) is a finite group:

**Proposition 4.7.** If \( F \) is a finite field with \( q \) elements, then \( \text{GL}_n(F) \) is finite, with

\[
|\text{GL}_n(F)| = (q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})
= q^{n(n-1)/2} (q^n - 1)(q^{n-1} - 1) \cdots (q - 1).
\]

In fact, certain quotients of subgroups of general linear groups of finite fields form an important family of finite simple groups. And understanding the finite simple groups is an important component in understanding finite groups.

The following is immediate from the results in Section 1. Once again, the field \( F \) is arbitrary.

**Proposition 4.8.** Let \( A \in \text{GL}_n(F) \). Then \( k \) is an exponent for \( A \) if and only if \( A \) is a matrix root of \( x^k - 1 \). Thus, \( k \) is an exponent for \( A \) if and only if \( \min_A(x) \) divides \( x^k - 1 \).
Thus, if we know the prime decomposition of $x^k - 1$ in $\mathbb{F}[x]$, rational canonical form will allow us to calculate all similarity classes of matrices of exponent $k$.

In fact, every matrix root, $A$, of $x^k - 1$ is invertible, as $A^{k-1}A = AA^{k-1} = I_n$. In $\text{GL}_n(\mathbb{F})$, the notion of similarity coincides with the group theoretic notion of conjugacy: $g, h \in G$ are conjugate if $h = xgx^{-1}$ for some $x \in G$. Conjugacy preserves order and preserves many other important group theoretic properties.

Determining the conjugacy classes of the elements of different orders in a finite group is valuable in understanding that group. Thus, understanding the prime factorization of $x^k - 1$ in $\mathbb{F}[x]$ is important in understanding the group structure of $\text{GL}_n(\mathbb{F})$ when $\mathbb{F}$ is finite. Thus, a discussion of finite fields will be valuable.

5. Finite fields

Let $\mathbb{F}$ be a finite field. Then the additive subgroup generated by 1 is a subring of $\mathbb{F}$, and hence is an integral domain, and may be identified in a unique way with the ring $\mathbb{Z}_n$. But the only values of $n$ for which $\mathbb{Z}_n$ is a domain are the prime numbers. So there is a uniquely identified copy of $\mathbb{Z}_p$ contained in $\mathbb{F}$, additively generated by 1. We say $\mathbb{F}$ has characteristic $p$ and that $\mathbb{Z}_p$ is its prime subfield.

Whenever one field is a subfield of another, the larger field is a vector space over the smaller one in the obvious way. Since $\mathbb{F}$ is finite, it is a finite-dimensional vector space over $\mathbb{Z}_p$, and hence has $p^r$ elements for some $r \geq 0$. We then say $\mathbb{F}$ has degree $r$ over $\mathbb{Z}_p$.

We shall give some results from Galois theory that characterize finite fields. The following depends on the theory of splitting fields.

**Proposition 5.1.** Let $p$ be a prime and $r \geq 1$. Then there exists a field $\mathbb{F}$ with $p^r$ elements. Moreover, it is unique in the sense that if $\mathbb{F}_1$ is another field with $p^r$ elements there is an isomorphism $\nu : \mathbb{F} \to \mathbb{F}_1$ that restricts to the identity on the prime subfield. It is customary to write $\mathbb{F}_{p^r}$ for this field.

Finally, if $s$ divides $r$, then $\mathbb{F}_{p^s}$ is the subfield of $\mathbb{F}_{p^r}$ consisting of all $\alpha \in \mathbb{F}_{p^r}$ such that $\alpha^{p^s} = \alpha$.

Here, an isomorphism of fields is a ring homomorphism (i.e., $\nu$ preserves addition and multiplication and takes 1 to 1) that is 1-1 and onto.

There is an important notion of minimal polynomials of elements of an extension of fields. Let $\mathbb{E}$ be a field containing $\mathbb{F}$ as a subfield such that $\mathbb{E}$ is finite-dimensional as a vector space over $\mathbb{F}$. We call such an $\mathbb{E}$ a finite extension of $\mathbb{F}$. The degree of $\mathbb{E}$ over $\mathbb{F}$, written $[\mathbb{E} : \mathbb{F}]$, is the dimension of $\mathbb{E}$ as a vector space over $\mathbb{F}$.

If $\alpha \in \mathbb{E}$ and $f \in \mathbb{F}[x]$, say, $f = a_nx^n + \cdots + a_0$, we define $f(\alpha)$ to be $a_n\alpha^n + \cdots + a_0$. We say $\alpha$ is a root of $f$ if $f(\alpha) = 0$. It is not hard to show that there exist monic polynomials $f$ having $\alpha$ as a root. We define $\text{min}_\alpha(x)$
to be the monic polynomial of lowest degree for which $\alpha$ is a root. As in
the case of matrices, we then have that $\alpha$ is a root of $f \in \mathbb{F}[x]$ if and only if
$\text{min}_{\alpha}(x)$ divides $f$. The following is basic.

**Proposition 5.2.** $\text{min}_{\alpha}(x)$ is irreducible, $\deg \text{min}_{\alpha}(x) \leq [E:F]$ with equal-
ity if and only if there is no proper subfield of $E$ containing both $F$ and $\alpha$. In fact, $\deg \text{min}_{\alpha}(x)$ is equal to the degree over $F$ of the smallest subfield of
$E$ containing $F$ and $\alpha$. This implies that $\deg \text{min}_{\alpha}(x)$ divides $[E:F]$.

**Examples 5.3.**

1. $\alpha$ is a root of the degree 1 monic $x - a \in \mathbb{F}[x]$ if and only if $\alpha = a \in \mathbb{F}$. Thus $\text{min}_{\alpha}(x)$ has degree 1 if and only if $\alpha \in \mathbb{F}$.
2. If $[E:F] = 2$ and if $\alpha \notin \mathbb{F}$, then $\deg \text{min}_{\alpha}(x) = 2$. Thus if we regard $\text{min}_{\alpha}(x)$ as an element of $E[x]$ via the natural inclusion, $x - \alpha$
divides it. Since the complementary factor is monic of degree 1, $\text{min}_{\alpha}(x)$ factors as

$$\text{min}_{\alpha}(x) = (x - \alpha)(x - \beta)$$

in $E[x]$ for some $\beta \in E$. It could happen that $\beta = \alpha$ but only in
what’s known as a nonseparable extension in characteristic 2. We
will avoid that case.

There is a standard construction, similar to the one used to construct
the rings $\mathbb{Z}_n$ from the integers, that, given a monic irreducible polynomial

$f \in \mathbb{F}[x]$ constructs an extension $E$ of $\mathbb{F}$, with $[E:F] = \deg f$, such that

$f$ has a root in $E$. Since the minimal polynomial of that root then divides $f$ and since $f$ is irreducible, this implies that $f$ is the minimal polynomial of
any of its roots in an extension field of $\mathbb{F}$. This, together with the fact
that $\mathbb{F}_{p^r}$ is what’s known as a Galois extension of $\mathbb{Z}_p$ allows us to deduce the
following.

**Theorem 5.4.** Let $p$ be prime, $r > 0$, and let $d$ be a divisor of $r$. Let $f$
be a monic irreducible polynomial of degree $d$ in $\mathbb{Z}_p[x]$. Then regarding $f$
as a polynomial in $\mathbb{F}_{p^r}[x]$ via the natural inclusion of $\mathbb{Z}_p$ in $\mathbb{F}_{p^r}$, we obtain a
factorization

$$f = (x - \alpha_1) \ldots (x - \alpha_d),$$

where $\alpha_1, \ldots, \alpha_d$ are distinct elements of $\mathbb{F}_{p^r}$. In particular, $f$ has the $d$
distinct roots $\alpha_1, \ldots, \alpha_d$ in $\mathbb{F}_{p^r}$, and hence $f$ is the minimal polynomial of
each of these roots over $\mathbb{Z}_p$.

Finally, the minimal polynomial over $\mathbb{Z}_p$ of any $\alpha \in \mathbb{F}_{p^r}$ has degree divid-
ing $r$.

\footnote{Note this important difference between the minimal polynomial of a matrix over $\mathbb{F}$
and the minimal polynomial of an element in a field containing $\mathbb{F}$. In the former case, the
minimal polynomial need not be irreducible. In the latter, it must be.}
Since the monic irreducible \( f \in \mathbb{Z}_p[x] \) is the minimal polynomial of any of its roots in an extension of \( \mathbb{Z}_p \), distinct monic irreducibles cannot share a root in any extension. So the elements of \( \mathbb{F}_{p^r} \) are partitioned by the sets of roots of their minimal polynomials, and those minimal polynomials range over all monic irreducibles in \( \mathbb{Z}_p[x] \) of degree dividing \( r \). If the degree, \( d \), is less than \( r \), then its roots lie in a proper subfield of \( \mathbb{F}_{p^r} \).

**Example 5.5.** In \( \mathbb{F}_{3^2} \), only 0 and 1 lie in the prime field, and the minimal polynomial of every other element has degree 5. Each irreducible of degree 5 in \( \mathbb{Z}_2[x] \) has 5 roots in \( \mathbb{F}_{3^2} \). Since there are 30 elements of \( \mathbb{F}_{3^2} \) not in the ground field, that gives six irreducibles of degree 5 in \( \mathbb{Z}_2[x] \) (all nonzero polynomials over \( \mathbb{Z}_2 \) are monic). Exercise: find them.

Finally, we'll make use of the following.

**Theorem 5.6.** Let \( \mathbb{F} \) be a field and let \( H \) be a finite subgroup of \( \mathbb{F}^\times \). Then \( H \) is cyclic. In particular, if \( \mathbb{F} \) itself is finite, then \( \mathbb{F}^\times \) is cyclic.

### 6. Matrices whose order is a power of the characteristic

In this section, \( \mathbb{F} \) is a field of characteristic \( p > 0 \), e.g., \( \mathbb{F} = \mathbb{Z}_p \). In particular, the additive subgroup generated by 1 is a subfield we may canonically identify with \( \mathbb{Z}_p \).

In particular the sum of \( p \) copies of 1 gives 0. By the distributive law, the sum of \( p \) copies of anything gives 0. Since the binomial coefficient \( \binom{p}{i} \) is divisible by \( p \) for \( 0 < i < p \), the binomial theorem gives the following.

**Lemma 6.1.** Let \( \mathbb{F} \) be a field of characteristic \( p > 0 \) and let \( f, g \in \mathbb{F}[x] \). Then \( (f + g)^p = f^p + g^p \). Inductively, \( (f + g)^{ps} = f^{ps} + g^{ps} \) for all \( s \geq 1 \). Note this includes the case \( f \) and \( g \) have degree \( \leq 0 \), i.e., are elements of \( \mathbb{F} \).

This makes rational canonical forms incredibly simple when the exponent is a power of the characteristic of the field.

**Corollary 6.2.** Let \( \mathbb{F} \) be a field of characteristic \( p \). Then the prime factorization of \( x^{ps} - 1 \) in \( \mathbb{F}[x] \) is

\[
x^{ps} - 1 = (x - 1)^{ps}.
\]

Thus, the rational canonical forms for matrices in \( \text{GL}_n(\mathbb{F}) \) whose order is a power of \( p \) are given by

\[
C((x - 1)^{r_1}) \oplus \cdots \oplus C((x - 1)^{r_k}),
\]

where \( r_1 \geq \cdots \geq r_k \) and \( r_1 + \cdots + r_k = n \). In particular, the number of conjugacy classes is equal to the number of partitions of \( n \).

In particular, for a matrix whose rational canonical form is \( (6.2) \), the minimal polynomial is \( (x - 1)^{r_1} \), which in turn implies that the order of the matrix in question is the smallest power of \( p \) greater than or equal to \( r_1 \).
Note this is true for an arbitrary field of characteristic $p$ and hence each conjugacy class in $\text{GL}_n(F)$ whose order is a power of $p$ has a representative coming from $\text{GL}_n(Z_p)$. Indeed, we have a precise formula coming from the companion matrices. But actually writing down the matrices depends on computing all the binomial coefficients $\binom{r}{j}$ mod $p$. Since the prime factors of the invariant factors all have degree 1, these matrices can actually be put in Jordan canonical form. That is much cleaner and easier to write down. In particular, for any field $F$, any $a \in F$ and any $m \geq 1$, $C((x - a)^m)$ is similar to the $m \times m$ Jordan block, $J(a, m)$, with eigenvalue $a$:

$$J(a, m) = \begin{bmatrix}
a & 0 & \cdots & 0 \\
1 & a & \cdots & 0 \\
\vdots & & \ddots & \vdots \\
0 & 0 & \cdots & 1\end{bmatrix} = aI_n + \begin{bmatrix}0 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 0 \end{bmatrix} = aI_n + C(x^n).$$

Proof of Corollary 6.2. If the order of $A \in \text{GL}_n(F)$ is a power of $p$, then the minimal polynomial of $A$ is a factor of $x^{ps} - 1 = (x - 1)^{ps}$ for some $s$, and hence must be equal to $(x - 1)^{r_1}$ for some $r_1 \leq s$ by the uniqueness of prime factorization in $F[x]$. Since each invariant factor divides the previous one, uniqueness of factorization shows the rational canonical form must have the form given in (6.2).

The exponents $r_i$ add up to $n$ because the matrix $C((x - 1)^{r_1})$ is $r_1 \times r_1$. A sequence $r_1, \ldots, r_k$ with $r_1 \geq \cdots \geq r_k$ and $r_1 + \cdots + r_k = n$ is precisely a partition of $n$. Since the minimal polynomial divides $x^{ps} - 1$ the order divides $p^s$, and hence will be the smallest power, $p^t$ of $p$ such that $x^{p^t} - 1$ is divisible by $(x - 1)^{r_1}$. Since $x^{p^t} - 1 = (x - 1)^{p^t}$, unique factorization says $p^t$ is the smallest power of $p$ greater than or equal to $r_1$. □

7. Cyclotomic polynomials

The main result here is the following:

**Theorem 7.1.** There are monic polynomials $\Phi_m(x) \in Z[x]$, $m \geq 1$ such $\Phi_m(x)$ is irreducible in $Q[x]$ of degree $\phi(m)$, the Euler $\phi$-function of $m$, and for $k \geq 1$,

$$(7.1) \quad x^k - 1 = \prod_{d|k} \Phi_d(x),$$

where the product is taken over all $d$ dividing $k$. The equality holds in $Z[x]$ and gives the prime decomposition of $x^k - 1$ in $Q[x]$.

$\Phi_m(x)$ is called the $m$th cyclotomic polynomial. Since (7.1) holds in $Z[x]$, it holds in $F[x]$ for any field $F$. Therefore, calculating the prime decomposition of $\Phi_d(x)$ in $F[x]$ for every $d$ dividing $k$ will give us the prime decomposition of $x^k - 1$ in $F[x]$, and hence will determine the conjugacy classes of matrices of exponent $k$ in $\text{GL}_n(F)$ via rational canonical form. In fact, knowing which of these prime factors divide which cyclotomic polynomials...
will help us determine not only the exponent but the order of a particular companion matrix (and hence of the full rational canonical form).

Thus, it would be helpful to be able to actually calculate the cyclotomic polynomials. This may be done inductively via (7.1). The starting point is noting that since 1 has no proper divisors, \( x - 1 = \Phi_1(x) \). This now gives and immediate calculation of \( \Phi_p(x) \) for \( p \) prime: since \( x^p - 1 = \Phi_1(x)\Phi_p(x) \), we obtain

\[
(7.2) \quad \Phi_p(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + 1.
\]

The latter expression may be obtained by polynomial division or by the straightforward multiplication that \( (x^{p-1} + x^{p-2} + \cdots + 1)(x-1) = x^p - 1 \) as in the derivation of the geometric series.

Continuing in this manner, we see that \( x^{pr} - 1 = \Phi_1(x)\cdots\Phi_{pr-1}(x) \cdot \Phi_{pr}(x) \)

\[
= (x^{pr-1} - 1) \cdot \Phi_{pr}(x)
\]

by (7.1), so

\[
(7.3) \quad \Phi_{pr}(x) = \frac{x^{pr} - 1}{x^{pr-1} - 1} = \frac{(x^{pr-1})^p - 1}{x^{pr-1} - 1}
\]

\[
= (x^{pr-1})^{p-1} + (x^{pr-1})^{p-2} + \cdots + 1,
\]

just substituting \( x^{pr-1} \) for \( x \) in (7.2). Note the final expression in (7.3) is just the result of substituting \( x^{pr-1} \) for \( x \) in \( \Phi_p(x) \), so we obtain

\[
(7.4) \quad \Phi_{pr}(x) = \Phi_p(x^{pr-1}).
\]

Thus, for example, \( \Phi_3(x) = x^2 + x + 1 \) and \( \Phi_9(x) = x^6 + x^3 + 1 \).

When \( m \) is divisible by more than one prime, the calculation of \( \Phi_m(x) \) is more complicated, but may be carried out inductively as above. For instance,

\[
x^6 - 1 = (\Phi_1(x)\Phi_3(x))\Phi_2(x)\Phi_6(x) = (x^3 - 1)\Phi_2(x)\Phi_6(x),
\]

and \( \Phi_2(x) = x + 1 \), so

\[
\Phi_6(x) = \frac{x^6 - 1}{x^3 - 1} \cdot \frac{1}{x + 1} = (x^3 + 1) \cdot \frac{1}{x + 1} = x^2 - x + 1.
\]

We may think of the above calculation as taking place in the field of rational functions (i.e., quotients of polynomials) over \( \mathbb{R} \) (or \( \mathbb{Q} \)), often studied in calculus or precalculus.

Using methods like this it is possible to grind out formulas for the \( \Phi_m(x) \) using polynomial division. But general formulas are going to be hard to obtain.

Theorem 7.1 tells us \( \Phi_m(x) \) is irreducible in \( \mathbb{Q}[x] \) for all \( m \). Here is an example of a factorization mod \( p \).
Example 7.2. By (7.3), $\Phi_8(x) = x^4 + 1$. Straightforward calculation shows that $\Phi_8$ factors in $\mathbb{Z}_7[x]$ as

$$x^4 + 1 = (x^2 + 3x + 1)(x^2 + 4x + 1) \in \mathbb{Z}_7[x],$$

and that each of the displayed factors is irreducible in $\mathbb{Z}_7[x]$.

The prime factorization of the cyclotomic polynomials is a very important issue over fields containing $\mathbb{Q}$. But we shall not address that here. We shall now restrict attention to $\mathbb{F} = \mathbb{Z}_p$.

8. Cyclotomic polynomials over $\mathbb{Z}_p$

Note we have seen in (6.1) that for $k = p^s$, the prime decomposition of $x^{p^s} - 1$ is $(x - 1)^{p^s}$. (7.1) together with uniqueness of prime decomposition in $\mathbb{F}[x]$ then implies that $\Phi_{p^s}(x) = (x - 1)^{\varphi(p^s)} = (x - 1)^{p^s-1(p-1)}$ for all $t$.

Things become more interesting when $k$ is relatively prime to $p$. We shall make use of two facts quoted above: that there exist fields of arbitrary degree over $\mathbb{Z}_p$ and that the unit group of a finite field is cyclic. The former relies on the theory of splitting fields, which I believe to be required for any treatment of this material. But when used in conjunction with the latter it obviates a lot of field theory one might use for this study.

Lemma 8.1. Let $(k, p) = 1$. Then $\mathbb{F}_p^\times$ contains an element of order $k$ if and only if $s$ is divisible by the order of $p$ in $\mathbb{Z}_k^\times$.

Proof. Since $\mathbb{F}_p^\times$ is cyclic of order $p^s - 1$ and since a cyclic group contains elements of every order dividing the order of the group, there is an element of order $k$ in $\mathbb{F}_p^\times$ if and only if $k$ divides $p^s - 1$. But that is equivalent to saying that $p^s \equiv 1 \mod k$, i.e., $s$ is an exponent for $p$ in $\mathbb{Z}_k^\times$. □

Theorem 8.2. Let $(p, k) = 1$ and $r$ be the order of $p$ in $\mathbb{Z}_k^\times$. Then $x^k - 1$ has $k$ distinct roots in $\mathbb{F}_{p^r}$. These roots are the elements of the unique cyclic subgroup of $\mathbb{F}_{p^r}^\times$ of order $k$. For $d$ dividing $k$, the roots of order $d$ are precisely the roots of $\Phi_d(x)$. The degree of the minimal polynomial over $\mathbb{Z}_p$ of a root of order $d$ is equal to the order, $\psi(p,d)$, of $p$ in $\mathbb{Z}_d^\times$. In particular, the irreducible factors of $\Phi_d(x)$ in $\mathbb{Z}_p[x]$ all have degree $\psi(p,d)$ when $(p,d) = 1$.

Proof. For each $d$ dividing $k$ there are exactly $\phi(d)$ elements of order $d$ in the cyclic group $\mathbb{F}_{p^r}^\times$. If $\alpha \in \mathbb{F}_{p^r}^\times$ has order $k$ then there are exactly $\phi(d)$ elements of order $d$ in the cyclic group $\langle \alpha \rangle$. Thus, there are exactly $k$ elements of exponent $k$ in $\mathbb{F}_{p^r}^\times$, and they are precisely the elements of $\langle \alpha \rangle$. In particular, the roots of $x^k - 1$ in $\mathbb{F}_{p^r}$ are precisely the powers of $\alpha$ and we obtain the prime decomposition

$$x^k - 1 = (x - \alpha)(x - \alpha^2) \ldots (x - \alpha^k) \in \mathbb{F}_{p^r}[x].$$

By (7.1) and uniqueness of prime decomposition, each of these powers of $\alpha$ is a root of $\Phi_d(x)$ for precisely one $d$ dividing $k$. Now, $\Phi_d(x)$ divides $x^d - 1$,
so any root of $\Phi_d(x)$ has order dividing $d$. Thus, the $\phi(k)$ elements of $\langle \alpha \rangle$ of order $k$ are precisely the $\phi(k)$ roots of $\Phi_k(x)$ in $\mathbb{F}_{p^r}$. The same reasoning shows that the $\phi(d)$ elements of order $d$ in $\langle \alpha \rangle$ are precisely the $\phi(d)$ roots of $\Phi_d(x)$ in $\mathbb{F}_{p^r}$.

If $\beta \in \langle \alpha \rangle$ has order $k$, then $\deg \min_\beta(x)$ must be equal to $r$ (where $\min_\beta(x)$ is the minimal polynomial of $\beta$ over $\mathbb{Z}_p$) by Lemma 8.1. Moreover, since $\beta$ is a root of $\Phi_k(x)$, $\min_\beta(x)$ divides $\Phi_k(x)$, and hence is an irreducible factor of $\Phi_k(x)$ in $\mathbb{Z}_p[x]$.

If $\beta \in \langle \alpha \rangle$ has order $d < k$, let $s = \psi(p, d)$. Then $s$ divides $r$ by Lemma 8.1, and $\mathbb{F}_{p^s}$ is a subfield of $\mathbb{F}_{p^r}$ by Proposition 5.1. Moreover, $\beta$ must lie in this subfield. We can now apply the same reasoning given above to the roots of $x^{d} - 1$ in $\mathbb{F}_{p^s}$. The result follows.

Example 8.3. $\mathbb{F}_6^6 = \mathbb{F}_{2^6}$ has order 63. Since 1, 2 and 3 are the proper divisors of 6, the subfields of $\mathbb{F}_6^6$ are $\mathbb{Z}_2$, $\mathbb{F}_4$ and $\mathbb{F}_8$. Checking orders in these subfields, we see that $\psi(2, 3) = 2$, $\psi(2, 7) = 3$, and $\psi(2, k) = 6$ for $k = 9, 21$ and 63. In other words, the irreducible factors of $\Phi_9$, $\Phi_{21}$ and $\Phi_{63}$ in $\mathbb{Z}_2[x]$ all have degree 6, while the irreducible factors of $\Phi_7$ over $\mathbb{Z}_2$ have degree 3. $\Phi_3$ itself has degree 2, so it is either irreducible over a given field $\mathbb{F}$ or its roots must lie in $\mathbb{F}$. In this case, it is irreducible.

In particular, note that since $\Phi_9$ has degree 6, it is irreducible in $\mathbb{Z}_2[x]$, and the elements of order 9 in $\mathbb{F}_6^6$ are precisely the roots of $\Phi_9$ there.

If $k$ is divisible by more than one prime, one can find matrices $A$ of order $k$ whose minimal polynomials are not divisible by any irreducible factor of $\Phi_k$. But this phenomenon cannot occur if $k$ is a prime power.

Corollary 8.4. Let $p$ and $q$ be distinct primes and let $k = q^r$. Then $A \in \text{GL}_n(\mathbb{Z}_p)$ has order $k$ if and only if $\min_A$ divides $x^k - 1$ and at least one of the irreducible factors of $\min_A$ is an irreducible factor of $\Phi_k$.

Proof. If $\min_A$ divides $x^{q^r} - 1$ but none of the irreducible factors of $\min_A$ divide $\Phi_{q^r}$, then $\min_A$ divides $x^{q^r-1} - 1$ and hence $A$ has order strictly less than $q^r$. □

Exercises.

1. Find representatives for the conjugacy classes in $\text{GL}_2(\mathbb{Z}_p)$ of elements whose order is a power of $p$, for various odd primes $p$.

2. In calculating conjugacy classes of elements of two-power order in $\text{GL}_2(\mathbb{Z}_p)$, what distinctions do you see depending on the congruence class of $p$ mod 4? What about the congruence class mod 8?

3. What happens if you replace 2 by 3 or 5 in the preceding problems?

4. What happens if you let the size of the matrices grow in the preceding questions?

5. What can you say about the conjugacy classes of elements of $p$-power order in $\text{GL}_n(\mathbb{Z}_2)$?

6. What is the order of $C(\Phi_9^k)$ in $\text{GL}_{2k}(\mathbb{Z}_2)$? Generalize this phenomenon as much as you can.
7. When does $\Phi_8$ factor in $\mathbb{Z}_p[x]$? For which primes do the factors have degree 2? And for which of those can you actually execute the factorization?

8. When does $\Phi_5$ factor in $\mathbb{Z}_p[x]$? For which primes do the factors have degree 2? And for which of those can you actually execute the factorization?

9. For which primes $p$ is $\Phi_9$ irreducible in $\mathbb{Z}_p[x]$? Find a prime $p$ for which $\Phi_3$ is irreducible but $\Phi_9$ reducible in $\mathbb{Z}_p[x]$.

10. Let $q$ be prime. For which primes $p$ is $\Phi_{q^2}$ irreducible. What then can you deduce about $\Phi_{q^r}$ in $\mathbb{Z}_p[x]$?