An Introduction to Abstract Algebra

Andrew Klapper[1] and Mark Goresky[2]

[1] Department of Computer Science, 779A Anderson Hall, University of Kentucky, Lexington, KY, 40506. www.cs.uky.edu/∼klapper

[2] School of Mathematics, Inst. for Advanced Study, Princeton, NJ,08540. www.math.ias.edu/∼goresky

This document is an introduction to a variety of topics in modern algebra. It is extracted from a book on algebraically defined pseudorandom sequences and the set of topics is geared to that purpose. There is an emphasis, for example, on finite fields and adic rings. The beginning sections, however, are quite general and can serve as an introduction to the algebra needed for such topics as coding theory and cryptography. There is a bibliography that contains many general books on algebra.

Contents

Chapter 1 Abstract Algebra

1.1 Group Theory

Groups are basic building blocks of modern algebra. They arise in a vast range of applications, including coding theory, cryptography, physics, chemistry, and biology. They are commonly used to model symmetry in structures or sets of transformations. They are also building blocks for more complex algebraic constructions such as rings, fields, vector spaces, and lattices.

1.1.a Basic properties

Definition 1.1.1 *A* group *is a set G with a distinguished element e (called the* identity*) and a binary operation $*$ satisfying the following axioms:*

1. *(Associative law) For all $a, b, c \in G$, $(a * b) * c = a * (b * c)$.*
2. *(Identity law) For all $a \in G$, $a * e = e * a = a$.*
3. *(Inverse law) For all $a \in G$, there exists $b \in G$ such that $a * b = e$. The element b is called an* inverse *of a.*

A group G is said to be commutative *or* Abelian *if it satisfies the following*

4. *(Commutative law) For all $a, b \in G$, $a * b = b * a$.*

Proposition 1.1.2 *Let G be a group. Then the following statements hold.*

1. *If $a, b \in G$ and $a * b = e$ then $b * a = e$.*
2. *Every $a \in G$ has a* unique *inverse.*
3. *The identity $e \in G$ is unique.*

Proof: To prove the first claim, suppose $a*b = e$. Let c be an inverse of b. By associativity we have $(b*a)*b = b*(a*b) = b*e = b$. Therefore $e = b*c = ((b*a)*b)*c = (b*a)*(b*c) = (b*a)*e = b*a$.

To prove the second claim, suppose $a * b = e = a * c$. Then $b = e * b = (b * a) * b = b * (a * b) = b * (a * c) = (b * a) * c = e * c = c$.

To prove the third claim, suppose e and f are both identities in G. That is, for all $a \in G$, $a * e = e * a = a * f = f * a = a$. Then (taking $a = f$) $f * e = f$. But also (taking $a = e$) $f * e = e$. Thus $f = e$. □

Sometimes we use multiplicative notation and write a^{-1} to denote the inverse of a, ab for $a * b$, $a^0 = e$, and $a^n = aa^{n-1}$ for $n \geq 1$. Then $a^n a^m = a^{n+m}$ and $(a^n)^m = a^{nm}$. If G is Abelian, it is common to use *additive notation* in which we write $+$ instead of $*$, $-a$ instead of a^{-1}, $a - b$ for $a + (-b)$, and 0 instead of e. We sometimes write $e = e_G$ when considering several different groups.

Examples:

1. The integers $\mathbf{Z}$ with identity 0 and addition as operation is an Abelian group.

2. The rational numbers $\mathbf{Q}$ with identity 0 and addition as operation is an Abelian group.

3. The nonzero rational numbers $\mathbf{Q} - \{\mathbf{0}\}$ with identity 1 and multiplication as operation is an Abelian group.

4. If S is any set, the set of permutations of S is a (non-Abelian if $|S| \geq 3$) group with composition as operation and the identity function as identity.

5. For any $n \geq 1$, the set of invertible $n \times n$ matrices (that is, with nonzero determinant) with rational entries is a (non-Abelian if $n \geq 2$) group with multiplication as operation and the $n \times n$ identity matrix as identity.

6. If $m \geq 2$ is an integer, then a is congruent to b modulo m, written $a \equiv b \pmod{m}$, if m divides $a - b$. This is an equivalence relation on $\mathbf{Z}$. Let $\mathbf{Z}/m\mathbf{Z}$ denote the set of equivalence classes for this relation. That is, $\mathbf{Z}/m\mathbf{Z}$ is the set of sets of the form

$$a + m\mathbf{Z} = \{a + mb : b \in \mathbf{Z}\}.$$

Then $\mathbf{Z}/m\mathbf{Z}$ is an Abelian group with the operation $(a + m\mathbf{Z}) + (b + m\mathbf{Z}) = (a + b) + m\mathbf{Z}$ and $0 + \mathbf{Z}$ as identity. To prove this it suffices to show that this definition of addition is independent of the choice of representatives a and b (that is, if $a + m\mathbf{Z} = c + m\mathbf{Z}$ and $b + m\mathbf{Z} = d + m\mathbf{Z}$, then $(a + b) + m\mathbf{Z} = (c + d) + m\mathbf{Z}$) and that the group axioms for $\mathbf{Z}/m\mathbf{Z}$ follow immediately from the group axioms for $\mathbf{Z}$. The set of equivalence classes of elements that are relatively prime to m, denoted $(\mathbf{Z}/m\mathbf{Z})^*$, is also an Abelian group, with multiplication as operation and 1 as unit.

Following is a basic fact about groups that we shall use later.

Theorem 1.1.3 *If G is a finite group and $a \in G$, then $a^{|G|} = e$.*

Proof: First suppose that G is Abelian. Let us define a function from G to itself by $f(b) = ab$. This function is one-to-one (if $ab = ac$ then multiplying by a^{-1} gives $b = b$), so it is also onto. Therefore

$$\prod_{b \in G} b = \prod_{b \in G} ab = a^{|G|} \prod_{b \in G} b.$$

Multiplying by the inverse of $\prod_{b \in G} b$ gives the result of the theorem.

Now suppose that G is arbitrary. It is nonetheless the case that $H = \{a^i : i = 0, 1, \cdots\}$ is an Abelian group, so $a^{|H|} = e$. Thus it suffices to show that $|H|$ divides $|G|$. Consider the cosets bH with $b \in G$. Suppose two of these have a nonempty intersection, $bH \cap cH \neq \emptyset$. Then there are integers i, j so that $ba^i = ca^j$. It follows from this that every ba^k is in cH and every ca^k is in bH. That is, $bH = cH$. This implies that the set of all cosets bH forms a partition of G. Since each bH has cardinality $|H|$, $|G|$ is a multiple of $|H|$ as desired. □

1.1.b Subgroups

Definition 1.1.4 *If G is a group, then a subset $H \subseteq G$ is a subgroup if it is a group with the same operation as G and the same identity as G.*

This means that H is a subset of G such that (1) $e \in H$; (2) if $a, b \in H$, then $a + b \in H$; and (3) if $a \in H$, then $a^{-1} \in H$. Then the group axioms hold in H. Also, if G is Abelian then H is Abelian. The *order* of a group G is its cardinality as a set.

For example, the additive group of integers is a subgroup of the additive group of rational numbers. The set of cyclic permutations of $\{1, 2, \cdots, n\}$ is a subgroup of the group of all permutations.

If G_1 and G_2 are groups with operations $*_1$ and $*_2$ and identities e_1 and e_2, then their *direct product* $G_1 \times G_2 = \{(a, b) : a \in G_1, b \in G_2\}$ is a group with operation $(a, b) * (c, d) = (a * c, b * d)$ and identity (e_1, e_2). More generally, if $\{G_i : i \in I\}$ is any collection of groups, indexed by a set I, then the Cartesian product

$$\prod_{i \in I} n G_i$$

is a group, again called the direct product of $\{G_i : i \in I\}$. The group operation is defined coordinate-wise. If all the groups are Abelian, then so is the product. If $I = \{1, 2, \cdots, n\}$ for some natural number n, then we write

$$\prod_{i \in I} n G_i = \prod_{i=1}^{n} G_i = G_1 \times G_2 \times \cdots \times G_n.$$

If $a \in G$ then we let $\langle a \rangle$ denote $\{a^i : i \in \mathbf{Z}\}$. This set is an Abelian subgroup, called the *subgroup generated by* a. If $\langle a \rangle$ has finite order then we say the *order of* a is the order of $\langle a \rangle$. Otherwise we say a has infinite order. Equivalently, the order of a is the least $k > 0$ such that $a^k = e$, if such a k exists. A group is *cyclic* if $G = \langle a \rangle$ for some a and then a is called a generator of G. So every infinite cyclic group is isomorphic to the integers $\mathbf{Z}$ and every finite cyclic group is isomorphic to the (additive) group $\mathbf{Z}/(n)$ where n is the order of any generator.

Theorem 1.1.5 *Every subgroup of a cyclic group is cyclic. Suppose $\langle a \rangle$ is a finite cyclic group with order n.*

1. *If k is a positive integer, then $\langle a^k \rangle$ is a subgroup of $\langle a \rangle$ of order $n/\gcd(n,k)$.*
2. *If $d|n$ and $d > 0$, then $\langle a \rangle$ contains one subgroup of order d.*
3. *If $d|n$ and $d > 0$, then $\langle a \rangle$ contains $\phi(d)$ elements of order d. ($\phi(d)$ is Euler's phi function, the number of positive integers less than d and relatively prime to d.)*
4. *$\langle a \rangle$ contains $\phi(n)$ generators.*

Proof: Let H be a nontrivial subgroup of $\langle a \rangle$. H contains some a^k with $k > 0$. Let k be the smallest positive integer with $a^k \in H$ and let $a^m \in H$. Suppose k does not divide m. Then $\gcd(k,m) < k$ and $\gcd(k,m) = sk + tm$ for some integers s, t. Then

$$a^{\gcd(k,m)} = (a^k)^s (a^m)^t \in H,$$

which is a contradiction. Therefore $H = \langle a^k \rangle$. Thus every subgroup of $\langle a \rangle$ is cyclic.

(1) Let $H = \langle a^k \rangle$ and $d = \gcd(n,k)$. We have $(a^k)^r = e$ if and only if $n|kr$. Thus the order of H is the least positive r such that $n|kr$. This is equivalent to $(n/d)|(k/d)r$, and this is equivalent to $(n/d)|r$. That is, the order of H is n/d.

(2) By (1), a subgroup $H = \langle a^k \rangle$ has order $d|n$ if and only if $d = n/\gcd(n,k)$, or, equivalently, $d \cdot \gcd(n,k) = n$. Let $f = \gcd(n,k) = sn + tk$ for some $s, t \in \mathbf{Z}$. Then $e = a^n \in H$, so $a^f \in H$ as above. Since $f|k$, we also have $H = \langle a^f \rangle$. But $f = n/d$ so f is unique. Conversely, $\langle a^{n/d} \rangle$ is always a subgroup of order d.

(3) Let $n = df$. By (1), an element a^k has order d if and only if $\gcd(n,k) = n/d = f$. This holds precisely when $k = gf$ with g relatively prime to $n/f = d$ and $0 < k < n$. That is, $0 < g < d$. The number of such g is $\phi(d)$.

(4) Follows immediately from (3) with $d = n$. □

For example, the group $\mathbf{Z}$ is cyclic (with generator 1) so every subgroup is of the form $m\mathbf{Z} = \{mk : k \in \mathbf{Z}\}$ for some integer m.

1.1.c Homomorphisms

More generally, relationships between groups often arise as functions from one group to another that preserve all the relevant algebraic structures and operations.

Definition 1.1.6 *Let G and H be two groups. A function $\varphi : G \to H$ is a* homomorphism *if it preserves the group operations. That is, if for every $a, b \in G$ we have $\varphi(ab) = \varphi(a)\varphi(b)$. The* image *of φ, denoted by Im(φ), is the set of $b \in H$ such that there is $a \in G$ with $\varphi(a) = b$. The* kernel *of φ, denoted by ker(φ), is the set of $a \in G$ such that $\varphi(a) = e_H$. The homomorphism φ is an* endomorphism *if $G = H$. It is an* epimorphism *or is* surjective *if it is onto as a set function. It is a* monomorphism *or is* injective *if it is one-to-one as a set function. It is an* isomorphism *if it is both injective and surjective. It is an* automorphism *if it is an endomorphism and an isomorphism.*

Proposition 1.1.7 *Let $\varphi : G \to H$ be a homomorphism. Then φ preserves identities and inverses. Morever ker(φ) is a subgroup of G and Im(φ) is a subgroup of H.*

Proof: To see that φ preserves identities observe that $\varphi(e_G) = \varphi(e_G e_G) = \varphi(e_G)\varphi(e_G)$. Multiplying by $\varphi(e_G)^{-1}$ then gives $e_H = \varphi(e_G)$. To see that φ preserves inverses, let $a \in G$. Then $e_H = \varphi(e_G) = \varphi(aa^{-1}) = \varphi(a)\varphi(a^{-1})$ so $\varphi(a)^{-1} = \varphi(a^{-1})$ by uniqueness. The remaining statements are equally easy. □

Proposition 1.1.8 *If $f : F \to G$ and $g : G \to H$ are homomorphisms, then the composition $g \circ f : F \to H$ is a homomorphism.*

Proof: For all $a, b \in F$, we have $(g \circ f)(a + b) = g(f(a + b)) = g(f(a) + f(b)) = g(f(a)) + g(f(b))$. Similarly for multiplication. □

Definition 1.1.9 *A pair of homomorphisms $\phi : F \to G$ and $\psi : G \to H$ is* exact *(at G) if the kernel of ψ equals the image of ϕ. A sequence of maps*

$$1 \to F \to G \to H \to 1 \tag{1.1}$$

is a short exact sequence *if it is exact at F, G, and H. Here 1 denotes the trivial group with a single element.*

The short exact sequence in (1.1) splits *if there is a homomorphism $h : H \to G$ so that $g \cdot h$ is the identity.*

Proposition 1.1.10 *If $1 \to F \to G \to H \to 1$ is a short exact sequence and all three groups are finite, then $|G| = |F| \cdot |H|$.*

Proof: Let ϕ denote the homomorphism from F to G, and let ψ denote the homomorphism from G to H. Since ψ is surjective, there is a subset U of G that maps one to one and onto H. If g is any element of G, then there is some $u \in U$ so that $\psi(u) = \psi(g)$. Then gu^{-1} maps to the identity in H, so $gu^{-1} = f \in \mathrm{Im}(\phi)$. Thus we can write $g = fu$ with $f \in \mathrm{Im}(\phi)$. Suppose that $fu = f'u'$ for some $f, f' \in \mathrm{Im}(\phi)$ and $u, u' \in U$. Then $uu^{-1} = (f')^{-1}f \in \mathrm{Im}(\phi)$. It follows that $\psi(uu^{-1}) = e_H$, so $\psi(u) = \psi(u')$. By the choice of U, we have $u = u'$. Then also $f = f'$. It follows that for each g there is a unique representation of u in the form $g = fu$. The proposition is immediate from this. □

Proposition 1.1.11 *Suppose* $1 \to F \to G \to H \to 1$ *is a short exact sequence with* $\phi : F \to G$ *and* $\psi : G \to H$*, all three groups are Abelian, and the short exact sequence splits via a homomorphism* $\mu : H \to G$*, then there is an isomorphism between* $F \times H$ *and* G *given by* $(a, b) \mapsto \phi(a)\mu(b)$*. Conversely, if* $G = F \times H$*, then there is a short exact sequence as in (1.1), where* g *is the projection map and* f *maps* a *to* $(a, 1)$*.*

Proof: In Proposition 1.1.10 we can take U to be the image of μ to prove the first statement. The converse is trivial. □

1.1.d Quotients

Recall that $\mathbf{Z}$ is a group. If m is any positive integer, then the set of multiples of m, $m\mathbf{Z}$, is a subgroup. We defined an equivalence relation by saying $a \equiv b \pmod{m}$ if $a - b \in m\mathbf{Z}$. We then formed a group $\mathbf{Z}/m\mathbf{Z}$ whose elements are the equivalence classes for this relation.

More generally, suppose G is any group and H is a subgroup of G. We can form an equivalence relation by saying $a \sim b$ if there is an $h \in H$ such that $b = ah$ (The proof that this is an equivalence relation is left as an exercise). The equivalence class of a is called the *left coset of* a. It is exactly aH. We would like to form a group consisting of the equivalence classes $\{aH : a \in G\}$. Unfortunately, this does not work in general In fact, we could have started by defining $a \sim' b$ if there is an $h \in H$ such that $b = ha$. This is also an equivalence relation. The equivalence class of a with respect to this relation is called the *right coset of* a. It is exactly Ha. We can form a group out of the equivalence classes exactly when the left and right cosets are the same.

Definition 1.1.12 *If* H *is a subgroup of* G*, then* H *is* normal in G *if for every* $a \in G$*, we have* $aH = Ha$*.*

Equivalently, H is normal in G if for every $a \in G$ and $h \in H$, we have $aha^{-1} \in H$. If the group G is Abelian then every subgroup H is normal in G.

Theorem 1.1.13 *If H is normal in G, then the set of left cosets of G modulo H, denoted G/H, is a group under the operation $(aH)(bH) = abH$.*

Proof: Left as an exercise. □

In this case, G/H is called the *quotient group of G modulo H*. The natural mapping $G \to G/H$ (given by $a \mapsto aH$) is a homomorphism. If the set of left cosets is finite, then we say H has finite index in G. The number of left cosets (which equals the number of right cosets) is called the *index of H in G*. Thus if H is normal in G and of finite index, then G/H is finite and $|G/H|$ equals the index of H in G. If G is finite, so is G/H, and we have $|G/H| = |G|/|H|$.

Theorem 1.1.14 *If $\varphi : G \to G'$ is a homomorphism then the following statements hold.*

1. *ker(φ) is normal in G.*
2. *The quotient $G/$ker(φ) is isomorphic to Im(φ).*
3. *Conversely, if H is a normal subgroup of G, then the natural mapping $a \mapsto aH$ is a surjection from G to G/H with kernel equal to H.*

Proof: Left as an exercise. □

1.1.e Finitely generated Abelian groups

An Abelian group G is *finitely generated* if there is a finite set $V \subseteq G$ such that every element of G is equal to a finite product of elements of V. We state without proof the fundamental theorem of finite Abelian groups (See, for example, Lang [14, p. 46]):

Theorem 1.1.15 *Let G be a finitely generated Abelian group. Then G is isomorphic to a direct product of cyclic groups.*

Corollary 1.1.16 *Let G be a finite Abelian group with nm elements, where n and m are relatively prime positive integers. Then there are groups H_1 and H_2 with n and m elements, respectively, so that G is isomorphic to $H_1 \times H_2$.*

An element g in an Abelian group G is a *torsion* element if $g \neq 0$ and if some finite sum $g + g + \cdots + g = 0$ vanishes. That is, if it has finite order. The group G is *torsion-free* if it contains no torsion elements.

Corollary 1.1.17 *Let G be a finitely generated torsion-free Abelian group. Then G is isomorphic to a direct product of finitely many copies of $\mathbf{Z}$.*

1.2 Rings and Fields

Many important algebraic structures come with two interrelated operations. For example, addition and multiplication of integers, rational numbers, real numbers, and complex numbers; AND and XOR of Boolean valued functions; and addition and multiplication of $n \times n$ matrices of integers, etc.

Definition 1.2.1 *A ring R is a set with two binary operations $+$ and $\cdot$ and two distinguished elements $0, 1$ which satisfy the following properties for all $a, b, c \in R$:*

1. *R is an Abelian group with operation $+$ and identity 0;*
2. *$a \cdot (b \cdot c) = (a \cdot b) \cdot c$ and $1 \cdot a = a \cdot 1 = a$; and*
3. *$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ and $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$ (the distributive law).*

It follows that $a \cdot 0 = 0$ for all a, since $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. If $0 = 1$ then $R = \{0\}$ is the zero ring. It is common to denote by R^+ the Abelian group that is obtained from R by forgetting the multiplication.

A ring R is *commutative* if $a \cdot b = b \cdot a$ for all $a, b \in R$. *Throughout this book, all rings are commutative unless otherwise stated.* We generally write ab for the product $a \cdot b$.

1.2.a Units and zero divisors

Let R be a commutative ring. An element $a \in R$ is a *unit* if it is invertible, that is, if there exists $b \in R$ so that $ab = 1$. In this case b is unique. The collection of all units in R is denoted $R^\times$. It forms an Abelian group (under multiplication). An element $a \in R$ is a *zero divisor* if there exists a nonzero element $b \in R$ such that $ab = 0$. The ring of integers $\mathbf{Z}$ has no zero divisors, but only 1 and -1 are units. However if the ring R is finite then a given element is either a unit or a zero divisor. Indeed, let $\varphi_a : R \to R$ be the mapping which is given by multiplication by a. If φ_a is one to one, then it is also onto, hence a is invertible. If φ_a is not one to one, then there exist $b \neq c$ so that $ab = ac$ or $a(b - c) = 0$, so a is a zero divisor.

Definition 1.2.2 *An integral domain is a commutative ring with no zero divisors. A field is a commutative ring in which every nonzero element is invertible.*

In particular, a finite integral domain is necessarily a field. Every commutative ring R embeds in a ring $S^{-1}R$ which has the property that every element is either a zero divisor or is invertible, cf. §1.2.d.

1.2.b Ideals and Quotients

Definition 1.2.3 *A* subring *S of a ring R is a subset of R, which is a ring under the same operations as R, and with the same zero and identity.*

If I is an additive subgroup of R (meaning that if $a, b \in I$ then $a + b \in I$ and $-a \in I$) then the quotient R/I is the set of equivalence classes under the equivalence relation $a \sim b$ if $a-b \in I$. The equivalence class containing $a \in R$ is the coset $a+I$. Then R/I is an Abelian group under addition: $(a + I) + (b + I) = a + b + I$. However, the multiplication operation on R does not necessarily induce a well defined multiplication on R/I. For if $a' \sim a$, say, $a' = a+c$ and if $b' \sim b$, say, $b' = b+d$ (where $c, d \in I$) then $a'b' = ab+ad+bc+cd$ which is not necessarily equivalent to ab unless $ad+bc+cd \in S$. The following definition is necessary and sufficient to ensure this holds for all $a, b \in R$ and $c, d \in I$.

Definition 1.2.4 *An ideal is an additive subgroup $I \subset R$ such that for any $a \in I$ and for any $b \in R$ we have: $ab \in I$.*

It follows that the set of equivalence classes R/I inherits a ring structure from R if and only if I is an ideal. Two elements $a, b \in R$ are said to be *congruent modulo I* if they are in the same equivalence class. That is, if $a - b \in I$. Each equivalence class is called a *residue class modulo I*.

An ideal I is *proper* if $I \neq R$, in which case it does not contain any units. An ideal I is *principal* if there exists an element $a \in R$ such that $I = \{ar : r \in R\}$, in which case we write $I = (a)$. If I, J are ideals then the sum $I + J$ is the set of all sums $a + b$ where $a \in I$ and $b \in J$. It is the smallest ideal containing both I and J. The intersection $I \cap J$ is also an ideal. The product ideal IJ is the set of all finite sums $\sum a_i b_i$ where $a_i \in I$ and $b_i \in J$. An ideal $I \subset R$ is *maximal* if $I \neq R$ and if I is not a proper subset of any other proper ideal. An ideal I is *prime* if $ab \in I$ implies $a \in I$ or $b \in I$. An ideal $I \subset R$ is *primary* if $I \neq R$ and whenever $ab \in I$, then either $a \in I$ or $b^n \in I$ for some $n \geq 1$.

A field contains only the ideals (0) and (1).

Theorem 1.2.5 *Let R be a commutative ring. Then the following statements hold.*

1. *An ideal $P \subset R$ is maximal if and only if R/P is a field (called the* residue field with respect to *P).*
2. *An ideal $P \subset R$ is prime if and only if R/P is an integral domain. (See Definition 1.2.10.)*
3. *Every maximal ideal is prime.*

Proof: (1) Let P be maximal and $a \in R - M$. Then $J = \{ab + c : b \in R, c \in P\}$ is closed under addition and under multiplication by elements of R. It contains P (take $b = 0$) and a (take $b = 1$ and $c = 0$) so it properly contains P. Then by maximality it is not an ideal, so it must not be a proper subset of R. That is, $J = R$. In particular, $1 \in J$, so $1 = ab + c$ for some $b \in R$ and $c \in P$. Therefore $(a + P)(b + P) = ab + P = 1 - c + P = 1 + P$ so $a + P$ is invertible in R/P. Thus R/P is a field. On the other hand, suppose R/P is a field and J is an ideal containing P. Let $a \in J - P$. Then $a + P$ is invertible in R/P, so there is a $b \in R$ such that $(a + P)(b + P) = 1 + P$. That is, such that $ab = 1 + cm$ for some $c \in P$. But then $1 = ab - c \in J$. By closure under multiplication by R, we have $R \subseteq J$. But this contradicts the fact that J is an ideal. Therefore P is maximal.

(2) Let $a, b \in R$. Then $(a + P)(b + P) = 0$ in R/P if and only if $ab \in P$. If P is prime, this says $(a+P)(b+P) = 0$ implies $a \in P$ or $b \in P$, which implies $a+P = 0$ or $b+P = 0$ in R/P, so R/P is an integral domain. Conversely, if R/P is an integral domain, then $ab \in P$ implies $(a + P)(b + P) = 0$ which implies $a + P = 0$ or $b + P = 0$. That is, $a \in P$ or $b \in P$, so P is a prime ideal.

(3) This follows from (1) and (2). □

For example, consider the ring of ordinary integers $\mathbf{Z}$. Let I be an ideal containing a nonzero element. Multiplication by -1 preserves membership in I, so I contains a positive element. Let m be the least positive element of I. Suppose that $a \in I$ is any other element of I. Then $\gcd(m, a) = um + va$ for some integers u and v, so $\gcd(m, a) \in I$. We have $\gcd(m, a) \leq m$, so by the minimality of m, $\gcd(m, a) = m$. That is, m divides a. Since every multiple of m is in I, it follows that I consists exactly of the multiples of m. In particular, $I = (m)$ is principal.

The ideal (m) is contained in the ideal (n) if and only if m is a multiple of n. The ideal (m) is prime if and only if m is prime. In this case it is also maximal. It is primary if and only if m is a power of a prime.

Definition 1.2.6 *A function $\varphi : R \to S$ from a ring R to a ring S is a* ring homomorphism *if $\varphi(a + b) = \varphi(a) + \varphi(b)$ and $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in R$. The homomorphism φ is a* surjection *(or* epimorphism*) if it is onto. It is an* injection *(or* monomorphism*) if it is one to one. It is an* isomorphism *if it is both an injection and a surjection. It is an* endomorphism *if $R = S$. It is an* automorphism *if it is an endomorphism and an isomorphism.*

The set of automorphisms of a ring S forms a group under composition, denoted by $\text{Aut}(S)$. More generally, if R is a subring of S (we also say that S is an *extension of* R), then the set of automorphisms of S whose restrictions to R are the identity forms a subgroup $\text{Aut}_R(S)$.

The proof of the following theorem is left as an exercise.

Theorem 1.2.7 *If $\varphi : R \to S$ is a ring homomorphism, then*

$$ker(\varphi) = \{r \in R : \ \varphi(r) = 0\}$$

is an ideal of R, the image of φ is a subring of S, and φ induces an isomorphism between $R/ker(\varphi)$ and the image of φ. Conversely, if I is an ideal of R then the map $a \mapsto a + I$ is a surjective homomorphism from $R \to R/I$ with kernel I.

If F is a field and E is a ring, then the kernel of any non-zero homomorphism $F \to E$ is the zero ideal (the only ideal), so every homomorphism is an injection. We say that E is an *extension* of F. If $F \subset E$ are fields then the group $\mathrm{Aut}_F(E)$ of automorphisms of E which fix each element of F is the *Galois group* of E over F and it is denoted by $\mathrm{Gal}(E/F)$. In general, if G is a subgroup of the group of automorphisms of E, then the set of elements in E that are fixed by every automorphism in G (that is, $\sigma(a) = a$ for every $a \in E$ and every $\sigma \in G$) is denoted E^G. It is necessarily a field since it is closed under addition, multiplication, and inverse. If $G = \mathrm{Gal}(E/F)$, then $F \subseteq E^G$. If in fact $F = E^G$, then we say that *E is a Galois extension of F*. The general theory of Galois extensions is venerable, and its invention by Galois was a turning point in the understanding of the nature of algebraic equations.

1.2.c Characteristic

Let R be a commutative ring. If m is a nonnegative integer, we write $m \in R$ for the sum $1 + 1 \cdots + 1$ (m times). This defines a homomorphism from $\mathbf{Z}$ into R. That this function is a homomorphism can be shown by a series of induction arguments. In fact this is the unique homomorphism from $\mathbf{Z}$ into R, since any such homomorphism is completely determined by the fact that $1_{\mathbf{Z}}$ maps to 1_R, and the ring operations are preserved. The kernel of this homomorphism is an ideal in $\mathbf{Z}$, hence by the example in §1.2.b is of the form (m) for some nonnegative integer m. This integer is called the *characteristic* of R. For any $a \in R$, we have $ma = a + a + \cdots + a$ (m times). Hence if the characteristic is nonzero, it is the smallest positive integer m such that $ma = 0$ for all $a \in R$. If the characteristic is zero, then no such m exists and $\mathbf{Z}$ is isomorphic to a subring of R. Otherwise $\mathbf{Z}/(m)$ is isomorphic to a subring of R. If R is finite then its characteristic is positive since the sequence of elements $1, 1 + 1, 1 + 1 + 1, \cdots$ must eventually lead to a repetition.

Theorem 1.2.8 *If R is an integral domain then its characteristic is either 0 or is a prime number. In particular, the characteristic of any finite field is prime.*

Proof: Let $k > 0$ be the characteristic and suppose $k = mn$, with $m > 0$ and $n > 0$. Let $a \in R$ be the element $1 + \cdots + 1$ (m times) and let $b \in R$ be the element $1 + \cdots + 1$ (n times). Then $ab = 0$, so $a = 0$ or $b = 0$. Suppose $a = 0$. For any $c \in R$, the element $c + \cdots + c$ (m

times) is $ac = mc = 0$. By the minimality of k, we must have $m = k$ and $n = 1$. A similar argument holds when $b = 0$. It follows that k is prime. □

Lemma 1.2.9 *Let R be a commutative ring. If the characteristic k of R is a prime number, and if q is any positive power of k then*

$$(a+b)^q = a^q + b^q \in R \tag{1.2}$$

for every $a, b \in R$.

Proof: If k is prime and if $0 < m < k$, the binomial coefficient $\binom{k}{m} = k!/m!(k-m)!$ is divisible by k since k appears as a factor in the numerator but not in the denominator. Consequently $(a+b)^k = a^k + b^k$ and equation (1.2) follows by induction. □

If k is not prime, then equation (1.2) is generally false.

1.2.d Divisibility in rings

Let R be a commutative ring. If $a, b \in R$ then a is a *divisor* of b if there exists $c \in R$ such that $ac = b$, in which case we write $a|b$. The element a is a *unit* if it is a divisor of 1. Elements $a, b \in R$ are *associates* if $a = \epsilon b$ for some unit ϵ. A non-zero element $c \in R$ is a *common divisor* of a and b if $c|a$ and $c|b$. It is a *greatest common divisor* of a and b (written $c = \gcd(a, b)$) if it is a common divisor and if every other common divisor of a and b divides c. An element $c \neq 0$ is a *common multiple* of a and b if $a|c$ and $b|c$. It is a *least common multiple* (written $c = \mathrm{lcm}(a, b)$) if it is a common multiple and if it divides every other common multiple of a and b.

A nonzero element $r \in R$ is *prime* if (r) is a proper prime ideal. It is *primary* if (r) is primary. It is *irreducible* if it is not a unit and if $r = ab$ implies that a or b is a unit. Two nonzero non-units $r, s \in R$ are *coprime* if $(r) + (s) = R$.

Definition 1.2.10 *Let R be a commutative ring.*

1. *R is an* integral domain *(or is* integral*) if it has no zero divisors.*
2. *R is* principal *if every ideal in R is principal. It is a* principal ideal domain *or* PID *if it is principal and is an integral domain.*
3. *R is a* GCD ring *if every pair of elements has a greatest common divisor.*
4. *R is a* local ring *if it contains a unique maximal ideal.*
5. *R is a* unique factorization domain *(or* UFD*, or* factorial*) if it is an integral domain and every nonunit $a \in R$ has a factorization into a product*

$$a = \prod_{i=1}^{m} p_i \tag{1.3}$$

of irreducible elements (not necessarily distinct), which is unique up to reordering of the p_is and multiplication of the p_is by units. That is, if $a = \prod_{i=1}^{n} q_i$, then $m = n$ and there is a permutation σ of $\{1, \cdots, m\}$ so that p_i and $q_{\sigma(i)}$ are associates.

6. *R is* Euclidean *if there is a function $\delta : R \to \{0, 1, 2, \cdots\} \cup \{-\infty\}$ such that (1) for every $a, b \in R$ with a and b both nonzero, we have $\delta(ab) \geq \delta(a)$, and (2) for every $a, b \in R$ with $b \neq 0$ there exist $q, r \in R$ so that*

$$a = qb + r \quad \text{and} \quad \delta(r) < \delta(b).$$

(We say that a divided by b has quotient q and remainder r.)

Theorem 1.2.11 summarizes the various inclusions among the special types of rings that we have discussed. We have included the polynomial ring $R[x]$ for ease of reference although it will not be defined until Section 1.4.

Theorem 1.2.11 *Let R be a commutative ring and let $R[x]$ be the ring of polynomials with coefficients in R (see §1.4). Then we have the following diagram of implications between various possible properties of R.*

$$\begin{tikzcd}
\textit{field} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \textit{Euclidean} \arrow[r, Rightarrow] & \textit{PID} \arrow[r, Rightarrow] & \textit{UFD} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \textit{integral} \arrow[r, Rightarrow] & R[x]\textit{integral} \\
R[x]\textit{Euclidean} & & & \textit{GCD} & &
\end{tikzcd}$$

Proof: The properties of the polynomial ring $R[x]$ are proved in Lemma 1.4.1 and Theorem 1.4.2. If R is a field then it is Euclidean with $\delta(0) = -\infty$ and $\delta(r) = 0$ for all nonzero elements $r \in R$.

To show that every Euclidean ring is a PID, let R be Euclidean. Suppose $a \in R$ is nonzero. We can write $0 = qa + r$ with $\delta(r) < \delta(a)$. Suppose that q is nonzero. Then $\delta(r) = \delta(-qa) \geq \delta(a)$, which is a contradiction. Thus $q = 0$ so $r = 0$. But then we must have $\delta(0) < \delta(a)$ for every $a \neq 0$. In particular, $\delta(a) \geq 0$ if a is nonzero. Now let I be a nonzero ideal in R. Let $a \in I - \{0\}$ be an element such that $\delta(a)$ is minimal. There is at least one such element since $\delta(I - \{0\}) \subset \mathbf{N}$ has a least element (by the well ordering principal). We claim that $I = (a)$. Let b be any other element in I. Then $b = qa + r$ for some $q, r \in R$ such that $\delta(r) < \delta(a)$. But $r = b - qa \in I$, so $r = 0$. That is, $b = qa$, as claimed. Moreover, if $0 = ab$ for some nonzero a, then the argument above shows that $b = 0$, so R is an integral domain.

Now assume that R is a PID. If a and b are two elements of R, then the ideal (a, b) has a principal generator, $(a, b) = (c)$. Thus c divides both a and b, and $c = ua + vb$ for some $u, v \in R$. Therefore any common divisor of a and b divides c as well. That is, c is a GCD of a and b. It follows that R is a GCD ring. It also follows that the GCD c can be written in the form $c = ua + vb$.

Lemma 1.2.12 *If R is a PID then every (properly increasing) chain of ideals $(a_1) \subset (a_2) \subset \cdots$ is finite.*

Proof: The union of such a chain is again an ideal, hence is principal, say with generator a. Then the element a lies in one of the ideals in the chain, say $a \in (a_n)$. Hence $(a_n) \subset \cup_i (a_i) = (a) \subset (a_n)$ so the chain stops at (a_n). □

Next we show that R is a UFD. We first prove that every element $a \in R$ has a prime factorization. Let S be the set of elements of R that do not have prime factorizations, and suppose S is nonempty. Any chain $(a_1) \subseteq (a_2) \subseteq \cdots$ with every $a_i \in S$, is finite by Lemma 1.2.12. Thus there is an element $a \in S$ such that the generator of every ideal properly containing a has a factorization. The element a cannot be irreducible, so we have $a = bc$ with neither b nor c a unit. Hence (a) is a proper subideal of (b) and of (c), and b and c have prime factorizations. The product of these factorizations is a factorization of a, contradicting the fact that $a \in S$.

Next we prove uniqueness. Suppose $a \in R$ is irreducible and $a|bc$. If $a \not| b$, then 1 is a gcd of a and b, so we have

$$1 = ua + vb,$$

for some $u, v \in R$. Thus $c = uac + vbc$, so $a|c$. That is, if $a|bc$, then $a|b$ or $a|c$. In other words, a is prime if a is irreducible.

Suppose some nonunit $b \in R$ can be factored in two ways,

$$b = \prod_{i=1}^{k} p_i = \prod_{i=1}^{\ell} q_i.$$

Since b is not a unit, we have $k > 0$ and $\ell > 0$. We use induction on k. Since $p_k | \prod_{i=1}^{\ell} q_i$, we have $p_k | q_n$ for some n by the primality of p_k, say $q_n = dp_1$. By the irreducibility of p_k and q_n, d is a unit. Then $\prod_{i=1}^{k-1} p_i = d(\prod_{i=1}^{\ell} q_i)/q_n$, and the result follows by induction. This completes the proof that R is a UFD.

The implication UFD $\Longrightarrow$ GCD is obvious. Every UFD is integral by definition. □

In particular, in a PID, two elements are coprime if and only if 1 is a GCD. Note that for finite rings R these distinctions are irrelevant since a finite integral ring is a field.

Theorem 1.2.13 *Let R be a commutative ring and let $a, b \in R$. Then*

1. *The element a is prime if and only if it has the following property: if $a|cd$ then $a|c$ or $a|d$.*
2. *If a is prime and is not a zero divisor, then a is irreducible.*

3. *If R is a UFD, then a is prime if and only if a is irreducible.*
4. *If a and b are coprime, then every common divisor of a and b is a unit.*
5. *If R is a PID and if every common divisor of a and b is a unit, then a and b are coprime.*
6. *If R is a PID and $a \in R$, then a is prime if and only if (a) is maximal (if and only if $R/(a)$ is a field).*

Proof: Part (1) is just a restatement of the definition that (a) is a prime ideal.

Now suppose a is prime and is not a zero divisor, and suppose $a = cd$. Then either $c \in (a)$ or $d \in (a)$; we may assume the former holds. Then $c = ea$ for some $e \in R$, so $a = cd = ead$ or $a(1 - ed) = 0$. Since a is not a zero divisor, we have $ed = 1$ hence d is a unit. This proves (2).

For part (3), first suppose that $a \in R$ is irreducible and let $cd \in (a)$. Then $cd = ae$ for some element $e \in R$. The right side of this equation is part of the unique factorization of the left side, so a must divide either c or d. Therefore either $c \in (a)$ or $d \in (a)$. The converse was already proven in part (2). (Note that a UFD contains no zero divisors, due to the unique factorization of 0.)

For part (4), supposing a and b are coprime, we may write $1 = ac + bd$ for some $c, d \in R$. If $e|a$ and $e|b$ then $a = fe$ and $b = ge$ for some $f, g \in R$. This gives $1 = (fc + gd)e$ so e is invertible.

For part (5), Suppose R is a PID. Given a, b the ideal $(a) + (b)$ is principal, so it equals (c) for some $c \in R$, which implies that $c|a$ and $c|b$. Therefore c is a unit, so $(a) + (b) = (c) = R$.

For part (6), we have already shown, in Theorem 1.2.5 that (a) maximal implies that a is prime. For the converse, suppose that (a) is prime and that $(a) \subset (b) \neq R$. Then b is not a unit, and $a = cb$ for some $c \in R$. Since the ring R is also a UFD, the element a is irreducible, so c is a unit. Therefore $(a) = (b)$ hence (a) must be maximal. □

1.2.e Fractions

Let R be a commutative ring. A subset S of R is *multiplicative* if it contains 1, does not contain 0, and is closed under multiplication. For example, we could take S to be the collection of all elements of R which are not zero divisors. If S is any multiplicative subset of R, we define the ring $S^{-1}R$ to be the collection of all formal symbols a/b (where $a \in R$ and $b \in S$), under the following equivalence relation: $a/b \sim a'/b'$ if $ab' = ba'$. Addition and multiplication of fractions are defined by the usual formulas:

$$\frac{a}{b} + \frac{a'}{b'} = \frac{ab' + a'b}{bb'}$$

and

$$\frac{a}{b}\frac{a'}{b'} = \frac{aa'}{bb'}.$$

The natural mapping $R \to S^{-1}R$ (which takes a to $a/1$) is an injection if S does not contain any zero divisors. Every element of S that is not a zero divisor has become invertible in $S^{-1}R$. If S is the set of elements that are not zero divisors, then an element of $S^{-1}R$ is either a zero divisor or else it is invertible (exercise). In this case, the ring $S^{-1}R$ is called the *ring of fractions* of R. If R is an integral domain then its ring of fractions is therefore a field, which is called the *fraction field* of R. See for example, §2.2.9 and §4.2.

1.2.f Examples

Here are a few standard examples of rings.

1. The integers $\mathbf{Z}$ is a Euclidean domain with $\delta(a) = |a|$.

2. The rational numbers $\mathbf{Q}$, the real numbers $\mathbf{R}$, and the complex numbers $\mathbf{C}$ are fields.

3. If $k = mn$ is a composite integer (with $m, n \geq 2$) then $\mathbf{Z}/k\mathbf{Z}$ is not an integral domain since $m \cdot n = 0$.

4. If R is a ring and S is a nonempty set, then the set of functions from S to R is a ring with the operations $(f + g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$. The zero is the function $z(x) = 0$ for all x, and the identity is the function $i(x) = 1$ for all x.

5. If R is a ring then the collection $R[x]$ of polynomials with coefficients in R (see §1.4) is a ring.

6. Let G be an Abelian group with operation $*$ and identity e. The set E of endomorphisms of G is a ring with the operations $+_E =$ "product" and $\cdot_E =$ "composition". The zero is the function $z(a) = e$ for all a, and the identity is the function $i(a) = a$ for all a.

7. If R_1 and R_2 are rings then their Cartesian product $R_1 \times R_2$ is a ring under the coordinate-wise operations of addition and multiplication.

Theorem 1.2.14 *(Chinese Remainder Theorem) Let R be a ring and let $I_1, \cdots, I_k$ be ideals such that $I_i + I_j = R$ for every $i \neq j$. Then for every $a_i, \cdots, a_k \in R$ there is an element $a \in R$ such that for every i, $a \equiv a_i \pmod{I_i}$. Furthermore, if $I = \cap_{j=1}^{k} I_j$, then*

$$R/I \cong \prod_{j=0}^{k} R/I_j.$$

Proof: For $k = 1$ the statement is trivial. If $k = 2$, then there are elements $b_1 \in I_1$ and $b_2 \in I_2$ so that $1 = b_1 + b_2$. Let $a = a_1 b_2 + a_2 b_1$.

Now suppose $k > 2$. For every i let $J_i = \prod_{j \neq i} I_j$. For every $i \geq 2$ there are elements $c_i \in I_1$ and $b_i \in I_i$ such that $1 = c_i + b_i$. In particular, $\prod_{i=2}^{k}(c_i + b_i) = 1$. This product is in $I_1 + J_1$, so $R = I_1 + J_1$. Similarly, $R = I_j + J_j$ for every j. By the theorem in the case of two ideals, there is an element $d_j \in R$ such that $d_j \equiv 1 \pmod{I_j}$ and $d_j \equiv 0 \pmod{J_j}$. Then $a = a_1 d_1 + \cdots + a_k d_k$ satisfies our requirements

For each i there is a reduction homomorphism φ_i from R/I to R/I_i. This induces a homomorphism φ from R/I to $\prod_{j=1}^{k} R/I_j$ whose kernel is $I = \cap_{j=1}^{k} I_j$. Thus φ is injective. By the first part it is surjective, hence an isomorphism. □

Corollary 1.2.15 *Suppose R is a Euclidean domain and $b_1, \cdots, b_k \in R$ are pairwise relatively prime. If $a_1, \cdots, a_k \in R$, then there exists an element $a \in R$ such that for every i, $a \equiv a_i \pmod{b_i}$.*

Proof: By Theorem 1.2.14 it suffices to show that for each $i \neq j$ we have $(b_i) + (b_j) = R$. Suppose not. Then $(b_i) + (b_j)$ is an ideal. Since Euclidean $\Longrightarrow$ PID (Theorem 1.2.11, there is some $b \in R$ so that $(b_i) + (b_j) = (b)$. This says that b is a common divisor of b_i and b_j, which is false by assumption. □

The case when $R = \mathbf{Z}$ is the classical Chinese Remainder Theorem.

1.2.g Vector Spaces

In many settings we have a notion of one algebraic object "acting on" another by multiplication. For example, a real number r acts on the set of points in the plane by $(x, y) \mapsto (rx, ry)$.

Definition 1.2.16 *A* vector space *over a field F is a set V such that V is an Abelian group with an operation $+$, and there is a function $\cdot$ from $F \times V$ to V such that for all $a, b \in F$ and $u, v \in V$*

1. *$a \cdot (u + v) = (a \cdot u) + (a \cdot v)$;*
2. *$(ab) \cdot u = a \cdot (b \cdot u)$;*
3. *$(a + b) \cdot u = (a \cdot u) + (b \cdot u)$; and*
4. *$1 \cdot u = u$.*

It follows from these axioms that for every $u \in V$, $0 \cdot u = 0$.

For example, the set of points in the real plane is a vector space over the real numbers. If F is a field which is a subring of a ring R, then R is vector space over F (just use the multiplication in R for the action of F on R). If F is a field and S is a nonempty set, then the set of functions from S to F is a vector space over F with the operations $(f+g)(x) = f(x)+g(x)$ and $(a \cdot f)(x) = af(x)$ for $a \in F$, $x \in S$, and $f, g : S \to F$. Various restrictions can be put on the functions to produce interesting vector spaces (e.g., continuity if $S = F = \mathbf{R}$).

Let V be a vector space over a field F. The elements of V are called *vectors*. A *linear combination* of vectors $v_1, v_2, \cdots, v_k \in V$ is a vector $a_1v_1+a_2v_2+\cdots+a_kv_k$ with $a_1, a_2, \cdots, a_k \in F$. A set of vectors $S \subseteq V$ is *linearly independent* if the only linear combination of elements of S that is zero is the one with all the coefficients a_i equal to zero. S *spans* V if every vector can be written as a linear combination of elements of S. S is a *basis* for V if it spans V and is linearly independent.

Theorem 1.2.17 *Let V be a vector space over a field F. If V has more than one element then it has a nonempty basis. If S is a basis, then every vector can be written uniquely as a linear combination of elements of S.*

Proof: Left as an exercise. □

If V has a basis S with a finite number of elements, then we say V is *finite dimensional with dimension* $= |S|$. In this case it can be shown that every basis has the same number of elements. In the important case when F is a subfield of a field E, E is called an *extension field.* If E is finite dimensional as a vector space over F, then its dimension is called *the degree of the extension* and is denoted $[E : F]$.

Theorem 1.2.18 *If F is a finite field and V is a finite dimensional vector space over F with dimension d, then $|V| = |F|^d$.*

Proof: Let S be a basis for V. Thus $|S| = d$. That is $S = \{v_1, v_2, \cdots, v_d\}$ for some $v_1, v_2, \cdots, v_d$. By the previous theorem, the elements of V are in one-to-one correspondence with the linear combinations $\sum_{i=1}^{d} a_iv_i$, $a_i \in F$. There are exactly $|F|^d$ such linear combinations. □

Definition 1.2.19 *If F is a field and V and W are vector spaces over F, then a function $L : V \to W$ is a* homomorphism *or is F-*linear *if it is a group homomorphism and for all $a \in F$ and $v \in V$ we have $L(av) = aL(v)$.*

If $S = \{v_1, v_2, \cdots, v_d\}$ is a basis for V, then an F-linear function L is completely determined by its values on the elements of S (because $L(\sum_i a_iv_i) = \sum_i a_iL(v_i)$). On the other hand, any choice of values for the $L(u_i)$ determines an F-linear function L. Furthermore, if $T = \{w_1, w_2, \cdots, w_e\}$ is a basis for W, then each value $L(v_i)$ can be expressed as a linear combination $L(v_i) = \sum_{j=1}^{e} b_{ij}w_j$ with $b_{ij} \in F$.

Theorem 1.2.20 *If F is finite and V and W are finite dimensional with dimensions d and e, respectively, then there are $|F|^{de}$ F-linear functions from V to W.*

The image and kernel of L are Abelian groups, and it is straightforward to check that they are also vector spaces over F. Their dimensions are called the *rank* and *co-rank* of L, respectively. We leave it as an exercise to show that the rank plus the co-rank equals the dimension of V.

We can identify an element $\sum_i a_i v_i \in V$ with the column vector $(a_1, \cdots, a_d)^t$, and similarly for an element of W. Then the linear function L is identified with ordinary matrix multiplication by the matrix $B = [b_{ij}]$. The rank of L is the size of a maximal set of independent columns or independent rows of B.

If B is a square matrix, then the determinant of B is defined as usual in linear algebra. In this case the kernel is nonempty if and only if the determinant is zero.

1.2.h Modules and Lattices

The notion of a vector space over a field can be generalized to rings.

Definition 1.2.21 *Let $(R, +, \cdot, 0, 1)$ be a commutative ring. A* module over R *is an Abelian group $(M, +, 0_M)$ with an operation $\cdot$ from $R \times M$ to M such that for all $a, b \in R$ and $u, v \in M$*

1. $a \cdot (u + v) = (a \cdot u) + (a \cdot v)$;
2. $(ab) \cdot u = a \cdot (b \cdot u)$;
3. $(a + b) \cdot u = (a \cdot u) + (b \cdot u)$; *and*
4. $1 \cdot u = u$.

Again, it follows from these axioms that for every $u \in V$, $0 \cdot u = 0$.

For example, every Abelian group is a module over the integers (if $n \in \mathbf{Z}^+$, then $n \cdot a$ equals the sum of n copies of a). If f is a homomorphism from a ring R to a ring S, then S is a module over R with the operation $a \cdot u = f(a)u$.

It is apparent that the notion of basis no longer makes sense for modules in general – even a single element of a module may not be linearly independent. However, if there is a finite set of elements $m_1, \cdots, m_k \in M$ such that every element of M can be written (perhaps not uniquely) as a linear combination $a_1 m_1 + \cdots + a_k m_k$ with $a_1, \cdots, a_k \in R$, then we say that M is *finitely generated over* R. If M is finitely generated, then the size of the smallest set of generators for M over R is called the *R-rank* or simply the rank of M.

A module M over a ring R is *free* if M is isomorphic to the Cartesian product of a finite number of copies of R. That is, M is free if there are elements $m_1, \cdots, m_k \in M$ such that every element $m \in M$ can be represented uniquely in the form $\sum_{i=1}^{k} c_i m_i$ with $c_i \in R$. In this case the set $m_1, \cdots, m_k$ is called a basis of M over R. A free $\mathbf{Z}$-module that is a subset of $\mathbf{R}^n$ for some n is called an *integer lattice* or $\mathbf{Z}$-lattice. If $n = k$, then it is said to be a full lattice. In this case a basis for M is also a basis for $\mathbf{R}^n$. If M is a full lattice in $\mathbf{R}^n$ with basis $m_1, \cdots, m_n$, then the set

$$P^c = \left\{ \sum_{i=1}^{n} z_i m_i : \ z_i \in \mathbf{R} \text{ and } -1 \leq z_i \leq 0 \right\} \subset R^n \tag{1.4}$$

is called the *fundamental (or closed) parallelepiped* of M. The set

$$P = \left\{ \sum_{i=1}^{n} z_i m_i : \ z_i \in \mathbf{R} \text{ and } -1 < z_i \leq 0 \right\} \subset R^n \tag{1.5}$$

is the half-open parallelepiped of M, and the set

$$P^o = \left\{ \sum_{i=1}^{n} z_i m_i : \ z_i \in \mathbf{R} \text{ and } -1 < z_i < 0 \right\} \subset R^n \tag{1.6}$$

is the open parallelepiped of M.

Next we show that the number of points of a lattice in any bounded set is finite. Let $||(x_1, \cdots, x_n)|| = (\sum_i x_i^2)^{1/2}$ be the Euclidean norm on $\mathbf{R}^n$. Recall that Schwartz's inequality says that for any real vectors $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$ of length n we have $\langle x, y \rangle \leq ||x|| \cdot ||y||$ where $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ is the ordinary inner product.

Theorem 1.2.22 *If $L \subseteq \mathbf{R}^n$ is an integer lattice of rank at most n, then $L \cap \{x : ||x|| < c\}$ is finite for every $c \in \mathbf{R}$.*

Proof: We can extend any non-full lattice to a full one, and this cannot decrease the number of points of the lattice with norm at most c, so we may assume L is full, with basis $m_1, \cdots, m_n$. The conditions $\langle x, m_2 \rangle = 0, \cdots, \langle x, m_n \rangle = 0$ amount to a system of $n-1$ independent linear equations in n variables (the coordinates of x), so there is at least one nonzero solution x to these equations. It cannot then hold that $\langle x, m_1 \rangle = 0$. Let $z_1 = (1/\langle x, m_1 \rangle)x$, so that $\langle z_1, m_1 \rangle = 1$ and $\langle z_1, m_i \rangle = 0$ for $i \neq 1$. Similarly we can find vectors $z_2, \cdots, z_n$ so that

$$\langle z_j, m_i \rangle = \begin{cases} 0 & \text{if } j = i \\ 1 & \text{if } j \neq i. \end{cases}$$

Now suppose that $w = a_1 m_1 + \cdots + a_n m_n \in L \cap \{x : ||x|| < c\}$. We have $a_i = \langle w, z_i \rangle$, so by Schwartz's inequality

$$|a_i| = |\langle w, z_i \rangle| \leq ||w|| \cdot ||z_i|| \leq c||z_i||.$$

But a_i is an integer, so there are only finitely many possible values for each a_i, and thus there are only finitely many such w. □

Sometimes a module M over a ring R has the structure of a commutative ring. If the function $a \mapsto a \cdot 1_M$ is a ring homomorphism, then we say that M is a (commutative) S-algebra. For example, every commutative ring is a $\mathbf{Z}$-algebra. If R is a subring of a ring R', then R' is an R-algebra. If R is commutative ring and S is a multiplicative set in R, then $S^{-1}R$ is an R-algebra. More generally, if I is an ideal of R and R/I is a subring of a ring R', then R' is an R-algebra.

1.3 Characters and Fourier transforms

The Fourier transform can be defined in tremendous generality. In this section we describe the main properties of the Fourier transform for finite Abelian groups.

1.3.a Basic properties of characters

Definition 1.3.1 *A (complex)* character *of an Abelian group G is a group homomorphism from G to the multiplicative group $\mathbf{C}^\times = \mathbf{C} - \{0\}$ of the complex numbers. That is, it is a function $\chi : G \to \mathbf{C}$ such that $\chi(a+b) = \chi(a)\chi(b)$ for all $a, b \in G$. Such a character is* nontrivial *if $\chi(a) \neq 1$ for some a. The trivial character is denoted* 1*, and the collection of all characters of G is denoted $\widehat{G}$.*

The group operation in an Abelian group is usually denoted "+", and this can lead to some confusion since a character takes values in a multiplicative group. In particular, if χ is a character of G then $\chi(mg) = \chi(g)^m$ (for any integer m), and $\chi(0) = 1$. For example, if $G = \mathbf{Z}/(2)$ then there is a unique nontrivial character χ and it is given by $\chi(0) = 1$ and $\chi(1) = -1$. That is, it converts $\{0, 1\}$ sequences into $\{\pm 1\}$ sequences. If G is a finite Abelian group then $|\chi(g)| = 1$ for all $g \in G$ (since $\chi(g)^{|G|} = 1$). It follows that $\chi(-g) = \overline{\chi}(g)$ (complex conjugate) for all $g \in G$.

If $G = \mathbf{Z}/(N)$ is the additive group of integers modulo N then the group $\widehat{G}$ of characters is also cyclic and is generated by the primitive character $\chi(i) = e^{2\pi i/N}$. If $G = G_1 \times G_2$ is a product of two groups then $\widehat{G} = \widehat{G}_1 \times \widehat{G}_2$. In other words, if χ is a character of G then there are unique characters χ_1, χ_2 of G_1, G_2 (respectively) such that $\chi(g_1, g_2) = \chi_1(g_1)\chi_2(g_2)$, namely $\chi_1(g_1) = \chi(g_1, 1)$ and $\chi_2(g_2) = \chi(1, g_2)$ (for any $g_1 \in G_1$ and $g_2 \in G_2$). From this, together with the fundamental theorem for finite Abelian groups 1.1.15, it follows that the collection $\widehat{G}$ of characters of a finite Abelian group G is itself a finite Abelian group which is isomorphic to G. (The corresponding statement for infinite Abelian groups is false: any nonzero $x \in \mathbf{C}$ defines a character of the integers $\mathbf{Z}$ by setting $\chi(m) = x^m$.)

Proposition 1.3.2 *Let G be a finite Abelian group, let $\chi : G \to \mathbf{C}^\times$ be a character, and let $g \in G$. Then*

$$\sum_{h \in G} \chi(h) = \begin{cases} 0 & \textit{if} \;\; \chi \neq 1 \\ |G| & \textit{if} \;\; \chi = 1 \end{cases} \tag{1.7}$$

and

$$\sum_{\psi \in \widehat{G}} \psi(g) = \begin{cases} 0 & \textit{if} \;\; g \neq 0 \\ |G| & \textit{if} \;\; g = 0. \end{cases} \tag{1.8}$$

Proof: If χ is nontrivial, there exists $ah \in G$ with $\chi(a) \neq 1$. Then

$$\chi(a) \sum_{h \in G} \chi(h) = \sum_{h \in G} \chi(ah) = \sum_{h' \in G} \chi(h')$$

so $(1 - \chi(a)) \sum_{h \in G} \chi(g) = 0$. For the second statement, note that g determines a character ψ_g of $\widehat{G}$ by the equation $\psi_g(\chi) = \chi(g)$. This character is nontrivial precisely when $g \neq 0$. In this case, the sum is $\sum_{\chi \in \widehat{G}} \psi_g(\chi)$, which is zero by the first part of the lemma. □

Corollary 1.3.3 *If G is a finite Abelian group and if $g, h \in G$ with $g \neq h$, then there exists a character χ such that $\chi(g) \neq \chi(h)$.*

Proof: If $\chi(g - h) = 1$ for every $\chi \in \widehat{G}$, then summing over all characters gives $|G|$. By equation (1.8) we conclude that $g - h = 0$. □

Corollary 1.3.4 *(Orthogonality relations) If G is a finite Abelian group and if $\psi, \chi \in \widehat{G}$ are distinct characters then*

$$\sum_{g \in G} \psi(g) \overline{\chi}(g) = 0. \tag{1.9}$$

If $g, h \in G$ are distinct elements then

$$\sum_{\chi \in \widehat{G}} \chi(g) \overline{\chi}(h) = 0. \tag{1.10}$$

Proof: The first equation follows by applying Proposition 1.3.2 to the character $\psi\chi^{-1}$. The second equation is $\sum_\chi \chi(g - h) = 0$, also by Proposition 1.3.2. □

1.3.b Fourier Transform

Let G be a finite Abelian group and $f : G \to \mathbf{C}$ be a function. We define its *Fourier transform* $\widehat{f} : \widehat{G} \to \mathbf{C}$ by

$$\widehat{f}(\chi) = \sum_{g \in G} \chi(g) f(g).$$

The *Fourier inversion formula*

$$f(g) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi) \overline{\chi}(g) \tag{1.11}$$

expresses an arbitrary function f as a linear combination of characters. Equation (1.11) follows immediately from the orthogonality relation for characters, for the sum on the right hand side is

$$\frac{1}{G} \sum_{\chi \in \widehat{G}} \sum_{h \in G} f(h) \chi(h) \overline{\chi}(g) = \frac{1}{|G|} \sum_{h \in G} f(h) \sum_{\chi \in \widehat{G}} \chi(h-g) = f(g)$$

by equation (1.8). Equation (1.11) implies that the characters span the group $\mathbf{C}[G]$ of complex-valued functions on G.

Proposition 1.3.5 *(Parseval's formula) Let $f : G \to \mathbf{C}$. Then*

$$|G| \sum_{g \in G} |f(g)|^2 = \sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)|^2. \tag{1.12}$$

Proof: Multiply $\widehat{f}(\chi) = \sum_g \chi(g) f(g)$ by its conjugate, $\sum_h \overline{\chi}(h) \overline{f}(h)$ to get

$$\sum_{\chi} |\widehat{f}(\chi)|^2 = \sum_{\chi} \sum_{g} \sum_{h} f(g) \overline{f}(h) \chi(g) \overline{\chi}(h) = \sum_{g,h} f(g) \overline{f}(h) \sum_{\chi} \chi(g) \overline{\chi}(h).$$

The inner sum vanishes unless $g = h$, which leaves $|G| \sum_g f(g) \overline{f}(g)$ as claimed. □

If $G \cong \mathbf{Z}/(N)$ is a cyclic group then a choice $\zeta \in \mathbf{C}$ of primitive N-th root of unity determines an isomorphism $G \cong \widehat{G}$ which takes 1 to the character χ_1 with $\chi_1(k) = \zeta^k$. The other nontrivial characters χ_m are powers of this: $\chi_m(k) = \zeta^{mk}$. Thus, if $f : G \to \mathbf{C}$ is a function, its Fourier transform $\widehat{f}$ may be considered as a function $\widehat{f} : G \to \mathbf{C}$ by writing $\widehat{f}(m)$ rather than $\widehat{f}(\chi_m)$. Thus

$$\widehat{f}(m) = \sum_{k=0}^{N-1} \zeta^{mk} f(k). \tag{1.13}$$

1.4 Polynomials

In this section we describe some of the basic properties of the ring of polynomials. The polynomial ring is among the most fundamental algebraic constructions. It is needed for much of the analysis of shift register sequences.

1.4.a Polynomials over a ring

Throughout this section R denotes a commutative ring. A *polynomial over* R is an expression

$$f = f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_dx^d = \sum_{i=0}^{d} a_ix^i$$

where $a_0, a_1, \cdots, a_d \in R$ and x is an indeterminate. The a_i are called the *coefficients of* R. When writing polynomials we may omit terms whose coefficients equal zero. We may also write the terms in a different order. If $a_d \neq 0$, then we say that $f(x)$ has degree $d = \deg(f(x))$. In this case a_d is called the *leading coefficient* of $f(x)$. We say $\deg(0) = -\infty$. If $\deg(f(x)) = 0$ then $f(x)$ is a *constant polynomial.* If $a_d = 1$ then $f(x)$ is *monic.* The term a_0 is called the *constant term.* The value of $f(x)$ at an element $b \in R$ is $f(a) = \sum_{i=0}^{d} a_ib^i$. An element $a \in R$ is a *root of* $f(x)$ if $f(a) = 0$. If $g(x) = \sum_{i=0}^{e} b_ix^i$ is a second polynomial over R, then we define

$$(f+g)(x) = f(x) + g(x) = \sum_{i=0}^{\max(d,e)} (a_i + b_i)x^i$$

(where we may have to extend one of the polynomials with zero coefficients so that this makes sense) and

$$(fg)(x) = f(x)g(x) = \sum_{i=0}^{d+e} \left(\sum_{j=\max(0,i-e)}^{\min(d,i)} a_jb_{i-j} \right) x^i.$$

The set of polynomials over R is denoted $R[x]$. The operations of addition and multiplication make $R[x]$ into a ring whose zero is the polynomial with every $a_i = 0$, and whose identity is the polynomial with $a_0 = 1$ and $a_i = 0$ for $i \geq 1$. The proof of the following lemma is straightforward.

Lemma 1.4.1 *If* $f(x), g(x) \in R[x]$, *then* $\deg(f+g) \leq \max(\deg(f), \deg(g))$ *with equality if* $\deg(f) \neq \deg(g)$. *Also,* $\deg(fg) \leq \deg(f) + \deg(g)$, *and equality can fail only when the product of the leading coefficients of* f *and* g *equals zero. In particular, if* R *is an integral domain then so is* $R[x]$.

If R is an integral domain, then the units in $R[x]$ are exactly the polynomials with degree zero. This is false in general. For example, if $R = (\mathbf{Z}/(4)$, then $(1+2x)^2 = 1$, so $1+2x$ is a unit with degree one.

The following result says that sometimes we can perform division with remainder in $R[x]$.

Theorem 1.4.2 *(Division Theorem for f/g) Let $f(x), g(x) \in R[x]$. Suppose the leading coefficient of g is invertible. Then there exist unique polynomials $q, r \in R[x]$ such that $\deg(r) < \deg(g)$ and*

$$f(x) = q(x)g(x) + r(x).$$

Proof: By induction on the degree d of f. If $\deg(f) < \deg(g)$, take $q = 0$ and $r = f$. Otherwise, suppose f has leading coefficient a_d. Suppose g has degree $e \le d$ and leading coefficient b_e. Then we have $f(x) = a_d b_e^{-1} x^{d-e} g(x) + f'(x)$ for some polynomial f'. The degree of f' is less than the degree of f, so by induction we have $f' = q'g + r$. It follows that $f = (a_d b_e^{-1} + q')x^{d-e} g + r$. For uniqueness, suppose $f = q_1 g + r_1 = q_2 g + r_2$ with $\deg(r_i) < \deg(g)$. Then $0 = (q_1 - q_2)g + (r_1 - r_2)$. The leading coefficient of g is invertible, and $\deg(r_1 - r_2) < \deg(g)$. It follows that the leading coefficient of $q_1 - q_2$ is zero, that is, $q_1 - q_2 = 0$. Therefore $r_1 - r_2 = 0$. □

Theorem 1.4.3 *If a is a root of $f(x) \in R[x]$, then there exists a polynomial $q(x) \in R[x]$ such that*

$$f(x) = (x-a)q(x).$$

If R is an integral domain, then the number of distinct roots of f is no more than the degree of f (but see exercise 16).

Proof: Use the division theorem (Theorem 1.4.2) with $g = x - a$. The remainder r has degree zero but has a as a root. Thus r is zero. If R is an integral domain and if $b \ne a$ is another root of $f(x)$ then b is necessarily a root of $q(x)$. So the second statement follows by induction. □

A root a of polynomial f is said to be *simple* if a is not a root of $f(x)/(x-a)$.

Lemma 1.4.4 *Let $q = \sum_{i=0}^{m} q_i x^i \in R[x]$ be a polynomial with coefficients in R. Consider the following statements*

1. *q_0 is invertible in R.*
2. *The polynomial x is invertible in the quotient ring $R[x]/(q)$.*
3. *The polynomials $q(x)$ and x are relatively prime in the ring $R[x]$.*
4. *There exists an integer $T > 0$ such that $q(x)$ is a factor of $x^T - 1$.*

5. There exists an integer $T > 0$ such that $x^T = 1$ in the ring $R[x]/(q)$.

Then statements (1), (2), and (3) are equivalent and

$$x^{-1} = -q_0^{-1}(q_1 + q_2 x + \cdots + q_m x^{m-1})$$

in $R[x]/(q)$. Statements (4) and (5) are equivalent (and the same T works for both) and $x^{-1} = x^{T-1}$ in $R[x]/(q)$. Statement (4) (or (5)) implies (1), (2), (3). If R is finite then (1) (or (2) or (3)) implies (4),(5).

Proof: The statements are all straightforward except (possibly) the last one. Suppose that R is finite. Then the quotient ring $R[x]/(q)$ also contains finitely many elements so the powers $\{x^n\}$ of x in this ring cannot all be different. Hence there exists T such that $x^{n+T} \equiv x^n \pmod{q}$ for all sufficiently large n. Under assumption (2) this implies that $x^T \equiv 1 \pmod{q}$. In other words, q divides the polynomial $x^T - 1$, as claimed. □

When condition (4) (or (5)) in Lemma 1.4.4 holds, the smallest T such that $q(x)|(x^T - 1)$ is called the *order* of the polynomial q. (Otherwise one may say that q does not have an order, or that its order is infinite. The terminology is confusing: it should be called the order of $x \pmod{q}$ for consistency with the terminology of group theory.)

1.4.b Polynomials over a field

Theorem 1.4.5 *If F is a field, then $F[x]$ is Euclidean with $\delta(f) = \deg(f)$. Every ideal in $F[x]$ has a unique monic principal generator. Any $f(x) \in F[x]$ can be written in the form*

$$f(x) = a p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$$

where $a \in F$, the p_i are distinct monic irreducible elements of $F[x]$, and the e_i are positive integers. This representation is unique apart from changing the order of the p_i.

Proof: It follows from Theorem 1.4.2 that $F[x]$ is Euclidean. It is also principal and is a UFD by Theorem 1.2.11. Each irreducible polynomial has a unique monic associate (divide by the leading coefficient). This accounts uniquely for a. □

It also follows from Theorem 1.2.11 that $F[x]$ is a GCD ring, but to be precise we have:

Theorem 1.4.6 *Let F be a field and $f_1, \cdots, f_k \in F[x]$, not all zero. There is a unique monic $g \in F[x]$ such that (1) g divides every f_i and (2) if h divides every f_i then h also divides g. Moreover, g can be written in the form*

$$g = h_1 f_1 + h_2 f_2 + \cdots + h_k f_k \tag{1.14}$$

for some $h_1, h_2, \cdots, h_k \in F[x]$.

Proof: Let $I = \{h_1f_1 + h_2f_2 + \cdots + h_kf_k : h_1, h_2, \cdots, h_k \in F[x]\}$. Then I is an ideal in $F[x]$, so by Theorem 1.4.5, I has a unique monic generator g. Since $g \in I$, g can be written in the form in equation (1.14). It follows that any h that divides every f_i also divides g. Since $f_i \in I$, g divides f_i. □

We write $g = \gcd(f_1, \cdots, f_k)$. It can be found by the usual Euclidean algorithm by repeatedly using Theorem 1.4.2. There is also a notion of least common multiple in $F[x]$. The following theorem later allows us to construct finite fields of all possible sizes. The proof is omitted.

Theorem 1.4.7 *If F is a finite field and d is a positive integer, then there is at least one irreducible polynomial of degree d in $F[x]$.*

If $F \subseteq E$ are fields and if $a \in E$ is an element that is the root of some polynomial with coefficients in F, then we say a is *algebraic over F*. A polynomial $f \in F[x]$ is called a *minimal polynomial* of a (over F) if it is monic, if $f(a) = 0$ and if it is a polynomial of smallest degree with these properties.

Theorem 1.4.8 *Suppose a is algebraic over F. Then it has a unique minimal polynomial $f \in F[x]$. The minimal polynomial f is also the unique monic irreducible polynomial in $F[x]$ having a as a root. If $g \in F[x]$ is any other polynomial such that $g(a) = 0$ then f divides g in $F[x]$.*

Proof: If two monic polynomials $f, g \in F[x]$ have the same (minimal) degree and both have a as a root then $f - g$ has smaller degree, which is a contradiction. Now suppose f is a monic irreducible polynomial such that $f(a) = 0$. The set $J = \{h \in F[x] : \ h(a) = 0\}$ is an ideal, so it is principal. It contains f, but f is irreducible, so $J = (f)$ is the ideal generated by f, and f is the unique monic polynomial with this property. If $g(a) = 0$ then $g \in J$ so g is a multiple of f. □

1.5 Exercises

1. Prove that if G_1 and G_2 are groups, then the direct product $G_1 \times G_2$ is a group. Prove that $G_1 \times G_2$ is Abelian if G_1 and G_2 are Abelian.

2. Describe the set of all subgroups of the group $\mathbf{Z}/m\mathbf{Z}$.

3. Let $\varphi : G \to H$ be a group homomorphism. Prove that $\ker(\varphi)$ is a subgroup of G and $\mathrm{Im}(\varphi)$ is a subgroup of H.

4. Let G be a group and let H be a subgroup of G. Prove that the relation defined by $a \sim b$ if there is an $h \in H$ such that $b = ah$ is an equivalence relation. Find an example where the definition $aHbH = abH$ does not make the set of equivalence classes into a group.

5. Prove that a subgroup H of a group G is normal if and only if for every $a \in G$ and $h \in H$, we have $aha^{-1} \in H$.

6. Theorem 1.1.14: Let $\varphi : G \to G'$ be a homomorphism.

1. Prove that $\ker(\varphi)$ is normal in G.
2. Prove that the quotient $G/\ker(\varphi)$ is isomorphic to $\mathrm{Im}(\varphi)$.
3. Conversely, prove that if H is a normal subgroup of G, then the map $a \mapsto aH$ is a surjection from G to G/H with kernel equal to H.

7. Show that the set of endomorphisms of an Abelian group is a ring.

8. Theorem 1.2.7:

1. Suppose $\varphi : R \to S$ is a ring homomorphism. Prove that $\ker(\varphi)$ is an ideal of R and φ induces an isomorphism between $R/\ker(\varphi)$ and the image of f.
2. Prove that if I is an ideal of R, then the map $a \mapsto a + I$ is a homomorphism from R onto R/I with kernel I.

9. Prove that a GCD ring with no infinite chain of proper ascending ideals is also a LCM ring.

10. Let $\{R_s : s \in S\}$ be a family of rings. Prove that R_S is the unique (up to isomorphism) ring such that if T is any ring and $\psi_s : T \to R_s$ any set of homomorphisms, then there is a homomorphism $g : T \to R_S$ such that $\psi_s = \varphi_s \circ g$ for every $s \in S$.

11. Prove that if V is a vector space over a field F, then for every $u \in V$ we have $0 \cdot u = 0$.

12. Theorem 1.2.17:

1. Prove that every vector space has a basis. (Hint: use Zorn's Lemma.)
2. Prove that if S is a basis for a vector space V, then every vector can be written uniquely as a linear combination of elements of S.

13. Develop a theory of characters as functions with values in an arbitrary field F rather than $\mathbf{C}$. For certain parts you will need to assume that F contains the n-th roots of unity.

14. Prove that the Hadamard transform is $\sum_x \zeta^{a \cdot x} f(x)$.

15. Prove that the Walsh transform is something else.

16. Let $R = \mathbf{Z} \times \mathbf{Z}$. Let $f(x) = (1,0)x - (1,0) \in R[x]$. Show that f has infinitely many roots in the ring R.

Chapter 2 Special Fields

2.1 Finite Fields

In this section we analyze the structure of finite fields. For a more complete treatment see the excellent reference by Lidl and Niedereitter [15]. Our first task is identify all finite fields and all inclusion relations among them.

2.1.a Basic properties

Theorem 2.1.1 *Let p be a prime number. For each $d > 0$ there is (up to isomorphism) a unique field $\mathbf{F}_{p^d}$ with p^d elements. These account for all finite fields. If $e > 0$ is another integer, then there is an inclusion $\mathbf{F}_{p^d} \subseteq \mathbf{F}_{p^e}$ if and only if d divides e. That is, the lattice of finite fields with characteristic p under inclusion is isomorphic to the lattice of whole numbers under divisibility. The subfield $\mathbf{F}_{p^d}$ consists of those elements a of $\mathbf{F}_{p^e}$ satisfying $a^{p^d} = a$.*

The field $\mathbf{F}_{p^d}$ is sometimes denoted $GF(p^d)$ (for "Galois field"). The proof of Theorem 2.1.1 will occupy the rest of §2.1.a.

Suppose d is a positive integer and F is a finite field with q elements. Let $f(x)$ be an irreducible polynomial over F with degree d. Then by Theorem 1.2.5.4, $F[x]/(f(x))$ is a field. It has q^d elements. In particular, if p is a prime integer and we take $F = \mathbf{Z}/(p)$, then this together with Theorem 1.4.7 shows that there exists a finite field of order p^d for every prime p and positive integer d.

Next suppose F is a finite field with characteristic $p > 0$. Recall that we showed in Theorem 1.2.8 that p is prime. It follows that the mapping $\mathbf{Z}/(p) \to F$ which takes an element n to $1 + 1 + \cdots + 1$ (n times) is a ring homomorphism. So we can view $\mathbf{Z}/(p)$ as a subfield of F. Hence F has the structure of a finite dimensional vector space over $\mathbf{Z}/(p)$. By Theorem 1.2.18, F has p^d elements for some d.

Proposition 2.1.2 *If $F \subseteq E$ are two finite fields, then E and F have the same characteristic. If p is the characteristic, then $|F| = p^d$ and $|E| = p^e$ for some integers d and e such that d divides e.*

Proof: If F has characteristic p and E has characteristic r, then $|F| = p^d$ and $|E| = r^e$ for some d and e. But E is a vector space over F, so $r^e = (p^d)^k$ for some k. Thus $r = p$ and $e = dk$. □

To complete the picture of the set of finite fields we want to show that there is, up to isomorphism, a unique finite field of a given cardinality. First we need a lemma.

Lemma 2.1.3 *If F is a finite field, then every $a \in F$ is a root of the polynomial $x^{|F|} - x$ and we have*

$$x^{|F|} - x = \prod_{a \in F} (x - a).$$

No other element of any extension field of F is a root of this polynomial.

Proof: The multiplicative group of F has order $|F| - 1$, so by Theorem 1.1.3 any nonzero element $a \in F$ satisfies $a^{|F|-1} = 1$. Therefore any element $a \in F$ satisfies $a^{|F|} = a$. That is, every a is a root of the polynomial $x^{|F|} - x$. It follows that $x - a$ divides $x^{|F|} - x$. Furthermore, the degree of $x^{|F|} - x$ equals $|F|$, so there are no other roots of this polynomial in E. The factorization follows from Theorem 1.4.3. □

Corollary 2.1.4 *Suppose E is a field, p is a prime number, and d is a positive integer. Then E contains at most one subfield of order p^d.*

Proof: Suppose F is a subfield of E of order p^d. By Lemma 2.1.3 every $a \in F$ is a root of $x^{p^d} - x$, and there are no other roots of this polynomial in E.

Now suppose F' is another subfield of E of order p^d. The same reasoning applies to F'. Thus $F = F'$. □

Proposition 2.1.5 *Let p be a prime number and let $d > 0$ be an integer. Any two finite fields with p^d elements are isomorphic.*

Proof: Let $E = (\mathbf{Z}/(p))[x]/(f(x))$, where $f(x)$ is an irreducible polynomial with degree d and coefficients in $\mathbf{Z}/(p)$. It is enough to show that any field F with p^d elements is isomorphic to E.

By Lemma 2.1.3, every $a \in E$ satisfies $a^{p^d} = a$. In particular, $x^{p^d} - x = 0$ in E, so $f(x)$ divides $x^{p^d} - x$ as polynomials. That is, $x^{p^d} - x = f(x)g(x)$ for some $g(x) \in (\mathbf{Z}/(p))[x]$.

On the other hand, we can think of $x^{p^d} - x$ as a polynomial over F. By the same reasoning, every element of F is a root of this polynomial, so

$$f(x)g(x) = x^{p^d} - x = \prod_{a \in F} (x - a).$$

In particular, $f(x)$ factors into linear factors over F. Let a be a root of $f(x)$ in F. If the elements $\{1, a, a^2, \cdots, a^{d-1}\}$ were linearly dependent over $(\mathbf{Z}/(p))[x]$, a would be a root of a lower degree polynomial, and this polynomial would divide $f(x)$. That would contradict the irreducibility of $f(x)$. Thus they are linearly independent and hence a basis (F has dimension d over $(\mathbf{Z}/(p))[x]$). That is, every b in F can be written

$$b = \sum_{i=0}^{d-1} c_i a^i,$$

with $c_i \in (\mathbf{Z}/(p))[x]$. We define a function

$$L(\sum_{i=0}^{d-1} c_i a^i) = \sum_{i=0}^{d-1} c_i x^i$$

from F to E. This function is one-to-one and it can be checked that it preserves multiplication and addition. Hence it is an isomorphism. □

Thus for each prime power $q = p^d$ there is a unique field $\mathbf{F}_q$ with q elements.

Proposition 2.1.6 *Let p be prime, let d, e be positive integers, and suppose that d divides e. Then the field $\mathbf{F}_{p^d}$ may be realized as a subfield of $\mathbf{F}_{p^e}$.*

Proof: Let $F = \mathbf{F}_{p^d}$ and set $q = p^d = |F|$. Let $E = \mathbf{F}_{p^e}$. Assume $e = dk$ for some integer $k > 0$. Then $|E| = q^k$. Recall from Lemma 2.1.3 that E consists of the distinct roots of the polynomial $x^{p^e} - x = x^{q^k} - x$. This polynomial is divisible by the polynomimal $x^q - x$, for the quotient is

$$x^{(q^k-1)-(q-1)} + x^{(q^k-1)-2(q-1)} + \cdots + x^{q-1} + 1.$$

Thus E contains a set S of q distinct roots of the polynomial $(x^q - x)$. By Lemma 1.2.9, both addition and multiplication commute with raising to the qth power, so the subset $S \subset E$ is a field. Therefore it is isomorphic to the field $F = \mathbf{F}_q$. □

Suppose $q \in F[x]$ is irreducible. Recall that in the terminology of §1.4.a, the order of q is the smallest T such that $q(x)|(x^T - 1)$. This is the order of x in the group of units of $F[x]/(q)$, a group that has $|F|^{\deg(q)} - 1$ elements. Thus by Theorem 1.1.3 the order of q divides $|F|^{\deg(q)} - 1$.

This completes our picture of the set of finite fields and the proof of Theorem 2.1.1.

2.1.b Galois groups

Some of the preceding notions can be understood in terms of Galois groups (see §1.2.a for the definition of Galois groups). From the proof of Proposition 2.1.6 we see that $\sigma(a) = a^p$ defines an automorphism (that is, an invertible homomorphism that is both additive and multiplicative) of any finite field $\mathbf{F}_{p^d}$. If a is in $\mathbf{F}_{p^d}$, then we have $\sigma^d(a) = a^{p^d} = a$, so σ^d is the identity on $\mathbf{F}_{p^d}$. If σ^e is the identity on $\mathbf{F}_{p^d}$ for any e, then $a^{p^e} = \sigma^e(a) = a$ for every a, so $\mathbf{F}_{p^d} \subseteq \mathbf{F}_{p^e}$. In particular, $d < e$ so the powers of σ constitute a cyclic group of order d.

Furthermore, if c is a divisor of d, then $\sigma^c(a) = a$ if and only if $a \in \mathbf{F}_{p^c}$. That is,

$$\mathbf{F}_{p^c} = \left\{a \in \mathbf{F}_{p^d} : \ \sigma^c(a) = a\right\}$$

is the subfield of $\mathbf{F}_{p^d}$ fixed by the group generated by σ^c. Thus $\mathbf{F}_{p^c} \subset \mathbf{F}_{p^d}$ is a *Galois extension.*

Theorem 2.1.7 *The Galois group* $\mathrm{Gal}(\mathbf{F}_{p^d}/\mathbf{F}_{p^c})$ *is a cyclic group of order d/c, generated by the automorphism $\sigma^c : a \mapsto a^{p^c}$.*

Proof: Suppose that τ is any automorphism of $\mathbf{F}_{p^d}$. It suffices to show that $\tau = \sigma^i$ for some i. The theorem then follows from the fact that $\mathbf{F}_{p^c}$ is the fixed field of the subgroup generated by σ^c. Let f be an irreducible polynomial over $\mathbf{F}_p$ with degree d, and let a be a root of f. Then $\mathbf{F}_{p^d} = \mathbf{F}_p[a]$ and $1, a, a^2, \cdots, a^{d-1}$ is a basis for $\mathbf{F}_{p^d}$ over $\mathbf{F}_p$. Thus to show that two automorphisms are equal, it suffices to show that they are equal on a. We have that $\sigma^i(f) = f$ for every i, so $\sigma^i(a)$ is a root of f. Similarly, $\tau(a)$ is a root of f. The $\sigma^i(a)$ are distinct – otherwise a and hence $\mathbf{F}_{p^d}$ are in a proper subfield, which is a contradiction. Thus there are $d = \deg(f)$ of them, and they account for all the roots of f. In particular, $\tau(a) = \sigma^i(a)$ for some i. So $\tau = \sigma^i$, proving the theorem. □

Thus we have an inclusion reversing correspondence between the lattice of subfields of $\mathbf{F}_{p^d}$ and the lattice of subgroups of $\mathrm{Gal}(\mathbf{F}_{p^d}/\mathbf{F}_p)$. The main theorem of Galois theory describes the solutions of a polynomial equation in terms of the Galois group.

Theorem 2.1.8 *Let F be a finite field and $f(x) \in F[x]$ be a polynomial of degree d with coefficients in F. Let E be an extension field of F and suppose $\alpha \in E$ is a root of f. Then for any $\sigma \in Gal(E/F)$, the element $\sigma(\alpha) \in E$ is also a root of f. If f is irreducible in $F[x]$ and if E is the degree d extension of F then all the roots of f are contained in E. They consist exactly of the Galois conjugates,*

$$\sigma_i(\alpha) = \alpha^{q^i},$$

where $0 \le i \le d-1$. That is, where σ_i ranges over all elements of $Gal(E/F)$.

Proof: Let $q = |F|$. The Galois group $Gal(E/F)$ is cyclic and it is generated by the mapping $\sigma : E \to E$ given by $\sigma(a) = a^q$. If $f(x) = \sum_{i=0}^{d} a_i x^i$ and if $\alpha \in E$ is a root of f, then

$$0 = \sigma(f(\alpha)) = \left(\sum_{i=0}^{d} a_i \alpha^i\right)^q = \sum_{i=0}^{d} a_i^q \alpha^{iq} = \sum_{i=0}^{d} a_i \sigma(\alpha) = f(\sigma(\alpha))$$

(by Lemma 1.2.9), so $\sigma(\alpha)$ is also a root of f.

Now suppose f is irreducible and, without loss of generality, monic). Then it is the minimal polynomial of α by Theorem 1.4.8. But the polynomial $g(x) = \prod_{g \in Gal(E/F)}(x - g(\alpha)) \in E[x]$ has the same degree as f, and it is clearly fixed under each element of $Gal(E/F)$. So $g \in F[x]$, and it has α as a root. Therefore $g = f$, so the roots of f are all the Galois conjugates of α. □

2.1.c Primitive elements

To work within a particular finite field F, it is useful to have some structural information. An element $a \in F$ is called *primitive* if every nonzero element of F can be written as a power of a. A polynomial $f \in \mathbf{F}_p[x]$ of degree d is primitive if it is irreducible and if one (and hence all) of its roots in $\mathbf{F}_{p^d}$ are primitive elements.

The following lemma will be used in §**??**.)

Lemma 2.1.9 *Let $F = \mathbf{F}_q$ be the field with q elements. Let $f \in F[x]$ be a polynomial. Then f is primitive if and only if its order is $q^{\deg(f)} - 1$.*

Proof: In the ring $F[x]/(f)$ the element x is a root of the polynomial $f(x)$. If x is primitive then the order of x is $T = |F| - 1 = q^{\deg(f)} - 1$. Thus T is the smallest integer such that $x^T = 1 \pmod{f}$, which is to say that T is the smallest integer such that f divides $x^T - 1$. Thus the order of f is T. The converse is similar. □

We next show that every finite field has primitive elements. This implies that the multiplicative group of a finite field is cyclic.

Proposition 2.1.10 *The finite field $\mathbf{F}_{p^d}$ has $\phi(p^d - 1)$ primitive elements.*

Proof: Suppose that $a \in \mathbf{F}_{p^d}$ has order e. That is, $a^e = 1$ and no smaller positive power of a equals 1. Then the elements $1, a, a^2, \cdots, a^{e-1}$ are distinct and are all roots of $x^e - 1$. That is,

$$x^e - 1 = (x-1)(x-a)(x-a^2)\cdots(x-a^{e-1}).$$

It follows that every element whose eth power equals 1 is a power of a, and an element $b = a^i$ has order e if and only if $\gcd(i, e) = 1$. Thus if there is at least one element of order e, then there are exactly $\phi(e)$. That is, for every e there are either 0 or $\phi(e)$ elements of order e.

Furthermore, by Lemma 2.1.3 every nonzero $a \in F$ is a root of the polynomial $x^{p^d - 1} - 1$. Thus if there is an element in F with order e, then e divides $p^d - 1$. It is a fact from number theory that for any positive integer k

$$\sum_{e|k} \phi(e) = k.\textbf{CITATION}?$$

Thus we have

$$\begin{aligned} p^d - 1 &= \sum_{e|p^d-1} |\{a \in F : \text{ the order of } a = e\}| \\ &\le \sum_{e|p^d-1} \phi(e) = p^d - 1. \end{aligned}$$

Therefore the two sums are equal. Since each term in the first sum is less than or equal to the corresponding term in the second sum, each pair of corresponding terms must be equal.

In particular, the number elements with order $p^d - 1$ equals $\phi(p^d - 1) > 0$. □

In fact, it can be shown that every finite field $\mathbf{F}_{p^d}$ has a *primitive normal basis* over a subfield $\mathbf{F}_{p^c}$. This is a basis of the form $a, a^{p^c}, \cdots, a^{p^{d-c}}$ with a primitive. The interested reader can find the details in [15, §2.3].

2.1.d The Trace Function

The trace function is an important function from a field to a subfield. It is used, for example, in the construction of binary sequences for a variety of engineering applications such as radar ranging, spread spectrum communication, Monte Carlo simulation, and stream ciphers. We define it here just for finite fields.

Definition 2.1.11 *Let d and e be positive integers with d dividing e. The* trace function from $\mathbf{F}_{p^e}$ to $\mathbf{F}_{p^d}$ *is defined by*

$$Tr_{p^d}^{p^e}(a) = a + a^{p^d} + a^{p^{2d}} + \cdots + a^{p^{e-d}}.$$

Lemma 2.1.12 *Let d and e be positive integers with d dividing e. If $a \in \mathbf{F}_{p^e}$ then*

$$a + a^{p^d} + a^{p^{2d}} + \cdots + a^{p^{e-d}} = \sum_{\sigma \in \mathrm{Gal}(\mathbf{F}_{p^e}/\mathbf{F}_{p^d})} \sigma(a)$$

is in $\mathbf{F}_{p^d}$.

Proof: Left as an exercise. □

Thus the trace function from $\mathbf{F}_{p^e}$ to $\mathbf{F}_{p^d}$ does indeed have values in $\mathbf{F}_{p^d}$. If there is no possibility of confusion, we simply write Tr for the trace function.

Theorem 2.1.13 *Let d and e be positive integers with d dividing e.*

1. *For all $a, b \in \mathbf{F}_{p^e}$ and $c \in \mathbf{F}_{p^d}$ we have $Tr(a+b) = Tr(a) + Tr(b)$ and $Tr(ca) = c\,Tr(a)$. That is, Tr is $\mathbf{F}_{p^d}$-linear.*
2. *For all $c \in \mathbf{F}_{p^d}$, we have $|\{a \in \mathbf{F}_{p^e} : Tr(a) = c\}| = p^{e-d}$.*
3. *For all $a \in \mathbf{F}_{p^e}$ we have $Tr(a^p) = Tr(a)^p$.*
4. *$Tr(1) \in \mathbf{F}_p$ and $Tr(1) \equiv e/d \pmod{p}$.*
5. *If $L : \mathbf{F}_{p^e} \to \mathbf{F}_{p^d}$ is an $\mathbf{F}_{p^d}$-linear function, then there is an element $a \in \mathbf{F}_{p^e}$ such that for every $b \in \mathbf{F}_{p^e}$, we have $L(b) = Tr(ab)$. We denote this function by $L_a(b)$.*

Proof:

1. Since $(a+b)^p = a^p + b^p$ in any field of characteristic p, and $a^{p^d} = a$ for any $a \in \mathbf{F}_{p^d}$, Tr is a sum of $\mathbf{F}_{p^d}$-linear functions.
2. For any $c \in \mathbf{F}_{p^d}$, the expression $\mathrm{Tr}(x) - c$ is a polynomial of degree p^{e-d}. Thus it has at most p^{e-d} roots. Thus the total number of roots of all these polynomials is at most p^e, with equality only if every such polynomial has exactly p^{e-d} roots. But every element of $\mathbf{F}_{p^e}$ is a root of exactly one such polynomial. Thus the total number of roots of all these polynomials is equal to p^e. It follows that there are exactly p^{e-d} elements a of $\mathbf{F}_{p^e}$ such that $\mathrm{Tr}(a) = c$.
3. All the operations used to define Tr commute with raising to the pth power.
4. We have $\mathrm{Tr}(1) = 1 + 1^{p^d} + \cdots + 1^{p^{e-d}} = 1 + 1 + \cdots + 1$, with e/d terms.
5. We prove this by counting. The field $\mathbf{F}_{p^e}$ has dimension e/d over $\mathbf{F}_{p^d}$, so by Theorem 1.2.20, there are p^e distinct $\mathbf{F}_{p^d}$-linear functions from $\mathbf{F}_{p^e}$ to $\mathbf{F}_{p^d}$. On the other hand, each function $L_a : \mathbf{F}_{p^e} \to \mathbf{F}_{p^d}$ is $\mathbf{F}_{p^d}$-linear. All that remains is to show that these are distinct as a varies. So, suppose that for some $a, b \in \mathbf{F}_{p^e}$ we have $\mathrm{Tr}(ac) = \mathrm{Tr}(bc)$ for every $c \in \mathbf{F}_{p^e}$. Then also $\mathrm{Tr}((a-b)c) = 0$ for every c. But if $a \neq b$, this implies that $Tr(x) = 0$ for all x, which is false. Thus $a = b$.

This completes the proof. □

There is an alternative definition of the trace function. If $a \in \mathbf{F}_{p^e}$, then the function $\kappa_a : x \mapsto ax$ is $\mathbf{F}_{p^d}$-linear, and thus can be represented by an $(e/d) \times (e/d)$ matrix M_a over R once a basis for $\mathbf{F}_{p^e}$ has been chosen. We denote by $T(a)$ the trace of this matrix.

Proposition 2.1.14 *For all $a \in \mathbf{F}_{p^e}$ we have $Tr^{p^e}_{p^d}(a) = T(a)$.*

Proof: Both maps T and Tr are R-linear, hence are equal if and only if they are equal on a basis. This holds in particular for a primitive normal basis $a, a^{p^d}, a^{p^{2d}}, \cdots, a^{p^{e-d}}$. In particular, we may assume that every element of the basis is a root of an irreducible polynomial of degree e/d over $\mathbf{F}_{p^d}$. Thus it suffices to prove that $T(a) = \mathrm{Tr}(a)$ for every element a that is a root of an irreducible polynomial of degree e/d over $\mathbf{F}_{p^d}$.

Let a be such an element, with minimal polynomial $f(x) = x^d + \sum_{i=0}^{e/d-1} a_i x^i$. Then M_a is the matrix with 1 in each entry of the subdiagonal, $a_0, \cdots, a_{d-1}$ in the last column, and 0s elsewhere. Its trace is a_{d-1}. On the other hand, we have

$$f(x) = \prod_{\sigma \in \mathrm{Gal}(\mathbf{F}_{p^e}/\mathbf{F}_{p^d})} (x - \sigma(a)),$$

so $a_{d-1} = \sum_{\sigma \in \mathrm{Gal}(\mathbf{F}_{p^e}/\mathbf{F}_{p^d})} \sigma(a) = \mathrm{Tr}(a)$. □

Similarly, the *norm* of an element $a \in \mathbf{F}_{p^e}$ is defined to be

$$N_{p^d}^{p^e}(a) = \prod_{\sigma \in \mathrm{Gal}(\mathbf{F}_{p^e}/\mathbf{F}_{p^d})} \sigma(a) = a^{(p^e-1)/(p^d-1)} \in \mathbf{F}_{p^d}.$$

It is a multiplicative function (that is, $N_{p^d}^{p^e}(ab) = N_{p^d}^{p^e}(a)N_{p^d}^{p^e}(b)$). The norm can also be defined in terms of the matrix M_a – it is simply the determinant. This can be seen by checking it for any primitive element a, which is straightforward since the matrix of such an element is as described in the proof of Proposition 2.1.14.

2.1.e Characters of finite fields

Let F be a finite field, say, $|F| = p^r$ where p is a prime number. Let $F^\times$ be the group of all nonzero elements of F under multiplication and let F^+ be the group of all elements of F under addition. A character of F^+ is called an *additive* character. If χ is a nontrivial additive character then every additive character is of the form $\psi(x) = \chi(Ax)$ for some element $A \in F$. (Different values of A give distinct characters, and there are $|F|$ of them, which therefore account for all additive characters.)

A character of $F^\times$ is called a *multiplicative* character of F. It is common to extend each multiplicative character $\psi : F^\times \to \mathbf{C}$ to all of F by setting $\psi(0) = 0$. There is a notion of Fourier transform with respect to either the additive or the multiplicative structure. Since the prime field $F_p = \mathbf{Z}/(p)$ is cyclic, and since the multiplicative group $F^\times$ is cyclic, equation (1.13) gives explicit formulae for these Fourier transforms. In this case they are sometimes called the Hadamard and Walsh transforms (respectively). (See exercises?)

If ψ is a multiplicative character one can take its Fourier transform $\widehat{\psi}$ with respect to the additive structure to obtain

$$\widehat{\psi}(\chi) = \sum_{g \in F} \chi(g)\psi(g) = \sum_{g \in F^\times} \chi(g)\psi(g) \tag{2.1}$$

for any additive character χ. Conversely, equation (2.1) may be interpreted as the Fourier transform $\widehat{\chi}$ of the additive character χ evaluated on the multiplicative character ψ. This sum is called a *Gauss sum* and is denoted $G(\psi, \chi)$. The results in §1.3.b therefore give a number of simple facts concerning Gauss sums. For example, the Fourier expansion of a multiplicative character ψ in terms of additive characters (1.11) gives

$$\psi(g) = \frac{1}{|F|}\sum_{\chi} G(\psi,\chi)\overline{\chi}(g) = \frac{1}{|F|}\sum_{\chi} G(\psi,\overline{\chi})\chi(g).$$

Other basic properties of Gauss sums are described in [15] §5.2.

2.1.f The Discrete Fourier Transform

We can generalize the notion of a Fourier transform by generalizing equation (1.13). Suppose that f is a function from $\mathbf{Z}/(N)$ into a finite field $F = \mathbf{F}_q$, $q = p^e$ with p prime, and suppose that N is relatively prime to p. Then there is a primitive Nth root of unity b in some extension field $\mathbf{F}_r$, $r = q^d$. This is true since q is a unit in $\mathbf{Z}/(N)$, hence has finite order d. That is, N divides $q^d - 1$ for some d. If c is a primitive element in $\mathbf{F}_{q^d}$, then c has order $q^d - 1$, so $b = c^{(q^d-1)/N}$ has order N.

Definition 2.1.15 *The* discrete Fourier transform of f *is defined to be the function*

$$\widehat{f}(m) = \sum_{k=0}^{N-1} b^{mk} f(k) \in \mathbf{F}_{q^d}.$$

This transform behaves similarly to the Fourier transform for characters defined in Section 1.3.b. The discrete Fourier transform $\widehat{f}$ is itself a function from $\mathbf{Z}/(N)$ to $\mathbf{Z}_{q^d}$. As such we can take its discrete Fourier transform.

Theorem 2.1.16 *Let b be a primitive Nth root of unity in $\mathbf{F}_{q^d}$. If f is any function from $\mathbf{Z}/(N)$ into $\mathbf{F}_{q^d}$, then*

$$f(g) = \frac{1}{N}\sum_{m=0}^{N-1} \widehat{f}(m) b^{-mg}. \tag{2.2}$$

This equation is known as the Fourier inversion formula

Proof: We have

$$\begin{aligned}
\frac{1}{N}\sum_{m=0}^{N-1} \widehat{f}(m) b^{-mg} &= \frac{1}{N}\sum_{m=0}^{N-1}\sum_{k=0}^{N-1} b^{mk} f(k) b^{-mg} \\
&= \frac{1}{N}\sum_{k=0}^{N-1} f(k) \sum_{m=0}^{N-1} b^{m(k-g)}.
\end{aligned}$$

If $k \neq g$, then $b^{k-g} \neq 1$ is a root of $(x^N - 1)/(x-1) = 1 + x + \cdots x^{N-1}$, so the inner sum is N if $k = g$ and is zero otherwise. The theorem follows from this. □

2.2 Algebraic Number Fields

2.2.a Basic properties

So far our examples of fields have consisted of finite fields and the familiar fields $\mathbf{Q}$, the rational numbers, $\mathbf{R}$, the real numbers, and $\mathbf{C}$, the complex numbers. Recall that we we obtain the various finite fields of characteristic $p > 0$ from the prime field $\mathbf{F}_p$ by constructing the quotient $\mathbf{F}_p[x]/(f(x))$ where $f(x)$ is an irreducible polynomial. We can think of this construction as adjoining a root (the variable x) of $f(x)$ to the field $\mathbf{F}_p$. Similarly, we obtain the complex numbers from the real numbers by adjoining a root of the polynomial $x^2 + 1$.

In this section we study a class of fields, called *algebraic number fields* that are obtained in the same way from the rational numbers. For the most part we omit proofs and leave the interested reader to find them in other references.

Definition 2.2.1 *An algebraic number field E is a finite extension of the rational numbers $\mathbf{Q}$.*

This means that E is a field that contains $\mathbf{Q}$ and that as a vector space over $\mathbf{Q}$ it is finite dimensional.

A complex number $a \in \mathbf{C}$ is said to be *algebraic over* $\mathbf{Q}$, or simply *algebraic*, if it is a root of some polynomial $f(x) \in \mathbf{Q}[x]$ with coefficients in $\mathbf{Q}$. In this case, there exists a unique monic polynomial $f(x) \in \mathbf{Q}[x]$, irreducible in $\mathbf{Q}[x]$, such that $f(a) = 0$. It is called the *minimal polynomial* (over $\mathbf{Q}$) of a; see Theorem 1.4.8. If $\mathbf{Q}(a) \subset \mathbf{C}$ denotes the smallest field that contains both $\mathbf{Q}$ and a then the mapping $\mathbf{Q}[x] \to \mathbf{Q}(a)$ which takes x to a induces an isomorphism

$$\mathbf{Q}[x]/(f) \to \mathbf{Q}(a),$$

where f is the minimal polynomial of a. The proof is left as an exercise. An important result is the following:

Theorem 2.2.2 *Suppose that E and F are algebraic number fields with $F \subseteq E$. Then there is an element $a \in E$ such that $E = F(a)$. In particular, every algebraic number field is of the form $\mathbf{Q}(a)$ for some algebraic number a.*

A field F is said to be *algebraically closed* if every element that is algebraic over F is already in F. This is equivalent to saying that every polynomial with coefficients in F splits as a product of linear factors. Every field is contained in an algebraically closed field, and any two minimal algebraically closed fields containing a given field F are isomorphic. Thus in general we may speak of *the* algebraic closure of a field F.

For example, $\mathbf{C}$ is algebraically closed. The set $\overline{\mathbf{Q}}$ of all algebraic numbers over $\mathbf{Q}$ is an algebraically closed subfield of $\mathbf{C}$, and we shall refer to this particular field as the algebraic closure of $\mathbf{Q}$. It is not a finite extension of $\mathbf{Q}$, so it is not a number field. However, this observation allows us to embed any algebraic number field in the complex numbers. For any prime number p, the set $\mathbf{F}_{p^\infty} = \cup_d \mathbf{F}_{p^d}$ is a field. It is the algebraic closure of every $\mathbf{F}_{p^d}$.

Theorem 2.2.3 *Let F be a number field. Then there are exactly $[F : \mathbf{Q}]$ embeddings of F in $\mathbf{C}$.*

Proof: Let $F = \mathbf{Q}(a)$ and suppose there are k distinct embeddings of F in $\mathbf{C}$. An embedding σ of F in $\mathbf{C}$ is completely determined by its value on a. The image $\sigma(a)$ is a root of the minimal polynomial $f \in \mathbf{Q}[x]$ of a over $\mathbf{Q}$ (thinking of f as a polynomial over $\mathbf{C}$). It is straightforward to check that every root of f determines an embedding. The number of roots of f is exactly its degree, since $\mathbf{C}$ is algebraically closed. Thus the number of embeddings of F in $\mathbf{C}$ is exactly the degree of f, which equals $[F : \mathbf{Q}]$. □

More generally we can consider *extensions* of embeddings. If K is a subfield of the algebraic number field F and τ is an embedding of K in $\mathbf{C}$, then an extension of τ to F is an embedding σ of F in $\mathbf{C}$ such that $\sigma(b) = \tau(b)$ for all $b \in K$.

Theorem 2.2.4 *If K is a subfield of an algebraic number field F, then every embedding of K in $\mathbf{C}$ extends to $[F : K]$ distinct embeddings of F in $\mathbf{C}$.*

Proof: Left as an exercise. □

Definition 2.2.5 *Let F be a number field and let $\sigma_1, \cdots, \sigma_d$ be the distinct embeddings of F in $\mathbf{C}$. Then the trace and norm of an element $b \in F$ are defined as follows.*

1. $Tr_F(b) = \sigma_1(b) + \sigma_2(b) + \cdots + \sigma_d(b)$.
2. $N_F(b) = \sigma_1(b)\sigma_2(b) \cdots \sigma_d(b)$.

Theorem 2.2.6 *Let F be a number field with $[F : \mathbf{Q}] = d$.*

1. *If $b \in F$ and $[\mathbf{Q}(b) : \mathbf{Q}] = e$, then $Tr_F(b) = (d/e)Tr_{\mathbf{Q}(b)}(b) \in \mathbf{Q}$ and $N_F(b) = N_{\mathbf{Q}(b)}(b)^{d/e} \in \mathbf{Q}$. Furthermore, the minimal polynomial (over $\mathbf{Q}$) of b is*
$$f(x) = x^e - Tr_{\mathbf{Q}(b)}(b)x^{e-1} + - \cdots \pm N_{\mathbf{Q}(b)}(b).$$
2. *For every $b, c \in F$, $Tr_F(b + c) = Tr_F(b) + Tr_F(c)$.*
3. *For every $b \in F$ and $u \in \mathbf{Q}$, $Tr_F(ub) = u\,Tr_F(b)$.*
4. *For every $b, c \in F$, $N_F(bc) = N_F(b)N_F(c)$.*
5. *For every $u \in \mathbf{Q}$, $N_F(u) = u^d$.*

Proof: By Theorem 2.2.4, for each embedding τ of $\mathbf{Q}(b)$ in $\mathbf{C}$, the number $\tau(b)$ occurs exactly d/e times as a summand in the definition of $\mathrm{Tr}_F(b)$ and as a factor in the definition of $N_F(b)$. Thus $\mathrm{Tr}_F(b) = (d/e)\mathrm{Tr}_{\mathbf{Q}(b)}(b)$ and $N_F(b) = N_{\mathbf{Q}(b)}(b)^{d/e}$. It is apparent from the proof of Theorem 2.2.3 that if $\tau_1, \cdots, \tau_e$ are the distinct embeddings of $\mathbf{Q}(b)$ in $\mathbf{C}$, then the minimal polynomial (over $\mathbf{Q}$) of b is

$$f(x) = \prod_{i=1}^{e}(x - \tau_i(b)) = x^e - \mathrm{Tr}_{\mathbf{Q}(b)}(b)x^{e-1} + - \cdots \pm N_{\mathbf{Q}(b)}(b).$$

It follows that the trace and norm are in $\mathbf{Q}$.

The arithmetic properties of the trace and norm follow from the properties of an embedding σ of F in $\mathbf{C}$: $\sigma(b + c) = \sigma(b) + \sigma(c)$ and $\sigma(bc) = \sigma(b)\sigma(c)$ if $b, c \in F$, and $\sigma(u) = u$ if $u \in \mathbf{Q}$. □

2.2.b Algebraic Integers

Just as algebraic number fields are generalizations of the rational numbers, there is a generalization of the rational integers $\mathbf{Z}$.

Definition 2.2.7 *An algebraic number a is an* algebraic integer *or is* integral *if its minimal polynomial $f \in \mathbf{Q}[x]$ over $\mathbf{Q}$ has all its coefficients in $\mathbf{Z}$.*

Theorem 2.2.8 *The following are equivalent*

1. *a is an algebraic integer.*
2. *$\mathbf{Z}[a]$ is a finitely generated $\mathbf{Z}$-module.*
3. *$a \in R$ for some ring $R \subseteq \mathbf{C}$ that is a finitely generated $\mathbf{Z}$-module.*
4. *$aM \subseteq M$ for some finitely generated $\mathbf{Z}$-module $M \subseteq \mathbf{C}$.*

Proof: If a is an algebraic integer, then a^d is a linear combination of $1, a, \cdots, a^{d-1}$ with integer coefficients, and it follows that $\mathbf{Z}[a]$ is generated as a $\mathbf{Z}$-module by $1, a, \cdots, a^{d-1}$. The implications (2) $\Longrightarrow$ (3) $\Longrightarrow$ (4) are straightforward.

To prove that (4) implies (1), suppose that M is generated by $m_1, \cdots, m_k$. Thus for $j = 1, \cdots, k$, we have

$$am_j = \sum_{i=1}^{k} b_{i,j}m_j \tag{2.3}$$

with $b_{i,j} \in \mathbf{Z}$. Let $c_{i,j} = b_{i,j}$ if $i \neq j$, and $c_{i,i} = b_{i,i} - x$. It follows from equation (2.3) that the determinant of the matrix $[c_{ij}]$ is zero at $x = a$. But the determinant of this matrix is a monic polynomial with integer coefficients, so a is algebraic. □

2.2.c Orders

Let F be an algebraic number field. If $R \subset F$ is a sub-ring, then it is automatically an integral domain. An *order* $R \subset F$ is a subring of F such that its additive group R^+ (meaning that we forget about the multiplication for the moment) is finitely generated and has maximal rank in F. In this case, Corollary 1.1.17 implies that R^+ is isomorphic to $\mathbf{Z}^m$ for some integer m. A standard result is the following.

Theorem 2.2.9 *A sub-ring R in a number field F is an order in F if and only if it satisfies the following three conditions,*

1. $R \cap \mathbf{Q} = \mathbf{Z}$
2. *The fraction field (§1.2.e) of R is F.*
3. *R^+ is finitely generated, as an Abelian group.*

Except when $F = \mathbf{Q}$ there are infinitely many orders in F. Every order $R \subset F$ consists entirely of algebraic integers and in fact the intersection $\mathbf{Z}_F = F \cap \mathbf{A}$ (where $\mathbf{A}$ denotes the set of all algebraic integers) is an order which contains all the other orders in F. This maximal order $\mathbf{Z}_F$ is called the *ring of integers* of F.

The ring of integers of $\mathbf{Q}$ is $\mathbf{Z}$; the ring of integers of $\mathbf{Q}[i]$ is $\mathbf{Z}[i]$. However the ring of integers of $\mathbf{Q}[\sqrt{5}]$ is larger than $\mathbf{Z}[\sqrt{5}]$ (which is an order). Rather, the ring of integers consists of all integer combinations of $(1+\sqrt{5})/2$ and $(1-\sqrt{5})/2$. For any number field F the maximal order $\mathbf{Z}_F$ has several particularly nice properties (it is a Dedekind ring, for example). However in §**??** we will consider algebraic shift registers whose entries come from an arbitrary order in an arbitrary number field.

2.3 Local fields

There are two more types of fields that we will encounter: *function fields* and *p-adic fields*, both of which contain a local ring R of "integers". These will be discussed in more detail in Chapter 4, however here is a preview. If F is a field, then the (local) *function field* $F((x))$ consists of all *formal Laurent series* $\sum_{i=-k}^{\infty} a_i x^i$, with $a_i \in F$. Such a series has finitely many terms of negative degree and possibly infinitely many terms of positive degree. Its ring of "integers" is the sub-ring $F[[x]]$ of formal power series, that is, sums with no terms of negative degree. Every formal Laurent series $a(x) \in F((x))$ may be expressed as a quotient $a(x) = f(x)/g(x)$ of two formal power series $f, g \in F[[x]]$ and in fact the denominator $g(x)$ may be chosen to be a power of x. Addition and multiplication in $F((x))$ are performed in the obvious way, analogous to that of addition and multiplication of polynomials.

Let p be a prime number. The p-adic field $\mathbf{Q}_p$ consists of all formal Laurent series $\sum_{i=-k}^{\infty} a_i p^i$ (with finitely many terms of negative degree and possibly infinitely many terms of positive degree), where $0 \le a_i \le p-1$, and where addition and multiplication are performed “with carry”. It contains a ring $\mathbf{Z}_p$ of “integers” consisting of formal power series with no terms of negative degree. Every $a \in \mathbf{Q}_p$ can be expressed as a fraction f/g with $f, g \in \mathbf{Z}_p$ and in fact the denominator g may be chosen to be a power of p.

2.4 Exercises

1. Lemma 2.1.12: Let d and e be positive integers with d dividing e. Prove that if $a \in \mathbf{F}_{p^e}$, then $a + a^{p^d} + a^{p^{2d}} + \cdots + a^{p^{e-d}} \in \mathbf{F}_{p^a}$.

2. Suppose p is prime and c, d, and e are integers with $c|d|e$. Prove that $\mathrm{Tr}^{p^d}_{p^c} \circ \mathrm{Tr}^{p^e}_{p^d} = \mathrm{Tr}^{p^e}_{p^c}$.

3. Develop an alternate definition of the trace function for a finite field F in terms of embeddings of F in its algebraic closure. Prove that your definition agrees with the previous one.

Chapter 3 Finite Rings and Galois Rings

3.1 Finite Local Rings

In this section we examine the structure of a commutative ring (with identity) which has finitely many elements. The standard reference for this section is [17]. During the last decade a considerable amount of effort has been directed towards developing linear feedback shift register sequences based on a finite local ring R. The analysis of these sequences depends on an understanding of the units in R (see Theorem **??**).

Let R be a commutative ring. Recall from Definition 1.2.10 that R is said to be a *local ring* if it contains a unique maximal ideal $\mathfrak{m}$. In this case (see §1.2.a), the maximal ideal $\mathfrak{m}$ consists precisely of the non-units of R. The quotient $F = R/\mathfrak{m}$ is called the *residue field* of R. For each $i \geq 0$ the quotient $\mathfrak{m}^{i-1}/\mathfrak{m}^i$ is naturally a vector space over F (because R acts on this quotient by multiplication, and $\mathfrak{m}$ acts trivially). For the remainder of this section we assume that R is a finite local ring. The following are examples of finite local rings.

- any finite (Galois) field.
- $\mathbf{Z}/(p^n)$ for any prime number p, with maximal ideal (p) and residue field $\mathbf{Z}/(p)$.
- $\mathbf{F}[x]/(f^n)$, where $\mathbf{F}$ is a finite field and f is an irreducible polynomial, with maximal ideal (f) and residue field $\mathbf{F}[x]/(f)$.
- $R[x]/(f^n)$ where R is a finite local ring and f is a basic irreducible polynomial (see below).

Any commutative finite ring may be expressed as a direct sum of finite local rings.

Basic irreducible polynomials: Let R be a finite local ring with maximal ideal $\mathfrak{m}$. Let $\mu : R \to F = R/\mathfrak{m}$ be the projection. Applying μ to each coefficient of a polynomial gives a mapping which we also denote by $\mu : R[x] \to F[x]$. A polynomial $f(x) \in R[x]$ is *regular* if it is not a zero divisor, which holds if and only if $\mu(f) \neq 0$. Let $f(x) \in R[x]$. If $\mu(f)$ is nonzero and is irreducible in $F[x]$ then f is irreducible in $R[x]$, and we refer to f as a *basic irreducible polynomial.* In this case $R[x]/(f^n)$ is again a local ring for any $n > 0$ (see ([17], XIV.10). Its maximal ideal is $\mathfrak{m}[x] + (f)$ and its residue field is $F[x]/(\mu(f))$, where $\mathfrak{m}[x]$ is the collection of those polynomials $f \in R[x]$ all of whose coefficients are in $\mathfrak{m}$.

If the leading term of a basic irreducible polynomial $f(x) \in R[x]$ is in the maximal ideal $\mathfrak{m}$ then the degree of the reduction $\mu(f) \in F[x]$ will be less than $\deg(f)$. If $f(x)$ is a *monic* polynomial then $\deg(f) = \deg(\mu(f))$ since the leading term is 1. For this reason we will often consider monic basic irreducible polynomials.

Lemma 3.1.1 *Let $f \in R[x]$ be a regular polynomial and suppose $\bar{\alpha} \in F$ is a simple zero of $\mu(f) \in F[x]$. Then f has one and only one root $\alpha \in R$ such that $\mu(\alpha) = \bar{\alpha}$.*

Proof: This is proven in Lemma (XV.1) of [17]. □

Further properties of polynomials over R are described in §3.3. The following is a powerful tool for studying local rings.

Theorem 3.1.2 *(Nakayama's Lemma for local rings [17], [16, p. 11]) Let R be a finite local ring with maximal ideal $\mathfrak{m}$. Let M be a module over R.*

1. *If M is finite and $\mathfrak{m}M = M$, then $M = 0$.*
2. *If N is a submodule of M and $M = N + \mathfrak{m}M$, then $N = M$.*

3.1.a Units in a finite local ring

Let R be a finite local ring with maximal ideal $\mathfrak{m}$ and residue field F. Let $R^\times$ be the set of invertible elements in R. Let $1 + \mathfrak{m} = \{1 + a : \ a \in \mathfrak{m}\}$. By [17], Theorem (V.1) and Proposition (IV.7),

- the ideal $\mathfrak{m}$ consists precisely of the non-units of R,
- for every $a \in R$, at least one of a and $1 + a$ is a unit, and
- there is a positive integer n such that $\mathfrak{m}^n = 0$.

The details are left as an exercise.

Proposition 3.1.3 *There exists an isomorphism of Abelian groups*

$$R^\times \cong F^\times \times (1 + \mathfrak{m}) \tag{3.1}$$

Proof: Let n be the smallest integer such that $\mathfrak{m}^n = 0$. It is called the *degree of nilpotency* of $\mathfrak{m}$. As in [17] Exercise (V.9), we have a sequence of surjective ring homomorphisms

$$R = R/\mathfrak{m}^n \xrightarrow{\sigma_n} R/\mathfrak{m}^{n-1} \xrightarrow{\sigma_{n-1}} \cdots \xrightarrow{\sigma_2} R/\mathfrak{m} = F.$$

For $2 \le i \le n$, the kernel $\ker(\sigma_i) = \mathfrak{m}^{i-1}/\mathfrak{m}^i$ is a vector space over F. If $|F| = q$ it follows by induction that there exists an integer j such that

$$|\mathfrak{m}| = q^j \ \text{ and } \ |R| = q^{j+1}. \tag{3.2}$$

The natural ring homomorphism $\mu : R \to F = R/\mathfrak{m}$ gives an exact sequence of (multiplicative) Abelian groups,

$$1 \to 1 + \mathfrak{m} \to R^\times \to F^\times \to 1.$$

The Abelian group $F^\times$ is cyclic of order $q-1$, and $1+\mathfrak{m}$ has order q^j, which is relatively prime to $q-1$. It follows (from the structure theorem for finite Abelian groups, Theorem 1.1.15) that there is a splitting $\iota : F^\times \to R^\times$ and this gives the isomorphism (3.1). □

The structure of $1+\mathfrak{m}$ is often very complicated. However it is possible to identify the cyclic group $F^\times$ as a subgroup of $R^\times$.

Lemma 3.1.4 *There is a unique (group homomorphism) splitting $\iota : F^\times \to R^\times$ of the projection μ, and its image consists of all elements $\alpha \in R$ such that $\alpha^{q-1} = 1$.*

Proof: Every element $a \in F^\times$ satisfies $a^{q-1} = 1$ so if ι exists, the same must be true of $\iota(a)$. Let $g(x) = x^{q-1} - 1$. Then every element of $F^\times$ is a (simple) root of $\mu(g) \in F[x]$. Therefore g is a regular polynomial, and Lemma 3.1.1 implies that every element $a \in F^\times$ has a unique lift $\iota(a) \in R$ such that $\iota(a)^{q-1} = 1$. Hence the splitting ι exists, and there is only one such. □

3.2 Examples

3.2.a $\mathbf{Z}/(p^m)$

Fix a prime number $p \in \mathbf{Z}$ and let $R = \mathbf{Z}/(p^m)$. This is a finite local ring with maximal ideal $\mathfrak{m} = (p)$ and residue field $F = \mathbf{Z}/(p)$. The multiplicative group $F^\times$ is cyclic, of order $p-1$. By Proposition 3.1.3 the group of units $R^\times$ is the product $F^\times \times (1+\mathfrak{m})$.

Proposition 3.2.1 *If $p > 2$ then $1+\mathfrak{m}$ is a cyclic group of order p^{m-1} so $R^\times \cong \mathbf{Z}/(p-1) \times \mathbf{Z}/(p^{m-1}) \cong \mathbf{Z}/(p^{m-1}(p-1))$. If $p = 2$ and if $m \geq 3$ then $1+\mathfrak{m}$ is a product of two cyclic groups, one of order 2 (generated by the element -1), the other of order 2^{m-2} (generated by the element 5).*

Proof: The order of the group of units is easy to calculate: since every pth integer is a multiple of p, there are $p^m/p = p^{m-1}$ non-invertible elements in R. So there are $p^m - p^{m-1} = (p-1)p^{m-1}$ units. It follows that $1+\mathfrak{m}$ contains p^{m-1} elements.

Now consider the case $p \geq 3$. Define $E : \mathbf{Z} \to R = \mathbf{Z}/(p^m)$ by $E(a) = \exp(pa) \pmod{p^m}$. That is,

$$E(a) = 1 + pa + \frac{p^2a^2}{2!} + \frac{p^3a^3}{3!} + \cdots \pmod{p^m} \tag{3.3}$$

Consider the nth term, $a^np^n/n!$. The number $n!$ is not necessarily invertible in $\mathbf{Z}/(p^m)$ but the number $p^n/n!$ does make sense in $\mathbf{Z}/(p^n)$ if we interpret it to mean that the factor p^e which occurs in the prime decomposition of $n!$ should be canceled with the same factor p^j

which occurs in the numerator. In fact, the prime p occurs in the prime decomposition of $n!$ fewer than $n/p+n/p^2+n/p^3\cdots = n/(p-1)$ times. Since it occurs in the numerator n times, it is possible to cancel all occurrences of p from the denominator. This leaves a denominator which is relatively prime to p and hence is invertible in $\mathbf{Z}/(p^m)$. It follows, moreover, that after this cancellation the numerator still has at least $n(p-2)/(p-1)$ factors of p. So if $n \ge m(p-1)/(p-2)$ the term $a^n p^n/n!$ is 0 in $\mathbf{Z}/(p^m)$. Therefore the sum (3.3) is finite.

Since $E(a+b) = E(a)E(b)$, the mapping E is a group homomorphism. Moreover $E(a) = 1$ if and only if a is a multiple of p^{m-1}. So E induces to an injective homomorphism

$$E : \mathbf{Z}/(p^{m-1}) \to 1 + \mathfrak{m}.$$

This mapping is also surjective because both sides have p^{m-1} elements.

Now consider the case $R = \mathbf{Z}/(2^m)$ with $m \ge 3$. The element $\{-1\}$ generates a cyclic subgroup of order 2. The element 5 generates a cyclic subgroup of order 2^{m-2}. To show this, first verify by induction that

$$5^{2^{m-3}} = (1+2^2)^{2^{m-3}} \equiv 1 + 2^{m-1} \pmod{2^m}$$

so this number is not equal to 1 in $\mathbf{Z}/(2^m)$. However

$$5^{2^{m-2}} \equiv (1+2^{m-1})^2 \equiv 1 \pmod{2^m}.$$

So 5 has order 2^{m-2} in R. Since -1 is not a power of 5 $\pmod 4$ it is also not a power of 5 $\pmod{2^m}$. Therefore the product of cyclic groups $\langle -1\rangle\langle 5\rangle$ has order 2^{m-1}, and it consequently exhausts all the units. □

3.2.b $F[x]/(x^m)$

Let F be a field and let $R = F[x]/(x^m)$. Then R is a local ring with maximal ideal $\mathfrak{m} = (x)$ and with residue field F. The mapping $\mu : R \to F$ (which associates to each polynomial its constant term) takes $R^\times$ surjectively to $F^\times$. This mapping has a splitting $F^\times \to R^\times$ which assigns to any nonzero $a \in F$ the polynomial $a + 0x$. This gives an isomorphism $R^\times \cong F^\times \times (1+\mathfrak{m})$, where $1+\mathfrak{m}$ is the (multiplicative) group of all polynomials of the form $1 + xh(x)$, $h(x)$ a polynomial of degree $\le m-2$. In the case that F is a finite field, the ring R is a finite local ring, and we have recovered Proposition 3.1.3. The structure of the group $1+\mathfrak{m}$ is fairly complicated in general, but it can be described simply in some cases.

Proposition 3.2.2 *If char$(F) = 0$ or if $m <$ char(F) then the group $1+\mathfrak{m}$ is isomorphic to the additive group F^{m-2}.*

Proof: If $n < \mathrm{char}(F)$ or $\mathrm{char}(F) = 0$, then the number $n!$ is invertible in F and, for any $a \in F$ we may define

$$\exp(ax) = 1 + ax + a^2x^2/2 + \cdots + a^{m-1}x^{m-1}/(m-1)! \pmod{x^m} \in 1 + \mathfrak{m}.$$

This mapping $a \mapsto \exp(ax)$ is a homomorphism from (the additive group) F into (the multiplicative group) $1 + \mathfrak{m}$ whose inverse $\pi_1 : (1 + \mathfrak{m}) \to F$ assigns to any polynomial $h(x) = 1 + h_1x + h_2x^2 + \cdots$ the coefficient h_1. The kernel of π_1 is the subgroup $1 + \mathfrak{m}^2$ of $R^\times$ consisting of all polynomials of the form $1 + x^2h(x)$. The mapping $F \to 1 + \mathfrak{m}^2$ which is given by $a \mapsto \exp(ax^2)$ is again a homomorphism, whose inverse $\pi_2 : 1 + \mathfrak{m}^2 \to F$ assigns to any polynomial $h(x) = 1 + h_2x^2 + \cdots$ the coefficient h_2. Continuing in this way we construct an isomorphism $F^{m-2} \to 1 + \mathfrak{m}$ given by

$$(a_1, a_2, \cdots, a_{m-2}) \mapsto \exp(a_1x)\exp(a_2x^2)\cdots\exp(a_{m-2}x^{m-2}).$$

This completes the proof of the Proposition. □

If $m \geq \mathrm{char}(F) > 0$ this argument fails. However it is still possible to describe the structure of $1 + \mathfrak{m}$ **Citation?**.

more to say here???

3.2.c $F[x]/(f^m)$

Let F be a finite field and let $f \in F[x]$ be an irreducible polynomial. Fix $m \geq 1$. The ring $R = F[x]/(f^m)$ is a finite local ring with maximal ideal (f) and with quotient field $K = F[x]/(f)$. Let

$$\mu : R = F[x]/(f^m) \to K = F[x]/(f)$$

be reduction modulo f. It is a surjective ring homomorphism.

Proposition 3.2.3 *There is a unique splitting of μ. That is, there is a unique injective ring homomorphism $\varphi : K \to R$ so that $\mu(\varphi(a)) = a$ for all $a \in K$. Moreover the mapping φ extends to a mapping $\varphi : K[y] \to R$ by setting $\varphi(y) = f$. The resulting mapping*

$$\bar{\varphi} : K[y]/(y^m) \to R$$

is an isomorphism of rings.

Proof: Let q denote the number of elements in F and let $Q = q^d$ denote the number of elements in K, where $d = \deg(f)$. First we show that the set

$$Z_m = \left\{g \in F[x]/(f^m) : \ g^Q = g\right\}$$

is a lift of the field K into R and is therefore a candidate for the image of φ.

The set Z_m is closed under addition and multiplication, because if $g_1, g_2 \in Z_m$ then $(g_1 + g_2)^Q = g_1^Q + g_2^Q = g_1 + g_2$. Moreover the restriction $\mu : Z_m \to K$ is an injection, for if $g \in Z_m$ lies in the kernel of μ and if $\dot{g} \in F[x]$ is any lift of g, then f divides $\dot{g}$. However f^m divides $\dot{g}^Q - \dot{g} = (\dot{g}^{Q-1} - 1)(\dot{g})$. Since these two factors are relatively prime, it follows that f^m divides $\dot{g}$, which says that $g = 0$ in R. Now let us show that the restriction $\mu : Z_m \to K$ is surjective. Fix $a \in K$. We need to find $g \in Z_m$ so that $\mu(g) = a$. We use induction on m, and the case $m = 1$ holds trivially. So let m be arbitrary and consider the mapping $\mu_m : F[x]/(f^m) \to F[x]/(f^{m-1})$. By induction, there exists $g' \in F[x]/(f^{m-1})$ so that $(g')^Q = g'$ and so that g' maps to the given element $a \in K$, that is, $g' \pmod{f} = a$. Let $\dot{g}' \in F[x]$ be any lift of g'; then f^{m-1} divides $(\dot{g}')^Q - \dot{g}'$, or

$$(\dot{g}')^Q - \dot{g}' = f^{m-1}h$$

for some polynomial $h \in F[x]$. Set $g = \dot{g}' + hf^{m-1}$. Then

$$(g)^Q - g = (\dot{g}')^Q - \dot{g}' + h^Q f^{(m-1)Q} - hf^{m-1} = h^Q f^{(m-1)Q}$$

which is divisible by f^m. This says that the class $[g] \in F[x]/(f^m)$ lies in the set Z_m and that $g \pmod{f} = a$ as needed.

We have shown that there is a unique injective homomorphism $\varphi : K \to R$. This function extends to a function $\varphi : K[y] \to R$ by mapping y to f. We claim that the kernel of φ is (y^m) and that φ is onto. The kernel contains (y^m) since $f^m = 0$ in R. Let $g(y) = \sum_{i=0}^{m-1} g_i y^i$ with $\varphi(g) = 0$. Thus

$$\sum_{i=0}^{m-1} g_i f^i = 0. \tag{3.4}$$

As a vector space over K the ring R has dimension m since $|R| = Q^m = |K|^m$. R is spanned over K by $\{1, f, f^2, \cdots, f^{m-1}\}$. Therefore these elements form a basis. As we have seen in the preceding paragraph, the projection $\mu_m : F[x]/(f^m) \to F[x]/(f^{m-1})$ takes Z_m to Z_{m-1} (both of which are lifts of the field K). Applying the projection μ_m to equation (3.4) gives

$$\sum_{j=0}^{m-2} g_j f^j = 0$$

and by induction we conclude that $g_0 = g_1 = \ldots = g_{m-2} = 0$. This leaves $g_{m-1}f^{m-1} = 0$ in the ring R, which means that f^m divides $g_{m-1}f^{m-1}$ in the polynomial ring $F[x]$. But $F[x]$ is an integral domain, so we conclude that f divides g_{m-1}, hence $g_{m-1} = 0$ as an element of K.

In conclusion, we obtain a well defined surjective ring homomorphism $K[y] \to R$ by sending y to f. The kernel of this homomorphism is clearly the ideal (y^m) so we obtain an isomorphism $K[y]/(y^m) \to R$. □

3.3 Divisibility in $R[x]$

Throughout this subsection, R denotes a finite local ring with $\mu : R \to F = R/\mathfrak{m}$ the projection to its residue field. Let $f, g \in R[x]$.

1. f is *nilpotent* if $f^n = 0$ for some $n \geq 0$.
2. f is a *unit* if there exists $h \in R[x]$ so that $fh = 1$.
3. f is *regular* if f is not a zero divisor.
4. f is *prime* if the ideal (f) is a proper prime ideal.
5. f is *irreducible* if f is not a unit and, whenever $f = gh$ then g or h is a unit.
6. f and g are *coprime* if $R[x] = (f) + (g)$.

In [17] the following results are proven.

Theorem 3.3.1 *Let $f = a_0 + a_1x + \cdots + a_dX^d \in R[x]$. Then*

1. *The following are equivalent:*
   - *(a) f is a unit.*
   - *(b) $\mu(f) \in F[x]$ is a unit.*
   - *(c) a_0 is a unit and the remaining coefficients $a_1, \cdots, a_d$ are nilpotent.*
2. *The following are equivalent:*
   - *(a) f is nilpotent.*
   - *(b) $\mu(f) = 0$.*
   - *(c) All the a_i are nilpotent.*
   - *(d) f is a zero divisor.*
   - *(e) there exists $a \neq 0$ in R such that $af = 0$.*
3. *The following are equivalent:*
   - *(a) f is regular.*
   - *(b) $\mu(f) \neq 0$.*
   - *(c) a_i is a unit for some i $(0 \leq i \leq d)$.*
4. *f and g are coprime if and only if $\mu(f)$ and $\mu(g)$ are coprime. In this case, f^i and g^j are coprime for all $i, j \geq 1$.*
5. *If $\mu(f)$ is irreducible then f is irreducible. If f is irreducible then $\mu(f) = ag^n$ where $a \in F$ and $g \in F[x]$ is a monic irreducible polynomial.*

6. *(Euclidean algorithm) If* $f \neq 0$ *and if* $g \in R[x]$ *is regular then there exist (not necessarily unique) elements* $q, r \in R[x]$ *such that* $\deg r < \deg g$ *and* $f = gq + r$.
7. *If* f *and* g *are monic and regular and if* $(f) = (g)$ *then* $f = g$.

Recall that an ideal $I \subset R[x]$ is *primary* if $I \neq R[x]$ and whenever $ab \in I$, then either $a \in I$ or $b^n \in I$ for some $n \geq 1$. An element $g \in R[x]$ is primary if (g) is primary.

Proposition 3.3.2 *An element* $f \in R[x]$ *is a primary regular non-unit if and only if* $f = ug^n + h$ *where* $u \in R[x]$ *is a unit,* $g \in R[x]$ *is a basic irreducible,* $n \geq 1$*, and* $h \in \mathfrak{m}[x]$ *(that is, all the coefficients of* h *lie in* $\mathfrak{m}$*).*

Although $R[x]$ is not necessarily a unique factorization domain, the following theorem ([17] Thm. XIII.11) states that regular polynomials have unique factorization.

Theorem 3.3.3 *Let* $f \in R[x]$ *be a regular polynomial. Then there exist unique (up to reordering and multiplication by units) regular coprime primary polynomials* $g_1, g_2, \cdots, g_n \in R[x]$ *so that* $f = g_1 g_2 \cdots g_n$.

3.4 Tools for Local Rings

In this section we develop several tools for the analysis of finite local rings – Galois theory, the trace and norm, and primitive elements. These are all generalizations of the similarly named tools for analyzing finite fields, and in most cases we use the finite field versions to help construct the finite local ring version.

3.4.a Galois theory of local rings

In the next few paragraphs we will see that a finite local ring R has a distinguished collection of *Galois extensions* $\mathrm{GR}(R, n)$, one for each positive integer n, which are themselves local rings and for which many of the familiar properties of Galois fields continue to hold.

Extensions. Let R be a finite local ring. An *extension* ring is a finite local ring S which contains R. Any extension S of R has the structure of an R-algebra, that is, R acts on S such that $a(c + d) = ac + ad$ and $a(cd) = (ac)d$ for all $a \in R$ and all $c, d \in S$. A (ring) homomorphism $\varphi : S \to S$ is said to be an R-algebra automorphism of S provided it is both surjective and injective, and provided $\varphi(ac) = a\varphi(c)$ for all $a \in R$ and $c \in S$. Define the *Galois group*

$$G = \mathrm{Gal}(S/R) = \mathrm{Aut}_R(S)$$

to be the set of R-algebra automorphisms of S. The Galois group G acts on S. Let S^G denote the set of elements which are fixed under the action of G (hence $R \subset S^G$). An extension S of R is *unramified* if the maximal ideal $\mathfrak{m}$ of R generates the maximal ideal $\mathfrak{M}$ of S; otherwise it is said to be *ramified.* If S is an unramified extension of R then $\mathfrak{m}^i$ generates $\mathfrak{M}^i$ so the degree of nilpotency of $\mathfrak{m}$ equals the degree of nilpotency of $\mathfrak{M}$. An unramified extension $R \subset S$ is said to be a *Galois extension* if $R = S^G$.

Example Let R be a finite local ring with maximal ideal $\mathfrak{m}$. Let $f \in R[x]$ be a monic basic irreducible polynomial. The extension $S = R[x]/(f^m)$ is again a finite local ring (see §3.1). Its maximal ideal is $\mathfrak{M} = \mathfrak{m} + (f)$. If $m > 1$ then S is a *ramified* extension of R. If $m = 1$ then S is an unramified extension and $\mathfrak{M} = \mathfrak{m}S$ is generated by $\mathfrak{m}$.

The following result is the main theorem in the Galois theory of finite local rings. The proof may be found in [17].

Theorem 3.4.1 *Let R be a finite local ring. Then every unramified extension $R \subset S$ is a Galois extension. Suppose $R \subset S$ is such an extension, with corresponding maximal ideals $\mathfrak{m} \subset \mathfrak{M}$. Then the following diagram*

$$\begin{tikzcd} S \arrow[r, "\nu"] & K = S/\mathfrak{M} \\ R \arrow[u, phantom, "\bigcup"] \arrow[r, "\mu"] & F = R/\mathfrak{m} \arrow[u, phantom, "\bigcup"] \end{tikzcd} \tag{3.5}$$

induces an isomorphism Gal(S/R) $\cong$ Gal(K/F) which is therefore a cyclic group. There exists $h \in S$ so that $S = R[h]$. The mapping determined by $h \mapsto h^{|F|}$ generates Gal(S/R). Let $h = h_1, h_2, \ldots, h_d$ be the distinct images of h under Gal(S/R). Then the following polynomial

$$f(x) = (x - h_1)(x - h_2) \cdots (x - h_d) \tag{3.6}$$

actually lies in $R[x]$. It is a (monic) basic irreducible polynomial of degree $d = |Gal(S/R)|$. The mapping $R[x]/(f) \to S$ which takes $x \in R[x]$ to $h \in S$ is an isomorphism of rings (and of R-algebras). The ring S is a free module of rank d over the ring R, hence $|S| = |R|^d$ and we say that S is an extension of degree d. *The above diagram induces a lattice preserving bijection between the Galois extensions of R which are contained in S and the field extensions of F which are contained in K. The ring S is a field if and only if the ring R is a field. If $f' \in R[x]$ is another monic basic irreducible polynomial of the same degree d then there exists an R-algebra isomorphism $S \cong R[x]/(f')$. In particular, f' also splits into linear factors over S.*

Corollary 3.4.2 *Let R be a finite local ring, let S be an unramified degree d extension of R, and let $f \in R[x]$ be a monic basic irreducible polynomial of degree d. Let $\alpha \in S$ be a root of f. Then the collection $\left\{1, \alpha, \alpha^2, \cdots, \alpha^{d-1}\right\}$ forms a basis of S over R. The element α is invertible in S.*

Proof: According to Theorem 3.4.1, we may replace S with $R[x]/(f)$ and we may replace α with x. But it is clear that the set $\left\{1, x, x^2, \cdots, x^{d-1}\right\}$ forms a basis of $R[x]/(f)$ over R. If $f(x) = a_0 + a_1 x + \cdots + a_d x^d$ then $\mu(a_0) \neq 0$ since $\mu(f)$ is irreducible. Therefore a_0 is invertible in S and

$$x^{-1} = \frac{-1}{a_0}(a_1 + a_2 x^2 + \cdots + a_d x^{d-1})$$

in $R[x]/(f)$. □

3.4.b The trace

Let $R, \mathfrak{m}, F = R/\mathfrak{m}$ be a finite local ring with $\mu : R \to F$ the reduction map. Let $S, \mathfrak{M}, K = S/\mathfrak{M}$ be a Galois extension of degree d with $\nu : S \to K$ the reduction map. Let $a \in S$. The *trace* $\mathrm{Tr}_{S/R}(a) \in R$ and *norm* $N_{S/R}(a) \in R$ of a are defined to be

$$\mathrm{Tr}_{S/R}(a) = \sum_{\sigma \in \mathrm{Gal}(S/R)} \sigma(a)$$

and

$$N_{S/R}(a) = \prod_{\sigma \in \mathrm{Gal(S/R)}} \sigma(a).$$

Let $\sigma_S \in \mathrm{Gal}(S/R)$ be a generator of the Galois group. Then $N_{S/R}(a) = 1$ if and only if there is a unit $b \in S$ so that $a = b\sigma(b)^{-1}$, and $\mathrm{Tr}_{S/R}(a) = 0$ if and only if there exists $c \in S$ such that $a = c - \sigma(c)$.

Consider the mapping $\kappa_a : S \to S$ which is given by multiplication by a. Since S is a free module over R it has a basis consisting of d elements, and the mapping κ_a may be expressed as a $d \times d$ matrix M_a. Then the trace and norm of a equal the trace and determinant (respectively) of this matrix (which are thus independent of the choice of basis).

Lemma 3.4.3 *$Tr_{S/R}(a)$ equals the trace of M_a and $N_{S/R}(a)$ equals the determinant of M_a. Also, we have $\mu \circ Tr_{S/R} = Tr_{K/F} \circ \nu$ and $\mu \circ N_{S/R} = N_{K/F} \circ \mu$.*

Proof: The last statement of the theorem follows from Theorem 3.4.1. We know the first statement concerning the trace is true for the fields K and F by Proposition 2.1.14. Let N be the set of elements a of S such that the trace of a equals the trace of M_a. Then N is an R-submodule of S since the mapping from a to the trace of M_a is R-linear. Moreover $S = N + \mathfrak{M}S = N + \mathfrak{m}S$. By Nakayama's lemma (Theorem 3.1.2) we have $S = N$, which proves the claim.

Next we consider the norm. Let us denote the determinant of M_a by $D(a)$. We want to show that $D(a) = N_{S/R}(a)$ for every $a \in S$. Since both $N_{S/R}$ and D are multiplicative, it suffices to show this for a set V such that every element of S is a product of elements of V.

If $a \in R$, then $M_a = aI$ so $D(a) = a^d$, and $N_{S/R}(a) = a^d$.

Suppose that $a \in S$ reduces to a primitive element of K modulo $\mathfrak{M}$. If N is the R-submodule of S spanned by $1, a, \cdots, a^{d-1}$, then $S = N + \mathfrak{M}$, so by Nakayama's lemma $S = N$. That is, $1, a, \cdots, a^{d-1}$ is an R-basis for S. With respect to this basis M_a has the form described in Proposition 2.1.14. If $f(x) = x^d + \sum_{i=0}^{e/d-1} a_i x^i$ is the minimal polynomial of a over R, then $D(a) = a_0 = N_{S/R}(a)$. Thus $D(a^i) = N_{S/R}(a^i)$ for every i. If n is the degree of nilpotency of S and R, then $|S| = |K|^n$. We have thus far accounted for the $(|K|-2)|K|^{n-1}$ elements of S that are congruent to some a^i, $i = 1, \cdots, |K|-2$. We also have $D((a+b)/a) = N_{S/R}((a+b)/a)$ if $b \in \mathfrak{M}$. This accounts for the $|\mathfrak{M}| = |K|^{n-1}$ elements in $1 + \mathfrak{M}$, and hence for all the units. Finally, since $\mathfrak{M} = \mathfrak{m}S$, every element of $\mathfrak{M}$ can be written in the form cb with $c \in \mathfrak{m}^i$ for some i and b a unit. Using multiplicativity again completes the proof. □

Corollary 3.4.4 *The trace $Tr_{S/R} : S \to R$ is surjective.*

Proof: First we show there exists an element $s \in S$ so that $\mathrm{Tr}(s)$ is invertible in R. If this were false, then we would have $\mathrm{Tr}(s) \in \mathfrak{m}$ for all $s \in S$ which would imply that the induced mapping $S/\mathfrak{M} \to R/\mathfrak{m}$ is 0. This would contradict the above lemma which states that this induced mapping is the trace, $\mathrm{Tr}_{K/F}$, which is surjective. So choose $c \in S$ so that $\mathrm{Tr}_{S/R}(c)$ is invertible and let $a \in R$ denote its inverse. Then for any $b \in R$ we have $\mathrm{Tr}_{S/R}(bac) = baTr_{S/R}(c) = b$. □

Suppose $L : S \to R$ is any R-linear mapping. Then for any $i \geq 1$ we have $L(\mathfrak{M}^i) \subset \mathfrak{m}^i$. (Since $\mathfrak{M} = \mathfrak{m}S$, any element in $\mathfrak{M}^i$ may be expressed as ac with $a \in \mathfrak{m}^i$ and $c \in S$, in which case $L(ac) = aL(c) \in \mathfrak{m}^i$.) In particular, L induces an F-linear mapping $\bar{L} : K = S/\mathfrak{M} \to F = R/\mathfrak{m}$ and the diagram

$$\begin{tikzcd} S \arrow[r, "\nu"] \arrow[d, "L"'] & K = S/\mathfrak{M} \arrow[d, "\bar{L}"] \\ R \arrow[r, "\mu"] & F = R/\mathfrak{m} \end{tikzcd} \tag{3.7}$$

commutes. Let us say that L is *nonsingular* if this mapping $\bar{L}$ is surjective. This is equivalent to saying that $\bar{L}$ is not the zero map.

Theorem 3.4.5 *Let $L : S \to R$ be an R linear mapping. Then*

1. *The mapping $L : S \to R$ is surjective if and only if L is nonsingular. (In particular, the trace $Tr_{S/R}$ is nonsingular.)*
2. *If L is nonsingular, then $L(\mathfrak{M}^i) = \mathfrak{m}^i$ for any $i \geq 1$.*
3. *If L is nonsingular, $b \in S$ and $L(ab) = 0$ for all $a \in S$, then $b = 0$.*

4. *There exists $b \in S$ so that $L(a) = Tr(ba)$ for all $a \in S$. The element b is invertible if and only if L is nonsingular.*

Proof: If L is surjective then it is nonsingular by diagram (3.7). On the other hand, if L is nonsingular then (as above) there exists $b \in S$ such that $L(b)$ is invertible in R. If $a = L(b)^{-1}$ then, for any $c \in R$, $L(cab) = c$ so L is surjective. This proves (1). We already know that $L(\mathfrak{M}^i) \subset \mathfrak{m}^i$ so let $c \in \mathfrak{m}^i$ and, by part (1), let $a_0 \in S$ be an element such that $L(a_0) = 1$. Then $ca_0 \in \mathfrak{M}^i$ and $L(ca_0) = c$, which proves (2).

To prove (3), let n be the degree of nilpotency of $\mathfrak{m}$. That is, $\mathfrak{m}^n = 0$ but $\mathfrak{m}^{n-1} \neq 0$. Then n is also the degree of nilpotency of $\mathfrak{M}$. Let $b \neq 0 \in S$ and suppose that $L(ab) = 0$ for all $a \in S$. Let $m < n$ be the largest integer so that $b \in \mathfrak{M}^m$. Then $b = db_1$ with $d \in \mathfrak{m}^m - \mathfrak{m}^{m+1}$ and b_1 a unit in S. Therefore for all $a \in S$ we have $0 = L(da) = dL(a)$. But $m < n$ so we must have $L(a) \in \mathfrak{M}$ which contradicts the nonsingularity of L, proving (3).

To prove (4), consider the mapping $S \to \mathrm{Hom}_R(S, R)$ which assigns to any $b \in S$ the R linear mapping $a \mapsto \mathrm{Tr}_{S/R}(ab)$. This mapping is injective, for if $b' \in S$ and $\mathrm{Tr}_{S/R}(ab) = \mathrm{Tr}_{S/R}(ab')$ for all $a \in S$, then by part (3) this implies $b = b'$. Since S is a free module over R of some rank d, there are $|R|^d$ elements in $\mathrm{Hom}_R(S, R)$. But this is the same as the number of elements in S. Therefore every R-linear mapping $L : S \to R$ is of the form $a \mapsto \mathrm{Tr}_{S/R}(ab)$ for some $b \in S$. If b is invertible, then the mapping L is nonsingular, whereas if $b \in \mathfrak{M}$ then $L(ab) \in \mathfrak{m}$ so the resulting mapping $\bar{L} : S/\mathfrak{M} \to R/\mathfrak{m}$ is zero. □

3.4.c Primitive polynomials

Let R be a finite local ring with maximal ideal $\mathfrak{m}$ and residue field $\mu : R \to F = R/\mathfrak{m}$. Let S be a degree d Galois extension of R, with maximal ideal $\mathfrak{M}$ and residue field $\nu : S \to K = S/\mathfrak{M}$ as in (3.5). Let $f \in R[x]$ be a basic irreducible polynomial of degree d. Then f is said to be *primitive* if the polynomial $\bar{f} = \mu(f) \in F[x]$ is primitive. That is, if for some (and hence for any) root $\bar{a} \in K$ of $\bar{f}$, the distinct powers of $\bar{a}$ exactly account for all the nonzero elements in K. Unfortunately this is not enough to guarantee that each root $a \in S$ of f generates the cyclic group $\iota(K^\times) \subset S$.

Lemma 3.4.6 *Let $f \in R[x]$ be a basic irreducible polynomial of degree d and let S be a degree d Galois extension of R, so that f splits into linear factors over S. Let $a \in S$ be a root of f. If $\mu(f)$ is primitive (in $F[x]$) then the elements $\left\{1, a, a^2, \cdots, a^{Q-2}\right\}$ are distinct, where $Q = |K| = |F|^d$. The roots of f lie in $\iota(K^\times) \subset S^\times$ if and only if f divides $x^Q - 1$. Thus, if $\mu(f)$ is primitive and f divides $x^Q - 1$, then $\iota(K^\times) \subset S^\times$ consists of the $Q - 1$ distinct powers $\left\{1, a, a^2, \cdots, a^{Q-2}\right\}$ of a.*

Proof: The element $\mu(a) \in K$ is a root of $\mu(f) \in F[x]$. If $\mu(f)$ is primitive, then $\mu(a)$ is a primitive element in K and the elements $\mu(a)^i$ $(0 \leq i \leq Q - 2)$ are distinct, so the same is true of the elements a^i $(0 \leq i \leq Q-2)$. By 3.1.4 the polynomial $g(x) = x^{Q-1} - 1$ factors completely in S as

$$g(x) = \prod_{b \in K^\times} (x - \iota(b)).$$

Since f also factors completely over S, we see that the roots of f lie in $\iota(K^\times)$ if and only if f divides $g(x)$. □

3.5 Galois rings

Let $p \in \mathbf{Z}$ be a prime number. According to Theorem 3.4.1, for each $n, d \geq 1$ the ring $\mathbf{Z}/(p^n)$ has a unique Galois extension of degree d. This extension $S = GR(p^n, d)$ is called the *Galois ring* of degree d over $\mathbf{Z}/(p^n)$. For $n = 1$ it is the Galois field $\mathbf{F}_{p^d}$. For $d = 1$ it is the ring $\mathbf{Z}/(p^n)$. Let us review the general facts from §3.4 for the case of a Galois ring S.

The Galois ring $S = GR(p^n, d)$ is isomorphic to the quotient ring $\mathbf{Z}/(p^n)[x]/(f)$ where $f \in \mathbf{Z}/(p^n)[x]$ is a monic basic irreducible polynomial. That is, it is a monic polynomial such that its reduction $f \pmod p \in \mathbf{Z}/(p)[x]$ is irreducible. The ring S contains p^{nd} elements. For each divisor e of d the Galois ring S contains the ring $GR(p^n, e)$ and this accounts for all the subrings of S. For any $m \leq n$ there is a projection $S \to GR(p^m, d)$ whose kernel is the ideal (p^m), and this accounts for all the nontrivial ideals in S. In particular the maximal ideal $\mathfrak{M} = (p) = pS$ consists of all multiples of p. The quotient $S/\mathfrak{M} \cong \mathbf{F}_{p^d}$ is isomorphic to the Galois field with p^d elements. If μ denotes the projection to this quotient, then it is compatible with the trace mapping in the sense that the following diagram commutes,

$$\begin{tikzcd}
S = GR(p^n, d) \arrow[r, "\mu"] \arrow[d, "\mathrm{Tr}"'] & K = \mathbf{F}_q \arrow[d, "\mathrm{Tr}"] \\
\mathbf{Z}/(p^n) \arrow[r, "\mu"'] & \mathbf{F}_p
\end{tikzcd}$$

where $q = p^d$. There is a natural (multiplication-preserving) splitting $\iota : K \to S$ of the mapping μ whose image is the set all elements $x \in S$ such that $x^q = x$. The group of units of S is the product

$$S^\times = \iota(K^\times) \times (1 + \mathfrak{M}).$$

If $p \geq 3$ then

$$1 + \mathfrak{M} \cong \mathbf{Z}/(p^{n-1}) \times \cdots \times \mathbf{Z}/(p^{n-1}) \ \ (d \text{ times}).$$

If $p = 2$ and $n \geq 3$ then

$$1 + \mathfrak{M} \cong \left(\mathbf{Z}/(2^{n-1})\right)^{d-1} \times \mathbf{Z}/(2^{n-2}) \times \mathbf{Z}/(2)$$

If $p = 2$ and $n = 1, 2$ then in this equation, each factor $\mathbf{Z}/(2^m)$ should be dropped whenever $m \leq 0$.

It follows that, in general, $S^\times$ contains cyclic subgroups of order $(p^d - 1)p^{n-1}$ and that $|S^\times| = (p^d - 1)p^{d(n-1)}$.

Lemma 3.5.1 *For any $x \in S$ there are unique elements $a_0, a_1, \cdots, a_{n-1} \in \iota(K)$ such that*

$$x = a_0 + a_1 p + \cdots + a_{n-1}p^{n-1}. \tag{3.8}$$

The coefficients $a_0, a_1, \cdots a_{n-1}$ in (3.8) are called the coordinates *of x, and the expansion (3.8) is called the* p-adic expansion *of x.*

Proof: First note that if $t \in \iota(K)$ and if $1 - t$ is not a unit, then $t = 1$. Next, according to the comments in the first paragraph of this section, $|\mathfrak{M}^i/\mathfrak{M}^{i+1}| = q$ for $1 \leq i \leq n-1$. We claim that every element of $\mathfrak{M}^i/\mathfrak{M}^{i+1}$ has a unique representative of the form ap^i where $a \in \iota(K)$. Certainly $ap^i \in \mathfrak{M}^i$ and there are no more than q such elements, so we need to show these elements are distinct modulo $\mathfrak{M}^{i+1}$. Suppose $ap^i \equiv bp^i \pmod{\mathfrak{M}^{i+1}}$ with $a, b \in \iota(K)$. Then $p^i(1 - ba^{-1}) \in \mathfrak{M}^{i+1}$ from which it follows that $1 - ba^{-1} \in \mathfrak{M}$. But $ba^{-1} \in \iota(K)$ so the above note implies that $a = b$.

It now follows by induction that every $x \in \mathfrak{M}^i$ has a unique expression $x = p^i(a_0 + a_1 p + \cdots + a_{n-i-1}p^{n-i-1})$ with $a_i \in \iota(K)$. The coefficient a_0 is the unique representative of $x \pmod{\mathfrak{M}^{i+1}}$, while the inductive step applies to $x - p^i a_0 \in \mathfrak{M}^{i+1}$. □

The advantage of Lemma 3.5.1 is that multiplication by elements in $\iota(K)$ is described coordinatewise. That is, if $b \in \iota(K)$ and if x is given by 3.8, then $ba_0 + ba_1 p + \cdots + ba_{n-1}p^{n-1}$ is the p-adic expansion of bx. Multiplication by p is given by a "shift" of the coefficients a_i. However addition is described using a generalized "carry" procedure: if $a, b \in \iota(K)$ and if $a + b = c_0 + c_1 p + \cdots + c_{n-1}p^{n-1}$ is the p-adic expansion of $a + b$ then we may think of the coefficient c_0 as the "sum" and the coefficients c_i (for $i \geq 1$) as being higher "carries".

3.6 Exercises

1. Let R be a finite local ring with maximal ideal $\mathfrak{m}$. Show that

   a.the ideal $\mathfrak{m}$ consists precisely of the non-units of R,
   b.for every $a \in R$, at least one of a and $1 + a$ is a unit, and
   c.there is a positive integer n such that $\mathfrak{m}^n = 0$.

2. Let R be a finite local ring with maximal ideal $\mathfrak{m}$ and residue field $F = R/\mathfrak{m}$. Show that $\mathfrak{m}^{i-1}/\mathfrak{m}^i$ naturally admits the structure of a vector space over F.

3. If R is a local ring and $g \in R[x]$ is regular, then use Nakayama's Lemma to show that for every $f \in R[x]$ there exist $q, r \in R[x]$ with $f = gq + r$ and $\deg(r) < \deg(g)$.

4. Show that for $p = 3$ and $m = 3$, the mapping $E : \mathbf{Z}/(3^2) \to \mathbf{Z}/(3^3)$ of §3.2.a is given by

$$E(a) = 1 + 3a + 18a^2 + 18a^3.$$

Chapter 4 Sequences, Power Series and Adic Rings

The central theme of this work is the design and analysis of sequences by identifying them with algebraic structures. The most common example of such a structure is a *generating function.* This is a power series whose coefficients are the elements of the sequence. Generating functions have been used to analyze sequences that arise in probability theory, cryptography, analysis of recurrences, combinatorics, random number generation, algebraic topology, and many other areas. In this chapter we develop an algebraic framework for generalizing generating functions.

4.1 Sequences

In this section we review the basic combinatorial notions concerning sequences. See also §**??**.

4.1.a Periodicity

Let A be a set and let $\mathbf{a} = (a_0, a_1, a_2, \cdots)$ be a sequence of elements $a_i \in A$. If the set A is discrete (meaning that it is finite or countable) then we refer to A as the *alphabet* from which the *symbols* a_i are drawn. The sequence $\mathbf{a}$ is *periodic* if there exists an integer $T > 0$ so that

$$a_i = a_{i+T} \tag{4.1}$$

for all $i = 0, 1, 2, \cdots$. Such a T is called *a period* of the sequence $\mathbf{a}$ and the least such T is called *the period*, or sometimes the *least period* of $\mathbf{a}$. The sequence $\mathbf{a}$ is eventually periodic if there exists $N > 0$ and $T > 0$ so that equation (4.1) holds, for all $i \geq N$. To emphasize the difference, we sometimes refer to a periodic sequence as being *purely periodic* or *strictly periodic.* A *period* (resp. the *least period*) of an eventually periodic sequence refers to a period (resp. least period) of the periodic part of $\mathbf{a}$.

Lemma 4.1.1 *Suppose* $\mathbf{a}$ *is a periodic (or eventually periodic) sequence with least period* T*. Then every period of* $\mathbf{a}$ *is a multiple of* T*.*

Proof: If T' is a period of $\mathbf{a}$, then dividing by T gives $T' = qT + r$ for some quotient $q \geq 1$ and remainder r with $0 \leq r \leq T-1$. Since both T and T' are periods, $a_{i+T'} = a_{i+qT+r} = a_{i+r}$ for all $i \geq 0$. Therefore r is a period also, but $r < T$ which contradicts the minimality of T. Therefore $r = 0$. □

4.1.b Distinct sequences

Let A be an alphabet and let $\mathbf{a} = (a_0, a_1, \cdots)$ and $\mathbf{b} = (b_0, b_1, \cdots)$ be periodic sequences of elements of A with the same period. We say that $\mathbf{b}$ is a *cyclic shift* of $\mathbf{a}$ if there exists $\tau \geq 0$ so that$b_i = a_{i+\tau}$ for all $i \geq 0$. If no such shift τ exists then we say that $\mathbf{a}$ and $\mathbf{b}$ are *cyclically distinct*. We say that $\mathbf{a}$ and $\mathbf{b}$ are *isomorphic* if there exists a (single) permutation $\sigma : A \to A$ so that $b_i = \sigma(a_i)$ for all $i \geq 0$. We say they are *isomorphic up to a shift* if there exists a permutation $\sigma : A \to A$ and a shift τ such that $b_i = \sigma(a_{i+\tau})$ for all $i \geq 0$. If no such pair σ, τ exists then we say that $\mathbf{a}$ and $\mathbf{b}$ are *strongly distinct* sequences, or that they are *non-isomorphic, even after a possible shift*. Similarly if $\mathbf{a} = (a_0, a_1, \cdots)$ is a periodic sequence taken from an alphabet A and if $\mathbf{b} = (b_0, b_1, \cdots)$ is a periodic sequence taken from an alphabet B then we say that $\mathbf{a}$ and $\mathbf{b}$ are *isomorphic up to a shift* if there exists a mapping $\sigma : A \to B$ and a shift τ such that $b_i = \sigma(a_{i+\tau})$ for all $i \geq 0$. If no such σ, τ exists then $\mathbf{a}$ and $\mathbf{b}$ are *strongly distinct*.

4.1.c Sequence generators and models

The sequences described in this book are generated by algebraic methods involving rings. We formalize constructions of this type by defining a *sequence generator*. In the models we encounter, the state space of the sequence generator usually corresponds to a cyclic subgroup of the group of units in a ring.

Definition 4.1.2 *A* sequence generator*, or* discrete state machine with output U *consists of a set* Σ *of* states*, an alphabet* A *of* output values*, a state transition function* $\tau : \Sigma \to \Sigma$ *and an output function* $\mathsf{out} : \Sigma \to A$*.*

Such a generator is depicted as follows:

$$\begin{tikzcd} \Sigma \arrow[loop left, "\tau"] \arrow[r, "\mathsf{out}"] & A \end{tikzcd}$$

The set Σ of states is assumed to be *discrete*, meaning that it is either finite or countably infinite. We also assume the alphabet A of possible output values is discrete. Given an initial state $\mathbf{s} \in \Sigma$, such a sequence generator outputs an infinite sequence

$$U(\mathbf{s}) = \mathsf{out}(\mathbf{s}), \mathsf{out}(\tau(\mathbf{s})), \mathsf{out}(\tau^2(\mathbf{s})), \cdots$$

with elements in A. A state $\mathbf{s} \in \Sigma$ is *aperiodic* if, starting from $\mathbf{s}$, the generator never returns to this state. The state $\mathbf{s}$ is periodic of period L if starting from $\mathbf{s}$, after L steps, the generator returns to the state $\mathbf{s}$, that is, $\tau^L \mathbf{s} = \mathbf{s}$. The *least period* of such a periodic state is the least such $L \geq 1$. A state $\mathbf{s}$ is *eventually periodic* if, starting from $\mathbf{s}$, after a finite number of steps, the generator arrives at a periodic state. If Σ is finite then every state is eventually periodic. We say a set of states is *closed* if it is closed under state change. It is *complete* if it consists of all the periodic states.

If R is a ring, and if $b \in R$ denote by $\kappa_b : R \to R$ the multiplication by b, that is, $\kappa_b(x) = bx$.

Definition 4.1.3 *Let* $U = (\Sigma, A, \tau, \mathsf{out})$ *be a sequence generator. An* algebraic model *or simply a* model *for* U *is a ring* R*, an element* $b \in R$*, a mapping* $\psi : R \to \Sigma$*, and an output mapping* $T : R \to A$ *such that the following diagram commutes:*

$$\begin{tikzcd}
R \arrow[loop above, "\kappa_b"] \arrow[rr, "\psi"] \arrow[dr, "T"'] & & \Sigma \arrow[loop above, "\tau"] \arrow[dl, "\mathsf{out}"] \\
& A &
\end{tikzcd}$$

which means that $\mathsf{out}(\psi(a)) = T(a)$ *and* $\psi(ba) = \tau(\psi(a))$ *for all* $a \in R$*.*

Each $a \in R$ then corresponds to an initial state $\psi(a) \in \Sigma$, and the output sequence generated from this initial state is then described by the following *exponential representation*,

$$T(a), T(ba), T(b^2a), \cdots .$$

If the ring R is finite and if it is an integral domain then every such sequence is strictly periodic (because $b^k a = b^{k+r} a$ implies that $a = b^r a$). We say the model is *complete* if every periodic state $\mathbf{s} \in \Sigma$ may be realized as the image $\mathbf{s} = \psi(a)$ of some element $a \in R$. A complete model, if one exists, allows us to analyze the behavior of the sequence generator using the algebraic structure of the ring R. In this book we will encounter many different types of sequence generators and their models.

In some circumstances it is more convenient to specify a mapping $\phi : \Sigma \to R$ (rather than the other way around) so that the corresponding diagram commutes:

$$\begin{tikzcd}
R \arrow[loop above, "\kappa_b"] \arrow[dr, "T"'] & & \Sigma \arrow[loop above, "\tau"] \arrow[ll, "\phi"'] \arrow[dl, "\mathsf{out}"] \\
& A &
\end{tikzcd}$$

To distinguish between these two types of models, we will sometimes refer to the first one as an *injective* and the second as a *projective model.* If (R, ψ) is a complete injective model, then the inverse mapping $\phi = \psi^{-1}$ is a complete projective model (and vice versa). However it may require a nontrivial amount of computation to describe the inverse mapping, particularly when attempting to describe the initial state of the generator, cf. (**??**), (**??**), (**??**).

4.2 Power Series

4.2.a Definitions

Throughout this section we fix a commutative ring R.

Definition 4.2.1 *A (formal)* power series over *R is an infinite expression*

$$a(x) = \sum_{i=0}^{\infty} a_i x^i,$$

where x is an indeterminate and $a_0, a_1, \cdots \in R$. As with polynomials, the a_is are called coefficients. *The sequence $(a_0, a_1, \cdots)$ of coefficients of a power series $a(x)$ is denoted* $\mathbf{seq}(a)$*. If $b(x) = \sum_{i=0}^{\infty} b_i x^i$ is a second power series over R, then define*

$$(a+b)(x) = a(x) + b(x) = \sum_{i=0}^{\infty} (a_i + b_i) x^i$$

and

$$(ab)(x) = a(x)b(x) = \sum_{i=0}^{\infty} (\sum_{j=0}^{i} a_j b_{i-j}) x^i.$$

The set of power series over R is denoted $R[[x]]$. The order *of a non-zero power series $a(x) = \sum_{i=0}^{\infty} a_i x^i$ is the least index i such that $a_i \neq 0$. The order of 0 is ∞.*

These operations make $R[[x]]$ into a ring with zero given by the power series all of whose coefficients are zero, and with identity (1) given by the power series $1 + 0x + 0x^2 + \cdots$. The set of polynomials over R is the sub-ring of $R[[x]]$, consisting of those power series with finitely many nonzero coefficients. In fact there is a tower of sub-rings,

$$R \subset R[x] \subset E \subset R_0(x) \subset R[[x]] \subset R((x))$$

which we now describe.

Definition 4.2.2 *The ring $R((x))$ of* formal Laurent series *consists of infinite sums*

$$a(x) = a_{-m}x^{-m} + a_{-m+1}x^{-m+1} + \cdots + a_0 + a_1 x + \cdots$$

with coefficients $a_i \in R$ and at most finitely many non-zero terms of negative degree. Addition and multiplication are defined as with power series.

The ring $R((x))$ is obtained from $R[[x]]$ by inverting x, that is, $R((x)) = S^{-1}R[[x]]$ where $S = \{x, x^2, x^3, \cdots\}$.

4.2.b $R_0(x)$

Lemma 4.2.3 *Let $b(x) = \sum_{i=0}^{\infty} b_i x^i \in R[[x]]$ be a power series. Then the following statements are equivalent: (1) b is invertible in $R[[x]]$, (2) the constant term $b_0 \in R$ is invertible in R, and (3) the elements b and x are relatively prime in $R[[x]]$.*

Proof: The proof is straightforward except possibly for (2) $\implies$ (1). If b_0 is invertible then the equation $b(x)c(x) = 1$ may be solved inductively for $c(x) = \sum_{i=0}^{\infty} c_i x^i$ because $c_0 = b_0^{-1}$ and

$$c_i = -b_0^{-1}\left(b_1 c_{i-1} + b_2 c_{i-1} + \cdots + b_i c_0\right). \qquad \square$$

Let $S \subset R[x]$ denote the multiplicative subset consisting of all polynomials $b(x)$ such that the constant term $b_0 = b(0) \in R$ is invertible in R. Then the ring of fractions (§1.2.d)

$$R_0(x) = S^{-1}R[x]$$

consists of all formal symbols $a(x)/b(x)$ with $b(x) \in S$, under the equivalence relation that $a(x)/b(x) \sim a'(x)/b'(x)$ if $a(x)b'(x) = a'(x)b(x)$. We obtain an injective homomorphism $\psi : R_0(x) \to R[[x]]$ by mapping $a(x)/b(x)$ to the product $a(x)c(x)$ where $c(x) \in R[[x]]$ is the power series inverse of $b(x)$ which was constructed in Lemma 4.2.3. Henceforth we identify $R_0(x)$ with its image in $R[[x]]$. If R is a field then every nonzero element is invertible, so $R_0(x)$ consists of all fractions $a(x)/b(x)$ with $b(0) \neq 0$. In this case, $R_0(x)$ is a field; it is usually denoted $R(x)$ and is referred to as the *field of rational functions* over R.

Definition 4.2.4 *Let $a(x), b(x) \in R[x]$ and suppose $b(0)$ is invertible in R. We refer to the power series $\psi(a(x)/b(x)) \in R[[x]]$ as the* power series expansion *of the fraction $a(x)/b(x)$.*

Definition 4.2.5 (See also §**??**.) *A sequence $\mathbf{a} = a_0, a_1, \cdots$ of elements of R is* linearly recurrent *(of degree d) if there exists $q_1, \cdots, q_d \in R$ such that for all $n \geq d$ we have*

$$a_n = q_1 a_{n-1} + \cdots + q_d a_{n-d}. \qquad (4.2)$$

Theorem 4.2.6 *Let $a = a_0 + a_1 x + \cdots \in R[[x]]$ be a formal power series. Then $a \in R_0(x)$ (that is, $a(x)$ is a quotient $f(x)/g(x)$ of two polynomials, where $g(0)$ is invertible in R) if and only if the sequence of coefficients $a_n, a_{n+1}, a_{n+2}, \cdots$ satisfies a linear recurrence, for n sufficiently large.*

Proof: First suppose that $a(x) = f(x)/g(x)$ with $g(x) = g_0 + g_1 x + \cdots + g_d x^d$. Then $f(x) = a(x)g(x)$ which gives $f_n = \sum_{i=0}^{d} g_i a_{n-i}$. Since $f(x)$ is a polynomial, these coefficients vanish for sufficiently large n which leaves

$$a_n = -g_0^{-1}\left(g_1 a_{n-1} + g_2 a_{n-2} + \cdots + g_d a_{n-d}\right)$$

which is a linear recurrence (of degree $d = \deg g$). Conversely, suppose the coefficients of f satisfy a linear recurrence $a_n = g_1 a_{n-1} + \cdots + g_d a_{n-d}$ for all $n \geq N$. Let $g(x) = -1 + g_1 x + \cdots + g_d x^d$ (so $g_0 = -1$.) Then the product $f(x) = g(x)a(x)$ is a polynomial, because for sufficiently large n its term of degree n is

$$\sum_{i=0}^{d} g_i a_{n-i} = 0.$$

Consequently $a(x) = f(x)/g(x)$ and g_0 is invertible. □

4.2.c Eventually periodic power series

Definition 4.2.7 *The ring $E \subset R[[x]]$ is the collection of all power series $a(x) = \sum_{i=0}^{\infty} a_i x^i$ such that the sequence of coefficients $\mathbf{seq}(a) = (a_0, a_1, \cdots)$ is eventually periodic.*

Theorem 4.2.8 *Let $a(x) = \sum_{i=0}^{\infty} a_i x^i$ be a power series over a ring R and let $n \geq 1$. Then the following are equivalent. (See also Lemma 1.4.4.)*

1. *The sequence $\mathbf{seq}(a) = (a_0, a_1, \cdots)$ is eventually periodic and n is a period of $\mathbf{seq}(a)$.*
2. *$a(x) = h(x)/(x^n - 1)$ for some $h(x) \in R[x]$.*
3. *$a(x) = f(x)/g(x)$ for some $f, g \in R[x]$ such that $g(x)$ is monic and $g(x)|(x^n - 1)$.*
4. *$a(x) = f(x)/g(x)$ for some $f, g \in R[x]$ such that $g(x)|(x^n - 1)$.*

These statements imply

5. *$a(x) = f(x)/g(x)$ for some $f, g \in R[x]$ such that $g(0)$ is invertible in R,*

hence $E \subset R_0(x)$. The eventual period is the least n for which (2), (3), or (4) holds. If R is finite then statement (5) implies the others (for some $n \geq 1$.)

The sequence $\mathbf{seq}(a)$ is purely periodic if and only if (2) holds with $\deg(h(x)) < n$ or (3) or (4) holds with $\deg(f(x)) < \deg(g(x))$.

Proof: To see that condition (1) implies condition (2), suppose $a(x)$ is eventually periodic with $a_i = a_{i+n}$ for all $i \geq N$. Then we have

$$\begin{aligned} a(x) &= \sum_{i=0}^{N-1} a_i x^i + x^N \sum_{j=0}^{\infty} (\sum_{k=0}^{n-1} a_{nj+i+N} x^i) x^{nj} \\ &= \sum_{i=0}^{N-1} a_i x^i + \frac{x^N \sum_{k=0}^{n-1} a_{nj+i+N} x^i}{1 - x^n}. \end{aligned}$$

This can be written as a rational function with denominator $x^n - 1$.

Conditions (2), (3) and (4) are clearly equivalent. In case (3) or (4), $\deg(b(x)f(x)) < n$ if and only if $\deg(f(x)) < \deg(g(x))$, which reduces the statements about purely periodic power series to the statement about purely periodic power series in case (2).

To see that condition (2) implies condition (1), suppose $a(x) = h(x)/(x^n - 1)$ with $h(x) \in R[x]$. By the division theorem we can write $h(x) = (x^n - 1)u(x) + v(x)$ with $u(x), v(x) \in R[x]$ and $\deg(v(x)) < n$. Thus

$$\begin{aligned} a(x) &= u(x) + \frac{v(x)}{x^n - 1} \\ &= u(x) + (v(x) + x^n v(x) + x^{2n} v(x) + \cdots). \end{aligned}$$

The power series $v(x) + x^n v(x) + x^{2n} v(x) + \cdots$ is strictly periodic since there is no overlap among the degrees of the monomials in any two terms $x^{in} v(x)$ and $x^{jn} v(x)$. The addition of $u(x)$ only affects finitely many terms, so the result is eventually periodic. Also, the sequence is periodic if and only if $u(x) = 0$, which is equivalent to $\deg(h(x)) < n$.

It follows immediately that the eventual period is the least n for which (2), (3), or (4) holds. Lemma 1.4.4 says that (4) implies (5), and if R is finite, then (5) implies (4) (for some n). □

4.2.d When R is a field

Theorem 4.2.9 *Suppose R is a field. Then the rings E and $R_0(x) = R(x)$ coincide. (In other words, every eventually periodic sequence eventually satisfies a linear recurrence and vice versa.) The only non-trivial ideals in $R[[x]]$ are the principal ideals (x^m) for $m \geq 1$. Moreover, $R(x)$ and $R((x))$ are also fields: they are the fraction fields of $R[x]$ and of $R[[x]]$ respectively.*

Proof: The only nontrivial statement in this theorem concerns the ideal structure of $R[[x]]$. Suppose that I is a non-zero ideal in $F[[x]]$. Let $a(x)$ be an element of I whose order n is as small as possible. Then we have $a(x) = x^n b(x)$ for some $b(x) \in F[[x]]$, and the constant term of $b(x)$ is nonzero. By Lemma 4.2.3, $b(x)$ is invertible in $F[[x]]$. Hence $x^n \in I$. Moreover, every element of I has order at least n, so can be written as $x^n c(x)$ for some $c(x) \in F[[x]]$. Hence $I = (x^n)$. □

4.2.e $R[[x]]$ as an inverse limit

The quotient ring $R[x]/(x^\ell)$ may be identified with the collection of all polynomials of degree $\leq \ell - 1$. Let $\phi_\ell : R[[x]] \to R[x]/(x^\ell)$ be the homomorphism that associates to each $a = \sum_{i=0}^{\infty} a_i x^i$ the partial sum (that is, the polynomial) $\sum_{i=0}^{\ell-1} a_i x^i$. These homomorphisms are compatible in the sense that if $k \leq \ell$ then $T_k^\ell(\phi_\ell(a)) = \phi_k(a)$ where $T_k^\ell : R[x]/(x^\ell) \to R[x]/(x^k)$ is reduction modulo x^k. The next lemma says that every element of $R[[x]]$ can be described in terms of such a sequence of partial sums.

Lemma 4.2.10 *Suppose $s_1, s_2, \cdots$ is a sequence with $s_i \in R[x]/(x^i)$. Assume these elements are* compatible *in the sense that $T_k^\ell(s_\ell) = s_k$ for every pair $k \leq \ell$. Then there is a unique element $a \in R[[x]]$ such that $\phi_i(a) = s_i$ for all $i \geq 1$.*

Proof: The element $a = \sum_{i=0}^{\infty} a_i x^i$ is given by $a_\ell = (\phi_{\ell+1}(a) - \phi_\ell(a))/x^\ell$. □

For the readers who knows about limits, this lemma says that

$$R[[x]] = \varprojlim \{R[x]/(x^i)\}$$

is the inverse limit of the system of rings $R[x]/(x^i)$.

4.3 N-Adic Numbers

4.3.a Definitions

In this section we see a somewhat different way to identify an infinite sequence with an algebraic object. Fix an integer $N \geq 2$.

Definition 4.3.1 *An N-*adic number *is an infinite expression*

$$a = \sum_{i=0}^{\infty} a_i N^i,$$

where $a_0, a_1, \cdots \in \{0, 1, \cdots, N-1\}$. The set of N-adic numbers is denoted by $\mathbf{Z}_N$. The order *of an non-zero N-adic number $a = \sum_{i=0}^{\infty} a_i N^i$ is the least index i such that $a_i \neq 0$. The order of 0 is ∞.*

Again, the a_i are called *coefficients*. When writing N-adic numbers we may omit terms whose coefficient is zero. We may also write the terms in a different order.

So far, N-adic numbers look just like power series. The difference lies in the algebra. Addition and multiplication are defined so as to take into account the "carry" operation. If $b = \sum_{i=0}^{\infty} b_i N^i$ is a second N-adic number, then the sum $a + b$ is the N-adic number $c = \sum_{i=0}^{\infty} c_i N^i$ defined as follows. There exists a unique c_0 $(0 \leq c_0 \leq N-1)$ and $t_0 \geq 0$ so that $a_0 + b_0 = c_0 + Nt_0$ (namely $c_0 = (a_0 + b_0) \pmod{N}$ and $t_0 = (a_0 + b_0) \text{ (div } N)$, where we have identified $\mathbf{Z}/(N)$ with the set $\{0, 1, 2, \cdots, N-1\}$ and where $x \text{ (div } N) = \lfloor x/N \rfloor$. The quantity t_0 is the "carry" at the zeroth stage. Assume by induction that $c_0, c_1, \cdots, c_{n-1}$ and $t_0, t_1, \cdots, t_{n-1}$ have been found with $0 \leq c_i \leq N-1$ and $t_i \geq 0$ and $a_i + b_i + t_{i-1} = c_i + Nt_i$. Then there exist unique c_n, t_n such that $0 \leq c_n \leq N-1$; $t_n \geq 0$, and

$$a_n + b_n + t_{n-1} = c_n + Nt_n,$$

namely $c_n = (a_n + b_n + t_{n-1}) \pmod{N}$ and $t_n = (a_n + b_n + t_{n-1}) \text{ (div } N)$. The product $ab = c$ is defined similarly with

$$\sum_{i=0}^{n} a_i b_{n-i} + t_{n-1} = c_n + Nt_n. \tag{4.3}$$

It is easy to see that these operations make $\mathbf{Z}_N$ into a ring. As with power series, we refer to the sequence $(a_0, a_1, \cdots)$ of coefficients as $\mathbf{seq}_N(a)$. We say that a is periodic (resp. eventually periodic) if the sequence $\mathbf{seq}_N(a)$ of coefficients is periodic (resp. eventually periodic).

If $a = \sum_{i=0}^{\infty} a_i N^i$ is an N-adic number, then the coefficient a_0 is called the *reduction of a modulo N* and it is denoted $a_0 = a \pmod{N}$. This gives a ring homomorphism $\mathbf{Z}_N \to \mathbf{Z}/(N)$. We also refer to $\sum_{i=0}^{\infty} a_{i+1} N^i = (a - a_0)/N$ as the *integral quotient* of a by N, denoted $\text{quo}(a, N)$ or $a \text{ (div } N)$. Thus

$$a = (a \pmod{N}) + N\text{quo}(a, N).$$

In the ring $\mathbf{Z}_N$ we have an identity,

$$-1 = (N-1) + (N-1)N + (N-1)N^2 + \cdots,$$

which can be verified by adding 1 to both sides. There is an explicit formula for multiplication by -1. If $a = N^d(1 + \sum_{i=0}^{\infty} a_i N^i)$ then

$$-a = N^d((N - a_0) + \sum_{i=1}^{\infty} (N - a_i - 1)N^i) \tag{4.4}$$

which may be verified by adding a to both sides of the equation.

4.3.b The ring $\mathbf{Z}_{N,0}$

The nonnegative integers may be identified with the set of N-adic numbers with finitely many nonzero coefficients. Since we have negation, this identification extends to a ring homomorphism $\mathbf{Z} \to \mathbf{Z}_N$. Its kernel is an ideal that does not contain any positive integers, so it must be (0). So this homomorphism is an injection, and we may view the integers as a sub-ring of $\mathbf{Z}_N$. As with the case of power series, there is an intermediate ring,

$$\mathbf{Z} \subset \mathbf{Z}_{N,0} \subset \mathbf{Z}_N$$

which we will now describe.

Lemma 4.3.2 *Let $a \in \mathbf{Z}$. Then the following statements are equivalent: (1) a is relatively prime to N, (2) N is invertible in $\mathbf{Z}/(a)$, (3) a is invertible in $\mathbf{Z}/(N)$, (4) there exists $n \geq 0$ so that $a|(N^n - 1)$.*

Proof: The proof is the same as the proof of Lemma 1.4.4. □

Lemma 4.3.3 *Let $a = \sum_{i=0}^{\infty} a_i N^i \in \mathbf{Z}_N$. Then a is invertible in $\mathbf{Z}_N$ if and only if a_0 is relatively prime to N.*

Proof: The proof is essentially the same as that of Lemma 4.2.3. Suppose a_0 is relatively prime to N. We search for $b = \sum_{i=0}^{\infty}$ so that $ab = 1$, and $0 \leq b_i \leq N-1$. By equation 4.3 this means $a_0 b_0 = 1 + N t_0$ (which has the unique solution $b_0 = a_0^{-1} \pmod{N}$ and $t_0 = a_0 b_0 - 1$ (div N)) and $\sum_{i=0}^{n} a_i b_{n-i} + t_{n-1} = c_n + N t_n$, which has the (unique) solution

$$\begin{aligned} b_n &= a_0^{-1}\left(c_n - t_{n-1} - \sum_{i=1}^{n} a_i b_{n-i}\right) \pmod{N} \\ t_n &= \left(\sum_{i=0}^{n} a_i b_{n-i} - c_n\right) \text{ (div } N). \end{aligned}$$

This completes the proof of Lemma 4.3.3. □

Definition 4.3.4 *Let $\mathbf{Z}_{N,0}$ denote the set of all rational numbers $a/b \in \mathbf{Q}$ (in lowest terms) such that b is relatively prime to N.*

Lemma 4.3.3 says that $\mathbf{Z}_{N,0}$ is naturally contained in the N-adic numbers $\mathbf{Z}_N$. It is easy to see that it forms a sub-ring of $\mathbf{Z}_N$. The next theorem says that this ring of fractions f/g (with g relatively prime to N) is exactly the collection of N-adic numbers $a \in \mathbf{Z}_N$ such that $\mathbf{seq}_N(a)$ is eventually periodic.

Theorem 4.3.5 *Let $a = \sum_{i=0}^{\infty} a_i N^i \in \mathbf{Z}_N$ and let $n \geq 1$. Then the following statements are equivalent.*

1. *$\mathbf{seq}_N(a)$ is eventually periodic and n is a period of a.*
2. *$a = h/(N^n - 1)$ for some $h \in \mathbf{Z}$.*
3. *$a = f/g$ for some $f, g \in \mathbf{Z}$ such that $g|(N^n - 1)$.*

The eventual period is the least n for which (2) or (3) holds. The N-adic number a is purely periodic if and only if $-(N^n - 1) \leq h \leq 0$ in case (2) or $-g \leq f \leq 0$ in case (3).

Proof: To see that condition (1) implies condition (2), suppose $\mathbf{seq}_N(a)$ is eventually periodic with $a_i = a_{i+n}$ for all $i \geq M$. Then we have

$$\begin{aligned} a &= \sum_{i=0}^{M-1} a_i N^i + N^M \sum_{j=0}^{\infty} (\sum_{k=0}^{n-1} a_{nj+i+M} N^i) N^{nj} \\ &= \sum_{i=0}^{M-1} a_i N^i + \frac{N^M \sum_{k=0}^{n-1} a_{nj+i+M} N^i}{1 - N^n}. \end{aligned}$$

This is can be written as a rational number with denominator $N^n - 1$.

Condition (2) trivially implies condition (3). If $a = f/g$ and $g|N^n - 1$, then $bg = N^n - 1$ for some $b \in \mathbf{Z}$, so $a = bf/(N^n - 1)$. Thus condition (3) implies condition (2). In case (3) $-(N^n - 1) \leq bf \leq 0$ if and only if $-g \leq f \leq 0$, which reduces the statements about purely periodic N-adic numbers to the statement about purely periodic N-adic numbers in case (2).

Now suppose case (2) holds, that is, $a = h/(N^n - 1)$ with $h \in \mathbf{Z}$. By the division theorem we can write $h = (N^n - 1)m - k$ with $m, k \in \mathbf{Z}$, and $0 \leq k < N^n - 1$. Thus

$$\begin{aligned} a &= m - \frac{k}{N^n - 1} \\ &= m + (k + N^n k + N^{2n} k + \cdots). \end{aligned} \tag{4.5}$$

The N-adic number $k + N^n k + N^{2n} k + \cdots = \sum_{i=0}^{\infty} k N^{in}$ is strictly periodic since there is no overlap among the exponents in two terms $N^{in} k$ and $N^{jn} k$. In particular we see that the N-adic expansion of $a = h/(N^n - 1)$ is strictly periodic if and only if (a) $m = 0$ or (b) $k = 0$ and $m = 0, -1$. These conditions are equivalent to the statement that $-(N^n - 1) \leq h \leq 0$.

Now let us prove that (2) $\Longrightarrow$ (1). There are several cases to consider.

First, suppose m is positive. Then the addition of m affects only finitely many terms in equation (4.5) because the carries eventually reach a place where $k \leq N^n - 2$, and there is no carry beyond this place . Thus in this case the result is eventually periodic.

Suppose m is negative and k is not zero. Then there is an integer j, such that $-m < \sum_{i=0}^{j} kN^{in}$. Thus

$$a = (\sum_{i=0}^{j} kN^{in} + m) + \sum_{i=j+1}^{\infty} kN^{in}.$$

The first expression on the right hand side has an N-adic expansion with terms only up to degree jn, so $\mathbf{seq}_N(a)$ is eventually periodic.

If m is negative and $k = 0$, then a is a negative. It follows from equation (4.4) that the N-adic expansion of a negative integer eventually becomes $(N-1)N^i + (N-1)N^{i+1} + \cdots$, which is periodic. This completes the proof that (2) $\implies$ (1). It follows immediately that the eventual period is the least n for which (2) or (3) holds. □

Corollary 4.3.6 *If* $\gcd(f, g) = 1$*, then the period of the* N*-adic expansion* $\mathbf{seq}_N(f/g)$ *is the multiplicative order of* N *modulo* g*.*

4.3.c Structure of $\mathbf{Z}_N$

Let $\phi_\ell : \mathbf{Z}_N \to \mathbf{Z}/(N^\ell)$ be the homomorphism that associates to each $a = \sum_{i=0}^{\infty} a_i N^i$ the partial sum $\sum_{i=0}^{\ell-1} a_i N^i$. These homomorphisms are compatible in the sense that if $k \le \ell$ then ${}_NT_k^\ell(\phi_\ell(a)) = \phi_k(a)$ where ${}_NT_k^\ell : \mathbf{Z}/(N^\ell) \to \mathbf{Z}/(N^k)$ is reduction modulo N^k. The next lemma says that every N-adic number can be described as such a sequence of partial sums. It is an exact parallel of Lemma 4.2.10.

Lemma 4.3.7 *Suppose* $s_1, s_2, \cdots$ *is a sequence with* $s_i \in \mathbf{Z}/(N^i)$*. Assume these elements are* compatible *in the sense that* ${}_NT_k^\ell(s_\ell) = s_k$ *for every pair* $k \le \ell$*. Then there is a unique* N*-adic number* $a \in \mathbf{Z}_N$ *such that* $\phi_i(a) = s_i$ *for all* $i \ge 1$*.*

Proof: The desired number $a = \sum_{i=0}^{\infty} a_i N^i$ is given by $a_\ell = (\phi_{\ell+1}(a) - \phi_\ell(a))/N^\ell$. □

(This lemma says that $\mathbf{Z}_N$ is the inverse limit $\varprojlim\{\mathbf{Z}/(N^i)\}$ of the system of rings $\mathbf{Z}/(N^i)$.)

Theorem 4.3.8 *Let* $N = \prod_{i=1}^{k} p_i^{n_i}$ *be the prime factorization of* N *(where the* p_i *are distinct primes and* $n_i \ge 1$*). Then the ring* $\mathbf{Z}_N$ *is isomorphic to the product of rings,* $\prod_{i=1}^{k} \mathbf{Z}_{p_i}$*.*

Proof: First suppose that $N = AB$ with A, B relatively prime. We construct an isomorphism $\psi : \mathbf{Z}_N \cong \mathbf{Z}_A \times \mathbf{Z}_B$ as follows. For each $\ell \ge 1$ there is a mapping

$$H_\ell : \mathbf{Z}/(N^\ell) \to \mathbf{Z}/(A^\ell) \times \mathbf{Z}/(B^\ell)$$

given by $s \mapsto (s \pmod{A})^\ell, s \pmod{B})^\ell)$. Let $a = \sum_{i=0}^{\infty} a_i N^i \in \mathbf{Z}_N$. The sequence of partial sums $s_i = \phi_i(a) \in \mathbf{Z}/(N^i)$ therefore correspond to pairs of elements

$$H_i(s_i) = (x_i, y_i) = (s_i \pmod{A})^i, s_i \pmod{B})^i) \in \mathbf{Z}/(A^i) \times \mathbf{Z}/(B^i),$$

and it is easy to see that these elements are compatible in the sense that ${}_AT_k^\ell(x_\ell) = x_k$ and ${}_BT_k^\ell(y_\ell) = y_k$ for any $k \le \ell$. It follows from Lemma 4.3.7 that the pairs (x_i, y_i) determine a unique element $\psi(a) \in \mathbf{Z}_A \times \mathbf{Z}_B$. This mapping ψ is both injective and surjective because, according to Theorem 1.2.14 (Chinese Remainder Theorem) each of the mappings H_ℓ is an isomorphism.

By repeatedly applying the isomorphism ψ corresponding to different prime factors of N we obtain an isomorphism or rings $\mathbf{Z}_N \cong \prod_{i=1}^{k} \mathbf{Z}_{p_i^{n_i}}$. This reduces the theorem to the case where $N = p^n$ for some prime p. However it is easy to see why $\mathbf{Z}_{p^n}$ is the same as $\mathbf{Z}_p$: Given any element

$$a = a_0 + a_1 p^n + a_2 p^{2n} + \cdots \in \mathbf{Z}_{p^n} \tag{4.6}$$

(with $0 \le a_i \le p^n - 1$) just expand each $a_i = a_{i,0} + a_{i,1}p + \cdots + a_{i,n-1}p^{n-1}$ and substitute this into equation (4.6) to obtain an element of $\mathbf{Z}_p$. The inverse mapping $\mathbf{Z}_p \to \mathbf{Z}_{p^n}$ is obtained by grouping the terms of a p-adic number, n at a time. □

There are many irrational algebraic numbers in $\mathbf{Z}_N$. For example, suppose that $u(x)$ is a polynomial with integer coefficients that has a root modulo N. Then $u(x)$ has a root in $\mathbf{Z}_N$. This is proved in the next section using Hensel's Lemma.

If p is a prime number then the ring $\mathbf{Z}_p$ is an integral domain so its ring of fractions (denoted $\mathbf{Q}_p$) (cf. §1.2.e) is a field. It is called the field of p-adic numbers. Elements of $\mathbf{Q}_p$ can be expressed as fractions $p^{-r}a$ (where $a \in \mathbf{Z}_p$ and r is an integer) or, alternatively, as formal Laurent series $\sum_{i=-r}^{\infty} a_i p^i$. Because $\mathbf{Q}_p$ is a field, and because of Theorem 4.3.8, the ring $\mathbf{Z}_N$ (for composite N) is seldom encountered in the mathematical literature. However we will make use of it when studying sequences generated by an FCSR in Chapter **??**.

4.4 π-Adic Numbers

In this section put the constructions from Subsections 4.2 and 4.3 into a larger context that enables us to build very general algebraic sequence generators. Let R be an integral domain with field of fractions F. Let $\pi \in R$.

Definition 4.4.1 *A* pre-π-adic number over R *is an infinite expression*

$$a = \sum_{i=0}^{\infty} a_i \pi^i,$$

with $a_0, a_1, \cdots \in R$*.*

Again, the a_is are called *coefficients* and the sequence $(a_0, a_1, \cdots)$ is referred to as $\mathbf{seq}_N(a)$. When writing π-adic numbers we may omit terms whose coefficient is zero. We may also write the terms in a different order.

If $b = \sum_{i=0}^{\infty} b_i\pi^i$ is a second pre-π-adic number, then we let $a + b = \sum_{i=0}^{\infty}(a_i + b_i)\pi^i$, $-a = \sum_{i=0}^{\infty} -a_i\pi^i$, and $ab = \sum_{i=0}^{\infty} \sum_{j=0}^{i}(a_j b_{i-j})\pi^i$. It is straightforward to see that these operations make the set R'_π of pre-π-adic numbers into a ring whose zero is the element all of whose coefficients are 0, and with identity the element with 0th coefficient 1 and all remaining coefficients zero (it is really nothing more than the ring of power series over R).

Let

$$I = \left\{ \sum_{i=0}^{\infty} a_i\pi^i : \forall n : \pi^n | \sum_{i=0}^{n-1} a_i\pi^i \right\}.$$

Then I is closed under addition and if $a \in I$ and b is arbitrary, then we have

$$\sum_{i=0}^{n-1}(ab)_i\pi^i \equiv \left(\sum_{i=0}^{n-1} a_i\pi^i\right)\left(\sum_{i=0}^{n-1} b_i\pi^i\right) \pmod{\pi^n},$$

so $ab \in I$. Thus I is an ideal.

Definition 4.4.2 *The ring of* π-adic numbers over R *is the ring of pre-π-adic numbers modulo the ideal I. This ring is denoted by R_π.*

If the context is clear we may simply refer to a π-adic number. There is a homomorphism from R to R_π – map an element a to the π-adic number $\sum_{i=0}^{\infty} a_i\pi^i$ with $a_0 = a$ and $a_i = 0$ for $i \geq 1$. The kernel of this homomorphism is the set of $a \in R$ such that $\pi^n | a$ for all n. Thus this homomorphism is injective if and only if

$$\bigcap_{i=0}^{\infty}(\pi^i) = (0). \tag{4.7}$$

In studying sequences we are generally only interested in rings that satisfy equation (4.7) since we could replace R by $R/\cap_{i=0}^{\infty}(\pi^i)$.

The element π of course generates an ideal in R_π, and $R_\pi/(\pi^n) \cong R/(\pi^n)$. To see this consider the homomorphism from R to R_π. This induces an injection from $R/(\pi^n)$ to $R_\pi/(\pi^n)$. Any $a = \sum_{i=0}^{\infty} a_i\pi^i \in R_\pi/(\pi^n)$ is the image of $\sum_{i=0}^{n-1} a_i\pi^i \in R/(\pi^n)$, so it is also a surjection, hence an isomorphism.

There are ways to represent π-adic numbers that are sometimes more convenient. By a *complete set of representatives for R modulo π* we mean a set S such that for all $a \in R$ there is a unique $b \in S$ so that $a \equiv b \pmod{\pi}$.

Theorem 4.4.3 *Let S be a complete set of representatives for R modulo π. Then for every π-adic number there is a unique representative all of whose coefficients are in S.*

Proof: Let $a = \sum_{i=0}^{\infty} a_i \pi^i \in R_\pi$. We need to construct a sequence $b_0, b_1, \cdots \in S$ so that for all n,

$$\pi^n | \sum_{i=0}^{n-1} (a_i - b_i)\pi^i. \tag{4.8}$$

Let $b_0 \in S$ be the unique element so that $a_0 \equiv b_0 \pmod{\pi}$. Inductively assume that we have found $b_0, \cdots, b_n$ so that equation (4.8) holds. Then $\sum_{i=0}^{n-1}(a_i - b_i)\pi^i = \pi^n c$ for some $c \in R$. Let $b_n \in S$ be congruent to $a_n + c$ modulo π, so there is a $d \in R$ such that $a_n + c = b_n + \pi d$. Then

$$\begin{aligned} \sum_{i=0}^{n}(a_i - b_i)\pi^i &= (a_n - b_n)\pi^n + \sum_{i=0}^{n-1}(a_i - b_i)\pi^i \\ &= (a_n - b_n)\pi^n + c\pi^n \\ &= d\pi^{n+1}. \end{aligned}$$

This proves the existence part of the theorem.

Suppose $c_0, c_1, \cdots \in S$ is a second set of coefficients such that $\pi^n | \sum_{i=0}^{n-1}(a_i - c_i)\pi^i$ for all n. Then also $\pi^n | \sum_{i=0}^{n-1}(b_i - c_i)\pi^i$ for all n. Then $\pi | (b_0 - c_0)$ which implies $b_0 = c_0$. Inductively suppose that $b_i = c_i$ for $i < n$. Then $\pi^{n+1} | (b_n - c_n)\pi^n$. But R is an integral domain, so $\pi | (b_n - c_n)$, so $b_n = c_n$. □

From this theorem it is apparent that the power series ring $R[[x]]$ over a ring R is the ring of x-adic numbers over the polynomial ring $R[x]$. Also, the ring of N-adic numbers (in the terminology of the preceding section) is the ring of N-adic numbers over **Z** (in the terminology of this section).

Relative to a fixed complete set of representatives S for R modulo π, there is a well defined notion of the reduction of an element of R_π modulo π in R, and of the integral quotient of an element of R_π by π. If

$$a = \sum_{i=0}^{\infty} a_i \pi^i,$$

is a π-adic number with $a_0, a_1 \cdots \in S$, then the *reduction of a modulo π* is a_0 and the *integral quotient* of a by π

$$\mathrm{quo}_S(a, \pi) = \sum_{i=0}^{\infty} a_{i+1}\pi^i.$$

If the set S is clear we simply write $\mathrm{quo}(a,\pi)$. Thus in general

$$a = (a \pmod{\pi}) + \pi \mathrm{quo}(a,\pi).$$

Note that if $a \in R$, then $\mathrm{quo}(a,\pi) \in R$.

4.5 Examples

4.6 Alternate Definitions

In this section we describe several other ways to define the π-adic numbers over a ring R.

4.6.a Inverse Limits

The ring R_π can be defined using inverse limits. The set of rings $\{R^i = R/(\pi^{i+1}) : 0 \leq i < \infty\}$ is a directed system with the reduction functions $\tau_i : R^i \to R^{i-1}$. We also have $\psi_i : R_\pi \to R^i$ by reduction modulo π^{i-1}, and $\psi_{i-1} = \tau_i \circ \psi_i$. Thus there is a homomorphism ψ from R_π to $\varprojlim\{R^i\}$ so that if $\varphi_i : \varprojlim\{R^i\} \to R^i$ is the projection function, then $\psi_i = \varphi_i \circ \psi$. We claim that ψ is an isomorphism. If $a = \sum_{i=0}^{\infty} a_i \pi^i \in R_\pi$ is nonzero, then π^n does not divide $\sum_{i=0}^{n-1} a_i \pi^i$ for some n. Thus $\psi_n(a) \neq 0$, and therefore $\psi(a) \neq 0$. This implies ψ is injective. Let $b = (b_0, b_1, \cdots) \in \varprojlim\{R/(\pi^i)\}$. For each i let $c_i \in R$ reduce to b_i modulo π^i. Thus $\pi^i | c_i - c_{i-1}$. Let $a_i = (c_i - c_{i-1})/\pi^i$. Then $a = \sum_{i=0}^{\infty} a_i \pi^i$ reduces to b_i modulo π^{i+1} for every i. That is, $\psi(a) = b$ and ψ is a surjection and an isomorphism.

4.6.b Valuations

In some cases we can also describe R_π in terms of *discrete valuations*. This notion is central to much of algebraic geometry where valuations are used to explain the local structure of an algebraic variety.

Definition 4.6.1 *Let F be a field. A* valuation on F *is a function $\nu : F \to \mathbf{Z} \cup \{\infty\}$ such that for all $a, b \in F$*

1. $\nu(a+b) \geq \min(\nu(a), \nu(b))$.
2. $\nu(ab) = \nu(a) + \nu(b)$.
3. $\nu(a) = \infty$ *if and only if* $a = 0$.

If F is the field of fractions of a ring R, then to define a valuation on F it is sufficient to define it on R since it will then extend to F by $\nu(a/b) = \nu(a) - \nu(b)$.

It follows from the second axiom that $\nu(1) = 0$. It then also follows that if a has finite order (i.e., $a^k = 1$ for some k), then $\nu(a) = 0$. Also, $\nu(a^{-1}) = -\nu(a)$ for every $a \in F$. Thus if ν is a valuation on a field F, then the set $R_\nu = \{a : \nu(a) \geq 0\}$ is a ring, and the set $I_\nu = \{a : \nu(a) > 0\}$ is a maximal ideal in R_ν. We denote the residue field R_ν/I_ν by K_ν. An element a in R_ν is a unit if and only if $\nu(a) = 0$.

Examples:

1. Let p be a prime integer. If a is a nonnegative integer then we have $a = p^n b$ for some nonnegative integer n and integer b that is relatively prime to p. If we define $\nu_p(a) = n$, then ν_p is a valuation on $\mathbf{Z}$.

2. More generally, let R be a UFD and let $\pi \in R$ be prime. If $a \in R$, then $a = \pi^n b$ for some nonnegative integer n and some $b \in R$ such that the gcd of π and b is 1. If we define $\nu_\pi(a) = n$, then ν_π is a valuation on R.

It is not in general the case that $R_\nu = R$. For instance, in the first example $R_\nu = \{a/b : a, b \in \mathbf{Z}, \gcd(b, p) = 1\}$, a ring we encountered when we studied N-adic numbers (with $N = p$). Notice that it is essential to take p or π prime. For suppose $\pi = ab$ and a and b are not units, and suppose that the function ν as defined in example 2 is a valuation. Then $\nu(a) = \nu(b) = 0$, so a and b are units in R_ν. But it follows that $\pi = ab$ is also a unit, which is impossible since it is a generator of the maximal ideal I_ν.

We observe that in any valued field F with valuation ν, there is an element π of I_ν whose valuation is minimal, say $\nu(\pi) = c$. If $x \in R_\nu$, then $\nu(x)$ is a multiple of c, for otherwise we would have $\nu(x) = ac + d$ with $0 < d < c$ and $\nu(x/\pi^a) = d$. Thus $\nu(x) = ac$ for some a, and $\nu(x/\pi^a) = 0$, so x/π^a is a unit. Furthermore, π is prime in R_ν. For if $\pi = uv$ with neither u nor v a unit, then $\nu(\pi) = \nu(u) + \nu(v) \geq 2c$, a contradiction. Thus the second example is completely general. It follows also that in any valued field F, every element is of the form $\pi^a x$ with $\nu(x) = 0$.

Let us recall the definition of a metric space.

Definition 4.6.2 *A* metric space *is a set X with a function $\delta : X \times X \to \mathbf{R}$ (called a* metric function*) such that for all $a, b, c \in X$*

1. *$\delta(a, b) = 0$ if and only if $a = b$.*
2. *$\delta(a, b) \leq \delta(a, c) + \delta(c, b)$ (triangle inequality).*

A sequence of points $x_1, x_2, \cdots$ in a metric space X is a *Cauchy sequence* if for every $\epsilon > 0$ there exists a k so that $x_i - x_j < \epsilon$ if $i, j \geq k$. A sequence $x_1, x_2, \cdots$ *converges* if there is some element $z \in X$ such that for every $\epsilon > 0$ there exists a k so that $x_i - z < \epsilon$ if $i \geq k$. A metric space is *complete* if every Cauchy sequence converges.

Theorem 4.6.3 *Let ν be a valuation on a field F and let $q > 1$ be a positive real number. Then $\delta(a, b) = q^{-\nu(a-b)}$ is a metric function on F.*

Proof: For any $a, b \in F$ we have $0 = \delta(a, b) = q^{-\nu(a-b)}$ if and only if $\nu(a - b) = \infty$, which holds if and only if $a = b$. For any $a, b, c \in F$, we have

$$\begin{aligned}
\delta(a,b) &= q^{-\nu(a-b)} \\
&= q^{-\nu(a-c+c-b)} \\
&\le q^{-\min(\nu(a-c),\nu(c-b))} \\
&= \max(q^{-\nu(a-c)}, q^{-\nu(ac-b)}) \\
&\le q^{-\nu(a-c)} + q^{-\nu(ac-b)} \\
&= \delta(a,c) + \delta(c,b).
\end{aligned}$$

□

The particular choice of $q > 1$ does not matter in what follows. To say that $x = x_1, x_2, \cdots$ is a Cauchy sequence in F amounts to saying that for all $n \in \mathbf{Z}$ there is a k such that $\nu(x_i - x_j) > n$ if $i, j \ge k$.

Theorem 4.6.4 *Let F be a field with a discrete valuation ν. There is a field $\hat{F}$ containing F such that the following hold.*

1. *ν extends to $\hat{\nu}$ on $\hat{F}$.*
2. *$\hat{F}$ is complete with respect to $\hat{\nu}$.*
3. *Suppose E is a field, μ is a valuation on E, E is complete with respect to μ, and there is a homomorphism $\varphi : F \to E$ such that $\mu(\varphi(a)) = \nu(a)$ for all $a \in F$. Then there is a homomorphism $\hat{\varphi} : \hat{F} \to E$ such that $\hat{\varphi}(a) = \varphi(a)$ for all $a \in F$ and $\mu(\hat{\varphi}(a)) = \hat{\nu}(a)$ for all $a \in \hat{F}$.*

In this case $\hat{F}$ is unique in the sense that any other ring satisfying (1), (2), and (3) is isomorphic to $\hat{F}$, and we say that $\hat{F}$ is the completion of F with respect to ν.

Proof: Let $R = R_\nu$ and let π have minimal valuation c in I_ν. Let T be the set of all Cauchy sequences in F. Then T is a subring of the product of infinitely many copies of R. The set of Cauchy sequences with limit 0 is an ideal I in T, and we let $\hat{R} = T/I$. Thus two sequence $\mathbf{x} = x_1, x_2, \cdots$ and $\mathbf{y} = y_1, y_2, \cdots$ are equivalent if for every $\epsilon > 0$ there is a k such that $\delta(x_i, y_i) < \epsilon$ if $i > k$. The ring R embeds in $\hat{R}$ as the set of constant sequences, and so F embeds in the field of fractions $\hat{F}$ of $\hat{R}$.

We extend ν to $\hat{R}$ as follows. Let $\mathbf{x} = x_1, x_2, \cdots \in \hat{R}$. If $\nu(x_i)$ tends to infinity as i tends to infinity, we let $\nu(\mathbf{x}) = \infty$. Otherwise there is some $n \in \mathbf{Z}$ such that $\nu(x_i) \le n$

for infinitely many i. But $\nu(x_i) \geq 0$, so we may assume that $\nu(x_i) = n$ for infinitely many i. Let k be large enough that $\nu(x_i - x_j) \geq n + 1$ if $i, j \geq k$, and suppose that $i \geq k$ with $\nu(x_i) = n$. Let $j \geq k$. Then $\nu(x_j) = \nu(x_i + (x_j - x_i)) \geq \min(\nu(x_i), \nu(x_j - x_i)) = n$. Also, $\nu(x_i) = \nu(x_j + (x_i - x_j)) \geq \min(\nu(x_j), \nu(x_i - x_j))$. That is, $n \geq \min(\nu(x_j), n + 1)$. Thus $n \geq \nu(x_j)$, so $\nu(x_j) = n$. Note in particular that for every Cauchy sequence $\mathbf{x}$, the limit of $\nu(x_i)$ exists. We let $\hat{\nu}(\mathbf{x}) = n$. It is straightforward to verify that $\hat{\nu}$ is a valuation and agrees with ν on F.

To see that $\hat{F}$ is complete, let $\hat{\mathbf{z}} = \mathbf{z}_1, \mathbf{z}_2, \cdots$ be a Cauchy sequence in $\hat{F}$. We have $\mathbf{z}_i = \pi^{a_i}\mathbf{x}_i$ with $\hat{\nu}(\mathbf{x}_i) = 0$. As we have seen, the a_is either have infinite limit or are constant after some point. In the former case the sequence $\hat{\mathbf{z}}$ converges to 0. In the latter case $\hat{\mathbf{x}} = \mathbf{x}_1, \mathbf{x}_2, \cdots$ is a Cauchy sequence and it suffices to show that $\hat{\mathbf{x}}$ has a limit $\mathbf{y}$. If the limit of the a_is is a, then the limit of $\hat{\mathbf{z}}$ is $\pi^a\mathbf{y}$.

So let $\hat{\mathbf{x}} = \mathbf{x}_1, \mathbf{x}_2, \cdots$ be a Cauchy sequence in $\hat{R}$, with $\mathbf{x}_i = x_{i1}, x_{i2}, \cdots$, and $x_{ij} \in R$. For each n there is a k_n so that $\hat{\nu}(\mathbf{x}_i - \mathbf{x}_j) > n$ if $i, j \geq k_n$. We may assume $k_n \leq k_{n+1}$ for all n. Since the sequence $\mathbf{x}_{k_n}$ is a Cauchy sequence, we also have k'_n so that $\nu(x_{k_n i} - x_{k_n j}) > n$ if $i, j \geq k'_n$. Furthermore, by the definition of $\hat{\nu}$ and the fact that there are finitely many $m < k_n$, we can choose k'_n so also $\nu(x_{mi} - x_{k_n i}) > n$ if $i \geq k'_n$ and $m \leq n$ and we may assume $k'_n \leq k'_{n+1}$ for all n. Let $y_n = x_{k_n k'_n}$. Then we claim that $\mathbf{y} = y_1, y_2, \cdots$ is a Cauchy sequence in R and $\mathbf{y}$ is the limit of $\hat{\mathbf{x}}$ in $\hat{R}$.

Suppose $n \leq m$. Then

$$\begin{aligned}
\nu(y_n - y_m) &= \nu(x_{k_n k'_n} - x_{k_m k'_m}) \\
&= \nu(x_{k_n k'_n} - x_{k_n k'_m} + x_{k_n k'_m} - x_{k_m k'_m}) \\
&\geq \min(\nu(x_{k_n k'_n} - x_{k_n k'_m}), \nu(x_{k_n k'_m} - x_{k_m k'_m})) \\
&\geq n.
\end{aligned}$$

Thus $\mathbf{y}$ is a Cauchy sequence.

Now let $i > k_n$. Then

$$\begin{aligned}
\hat{\nu}(\mathbf{x}_i - \mathbf{y}) &\geq \min(\hat{\nu}(\mathbf{x}_i - \mathbf{x}_{k_n}), \hat{\nu}(\mathbf{x}_{k_n} - \mathbf{y})) \\
&\geq \min(n + 1, \hat{\nu}(\mathbf{x}_{k_n} - \mathbf{y})) \\
&= \min(n + 1, \lim_{m \to \infty} \nu(x_{k_n m} - x_{k_n k'_n})) \\
&\geq \min(n + 1, \lim_{m \to \infty} \min(\nu(x_{k_n m} - x_{k_n k'_m}), \nu(x_{k_n k'_m} - x_{k_n k'_n}))) \\
&> n.
\end{aligned}$$

Thus $\hat{F}$ is complete.

Finally, suppose E, μ, and φ are as in the hypotheses. It suffices to extend φ to $\hat{\varphi}$ on F. Let $\mathbf{x} = x_1, x_2, \cdots$ be a Cauchy sequence in F. Then the sequence $\mathbf{w} = \varphi(x_1), \varphi(x_2), \cdots$

is a Cauchy sequence in E. By the completeness of E, $\mathbf{w}$ has a limit z in E. We define $\hat{\varphi}(x) = z$. It is straightforward to check that this function has the properties desired, and that it is in fact unique. The uniqueness of $\hat{F}$ also follows. □

Theorem 4.6.5 *Let R be a UFD with field of fractions F and a discrete valuation ν_π as in example 2 of §4.6.b. Then the field of fractions F_π of R_π is the completion of F with respect to ν_π.*

Proof: We extend ν_π to F_π as follows. Let S be a complete set of representatives in R modulo π. Let $a = \sum_{i=m}^{\infty} a_i\pi^i$, with $a_m \not\equiv 0 \pmod{\pi}$. Then $\nu_\pi(a) = m\nu_\pi(\pi)$. Alternatively, if $a \in R_\pi$ and we think of R_π as the inverse limit of $\{R/(\pi^{i+1})\}$, then $\nu_\pi(a)$ is $\nu_\pi(\pi)$ times the least i such that the image of a in $R/(\pi^{i+1})$ is nonzero. We leave it to the reader to verify that this is a valuation.

Let $\mathbf{x} = x_1, x_2, \cdots \in F_\pi$ be a Cauchy sequence. As in the proof of Theorem 4.6.4 in showing that $\mathbf{x}$ converges we may assume that each x_i is in R_π. Let $x_i = \sum_{i=0}^{\infty} x_{ij}\pi^j$ with $x_{ij} \in S$. Then $\nu_\pi(x_i - x_{i'})$ is the least j such that $x_{ij} \neq x_{i'j}$. Thus for every n there is a k_n so that $x_{ij} = x_{i'j}$ for all $i, i' \geq k_n$ and all $j \leq n$. Let $a_n = x_{k_n n}$. Then $a = \sum_{i=0}^{\infty} a_i\pi^i$ is the limit of $\mathbf{x}$. Therefore F_π is complete.

Finally we show that F_π is the completion of F. If E is the completion of F and $a = \sum_{i=0}^{\infty} a_i\pi^i$, we map a to the limit of the Cauchy sequence $\sum_{i=0}^{1} a_i\pi^i, \sum_{i=0}^{2} a_i\pi^i, \cdots$ in E. It is straightforward to check that this is a homomorphism that preserves valuations, and by uniqueness considerations it is an isomorphism. □

In order to choose representations that result in efficient implementations, we need an old result, known as Hensel's lemma, which allows factorizations of polynomials to be lifted from the residue field. Let F be a complete valued field with valuation ν. If $f(x)$ is a polynomial over R_ν, we denote by $\bar{f}(x)$ the reduction of $f(x)$ modulo I_ν.

Lemma 4.6.6 *(Hensel) Suppose $f(x)$ is a monic polynomial over R_ν, and $\bar{f}(x) = g_0(x)h_0(x)$ in $K_\nu[x]$, where $g_0(x)$ and $h_0(x)$ are monic and relatively prime. Then there exist monic polynomials $g(x)$ and $h(x)$ in $R_\nu[x]$ such that $f(x) = g(x)h(x)$, $\bar{g}(x) = g_0(x)$, and $\bar{h}(x) = h_0(x)$.*

Corollary 4.6.7 *With the same hypotheses, if $\bar{f}(x)$ has a simple root a_0, then $f(x)$ has a simple root a such that $a \pmod{I_\nu} = a_0$.*

Proofs of Hensel's Lemma and the corollary can be found in Jacobson's book [10, pp. 573-4].

4.6.c Adic Topology

The approach using valuations only works when R is a UFD and π is prime. More generally, we can use a topological approach. Recall that a *topology* on a set X is a collection $\mathcal{T}$ of subsets of X (called the *open sets*) such that $\emptyset$ and T are open sets; the intersection of any finite collection of open sets is an open set; and the union of an arbitrary collection of open sets is an open set. A set with a topology is called a *topological space.* The functions that are of interest in topology are those functions $\varphi : X \to Y$ from a topological set X to a topological set Y such that whenever U is open in Y, $\varphi^{-1}(U)$ is open in X. One way to specify a topology on a set X is to specify a *base.* That is, a subset $\mathcal{B} \subseteq \mathcal{T}$ such that every open set is a union of sets in $\mathcal{B}$. For example, the set of open intervals is a base for the standard topology on $\mathbf{R}$.

If R is a ring and I is an ideal, then we can construct a topology by taking $\mathcal{B} = \{x+I^n : x \in R, n \geq 1\}$ as a base. To see that the set of unions of sets in $\mathcal{B}$ is a topology, it suffices to see that the intersection of any two sets in $\mathcal{B}$ is again in $\mathcal{B}$. Indeed, suppose $(x+I^n)\cap(y+I^k)$ is nonempty, and without loss of generality assume $n \leq k$. Then there are elements $u \in I^n$ and $v \in I^k$ such that $x+u = y+v$. It follows that $x-y \in I^n$, so $y+I^k \subseteq x+I^n$. Thus $(x+I^n)\cap(y+I^k) = x+I^n \in \mathcal{B}$. This topology is called the *I-adic topology on R.* It is left as an exercise to show that addition, multiplication, and negation are continuous functions in the I-adic topology. Thus R is a *topological ring.* If $I = (\pi)$ is principal, we refer to the (π)-adic topology as the π-adic topology.

A sequence of points $x_1, x_2, \cdots$ in a ring R is a *Cauchy sequence* (with respect to the I-adic topology) if for every n there exists a k so that $x_i - x_j \in I^n$ if $i, j \geq k$. We are to think of the elements of I^n for large n as small, so $x_i - x_j \in I^n$ means that x_i and x_j are close to each other. The definition of Cauchy sequence says that beyond some place in the sequence, all pairs are arbitrarily close. A sequence $x_1, x_2, \cdots$ *converges* if there is some element z such that for every n there exists a k so that $x_i - z \in I^n$ if $i \geq k$. A topological ring is *complete* if every Cauchy sequence converges.

Finally, R is *separable* in the I-adic topology if equation (4.7) holds. This is equivalent to saying that for every pair of elements $x, y \in R$, there is an open set containing x and not containing y.

Theorem 4.6.8 *Let R be an integral domain, $\pi \in R$, and suppose R is separable in the π-adic topology. Then R_π is complete in the π-adic topology. Suppose Q is a ring, I is an ideal of Q, Q is complete in the I-adic topology, and there is a continuous homomorphism $\varphi : R \to Q$. Then there is a unique continuous homomorphism $\hat{\varphi} : R_\pi \to Q$ such that $\hat{\varphi}(a) = \varphi(a)$ for all $a \in R$. In particular, R_π is the smallest ring that is complete in the π-adic topology and contains R.*

Proof: Suppose that $x_1, x_2, \cdots$ is a Cauchy sequence in R_π. Then for every n there exists a k_n so that $\pi^n|x_i - x_j$ if $i, j \geq k_n$, and we may assume that the sequence $k_1, k_2, \cdots$ is increasing. Let

$$x = x_{k_1} + \left(\frac{x_{k_2} - x_{k_1}}{\pi}\right)\pi + \left(\frac{x_{k_3} - x_{k_2}}{\pi^2}\right)\pi^2 + \cdots.$$

Then for every n, $x \equiv x_n \pmod{\pi^n}$. For all $j \geq n$ we have

$$\begin{aligned} x - x_j &= x - x_{k_n} + x_{k_n} - x_j \\ &= \left(\left(\frac{x_{k_{n+1}} - x_{k_n}}{\pi^n}\right)\pi^n + \left(\frac{x_{k_{n+2}} - x_{k_{n+1}}}{\pi^{n+1}}\right)\pi^{n+1} + \cdots\right) + x_{k_n} - x_j. \end{aligned}$$

Thus π^n divides $x - x_j$ and the sequence converges to x.

Now suppose Q is a ring, I is an ideal of Q, Q is complete in the I-adic topology, and there is a continuous homomorphism $\varphi : R \to Q$. Let $a = \sum_{i=0}^{\infty} a_i\pi^i \in R_\pi$. For each n, let $x_j = \sum_{i=0}^{j-1} a_i\pi^i \in R$. Let $y_j = \varphi(x_j)$. Let n be any positive integer. Since φ is continuous, $\varphi^{-1}(I^n) = (\pi^m)$ for some m. For all $i, j \geq m$ we have $x_i - x_j \in (\pi)^m$. Therefore $y_i - y_j \in I^n$. That is, the sequence $y_1, y_2, \cdots$ is a Cauchy sequence in Q. By the completeness of Q, this sequence converges to some $y \in Q$. Define $\hat{\varphi}(a) = y$.

It remains to show that $\varphi(a)$ is independent of the representation of a, that this definition makes φ a continuous homomorphism from R_π to Q, and that $\hat{\varphi}$ is unique. We leave these details as an exercise. □

Thus we can refer to R_π as *the completion of R in the π-adic topology.* R_ν is separable in the I_ν-adic topology.

4.7 Exercises

1. Let F be a field and suppose that k is a positive integer that is invertible in F. Let $a(x) = \sum_{i=0}^{\infty} a_i x^i \in F[[x]]$ be a power series such that a_0 is a kth power in F. Show that a is a kth power in $F[[x]]$.

2. Let F be a field that is not algebraically closed. Show that $F[[x]]$ does not contain the algebraic closure of F.

3. If $a, b \in \mathbf{Z}_N$, make the definition of ab precise and show that $\mathbf{Z}_N$ is a ring.

4. Prove that $\mathbf{Z}_N = \varprojlim\{\mathbf{Z}/(N^i)\}$.

5. Use (4) to give an alternate proof that there is a surjective homomorphism

$$\{f/g : f, g \in \mathbf{Z}, \gcd(g, N) = 1\} \to \mathbf{Z}_N.$$

6. Complete the details of the proof of Theorem 4.3.8, showing that all the appropriate homomorphisms commute.

7. Finish the proof of Theorem 4.6.8.

8. Let R be a finite ring and let I be an ideal of R. Prove that the completion of R at I is a quotient ring of R.

Bibliography

[1] Z. I. Borevich and I. R. Shefarevich, *Number Theory*, Academic Press: New York, N.Y., 1966.

[2] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer Verlag, N. Y., 1993.

[3] H. D. Ebbinghaus et al, *Numbers.* Graduate Texts in Mathematics vol. 123, Springer Verlag, N. Y. (1990).

[4] C. F. Gauss, *Disquisitiones Arithmeticae*, 1801; reprinted in English translation by Yale Univ. Press, New Haven, CT. 1966.

[5] S. Golomb, *Shift Register Sequences.* Aegean Park Press, Laguna Hills CA, 1982.

[6] G. Hardy and E. Wright, *An Introduction to the Theory of Numbers.* Oxford University Press, Oxford UK, 1979.

[7] I.N. Herstein, *Topics in Algebra*, 2nd ed., 1975: Xerox College Publ., Lexington, MA.

[8] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, Springer Verlag, N.Y., 1990.

[9] N. Jacobson, *Basic Algebra I.* W.H. Freeman, San Francisco, 1974.

[10] N. Jacobson, *Basic Algebra II.* W.H. Freeman, San Francisco, 1980.

[11] N. Koblitz, *p-Adic Numbers, p-Adic Analysis, and Zeta Functions,* Springer-Verlag: New York, 1984.

[12] D. Knuth, *The Art of Computer Programming, Vol 2. Seminumerical Algorithms.* Addison-Wesley, Reading MA, 1981.

[13] N. Koblitz, *p-Adic Numbers, p-Adic Analysis, and Zeta Functions.* Graduate Texts in Mathematics Vol. 58, Springer Verlag, N. Y. 1984.

[14] S. Lang, *Algebra*, 2nd ed., 1984: Addison-Wesley, Reading, MA.

[15] R. Lidl and H. Niederreiter, *Finite Fields*, 2nd ed., 1997: Cambridge University Press, Cambridge, UK.

[16] H. Matsumura, *Commutative Algebra*, 1970: W. A. Benjamin, New York.

[17] B. MacDonald, *Finite Rings with Identity*, 1974: Marcel Dekker, New York.

[18] R. McEliece, *Finite Fields for Computer Scientists and Engineers*, 1987: Kluwer Academic Publishers, Norwell, MA.

[19] W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes* second edition, MIT Press, Cambridge MA, 1972.

[20] B. Schneier, *Applied Cryptography.* John Wiley & Sons, New York, 1996.