Rings and Modules

Jacob Denson

July 24, 2020

Table Of Contents

Chapter 1

Basic Definitions

Rings are algebraic structures closed under addition, subtraction, and multiplication, but not necessarily under division. They can be noncommutative, such as the ring of square matrices of a fixed dimension, or commutative, like the ring of integers. To be precise, a *ring* is a set R together with addition and multiplication operations $\cdot : R \times R \to R$ and $+ : R \times R \to R$, giving R the structure of an abelian group and the structure of a (not necessarily commutative) semigroup respectively. In addition, the operations must satisfy the *distributive law*

$$a(b+c) = ab + ac \quad \text{and} \quad (b+c)a = ba + ca,$$

for any $a, b, c \in R$. Note that one equation does not establish the other since the multiplication operation need not be commutative. It is often the case that rings have a multiplicative identity, and provided that it is not also the additive identity, we denote it by 1.

Example. *The integers* $\mathbf{Z}$ *form the classical example of a ring, as are the integers* $\mathbf{Z}_n$ *modulo* n*, and we find they exhibit most of the basic properties of rings. They have a nontrivial divisibility theory, yet still possess the property of unique factorization into prime elements, an idea we will study in the more general situation of unique factorization domains.*

Example. *All the number systems* $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{F}_p$, *and* $\mathbf{C}$ *are rings, in which case every nonzero element is invertible. Such rings are known as* division rings, *and if the multiplicative operation is commutative,* fields.

Example. *For any ring A, the family of $n \times n$ matrices $M_n(A)$ with entries in A forms a ring, extensively studied in linear algebra and representation theory. An important fact about these matrices is that $M_n(M_m(A))$ is isomorphic to $M_{nm}(A)$, because* block multiplication *works in these rings.*

Example. *A key way to analyze the algebraic structure of a ring A is to introduce encodings of the algebraic structure of A through the theory of polynomials $A[X]$ over that ring, formal sums of the form $a_0 + a_1 X + \cdots + a_N X^N$, where $a_n \in A$. We view two polynomials as being equal precisely when their coefficients are equal. More generally, we can discuss their multivariate counterparts $A[X_1, \ldots, X_n]$, the ring of formal sums in the monomials $X_1^{m_1} \ldots X_n^{m_n}$. The addition of two polynomials is defined by taking the sum over each monomial separately, and the product is obtained by expanding and multiplying monomials together in the obvious way. The polynomial ring is the 'most general' way to add a new family of 'commuting' elements to a particular ring; it has the universal property that for any homomorphism $f : A \to B$ of rings and any $x_1, \ldots, x_n \in C_B(f(A))$, there exists a unique extension of f to a homomorphism $f : A[X_1, \ldots, X_n] \to B$ with $f(X_i) = x_i$ for each $i \in \{1, \ldots, n\}$.*

Example. *If A is a ring, and M is a multiplicative semigroup, then we can consider a ring $A[M]$, known as the* monoid ring, *whose elements are finite formal sums of the form $\sum a_n x_n$, with $x_n \in A$, $x_n \in M$, with the obvious additive structure, and with multiplicative structure defined by multiplying elements of the semigroup termwise. We calculate that*

$$\left(\sum_{x \in M} a_x x \right) \left(\sum_{y \in M} b_y y \right) = \sum_{x,y \in M} a_x b_y (xy) = \sum_{z \in M} \left(\sum_{xy=z} a_x b_y \right) z.$$

Thus multiplication is given by a kind of convolution in the coefficients of the ring. In the case where $M = \mathbf{N}$ or $M = \mathbf{N}^n$, $A[M]$ is the polynomial ring in one or more variables. There are other variants of this group:

- *If M is a semigroup such that, for each $k \in M$, there are only finitely many pairs $g, h \in M$ such that $gh = k$, then we can consider the ring $A_\infty[M]$ of infinite formal sums $\sum a_n x_n$, with multiplication defined as in $A[M]$. For $M = \mathbf{N}$, the ring $A[[M]]$ can be viewed as the ring of formal power series with elements in A.*

- *If G is a locally compact group with Haar measure μ, then for $f, g \in L^1(G)$, we can define the* convolution function *$f * g$ by the formula*

$$(f * g)(x) = \int f(y) g(y^{-1}x)\, d\mu(y).$$

  Then $\|f * g\|_{L^1(M)} \leqslant \|f\|_{L^1(M)} \|g\|_{L^1(M)}$, so $L^1(M)$ can be viewed as a generalization of the ring $\mathbf{R}[G]$ to more analytical settings. If G is non-discrete, however, then $L^1(G)$ does not have an identity; to obtain an identity, we typically need to introduce the Dirac mass *δ at the identity, which is no longer an integrable function, but can be defined as a finite measure on G, namely part of the more general measure algebra $M(G)$ upon which we have an identity.*

If $M = G$ is a multiplicative group, then $R[G]$ is known as the group ring *associated with G.*

Example. *The quaternion division ring* **H**, *named after their creator, Hamilton, which we can informally describe as the family of formal quantities $a + bi + cj + dk$ with operations induced by the identities*

$$i^2 = j^2 = k^2 = ijk = -1$$

Formally, we can construct this ring by considering the group algebra $\mathbf{R}[Q]$, where Q is the quaternion group, whose elements we denote by

$$\{1, \overline{1}, i, \overline{i}, j, \overline{j}, k, \overline{k}\}.$$

This group algebra cannot be identified with **H**; *for instance, we would like $\overline{1}$ to be equal to -1 in this ring. To solve this problem, we quotient by the ideal $\mathfrak{a} = (\overline{1} + 1)$. For any $s \in \{1, i, j, k\}$, we therefore conclude that*

$$\overline{s} = \overline{1} \cdot s = -s + (\overline{1} + 1) \cdot s \in -s + \mathfrak{a}.$$

Thus any element of $\mathbf{R}[Q]/\mathfrak{a}$ can be written as $a + bi + cj + dk$ for some real numbers $a, b, c, d \in \mathbf{R}$. We claim no nontrivial elements of $a + bi + cj + dk$ are in $\mathfrak{a}$, so any element of $\mathbf{R}[Q]/\mathfrak{a}$ can be uniquely *expressed in this way. Indeed, if there is $x \in \mathbf{R}[Q]$ such that*

$$x(\overline{1} + 1) = a + bi + cj + dk$$

and if we write $x[s]$ for the coefficient of x corresponding to $s \in Q$, then we must have $x[s] = -x[\overline{s}]$ for each $s \in \{1, i, j, k\}$. But then

$$x(\overline{1} + 1) = \sum_{s \in \{1,i,j,k\}} (x[s] + x[\overline{s}])\, s = \sum_{s \in \{1,i,j,k\}} (x[s] - x[s])s = 0.$$

Thus $a + bi + cj + dk = 0$. Thus the quotient $\mathbf{R}[Q]/\mathfrak{a}$ can be identified with a ring structure on the set $\{a + bi + cj + dk : a, b, c, d \in \mathbf{R}\}$, which is the algebraic structure we want to study. Invented by the Irish mathematician Lord Hamilton in the mid 19th century to obtain algebraic characterization of rotation in three dimensional space, the quaternions have a special place in an algebraists heart, for they are the first truly strange algebraic structures obtained in the historical development of abstract algebra.

Example. *George Boole began the modern study of logic by studying the algebraic notions of truth. He saw that the logical operations of conjunction and disjunction behaved very similarily to the algebraic operations of multiplication and addition. If we consider the set of all equivalence classes of logical statements (two statements being equivalent if they both imply each other), and consider conjunction as a multiplication, and exclusive disjunction as an additive structure, then we obtain a ring satisfying $x^2 = x$ for all statements x, where 0 is a statement that is always false, and 1 a statement the is always true. In his honour, we call a ring* Boolean *if this equation is satisfied. In any Boolean ring, $x + x = 0$, since*

$$x + x = (x + x)^2 = x^2 + x + x + x^2 = x + x + x + x.$$

We say the ring has characteristic two. *Moreover, any Boolean ring is commutative, since*

$$x + y = (x + y)^2 = x^2 + xy + yx + y^2 = x + xy + yx + y$$

from which we obtain $xy + yx = 0$, so $yx = -xy = xy$. For any set X, the set of subsets of X form a Boolean ring, such that for $A, B \subset X$,

$$A + B = (A - B) \cup (B - A) \quad \textit{and} \quad A \cdot B = A \cap B.$$

The set X acts as the multiplicative identity, and the empty set $\emptyset$ is the additive identity. Boolean rings are essentially the same as Boolean algebras studied in logic and measure theory, and the exact correspondence is provided by the Stone representation theorem, employing tools from set theoretic topology!

Example. *The theory of rings arises very often in the study of functions. If A is a ring, and X is a set then one can make the set A^X of maps from X to A into a ring, by defininining addition and multiplication pointwise. Thus, for instance, the set $\mathbf{R}^{\mathbf{N}}$ of real valued sequences forms a ring, as does $\mathbf{R}^{\mathbf{R}}$. Subrings of these rings occur all the time in analysis. The ring $C_c(\mathbf{R})$ of compactly supported continuous functions on the real line provides our first natural example of a ring without identity, as does the ring $C_0(\mathbf{R})$ of continuous functions vanishing at infinity.*

Example. *If B is a commutative subring of a ring A, then for any subset S of A, we can consider the ring generated by B and S, denoted by $B[S]$. Interesting examples of these include the* Gaussian integers $\mathbf{Z}[i]$, *whose points form a lattice in the plane, and the* dyadic numbers $\mathbf{Z}[1/2]$, *which are the fractions expressible with a denominator a power of two, which form a dense subset of the real line. In algebraic number theory, one studies the ring $\mathbf{Z}[\sqrt{D}]$, where D is a squarefree integers. Such a ring can be described as the set of all numbers of the form $n + m\sqrt{D}$, where $n, m \in \mathbf{Z}$.*

Remark. There is only a single example of a ring with identity in which the multiplicative identity equals the additive identity. This is because for any a,

$$a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$$

From which we conclude $a \cdot 0 = 0$. But if 0 is also a multiplicative identity, we conclude that $a = a \cdot 0$. This means the only ring for which the additive and multiplicative identities correspond is the ring consisting of a single element: the number zero! We denote the zero ring by (0), to mirror the notation we will develop for ideals later on.

Remark. One might remark that one might wish to study algebraic structures $(R, +, \cdot)$ in which the *addition operation* is noncommutative. Such structures do exist, but are exceedingly rare in applications (they are known as *near rings*). Perhaps such rings do not show up often is that if a near ring has a multiplicative identity, the addition is automatically commutative. To see why, we note that if R is a near ring with identity, then for any $a, b \in R$, the two different applications of the distributive law imply that

$$(1 + 1) \cdot (a + b) = (1 + 1) \cdot a + (1 + 1) \cdot b = a + a + b + b$$

and

$$(1 + 1) \cdot (a + b) = 1 \cdot (a + b) + 1 \cdot (a + b) = a + b + a + b.$$

Combining these two equations gives $a + b = b + a$.

Rings arise naturally when we start studying symmetries of preexisting algebraic structures, like those that arise in group theory. Matrices are symmetries of vector spaces, which themselves can be seen as shifting symmetries of space. Polynomials are symmetries over a field of numbers, which themselves are also very well behaved symmetries. In fact, like groups, all rings can be represented as symmetries of some abelian group.

Example. *Let G be an abelian group, and consider the set $\mathrm{End}(G)$ of all homomorphisms from G to itself. We define a ring structure on this group. Given $f, g \in \mathrm{End}(G)$, we define $f + g$ to be the endomorphism on G defined by $(f + g)(x) = f(x) + g(x)$, and where composition $f \circ g$ is the multiplicative structure. The fact that $\mathrm{End}(G)$ satisfies the laws of a ring are trivial, with the identity behaving as 1, and the trivial homomorphism acting as 0.*

Theorem 1.1. *Suppose A is a ring such that there does not exist nonzero $a \in A$ with $ax = 0$ for all $x \in A$. Then A is isomorphic to a subring of $\mathrm{End}(G)$ for some G.*

Proof. Given a ring A, let A_+ denote the additive group structure of A. Then $\mathrm{End}(A_+)$ is a ring. Consider the map $\varphi : A \to \mathrm{End}(A_+)$, where for $a \in A$, $\varphi(a)$ acts as the map $x \mapsto ax$. The distributive law implies that such a map is an endomorphism. What's more φ is a ring homomorphism, since for each $x, y, z \in A$,

$$\varphi(x + y)(z) = (x + y)(z) = xz + yz = [\varphi(x) + \varphi(y)](z)$$

and

$$\varphi(xy)(z) = (xy)z = x(yz) = (\varphi(x) \circ \varphi(y))(z).$$

The map φ is injective, since if $\varphi(a) = 0$, then $ax = 0$ for all $x \in A$. Thus $\mathrm{End}(A)$ naturally contains A as a subring. □

The problem with this proof is that the theorem doesn't really give a 'nice' answer to what a ring really is. Groups are already abstract, so we may not necessarily be able to visualize what a symmetry of an arbitrary abstract object is. Alas, most general theories in mathematics do not have natural correspondences with a single object of study, unlike the niceities of group theory. This is to be expected, since ring theory arose from many fields of study, like number theory, geometry, and logic. We will just have

to accept this theorem as a little tidbit of intuition, and move on. We will return to this idea in the theory of modules, where one studies a ring 'acting' on an abelian group, just like Cayley's theorem hints at using group actions to understand the theory of group actions.

1.1 Units of a Ring

We begin with discussing an operation that seems left out of the definition of a ring – divisibility. In the ring of rational numbers, we can divide a rational number by any *non-zero* rational number, and still get a rational number. On the other hand, an integer divided by an integer is only in very special cases an integer. If A is a ring, the *units* are the elements x which possess a multiplicative inverse x^{-1} such that $xx^{-1} = 1 = x^{-1}x$; note both ends of the equation need to be satisfied since ab may not equal to ba. For example, when A is the ring of endomorphisms on a ring, $ab = 1$ implies b is injective, whereas $ba = 1$ implies b is surjective, and when the set is infinite injectivity is not equivalent to surjectivity. Thus there are subtle differences between *left invertibility* and *right invertibility*; however, as might be expected from the endomorphism model case, if an element $a \in A$ has both a left inverse and a right inverse, then both inverses agree, and a is a unit (since if $ba = ac = 1$, then $ab = abac = ac = 1$. We let $U(A)$ denote the set of all units in a ring. It forms a multiplicative group. Here are some examples:

- In the ring of integers, $U(\mathbf{Z}) = \{\pm 1\}$.

  Proof: For any $n, m \in \mathbf{Z}$, $|nm| \geqslant \max(n, m)$, so if $nm = 1$, $n = \pm 1$.

- In the ring of dyadic numbers, $U(\mathbf{Z}[1/2]) = \{\pm 2^n : n \in \mathbf{Z}\}$.

  Any element of $\mathbf{Z}[1/2]$ is of the form $n2^m$ for some integers n, m, where n is odd. But $n2^m \in U(\mathbf{Z}[1/2])$ if and only if $1/n \in \mathbf{Z}[1/2]$. But if we can write $1/n = r2^k$ for integers r and k with r odd, then $nr = 2^k$, which is only possible if $n, r \in \{\pm 1\}$.

- In the ring of Gaussian integers, $U(\mathbf{Z}[i]) = \{\pm 1, \pm i\}$.

  Let $x \in U(\mathbf{Z}[i])$, and suppose $x = n + im$. Then there is $k, r \in \mathbf{Z}$ such that $(n + im)(k + ir) = 1$. But this implies that

$$|(n + im)(k + ir)|^2 = (n^2 + m^2)(k^2 + r^2) = 1$$

so $n^2 + m^2 = 1$, implying $n \in \{\pm 1, \pm i\}$.

This technique can be generalized to calculate the group of units of $U(\mathbf{Z}[\sqrt{D}])$ for any squarefree integer D. We define $N : \mathbf{Z}[\sqrt{D}] \to \mathbf{Z}$ by setting

$$N(n + m\sqrt{D}) = \left(n + m\sqrt{D}\right)\left(n - m\sqrt{D}\right) = n^2 - Dm^2$$

A perhaps surprising fact, following from a short calculation, is that N is multiplicative, i.e. $N(xy) = N(x)N(y)$ for any $x, y \in \mathbf{Z}[\sqrt{D}]$. Thus if $x \in U(\mathbf{Z}[\sqrt{D}])$, then $1 = N(xx^{-1}) = N(x)N(x^{-1})$, so $N(x) \in \{\pm 1\}$. Conversely, if $N(x) = \pm 1$, then $x^{-1} = \pm(n - m\sqrt{D})$, so $x \in U(\mathbf{Z}[\sqrt{D}])$. Calculating the units of these rings thus reduces to counting integer solutions to the Diophantine equation $n^2 - Dm^2 = \pm 1$.

The value $D = -1$ gives the Gaussian integers. For any integer $K > 1$, the equation $n^2 + Km^2 = \pm 1$ only has the trivial solutions $n = \pm 1$, $m = 0$, so $U(\mathbf{Z}[\sqrt{D}]) = \{\pm 1\}$ for $D < -1$. For any $D \geqslant 2$, $U(\mathbf{Z}[\sqrt{D}])$ is infinite; for instance, $(1 + \sqrt{2})^n \in \mathbf{Z}[\sqrt{2}]$ for any $n \geqslant 0$.

- The group of units in $\mathbf{Z}_n$ is the set of equivalence classes of integers coprime to n. This is because if m is coprime to n, then Bezout's theorem implies that there exists integers $k, k' \in \mathbf{Z}$ such that $km + k'n = 1$, and then $km = 1$ in $\mathbf{Z}_n$. Thus $U(\mathbf{Z}_n)$ contains $\phi(n)$ elements.

- If A is a unital ring with no zero divisors, then the units of the polynomial ring $A[X_1, \ldots, \ldots, X_n]$ are precisely the units of A. Similarily, a power series in $A[[X]]$ is a unit in $A[[X]]$ if and only if the constant term in the power series is a unit in A.

  One way to verify this is using the degree formula. If A has no zero divisors, then $\deg(fg) = \deg(f) + \deg(g)$ for any polynomials $f, g \in A[X_1, \ldots, X_n]$. Thus if $fg = 1$, then $\deg(f) = \deg(g) = 0$, which implies $f, g \in A$, and thus $f, g \in U(A)$. Similarily, we see that for two power series f and g, the constant term of fg is the product of the constant terms of f and g. Thus if $fg = 1$, then the constant term of f is in $U(A)$. Conversely, if

$$f = \sum_{n=0}^{\infty} a_n X^n,$$

and a_0 is a unit, then we can define an inverse to f by defining

$$g = \sum_{n=0}^{\infty} b_n X^n,$$

by defining $b_0 = a_0^{-1}$, and then inductively defining $b_n = -a_0^{-1}(a_1 b_{n-1} + \cdots + a_{n-1} b_0)$. Then $fg = 1$. But we can similarily define a left inverse to f, so f is invertible.

- If A is a ring with identity, then Cramer's rule tells us that the units of $M_n(A)$ are precisely those matrices whose determinant is a unit in A. In particular, if k is a field, then the group of units in $M_n(k)$ is the set of matrices with nonzero determinant, which gives the general linear group $GL_n(k)$.

- In number theory it is common to study the monoid ring $\mathbf{C}[[\mathbf{N}]]$ as the ring of functions $f : \mathbf{N} \to \mathbf{C}$, with addition given pointwise, and with multiplication given by *Dirichlet convolution*

$$(f * g)(k) = \sum_{nm=k)} f(n)g(m)$$

Then $\mathbf{C}[[\mathbf{N}]]$ has a multiplicative identity

$$\delta(n) = \begin{cases} 1 & n = 1 \\ 0 & \text{otherwise} \end{cases}$$

If μ is the *Möbius function*

$$\mu(n) = \begin{cases} 0 & p^2 \text{ divides } n \text{ for some prime } p \\ (-1)^k & n \text{ has } k \text{ distinct prime factors} \end{cases}$$

and 1 is the constant function with $1(n) = 1$ for all $n \in \mathbf{N}$, then it is easy to see that $1 * \mu = \delta$. The Möbius inversion formula states that $f = g * \mu$ if and only if $f * 1 = g$, and this is nothing more than the fact than saying that μ is a unit in the convolution ring, with 1 as it's inverse.

As with groups, one may consider subrings of a ring, and homomorphisms between rings. After studying group theory, you should be able to figure out the definitions yourself, but for completeness, we now specify them. A *subring* of a ring is a subset of a ring which also possesses a ring structure. That is, a subring is closed under addition and multiplication.

Example. *The most classical chain of commutative subrings is*

$$\mathbf{Z} < \mathbf{Q} < \mathbf{R} < \mathbf{C} < \mathbf{H}$$

The other subrings in this chain are still actively researched today, most importantly in the theory of algebraic number theory.

Example. *If A is a ring, the center $Z(A)$ is defined to be the set of elements a such that, for all $b \in A$, $ab = ba$. Then $Z(A)$ is a commutative subring of A. Similarily, for any $a \in A$, the center $C(a) = \{x \in A : ax = xa\}$ is a subring of A, but is not necessarily commutative.*

A particularly important example of this occurs in the case when we study the group algebra $A[G]$. One can verify that if G is a finite group, and K is a conjugacy class of G containing elements $k_1, \ldots, k_n \in G$, then

$$k(K) = k_1 + \cdots + k_n$$

lies in $Z(A[G])$. This is because for any $a \in A$ and $g \in G$,

$$\begin{aligned} ag(k_1 + \cdots + k_n) &= a(gk_1g^{-1})g + \cdots + a(gk_ng^{-1})g \\ &= [(gk_1g^{-1}) + \cdots + (gk_ng^{-1})](ag) \\ &= (k_1 + \cdots + k_n)(ag), \end{aligned}$$

where the last equality follows because the map $k_i \mapsto gk_ig^{-1}$ permutes K. A similar calculation verifies that any elements of $Z(A[G])$ must be constant on any conjugacy class of G. Thus if $K_1, \ldots, K_N$ are the conjugacy classes of G, then

$$Z(A[G]) = \{a_1k(K_1) + \cdots + a_Nk(K_N) : a_1, \ldots, a_N \in Z(A)\}.$$

As verified in representation theory, if $A = \mathbf{C}$, then the irreducible characters of G form an orthonormal basis for $Z(A[G])$.

Example. *The continuous functions form a subring of $\mathbf{R}^{\mathbf{R}}$, as do the polynomial functions, or differentiable functions, and so on and so forth.*

Example. *We have a homomorphism from the semigroup ring $R[M]$ to R by setting*

$$\varphi\left(\sum a_i x_i\right) = \sum a_i.$$

The kernel of this map is called the augmented ideal *of the semigroup ring.*

Example. *For each complex number $z \in \mathbf{C}$, then, identifying $\mathbf{C}$ with $\mathbf{R}^2$, we obtain a linear transform $T_z : \mathbf{R}^2 \to \mathbf{R}^2$ given by multiplication on the left by z, which, if $z = u + iv$, we can identify with the 2×2 matrix*

$$\begin{pmatrix} u & -v \\ v & u \end{pmatrix}.$$

The algebra of complex multiplication implies that the map $\varphi : \mathbf{C} \to M_2(\mathbf{R})$ obtained by setting

$$\varphi(u + iv) = \begin{pmatrix} u & -v \\ v & u \end{pmatrix}$$

is an injective homomorphism. Similarily, we can obtain an injective homomorphism $\varphi : \mathbf{H} \to M_4(\mathbf{R})$ by setting

$$\varphi(a + bi + cj + dk) = \begin{pmatrix} +a & -b & -c & -d \\ +b & +a & -d & +c \\ +c & +d & +a & -b \\ +d & -c & +b & +a \end{pmatrix}.$$

Thus we can view $\mathbf{C}$ and $\mathbf{H}$ as certain subrings of matrices.

The family $SO(2) \in M_2(\mathbf{R})$ can be identified precisely with the set of unitary *complex numbers, i.e. those $z \in \mathbf{C}$ with $|z| = 1$. In some sense, the complex numbers as we know them today were invented precisely to provide a coordinate system for rotations and dilations in space in this manner. The quaternions were invented to analyze rotations and dilations in three dimensional space, but things do not work out quite as nicely in this setting. We note that $\varphi(z) \in SO(4)$ for each unitary $z \in \mathbf{H}$ with $|z| = 1$. If we set G to be the group of all unit quaternions, then we obtain a group homomorphism $\psi : G \to SO(4)$ by setting $(\psi z)(w) = zwz^{-1}$, where $w \in \mathbf{H}$ is identified canonically with a unit vector in $\mathbf{R}^4$. Note that $(\psi z)(1) = 1$, so $\mathbf{R}$ is an invariant subspace of ψz. But since $\psi z \in SO(4)$, if we define $\mathbf{H}_0 = \{ai + bj + ck : a, b, c \in \mathbf{R}\}$ to be the set of* pure quaternions, *then ψz acts as an orthogonal transformation on $\mathbf{H}_0$. Thus we obtain an induced homomorphism from G to $SO(3)$. Geometrically,*

ψz can be described in the following way: Any element z of G can be written as $\cos\theta + \sin\theta z_0$, where z_0 is a pure unit quaternion. The rotation ψz is then a counterclockwise rotation of 2θ about z_0 in $\mathbf{H}_0$. Certainly z_0 is fixed by z, because

$$zz_0z^{-1} = (\cos\theta + \sin\theta z_0)z_0(\cos\theta - \sin\theta z_0) = z_0.$$

Thus z does rotate about z_0 at a certain angle. Let us see that this angle is 2θ is the angle in the special case where $z = \cos\theta + \sin\theta i$. Then

$$\begin{aligned} zjz^{-1} &= (\cos\theta + \sin\theta \cdot i)j(\cos\theta - \sin\theta \cdot i) \\ &= (\cos\theta \cdot j + \sin\theta \cdot k)(\cos\theta - \sin\theta \cdot i) \\ &= ((\cos^2\theta - \sin^2\theta) \cdot j + 2\sin(\theta)\cos(\theta) \cdot k) \\ &= (\cos(2\theta)j + \sin(2\theta)k). \end{aligned}$$

Thus we see the angle 2θ is correct here. But this claim is now true in general, because we can write any pure unit quaternion z_0 as $w_0iw_0^{-1}$ for some other unit quaternion w_0; the unit vector $w_0jw_0^{-1}$ is then orthogonal to z_0, and one can calculate that

$$z(w_0^{-1}jw_0)z^{-1} = (\cos(2\theta)(w_0^{-1}jw_0) + \sin(2\theta)(w_0^{-1}kw_0)).$$

This geometric description makes it easy to see that $\psi : G \to SO(3)$ is a double cover, *with kernel equal to -1. More generally, ψ extends to a map from the nonzero elements of $\mathbf{H}$ to the group of oriented rotations and dilations.*

1.2 Homomorphisms and Ideals

A *ring homomorphism* from a ring A to a ring B is a function $\phi : A \to B$ which is a homomorphism of abelian groups, and a homomorphism of the multiplicative semigroup structure on the two spaces. If the rings are unital, we also assume that they map the identity to the identity. As with groups and vector spaces, the kernel $\mathrm{Ker}(\phi)$ of the map ϕ is defined to be the set of all a such that $\phi(a) = 0$. As with groups, determining the structure of the kernel of a homomorphism will enable us to obtain a variant of the isomorphism theorems for rings, which we carry out in the next section.

Example. *If A is a unital ring, there is a unique homomorphism of unital rings from $\mathbf{Z}$ to A, since $\mathbf{Z}$ is generated by the unit. We identify elements of*

$\mathbf{Z}$ *with elements of A. Even if A is non-unital, we can still define an action of* $\mathbf{Z}$ *on the additive structure of R by defining* $nx = x + \cdots + x$ *as the n fold sum of x's. Thus we can consider the additive group* $\mathbf{Z} \oplus R$ *with an additional multiplication operation* $(n \oplus x)(m \oplus y) = nm \oplus (mx + ny + xy)$*. One verifies that this gives the structure of a ring, which now has an identity* $1 \oplus 0$*. The resultant ring is known as the* unitization *of R, and has the universal property that every homomorphism of R into a unital ring extends to a homomorphism of the unitization of R. This process often enables us to extend facts about unital rings to facts about non-unital rings, which is why it is often fine to assume that a ring has unity for convenience. D.D. Anderson has written a useful article on the correspondence between unital and non-unital rings, entitled* Commutative Rngs.

We wish to establish a quotient structure on rings, and obtain analogies of the isomorphism theorems for groups. Let's consider $\mathfrak{a}$ as a subset of a ring A, and try to determine which properties allow the cosets $A/\mathfrak{a}$ of the form $x+\mathfrak{a}$ allow the operations on A to be well defined on the quotient. In order to even define these cosets, we first need $\mathfrak{a}$ to be an additive subgroup of the additive group structure on $\mathfrak{a}$. Since all subgroups of abelian groups are normal, this means the operation of addition on the quotient is well defined. In order for multiplication to be well defined, we need to conclude $(a + \mathfrak{a})(b + \mathfrak{a}) = (ab + \mathfrak{a})$. In terms of sets, this says

$$\{(a + x)(b + y) = ab + xb + ay + xy : x, y \in \mathfrak{a}\} = \{ab + x : x \in \mathfrak{a}\}$$

Thus we require $xb + ay + xy \in \mathfrak{a}$ for any $x, y \in \mathfrak{a}$. This implies that $\mathfrak{a}$ not *only* needs to be closed under multiplication, but also closed under multiplication by an element of A, both on the left and the right. We say $\mathfrak{a}$ is an *ideal* if it is an additive subgroup of R closed under multiplication on the left and the right. In a commutative ring, of course, we need only prove that an ideal is closed by multiplication on the left.

Example. *As should be expected, if* $\phi : A \to B$ *is a ring homomorphism, then the kernel* $Ker(\phi)$ *is a double sided ideal of A. Conversely, if* $\mathfrak{a}$ *is a two-sided ideal, then* $A/\mathfrak{a}$ *is a ring, and the projection* $\pi : A \to A/\mathfrak{a}$ *is a homomorphism with kernel* $\mathfrak{a}$*. A ring homomorphism is an isomorphism if and only if the kernel is trivial.*

Example. *Every additive subgroup of* $\mathbf{Z}$ *is a set of multiples of some number n, which we denote by* (n)*. It is also an ideal of* $\mathbf{Z}$*, and so an ideal in* $\mathbf{Z}$ *is in*

one correspondence with the set of integers. In general, if A is a commutative ring, and $x \in A$, we find $(x) = Ax = \{ax : a \in A\}$ is an ideal, known as a principal ideal. *If A is a ring such that every ideal is principal, we say that A is a* principal ideal ring. *The integers are an example, as we have just argued. We can also consider the ideal $(x_1, \ldots, x_n)$, the smallest ideal containing the elements $x_1, \ldots, x_n$. An ideal of this form is known as a* finitely generated ideal.

Example. *Let A be a division ring, and suppose $\mathfrak{a}$ is an ideal containing some $x \neq 0$. Then for any $a \in \mathfrak{a}$, $aax^{-1}x \in \mathfrak{a}$. Thus $\mathfrak{a} = A$. Thus the only ideals of A are (0) and (1). In particular, we conclude that the only homomorphisms from A to any other ring B are injective, or identically zero.*

Example. *Suppose A is a ring with unity. Then the only ideals of $M_n(A)$ are those of the form $M_n(\mathfrak{a})$ for some ideal $\mathfrak{a}$ in A. Indeed, if I is an ideal in $M_n(A)$, then for each i, j,*

$$\mathfrak{a}_{ij} = \{M_{ij} : M \in I\}$$

forms an ideal in A. It is certainly an additive subgroup. Moreover, if $a \in A$, and $a_0 \in \mathfrak{a}_{ij}$, we can find $M \in I$ with $M_{ij} = a_0$. If I_n is the identity matrix in $M_n(A)$, then $[(aI_n)M]_{ij} = aa_0$ and $[M(aI_n)]_{ij} = a_0 a$, so $aa_0, a_0 a \in \mathfrak{a}_{ij}$. Multiplication by permutation matrices verifies that $\mathfrak{a}_{ij}$ does not really depend on i and j, so we denote it my $\mathfrak{a}$. And since

$$e_{ii} M e_{jj} = M_{ij} e_{ij},$$

it is clear that $I = M_n(\mathfrak{a})$. Note, in particular, that if A is a division ring, then $M_n(A)$ has no nontrivial ideals. *Thus any homomorphism with domain $M_n(A)$ is either identically zero, or injective.*

Example. *If $M \in M_n(\mathbf{R})$, then there exists two invertible square matrices N and K such that*

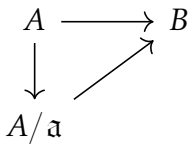
$$NMK = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$$

Then for any for $j \leqslant k$, $e_{ij} NMK = e_{ij}$. Similarily, if $i \leqslant k$, $NMK e_{ij} = e_{ij}$. In particular, we conclude that the two sided ideal generated by M over $M_n(K)$ contains all e_{ij}, so in particular $M_n(K)$ has only two ideals, (0) and (1). In particular, this means that every homomorphism from $M_n(K)$ to another ring A must be injective, or is identically zero. Nonzero rings with this property are called simple.

The phenomenon in the last example occurs rarely in the family of commutative unital rings. This is because A is commutative, nonzero, unital, and simple if and only if it is a field. If A has these properties and $x \in A$ is nonzero, then $(x) = A$, so there is $y \in A$ such that $xy = 1$. Thus x is invertible.

Now we have defined quotient rings and homomorphisms, we obtain the isomorphism theorems for rings, which are direct analogues to the isomorphism theorems for groups.

Theorem 1.2 (First Isomorphism Theorem)**.** *Let $\phi : A \to B$ be a homomorphism of rings. If $\mathfrak{a}$ is an ideal contained in the kernel of ϕ, then there is a unique morphism from $A/\mathfrak{a} \to B$ satisfying the commutative diagram*



If $\mathfrak{a}$ is equal to the kernel of ϕ, then the morphism from $A/\mathfrak{a}$ to B is injective.

Theorem 1.3 (Second Isomorphism Theorem)**.** *Let B be a subring of A, and $\mathfrak{a}$ an ideal of A. Then $B + \mathfrak{a}$ is a subring of A, $\mathfrak{a}$ is an ideal in $B + \mathfrak{a}$, $B \cap \mathfrak{a}$ is an ideal in B, and*

$$B/(B \cap \mathfrak{a}) \cong (B + \mathfrak{a})/\mathfrak{a}.$$

Theorem 1.4 (Third Isomorphism Theorem)**.** *Let $\mathfrak{a}, \mathfrak{b}$ be ideals of A, with $\mathfrak{a} \subset \mathfrak{b}$. Then $\mathfrak{b}/\mathfrak{a}$ is an ideal of $A/\mathfrak{a}$, and*

$$(A/\mathfrak{a})/(\mathfrak{b}/\mathfrak{a}) \cong A/\mathfrak{b}.$$

Theorem 1.5 (Fourth Isomorphism Theorem)**.** *If $\phi : A \to B$ is a surjective homomorphism, there is a bijection correspondence with subrings of B and subrings of A containing ϕ, and a subring of B is an ideal if and only if the corresponding subring of A is an ideal.*

Example. *If A is a unital ring, we can consider the unique unital homomorphism from $\mathbf{Z}$ to A, whose kernel is of the form (n) for some positive integer n. We call n the* characteristic *of the ring A. The ring A therefore contains a subring isomorphic to $\mathbf{Z}_n$ for some integer n, known as the* prime subring *of A. The reason for the terminology is that, if A has no zero divisors, then n must necessarily be prime.*

Example. *The first isomorphism theorem implies the ring* $\mathbf{Z}[i]$ *of Gaussian integers is isomorphic to* $\mathbf{Z}[X]/(X^2+1)$*, where the isomorphism is induced by the* $\mathbf{Z}$*-algebra homomorphism* $\phi : \mathbf{Z}[X] \to \mathbf{Z}[i]$ *satisfying* $\phi(X) = i$*. Two applications of the third isomorphism theorem implies that*

$$\begin{aligned}\mathbf{Z}[i]/(i-2) &\cong \mathbf{Z}[X]/(X^2+1, X-2) \\ &= \mathbf{Z}[X]/(5, X-2) \cong \mathbf{Z}_5[X]/(X-2) \cong \mathbf{Z}_5.\end{aligned}$$

Given a ring A, we let $\mathcal{I}(A)$ denote the set of all ideals of A. A morphism $\phi : A \to B$ induces a map $\phi^{-1} : \mathcal{I}(B) \to \mathcal{I}(A)$. We can also consider a map $\phi_* : \mathcal{I}(A) \to \mathcal{I}(B)$ which associates with each ideal $\mathfrak{a} \in \mathcal{I}(A)$ the smallest ideal in $\mathcal{I}(B)$ containing $\phi(\mathfrak{a})$. If $\mathfrak{k}$ is the kernel of ϕ, we calculate that

$$\phi^{-1}(\phi_*(\mathfrak{a})) = \mathfrak{a} + \mathfrak{k} \quad \text{and} \quad \phi_*(\phi^{-1}(\mathfrak{b})) = \mathfrak{b}$$

for all ideals $\mathfrak{a} \in \mathcal{I}(\mathfrak{a})$ and $\mathfrak{b} \in \mathcal{I}(B)$. In particular, ϕ^{-1} is always an injective map. Studying the Galois connection between ϕ_* and ϕ^{-1} is useful to understand almost any homomorphism ϕ, in particular, determining which ideals in $\mathcal{I}(A)$ occur as images of ϕ^{-1}.

1.3 Properties of Ideals

Let A be a ring. For any family of ideals $S \subset \mathcal{I}(A)$, $\bigcap S \in \mathcal{I}(A)$. A consequence of this is we can talk about a generating set of an ideal. We say a set S *generates* $\mathfrak{a}$ if $\mathfrak{a}$ is the smallest ideal containing S, and we denote $\mathfrak{a}$ by (S). Using this fact, we can define algebraic operations on ideals. Given $\mathfrak{a}, \mathfrak{b} \in \mathcal{I}(A)$, the set $\mathfrak{a} + \mathfrak{b}$ is an ideal, and is the smallest ideal containing $\mathfrak{a}$ and $\mathfrak{b}$. More generally, we can take infinite sums of ideals, often using the notation $\bigoplus \mathfrak{a}_\alpha$, which is just the smallest ideal containing all the ideals in the sum. Together with intersection, we find that the family of ideals forms a complete lattice on the subsets of a ring, just like the family of subgroups of a group. More interestingly, we have a product structure. Given two ideals $\mathfrak{a}, \mathfrak{b} \in \mathcal{I}(A)$, we can consider the product $\mathfrak{a}\mathfrak{b}$, which is the ideal *generated* by products of the form ab, with $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$ (CAUTION: the set $\{ab : a \in \mathfrak{a}, b \in \mathfrak{b}\}$ may not be an ideal).

With the operations of addition and multiplication, one might expect $\mathcal{I}(A)$ to have a ring structure, but this isn't true; nonetheless, $\mathcal{I}(A)$ does

form a monoid under addition and multiplication; the multiplicative identity is A, and the additive identity is (0). They are some useful algebraic relations here; for instance, we have the distributive laws

$$\mathfrak{a}(\mathfrak{b}+\mathfrak{c}) = \mathfrak{a}\mathfrak{b}+\mathfrak{a}\mathfrak{c} \quad \text{and} \quad (\mathfrak{a}+\mathfrak{b})\mathfrak{c} = \mathfrak{a}\mathfrak{c}+\mathfrak{b}\mathfrak{c}.$$

If we place products with intersections, then we obtain the partial distributive law

$$\mathfrak{a}\cap\mathfrak{b}+\mathfrak{a}\cap\mathfrak{c} \subset \mathfrak{a}\cap(\mathfrak{b}+\mathfrak{c}).$$

However, in the special case that $\mathfrak{a}\subset\mathfrak{b}$ or $\mathfrak{a}\subset\mathfrak{c}$, we do have equality here, a fact known as the *modular law*. Using unique factorization, we see that

$$(\mathfrak{a}+\mathfrak{b})(\mathfrak{a}\cap\mathfrak{b}) = \mathfrak{a}\mathfrak{b}$$

in principal ideal domains, but we only have

$$(\mathfrak{a}+\mathfrak{b})(\mathfrak{a}\cap\mathfrak{b}) \subset \mathfrak{a}\mathfrak{b}$$

in general rings. One trivially verifies that $\mathfrak{a}\mathfrak{b}\subset\mathfrak{a}\cap\mathfrak{b}$ if $\mathfrak{a},\mathfrak{b}\in\mathcal{I}(A)$. In a commutative unital ring, one verifies we have equality here provided $\mathfrak{a}$ and $\mathfrak{b}$ are '*relatively prime*'.

Lemma 1.6. *If A is a commutative unital ring, and $\mathfrak{a}_1,\dots,\mathfrak{a}_n$ are ideals of A which are* pairwise relatively prime, *in then sense that $\mathfrak{a}_i+\mathfrak{a}_j = A$ for $i\neq j$, then*

$$\mathfrak{a}_1\cdots\mathfrak{a}_n = \mathfrak{a}_1\cap\cdots\cap\mathfrak{a}_n.$$

Proof. Let us consider first the case $n=2$. It suffices to prove $\mathfrak{a}_1\cap\mathfrak{a}_2\subset\mathfrak{a}_1\mathfrak{a}_2$. But we may find $a_1\in\mathfrak{a}_1$ and $a_2\in\mathfrak{a}_2$ such that $a_1+a_2=1$ by assumption. Given $c\in\mathfrak{a}_1\cap\mathfrak{a}_2$, we can then write

$$c = c(a+b) = ca+cb\in\mathfrak{a}_1\mathfrak{a}_2. \qquad \square$$

This completes the proof in this case. In general, we apply induction to conclude that

$$\mathfrak{a}_1\dots\mathfrak{a}_n = (\mathfrak{a}_1\cap\cdots\cap\mathfrak{a}_{n-1})\,\mathfrak{a}_n.$$

To complete the proof, it suffices to verify that $\mathfrak{a}_1\dots\mathfrak{a}_{n-1}$ and $\mathfrak{a}_n$ are relatively prime. If for each $i\in\{1,\dots,n-1\}$ we find $x_i\in\mathfrak{a}_i$ and $y_i\in\mathfrak{a}_n$ such that $x_i+y_i=1$, then

$$1 = (x_1+y_1)\dots(x_{n-1}+y_{n-1})\in x_1\dots x_{n-1}+\mathfrak{a}_n\subset\mathfrak{a}_1\dots\mathfrak{a}_{n-1}+\mathfrak{a}_n,$$

which verifies what was needed.

Chapter 2

Divisibility in Commutative Rings

In this chapter, all rings are commutative unless stated overwise. Our goal is to study the *divisibility theory* of a ring. Given a ring A and $a_1, a_2 \in A$, we say a_1 *divides* a_2, written $a_1 \mid a_2$ if there is $c \in A$ such that $a_2 = ca_1$. Our goal is to build a theory of divisibility in commutative rings. This is clearly related to the theory of ideals in a ring, since the set of factors of an element x in a ring is precisely the ideal (x), and a factorization $x = x_1 \dots x_n$ with $x_1, \dots, x_n$ is equivalent to the condition that $(x) = (x_1) \cdots (x_n)$.

There are two useful assumptions on the rings we study which can make our lives easier; a minor assumption is that one is working with an *integral domains*, i.e. a nonzero ring without zero divisors. Such rings are also called *entire*. A much more powerful assumption is that all ideals are principal; such a ring is called a *principal ideal domain*, and then the divisor theory is essentially exactly the same as the theory of divisors. For instance, if A is a ring, then for a family of elements $a_1, \dots, a_n \in A$, we say $a \in A$ is a *greatest common divisor* if $a \mid a_1, \dots, a_n$, and if any $x \in A$ divides each of $a_1, \dots, a_n$, then $x \mid a$. A *least common multiple* of $a_1, \dots, a_n \in A$ is $a \in A$ such that $a_1, \dots, a_n \mid a$, and if $x \in A$ is any element divisible by all of $a_1, \dots, a_n$, then $a \mid x$. Greatest common divisors and least common divisors exist in any principal ideal domain; given $a_1, \dots, a_n \in A$, if $(a) = (a_1, \dots, a_n)$, then a is the greatest common divisor of $a_1, \dots, a_n$, and if $(a) = (a_1) \cap \cdots \cap (a_n)$, then a is the least common multiple of $a_1, \dots, a_n$.

An ideal $\mathfrak{p}$ of a ring A is *prime* if $\mathfrak{p} \neq A$, and $ab \in \mathfrak{p}$ implies $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$. This mimics the definition of prime integers as those integers p such that

if p divides the product of two integers nm, then p divides either n or m. It is clear from the definition that $\mathfrak{p}$ is a prime ideal if and only if $A/\mathfrak{p}$ is an integral domain.

Theorem 2.1. *If $f : A \to B$ is a homomorphism, and $\mathfrak{p}$ is a prime ideal of B, then $f^{-1}(\mathfrak{p})$ is a prime ideal of A.*

Proof. If $ab \in f^{-1}(\mathfrak{p})$, then $f(a)f(b) \in \mathfrak{p}$, hence either $f(a)$ or $f(b)$ is in $\mathfrak{p}$, which implies either $a \in f^{-1}(\mathfrak{p})$ or $b \in f^{-1}(\mathfrak{p})$. □

Remark. This theorem implies that if A is a subring of a ring B, and $\mathfrak{p}$ is a prime ideal of B, then $\mathfrak{p} \cap A$ is a prime ideal of A.

One of the tenants of commutative ring theory is that one can obtain powerful control over a ring by understanding it's prime ideals. Consider the following example.

Theorem 2.2. *If all prime ideals of a ring A are principal, then A is principal.*

Proof. Let $\mathcal{I}$ be the set of ideals in A which are not principal. If we have an infinite chain of ideals $\{\mathfrak{a}_\alpha\}$ contained in $\mathcal{I}$, then $\bigcup \mathfrak{a}_\alpha$ is an ideal; if this ideal was principle it would be generated by some $a \in \mathfrak{a}_{\alpha_0}$ for some index α_0. But this would imply that the chain eventually terminated. Thus Zorn's lemma implies that there is a maximal ideal $\mathfrak{a}$ of $\mathcal{I}$. Since $\mathfrak{a}$ is not principal, it cannot be prime, so we can find $x, y \notin \mathfrak{a}$ such that $xy \in \mathfrak{a}$. Let $\mathfrak{a}_x = (x) + \mathfrak{a}$, $\mathfrak{a}_y = (y) + \mathfrak{a}$, and let $\mathfrak{b}$ be the set of all $a \in A$ such that $(a)\,\mathfrak{a}_x \subset \mathfrak{a}$. Then $\mathfrak{a}_y \subset \mathfrak{b}$, so $\mathfrak{a}_x$, $\mathfrak{a}_y$, and $\mathfrak{b}$ are all principal ideals, generated by some elements x_0, y_0, and z_0 in A respectively. Thus z_0 divides y_0, and $x_0 z_0 \in \mathfrak{a}$. Since $\mathfrak{a} \subset \mathfrak{a}_x$, if $a \in \mathfrak{a}$, we can write $a = tx_0$ for some $t \in A$. But then $t\,\mathfrak{a}_x \subset \mathfrak{a}$, so $t \in \mathfrak{b}$. Thus $a = sx_0 z_0$ for some $s \in A$. Thus we conclude that $\mathfrak{a} = (x_0 z_0)$, which gives a contradiction. Thus $\mathcal{I}$ must be empty, which implies all ideals in A are principal. □

2.1 Maximal Ideals

An ideal $\mathfrak{m}$ is *maximal* if $\mathfrak{m} \neq A$, and there is no ideal strictly containing $\mathfrak{m}$ except the entire ring. and there is no ideal containing it but the entire ring, and the ideal itself is not the entire ring. Using Zorn's lemma in the classical manner, one may verify that any proper ideal of a ring is contained in some maximal ideal. The most useful fact about maximal

ideals to use in basic proofs is to use the fact that if $a \notin \mathfrak{a}$, then $(a) + \mathfrak{m} = A$. Thus (0) is a prime ideal if and only if A is entire to begin with.

Example. *The maximal ideals of* $\mathbf{Z}$ *are* $p\mathbf{Z}$*, where* p *is a prime number.*

Theorem 2.3. *If* A *is unital, then every maximal ideal is prime.*

Proof. Suppose $\mathfrak{m}$ is maximal, and let $ab \in \mathfrak{m}$. If $a \notin \mathfrak{m}$, then $(a) + \mathfrak{m} = A$, and so we can write $xa + m = 1$ for some $x \in A$, $m \in \mathfrak{m}$. But this implies that $b = 1 \cdot b = xab + mb \in \mathfrak{m}$. □

Remark. This statement is not true for non unital rings. For instance, one maximal ideal in the non unital ring $2\mathbf{Z}$ is $4\mathbf{Z}$, yet $2\mathbf{Z}/4\mathbf{Z}$ contains zero divisors, so $4\mathbf{Z}$ is not prime.

Example. *Every prime ideal in a Boolean ring with identity is maximal. Let* A *be a Boolean ring, and* $\mathfrak{a}$ *a prime ideal. Then* $A/\mathfrak{a}$ *is a Boolean integral domain. Since* $a^2 - a = 0$ *for all* $a \in A$*, this implies that for each* $a \in A$*, either* $a \in \mathfrak{a}$*, or* $a - 1 \in \mathfrak{a}$*. Since* $\mathfrak{a} \neq A$*,* $A/\mathfrak{a}$ *is a nontrivial integral domain with two elements, and thus a field, so* $\mathfrak{a}$ *is maximal.*

Theorem 2.4. *If* A *is unital, then an ideal* $\mathfrak{m}$ *is maximal iff* $A/\mathfrak{m}$ *is a field.*

Proof. Suppose $\mathfrak{m}$ is maximal, and $a \notin \mathfrak{m}$. Then $(a) + \mathfrak{m} = A$, and so we can write $xa + m = 1$, which implies that $xa \cong 1$ modulo $\mathfrak{m}$. This verifies that all nonzero residues in the quotient ring have inverses. On the other hand, the third isomorphism theorem says there is a one to one correspondence between ideals in $A/\mathfrak{m}$ and ideals in A containing $\mathfrak{m}$. If $A/\mathfrak{m}$ is a field, then the only ideals are (0) and (1), implying that the only ideals containing $\mathfrak{m}$ are $\mathfrak{m}$ and A. Thus $\mathfrak{m}$ is maximal if $A/\mathfrak{m}$ is a field. □

Over non-commutative rings, this need not be the case, for instance, the only maximal ideal of the ring of $n \times n$ matrices $M_n(\mathbf{C})$ is (0), but $M_n(\mathbf{C})$ is not a field, nor even a division algebra.

Example. *The ideal* $(2, X)$ *is a maximal ideal of* $\mathbf{Z}[X]$*, since*

$$\mathbf{Z}[X]/(2, X) \cong \mathbf{Z}_2$$

which is a field.

As one may have noticed, over a principal ideal domain, *all* prime ideals are maximal.

Theorem 2.5. *If A is a unital P.I.D, then every prime ideal is maximal.*

Proof. Let $p \in A$ and suppose (p) is a prime ideal of A. Then (p) is contained in some maximal ideal (m), and thus m divides p, so we can write $p = xm$ for some $x \in A$. But then either p divides x or p divides m. If p divides m, then $(p) = (m)$, and the theorem is proved. Otherwise, we can write $x = py$ for some $y \in A$. Thus $p = pym$, hence $1 = ym$, implying that m is a unit, which is impossible. □

2.2 Euclidean Domains

If A is an integral domain, a *Euclidean function* is a positive integer valued function N on non-zero values of A such that if $a, b \neq 0$, then there is q, r such that $a = qb + r$, where $r = 0$ or $N(r) < N(q)$. A *Euclidean domain* is an integral domain possessing a Euclidean function. For convinience, we define $\text{ord}(0) = -\infty$.

Example. *The integers* $\mathbf{Z}$ *is a Euclidean domain, with order function*

$$N(n) = |n|.$$

Indeed, given integers n, m, which we both may assume without loss of generality to be positive, if r is the smallest positive integer such that there is k such that $n = km + r$, then $0 \leqslant r < m$, since if $r > m$, then $n = (k+1)m + (r-m)$, and $0 \leqslant r - m \leqslant r$.

Example. *If k is a field, then $k[X]$ is a Euclidean domain, with order function*

$$N(f) = \deg(f).$$

If $f(X) = a_0 + \cdots + a_n X^n$, $g(X) = b_0 + \cdots + b_m X^m$, and $n \geqslant m$, then we can write

$$f(X) = b_m^{-1} X^{n-m} g(X) + (f - b_m^{-1} X^{n-m} g(X))$$

and $N(f - b_m^{-1} X^{n-m} g(X)) < n$, so continuing this process inductively results in the required remainder.

Example. *The ring $\mathbf{Z}[i]$ of Gaussian integers of the form $n + im$, where $n, m \in \mathbf{Z}$, is a Euclidean domain if we define $N(z) = |z|^2$. To verify this, given nonzero*

$z, w \in \mathbf{Z}[i]$ with $|z| \geqslant |w|$, pick $u \in \{w, iw, -w, -iw\}$ with an angle of $\theta \leqslant \pi/4$ with z. Then

$$\begin{aligned} |z-u|^2 &= |z|^2 + |u|^2 - 2\langle z, u\rangle \\ &\leqslant |z|^2 + |w|^2 - 2\cos(\theta)|z||w| \\ &\leqslant |z|^2 + |w|^2 - \sqrt{2}\cdot|z||w| \\ &\leqslant |z|^2 + (|w| - \sqrt{2}\cdot|z|)|w| \leqslant |z|^2 - \sqrt{2}|w|^2 \end{aligned}$$

and so $N(z-u) < N(z)$. Applying this process inductively thus gives the required decomposition.

Example. *If K is a field, a* discrete valuation *on K is a surjective homomorphism $\nu : U(K) \to \mathbf{Z}^+$ such that if $x + y \neq 0$, then $\nu(x+y) \geqslant \min(\nu(x), \nu(y))$. Then the set*

$$A = \{x \in K : \nu(x) \geqslant 0\}$$

is a subring of K, called a discrete valuation ring*. It forms a Euclidean domain under the map ν, for if $\nu(x) \leqslant \nu(y)$, then there exists $c \in K$ such that $y = cx$, and then $\nu(c) = \nu(y) - \nu(x) \geqslant 0$, so $c \in A$. This fact identifies A as a* local ring*, since it shows that A has a maximal ideal $\mathfrak{m} = \{x \in A : \nu(x) \geqslant 1\}$. More generally, the only nontrivial ideals in A are of the form $\mathfrak{m}^n$ for some $n \geqslant 1$.*

Theorem 2.6. *If A if a Euclidean domain, then A is principal.*

Proof. We mimic the proof that $\mathbf{Z}$ is principal. Let $\mathfrak{a}$ be a nonzero ideal in A, and let a be an element of smallest order. If $b \in \mathfrak{a}$, then we can write $b = qa + r$, where $N(r) < N(a)$. But $r = b - qa \in \mathfrak{a}$, so $r = 0$, and so a divides b. □

The fact that $\mathbf{Z}$ was a principal ideal domain was known since the time of the Greeks. Gauss was the first to realise that $\mathbf{Z}[i]$ was a principal ideal domain, and he used it to prove some interesting results about the ordinary integers, solving congruences modulo primes.

It is not true that every principal ideal domain is a Euclidean domain. However, principal ideal domains do exist a less powerful version of a norm, known as a *Dedekind Hasse* norm, which often means that working with Euclidean domains is not that much more powerful than working with principal ideal domains. If A is a ring, a Dedekind Hasse norm is a positive integer valued function N on non-zero values of A such that for any $x, y \in A$, either x divides y, or there is $s, t \in A$ such that $sx + ty \neq 0$ and $N(sx + ty) < N(x)$.

Theorem 2.7. *If A has a Dedekind-Hasse norm, then A is principal.*

Proof. Given an ideal $\mathfrak{a}$ in A, let x be a nonzero element of $\mathfrak{a}$ with smallest norm. Then for any $y \in \mathfrak{a}$, any element of $s, t \in A$, either $sx + ty \neq 0$ or $N(sx + ty) \geqslant x$. Since N is a Dedekind Hasse norm, this implies that x divides y, so that $\mathfrak{a} = (x)$. □

Example. *The ring $A = \mathbf{Z}[(1 + \sqrt{-19})/2]$ has a Dedekind-Hasse norm, from which it follows the ring is principal. We define*

$$N(z) = |z|^2.$$

Then $N(z)$ is a positive integer for each $z \in A$. Indeed,

$$\begin{aligned}\left(n + m\frac{1+\sqrt{-19}}{2}\right)\left(n + m\frac{1-\sqrt{-19}}{2}\right) &= (n + m/2)^2 + 19(m^2/4) \\ &= n^2 + nm + 5m^2.\end{aligned}$$

We claim that N is a Hasse Dedekind norm, hence A is principal. Suppose α, β are nonzero elements of A, but β does not divide α in A. We must show that there are $s, t \in A$ such that $0 < N(st + xy) < N(x)$. This is equivalent to showing

$$0 < N\left((\alpha/\beta)s - t\right) < 1.$$

Find $a, b, c \in \mathbf{Z}$ with no common divisor such that

$$\alpha/\beta = \frac{a + b\sqrt{-19}}{c}.$$

Then $c > 1$, for otherwise $\alpha/\beta \in A$ and so β divides α. Then we can find integers $x, y, z \in \mathbf{Z}$ such that $ax + by + cz = 1$. Write $ay - 19bx = cq + r$ for integers q, r with $|r| \leqslant c/2$, and let $s = y + x\sqrt{-19}$ and $t = q - z\sqrt{-19}$. Then

$$\begin{aligned}(\alpha/\beta)s - t &= \frac{(a + b\sqrt{-19})(y + x\sqrt{-19})}{c} - (q - z\sqrt{-19}) \\ &= \frac{ay - 19bx - cq}{c} + \frac{(ax + by + cz)\sqrt{-19}}{c} \\ &= \frac{r}{c} + \frac{\sqrt{-19}}{c}.\end{aligned}$$

This shows $(\alpha/\beta)s - t \neq 0$, *and*

$$N((\alpha/\beta)s - t) = \frac{r^2 + 19}{c^2} \leqslant 1/4 + 19/c^2.$$

Provided $c \geqslant 5$, *this completes the calculation, and so we address the remaining cases on a case by case basis for* $c \in \{2, 3, 4\}$.

Suppose $c = 2$. *Then* a *and* b *cannot be both even or both odd, for then* $\alpha/\beta \in A$ *(*A *consists precisely of* $(n + m\sqrt{-19})/2$ *such that* $n - m$ *is even). But then*

$$\frac{(a-1) + b\sqrt{-19}}{2} \in A$$

and

$$\frac{\alpha}{\beta} - \frac{(a-1) + b\sqrt{-19}}{2} = \frac{1}{2}$$

So we can set $s = 1$ *and* $t = [(a-1) + b\sqrt{-19}]/2$.

Now consider the case $c = 3$. *Then* $a^2 + 19b^2$ *is not divisible by 3, for otherwise* $a^2 + b^2$ *is divisible by 3, which implies* a *and* b *are divisible by* 3. *Write* $a^2 + 19b^2 = 3q + r$, *where* $r = 1$ *or* $r = 2$. *Then*

$$\frac{a + b\sqrt{-19}}{3}(a - b\sqrt{-19}) - q = r/3 < 1.$$

so we may pick $s = a - b\sqrt{-19}$ *and* $t = q$.

Finally, we consider the case $c = 4$. *Then* a *and* b *are not both even. If* a *and* b *are both odd, then* $a^2 + 19b^2$ *is congruent to 4 modulo 8. Thus we can write* $a^2 + 19b^2 = 8q + 4$. *Then*

$$\frac{a + b\sqrt{-19}}{4}\frac{a - b\sqrt{-19}}{2} - q = \frac{1}{2}$$

so we can set $s = (a - b\sqrt{-19})/2$ *and* $t = q$. *If one of* a *and* b *is odd, and the other is even, then* $a^2 + 19b^2$ *is odd, then we can write* $a^2 + 19b^2 = 4q + r$ *with* $0 < r < 4$. *Then*

$$\frac{a + b\sqrt{-19}}{2}(a - b\sqrt{-19}) - q = r/4.$$

Thus we can set $s = a - b\sqrt{-19}$ *and* $t = q$.

We claim $\mathbf{Z}[(1+\sqrt{-19})/2]$ is a PID which is not a Euclidean domain. To begin with, we must find a property that Euclidean domains have but which PIDs do not necessarily have. Given a ring A, let $V(A) = U(A) \cup \{0\}$. An element $u \in A - V(A)$ is called a *universal side divisor* if for every $x \in A$ there exists $a \in V(A)$ such that u divides $x - a$.

Theorem 2.8. *If A is a Euclidean domain, it possesses universal side divisors.*

Proof. Let u be the element of $A - V(A)$ with smallest norm. Then u is a universal side divisor, because if $x \in A$, we can write $x = cu + r$, where $N(r) < N(u)$, so $r \in V(A)$. Thus u divides $x - r$. □

Example. *Recall the ring $A = \mathbf{Z}[(1+\sqrt{-19})/2]$ is a principal ring. We now show it is not a Euclidean domain, because it does not possess universal side divisors. Recall the Hasse Dedekind norm N constructed earlier. Since N is multiplicative, it is easy to see that the units of A are precisely ± 1, since the only integer solutions to the Diophantine equation*

$$n^2 + nm + 5m^2 = 1$$

are given by setting $n = \pm 1$, $m = 0$. However, A does not have a universal side divisor, from which it follows that A cannot be a Euclidean domain. Suppose u is a universal side divisor. Then u is a side divisor of 2, so u divides one of $\{1, 2, 3\}$. Since u is not a unit, u must either divide 2 or 3. If $2 = uv$, then $N(uv) = N(u)N(v) = 4$, which implies $u, v \in \mathbf{Z}$ since $N(u) \geqslant 5$ for any $u \notin \mathbf{Z}$, and so without loss of generality we find $u = 2$. Similarily, if u divides 3, with $3 = uv$, then $N(uv) = 9$. Since there are no $u \in A$ with $N(u) = 3$, we conclude $N(u) = 9$. Note that for any integers n, m,

$$n^2 + nm + 5m^2 \geqslant \min(5m^2, n^2 + 4m^2) \geqslant 4m^2,$$

we can see that the only solutions to

$$n^2 + nm + 5m^2 = 9$$

have $|m| \leqslant 1$. If $m = 0$, then $n = 3$. If $m = 1$, then $n(n+1) = 4$, which has no solutions. If $m = -1$, then $n(n-1) = 4$, which again has no solutions. Thus we conclude that if u is a side divisor of A, then $u = \pm 3$ or $u = \pm 2$. But none of

$$\frac{-1+\sqrt{19}}{2}, \quad \frac{1+\sqrt{-19}}{2}, \quad \text{or} \quad \frac{3+\sqrt{-19}}{2}$$

are divisible by 2 or 3, which shows u is not a side divisor. Thus A is not a Euclidean domain.

Let us now show that all principal ideal domains have Dedekind-Hasse norms.

Theorem 2.9. *Any principal ideal domain has a Dedekind-Hasse norm.*

Proof. If A is a principal ideal domain, then we can define a norm N by setting $N(u) = 1$ if $u \in U(A)$, and if $p_1, \ldots, p_n$ are primes, then define

$$N(p_1^{k_1} \ldots p_n^{k_n}) = 2^{k_1 + \cdots + k_n}.$$

Then N is multiplicative and positive. Now suppose a and b are nonzero elements of A. Then $(a, b) = (x)$ for some x. If

$$a = p_1^{t_1} \ldots p_n^{t_n}$$

and

$$b = u p_1^{s_1} \ldots p_n^{s_n},$$

Then we can choose

$$x = p_1^{\min(t_1, s_1)} \ldots p_n^{(t_n, s_n)}.$$

If a does not divide b, then there must exist some i with $s_i < t_i$, and then $N(x) < N(a)$. Thus we have shown N is a Dedekind-Hasse norm. □

An integral domain A is a *quasi-Euclidean domain* if there exists a function N such that for any $r_{-1}, r_0 \in A$, there exists k and $q_0, \ldots, q_{k-1} \in A$ and $r_1, \ldots, r_k \in A$ such that for each $i \in \{-1, \ldots, k-2\}$, $r_i = q_{i+1} r_{i+1} + r_{i+2}$, and $N(r_k) < N(r_0)$. In the fraction field of A, we may write this sequence in terms of a continued fraction, i.e. writing

$$\frac{r_{-1}}{r_0} = q_0 + \frac{1}{q_1 + \frac{1}{q_2 + \cdots + \frac{1}{q_{k-1} + r_k / r_{k-1}}}}.$$

We may write such a continued fraction as $[q_0, \ldots, q_{k-1}, r_{k-1}/r_k]$, where the bracket notation for continued fractions is defined inductively by setting $[a_1, a_2] = a_1 + 1/a_2$, and $[a_1, \ldots, a_k] = [a_1, [a_2, \ldots, a_k]]$. A ring is an n stage Euclidean domain if we can always choose $k \leqslant n$.

Lemma 2.10. *Any quasi-Euclidean domain is a Bezout domain.*

Proof. Fix $x_0, y_0 \in A$, and suppose r is the smallest nonzero element of (x_0, y_0) which respect to N. Then there is k such that we can perform the k stage Euclidean algorithm with $r_{-1} = x_0$ and $r_0 = y_0$, constructing $r_1, \ldots, r_k \in A$ and $q_1, \ldots, q_{k-1}$ such that $r_i = q_{i+1} r_{i+1} + r_{i+2}$ for each i, and $N(r_k) < N(y)$. One verifies quite easily that $(r_i, r_{i+1}) = (r_{i+1}, r_{i+2})$. Thus $(x_0, y_0) = (x_1, y_1)$, where $x_1 = r_{k-1}$ and $y_1 = r_k$. Thus $N(y_1) < N(y_0)$ or $y_1 = 0$. If $y_1 = 0$, then $(x_0, y_0) = (x_1)$. Otherwise, we may repeat this process, writing $(x_1, y_1) = (x_2, y_2)$ with $y_2 = 0$ or $N(y_2) < N(y_1)$. Clearly this process must terminate eventually, so that (x_0, y_0) is a principal ideal. □

2.3 Bezout Domains

An integral domain A is a *Bezout Domain* if every finitely generated ideal is principal. An inductive construction shows that this is equivalent to showing every ideal generated by two elements is principal, which means precisely that for every pair of elements $a_1, a_2 \in A$, the greatest common divisor of a_1 and a_2 can be written as $sa_1 + ta_2$ for some $s, t \in A$.

Example. *Every Boolean ring is a Bezout domain. Given a Boolean ring A, and $x, y \in A$, we have*

$$x(x + y - xy) = x^2 + xy - x^2 y = x + xy + xy = x$$

and

$$y(x + y - xy) = xy + y^2 - xy^2 = y + xy - xy = y.$$

Thus we conclude $(x, y) = (x + y - xy)$.

Example. *Let U be a connected, open subset of $\mathbf{C}$, and let $A(U)$ denote the family of all holomorphic functions on U. Then $A(U)$ is a Bezout domain. It is a consequence of the Weirstrass factorization theorem and the Mittag-Lefler theorem that if S is a discrete subset of U, and for each $s \in S$ we associate a natural number n_s and complex numbers $w_{s1}, \ldots, w_{sn_s}$, then there exists $f \in A(U)$ such that for all $s \in S$ and $k \leqslant n_s$, $f^{(k)}(s) = w_{sk}$, and the zeroes of f are contained in S. Moreover, if $f_1, f_2 \in A(U)$ are functions such that $ord_z(f_2) \geqslant ord_z(f_1)$ for all $z \in U$, then there is $g \in A(U)$ such that $f_2 = g f_1$. These analytical facts imply that $A(U)$ is a Bezout domain.*

To begin with, assume that $f_1, f_2 \in A(U)$ are holomorphic functions sharing no common zeroes. Then we can find a function $g_2 \in A(U)$ with specialized values and derivatives on the zeroes of f_1 such that all zeroes of f_1 are zeroes of $1 - f_2 g_2$, and the order of this zero for $1 - f_2 g_2$ is greater than f_1. It follows that there exists $g_1 \in A(U)$ such that $f_1 g_1 = 1 - f_2 g_2$, and so $(f_1, f_2) = 1$.

More generally, given f_1, f_2, we can find a function f with

$$ord_z(f) = \min(ord_z(f_1), ord_z(f_2))$$

for each $z \in \mathbf{C}$. Then there are $g_1, g_2 \in A(U)$ such that $f_1 = g_1 f$, $f_2 = g_2 f$. Thus $(f_1, f_2) \subset (f)$. But since g_1 and g_2 share no common zeroes, there is $h_1, h_2 \in A(U)$ such that $h_1 g_1 + h_2 g_2 = 1$, and so $f = (h_1 g_1 + h_2 g_2) f = h_1 f_1 + h_2 f_2 \in (f_1, f_2)$. Thus $(f_1, f_2) = (f)$, and we have shown all finitely generated ideals are principal.

We note that $A(U)$ is not *a unique factorization domain, which we will later see implies that $A(U)$ is not a principal ring. The units of $A(U)$ are precisely the functions with no zeroes, and the Weirstrass factorization theorem irreducible elements of $A(U)$ are precisely those elements with a single, simple zero. In particular this implies that any element of $A(U)$ which can be written as a finite product of irreducibles has only finitely many zeroes. On the other hand, the Weirstrass factorization theorem essentially says that all elements of $A(U)$ can be written as an* infinite product *of irreducible elements of $A(U)$, which is not measured in the algebraic theory of factorization.*

Example. *Let A be the subring of $\mathbf{Q}[X]$ consisting of polynomials with integer constant term. The units of A are precisely $\{\pm 1\}$. This means that the only irreducible elements of A are prime numbers, or irreducible polynomials in $\mathbf{Q}[X]$ with constant term 1. It is easy to see such elements are irreducible in A. Any nonconstant polynomial with constant term not equal to 1 is reducible. Conversely, if $f \in A$, and we can write $f = gh$, where $g, h \in \mathbf{Q}[X]$ are nonconstant polynomials. If $x = g(0)$ and $y = h(0)$, then $xy = 1$, so we can write $f = [g/x][xh]$, and $g/x, xh \in A$ since both have constant coefficient equal to 1, so f is reducible over A.*

In particular, this means that the polynomial X is not irreducible in A. On the other hand, the only irreducible factors of X in A are prime integers. Thus X cannot be factored into irreducibles, and thus A is not *a unique factorization domain. This implies the ring A is not Noetherian, and we can see this more explicitly from the infinite chain*

$$(X) \subsetneq (X/2) \supsetneq (X/4) \supsetneq \dots$$

is an infinite increasing chain of ideals. However, A is *a Bezout domain, as we now prove, so that all finitely generated ideals are principal.*

Let $f, g \in \mathbf{Q}[X]$, and suppose either $f(0)$ or $g(0)$ is nonzero. Let $h \in \mathbf{Q}[X]$ be a greatest common divisor of f and g, and scale h so that $\mathbf{Z}h(0) = \mathbf{Z}f(0) + \mathbf{Z}g(0)$. If we write $f = hf_0$ and $g = hg_0$, then $f_0, g_0 \in A$, since $f(0)$ and $g(0)$ are integer multiples of $h(0)$. Thus we conclude that $(f,g) \subset (h)$. On the other hand $f_0(0)$ and $g_0(0)$ must be relatively prime, because if $k_1, k_2 \in \mathbf{Z}$ are chosen such that $h(0) = k_1 f(0) + k_2 g(0)$, then this implies $1 = k_1 f_0(0) + k_2 g_0(0)$. We claim this implies that we can write $1 = af_0 + bg_0$ with $a, b \in A$, which would imply that $h = af + bg \in (f,g)$, so that $(h) \subset (f,g)$. Certainly we can write $1 = a_0 f_0 + b_0 g_0$ with $a_0, b_0 \in \mathbf{Q}[X]$. For any rational number m, if we write $a = a_0 + mg_0$ and $b = b_0 - mf_0$, then $1 = af_0 + bg_0$. And setting $m = (k_1 - a_0(0)/g_0(0)$ where k_1 is as in the last paragraph shows $a, b \in A$. Thus we conclude that $(f,g) = (h)$.

It is now easy to generalize this construction to the case where $f(0) = g(0) = 0$. For general $f, g \in A$, we can find an integer r such that $f = X^r f_0$, $g = X^r g_0$, and either $f_0(0)$ or $g_0(0)$ is nonzero. The previous techniques show the existence of $h_0 \in A$ such that $(f_0, g_0) = (h_0)$. If we let $h = X^r h_0$, we conclude that

$$(f,g) = (X^r)(f_0, g_0) = (X^r)(h_0) = (h),$$

which completes the general case.

Lemma 2.11. *Let K be the fraction field of a Bezout domain A. Then every element of K can be written as a/b where $a, b \in A$ and $(a) + (b) = A$.*

Proof. Any element of K can be written as x/y with $x, y \in A$, and suppose $(x,y) = (a)$ with a not a unit. Then we can find $x_1, y_1, t, s \in A$ such that

$$x = x_1 a, \quad y = y_1 a, \quad \text{and} \quad a = tx + sy.$$

Thus $a = a(tx_1 + sy_1)$, so $tx_1 + sy_1 = 1$. But then $(x_1, y_1) = (1)$, and $x/y = x_1/y_1$, completing the proof. □

Following the standard proof that PIDs are UFDs, we can conclude that any Bezout domain such that any element can be factored into irreducible elements is a factorial ring. In fact, this occurs if and only if the Bezout domain is a principal ideal domain.

Lemma 2.12. *Any factorial Bezout domain is principal.*

Proof. Let A be a factorial Bezout domain. For each $a \in A$, we can write a as $p_1^{k_1} \dots p_n^{k_n}$ for primes $p_1, \dots, p_n$. Define

$$N(a) = k_1 + \cdots + k_n.$$

We claim N is a Dedekind-Hasse norm. Indeed, if $a_1, a_2 \in A$, then there is $x \in A$ such that $(a_1, a_2) = (a)$. Then a is a greatest common divisor for a_1 and a_2. If

$$a_1 = p_1^{t_1} \dots p_n^{t_n} \quad \text{and} \quad a_2 = p_1^{s_1} \dots p_n^{s_n},$$

we therefore conclude that

$$a = p_1^{\min(t_1, s_1)} \dots p_n^{\min(t_n, s_n)}.$$

Thus either a_1 divides a_2, a_2 divides a_1, or $N(a) < \min(N(a_1), N(a_2))$. This verifies that N is a Dedekind-Hasse norm. □

2.4 Uniqueness of Congruences

In classical number theory, one takes a series of integers $k_1, \dots, k_m$ and values $a_1, \dots, a_m$, and asks to find an integer n such that $n \equiv a_i \pmod{k_i}$ for each i. The classical Chinese remainder theorem says that if the set of integers $\{k_i\}$ are pairwise coprime, such an element n can always be solved. These ideas can be extended to solve congruences over general rings. In the general setup, we are given a family of ideals $\mathfrak{a}_1, \dots, \mathfrak{a}_N$ over a commutative ring A, and we consider the corresponding projection $\pi : A \to A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_N$. The generalization of the Chinese remainder theorem is summarized in the next theorem. We saw two ideals $\mathfrak{a}$ and $\mathfrak{b}$ are *coprime* if $\mathfrak{a} + \mathfrak{b} = A$.

Theorem 2.13. *Let A be a commutative ring, fix $n \geqslant 2$, and suppose $\mathfrak{a}_1, \dots, \mathfrak{a}_n$ are pairwise coprime. Then the map $\pi : A \to A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n$ is surjective, and is an isomorphism if $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n = (0)$.*

Proof. Consider the case where we have two coprime ideals $\mathfrak{a}_1$ and $\mathfrak{a}_2$. Given any $x_1, x_2 \in A$, we wish to find $x \in A$ such that $x - x_1 \in \mathfrak{a}_1$ and $x - x_2 \in \mathfrak{a}_2$. Since $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are coprime, we can find $a_1 \in \mathfrak{a}_1$ and $a_2 \in \mathfrak{a}_2$ such that $a_1 + a_2 = x_2 - x_1$. Thus $\pi(a_1) = (0, x_2 - x_1)$ and $\pi(a_2) = (x_1 - x_2, 0)$. But this means that $\pi(x_1 + a_1) = (x_1, x_2)$. The kernel of the map π is $\mathfrak{a}_1 \cap \mathfrak{a}_2$, so the proof is completed in the case $n = 2$.

Let us now prove the remaining cases by induction. Suppose we could show that $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_{n-1}$ and $\mathfrak{a}_n$ are coprime. Then the Chinese remainder theorem for the case of two ideals, combined with the Chinese remainder theorem for $n-1$ ideals, would complete the proof. By an easy induction, it suffices to show that if $\mathfrak{a}, \mathfrak{b}$, and $\mathfrak{c}$ are pairwise coprime, then $\mathfrak{a} \cap \mathfrak{b}$ are coprime to $\mathfrak{c}$. Using the fact that $\mathfrak{a}$ and $\mathfrak{b}$ are relatively prime, and that $\mathfrak{b}$ and $\mathfrak{c}$ are relatively prime, we find

$$\begin{aligned}(\mathfrak{a} \cap \mathfrak{b}) + \mathfrak{c} &= \mathfrak{a}\mathfrak{b} + \mathfrak{c} \\ &= \mathfrak{a}\mathfrak{b} + (\mathfrak{a} + \mathfrak{b})\mathfrak{c} \\ &= \mathfrak{a}\mathfrak{b} + \mathfrak{a}\mathfrak{c} + \mathfrak{b}\mathfrak{c} \\ &= \mathfrak{a}(\mathfrak{b} + \mathfrak{c}) + \mathfrak{b}\mathfrak{c} \\ &= \mathfrak{a} + \mathfrak{b}\mathfrak{c}.\end{aligned}$$

Thus the ideal contains $\mathfrak{a}$ and $\mathfrak{c}$. But since $\mathfrak{a}$ and $\mathfrak{c}$ are relatively prime, we conclude $(\mathfrak{a} \cap \mathfrak{b}) + \mathfrak{c} = A$. □

Example. *Given an integer n, the units of $\mathbf{Z}_n$ are in one to one correspondence with the set of integers $1 \leqslant m \leqslant n$ which are relatively prime to n. Thus $\varphi(n) = \#(U(\mathbf{Z}_n))$, where φ is the number of such integers. If n and m are relatively prime, then $(n) + (m) = \mathbf{Z}$, and $(n) \cap (m) = (nm)$. Thus the Chinese remainder theorem applies, and we conclude that $\mathbf{Z}_{nm}$ is isomorphic to $\mathbf{Z}_n \times \mathbf{Z}_m$. If any pair of rings A and B, $U(A \times B) = U(A) \times U(B)$, so we conclude*

$$\phi(nm) = \#(U(\mathbf{Z}_{nm})) = \#(U(\mathbf{Z}_n) \times U(\mathbf{Z}_m)) = \phi(n)\phi(m).$$

This statement enables us to calculate $\phi(n)$ for any integer n. We note first that if p is prime, then

$$\varphi(p^n) = p^{n-1}(p-1),$$

because there are p^n integers between 1 and p^n, and they are all relatively prime to p^n, except for the multiples of p. Thus

$$\varphi(p_1^{n_1} \cdots p_m^{n_m}) = \prod_{i=1}^{m} \varphi(p_i^{n_i}) = \prod_{i=1}^{m} p_i^{n_i - 1}(p_i - 1).$$

Example. *For each $n, m \in \mathbf{Z}$, the map $f_m : \mathbf{Z}_n \to \mathbf{Z}_n$ given by setting $f_m(n) = nm$ is an endomorphism of $\mathbf{Z}_n$. It is clear that*

$$f_{n_1} \circ f_{n_2} = f_{n_1 n_2} \quad \textit{and} \quad f_{n_1 + n_2} = f_{n_1} + f_{n_2},$$

Thus we obtain a morphism from $\mathbf{Z}$ to $End(\mathbf{Z}_n)$, which is surjective since if $\varphi : \mathbf{Z}_n \to \mathbf{Z}_n$ is a morphism with $\varphi(1) = k$, then $\varphi = f_k$. The kernel of the morphism from $\mathbf{Z}$ to $End(\mathbf{Z}_n)$ is then clearly (m), so we have an isomorphism between $\mathbf{Z}_n$ and $End(\mathbf{Z}_n)$.

Example. *Let A be a nontrivial finite Boolean ring with identity. Let $I_1, \ldots, I_n$ be the set of all prime ideals of A. Since all prime ideals are maximal in Boolean rings, the ideals $\{I_1, \ldots, I_n\}$ are all pairwise coprime. Since A is a reduced ring, $I_1 \cap \cdots \cap I_n = (0)$. For each i, A/I_i is isomorphic to $\mathbf{Z}_2$, so the Chinese remainder theorem implies $A \cong A/I_1 \times \cdots \times A/I_n \cong \mathbf{Z}_2^n$. If A is a Boolean ring without identity, the unitization $\mathbf{Z} \times A$ has an ideal (2), and the quotient ring $\mathbf{Z}_2 \times A$ is a Boolean ring with identity. Thus $\mathbf{Z}_2 \times A$ is isomorphic to the Boolean ring of a finite set X, and so A is isomorphic to a maximal ideal of the Boolean ring of a finite set.*

2.5 Factorial Rings

Let A be a ring. An element $a \in A$ is *irreducible* if, given $x, y \in A$ such that $a = xy$, either x or y is a unit of A. A *factorial ring*, or *unique factorization domain*, is an integral domain A such that every nonzero $a \in A - U(A)$ can be written as $p_1 p_2 \ldots p_N$, for some irreducibles p_n, and such that if $p_1 \ldots p_N = q_1 \ldots q_M$, then $N = M$, and, after a permutation, each p_n differs from q_n by a unit.

Example. *The ring $\mathbf{Z}$ of integers is a factorial ring, as shown by the fundamental theorem of arithmetic. On the other hand, the ring $\mathbf{Z}[2i]$ is* not *factorial, since $2i$, $-2i$, and 2 are irreducible in $\mathbf{Z}[2i]$, so that*

$$4 = (2i) \cdot (-2i) = 2 \cdot 2$$

gives two distinct factorizations of 4. Similarily, $\mathbf{Z}[2\sqrt{2}]$ is not a factorial ring, since $8 = (2\sqrt{2})^2 = 2^3$.

Example. *Every field is trivially a unique factorization domain.*

We say a ring is *Noetherian* if every ideal is finitely generated (thus every principal ring is Noetherian); this is equivalent to saying the ring satisfies the *ascending chain condition*. That is, there do not exist an infinite

linear chain $\{\mathfrak{a}_\alpha\}$ of distinct ideals. This can be reworded by saying that every infinite chain of ideals

$$\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \dots$$

eventually becomes constant. Noetherian rings have a factorization theory, but this factorization need not be unique.

Theorem 2.14. *Every nonzero element of a Noetherian ring may be factored into irreducible elements.*

Proof. Fix some $x_0 \neq 0$. If x_0 is irreducible, we're done. Otherwise, we can write $x_0 = a_0 x_1$, where a and x_1 are both not units. If x_1 is not irreducible, we can write $x_1 = a_1 x_2$, where neither $a_1 x_2$ are not units. It is clear that if this process never stops, we can find a sequence $\{x_i\}$ with $x_{i+1} \mid x_i$ for each i, but such that x_i does not divide x_{i+1}. This corresponds to an infinite chain

$$(x_0) \subsetneq (x_1) \subsetneq (x_2) \subsetneq \dots$$

which is impossible since the ring is Noetherian. □

The fact that every ideal in $k[X_1, \dots, X_n]$ is finitely generated was discovered by Hilbert. However, Emmy Noether was one of the first to identify this property with the important commutative algebra has shown it to be.

Theorem 2.15. *If R is a Noetherian ring, then $R[X]$ is a Noetherian ring.*

Proof. For any ideal $\mathfrak{a}$ in $R[X]$, we can consider the *leading coefficient ideal* $\mathfrak{b}$ consisting of all elements of R which form the leading coefficients of polynomials in $\mathfrak{a}$, together with the zero element. The set $\mathfrak{b}_d$ of all elements of R which form the leading coefficients of polynomials in $\mathfrak{a}$ of degree d also forms an ideal, and $\mathfrak{b} = \lim \mathfrak{b}_d$. Since R is Noetherian, $\mathfrak{b} = \mathfrak{b}_N$ for some sufficiently large N. Each of the $\mathfrak{b}_1, \dots, \mathfrak{b}_N$ is finitely generated, and we can therefore consider a family of polynomials $f_1, \dots, f_M$ such that for each $b \in \mathfrak{b}_n$, there exists polynomials g_n such that $\sum f_n g_n$ has degree n, and has leading coefficient b. We claim that the $f_1, \dots, f_M$ generate $\mathfrak{a}$. If f is an element of minimal degree d in $\mathfrak{a} - (f_1, \dots, f_M)$, then the leading coefficient of f is an element of $\mathfrak{b}_d$, and so there exists polynomials g_n such that $\sum f_n g_n$ has degree n and has the same leading coefficient as g. In particular, this means that $f - \sum f_n g_n$ has degree less than f, implying it

cannot be an element of $\mathfrak{a} - (f_1, \ldots, f_M)$. But it certainly an element of $\mathfrak{a}$, hence an element of $(f_1, \ldots, f_M)$, but this implies that

$$f = (f - \sum f_n g_n) + \sum f_n g_n \in (f_1, \ldots, f_M) + (f_1, \ldots, f_M) = (f_1, \ldots, f_M)$$

which is impossible. □

The ideal of leading coefficients of an ideal $\mathfrak{a}$ in $R[X]$ is extremely useful to understanding $\mathfrak{a}$, and the higher dimensional analysis of such coefficients leads to a very rich area of computable operations on ideals of polynomial rings, known as the theory of Gröbner bases.

Theorem 2.16. *Every principal ideal domain is factorial.*

Proof. The fact that every principal entire ring *has* a factorization is justified because it is Noetherian. It now suffices to prove such a factorization is unique. Let $p_1 \ldots p_N = q_1 \ldots q_M$. We proceed by induction on N. If $p = q_1 \ldots q_M$, then p divides one of the quantities on the right, implying p must divide one of the q_n, which, without loss of generality, we may assume is q_M. Then $q_M = ap$, so, dividing by p on both sides of the equation, we conclude that $1 = aq_1 \ldots q_{M-1}$, so each q_n is a unit, which is a contradiction unless $M = 1$. Now in general, suppose $p_1 \ldots p_{N+1} = q_1 \ldots q_M$. Then p_{N+1} divides one of the quantities on the right, say q_M, so $q_M = ap_{N+1}$, hence, dividing out, we conclude $p_1 \ldots p_N = aq_1 \ldots q_{M-1}$, hence by induction, $N = M - 1$, and by permutation, we can assume $p_n = a_n q_n$. But then $p_1 \ldots p_{N+1} = aa_1 \ldots a_{M-1} p_1 \ldots p_N q_M$, hence $p_{N+1} = aa_1 \ldots a_{M-1} q_M$, so p_{N+1} differs from q_M by a unit. □

The *primes* of a factorial ring can be broken up into equivalence classes, where we identify two primes that differ by a unit. Thus if A is a factorial ring, and $\{p_\alpha\}$ is a family of representatives for the irreducible elements of A modulo $U(A)$, then any nonzer $a \in A$ may be *uniquely written as*

$$a = u p_{\alpha_1}^{k_1} \ldots p_{\alpha_n}^{k_n}$$

where $u \in U(A)$.

Example. Z *is a principal ideal domain, hence factorial. The group of units are* 1 *and* -1*, so the equivalence class of irreducibles in* **Z** *consist of integers* p *and* $-p$*, where* p *is a non-negative prime. It is canonical to take the positive*

primes as representatives, and so we find every positive integer can be uniquely decomposed as a product of positive prime number, and every negative integer is the negation of a product of primes.

If A is a ring of functions, then that ring being non-factorial normally indicates the presence of some singularity in the ring.

Example. *Consider the curve in* $\mathbf{A}^2 = K^2$ *defined as the locus of points satisfying the equation* $Y^2 = X^3$*. Then the curve has a singularity at the origin. The corresponding ring* $K[X,Y]/(Y^2 - X^3)$*, which is isomorphic to the ring of polynomial functions in two dimensions restricted to the curve, is not factorial, since* $X \neq Y$ *are both irreducible elements not differing by primes, yet* $X^3 = Y^2$*.*

Example. *The relation*

$$\sin^2 x = (1 + \cos x)(1 - \cos x)$$

indicates that the ring $\mathbf{R}[\sin x, \cos x]$ *of functions generated by* $\sin x$ *and* $\cos x$ *is not a factorial ring. Define the* trigonometric degree $\deg(f)$ *of a function*

$$f(x) = a + b_1 \cos(x) + \cdots + b_N \cos(Nx) + c_1 \sin(x) + \cdots + c_M \sin(Mx)$$

where $b_N, c_M \neq 0$*, as* $\max(N,M)$*. One can show* $\deg(fg) = \deg(f) + \deg(g)$ *directly, but it is convenient to extend our calculations to the complex algebra* $\mathbf{C}[e^{ix}]$*, which contains* $\mathbf{R}[\cos x, \sin x]$ *as a subring. We do this in the next paragraph, but note that this implies that* $\mathbf{R}[\sin x, \cos x]$ *has no zero divisors, and all degree one trigonometric polynomials are irreducible, which includes* $\sin x$*,* $1 + \cos x$*, and* $1 - \cos x$*.*

To see that the degree is well defined on $\mathbf{C}[e^{ix}]$*, it suffices to note that* e^{ix} *is trancendental over* $\mathbf{R}$*, so* $\mathbf{C}[e^{ix}]$ *is isomorphic to* $\mathbf{C}[X]$*. To see that* e^{ix} *is trancendental, we note that*

$$\int_{-\pi}^{\pi} e^{nix} e^{-mix} = \begin{cases} 2\pi & n = m \\ 0 & n \neq m \end{cases}$$

Thus, given $f(x) = \sum_{n=-N}^{N} a_n e^{nix}$*,*

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-nix}.$$

so if $f = 0$*, then* $a_n = 0$ *for all* n*.*

Lemma 2.17. *In an integral domain, every prime is irreducible.*

Proof. Suppose A is an integral domain, and suppose $p \in A$ is prime. Suppose $a, b \in A$ and $p = ab$. Since p is prime, either p divides a or p divides b. If we write $a = pa'$, and then $1 = a'b$, implying b is a unit. Thus p is irreducible. □

Conversely, in a factorial ring, every irreducible element is prime.

Theorem 2.18. *In a factorial ring, every irreducible element is prime.*

Proof. Suppose that A is a factorial ring with an equivalence class of irreducible elements $\{a_\alpha\}$, and fix some irreducible element a_{α_1} from this class. Suppose $x_1, x_2 \in A$ and a_{α_1} divides $x_1 x_2$. Then we can write $x_1 x_2 = x_3 a_{\alpha_1}$. Perform a factorization of each element, writing

$$x_1 = u_1 a_{\alpha_1}^{t_{11}} \dots a_{\alpha_n}^{t_{1n}},$$

$$x_2 = u_2 a_{\alpha_1}^{t_{21}} \dots a_{\alpha_n}^{t_{2n}},$$

and

$$x_3 = u_3 a_{\alpha_1}^{t_{31}} \dots a_{\alpha_n}^{t_{3n}}.$$

Thus we have

$$(u_1 u_2) a_{\alpha_1}^{t_{11}+t_{21}} \dots a_{\alpha_n}^{t_{1n}+t_{2n}} = u_3 a_{\alpha_1}^{1+t_{31}} \dots a_{\alpha_n}^{t_{3n}}.$$

Unique factorization implies that $t_{11} + t_{21} = 1 + t_{31}$, thus either $t_{11} > 0$ or $t_{21} > 0$. In particular, this implies that either a_{α_1} divides x_1 or a_{α_1} divides x_2. Thus a_{α_1} is prime. □

Remark. Suppose that A is a ring which every element can be uniquely factored (up to units) as a product of primes, but not necessarily uniquely as a product of irreducibles. If $a \in A$ is irreducible, and we write

$$a = u p_1^{k_1} \dots p_n^{k_n}$$

then we conclude that $n = 1$, $k_1 = 1$, and p_1 differs from a by a unit. Thus a is prime, and so A is really a factorial ring.

In a factorial domain A, the primes and units of the ring uniquely specify the multiplicative monoid structure of $A - \{0\}$. For instance, if $\mathbf{P} = \{2, 3, 5, \dots\}$ is the set of all primes in $\mathbf{Z}$, then for each permutation π of $\mathbf{P}$, we obtain a homomorphism $f : \mathbf{Z} - \{0\} \to \mathbf{Z} - \{0\}$ of the multiplicative monoid structure by setting

$$f(\pm p_1^{k_1} \dots p_n^{k_n}) = \pm \pi(p_1)^{k_1} \dots \pi(p_n)^{k_n}.$$

On the other hand, none of these permutations extend to ring homomorphisms, since the additive structure isn't reflected at all in these permutations.

Chapter 3

Polynomials

In a ring, we can add and multiply. It is natural then, to 'solve' equations of the form

$$5X^2 + 1 = 2 \qquad XYZ + 2Y = Z$$

Making an abstract concept into a precise mathematical object is often the key method to study mathematical phenomena. A polynomial is the static object representing the equations we can construct in a ring, which we can pin down and understand. In this Chapter, all rings will be assumed to be commutative unless stated otherwise.

3.1 Univariate Polynomials

We now provide a brief introduction to the ring of polynomials with coefficients in a ring. If A is a commutative ring, a *univariate polynomial* in the indeterminate X with coefficients in A is an abstract expression of the form $f(X) = a_0 + a_1X + \cdots + a_nX^n$, with $a_0, \ldots, a_n \in A$. The set of all univariate polynomials in X is denoted $A[X]$. We define a ring structure on $A[X]$ by letting

$$\sum a_k X^k + \sum b_k X^k = \sum (a_k + b_k) X^k$$

$$\left(\sum a_i X^i\right)\left(\sum b_j X^j\right) = \sum a_i b_j X^{i+j} = \sum_k \left(\sum a_i b_{k-i}\right) X^k$$

Since A embeds itself in $A[X]$ as the set of terms with no occurence of X, we can view $A[X]$ as an algebra over A.

If A is a subring of a ring B, then each polynomial

$$f = a_0 + a_1 X + \cdots + a_n X^n \in A[X]$$

gives rise to a function from B to itself, mapping $x \in B$ to

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \in B.$$

This gives a homomorphism from $A[X]$ to the ring A^A. The dual of this is the *evaluation homomorphism*. Given $x \in B$, we obtain a homomorphism $\mathrm{ev}_x : A[X] \to B$ mapping f to $f(x)$. Thus we can interpret $A[X]$ as the *free commutative A-algebra* generated by X, i.e. the 'most general' way of adding an additional element to the ring A.

Remark. If A is *not* a commutative ring, we may still define $A[X]$ as in the commutative case. But the evaluation maps are now *not* necessarily homomorphisms. For instance, over the Hamiltonian ring $\mathbf{H}$, in $\mathbf{H}[X]$ we find

$$(x+i)(x-i) = x^2 + 1$$

yet

$$(j+i)(j-i) = 2k \quad \text{and} \quad j^2 + 1 = 0.$$

In fact, for $x \in A$, $\mathrm{ev}_x : A[X] \to A$ is a homomorphism if and only if $x \in Z(A)$, since if ev_x is a homomorphism, then for any $a \in A$, the polynomial X, times the constant a, is equal to aX. Thus

$$\mathrm{ev}_x(Xa) = \mathrm{ev}_x(aX) = ax$$

whereas

$$\mathrm{ev}_x(X)\mathrm{ev}_x(a) = xa.$$

Thus $ax = xa$ for any $a \in A$.

If A and B are rings, then each homomorphism $\varphi : A \to B$ extends uniquely to a *reduction* homomorphism from $A[X]$ to $B[X]$ such that for each $a \in A$, the diagram below commutes

$$\begin{array}{ccc} A[X] & \longrightarrow & B[X] \\ \downarrow{\scriptstyle \mathrm{ev}_a} & & \downarrow{\scriptstyle \mathrm{ev}_{\varphi(a)}} \\ A & \xrightarrow{\varphi} & B \end{array}$$

The diagram forces us to define the mapping as

$$a_0 + a_1 X + \cdots + a_n X^N \mapsto \varphi(a_0) + \varphi(a_1)X + \cdots + \varphi(a_n)X^n$$

The most important case of the reduction homomorphism we will consider is when we consider an ideal $\mathfrak{a}$ in A, and then obtain the reduction homomorphism $A[X] \to (A/\mathfrak{a})[X]$; for instance, we can reduce an integer polynomial in $\mathbf{Z}[X]$ modulo some prime p to obtain a polynomial in $\mathbf{Z}_p[X]$. If we abuse notation, also writing $\mathfrak{a}$ for the ideal in $A[X]$ generated by $\mathfrak{a}$, then $A[X]/\mathfrak{a}$ is actually isomorphic to $(A/\mathfrak{a})[X]$. This is because a polynomial $a_0 + \cdots + a_n X^n$ is in the kernel of the reduction map if and only if $a_0, \ldots, a_n \in \mathfrak{a}$, which occurs if and only if $a_0 + \cdots + a_n X^n \in \mathfrak{a}$.

Corollary 3.1. *If $\mathfrak{a} \subset A$ is a prime ideal, then $\mathfrak{a} \subset A[X]$ is a prime ideal.*

3.2 The Euclidean Algorithm

If $f = a_0 + \cdots + a_n X^n \in A[X]$ is a non-zero polynomial with $a_n \neq 0$, we define the *degree* of f, denoted $\deg(f)$, to be n - the largest index with a nonzero coefficient in A. If $f = 0$, we define the degree of f to be $-\infty$. One can think of the degree of a polynomial as a measure of complexity of the corresponding arithmetic structure. The simplest polynomials are the *linear* polynomials, which have degree one. Upping the difficulty gives us the *quadratic* polynomials of degree two, the *cubic* polynomials of degree three, and so on and so forth. Looking at the operations defining polynomial addition and multiplication, it is easy to see that for any $f, g \in A[X]$,

$$\deg(f + g) \leqslant \max(\deg(f), \deg(g)),$$

and provided one of the leading coefficients of f or g is not a zero divisor,

$$\deg(fg) = \deg(f) + \deg(g)$$

thus multiplication of two polynomials always bilinearly magnifies the complexity of the polynomial. The multiplicative identity shows that the degree gives a *filtration* turning $A[X]$ into a *graded algebra*.

Remark. Note that the reason why we define $\deg(0) = -\infty$ is precisely so that these equations continuous to hold even if we do not assume the polynomials involved are nonzero.

Lemma 3.2. *If A is an integral domain, then $A[X]$ is an integral domain, and the units of $A[X]$ are precisely the units of A.*

Proof. The degree formula guarantees that if $fg = 0$, then $\deg(fg) = \deg(f) + \deg(g) = -\infty$, so either $\deg(f) = -\infty$, or $\deg(g) = -\infty$, which implies either $f = 0$ or $g = 0$. Thus $A[X]$ is an integral domain. Now suppose $f, g \in A[X]$ and $fg = 1$. Then $\deg(fg) = \deg(f) + \deg(g) = 0$. Thus $\deg(f) = \deg(g) = 0$, so $f, g \in A$. Thus $U(A[X]) = U(A)$. □

Corollary 3.3. *If A is a unital commutative ring, then $A[X]$ is unital, and the units of $A[X]$ are precisely polynomials of the form*

$$f = a_0 + a_1 X + \cdots + a_n X^n,$$

where a_0 is a unit of A, and $a_1, \ldots, a_n$ are nilpotent elements of A.

Proof. If a is a nilpotent element of a ring B, and u is a unit of B, then $u + a$ is a unit in B. This is because for any n,

$$(u + a)(u - a)(u^2 + a^2)\ldots(u^{2^n} - a^{2^n}) = u^{2^{n+1}} - a^{2^{n+1}}.$$

For suitably large n, $a^{2^{n+1}} = 0$, so

$$(u + a)(u - a)(u^2 + a^2)\ldots(u^{2^n} - a^{2^n}) = u^{2^{n+1}}$$

is invertible, which implies u is invertible. Thus if we can show that $a_1 X + \cdots + a_n X^n$ is nilpotent in $A[X]$ if $a_1, \ldots, a_n$ are nilpotent, then this will show $a_0 + \cdots + a_n X^n$ is a unit. But the nilpotent elements of a ring form an ideal, and since it is easy to see $a_i X^i$ is nilpotent if a_i is nilpotent, it is obvious that $a_1 X + \cdots + a_n X^n$ is nilpotent. Now suppose f is a unit in $A[X]$. Then for any prime $\mathfrak{p}$, f is a unit in $(A/\mathfrak{p})[X]$, which implies from the previous theorem that $a_i \in \mathfrak{p}$ for all $i > 0$. Thus $a_1, \ldots, a_n$ lies in every prime ideal; but this means that $a_1, \ldots, a_n$ are all nilpotent elements of the ring A. □

One of the most important facts about the degree of a univariate polynomial is that we can perform the Euclidean algorithm on them, which gives the ring $A[X]$ properties analogous to the ring of integers.

Theorem 3.4. *If $f, g \in A[X]$ and the leading coefficient of g is a unit, then there exists polynomials $h, r \in A[X]$ such that*

$$f = gh + r,$$

and $deg(r) < deg(g)$.

Proof. We prove the theorem by induction. If $\deg(f) < \deg(g)$, the theorem is trivial. Otherwise, write

$$f = a_0 + a_1 X + \cdots + a_n X^n \qquad g = b_0 + b_1 X + \cdots + b_m X^m$$

Then

$$\deg(f - a_n b_m^{-1} X^{n-m} g) < \deg(f)$$

so by induction,

$$f - a_n b_m^{-1} X^{n-m} g = hg + r$$

where $\deg(r) < \deg(g)$. But this implies

$$f = (h + a_n b_m^{-1} X^{n-m})g + r$$

so we have found an expansion for f. □

We have found that every polynomial ring is 'almost' a Euclidean domain, except that the expansion properties of the domain only hold for polynomials whose leading term is invertible. In particular, this means that if A is a field K, then *any* nonzero polynomial $A[X] = K[X]$ satisfies this property, and so the general argument for Euclidean domains gives the following corollary.

Corollary 3.5. *If K is a field, then $K[X]$ is a principal ideal domain.*

Proof. Let $\mathfrak{a}$ be a nonzero ideal of $K[X]$, and let g be a nonzero element of $\mathfrak{a}$ with smallest degree. Given any $f \in \mathfrak{a}$, the Euclidean algorithm enables us to find h and r with $f = gh + r$, where $\deg(r) < \deg(g)$. Since $r = f - gh \in \mathfrak{a}$, we conclude that $r = 0$. Thus we conclude that $\mathfrak{a} = (f)$. □

Corollary 3.6. *If K is a field, then $K[X]$ is factorial.*

Remark. For any nonzero ideal $\mathfrak{a}$ in $K[X]$, then $\mathfrak{a}$ is a generated by any f in $\mathfrak{a}$ where f is nonzero and has the smallest possible degree in $\mathfrak{a}$. But we can choose a unique generator by requiring $\mathfrak{a}$ to be monic, since if f and g are monic polynomials of smallest degree in $\mathfrak{a}$, then $\deg(f - g) < \deg(f)$, so we conclude $f - g = 0$, hence $f = g$.

Theorem 3.7. *Let A be an integral domain, and fix $f \in A[X]$.*

- *If $f(a) = 0$, then $X - a$ divides f.*

- *If $f \neq 0$, then f can have at most $\deg(f)$ roots in F.*

Proof. Since $X-a$ has degree 1, we may use the Euclidean alogrithm to find a polynomial $g \in A[X]$ and $r \in A$ such that we may write $f = g(X-a)+r$. Since $f(a) = r$, we conclude $r = 0$. If we have n distinct roots $a_1,\ldots,a_n$, we may apply induction to write $f = r(X-a_1)\ldots(X-a_n)$. The degree of the left hand side is n, and the degree of the right hand side is $n+\deg(r)$, hence $\deg(r) = 0$, so r is a nonzero constant. If $b \neq a_i$ for any i, then

$$f(b) = r \cdot (b-a_1)\ldots(b-a_n) \neq 0$$

Thus f can have at most n roots. □

Corollary 3.8. *If A is an integral domain, $f \in A[X]$, and $f(a) = 0$ for infinitely many $a \in A$, then $f = 0$.*

For finite integral domains, non-zero polynomials may still induce the zero function. For instance, if K is a finite field of order n, then $x^n = x$ for all $x \in K$. Thus the polynomial $X^n - X$ induces the zero function on K, yet $X^n - X$ is not formally the zero polynomial. This causes problems in certain problems where we must find a nonzero polynomial of low degree vanishing over a set of points in some K^n in an interesting way. Fortunately, the next lemma shows that these techniques generalize provided we can bound the degree of the nonzero terms.

Lemma 3.9. *Let K be a finite field with n elements. If $f \in K[X]$ induces the zero function on K, and $\deg(f) < n$, then $f = 0$.*

Proof. If f is nonzero but induces the zero function on K, then we obtain a contradiction by factoring out the linear terms corresponding to each element of K, which contradicts the degree of f. □

Now suppose K is a finite field with n elements, and $f \in K[X]$. Given f, the *reduced form* of f is a polynomial $g \in K[X]$ with $\deg(g) < n$ and $f(x) = g(x)$ for all $x \in K$. Repeatedly using the identity $x^n - x = 0$ in K shows that reduced forms always exist, and the above lemma shows they are unique.

Theorem 3.10. *If A is an integral domain, every finite subgroup of A^* is cyclic.*

Proof. Let G be such a subgroup. Since G is abelian, we can write it as the product of p groups, and so it suffices to prove the theorem by proving that a p-subgroup of A^* is abelian. Let x be an element of G of maximal period p^r. Then all elements of G are roots of the polynomial $X^{p^r} - 1$. But we know that there can only be at most p^r roots, and so G consists precisely of these roots, which are $x, x^2, \ldots, x^{p^r}$. □

Example. *If K is a finite field, then K^* is cyclic. In particular, $\mathbf{Z}_p^*$ is cyclic for each prime p; however, the proof above is not constructive, so we do not actually have efficient ways of* finding *generators for $\mathbf{Z}_p^*$ when p is a large prime.*

Example. *For each n, the set μ_n of n'th roots of unity, i.e. solutions to the equation $X^n - 1$ over $\mathbf{C}$, forms a finite subgroup of $\mathbf{C}^*$, and is therefore cyclic. The set $\mu = \bigcup_{n=1}^{\infty} \mu_n$ is a group, the group of all roots of unity. More generally, over any algebraically complete field K, we can conisder the groups $\mu_n(K)$ and $\mu(K)$. If K is a finite field with n elements, then $K^* = \mu(K)$, since all elements of K^* are roots of the polynomial $X^{n-1} - 1$.*

3.3 Algebraic and Trancendental Elements

Given a ring B with a subring A, we say $b \in B$ is *algebraic* over A if there is a nonzero polynomial $f \in A[X]$ such that $f(b) = 0$. Otherwise, b is called *trancendental*. It is fairly easy to show a particular element of a ring is algebraic (e.g. $\sqrt{2}$ is algebraic over $\mathbf{Q}$, since we can set $f(X) = X^2 - 2$), but it is often very difficult to show that an element of a ring is trancendental. We know that π and e are trancendental over $\mathbf{Q}$, but the proof is a difficult analytical argument. It is still an open question whether $\pi + e$ and π/e are trancendental; it is not even known whether they are irrational! For multivariate polynomial rings, the situation is even less understood. We say $b_1, \ldots, b_n \in B$, we say these elements are *algebraically independent* over A if there is no polynomial $f \in A[X_1, \ldots, X_n]$ with $f(b_1, \ldots, b_n) = 0$.

3.4 Multivariate Polynomials

We can study multivariate expressions in a commutative ring by using multivariate polynomials. Given n variables $X_1, \ldots, X_n$, we can consider

expressions of the form

$$\sum_{i_1,\ldots,i_n} a_{i_1\ldots i_n} X_1^{i_1} \ldots X_n^{i_n}$$

such that only finitely many $a_{i_1,\ldots,i_n}$ are non-zero. The set of all such expressions forms a ring over A, denoted $A[X_1,\ldots,X_n]$. Let us list some commonly used properties of this ring.

- One can reduce multi-dimensional polynomial rings to univariate polynomial rings by noticing that

  $$A[X_1,\ldots,X_n] = A[X_1,\ldots,X_{n-1}][X_n],$$

  because every polynomial can be uniquely written as $\sum f_k X_n^k$ for some $f_k \in A[X_1,\ldots,X_{n-1}]$, formed by factoring out the right powers of f_k.

- Given a tuple $b = (b_1,\ldots,b_n) \in B^n$, where A is a subring of a ring B, we can consider an evaluation morphism $\mathrm{ev}_b : A[X_1,\ldots,X_n] \to B$, as in the one-dimensional case.

- Given a homomorphism $f : A \to B$, there is a unique homomorphism from $A[X_1,\ldots,X_n]$ to $B[X_1,\ldots,X_n]$ causing the evaluation diagrams to commute.

- The polynomials $X_1^{i_1}\ldots X_n^{i_n}$ are known as *primitive monomials*. We define the degree of this primitive polynomial to be $i_1+i_2+\cdots+i_n$, and we define the degree of a general polynomial to be the maximal degree of the primitive polynomials in the expansion of the polynomial which have non-zero coefficients.

- Alternatively, if we want to focus on a particular variable, we define the degree of f with respect to X_n to be the degree of f viewed as an element of $A[X_1,\ldots,X_{n-1}][X_n]$.

- A polynomial $f \in A[X_1,\ldots,X_n]$ is *homogenous* of degree m if the only monomials $X^{i_1}\ldots X^{i_n}$ occuring in f satisfy $i_1+\cdots+i_n = m$. If A is a subring of B, and $u, t_1,\ldots,t_n \in B$, then we find

  $$f(ut_1,\ldots,ut_n) = u^m f(t_1,\ldots,t_n).$$

Homogenous polynomials are precisely those polynomials satisfying this equation, provided that there exists algebraically independent $u, t_1, \ldots, t_n$ in B over A, because then the fact that $f(ut_1, \ldots, ut_n) = u^m f(t_1, \ldots, t_n)$ implies

$$f(YX_1, \ldots, YX_n) = Y^m f(X_1, \ldots, X_n)$$

and looking at the terms in this expansion shows all monomials must have the same degree.

Just as in the univariate case, a nonzero multivariate polynomial cannot have too many zeroes.

Corollary 3.11. *Let $f \in A[X_1, \ldots, X_n]$, where A is an integral domain. If there exists infinite sets $S_1, \ldots, S_n \subset A$ such that $f(a_1, \ldots, a_n) = 0$ for each $a_1 \in S_1, \ldots, a_n \in S_n$, then $f = 0$.*

Proof. We prove by induction on n, the case $n = 1$ having already been proven. For each $a \in A$, we have an evaluation homomorphism

$$\mathrm{ev}_a : A[X_1, \ldots, X_n] \to A[X_1, \ldots, X_{n-1}]$$

obtained by setting $X_n = a$. By induction, we know $\mathrm{ev}_a(f) = 0$ for each $a \in S_n$. Now write

$$f = \sum_{i_1, \ldots, i_{n-1}} \left(\sum_{i_n} a_{i_1 \ldots i_n} X_n^{i_n} \right) X_1^{i_1} \ldots X_{n-1}^{i_{n-1}}$$

Since $\mathrm{ev}_a(f) = 0$ for each $a \in S_n$, then for each $i_1, \ldots, i_{n-1}$, the polynomial $\sum_{i_n} a_{i_1 \ldots i_n} X_n^{i_n}$ has infinitely may zeroes, and thus vanishes identically. Thus $f = 0$, completing the induction. □

For finite fields we obtain a similar result after applying a reduction.

Lemma 3.12. *Suppose K is a finite field with m elements. If $f \in K[X_1, \ldots, X_n]$ induces the zero function on K^n and has degree less than m in each variable $\{X_1, \ldots, X_n\}$, then $f = 0$. The ideal of functions vanishing on K^n is precisely*

$$(X_1^m - X_1, \ldots, X_n^m - X_n).$$

Proof. We proceed by induction. Write

$$\mathfrak{a} = (X_1^m - X_1, \ldots, X_n^m - 1).$$

Suppose $f \in K[X_1, \ldots, X_n]$ has degree less than m in each variable and induces the zero function on K^n. Write

$$f = \sum_{i_1, \ldots, i_{n-1}} \left(\sum_{i_n} a_{i_1 \ldots i_n} X_n^{i_n} \right) X_1^{i_1} \ldots X_{n-1}^{i_{n-1}}$$

The inductive case applies that for each $i_1, \ldots, i_{n-1}$, the polynomial $\sum_{i_n} a_{i_1 \ldots i_n} X^{i_n}$ induces the zero function on K. Since the degree in i_n is less than m, $a_{i_1 \ldots i_n} = 0$ for all $i_1, \ldots, i_n$. This completes the proof of the first property of this lemma.

Now write

$$\mathfrak{a} = (X_1^m - X_1, \ldots, X_n^m - X_m)$$

If $f \in K[X_1, \ldots, X_n]$ induces the zero function on K^n, then it is certainly equivalent modulo $\mathfrak{a}$ to a polynomial $g \in K[X_1, \ldots, X_n]$ with degree less than m in each variable. But then g induces the zero function on K^n, hence $g = 0$, so $f \in \mathfrak{a}$. □

3.5 Polynomials over a Factorial Ring

Let A be an integral domain. If K is the field of fractions of A, then $K[X]$ is a principal ideal domain, hence factorial. One might naturally ask what the relation is between the divisibility theory of $A[X]$ and the divisibility theory of $K[X]$. Normally this can be obtained by 'cancelling denominators' of equations in $K[X]$ to obtain equations in $A[X]$. Clearly we cannot use this fact to conclude $A[X]$ is factorial in general, since if $A[X]$ is factorial, A is factorial.

However, we can use this technique to prove $A[X]$ is factorial if A is factorial, a process we now carry out. Let A be a factorial ring. Since A is an integral domain, we may consider the field of fractions K. We shall show that $f \in A[X]$ is irreducible over $K[X]$ if and only if f is irreducible over $A[X]$, and if the greatest common denominator of the coefficients of f is equal to zero. For each prime $p \in A$, and non-zero $x \in K$, we may

uniquely write $x = p^r u$, where $r \in \mathbf{Z}$, and $p \nmid u$. We define the *order* of x at p to be r, and denote it by $\mathrm{ord}_p(x)$. Just as with polynomials, we have

$$\mathrm{ord}_p(x+y) \geqslant \min\left(\mathrm{ord}_p(x), \mathrm{ord}_p(y)\right) \qquad \mathrm{ord}_p(xy) = \mathrm{ord}_p(x) + \mathrm{ord}_p(y)$$

If $x = 0$, we define $\mathrm{ord}_p(x) = \infty$ so that these identities continue to hold. Any prime $p \in A$ is also a prime in $A[X]$, so we can define $\mathrm{ord}_p(f)$ for each $f \in A[X]$; if $f = a_0 + \cdots + a_n X^n$, then

$$\mathrm{ord}_p(f) = \min\left(\mathrm{ord}_p(a_0), \ldots, \mathrm{ord}_p(a_n)\right).$$

If A is a factorial ring, we define the *content* $\mathrm{cont}(f) \in A$ of a non-zero $f \in A[X]$ to be the greatest common denominator of the coefficients of f (technically, we must interpret $\mathrm{cont}(f)$ as a coset of A modulo it's units, but we abuse notation here). If we pick a prime p from each coset of primes identified up to units, then

$$\mathrm{cont}(f) = \prod_p p^{\mathrm{ord}_p(f)}.$$

If $f = 0$, define $\mathrm{cont}(f) = 0$. Then the content is unique up to a unit in A. We may always write $f = \mathrm{cont}(f)g$, where g is a polynomial in $A[X]$ with unit content (such polynomials are known as *primitive*.

Lemma 3.13 (Gauss)**.** *Let A be a factorial ring, and K it's field of fractions. Then for $f, g \in K[X]$, $cont(fg) = cont(f) \cdot cont(g)$.*

Proof. Assume without loss of generality that f and g have unit content. Then it suffices to prove fg has unit content. Let $p \in A$ be a prime, denote $A/(p)$ by B, and consider the reduction homomorphism $\varphi : A \to B$, which extends to a map $\varphi : A[X] \to B[X]$. Since $\varphi(f)$ and $\varphi(g)$ are nonzero polynomials, and p is prime, $\varphi(fg)$ is a nonzero polynomial. □

Corollary 3.14. *Suppose A is a factorial ring, let $f \in A[X]$ be primitive, and let K be the field of fractions of A. Then f is irreducible in $A[X]$ if and only if it is irreducible in $K[X]$.*

Proof. Suppose $f \in A[X]$ is primitive and irreducible over $A[X]$, and suppose $f = gh$, where $g, h \in K[X]$. Then we can find primitive polynomials $g_0, h_0 \in A[X]$ and $a_0, b_0 \in A$ such that $a_0 g = g_0$, $b_0 h = h_0$. If $c = a_0 b_0$, then $cf = g_0 h_0$. But then since g_0 and h_0 is primitive, we conclude by Gauss'

lemma that c is a unit in A. Thus $f = (g_0/c)h_0$, where $g_0/c, h_0 \in A[X]$, and so we conclude that either g_0 or h_0 is a unit in $A[X]$, and thus either g or h is a unit in $K[X]$.

Conversely, if $f \in A[X]$ is primitive and irreducible over $K[X]$, and if $f = gh$ for $g, h \in A[X]$, then either g or h is a unit in $K[X]$,which implies that either g or h is a constant. Since $\mathrm{cont}(g)\mathrm{cont}(h) = 1$, this implies that either g or h is a unit in A. □

Corollary 3.15. *If A is factorial, then $A[X_1, \ldots, X_n]$ is factorial.*

Proof. We just prove that $A[X]$ is factorial if A is, from which the general theorem holds by induction. The existence of a factorization is quite easy to show. Let K be the field of fractions of A. If $f \in A[X]$, we may write

$$f = g_1 \cdots g_n$$

where g_n are irreducible elements of $K[X]$. Now write $g_i = a_i g_i'$, where g_i' is a primitive polynomial in $A[X]$. Thus $f = (a_1 \ldots a_n) g_1' \ldots g_n'$. Each g_i' is an element of $A[X]$ which is irreducible over $K[X]$ and has unit content, so it is irreducible over $A[X]$. We may write

$$a_1 \ldots a_n = p_1^{k_1} \ldots p_m^{k_m}$$

where each p_i is an irreducible element of A (and thus irreducible over $A[X]$). Thus

$$f = p_1^{k_1} \ldots p_m^{k_m} g_1' \cdots g_n'$$

has been written as a product of irreducible elements in $A[X]$. If we have two different factorizations

$$p_1^{k_1} \ldots p_m^{k_m} g_1 \ldots g_n = q_1^{l_1} \ldots q_r^{l_r} h_1 \ldots h_t$$

Then by unique factorization in $K[X]$, we must have $t = n$, and after some rearranging, $f_i = u_i g_i$, for some nonzero $u_i \in K$. But since f_i and g_i are primitive, we may assume without loss of generality that u_i is a unit in A. Cancelling out appropriate factors, we conclude that

$$p_1^{k_1} \ldots p_m^{k_m} = (u_1 q_1^{l_1}) \ldots (u_r q_r^{l_r}),$$

and we may now apply unique factorization in A. □

Note that for $n \geqslant 2$, the ring $K[X_1,\ldots,X_n]$ is not principal. Indeed (X,Y) is an ideal in $K[X,Y]$ which cannot be principal, for no non unital element divides both X and Y. Thus the fact that these rings are factorial is truly a novel part of the proof above.

Corollary 3.16. *Let $f \in K[X_1,\ldots,X_n]$, and suppose we can find irreducibles $f_1,\ldots,f_N$ and integers $k_1,\ldots,k_N > 0$ such that $f = f_1^{k_1}\ldots f_N^{k_N}$, so*

$$K[X_1,\ldots,X_n]/(f) \cong K[X_1,\ldots,X_n]/(f_1)^{k_1} \times \cdots \times K[X_1,\ldots,X_n]/(f_N)^{k_N}.$$

Proof. The ideals $(f_i^{k_i})$ and $(f_j^{k_j})$ are coprime since they have no common factor. Thus

$$(f) = (f_1^{k_1})\ldots(f_N)^{k_N} = (f_1^{k_1}) \cap \cdots \cap (f_N^{k_N}),$$

and so we can apply the Chinese remainder theorem. □

3.6 Criterion for Irreducibility

It is actually quite tricky to determine whether a given polynomial is irreducible. For instance, X^4+4 does not have any roots in $\mathbf{Q}$, yet X^4+4 is reducible,

$$X^4+4 = (X^2+2X+2)(X^2-2X+2)$$

Some techniques can be used to determine when a polynomial is irreducible. We begin with a powerful result for polynomials over factorial rings.

Theorem 3.17 (Integral Root Test)**.** *Let A be a factorial ring, and K its quotient field. Let*

$$f = a_0 + a_1X + \cdots + a_nX^n$$

Suppose $f(b/d) = 0$, where b and d are relatively prime. Then b divides a_0, and d divides a_n. In particular, if $a_n = 1$, then the only roots of f are in A.

Proof. We have

$$a_0 + a_1(b/d) + \cdots + a_n(b/d)^n = 0$$

Then

$$d^n a_0 + a_1 b d^{n-1} + \cdots + a_n b^n = 0$$

which implies

$$b(a_1 d^{n-1} + \cdots + a_n b^{n-1}) = -d^n a_0$$

since b does not divide d, b does not divide d^n, and thus b divides a_0. Similarily, by factoring out d, we find d divides a_n. □

Example. *The polynomial $X^3 - 3X - 1$ is irreducible in $\mathbf{Z}[X]$. If the polynomial was reducible, it would have an integer root. But the integral root test implies that the only possible roots are either $+1$ or -1. Neither gives a root, completing the proof.*

Example. *For any prime $p \in \mathbf{Z}$, the polynomials $X^2 - p$ and $X^3 - p$ are irreducible in $\mathbf{Z}[X]$. To see this, they would only be reducible if they had an integer root. But the only possible integer roots are either p or $-p$ by the integral root test, completing the proof.*

Another way to prove a polynomial is irreducible is to reduce the polynomial's coefficients modulo an ideal, detailed in the next proposition. In the case $\mathbf{Z}[X]$, we reduce modulo a prime to obtain an element of $\mathbf{F}_p[X]$, and we can easily check this polynomial's properties since $\mathbf{F}_p$ is finite.

Theorem 3.18 (Reduction Criterion)**.** *Let A and B be integral domains and consider a surjective homomorphism $\varphi : A \to B$. If $f \in A[X]$, $\varphi(f)$ has the same degree in f, and $\varphi(f)$ cannot be factored into two polynomials of smaller degree in B, then f is irreducible.*

Example. *Consider the polynomial $X^2 + XY + 1 \in \mathbf{Z}[X, Y]$. View $\mathbf{Z}[X, Y]$ as $\mathbf{Z}[Y][X]$. Let $\phi : \mathbf{Z}[X, Y] \to \mathbf{Z}[X]$ be the homomorphism obtained by setting $Y = 0$. Then $\phi(X^2 + XY + 1) = X^2 + 1$ has the same degree in X. Moreover, $X^2 + 1$ is irreducible in $\mathbf{Z}[X]$, and so $X^2 + XY + 1$ is irreducible in $\mathbf{Z}[X, Y]$ by the reduction criterion.*

Theorem 3.19 (Eisenstein)**.** *Let A be an integral domain, and let $\mathfrak{a}$ be a prime ideal. Consider a polynomial*

$$f = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in A[X],$$

with $a_0, \ldots, a_{n-1} \in \mathfrak{a}$, but $a_0 \notin \mathfrak{a}^2$. Then f is irreducible in $A[X]$.

Proof. Let $\phi : A \to A/\mathfrak{a}$ denote the reduction homomorphism. If $f = gh$ for $g, h \in A[X]$, then $X^n = \phi(f) = \phi(g)\phi(h)$. Since $(A/\mathfrak{a})$ is an integral domain,

the only divisors of X^n in $(A/\mathfrak{a})[X]$ are powers of X^n multiplied by a unit, so $\phi(g) = tX^m$, $\phi(h) = t^{-1}X^l$ for some $t \in U(A/\mathfrak{a})$, where $m + l = n$. If $m, l > 0$, this gives a contradiction, for it implies $g(0), h(0) \in \mathfrak{a}$, and thus $a_0 = g(0)h(0) \in \mathfrak{a}^2$. This we may assume without loss of generality that $m = 0$. But then $\deg(h) = n$, $\deg(g) = 0$. Thus g is a constant, and since f is monic, this implies that g is actually an element of $U(A)$. □

Example. *Eisenstein's criterion's can often be used to determine when the polynomial $X^n - a \in \mathbf{Z}[X]$ is irreducible, i.e. when some prime p divides a, but p^2 does not. This shows $X^n - 6$ and $X^n - 4$ are irreducible. On the other hand, this cannot detect that $X^3 - 8 = (X - 2)(X^2 + 2X + 4)$ is irreducible.*

Example. *The polynomial $X^{p-1} + \cdots + X + 1$ is irreducible in $\mathbf{Q}$ if p is prime. Consider the transformation $X = Y + 1$. The transformation preserves irreducibility, since it is really an isomorphism of $\mathbf{Q}[X]$. Then*

$$(Y+1)^{p-1} + \cdots + (Y+1) + 1 = \frac{(Y+1)^p - 1}{Y} = \sum_{k=0}^{p-1} \binom{p}{k+1} Y^k$$

All coefficients of this polynomial are divisible by p except for the higher order term, which is equal to one, and the lowest term is $\binom{p}{1} = p$, so Eisenstein's criterion tells us the polynomial is irreducible.

Example. *Let K be a field, and consider the field of rational functions $K(X)$. The polynomial $Y^n - X$ is irreducible in $K(X)[Y]$. Note first that $Y^n - X$ has content one with respect to $K[X]$, so $Y^n - X$ is irreducible over $K(X)[Y]$ if and only if it is irreducible over $K[X][Y]$. But over $K[X][Y]$ we may apply Eisenstein's criterion, since X is a prime in $K[X]$, to conclude that $Y^n - X$ is irreducible.*

3.7 Partial Fractions

Theorem 3.20. *Let A be a factorial ring, and let K be its quotient field. Choose a representation $\{p_i\}$ of primes. Then for each $a/b \in K$ there is $a_i \in A$ and $j_i \in \mathbf{N}$ for each p_i such that almost all a_i are zero, and*

$$a/b = \sum_i \frac{a_i}{p^{j_i}}$$

Proof. First we show existence. Let $a, b \in A$ be relatively prime. Then we may write $ma + nb = 1$, so

$$\frac{1}{ab} = \frac{m}{b} + \frac{n}{a}$$

Thus for any $c \in A$,

$$\frac{c}{ab} = \frac{cm}{b} + \frac{cn}{a}$$

By induction, we may write

$$\frac{1}{p_1^{k_1} \dots p_{n+1}^{k_{n+1}}} = \sum \frac{a_i}{p_i^{k_i}}$$

Hence

$$\frac{c}{p_1^{k_1} \dots p_{n+1}^{k_{n+1}}} = \sum \frac{ca_i}{p_i^{k_i}}$$

□

Chapter 4

Modules

All groups are really sets of bijective maps in disguise. Regardless of the complex nature that grants us a specific group, we can still relate it back to some symmetric group, by Cayley's theorem. This leads to the study of group actions. It turns out that all rings can be seen as a set of endomorphisms over an abelian group. The counterpart to a group action on a G-set is then a ring action on an A-module. A representation of a ring A on an abelian group M is a ring homomorphism from A into $\mathrm{Hom}(M)$. If such a representation is fixed, we can define a 'scalar multiplication' structure on M by elements of A, for $x \in M$ and $a \in A$, letting ax denote the action of a on x via the representation. We obtain the relations

$$a(x+y) \qquad (ab)x = a(bx) \qquad (a+b)x = ax + bx$$

and if A has a multiplicative unit, we assume $1 \cdot x = x$. Conversely, any multiplication map from $A \times M$ to M satisfying these properties induces a representation of A on $\mathrm{Hom}(M)$, and we call any such M with a fixed scalar multiplication a (left) A *module*. If A is a field, then these are just the axioms which give M the structure of a *vector space* over A, and any such module over a field will be referred to as a vector space.

A *homomorphism* $f : M \to N$ between two A modules M and N is a homomorphism of abelian groups satisfying the additional requirement that $f(ax) = af(x)$ for each $x \in M$ and $a \in A$. The family of all morphisms between M and N is denoted $\mathrm{Hom}_A(M,N)$, or just $\mathrm{Hom}(M,N)$ if the underlying ring is obvious. Thus for each ring A, we have a category Mod_A of A-modules. It is simple to verify that $\mathrm{Hom}(M,N)$ is an abelian group; if A is commutative, then $\mathrm{Hom}(M,N)$ is even an A module if we define

$(af)(x) = a \cdot f(x)$, since we then find that

$$(af)(sx) = af(sx) = asf(x) = saf(x) = s(af)(x).$$

It is easy to see from this calculation that if A is not commutative, then af is a homomorphism of abelian groups which may not be a homomorphism of A modules. On the other hand, $\mathrm{End}(M)$ is always a ring, which forms an A algebra when A is commutative.

A *submodule* of an A module M is a subset N of M such which forms an abelian subgroup and such that $ax \in N$ for each $x \in N$ and $a \in A$. Submodules are the natural objects to quotient by; if N is a submodule of an A module M, then the quotient group M/N naturally has the structure of an A module. The natural analogues of the isomorphism theorems for abelian groups remain true for modules. In particular, one can use the first isomorphism theorem to pass through a quotient by the kernel of a homomorphism. The second isomorphism theorem implies if N_1 and N_2 are submodules, then $N_1 + N_2$ and $N_1 \cap N_2$ are submodules, and $N_1/(N_1 \cap N_2)$ is isomorphic to $(N_1 + N_2)/N_2$. For a submodule N of a module M, the third isomorphism theorem gives a correspondence with submodules of M/N and submodules of M containing N.

Remark. In the following, we always assume A is a untial ring. This is safe, because any module over a ring A without unity is automatically defined as a module over the unitization of A, since we can define $(n+a)x = nx+ax$.

Example. *If A is a ring, it is a module over itself. The same is true of A^n, known as the* free A module of rank n. *Submodules of A are precisely* left ideals *of A, i.e. subrings $\mathfrak{a}$ of A which are closed under left multiplication by elements of A. The homomorphisms in $\mathrm{Hom}(A^n, A^m)$ can be identified with the family $M_{n,m}(A)$ of $n \times m$ matrices over A, where compositions of homomorphisms act the same way as matrix multiplication does.*

Example. *Any abelian group M has the structure of a $\mathbf{Z}$ module, for we may define $nx = x + \cdots + x$. A $\mathbf{Z}$ module homomorphism is just a homomorphism of abelian groups, so that the category $\mathsf{Mod}_{\mathbf{Z}}$ is really just the category Ab in disguise.*

Example. *A module over a field is called a* vector space, *and the study of such modules forms the field of* linear algebra. *However, even in linear algebra one needs to study more general modules; given a vector space V over a field k and a*

linear transformation T, one usually gains deeper information on the structure of T by studying V as a $k[X]$-module, under the action $f \cdot v = f(T)(v)$. The study of invariant subspaces of a vector space under a linear transformation is really a discussion of the submodules of V, viewed as a $k[X]$ module.

More generally, let A be a commutative ring. If M is an A-module, then we have a homomorphism $\phi : A \to End(M)$. If $T \in End(M)$ is fixed, then ϕ extends to a homomorphism from $A[X]$ to $End(M)$ mapping X to T. Thus M naturally has the structure of an $A[X]$ module, where multiplication by $f \in A[X]$ acts as the endomorphism $f(T)$. Submodules of $A[X]$ are precisely the T invariant submodules of M. A more general fact is that $A[X]$ module structures on an A module M are in one to one correspondence with elements of $End(M)$, so studying a general endomorphism of an A module, when A is commutative, is exactly the same as studying an $A[X]$ module.

Example. *Let us consider some deeper examples of modules over polynomial rings. If $C^\infty(\mathbf{R}^d)$ is the real vector space of real-valued infinitely differentiable functions, then for each $i \in \{1, \ldots, n\}$ we can consider the linear operator*

$$D_i f = \frac{\partial f}{\partial x^i}.$$

This induces a natural $\mathbf{R}[X_1, \ldots, X_d]$ module structure such that $C^\infty(\mathbf{R}^d)$, where a polynomial corresponds to a constant coefficient differential operator. Similarily, we can consider the linear operators

$$(M_i f)(x) = 2\pi x_i f(x)$$

which induce a separate $\mathbf{R}[X_1, \ldots, X_d]$ module structure. The set of Schwartz functions $\mathcal{S} \subset C^\infty(\mathbf{R}^d)$ forms a $\mathbf{R}[X_1, \ldots, X_d]$ submodule under both representations. Moreover, the Fourier transform gives an isomorphism between both representations of $\mathbf{R}[X_1, \ldots, X_d]$ on $\mathcal{S}$, since for $f \in \mathcal{S}$,

$$M_i \widehat{f} = \widehat{D_i f}.$$

This is a powerful algebraic relation exploited in Harmonic anlaysis.

In operator theory, one studies bounded linear operators acting on Hilbert spaces H. If H is a complex Hilbert space and T is a bounded linear operator on H, then H naturally has the structure of a $\mathbf{C}[X]$ module, where each polynomial f acts as the bounded linear operator $f(T)$. The study of the spectral theory of such operators shows that we can actually extend this operation to give H the

natural structure of a $\mathcal{O}(\sigma(T))$ module, where $\mathcal{O}(\sigma(T))$ is the ring of functions analytic in a neighbourhood of the spectrum of T, each acting as a bounded linear operator on H. If T is self adjoint, then we can further extend this operation to give H the structure of a $C(\sigma(T))$ module, where each continuous function f on the spectrum of T acts as a bounded linear operator.

Both $k[X]$ and $\mathbf{Z}$ are principal ideal domains. Later on, we will develop a powerful theory which classifies finitely generated modules over principal ideal domains, which generalizes the classification of endomorphisms over a vector space by means of the Jordan normal form, and the classification of finite abelian groups.

Example. *Let A be a commutative ring. Then let $Tor(M)$ denote the submodule of all $x \in M$ such that there is a $a \neq 0$, which is* not *a zero divisor of A, such that $ax = 0$. Such elements are said to* have torsion. *A* torsion module *is a module M such that $Tor(M) = M$, and a module M is called* torsion free *if $Tor(M) = (0)$. If A is an integral domain, the ring $M/Tor(M)$ is always torsion free, and can be viewed as the most general torsion free quotient of M; if $\phi : M \to N$ is a homomorphism, then $\phi(Tor(M)) \subset Tor(N)$, so if A is an integral domain and N is torsion free, then ϕ automatically factors through $M/Tor(M)$ by the first isomorphism theorem.*

Example. *If N is a submodule of an A module M, we let $Ann(N)$ be the set of all $a \in A$ such that $aN = (0)$. Then $Ann(N)$ is a left ideal in A. Conversely, if $\mathfrak{a}$ is a left ideal of A, then we set $Ann(\mathfrak{a})$ to be the set of all $x \in M$ such that $\mathfrak{a}x = (0)$. We note that $Ann(N)$ is contained in $\mathfrak{a}$ if and only if N is contained $Ann(\mathfrak{a})$, so we obtain a Galois connection between left ideals of A and submodules of M.*

Example. *If $\mathfrak{a}$ is a left ideal of a ring A, and M is an A module, then the subgroup $\mathfrak{a}M$ of M generated by elements of the form ax, with $a \in \mathfrak{a}$ and $x \in M$, forms a submodule of M. If $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ are left ideals in A, then we can consider the natural homomorphism*

$$M \to M/\mathfrak{a}_1 M \times \cdots \times \mathfrak{a}_n M$$

which has kernel $\mathfrak{a}_1 M \cap \cdots \cap \mathfrak{a}_n M$. In fact, if the ideals $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ are coprime, *then this map is surjective, then $\mathfrak{a}_1 M \cap \cdots \cap \mathfrak{a}_n M = (\mathfrak{a}_1 \ldots \mathfrak{a}_n)M$, and so*

$$M/(\mathfrak{a}_1 \ldots \mathfrak{a}_n)M \cong M/\mathfrak{a}_1 M \times \cdots \times M/\mathfrak{a}_n M.$$

Thus we have a version of the Chinese remainder theorem for modules, and this more general version is proved essentially by the same techniques as in the case of the normal Chinese remainder theorem.

Example. *Let A be a ring, and fix $a \in Z(A)$. Given a module M, the set of all ax, for $x \in M$ forms a submodule of M, denoted aM, since $t(ax) = a(tx)$ for $t \in A$. This is no longer true if a is not an element of $Z(A)$. For instance, since*

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix},$$

On the other hand, for any ring A, the set $e_{11}M_2(A)$ is not a submodule of $M_2(A)$ when A is unital, because

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \notin e_{11}M_2(A).$$

4.1 Generators of Modules

Given elements $x_1, \ldots, x_n$ of a vector space V, we can talk about the span of these elements, the smallest subspace of V containing $x_1, \ldots, x_n$. Given an A-module M and a set $S \subset M$, we can consider the smallest submodule generated by S, which can be written as

$$AS = \{a_1 s_1 + \cdots + a_n s_n : a_1, \ldots, a_n \in A, s_1, \ldots, s_n \in S\}$$

A module is *finitely generated* if it is generated by a finite set, and *cyclic* if it is generated by a single element.

Remark. If A is non unital things can go very wrong here. For instance, S might not even be a subset of AS. Even weirder things can happen; as a module, A might not even be finitely generated over itself; this happens, for instance, if A is the ring of compactly supported continuous functions on $\mathbf{R}^d$.

It shall be useful to introduce some general constructions. The category of modules is closed under direct products and direct sums. We can construct a direct product for a family of modules $\{M_\alpha : \alpha \in I\}$ by consider the module M of all sequences $\{x_\alpha\}$ such that $x_\alpha \in M_\alpha$ for each α, with addition and multiplication taken componentwise. We must be slightly

more careful when constructing direct sums; we take the submodule of $\prod M_\alpha$ consisting of all sequences $\{x_\alpha\}$ such that only finitely many x_α are nonzero. One writes the direct product and direct sum as

$$\bigtimes_\alpha M_\alpha \quad \text{and} \quad \bigoplus_\alpha M_\alpha$$

respectively, and for finite families $M_1, \ldots, M_n$ of modules, we write the direct product and sum as $M_1 \times \cdots \times M_n$ and $M_1 \oplus \cdots \oplus M_n$ (note that in this case, the direct product is equal to the direct sum). One can also produce direct and inverse limits, just as in the case of abelian groups, but we leave this for the reader to construct.

A set S is a *basis* for a module M if every element of M can be *uniquely* written as $a_1 s_1 + \cdots + a_n s_n$ for some $a_1, \ldots, a_n \in A$ and distinct $s_1, \ldots, s_n \in S$, which is equivalent to the condition that for any distinct $s_1, \ldots, s_n$ and $a_1, \ldots, a_n \in A$, $a_1 s_1 + \cdots + a_n s_n = 0$ if and only if $a_1 = \cdots = a_n = 0$. If S is finite, containing n elements, then this implies precisely that M is isomorphic to A^n. More generally, M is isomorphic to the direct sum $\oplus_{s \in S} A$, since elements of M can be written freely as a finite sum of elements from the set S.

if $x_1, \ldots, x_n$ generate a module M, but are not a basis, we can still use the module A^n to understand M, because there is a natural surjective morphism from A^n to M, and so M is isomorphic to a quotient of A^n by some submodule. In particular, a cyclic module is isomorphic to $A/\mathfrak{a}$ for some left ideal $\mathfrak{a}$. A module M is *irreducible* if it contains no nontrivial submodules, and it is easy to see that such modules must be isomorphic to $A/\mathfrak{m}$ for some maximal left ideal $\mathfrak{m}$, or equal to (0). Irreducible modules play a crucial role in representation theory; note non-zero homomorphisms between irreducible modules must be isomorphisms, so in particular, if M is irreducible then $\mathrm{End}(M)$ is a division ring.

Even though not all modules over a ring are free, one might still like to ascribe a notion of *dimension* to the family of free modules. However, over certain exotic rings, even this may not be established, since we might have A^n isomorphic to A^m for $n \neq m$.

Example. *Let A be the ring of endomorphisms over a vector space with a countable basis $\{e_k\}$. Then as A modules, we claim that $A \oplus A$ is isomorphic to A, so in particular, any module isomorphic to a free module of rank n is automatically isomorphic to a free module of rank 1. Divide the basis $\{e_i\}$ into two*

infinite sets $\{f_i\}$ and $\{g_i\}$, and define two endomorphisms F and G by setting $F(f_n) = G(g_n) = e_n$, and $F(g_n) = G(f_n) = 0$. If $TF + SG = 0$, then

$$0 = (TF)(f_n) + (SG)(f_n) = T(e_n) + S(0) = T(e_n)$$

Thus $T = 0$. Conversely,

$$0 = (TF)(g_n) + (SG)(g_n) = T(0) + S(e_n) = S(e_n),$$

so that $S = 0$. Thus F and G are linearly independant over A. If T is an arbitrary endomorphism in A, we can find two endomorphisms T_0 and T_1 such that $T_0(e_n) = T(f_n)$ and $T_1(e_n) = T(g_n)$. Then $T = T_0F + T_1G$. Thus the map $(T_0, T_1) \mapsto T_0F + T_1G$ induces an isomorphism between $A \oplus A$ and A.

A ring A has the *invariant basis property* if A^n is not isomorphic to A^m for $n \neq m$. If M is a free module over a ring A with the invariant basis property, it follows that $\dim_A(M)$ is a well defined quantity. All commutative rings have the invariant basis property. To see this, if A^n is isomorphic to A^m, and $\mathfrak{m}$ is a maximal ideal of A, then we induce an isomorphism between $(A/\mathfrak{m})^n$ and $(A/\mathfrak{m})^m$ which is also an isomorphism of $A/\mathfrak{m}$ modules; since $A/\mathfrak{m}$ is a field, linear algebra implies that $n = m$. If A is an algebra over a field k, which is a finitely generated module over M, then A also has the invariant basis property; if A^n is isomorphic to A^m, and A is isomorphic to k^l as a vector space, then we conclude that k^{nl} is isomorphic to k^{ml}, hence $nl = ml$ and so $n = m$. This covers most of the rings one studies in the fundamentals of abstract algebra.

4.2 Algebras

It is often useful to discuss R-algebras, i.e. R-modules which also have the structure of a ring. To be more precise, a (left) R-algebra A is a (left) R-module together which is also a ring, and satisfies the identity

$$r(xy) = (rx)y = x(ry)$$

for all $x, y \in A$ and $r \in R$. If A is a *unital* R-algebra,

4.3 Exact Sequences and Homomorphisms

For a fixed M, the map $N \mapsto \mathrm{Hom}(M,N)$ is a covariant functor from the category of modules to the category of abelian groups, in the following way. Given a morphism $f : N \to L$, we obtain a morphism f_* from $\mathrm{Hom}(M,N)$ to $\mathrm{Hom}(M,L)$ by setting $f_*(T) = f \circ T$. On the other hand, the map $M \mapsto \mathrm{Hom}(M,N)$ for a fixed N is a contravariant functor, for given $M \to L$, we obtain $f^* : \mathrm{Hom}(L,N) \to \mathrm{Hom}(M,N)$ by setting $f^*(T) = T \circ f$. Almost all the structure of the modules over a ring can be seen through the structure of the homomorphism groups,

Over the category of modules, we have kernels and images of homomorphisms, and so we can consider exact sequences of modules. We have a relationship between exact sequences of modules and exact sequences of their morphisms. A functor is called exact if it maps exact sequences to exact sequences. The homomorphism functors are not exact, but preserve exactness in certain useful scenarios.

Theorem 4.1. *A sequence $M_0 \to M_1 \to M_2 \to 0$ is exact if and only if $0 \to Hom(M_2,N) \to Hom(M_1,N) \to Hom(M_0,N)$ is exact for all modules N, and a sequence $0 \to N_0 \to N_1 \to N_2$ is exact if and only if $0 \to Hom(M,N_0) \to Hom(M,N_1) \to Hom(M,N_2)$ is exact for all modules M.*

Proof. Exercise. □

We now consider short exact sequences of modules

$$0 \to M \to N \to L \to 0$$

given by maps $f : M \to N$ and $g : N \to L$. We say such a diagram *splits* if the following of three equivalent conditions hold:

- There exists a section $\psi : L \to N$ such that $g \circ \psi$ is the identity.
- There exists a section $\eta : N \to M$ such that $\eta \circ f$ is the identity.

We now prove that if either of these conditions hold, then the other must hold, and then N decomposes as $\mathrm{Ker}(g) \oplus \mathrm{Im}(\psi)$ and as $\mathrm{Im}(f) \oplus \mathrm{Ker}(\eta)$, and is therefore isomorphic to $M \oplus L$. Surely the existence of ψ implies the direct sum decomposition, because $g(\psi(x)) = x$, so $\mathrm{Ker}(g)$ is disjoint from $\mathrm{Im}(\psi)$, and $x - \psi(g(x)) \in \mathrm{Ker}(g)$. The second condition implies the second

decomposition in a similar manner. To prove the equivalence of the two splitting conditions, we note that if $N = \mathrm{Ker}(g) \oplus \mathrm{Im}(\psi) = \mathrm{Im}(f) \oplus \mathrm{Im}(\psi)$, we can define $\eta(f(x) + \psi(y)) = x$, since f is injective. If $N = \mathrm{Im}(f) \oplus \mathrm{Ker}(\eta) = \mathrm{Ker}(g) \oplus \mathrm{Ker}(\eta)$, since g is surjective, setting $\psi(x)$ to be the unique element of $\mathrm{Ker}(\eta)$ with $g(\psi(x)) = x$. Such an element exists because g is surjective, and such an element is unique since if $\eta(x) = \eta(y) = 0$ and $g(x) = g(y)$, then $x - y = f(z)$, and so $0 = \eta(x) - \eta(y) = z$, so $x = y$.

Remark. If we are considering the short exact sequence

$$0 \to M \to M \oplus L \to L \to 0$$

Then the existence of the splitting maps is obvious. The splitting argument above shows that this situation is essentially the only case where a short exact sequence can split.

A module P such that for any short exact sequence

$$0 \to M \to N \to P \to 0$$

splits is known as a *projective module*. These are the modules that are easy to define maps out of.

Theorem 4.2. *Fix a module P. Then the following are equivalent.*

- *For any map $f : P \to N$ and a surjective map $g : M \to N$, there exists a map $h : P \to M$ such that $g \circ h = f$.*
- *Any short exact sequence $0 \to M \to N \to P \to 0$ splits.*
- *There is a module N such that $N \oplus P$ is a free module.*
- *The functor $M \mapsto Hom(P, M)$ is exact.*

If any of these conditions are satisfied, we say P is a projective module.

Proof. Consider the first condition. Then given any $0 \to M \to N \to P \to 0$, we can take $N = P$ in the triangle condition to obtain the splitting map, showing P is projective. If any short exact sequence terminating at P splits, there certainly exists a free module M with a surjective map $M \to P \to 0$, and then M is isomorphic to the direct sum of P and the kernel of the homomorphism. If N is a free module, then the functor $M \mapsto \mathrm{Hom}(N, M)$

is obviously exact, and since $\mathrm{Hom}(N \oplus P, M) \cong \mathrm{Hom}(N, M) \oplus \mathrm{Hom}(P, M)$. Thus if we have $M_0 \to M_1 \to M_2$ is exact, then

$$\mathrm{Hom}(N \oplus P, M_0) \to \mathrm{Hom}(N \oplus P, M_1) \to \mathrm{Hom}(N \oplus P, M_2)$$

is exact. Restricting the domain of our homomorphisms gives an exact sequence

$$\mathrm{Hom}(P, M_0) \to \mathrm{Hom}(P, M_1) \to \mathrm{Hom}(P, M_2)$$

Finally, assume that the functor is exact. Then we have an exact sequence $M \to N \to 0$, so we have an exact sequence $\mathrm{Hom}(P, M) \to \mathrm{Hom}(P, N) \to 0$. In particular, for any homomorphism from P to N, there exists a homomorphism from $P \to M$ which completes the triangle, so the first property is proved. □

Since projective modules are those for which it is easy to define maps out of, all free modules are projective. Assuming certain circumstances on the ring underlying the module, we can prove that all projective modules are free.

4.4 Modules over Principal Ideal Domains

We could proceed to define more of the general theory of modules, but this gets kind of dry. Instead, to develop our intuition, we use our knowledge of principal ideal domains to come up with a complete classification of the finitely generated modules over a principal ideal domain. In particular, this has applications to modules over the polynomial rings $k[X]$, the integers $\mathbf{Z}$ (and therefore to abelian groups). Thus we assume a principal ideal domain A has been fixed, and we study modules over it.

If A is a principal ideal domain and M is a module over A that is annihilated by a nonzero ideal $\mathfrak{a}$, then we can write $\mathfrak{a} = (a)$ for some $a \neq 0$. Thus $ax = 0$ for all $x \in M$, and if $tx = 0$ for all $x \in M$, then $t = ua$ for some $u \in A$. Perform a factorization, writing $a = p_1^{k_1} \dots p_n^{k_n}$ for primes $p_1, \dots, p_n$. If we write $M_i = \mathrm{Ann}(p_i^{k_i})$ for each $i \in \{1, \dots, n\}$, then $M_i \cap M_j = (0)$ for $i \neq j$. This is because if $p_i^{k_i} x = p_j^{k_j} x = 0$ for $i \neq j$, then since p_i and p_j are distinct primes, there is $a_1, a_1 \in A$ such that $(a_1 p_i^{k_i} + a_2 p_j^{k_j}) = 1$, and so we conclude $x = 1 \cdot x = 0$, which gives $x = 0$. By the Chinese remainder theorem, we may find $a_i \in A$ such that $a_i - 1 \in (p_i^{k_i})$ for each i and if $j \neq i$, then

$a_i \in (p_j^{k_j})$. Then for each i, $p_i^{k_i} a_i \in (a)$, so that for each $x \in M$, $x_i = a_i x$ lies in M_i. But our assumptions imply that $a_1 + \cdots + a_n - 1$ is divisible by a, which implies that $x = x_1 + \cdots + x_n$. Thus we conclude that $M = M_1 \oplus \cdots \oplus M_n$. The module M_i is known as the *p_i primary component of M*. This works over any torsion module in a principal ideal domain.

Theorem 4.3. *Every submodule of a finite dimensional free module is free, with dimension less than the dimension of the module it contains.*

Proof. We prove by induction. Consider the case where the free module is generated by a single element, so it is isomorphic to A. Any submodule is therefore an ideal of A, and is therefore of the form (a). If $a = 0$, then the submodule is free with dimension 0, and if $a \neq 0$, since A has no zero divisors, then a forms the basis of (a). Thus the theorem is true for one dimensional free modules. For the induction, if M is generated by a basis $x_1, \ldots, x_N, x_{N+1}$, then for a submodule N, then by induction, $N' = N \cap (x_1, \ldots, x_N)$ is freely generated by $y_1, \ldots, y_M$. We can consider the ideal in $\mathfrak{a}$ consisting of all a_{N+1} such that there are a_n such that $a_1 x_1 + \cdots + a_n X_N + a_{N+1} x_{N+1} \in N$. Thus $\mathfrak{a} = (a)$. If $a = 0$, $N \subset (x_1, \ldots, x_N)$, and we are done. Otherwise, we can pick some $y_{M+1} = a_1 x_1 + \cdots + a_N x_N + a x_{N+1}$, so that for any $x \in N$, there is a unique b such that $x - b y_{M+1} \in N'$, and so $x - b y_{M+1}$ has a unique expression as a sum of $y_1, \ldots, y_M$, showing $y_1, \ldots, y_{M+1}$ is a basis for N. □

Remark. If A has an infinite basis, then by well ordering it, and letting N denote an ifninite ordinal in the proof above, the proof essentially extends to the case of an arbitrary free module, except for the case of limit ordinals. But this is proven fairly easily from the fact that a direct limit of free modules is free.

If A is a commutative ring such that every submodule of a free module is free, then A must be a principal ideal domain, for if $\mathfrak{a}$ is a left ideal of A, then $\mathfrak{a}$ cannot contain more than a single independant element, since for any a and b, $ba - ab = 0$. Thus if $\mathfrak{a}$ is free, it must be principal, and if this is true for all $\mathfrak{a}$, then A is a principal ideal domain.

Corollary 4.4. *Every submodule of a f.g. module is f.g.*

Proof. If M is finitely generated, then M is the quotient of a free module. The correspondence theorem shows that every submodule of M corre-

sponds to the quotient of a submodule of the free module, which is therefore the quotient of a free module. □

The classification of finitely generated modules over a principal ring is essentially a generalization of the classification of finitely generated abelian group. The generalization of a finite abelian group is a finitely generated torsion module.

Lemma 4.5. *If M is a finitely generated torsion free module, then M is free.*

Proof. Out of a series of generators $x_1, \ldots, x_N$ for M, select a finite independent set $y_1, \ldots, y_M$ of maximal cardinality. Then for any x_n, there exists constants $a_n, b_1, \ldots, b_M$ such that $a_n x_n + b_1 y_1 + \cdots + b_M y_M = 0$. We must have $a_n \neq 0$, and if we consider $a = a_1 \ldots a_n$, then we find that for any x_n, $ax_n \in (y_1, \ldots, y_M)$, so $aM \subset (y_1, \ldots, y_M)$. But the map $x \mapsto ax$ embeds M in aM, so M is isomorphic to aM as a module, and as a submodule of a free module, aM is free. □

Theorem 4.6. *If M is finitely generated, M decomposes as a direct sum of M_{tor} and submodule of M isomorphic to M/M_{tor}.*

Proof. We have a short exact sequence

$$0 \to M_{\text{tor}} \to M \to M/M_{\text{tor}} \to 0$$

Then M/M_{tor} is finitely generated and torsion free, hence free, and therefore splits this diagram, giving the result. □

Thus every finitely generated module is uniquely the direct sum of a free module of a certain finite dimension, and a finitely generated torsion module. The dimension of the free module is known as the *rank* of the module. Two modules will be isomorphic if they isomorphic torsion submodules, and they have the same rank. It therefore suffices to classify the finitely generated torsion modules over a principal ring.

The remainder of the proof essentially carries over exactly the same as the classification of finite abelian groups. First, we note that if x is fixed, the *annihilators* of x, the elements $a \in A$ such that $ax = 0$, form a left ideal of A, and therefore is of the form (a). We call a a *period* for x. We set $M(p)$ to be the submodule of M consisting of all x such that $p^n x = 0$. Thus the annihilators of any element of $M(p)$ are of the form (p^n) for some n.

Theorem 4.7. *Let M be a finitely generated torsion module. Then M is the direct sum of $M(p)$, where p ranges over all equivalence classes of primes in A.*

Proof. Let a be such that $aM = 0$, which exists since M is finitely generated and is a torsion module. If $a = a_0a_1$, with $(a_0, a_1) = 1$, then $b_0a_0 + b_1a_1 = 1$ for some b_0 and b_1. We claim $M \equiv M_{a_0} \oplus M_{a_1}$, where M_{a_0} is the submodule of elements annihilated by a_0, and M_{a_1} the submodule annihilated by a_1. Given any $x \in M$, $x = b_0a_0x + b_1a_1x$, and $a_1(b_0a_0x) = 0$, $a_0(b_1a_1x) = 0$,so certainly $M = M_{a_0} + M_{a_1}$. But if $a_0x = 0$ and $a_1x = 0$, then any element of $(a_0, a_1) = (1)$ annihilates x, so in particular $x = 0$. Thus we really do have a direct sum representation. Carrying out the entire prime decomposition gives the direct sum result. □

Lemma 4.8. *Let M be a torsion module with $p^nM = 0$, and suppose x has period p^n, and suppose there are $y_1, \ldots, y_M$ such that $M/(x)$ can be written as $(y_1 + (x)) \oplus \cdots \oplus (y_M + (x))$. Then we can select y_n having the same period as $y_n + (x)$ and with $M = (y_1) \oplus \cdots \oplus (y_M) \oplus x$.*

Proof. Suppose $y + (x)$ has period p^m. Then $p^m y \in (x)$, so $p^m y = p^k cx$, where c does not divide p. If $k = n$, y has the same period as $y + (x)$, since $p^{m-1}y \notin (x)$, and so in particular is nonzero. Otherwise, $p^m y$ has period p^{n-k}, and so y has period p^{m+n-k}. We must have $m + n - k \leqslant n$, hence $m \leqslant k$, and so $y' = y - p^{k-m}cx$ has y' the same period as $y + (x)$ and $y' + (x) = y + (x)$. Thus given $y_1 + (x), \ldots, y_M + (x)$, we may assume y_n has the same period as $y_n + (x)$. Suppose that $ax + a_1y_1 + \cdots + a_my_M = 0$. Then $a_ny_n \in (x)$ for all n, hence a_n is divisible by the period of y_n, and so in particular, $a_ny_n = 0$. Thus $ax = 0$. This completes the proof of the decomposition. □

Theorem 4.9. *For any finitely generated p module M, there is a sequence of integers such that*

$$M \equiv A/(p^{n_1}) \oplus \cdots \oplus A/(p^{n_m})$$

and $n_1 \geqslant \cdots \geqslant n_m$.

Proof. We prove by induction on the maximal exponent of M. If p is the exponent of M, then $M = M_p$ is a vector space over $A/(p)$, and therefore isomorphic to a direct sums of $A/(p)$. Now we prove the theorem for p^{N+1} by a second induction on the dimension of M_p over $A/(p)$. If $M_p = (0)$,

then $M(p)$ is isomorphic to $pM(p)$, which by induction has a required decomposition. Otherwise, consider x with maximal period p^N. If we consider $M' = M/(x)$, then we contend that M'_p has dimension strictly less than M_p over $A/(p)$. If we consider a basis $y_1 + (x), \ldots, y_M + (x)$ for M'_p, then we can be the last lemma assume $py_n = 0$, then $p^{N-1}x, y_1, \ldots, y_M$ are linearly independant over M_p. Thus by induction M' is the direct sum of $(x_1) \oplus \cdots \oplus (x_M)$ for some M, and the last lemma implies we can choose $x_1, \ldots, x_M$ such that $M = (x) \oplus (x_1) \oplus \cdots \oplus (x_M)$. The decreasing integer condition then holds. □

Such a decomposition is unique, as proved by this next lemma which reduces the argument to a different, equivalent decomposition of a finitely generated torsion module M.

Lemma 4.10. *If M is a finitely generated torsion module, then M is uniquely isomorphic to a direct sum of $A/(n_1) \oplus \cdots \oplus A/(n_N)$, where n_1 divides n_2, which divides n_3, and so on up to n_N.*

Proof. We decompose M into $M(p)$, and decompose $M(p)$ into the modules $A/(p_i^{n_{ij}})$, with n_{ij} increasing in j. Arrange these modules into a matrix with rows i and columns j. The existence of the result is obtained by taking the direct sum of modules over each column, since the direct sum of cyclic modules with relatively prime exponents is also cyclic. To prove uniqueness, consider a decomposition of M as $A/(n_1) \oplus \cdots \oplus A/(n_N)$. If $x = x_1 \oplus \cdots \oplus x_N$, then $x \in M_p$ if and only if $px_n = 0$ for all n, so M_p is the direct sum of $(A/n_m)_p$. In particular, the dimension of M_p over $A/(p)$ is precisely the number of n_m such that p divides n_m. Given another sequence $m_1, \ldots m_M$, such that M is decomposed as $A/(m_1) \oplus \cdots \oplus A/(m_M)$, consider a prime p such that p divides m_1, hence all m_n, we conclude that the dimension of M_p is M, and so $N \geqslant M$. By symmetry $N = M$, and also p divides all the n_n and m_n. The module pM is decomposed by removing a factor of p from each element in the decomposition. Once we continue this process, we reach relative primality, hence we must reach relative primality in the other decomposition by the dimensionality argument, and continuing this process shows that up to a unit, all the n_n were uniquely specified. □

Example. *Let G be a finitely generated abelian group. Then, since $\mathbf{Z}$ is principal, all of our results apply. Thus G has a rank n, and is isomorphic to the*

product of $\mathbf{Z}^n$ and a finite abelian group, i.e. a finitely generated torsion module over $\mathbf{Z}$. To classify the finite abelian groups, we can either prick integers n_1 divides n_2 dividing up to n_M, and considering $\mathbf{Z}_{n_1} \oplus \cdots \oplus \mathbf{Z}_{n_M}$, or we can decompose our abelian groups into products of cyclic groups of prime order.

Example. *Let T be an endomorphism of a vector space V. Then V has the structure of a $k[X]$ module. If V is finite dimensional, then V is a torsion module over $k[X]$, which is a principal ring, so there exists a unique decomposition of V into T invariant subspaces, and on each subspace W in the decomposition, there exists an irreducible polynomial f such that W is isomorphic to $k[X]/(f^N)$. If k is algebraically closed, then there is an element $\lambda \in k$ such that W is isomorphic to $k[X]/((X-\lambda)^N)$, and the elements v_n corresponding to $(X-\lambda)^n$ are a vector space basis over k of W, and $Tv_n = \lambda v_n + v_{n+1}$. Thus we obtain the Jordan normal form an endomorphism over an algebraically closed field. Over the real numbers, we either have a T invariant subspace of this form, or isomorphic as a $\mathbf{R}[X]$ module to $\mathbf{R}[X]/((X^2 + 2Re(\lambda)X + |\lambda|^2)^N)$ for some N and complex λ.. TODO: GIVING RISE TO THE REAL CANONICAL FORM OF A MATRIX, where the jordan blocks are formed by a basis $v_1, w_1, v_2, w_2, \ldots, v_n, w_m$, with $T(v_n + iw_n) = \lambda(v_n + iw_n) + (v_n + iw_n)$. Thus the Jordan blocks corresponds to rotation matrices and two by two identity matrices.*

Part I

Commutative Algebra

Chapter 5

Nilradicals

Recall that the nilradical of a commutative ring A is the ideal $\sqrt{A}$ of all *nilpotent* x, i.e. those elements with $x^n = 0$. The reason for the interest in a nilradical is that it removes nilpotent elements from the quotient. We now show a way to generalize the nilradical of a commutative ring to noncommutative cases, by showing the nilradical is equivalent to another construction. The *Jacobson radical* $J(R)$ of a (not necessarily commutative) ring R to be the intersection of all prime ideals in the ring. In the commutative case, we find $J(R) = \sqrt{R}$.

Theorem 5.1. *If A is a commutative ring, $J(A) = \sqrt{A}$.*

Proof. If $\mathfrak{a}$ is a prime ideal, and $x^n = 0$, then $x^n \in \mathfrak{a}$, hence $x \in \mathfrak{a}$, showing $\sqrt{A} \subset J(A)$. Conversely, suppose $x \notin \sqrt{A}$. Consider the set S of all powers x^n. Let L be the set of all (not necessarily prime) ideals in A disjoint from S. Then L is nonempty, since (0) is in L, and L is inductively ordered, so we can consider some maximal element $\mathfrak{a}^*$. Given $a, b \notin \mathfrak{a}^*$, $\mathfrak{a}^* + (a)$ and $\mathfrak{a}^* + (b)$ are both strictly larger than $\mathfrak{a}^*$, and so there is x_1, y_1 and x_2, y_2 such that $x_1 + ay_1 = x^n$ and $x_2 + by_2 = x^m$. But then

$$x^{m+n} \in (\mathfrak{a}^* + (a))(\mathfrak{a}^* + (b)) = \mathfrak{a}^* + (a)\,\mathfrak{a}^* + (b)\,\mathfrak{a}^* + (ab)$$

And therefore $ab \notin \mathfrak{a}^*$, so $\mathfrak{a}^*$ is prime, not containing x, and so $J(A)$ does not contain x. □

5.1 Direct and Inverse Limits

Let I be a directed index set, and consider a family of rings $\{A_i : i \in I\}$, and for each $i \leqslant j$ in I, a homomorphism $f_{ij} : A_i \to A_j$ such that $f_{jk} \circ f_{ij} = f_{ik}$. We can define a universal object from this construction, the *direct limit* $\lim_{i\to\infty} A_i$, together with morphisms $f_i : A_i \to \lim_{i\to\infty} A_i$, which has the universal property that any family of morphisms $\phi_i : A_i \to B$ such that $\phi_j \circ f_{ij} = \phi_i$ for each $i \leqslant j$, then there exists $\phi : \lim_{i\to\infty} A_i \to B$ such that $\phi_i = \phi \circ f_i$ for each i.

One can clearly consider the direct limit of the rings $\{A_i\}$ as Abelian groups, but this direct limit also has a ring structure, which can be defined as follows; given $x_1, x_2 \in \lim_i A_i$, there exists an index i and $a_1, a_2 \in A_i$ such that $f_i(a_1) = x_1$ and $f_i(a_2) = x_2$. We define $x_1 x_2 = f_i(a_1 a_2)$. This is independant of the choice of a_1 and a_2, since if $a_1', a_2' \in A_i$ are selected with $f_i(a_1) = f_i(a_1')$ and $f_i(a_2) = f_i(a_2')$, then there exists an index $k \geqslant i$ such that $f_{ik}(a_1) = f_{ik}(a_1')$, $f_{ik}(a_2) = f_{ik}(a_2')$, and then

$$\begin{aligned} f_i(a_1 a_2) &= f_k(f_{ik}(a_1 a_2)) \\ &= f_k(f_{ik}(a_1) f_{ik}(a_2)) \\ &= f_k(f_{ik}(a_1') f_{ik}(a_2')) \\ &= f_k(f_{ik}(a_1' a_2')) \\ &= f_i(a_1' a_2'). \end{aligned}$$

Moreover, the compatibility condition $f_j \circ f_{ij} = f_i$ ensure that this definition is independent of the index i selected. Thus this operation gives $\lim_i A_i$ a well defined ring structure such that each f_i is a ring homomorphism. If we consider a family of ring homomorphisms $\phi_i : A_i \to B$ compatible with the family of maps $\{f_{ij}\}$, then the induced group homomorphism $\phi : \lim_i A_i \to B$ is also a ring homomorphism, since given $x_1, x_2 \in \lim_i A_i$, there is an index i and $a_1, a_2 \in A_i$ such that $\phi_i(a_1) = x_1$, $\phi_i(a_2) = x_2$, so $\phi_i(a_1 a_2) = x_1 x_2$, and

$$\phi(x_1 x_2) = \phi_i(a_1 a_2) = \phi_i(a_1)\phi_i(a_2) = \phi(x_1)\phi(x_2).$$

Thus the direct limit as Abelian groups is the same as the direct limit as rings.

Example. *Consider a topological space X, and fix a point $x \in X$. For each open set U containin x, let $C(U)$ consist of the ring of scalar valued continuous*

functions on U. Given $V \subset U$, we have a map $R_{UV} : U \to V$, such that $R_{UV} f$ is the restriction of f to V. Then the family $\{R_{UV}\}$ are ring homomorphisms. Thus we can consider the direct limit of this family, often denoted $\mathcal{O}_x$, which is the ring of germs *of functions defined in a neighbourhood of x. The ring $\mathcal{O}_x$ is* local*, in the sense that it has a unique maximal ideal $\mathfrak{m}_x$, which consists of the germs of continuous function f defined in a neighbourhood of x with $f(x) = 0$. This is because every element of $\mathcal{O}_x - \mathfrak{m}_x$ is invertible in $\mathcal{O}_x$, since if f is a continuous function defined in a neighbourhood of x with $f(x) \neq 0$, then f has no zeroes in a neighbourhood of x, so $g(x) = 1/f(x)$ is well defined in this negihbourhood and fg equals one in a neighbourhood of the origin.*

Dual to the construction of direct limits is the construction of inverse limits. Given a family of rings $\{A_i : i \in I\}$ with maps $f_{ji} : A_j \to A_i$ for each $i \leqslant j$, we can construct an inverse limit $\lim_i A_i$ together with morphisms $f_i : \lim_i A_i \to A_i$, which has the universal property that for each families of morphisms $\phi_i : B \to A_i$ such that $f_{ji} \circ \phi_j = \phi_i$ for each $i \leqslant j$, there is a unique morphism $\phi : B \to \lim_i A_i$ such that $\phi_i \circ f_i = f$. Inverse limits exist in the category of abelian groups, and just as with direct limits it is easy to equip the inverse limit with an additional multiplication structure since all the maps involved here are ring homomorphisms.

Example. *For each prime p, let $A_i = \mathbf{Z}/p^i\mathbf{Z}$. For $i \leqslant j$, $(p^j) \subset (p^i)$, so we can consider a morphism $\mathbf{Z}/p^j\mathbf{Z} \to \mathbf{Z}/p^i\mathbf{Z}$. These morphisms are compatible, so they give rise to an inverse limit $\lim_i \mathbf{Z}_{p^i}$, which we denote by $\mathbf{Z}_p$ and call the ring of* p-adic integers*. For each $x \in \mathbf{Z}_p$, we can uniquely associate a formal series*

$$x = b_0 + b_1 p + b_2 p^2 + \dots$$

with $b_i \in \{0, \dots, p-1\}$, such that for each i, $f_i(x) = b_0 + \cdots + b_{i-1}p^{i-1}$. If we define $A = \mathbf{Z}[[X]]/(X - p)$, this this association gives a homomorphism from $\mathbf{Z}_p$ to A. Any element of A can be uniquely written as

$$\sum_{k=0}^{\infty} a_k p^k$$

for if

$$\left(\sum_{k=0}^{\infty} b_k X^k \right) (X - p) = \sum_{k=0}^{\infty} a_k p^k$$

then b_0

Let $A = \mathbf{Z}[[X]]/(X-p)$. *The surjective morphisms* $\phi_i : A \to \mathbf{Z}/p^i\,\mathbf{Z}$ *defined by setting*

$$\phi_i\left(\sum_{k=0}^{\infty} a_k X^k\right) = \sum_{k=0}^{i-1} a_k p^k$$

are well defined, and it is easy to see that $f_{ji} \circ \phi_j = \phi_i$. *Thus we obtain a map* $\phi : A \to \mathbf{Z}_p$. *Any element of* A *may be written as*

$$\sum_{k=0}^{\infty} a_k X^k,$$

where $a_0, \ldots, a_{i-1} \in \{0, \ldots, p-1\}$, *and if* p^i *divides*

$$\sum_{k=0}^{i-1} a_k p^k,$$

then we must conclude $a_0 = \cdots = a_{i-1} = 0$. *Thus the kernel of* ϕ_i *in* A *is* (X^i). *Since* $\bigcap_{i=1}^{\infty}(X^i) = (0)$, ϕ *is injective. But it is also surjective, since for any element* $x \in \mathbf{Z}_p$ *we may associate a sequence of integers* $b_0, b_1, \cdots \in \{0, \ldots, p-1\}$ *such that for each* i, $f_i(x) = b_0 + \cdots + b_{i-1}p^{i-1}$, *and then*

$$\phi\left(\sum_{k=0}^{\infty} b_k X^k\right) = x.$$

Thus A *is isomorphic to* $\mathbf{Z}_p$, *and so we may think of* $\mathbf{Z}_p$ *as the ring of 'formal power series' in* p, *of the form*

$$a_0 + a_1 p + a_2 p^2 + \ldots$$

with $a_i \in \{0, \ldots, p-1\}$ *for each* i.

The ideal $(X-p)$ *is prime in* $\mathbf{Z}[[X]]$. *To see this, since* $\mathbf{Z}[[X]]$ *is a unique factorization domain (implied by the fact that* $\mathbf{Z}$ *is principal), it suffices to show* $X-p$ *is irreducible. So suppose*

$$X - p = \left(\sum_{k=0}^{\infty} a_k X^k\right)\left(\sum_{k=0}^{\infty} b_k X^k\right).$$

Then $-p = a_0 b_0$, which implies either a_0 or b_0 is a unit since p is prime. But if a_0 is a unit, then $\sum_{k=0}^{\infty} a_k X^k$ is a unit in $\mathbf{Z}[[X]]$. Thus $X - p$ is irreducible, and so $\mathbf{Z}_p$ is an integral domain.

If we consider some element x of $\mathbf{Z}_p$ represented as the power series

$$\sum_{k=0}^{\infty} a_k p^k,$$

with $a_0 \neq 0$, then there exists some integer n such that na_0 is congruent to 1 modulo p, and so nx corresponds to a unit in $\mathbf{Z}[[X]]$. In particular, nx is invertible, so x is invertible in A. Thus we conclude that $\mathbf{Z}_p$ contains a unique maximal ideal of the form

$$\mathfrak{m}_p = \left\{ \sum_{k=0}^{\infty} a_k p^k : a_0 = 0 \right\}.$$

Thus $\mathbf{Z}_p$ is a local ring.

Now consider the equation $X^{p-1} - 1$ in $\mathbf{Z}_p$. Suppose a_0 is an integer not divisible by p. Then $a_0^{p-1} \equiv 1$ (modulo p). We now construct a sequence of integers $\{a_i\}$ such that for each i,

$$\left(\sum_{k=0}^{i-1} a_k p^k \right)^{p-1} \equiv 1 \quad (\textit{modulo } p^i) \tag{5.1}$$

We construct this sequence inductively. Suppose we have chosen $a_0, \ldots, a_{i-1}$ such that (5.1) holds. Then we must find an integer a_i such that

$$\left(\sum_{k=0}^{i} a_k p^k \right)^{p-1} \equiv 1 \quad (\textit{modulo } p^{i+1}). \tag{5.2}$$

Now

$$\left(\sum_{k=0}^{i} a_k p^k \right)^{p-1} = \sum_{j=0}^{p-1} \binom{p-1}{j} \left(a_i p^i \right)^j \left(\sum_{k=0}^{i-1} a_k p^k \right)^{p-1-j}$$

which is equivalent modulo p^{i+1} to

$$\left(\sum_{k=0}^{i-1} a_k p^k \right)^{p-1} + a_i (p-1) p^i \left(\sum_{k=0}^{i-1} a_k p^k \right)^{p-2}.$$

We can find an integer n such that

$$\left(\sum_{k=0}^{i-1} a_k p^k\right)^{p-1} = 1 + np^i.$$

The inductive case guarantees that $\sum_{k=0}^{i-1} a_k p^k$ *is relatively prime to p, hence this quantity has an inverse m modulo* p^{i+1}*, so* (5.2) *is equivalent to the equation*

$$p^i \left(n + a_i(p-1) \left(\sum_{k=0}^{i-1} a_k p^k\right)^{p-2} \right) \equiv 0 \quad (\textit{modulo } p^{i+1}).$$

Thus it clearly suffices to show that we can pick $a_i \in \{0, \ldots, p-1\}$ *such that*

$$n + a_i(p-1) \left(\sum_{k=0}^{i-1} a_k p^k\right)^{p-2} \equiv 0 \quad (\textit{modulo } p),$$

and this equation is solvable since $p-1$ *and* $\sum_{k=0}^{i-1} a_k p^k$ *are relatively prime to p. This completes the construction. Performing this construction for all* $a_1 \in \{1, \ldots, p-1\}$ *constructs* $p-1$ *distinct values* $x_1, \ldots, x_{p-1} \in \mathbf{Z}_p$ *such that* $x_i^{p-1} = 1$ *for each i. Thus the polynomial equation* $X^{p-1} - 1$ *splits over* $\mathbf{Z}_p$*.*

Chapter 6

Localization

In many situations, we study a commutative ring A with identity, and wish to invert elements of the ring which aren't necessarily units. Thus, given an element $a \in A$, we wish to consider the most 'free' morphism $f : A \to B$ such that $f(a)$ has an inverse.

may wish to embed A in a larger ring B in which a has an inverse. Unfortunately, this is not always possible. For instance, if $a^2 = 0$, then we cannot possibly embed A in such a way that makes a invertible. More generally, if $f : A \to B$ is a homomorphism in which $f(a)$ is invertible, and $ab = 0$, then we must have $f(b) = 0$. This implies that if we desire f to be injective, then the ring A cannot have any zero divisors. Nonetheless, if we remove this condition, then the only condition that prevents $f(a)$ from having an injective, then the only condition that prevents $f(a)$ from having an inverse is if $a = 0$. Identifying certain maps by a not necessarily injective map f is a process in algebra we now called localization.

The classical situation where we can localize is in the case where A is an integral domain, in which case the problems of zero divisors disappear completely. In this case, we can embed A into it's *field of fractions* B, which consist of formal quotients a/b, with $b \neq 0$, where a/b is identified with c/d if $ad - bc = 0$. After identification, we can define a multiplication and addition operation by setting $(a/b)(c/d) = ac/bd$, and by setting $(a/b) + (c/d) = (ad+bc)/bd$. It is simple to check these operations are well defined on B. Then B is given the structure of a commutative ring in which every nonzero element has an inverse. Thus B is not only and ring, but a field! We embed A by mapping a to the formal quotient $a/1$.

Example. *The localization of* $\mathbf{Z}$ *produces a field of fractions which is obviously just the rational numbers* $\mathbf{Q}$ *in disguise. Constructing the field of fractions over an integral domain is essentially just a generalization of this process.*

Example. *If* $A[X_1,\ldots,X_n]$ *is a polynomial ring with coefficients in some integral domain* A*, then the polynomial ring is an integral domain, and performing localization gives the field* $k(X_1,\ldots,X_n)$ *of rational functions over the field of fractions* A*, which consists of all finitary expressions of the form*

$$\frac{\sum a_\alpha X^\alpha}{\sum b_\beta X^\beta}$$

These can be considered as functions mapping certain 'nonsingular' elements of A *into it's field of fractions* k*. In particular,* f/g *is defined at* $x \in k$ *if* $g(a) \neq 0$*, because then the quotient* $f(a)/g(a) = f(a)g(a)^{-1}$ *is well defined. As an example, the field of fractions of* $\mathbf{Z}[X_1,\ldots,X_n]$ *is the field* $\mathbf{Q}(X_1,\ldots,X_n)$ *of rational functions over the rationals.*

Example. *Let* $A(D)$ *denote the complex algebra of functions holomorphic in some connected open region* D *of* $\mathbf{C}$*. Then* $A(D)$ *is an integral domain, for if* $fg = 0$*, where* $f, g \neq 0$*, then* $f^{-1}(0)$ *and* $g^{-1}(0)$ *are two discrete sets whose union is* D*, which is impossible. We may therefore form the field of fractions of* $A(D)$*, which is precisely the set of meromorphic functions on* D*. These functions* f/g *are defined except for certain points upon which* $g(z) = 0$*, except in the case that* z *is a removable singularity of* g*, which means that we can write* $f/g = f_1/g_1$*, where* $g_1(z) \neq 0$*.*

Considering this problem in a more general viewpoint, we consider a set $S \subset A$, and try to find the 'most general' homomorphism $f : A \to B$ such that $f(s)$ is invertible for each $s \in S$. If $f(s)$ and $f(t)$ are invertible, then $f(st) = f(s)f(t)$ is invertible, so we may assume from the outset that S is closed under multiplication. We may also assume that $1 \in S$, because $f(1)$ is always invertible. In this case, S is a multiplicative submonoid of A, which we call a *multiplicative set*. By 'localizing' S, we mean extending A to a space B in which all elements of S have an inverse. By a localization of A by S, we mean a ring $S^{-1}A$ together with a map $i : A \to S^{-1}A$ such that for any homomorphism $f : A \to B$ such that $f(s)$ is invertible for each $s \in S$, there is a unique homomorphism $S^{-1}f : S^{-1}A \to B$ for which $f = S^{-1}f \circ i$. This is an initial object in a certain category, and is therefore unique up to isomorphism.

More generally, suppose that a commutative ring A has zero divisors. Then forming the field of fractions is impossible – we cannot give every element of A an inverse simultaneously. More generally, we might hope to find the 'most general' homomorphism $i : A \to S^{-1}A$ such that $i(s)$ is invertible for each element s in some multiplicative set S. In particular, we hope to find an object i and $S^{-1}A$ such that for *any* homomorphism $f : A \to B$ into a commutative ring B such that $f(s)$ is invertible for each $s \in S$, there is a homomorphism $S^{-1}f : S^{-1}A \to B$ such that $f = S^{-1}f \circ i$. This is an initial object in the category of homomorphisms from A into some other ring B which map S to units, which means it is unique up to isomorphism.

Often, the correct technique to finding a universal object is to determine what properties the object must have, and then trying to form a formal structure based on these properties. Given what we know, this object will either fail to be constructed in general, in which case we must try and find more properties of the object, or the formal object we construct will often be the required universal object. Let us try and derive what our initial object $S^{-1}A$ should be 'forced to have'. Note that if $f : A \to S^{-1}A$ is the required morphism, then the set B of elements of $S^{-1}A$ of the form $i(a)i(s)^{-1}$, for $a \in A$ and $s \in S$ is a subring of $S^{-1}A$ (an easy calculation left to the reader). This means that $i : A \to B$ is a map in which each $f(s)$ is invertible, and so there must be a map $S^{-1}i : S^{-1}A \to B$ such that $i = S^{-1}i \circ i$. Clearly $S^{-1}i$ must be the identity map, which implies $B = S^{-1}A$. Now, let us determine when $i(a)i(s)^{-1} = i(b)i(t)^{-1}$. If this is true, then $i(at - bs) = 0$. One condition guaranteeing this to be true is if there is $u \in S$ for which $u(at - bs) = 0$, because then $f(u)f(at - bs) = 0$, and multiplying by $f(u)^{-1}$ gives the required property. It turns out that these properties are sufficient to formally define $S^{-1}A$.

Consider the set $S^{-1}A$ whose objects are fractions a/s, as in the field of fractions of an integral domain, but where $a \in A$ and $s \in S$. We identify two fractions a/s and b/t if there is an element $u \in S$ such that $u(at - bs) = 0$. We define multiplication by setting $(a/s)(b/t) = (ab/st)$, and addition by $a/s + b/t = (at + bs)/ts$. This gives $S^{-1}A$ a ring structure, and we have a map $i : A \to S^{-1}A$ given by $i(a) = a/1$, and then $i(s)^{-1} = 1/s$. If $f : A \to B$ is any ring homomorphism in which $f(s)$ is invertible for each $s \in S$, then we can define $S^{-1}f : S^{-1}A \to B$ by $S^{-1}f(a/s) = f(a)f(s)^{-1}$, and then it is a simple procedure to verify that the required diagram commutes, and that f is unique. Thus $S^{-1}A$ is exactly the initial object we required.

Example. *Let X be a topological space, and let $C(X)$ denote the ring of all (real/complex valued) continuous functions defined on X. If $p \in X$, then set the set S of all functions f with $f(p) \neq 0$ is a multiplicative set containing 1, closed under multiplication, and not containing 0. Thus we can consider the localization $S^{-1}C(X)$, which we denote by $C(X)_p$. Since $C(X)$ is almost never an integral domain, the map $C(X) \to C(X)_p$ will likely not be injective. Indeed, two functions f and g will be identified in $C(X)_p$ if there is a function h with $h(p) \neq 0$, and with $h(f-g) = 0$. Since $h(p) \neq 0$, the set of points q where $h(q) \neq 0$ contains an open neighbourhood of zero, and this implies that $(f - g)(q) = 0$ on this neighbourhood. Conversely, it suitably nice topological spaces (where Urysohn's theorem applies), if f agrees with g in a neighbourhood of p, we can find a function h such that h vanishes outside this neighbourhood, and then $h(f-g) = 0$. Thus functions are identified in $C(X)_p$ precisely when they are locally equal around p, and this is the context in which the term localization emerged, because localization takes a ring of functions, and identifies those functions which locally agree. More generally, if we set S to be the set of all functions with $f(p) \neq 0$ for all p in some $Y \subset X$, then $C(X)_Y$ consists of the equivalence class of all functions which agree on a neighbourhood of Y, provided we can construct functions vanishing outside of a neighbourhood of Y, with no zeroes on Y.*

Example. *Similarily, if M is a differentiable manifold, then the space $C^\infty(M)$ of (real/complex valued) differentiable functions on M forms a ring. For a fixed $p \in M$, the space of functions not vanishing at p forms a multiplicative set, and the corresponding localization corresponds to the equivalence class of differentiable functions which agree in a neighbourhood of p, known as the space of germs of differentiable functions at p. Viewed as a vector space over the real numbers, the dual space of germs of differentiable functions is used to construct the tangent space of a manifold at a point. A similar process is used to construct the germ of analytic functions on an analytic/holomorphic manifold, where we replace $C^\infty(M)$ with $C^\omega(M)$.*

Perhaps this formal approach is not so intuitive from a more geometric perspective. There is a more 'natural' approach to forming $S^{-1}A$, but it is much more messy. When learning fractions for the first time, you viewed them as ways to 'divide' certain integers into other integers. If you have 6 apples, you can 'apply' the fraction $1/2$ to divide the apples into two sets of three apples, the fraction $1/3$ to divide the 6 apples into three sets of two, but one cannot apply the fraction $1/5$. In other words, we can view a

fraction $1/n$ as a partial function on $\mathbf{Z}$ (defined on $n\mathbf{Z}$, to be precise), which outputs m when given input nm. Similarily, n/m is the partial function defined on the set of integers k such that nk is divisible by m, in which case applying n/m to k results in nk/m. It seems reasonable to set fractions equal if they agree on the common input upon which they are defined. That is, we should set $1/2 = 2/4$, because they have the same domain, and are equal to one another on this domain. To abstract these ideas to form $S^{-1}A$, we let Φ denote the set of all A-module homomorphisms from $(s) \to A$, for some $s \in S$. We then form a family of equivalence classes on Φ by identifying $f : (s) \to A$ and $g : (t) \to A$ if f and g agree on (st). On these equivalence classes, we can define addition between $f : (s) \to A$ and $g : (t) \to A$ by letting $f + g$ be the addition of the functions as morphisms from (st) to A. Similarily, we define fg to be $f \circ g$, once f and g are restricted to the proper ideals. We then embed A in Φ by mapping $a \in A$ to the 'multiplication by a' homomorphism from A to itself. Given $s \in S$, the inverse of s is the homomorphism with domain (s) mapping sa to a. Unfortunately, if A has zero divisors, then this approach does not work, in which case one must first quotient A by the ideal of all elements of A which are annihilated by elements of S.

Remark. Localization can be done in noncommutative rings. However, the resulting rings $S^{-1}A$ are extremely nontrivial to analyze, and as such we do not consider them. This follows because expressions of the form $rs^{-1}t + uv^{-1}w$ cannot in general be reduced to having a single common denominator. Thus one may have to repeat the process of localization many times to obtain inverses for all elements of S, and even if we repeat the process finitely many times we may still not end up with all the right inverses. What's more, even if A has no zero divisors, it can still be difficult to determine if the localization of A is nontrivial. However, one can in certain situations achieve success, by generalizing the 'partial homomorphism' technique of the last paragraph. The general technique is known as Ore localization, and is left for another time.

Finally, we remark that localization acts not only on a ring, but also on the modules over a ring. Given a module M over a ring A with multiplicative set S, we can find a natural module $S^{-1}M$ over $S^{-1}A$ which is the initial module in the category of A module maps $f : M \to N$ such that for any $s \in S$, the map $x \mapsto sx$ is an isomorphism of N. To construct this initial object, we consider elements of the form x/s, with $x \in M$ and $s \in S$,

and we identify x/s and y/s' if there is $s'' \in S$ such that $s''(xs' - ys) = 0$. The map $M \mapsto S^{-1}M$ is a functor from the category of A modules to $S^{-1}A$ modules, since if $f : M \to N$, then given the map $i : N \to S^{-1}N$, the map $f \circ i$ induces a morphism of $S^{-1}M$ with $S^{-1}N$.

As we have seen, a ring need not embed in its localization when the ring has zero divisors. In fact, zero divisors precisely describe when the embedding exists. Suppose that the annihilators of any $s \in S$ in M are trivial. Two elements $x, y \in M$ are identified in $S^{-1}M$ if and only if there is $s \in S$ such that $s(x - y) = 0$, so $x - y = 0$, hence $x = y$. Conversely, if $sx = 0$, then x is identified with zero in $S^{-1}M$. In particular, if S has no zero divisors in A, then A embeds in $S^{-1}A$.

6.1 Properties Preserved Under Localization

The universal description of the localization of a ring allows us to prove many useful properties of the localization. Given any family of A modules M_α, we find that $S^{-1}(\bigoplus M_\alpha)$ is isomorphic to $\bigoplus S^{-1}M_\alpha$ in a way that the embeddings of $\bigoplus M_\alpha$ in these rings corresponds with one another. A morphism $f : \bigoplus M_\alpha \to N$ corresponds to a unique sequence of morphisms $f_\alpha : M_\alpha \to N$. Because of the properties of localization, f_α extends uniquely to a morphism $f_\alpha : S^{-1}M_\alpha \to N$, and hence f extends unique to a morphism from $\bigoplus S^{-1}M_\alpha \to N$. Thus this direct sum has the properties of localization which $S^{-1}\bigoplus M_\alpha$ possesses, so they are both isomorphic in the required manner.

Remark. Since finite products and coproducts of modules correspond, this identity also holds if we swap the direct sum operation with a *finite* direct product. Unfortunately, this need not be true for infinite direct products. If we consider the fraction field of the integers generated by $S = \mathbf{Z} - \{0\}$, with $M_1 = M_2 = \cdots = \mathbf{Z}$ then the two rings we get from the direct product above are $S^{-1}(\mathbf{Z}^\infty)$ and $\mathbf{Q}^\infty$. The inclusion of $\mathbf{Z}^\infty$ in $\mathbf{Q}^\infty$ certainly identifies $S^{-1}(\mathbf{Z}^\infty)$ with a subspace of $\mathbf{Q}^\infty$, but this subspace is proper; it consists of all infinite sequences of rational numbers with bounded denominator. Since $S^{-1}(\mathbf{Z}^\infty)$ is countable, whereas $\mathbf{Q}^\infty$ is uncountable, these spaces cannot be isomorphic.

We have a map $\mathfrak{a} \to S^{-1}\mathfrak{a}$ from the ideals in A to the ideals in $S^{-1}A$,

such that $S^{-1}\mathfrak{a}$ is the ideal generated by $i(\mathfrak{a})$. We can also described it as

$$S^{-1}\mathfrak{a} = \{a/s : a \in \mathfrak{a}, s \in S\}$$

The map also has nice algebraic properties, in that it represents sums, products, and intersections of ideals, so

$$S^{-1}(\mathfrak{a}+\mathfrak{b}) = S^{-1}\mathfrak{a} + S^{-1}\mathfrak{b} \quad S^{-1}(\mathfrak{a}\mathfrak{b}) = (S^{-1}\mathfrak{a})(S^{-1}\mathfrak{b})$$

$$S^{-1}(\mathfrak{a}\cap\mathfrak{b}) = (S^{-1}\mathfrak{a})\cap(S^{-1}\mathfrak{b})$$

and respects inclusions. Every ideal in $S^{-1}A$ is of the form $S^{-1}\mathfrak{a}$, because

$$S^{-1}(i^{-1}(\mathfrak{a})) = \{a/b : a \in \mathfrak{a}, b \in S\} = \mathfrak{a}$$

Thus localization doesn't add any new ideal structure to a ring.

Proposition 6.1. *If A is principal, then $S^{-1}A$ is principal.*

Proof. This follows because all ideals in A are of the form $S^{-1}\mathfrak{a}$, and if $\mathfrak{a} = (a)$, then $S^{-1}\mathfrak{a} = (a)$. □

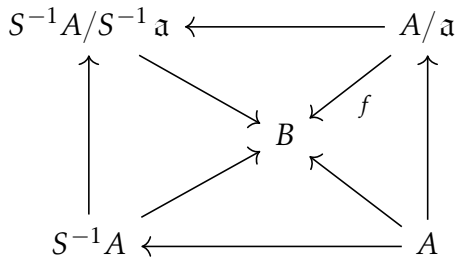
Proposition 6.2. *If A is Noetherian, then $S^{-1}A$ is Noetherian.*

Proof. If $S^{-1}(\mathfrak{a}_0) \subset S^{-1}(\mathfrak{a}_1) \subset \ldots$ is a chain of ideals in $S^{-1}A$, then $\mathfrak{a}_0 \subset \mathfrak{a}_1 \subset \ldots$, so eventually the $\mathfrak{a}_N$ are constant, so $S^{-1}\mathfrak{a}_N$ are constant. □

Because localization relates to a universal property of rings, it respects many of the useful transformations in the category of rings.

Proposition 6.3. *If A is a ring, S is a multiplicative subset, and $\mathfrak{a}$ is an ideal containing no elements in common with S, then $S^{-1}A/S^{-1}\mathfrak{a}$ is isomorphic to $(S/\mathfrak{a})^{-1}(A/\mathfrak{a})$, in a way which preserves the embedding of $A/\mathfrak{a}$ into the two sets.*

Proof. Let $f : A/\mathfrak{a} \to B$ be a ring homomorphism such that for each $s \in S/\mathfrak{a}$, $f(s)$ is invertible. Applying lifting techniques and universal properties, one can verify that given the canonical maps between the numerous rings associated with A, a function f on $A/\mathfrak{a}$ induces a unique diagram



where the left, bottom, and right triangles commute, as does the overall rectangle. But this implies that the top triangle, and thus the whole diagram, commutes, because we can make the upper triangle commute if we first apply the projection from A into $A/\mathfrak{a}$, and this map is surjective so the triangle itself must commute. Now conversely, any function from $S^{-1}A/S^{-1}\mathfrak{a}$ to B making the upper triangle commute induces a unique set of maps making the whole diagram above commute, so this map must be unique, and therefore $S^{-1}A/S^{-1}\mathfrak{a}$ is an initial object in the category defining the localized ring $(S/\mathfrak{a})^{-1}(A/\mathfrak{a})$, so the two rings must be isomorphic. □

Now let's show the localization of a factorial ring is factorial.

Lemma 6.4. *If A is entire, then $x \in A \cap U(S^{-1}A)$ if and only if $(x) \cap S \neq \varnothing$.*

Proof. If $x(m/n) = 1$, $xm = n \in S$. If $xm \in S$, then $x(m/xm) = 1$. □

Lemma 6.5. *If A is entire, and p is prime in A, then it is irreducible in $S^{-1}A$, provided it is not a unit.*

Proof. If $p = (m/n)(x/y)$, and it is not a unit, then $nyp = mx$, so $p \mid mx$. It follows that $p \mid m$ or $p \mid x$. In either case, we divide by p to conclude either m/n or x/y is a unit. □

Lemma 6.6. *Let A be factorial. Then a/b is irreducible if and only if $a/b = up$, where $u \in U(S^{-1}A)$, and p is irreducible in A and $S^{-1}A$.*

Proof. Let $a = p_1 \dots p_n$, and $b = q_1 \dots q_n$, where p_i and q_i are irreducible in A. Because a/b is irreducible, it follows that exactly one of the p_i is irreducible in $S^{-1}A$, and the other combined factors are units. But this means that p_i is irreducible in A as well. The converse is obvious. □

Lemma 6.7. *If y differs from x by a unit, and y is uniquely factorizable, then x is uniquely factorizable.*

Proof. Write $x = yu$, where y is factorizable, $y = p_1 \dots p_n$, then $x = up_1 \dots p_n$. Now suppose that x can be factorized in two ways

$$x = p_1 \dots p_n = q_1 \dots q_m$$

Then,

$$ux = (up_1)p_2 \dots p_n = p_1' \dots p_n' = (uq_1)q_2 \dots q_m = q_1' \dots q_n'$$

so, up to a permutation, $p_i' = u_i q_{\pi(i)}'$. But one verifies, by taking the vary cases, that this implies that $p_i = v_i q_{\pi(i)}$, where v_i is a unit. □

Theorem 6.8. *If A is factorial, and S is a multiplicative set with $0 \notin S$, then $S^{-1}A$ is factorial.*

Proof. Let a/b be given. We need only verify that a/b differs from a uniquely factorizable element by a unit. a differs from a/b by a unit. Write $a = p_1 \dots p_n$, where p_i is irreducible in A. We know that each p_i is either still irreducible, or a unit, so without loss of generality we may as well assume all p_i are irreducible in $S^{-1}A$. Suppose

$$p_1 \dots p_n = (u_1 q_1) \dots (u_m q_m) = (u_1 \dots u_m q_1) q_2 \dots q_m$$

Let $u_1 \dots u_m = x/y$. If $u_1 \dots u_m$ can be written as the quotient of two units in A, then we are done, for then the p_i and q_i differ by units in A, and thus the p_i differs from $u_i q_i$ by a unit. We show this is the only case that could happen, since we assume the p_i are irreducible in $S^{-1}A$.

If y is not a unit in A, write $y = y_1 \dots y_k$. If x is a unit in A, then when we apply unique factorization in A, we see y_1 differs from some p_i by a unit in A. But y_1 is a unit in $S^{-1}A$, so that p_i is a unit in $S^{-1}A$. If x is not a unit, then we may consider $x = x_1 \dots x_l$, and may assume no x_i and y_j differ by a unit (by cancelling like terms), so that when we apply unique factorization, y_1 is mapped to p_i again, contradicting the irreducibility of p_i. Thus y must be a unit in A, and when we expand x as we have already done, and write

$$(p_1/y) \dots p_n = x_1 \dots x_l q_1 \dots q_m$$

But then some x_i differs from a p_j by a unit in A, hence p_j is a unit in $S^{-1}A$. □

6.2 Local Rings

Originally, localization was used to construct the field of fractions of an integral domain. However, it has been studied in more detail to understand the *local rings*, which occur in areas such as complex analysis and algebraic geometry. A ring A is *local* if it is commutative, and has a unique, maximal ideal. This condition is equivalent to saying that the set $A - U(A)$ of non-invertible elements in A forms an ideal, because if A has a unique maximal ideal $\mathfrak{m}$, then for any $a \in A - U(A)$, (a) is an ideal not equal to A (because if $1 \in (a)$ then a is a unit), so $a \in (a) \subset \mathfrak{m}$. Another equivalent condition is that there exists a maximal ideal $\mathfrak{m}$ such that $1 + \mathfrak{m} \subset U(A)$, because if $x \notin \mathfrak{m}$, then there is y such that $xy \equiv 1$ modulo $\mathfrak{m}$, hence xy is invertible and in particular, x is invertible, so $\mathfrak{m} = U(A)^c$. Conversely, if, in a local ring, $1 + x$ is not invertible, where $x \in \mathfrak{m}$, then $1 + x \in \mathfrak{m}$, so $1 \in \mathfrak{m}$, which is absurd.

Recalling our intuition that maximal ideals in a ring of functions corresponds to a 'point' that the functions operate over, we see that a local ring can be seen as a ring of functions taking values in a unique ring, concentrated at a single point – this is the reason why local rings are called 'local', because they represent the properties of a ring of functions locally around a single point. Indeed, this means that, up to isomorphism, there is a unique field k, and a unique homomorphism from A into k. If a homomorphism $f : A \to k$ corresponds to some 'evaluation map' over elements of A, where k is some field, then we find that A has only a single evaluation map. The main context in which local rings occur is in the study of the localization of certain rings. If $\mathfrak{p}$ is a prime ideal, then $\mathfrak{p}^c$ is certainly a multiplicative subset of A containing 1, so we can form the localization with respect to $\mathfrak{p}^c$, which we denote by $A_\mathfrak{p}$, and call the local ring at $\mathfrak{p}$.

Theorem 6.9. *If $\mathfrak{p}$ is a prime ideal, then $A_\mathfrak{p}$ is a local ring.*

Proof. Since $\mathfrak{p}$ is an ideal, $U(A) \subset \mathfrak{p}^c$, and we can argue that no element of $\mathfrak{p}$ is invertible in $A_\mathfrak{p}$. If $a \in \mathfrak{p}$, and $ab = 1$ in $A_\mathfrak{p}$, then there is $u \in \mathfrak{p}^c$ such that $u(ab-1) = 0 \in \mathfrak{p}$. Since $\mathfrak{p}$ is prime, $ab-1 \in \mathfrak{p}$ and so we conclude $1 \in \mathfrak{p}$, which is impossible. Thus the set of elements of the form a/b with $a \in \mathfrak{p}$ is *precisely* the set $U(A_\mathfrak{p})^c$ of noninvertible elements. If $a \in \mathfrak{p}$, then $(a/b)(c/d) = ac/bd$, and $ac \in \mathfrak{p}$, so $ac/bd \notin \mathfrak{p}$. If $c \in \mathfrak{p}$, then $a/b+c/d = (ad+bc)/bd$, and $ad+bc \in \mathfrak{p}$, so $(ad+bc)/bd$ is not invertible. We conclude that $U(A_\mathfrak{p})^c$ is an ideal of $A_\mathfrak{p}$, so $A_\mathfrak{p}$ is a local ring. □

Example. *If $A(D)$ is the set of analytic functions on some open set D, then the set of functions $f \in A(D)$ such that $f(p) = 0$ forms a prime ideal, so we can form the local ring on this ideal, which is commonly denoted $\mathcal{O}_p(D)$. The invertible elements of $\mathcal{O}_p(D)$ are exactly those functions which are nonzero at p (or, viewing the functions as direct quotients, have a nonzero removable singularity at p). This ring is isomorphic to the subring of the ring $\mathbf{C}[[X - p]]$ of power series in $X - p$, consisting of elements which are convergent in a neighbourhood of p.*

Example. *On $\mathbf{Z}$, we can view elements $a \in \mathbf{Z}$ as functions on the set of prime integers, mapping a prime p to the congruence class of a modulo p in $\mathbf{F}_p$. Thus the integer $1984 = 2^6 \cdot 31$ is a function on the primes which has two zeros at 2 and 31, where 2 to a 'zero of multiplicity six'. This corresponds to the fact that 1984 is invertible in $\mathbf{Z}_{(p)}$ except for $p = 2$ and $p = 31$, where $1984/31$ is invertible in $\mathbf{Z}_{(1984)}$, and $1984/2^6$ is invertible in $\mathbf{Z}_{(2)}$.*

In modern commutative algebra, one takes the set of prime ideals in a space and views them as points, through which the elements of the ring act as functions mapping into integral domains.

Theorem 6.10. *If S is multiplicative, and $\mathfrak{p}$ is a maximal ideal not containing elements of S, then $\mathfrak{p}$ is prime.*

Proof. We claim $S^{-1}\mathfrak{p}$ is a maximal ideal. If $S^{-1}\mathfrak{p} \subsetneq S^{-1}\mathfrak{a}$, then $\mathfrak{p} \subsetneq \mathfrak{a}$, implying $\mathfrak{a}$ contains element of S, so $S^{-1}\mathfrak{a} = S^{-1}A$. Now we claim $i^{-1}(S^{-1}\mathfrak{p}) = \mathfrak{p}$. If $b = a/s$, where $a \in \mathfrak{p}$, $s \in S$, and $b \notin \mathfrak{p}$, then $(b) + \mathfrak{p}$ contains elements in S, hence $xb + y = t$, for $y \in \mathfrak{p}$, $t \in s$. But then $xa + ys = ts$, with the left hand in $\mathfrak{p}$, and the right hand side in S, contradicting the construction of $\mathfrak{p}$. Thus we conclude $\mathfrak{p}$ is prime. □

Proposition 6.11. *If A is local, and $f : A \to B$ a surjective homomorphism, then B is local.*

Proof. If $\mathfrak{m}$ is a maximal ideal in B, then $f^{-1}(\mathfrak{m})$ is an ideal, and the isomorphism theorem guarantees that $A/f^{-1}(\mathfrak{m}) \cong B/\mathfrak{m}$, and since $B/\mathfrak{m}$ is a field, we conclude $f^{-1}(\mathfrak{m}) = U(A)^c$ is the unique maximal ideal in A. If $\mathfrak{n}$ is another maximal ideal in B, then $f^{-1}(\mathfrak{m}) = f^{-1}(\mathfrak{n})$, implying $\mathfrak{m} = \mathfrak{n}$ because f is surjective. □

Local rings were originally designed to analyze rings of functions, such as the ring $\mathcal{O}_p(D)$ of meromorphic functions on an open, connected subset of D, defined at the point p. As discovered in single variable complex analysis, it is in this ring that the concept of orders of poles and zeroes occur. In particular, if f is a meromorphic function holomorphic in a neighbourhood of p, and if $f(p) = 0$, then we can write $f = (X - p)g$ for some meromorphic function g. Since $f \in \mathcal{O}_p(D)$ is non-invertible precisely when $f(p) = 0$, we conclude that the maximal ideal of non-invertible elements is principal, of the form $(X - p)$. More generally, we know that if f is a meromorphic function holomorphic in a neighbourhood of p, then there is a non-negative integer n such that we can write $f = (X - p)^n g$ for some meromorphic function g with $g(p) \neq 0$, and we call n the order of the zero at g. This implies that if $\mathfrak{a}$ is any proper ideal in $\mathcal{O}_p(D)$, then it is of the form $((X-p)^n)$ for some integer n, so $\mathcal{O}_p(D)$ is principal. Thus the smallest ideal in $A_\mathfrak{p}$ containing a function corresponds to it's order at the point p. Here's another example.

Example. *Let A be a factorial ring, and (p) a principal ideal, where p is prime. Then the ring A_p is principal, and also has the properties that $\mathcal{O}_p(D)$ has. Every principal ideal in A_p is of the form (p^N), because if $a = p^n q$, where $p \nmid q$, then $q \in U(A_p)$ and so $(a) = (p^n)$. But now if $\mathfrak{a}$ is any ideal, and we define the order of a to be the integer $ord(a)$ such that $(a) = (p^n)$, then*

$$\mathfrak{a} = \bigoplus_{a \in \mathfrak{a}} (a) = \bigoplus_{a \in \mathfrak{a}} (p^{ord(a)}) = (p^{\min ord(a)})$$

so every ideal is principal, and in particular, generated by a power of p. Thus the order of an element of the ring measures it's place in the linear heirarchy

$$(1) \supset (p) \supset (p^2) \supset \cdots \supset (0)$$

which consists of all ideals.

We want to consider rings where we can discuss the phenomenon of 'multiplicities of zeroes'. Since we are focusing on a ring, such a ring shuold be localized at the point where we want to measure zeroes, so our ring should be local. If the ring is Noetherian domain, but not a field, which maximal ideal is principal, we call the ring a *discrete valuation ring*. These are the rings having the properties we wish.

Proposition 6.12. *If A is a discrete valuation ring, then there exists an element $t \in A$ such that every nonzero element of A can be uniquely written as ut^n, where u is a unit in A.*

Proof. Let (t) be the maximal ideal of A. Suppose that $ut^n = vt^m$. If $n = m$, then $u = v$. Otherwise, if $n > m$, then $u = vt^{m-n}$, and this implies that (t) contains a unit, hence is not a maximal ideal. Thus it suffices to prove that every element of A has a required expansion of the form above. If $a \in A$ is a unit, we can write $a = at^0$, and we are done. If a is not a unit, then (a) is an ideal contained in (t), so we can write $a = a_1 t$ for some $a_1 \in A$. Then (a) is a proper subideal of (a_1), because if $a_1 = ba$, then $a = bat$, hence $1 = bt$, so t is invertible. If a_1 is a unit, we are done, otherwise we can write $a_1 = a_2 t$. Continuing this process, if this process does not terminate, we end up with an infinite ascending chain

$$(a) \subsetneq (a_1) \subsetneq (a_2) \subsetneq \dots$$

and this is impossible in a Noetherian ring. □

If A is a domain, the condition that we have a unique expansion of the form ut^n for each element of A is exactly the condition which guarantees that the ring is a discrete valuation ring. If this is true, then (t) is certainly a unique maximal ideal in A, so A is a local ring whose maximal ideal is principal. To prove that A is Noetherian, it suffices to notice that the proper ideals of A are exactly $(0), (t)$, (t^2), (t^3), and so on and so forth, so that the ring is actually principal. The element t in the theorem is known as a *uniformizing parameter* for A. Any other uniformizing parameter for A differs from t by a unit, so if $s = ut$ is another uniformizing parameter, then if $a = vt^n = rs^m$, then $rs^m = ru^m t^m$, so $v = ru^m$ and $n = m$. Since this value is invariant of the uniformizing parameter, it depends only on the element a, and we call this the *order of a*. We define the order of 0 to be ∞. If we consider the field B of fractions of A, then every nonzero element b of B can be written as ut^n for a unique integer $n \in \mathbf{Z}$, and we define this to be the order of b. If $n < 0$, we say that b has a pole of order $-n$.

Example. *Consider the ring $k[X] = k[\mathbf{A}^1]$. Then for any $a \in \mathbf{A}^1$, the ring $\mathcal{O}_a(\mathbf{A}^1)$ of rational functions defined at a (those polynomials f/g with $g(a) \neq 0$) is a discrete valuation ring. If we consider any function f/g with $g(a) \neq 0$, then $f = (X-a)^n h(X)$ for some $n \geqslant 0$ and since h with $h(a) \neq 0$. This gives us a*

decomposition $f/g = (h/g)(X-a)^n$, so $X-a$ is a uniformizing parameter, and $\mathcal{O}_a(\mathbf{A}^1)$ is a discrete valuation domain.

Example. *Consider the ring $\mathcal{O}_\infty(\mathbf{A}^1)$ of rational functions of the form $f/g \in k(X)$, with $\deg g \geqslant \deg f$. This rings models the set of rational functions which converges to a well defined quantity 'near infinity'. The only invertible functions in this ring are those with $\deg g = \deg f$, and so the noninvertible functions are generated by $(1/X)$, because if $\deg g - \deg f = n$, then $X^n(f/g) = (X^n f/g)$ is invertible, and contained in $\mathcal{O}_\infty(\mathbf{A}^1)$.*

Example. *If p is a prime number, then the local ring $\mathbf{Z}_{(p)}$ is a discrete valuation ring, because if $a/b \in \mathbf{Z}_{(p)}$, with $b \notin (p)$, we can write $a = p^n c$ with c and p relatively prime, and then $a/b = p^n(c/b)$ has c/b invertible. This gives an order function on $\mathbf{Q}$ defined by taking the order of a number $m = p^n(a/b)$ with respect to p to be n. This can be used to define a metric on $\mathbf{Q}$, and the completion is the field of p-adic numbers.*

The order function on the resulting field of fractions of a discrete valuation domain satisfies useful algebraic properties.

- $\operatorname{ord}(x) = 0$ if and only if $x = 0$.
- $\operatorname{ord}(xy) = \operatorname{ord}(x) + \operatorname{ord}(y)$.
- $\operatorname{ord}(x+y) \geqslant \min(\operatorname{ord}(x), \operatorname{ord}(y))$.

We will show that these properties are essentially the defining properties of a discrete valuation domain. Given any field k, an order function is a $\mathbf{Z} \cup \{\infty\}$ valued function φ on k with the properties above, and with $\varphi(x) = \infty$ if and only if $x = 0$.

Proposition 6.13. *For any order function φ on a field k, the ring A of elements $x \in k$ with $\varphi(x) \geqslant 0$ forms a discrete valuation domain, with k it's field of fractions.*

Proof. A is certainly closed under multiplication and addition. Since $\varphi(x) = \varphi(1 \cdot x) = \varphi(1) + \varphi(x)$, we conclude that $\varphi(1) = 0$. We use this to conclude that $\varphi(xx^{-1}) = \varphi(x) + \varphi(x)^{-1} = 0$, so an element $x \in A$ is invertible if and only if $\varphi(x) = 0$. This shows that the set of noninvertible elements forms an ideal, hence the ring A is local. The ring is certainly a domain. We may assume that there is $x \in k$ with $\varphi(x) = 1$, because otherwise every

noninfinite value of the order function is a multiple of some integer, and we obtain another order function by dividing by this integer. If $\varphi(x) = 0$, then for every $x \in A$, there is n such that $\varphi(xt^{-n}) = 0$, hence $xt^{-n} = u$ is a unit, and $x = ut^n$. We have justified that this proves A is a discrete valuation domain, and since $\varphi(x^{-1}) = -\varphi(x)$, every element of k is either an element of A, or of the form $1/x$ for some $x \in A$, showing that k is the field of fractions of A. □

Proposition 6.14. *If $ord(a) < ord(b)$, then $ord(a + b) = ord(a)$.*

Proof. $a = t^n u$, $b = t^m s$, then $a+b = t^n(u+t^{m-n}s)$, and $u+t^{m-n}s$ is invertible because it is congruent to u in the maximal ideal. This is analogous to the addition law for polynomials in $k[X]$. □

Often, a discrete valuation ring models the germ of functions around a point, and the evaluation map at this points gives us the maximal ideal, as well as an isomorphism between the ring of constant functions and the field upon which the functions are defined. In this situation, we can obtain some useful properties of the ring of constant functions, related to the Taylor expansion of functions around a point.

Proposition 6.15. *Suppose that a discrete valuation ring A contains a subfield k, such that if $\mathfrak{m}$ is the maximal ideal of A, then $k \to A \to A/\mathfrak{m}$ gives an isomorphism of fields. If t is a uniformizing parameter for A, then for any $n \geqslant 0$, every $x \in A$ has a unique expansion as $x = \lambda_0 + \lambda_1 t + \cdots + \lambda_n t^n + z_n t^{n+1}$, where $z_n \in A$.*

Proof. For any $x \in A$, there is $\lambda \in k$ such that x is congruent to λ modulo $\mathfrak{m} = (t)$, so $x = \lambda + z_0 t$. This gives the proposition for the case $n = 0$. For the inductive case, we write $x = \sum \lambda_i t^i + z_n t^{n+1}$. Then using the $n = 0$ case we can write $z_n = \lambda_{n+1} + z_{n+1} t$, and this gives the expansion for x one degree higher. To prove uniqueness, we note that if $\sum \lambda_i t^i + z_n t^{n+1} = 0$, then $\sum \lambda_i t^i = -z_n t^{n+1}$, and if $z_n \neq 0$, the right side has order greater than or equal to $n + 1$, whereas the right side has order equal to the minimum index i such that $\lambda_i \neq 0$, and these two values cannot be equal. □

The ring of formal power series over a field k is written $k[[X]]$, and is the ring of 'infinite power series' $\sum_{k=0}^{\infty} a_k X^k$, with $a_k \in k$. Then $k[[X]]$ is a ring containing $k[X]$ as a subring, and is a discrete valuation ring. To prove this, suppose $\left(\sum a_i X^i\right)$ is invertible, so there is a power series such

that $\left(\sum a_i X^i\right)\left(\sum b_i X^i\right) = 1$. This is equivalent to being able to solve the infinite series of equations

$$a_0 b_0 = 1 \quad a_1 b_0 + a_0 b_1 = 0 \quad a_2 b_0 + a_1 b_1 + a_0 b_2 = 0$$

The first equation guarantees that we must have $a_0 \neq 0$, but if this is true the first equation is uniquely solvable for b_0, and this value is nonzero. Once b_1 is fixed, the equation $a_0 b_1 = -a_1 b_0$ is uniquely solvable for b_1. Continuing this, we find that given that the previous equations are solvable, there is a unique value of b_n which satisfies the n'th equation, and so an element of $k[[X]]$ is invertible precisely when its constant coefficient is nonzero. This shows that the non-invertible elements of $k[[X]]$ are precisely (X), so the ring is local. We can write an arbitrary power series $\sum a_i X^i$ as $X^n \sum b_i X^i$, where $b_0 \neq 0$, so the ring is a discrete valuation domain, where the order function is precisely the degree corresponding to the smallest non-zero coefficient. The quotient field of $k[[X]]$ is denoted $k((X))$.

Assuming that we have an isomorphism $k \to A \to A/\mathfrak{m}$, the previous proposition shows that we have a natural injective homomorphism from A to $k[[X]]$. This shows that the class of discrete valuation domains which contain a field corresponding to the quotient by their maximal ideal are precisely the rings where we can consider 'power series' of elements. Furthermore, we obtain a map of k into $k((X))$, because the homomorphism is injective, and the order function on $k[[X]]$ agrees with the one induced from k. This essentially corresponds to the fact that all holomorphic functions can be expanded as power series, and here we also have additional analytic relationships between these expansions and their convergence around a point.

Example. *In complex analysis, one memorizes the power series expansion*

$$(1 - X)^{-1} = (1 + X + X^2 + \dots)$$

This equation holds in the ring $k[[X]]$ of power series over any field, because of

the telescoping series properties of $(1-X)(1+X+X^2+\dots)$. *Similarily,*

$$\begin{aligned}(1-X)(1+X^2)^{-1} &= (1-X)(1+iX)^{-1}(1-iX)^{-1}\\ &= (1-X)\left(\sum(-i)^kX^k\right)\left(\sum i^kX^k\right)\\ &= (1-X)\left(\sum(-1)^kX^{2k}\right)\\ &= (1-X-X^2+X^3+X^4-X^5-X^6+\dots)\end{aligned}$$

Proposition 6.16. *Suppose that A is a discrete valuation ring, with quotient field k. Then there are no local rings B with $A \subsetneq B \subset k$, such that the maximal ideal of B contains the maximal ideal of A.*

Proof. If a nonzero x is in k, but not in A, then x has some order $-n < 0$, so x^{-1} has order n, and is consequently in A. This means that $x^{-1} \in A$ for each $x \in A$. Iif the maximal ideal $\mathfrak{m}$ of B contains the maximal ideal $\mathfrak{n}$ of A, we claim that $\mathfrak{m} = \mathfrak{n}$. Otherwise, we can pick $x \in \mathfrak{m} - \mathfrak{n}$, and then $x^{-1} \in A$, so $1 = xx^{-1} \in \mathfrak{m}$, contradicting the fact that $B \neq k$. Now let t be a uniformizing parameter for A. Every element of k, and in particular B, can be written as xt^n, where x is a unit in A. In particular, if $B-A$ is nonempty, it contains some element ut^{-n}, where $n > 0$, and u is a unit in A. But then B contains t^{-n}, and hence all elements of the form $t^{k-mn} = t^k(t^{-n})^m$, so $B = k$, which is impossible. □

Example. *Using this theorem, we can classify the discrete valuation rings with quotient field $k(X)$ which contain k, where k is algebraically closed. Let A be a discrete valuation ring, and suppose the uniformizing parameter is some irreducible $t \in A$. If A contains X, then A contains $k[X]$, and the set of elements of $k[X]$ which are not invertible in A forms a prime ideal, which is therefore of the form (f) for some irreducible monic polynomial f. Since k is algebraically closed, $f(X) = X - a$, for some $a \in k$, and so A contains $\mathcal{O}_a(\mathbf{A}^1)$, implying the two are equal to one another. If A does not contain X, then A contains X^{-1}. Since the order of any nonzero $a \in k$ is zero, and the order of X^{-1} is greater than zero because it is not invertible, $a_0+a_1X^{-1}+\dots+a_nX^{-n} = (a_0X^n+\dots+a_n)/X^n$ is invertible in A, hence $X^n/(a_0X^n+\dots+a_n) \in A$. Multiplying by $b_0+b_1X^{-1}+\dots+b_nX^{-n}$, we conclude that $(b_0X^n+\dots+b_n)/(a_0X^n+\dots+a_n) \in A$ for any $a_0 \neq 0$. This shows that A contains $\mathcal{O}_\infty(\mathbf{A}^1)$, and if $f(X)/g(X)$ has $\deg g > \deg f$, then g/f is not in A, for otherwise we may write $g = (X-a_1)\dots(X-a_m)$, $f = (X-b_1)\dots(X-b_l)$, and then $h = (X-a_1)\dots(X-a_{m-1})/(X-b_1)\dots(X-b_l) \in A$,*

so $hg/f = X - a_m \in A$, implying $X \in A$, contradicting our assumption. Thus the maximal ideal of A contains the maximal ideal of $\mathcal{O}_\infty(\mathbf{A}^1)$, and this implies that A is in fact equal to $\mathcal{O}_\infty(\mathbf{A}^1)$.

Example. *The only discrete valuation rings with quotient field $\mathbf{Q}$ are the local rings $\mathbf{Z}_{(p)}$. If A is any such discrete valuation ring, then A contains all the integers $\mathbf{Z}$. Because A is a local ring, the set of non-invertible integers in A forms a prime ideal in $\mathbf{Z}$, and hence is of the form (p) for some prime integer. But then A contains $\mathbf{Z}_{(p)}$, which implies $A = \mathbf{Z}_{(p)}$.*

Similar techniques to the classifications above allow us to classify the set of all discrete valuation rings which are obtained from extensions of principal ideal domains. These valuation rings are exactly of the form A_p, where (p) is a prime ideal in the PID.

6.3 Tensor Products

Given two A modules M and N, the *tensor product* $M \otimes_A N$, or $M \otimes N$ if the module A is implicit (or over $\mathbf{Z}$, if M and N are just abelian groups), is the most general way we can form a 'bilinear space' corresponding to M and N. More specifically, $M \otimes N$ is the initial object in the category of bilinear maps $f : M \times N \to L$ into a module L. We can construct $M \otimes N$ by consider the quotient of the free module with basis elements $M \times N$, subject to the submodule generated by $(x+y,z) - (x,z) - (y,z)$, $a(x,y) - (ax,y)$, and $a(x,y) - (x,ay)$. We let the image of (x,y) in the quotient be denoted by $x \otimes y$, so that $(x+y) \otimes z = x \otimes z + y \otimes z$, $a(x \otimes y) = (ax) \otimes y = x \otimes (ay)$. The embedding of $M \times N$ in the free abelian group obviously descends to a bilinear map from $M \times N$ to $M \otimes N$, which is bilinear precisely because of the quotients defining $M \otimes N$. If $f : M \times N \to L$ is bilinear, then f extends uniquely to a map on the free group generated by $M \times N$. Furthermore, the relations which make f bilinear precisely mean that f descends to a map from $M \otimes N$ to L, so we get a unique morphism from $M \otimes N$ to L which represents f. However, this definition is the 'wrong' definition to use in most cases when understanding the tensor product, because it's quite a strange definition to work with.

Example. *Given the abelian groups $\mathbf{Z}_{10}$ and $\mathbf{Z}_{12}$, we find that $\mathbf{Z}_{10} \otimes \mathbf{Z}_{12}$ is isomorphic to $\mathbf{Z}_2$. We find that for any integers n, m,*

$$n \otimes m = (11n) \otimes m = n \otimes (11m) = -n \otimes m$$

Thus 2 annihilates all of $\mathbf{Z}_{10} \otimes \mathbf{Z}_{12}$*, and so we get the natural structure of* $\mathbf{Z}_{10} \otimes \mathbf{Z}_{12}$ *as a vector space over* $\mathbf{Z}_2$*. Yet*

$$n \otimes m = n(1 \otimes m) = (nm)(1 \otimes 1)$$

so the vector space is generated by a single element $1 \otimes 1$*. The element* $1 \otimes 1$ *doesn't equal to zero in* $\mathbf{Z}_{10} \otimes \mathbf{Z}_{12}$*. We have a bilinear map* $f : \mathbf{Z}_{10} \times \mathbf{Z}_{12} \to \mathbf{Z}_2$ *given by* $f(x,y) = xy$*, which is well defined because* $(x+10)y = x(y+12) = xy$ *modulo 2. Thus we have an induced map* $f_* : \mathbf{Z}_{10} \otimes \mathbf{Z}_{12} \to \mathbf{Z}_2$ *where* $f_*(1 \otimes 1) = f(1,1) = 1$*, which is different from* $f_*(0 \otimes 0) = 0$*, so* $1 \otimes 1 \neq 0$*. In particular, our calculation shows that for any bilinear map* $f : \mathbf{Z}_{10} \times \mathbf{Z}_{12} \to M$*, there exists a unique morphism* $g : \mathbf{Z}_2 \to M$ *such that* $f(x,y) = g(xy)$*.*

The tensor product is a covariant bifunctor on the category of modules, since if $f : M_0 \times M_1$ and $g : N_0 \times N_1$, then we have a unique morphism $(f \otimes g) : (M_0 \otimes N_0) \to (M_1 \times N_1)$ obtained by $(f \otimes g)(x \otimes y) = f(x) \otimes g(y)$. If we fix a module N on the right, then we obtain a covariant functor which is known as *right exact*. More specifically, given an exact sequence

$$M_0 \xrightarrow{f} M_1 \xrightarrow{g} M_2 \longrightarrow 0$$

The induced exact sequence

$$M_0 \otimes N \xrightarrow{f \otimes \mathrm{id}} M_1 \otimes N \xrightarrow{g \otimes \mathrm{id}} M_2 \otimes N \longrightarrow 0$$

is also exact. The surjectivity is easy to prove. Since g is surjective, given for any $y \in M_2$, there is $x \in M_1$ with $g(x) = y$, so $(g \otimes \mathrm{id})(x \otimes z) = y \otimes z$ for any $z \in N$. Now we prove that the image of $(f \otimes \mathrm{id})$ is the kernel of $(g \otimes \mathrm{id})$. Certainly the image, which we denote by L, is a subset of the kernel, which we denote by K. So we get an induced surjective map from $(M_1 \otimes N)/L \to (M_2 \otimes N)$. We claim it is an isomorphism, which would show that $L = K$. To define a left inverse, given $y \otimes z \in M_2 \otimes N$, choose $x \in M_1$ such that $g(x) = y$. The map $h(y \otimes z) = x \otimes z + L$ is a well defined map into the quotient, because if $x, x' \in M_1$ are such that $f(x) = f(x') = y$, then $x - x'$ is in the kernel of g, so $(x - x') \otimes z$ is in L. The map is clearly bilinear, and thus extends to a complete map h on $M_2 \otimes N$, and it is easy to check this is a left inverse on a generating set, hence everywhere.

Tensoring commutes with the direct sum operation, a fact easy to prove by the universal property. That is, we have $M \otimes \bigoplus N_\alpha$ isomorphic to

$\bigoplus(M \otimes N_\alpha)$. Any bilinear map f from $M \times \bigoplus N_\alpha$ to L corresponds to a unique family of bilinear maps f_α from $M \times N_\alpha$ to L, inducing a map from $M \otimes N_\alpha$ to L, which can be put together to form a unique map from $\bigoplus(M \otimes N_\alpha)$ to L. Thus multiplication and addition of modules is 'distributive'. Considering the tensor product of modules as a multiplication operation, and addition as a direct sum, the family of modules over an abelian group is given a sort of ring structure, which becomes very important in the field of K theory.

6.4 Dedekind Rings

In the understanding of integral solutions to polynomial equations such as $X^n - Y^n$ can be factored over $\mathbf{Z}[\zeta_n]$, where ζ_n is a primitive n th root of unity. In 1847 Gabriel Lumé used the fact that $\mathbf{Z}[\zeta_n]$ is a unique factorization domain to provide a proof of Fermat's last theorem, with one catch; $\mathbf{Z}[\zeta_n]$ is not always a unique factorization domain, and so his proof only works for certain values of n for which the ring is such a domain; in 1844 Ernst Kummer showed that $\mathbf{Z}[\zeta_{23}]$ is *not* a unique factorization domain. However, Ernst Kummer also showed that there are certain techniques which allow us to extend UFD type arguments to more general rings, including the rings $\mathbf{Z}[\zeta_{23}]$; rather than factorizing individual elements of a ring, we can factor ideals in the ring into prime ideal components.

Example. *Consider the ring* $\mathbf{Z}[\sqrt{-5}]$, *in which*

$$2 \cdot 3 = 6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$$

all of 2, 3, $1 + \sqrt{-5}$, *and* $1 - \sqrt{-5}$ *are irreducible in* $\mathbf{Z}[\sqrt{-5}]$, *because we know* $|a + b\sqrt{-5}|^2 = a^2 + 5b^2$, *and there are no solutions in* $\mathbf{Z}^2 + 5\mathbf{Z}^2$ *to the equations* $XY =$ 4, 9, *or* 6, *except for the trivial ones corresponding to a unit multiplied by a constant. Thus* $\mathbf{Z}[\sqrt{-5}]$ *is not a unique factorization domain. However, consider the corresponding relationship between the ideals, i.e.*

$$(2)(3) = (6) = (1 + \sqrt{-5})(1 - \sqrt{-5})$$

Even though these numbers are irreducible element of the ring, they are not prime elements, since, for instance, 2 *divides* $6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$, *but can't divide* $1 + \sqrt{-5}$ *or* $1 - \sqrt{-5}$. *On the other hand,* $(2, 1 + \sqrt{-5})$ *is a prime*

ideal, because $\mathbf{Z}[\sqrt{-5}]/(2,1+\sqrt{-5})$ *is isomorphic to* $\mathbf{Z}_2$*, which is obtained from the fact that the embedding of* $\mathbf{Z}$ *into* $\mathbf{Z}[\sqrt{-5}]/(2,1+\sqrt{-5})$ *is surjective, with kernel* (2)*, as is* $(3,1-\sqrt{-5})$*, and we have*

$$(6) = (2,1+\sqrt{-5})(3,1-\sqrt{-5})$$

which is a unique factorization of ideals.

A *Dedekind ring* is precisely a domain where one can factor ideals uniquely into products of prime ideals. An equivalent definition, more interesting, occurs in the theory of ideal class groups in algebraic number theory. If A if a domain with a field of fractions k, we say an A submodule $\mathfrak{a}$ of k is a *fractional ideal* if there is $x \in A$ with $x\mathfrak{a} \subset A$, so that $\mathfrak{a}$ has 'bounded denominator'. The family of fractional ideals forms a monoid, with A as the identity element, if we take products just as in the case of normal ideals, $\mathfrak{a}\mathfrak{b}$ is the subgroup of $K - \{0\}$ generated by elements of the form ab, for $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$. If the family of fractional ideals forms a group under this product, we will find that the ideals have a unique factorization theory.

To see this, let's explore some consequences of the group property. If $\mathfrak{a}$ is an ideal of A, this means there is a fractional ideal $\mathfrak{b}$ with $\mathfrak{a}\mathfrak{b} = A$, so that there are $x_1,\ldots,x_n \in \mathfrak{a}$, $y_1,\ldots,y_n \in \mathfrak{b}$ with $x_1y_1 + \cdots + x_ny_n = 1$. If $x \in \mathfrak{a}$ is arbitrary, then $x = x_1(y_1x) + \cdots + x_n(y_nx)$, and we know because of the product formula $\mathfrak{a}\mathfrak{b} = A$ that $y_kx \in A$, hence we have found $\mathfrak{a} = (x_1,\ldots,x_n)$. We conclude that any ring whose fractional ideals form a group is Noetherian.

6.5 Abelian Categories

If M and N are modules over the same ring, then $\mathrm{Hom}(M,N)$ is an abelian group. If $f,g \in \mathrm{Hom}(M,N)$, then define

$$(f+g)(x) = f(x) + g(x)$$

The zero homomorphism $0(x) = 0$ is the identity in this group. Given $\lambda \in \mathbf{R}$, we may define

$$(\lambda f)(x) = \lambda f(x)$$

but this is only in $\mathrm{Hom}(M,N)$ if R is commutative, so $\mathrm{Hom}(M,N)$ is an R module only if R is commutative. Given $f : M \to N$, and a fixed module X,

we obtain a morphism $f^* : \mathrm{Hom}(N, X) \to \mathrm{Hom}(M, X)$, mapping g to $g \circ f$. Similarily, we get a morphism $f_* : \mathrm{Hom}(X, M) \to \mathrm{Hom}(X, N)$, by letting $g \mapsto f \circ g$. This follows because composition is bilinear,

$$(f + g) \circ h = f \circ h + g \circ h \qquad f \circ (g + h) = f \circ g + f \circ h$$

It follows that Hom is a functor in two variables, contravariant in the first, and covariant in the second. We shall also make use of the relations

$$(g \circ f)_* = g_* \circ f_* \qquad (g \circ f)^* = f^* \circ g^*$$

Arrow theoretic arguments are very common in module theory. We consider exact sequences just as in group theory.

$$A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} \dots \xrightarrow{f_{n-1}} A_n$$

If $\ker(f_{i+1}) = \mathrm{im}(f_i)$ for each i.

Theorem 6.17. *If*

$$A \xrightarrow{f} B \xrightarrow{g} C \to 0$$

is exact, then

$$Hom(A, X) \xleftarrow{f^*} Hom(B, X) \xleftarrow{g^*} Hom(C, X) \leftarrow 0$$

is also exact.

Proof. Since $g \circ f = 0$, $(g \circ f)^* = 0$. Thus $\ker(f^*) \supset \mathrm{im}(g^*)$. Suppose that $f^*(T) = 0$. We claim that $T = g^*(S)$ for some $S \in \mathrm{Hom}(C, X)$. If $x = g(y)$, then define

$$Sx = Ty$$

This is well-defined, since if $g(y) = g(z)$, $g(y - z) = 0$, so there is some $a \in A$ such that $y - z = f(a)$. It then follows that

$$T(y - z) = (T \circ f)(a) = 0(a) = 0$$

Thus $Ty = Tz$. Since g is surjective, S is defined on all of C, is easily checked to be a module homomorphism, and satisfies $T = g^*(S)$.

We must also show g^* is injective. Suppose $T \circ g = 0$. If $x \in C$ is given, then there is $y \in b$ such that $g(y) = 0$. Then

$$0 = (T \circ g)(y) = T(x) = 0$$

so $T = 0$. □

Theorem 6.18. *If*

$$0 \to A \xrightarrow{f} B \xrightarrow{g} C$$

is exact, then

$$0 \to Hom(X,A) \xrightarrow{f_*} Hom(X,B) \xrightarrow{g_*} Hom(X,C)$$

is also exact.

Proof. We have the relation

$$g_* \circ f_* = (g \circ f)_* = 0_* = 0$$

Hence $\ker(g_*) \subset \mathrm{im}(f_*)$. Suppose $g \circ T = 0$. We claim $T = f \circ S$ for some $S \in \mathrm{Hom}(X,A)$. For each $x \in X$, define $Sx = y$, where $f(y) = Tx$. y must be necessarily unique, for f is injective, and exists because $g(Tx) = 0$, and the exactness of f and g. The map is easily checked to be a homomorphism, and satisfies $f_*(S) = T$.

Now we prove f_* is injective. Suppose $f \circ T = 0$. Then $f(T(x)) = 0$ for each x, implying $T(x) = 0$ since f is injective. Thus $T = 0$. □

A Category $\mathcal{C}$ is *Additive* if for any two objects X and Y, $\mathrm{Mor}(X,Y)$ is an abelian group, such that composition is bilinear, there exists an object 0 which is both initial and terminal, and finite products and coproducts exist. An additive category is *Abelian* if kernels and cokernels exist, and if 0 is the kernel of $f : X \to Y$, then f is the kernel of its cokernel, and if 0 is the cokernel of f, then f is the cokernel of its kernel, and if 0 is the kernel and cokernel of f, then f is an isomorphism. Most module arguments can be made into abelian categorical arguments, which is useful when other abelian categories appear, such as the category of chain complexes in homology theory.

Chapter 7

Algebras

TODO: Change the previous chapters so that all rings are unital, with unital homomorphisms. Then define algebras to be nonunital, so that the theory of nonunital rings is included in the theorem of $\mathbf{Z}$ algebras. Then one gets all the nice homomorphism theorems here. But then we must have a separate theory of modules over algebras?

7.1 Matrix Rings

Let R be a ring. Then the set of all endomorphisms from R^n to itself is the prime example of an R-module, and the set of endomorphisms from R^n to itself is an R-algebra. Every endomorphism $T : R^n \to R^n$ can be identified as an $n \times n$ matrix M with coefficients in R, such that $Mx = T(x)$. We denote the set of all $n \times n$ matrices as $M_n(R)$. The tractable case is really only when R is a commutative ring, those noncommutative examples do occur in certain problems. For now, we shall assume R is commutative.

The units of $M_n(R)$ are the invertible matrices, and the set of all matrices forms the general linear group $GL_n(R)$. The determinant operator

$\det : M_n(R) \to R$ still applies, and satisfies $\det(AB) = \det(A)\det(B)$, since

$$\begin{aligned}
\det(AB) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (A_{i1}B_{1\sigma(i)} + A_{i2}B_{2\sigma(i)} + \cdots + A_{in}B_{n\sigma(i)}) \\
&= \sum_{\tau \in S_n} \operatorname{sgn}(\tau) \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\tau^{-1}\sigma) \sum_{i=1}^{n} B_{\tau(i)\sigma(i)} \right) A_{1\tau(1)} \dots A_{n\tau(n)} \\
&= \sum_{\tau \in S_n} \operatorname{sgn}(\tau) \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \sum_{i=1}^{n} B_{i\sigma(i)} \right) A_{1\tau(1)} \dots A_{n\tau(n)} \\
&= \det(A)\det(B)
\end{aligned}$$

If $M \in GL_n(R)$, then $\det(M) \in U(R)$, because

$$\det(M)\det(M^{-1}) = \det(MM^{-1}) = \det(I) = 1$$

For instance, $M \in GL_n(\mathbf{Z})$ can only be invertible if $\det(M) = \pm 1$. In this case, we know by Cramer's rule that the inverse of M in $GL_n(\mathbf{R})$ is given by

$$\frac{1}{\det(M)} A$$

where the coefficient A_{ij} is the determinant of the submatrix of M obtained by removing row j and column i, multiplied by $(-1)^{i+j}$. This matrix lies in $GL_n(\mathbf{Z})$ if $\det(M) = \pm 1$, so $GL_n(\mathbf{Z})$ consists exactly of the matrices whose determinant is ± 1. We essentailly can apply Cramer's rule to all rings.

Theorem 7.1. *M is invertible in $M_n(R)$ if and only if $\det(M)$ is a unit in R.*

Proof. Consider the adjoint matrix A described above. Let M^{jk} be the matrix obtained by deleting row j and column k.

$$(MA)_{ij} = \sum_{k=1}^{n} M_{ik}A_{kj} = \sum_{k=1}^{n} (-1)^{j+k} M_{ik} \det(M^{jk})$$

If $i = j$, then this is just the Laplace expansion of the determinant, so $(MA)_{ii} = \det(A)$. If $i \neq j$, this is the Laplace expansion of the matrix obtained by replacing row j with row i, causing a repeated row, and so the Laplace expansion will be zero. Thus $MA = \det(A)$, and M is invertible provided $\det(A)$ is invertible, i.e. it is a unit. □

The group $GL_n(R)$, together with its action on R^n, make it somewhat tractable to study. In the field of representation theory, we try and understand all groups by their homomorphisms into $GL_n(R)$. The determinant allows us to understand some properties of the group. For instance, since the determinant is a group homomorphism from $GL_n(R)$ to $U(R)$, we have a normal subgroup $SL_n(R)$ consisting of matrices with determinant one, and since the map from $GL_n(R)$ to $U(R)$ is surjective, the index of $SL_n(R)$ in $GL_n(R)$ is the same as the number of invertible elements in R.

Theorem 7.2. *$M_n(M_m(R))$ is isomorphic $M_{nm}(R)$.*

Proof. The algebra $M_n(M_m(R))$ is isomorphic to the set of endomorphisms on $M_m^n(R)$. But the module $M_m^n(R)$ is isomorphic to $M^{nm}(R)$, so the set of endomorphisms on $M_m^n(R)$ is isomorphic to the set of endomorphisms on $M^{nm}(R)$. □

We note that the isomorphism from $M_{nm}(R)$ to $M_n(M_m(R))$ coagulates blocks of submatrices in a way which preserves the algebraic structure. For instance, $M_4(R)$ is isomorphic to $M_2(M_2(R))$, such that

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{N} & \mathbf{M} \\ \mathbf{O} & \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{AN} + \mathbf{BO} & \mathbf{AM} + \mathbf{BD} \\ \mathbf{CN} + \mathbf{DO} & \mathbf{CM} + \mathbf{DP} \end{pmatrix}$$

where the left side is multiplication in $M_4(R)$, and the algebra on the right side done over matrices in $M_2(R)$.

Chapter 8

Linear Algebra

Theorem 8.1. *Let $T : V \to V$ be an injective linear map. If W if a T stable subspace of V, and V/W and $W/T(W)$ is finite dimensional, then $V/T(V)$ is finite dimensional, and the dimension is equal to the dimension of $W/T(W)$.*

Proof. The map T induces a surjective map from V to $T(V)/T(W)$ whose kernel is W, so V/W is isomorphic to $T(V)/T(W)$ by the first isomorphism theorem. Since $W \subset W + T(V) \subset V$ and $T(W) \subset W \cap T(V) \subset T(V)$, we conclude that

$$\dim \frac{V}{W} = \dim \frac{V}{W + T(V)} + \dim \frac{W + T(V)}{W}$$

$$\dim \frac{T(V)}{T(W)} = \dim \frac{T(V)}{W \cap T(V)} + \dim \frac{W \cap T(V)}{T(W)}$$

The second isomorphism theorem tells us that $T(V)/[W \cap T(V)]$ is isomorphic to $[W + T(V)]/W$. Putting this together with the fact that V/W is isomorphic to $T(V)/T(W)$, we conclude that $\dim V/[W + T(V)] = \dim[W \cap T(V)]/T(W)$. But now, since $T(V) \subset W + T(V) \subset V$ and $T(W) \subset W \cap T(V) \subset W$, we conclude that

$$\dim \frac{V}{T(V)} = \dim \frac{W + T(V)}{T(V)} + \dim \frac{V}{W + T(V)}$$

$$\dim \frac{W}{T(W)} = \dim \frac{W \cap T(V)}{T(W)} + \dim \frac{W}{W \cap T(V)}$$

But $V/[W + T(V)]$ has the same dimension as $[W \cap T(V)]/T(W)$, and the second isomorphism theorem implies that $[W + T(V)]/T(V)$ is isomorphic

to $W/[W \cap T(V)]$, and we conclude that $V/T(V)$ has the same dimension as $W/T(W)$. $\square$

Chapter 9

Noetherian Rings

A ring A is *Noetherian* if every ideal of A is finitely generated, or equivalently, if every ascending chain of ideals terminates. The Noetherian property was used first by Hilbert, but it's importance as an axiom in algebraic geometry was discovered by Emmy Noether. In geometric terms, the next theorem says that every variety is the unique union of finitely many maximal irreducible subvarieties.

Theorem 9.1. *Let $\mathfrak{a}$ be an ideal in a Noetherian ring. Then among all prime ideals containing $\mathfrak{a}$, there are finitely many which are minimal.*

Proof. Since A is Noetherian, if this proposition fails, then their must be a maximal ideal $\mathfrak{a}$ in which it fails. We know that $\mathfrak{a}$ cannot be prime, so there is $a, b \notin \mathfrak{a}$ such that $ab \in \mathfrak{a}$. The ideals $\mathfrak{a}+Aa$ and $\mathfrak{a}+Ab$ are strictly larger than A, so there are only finitely many prime ideals containing $\mathfrak{a}+Aa$ and $\mathfrak{a}+Ab$ subject to inclusion. But if $\mathfrak{a} \subset \mathfrak{p}$, then since $ab \in \mathfrak{p}$, either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$, implying that $\mathfrak{a}+Aa \subset \mathfrak{p}$ or $\mathfrak{a}+Ab \in \mathfrak{p}$. Thus we reach a contradiction. □

Most of the rings we encounter in basic algebra are Noetherian. Proving that such rings are Noetherian is quite easy, because every finitely generated algebra over a Noetherian ring is also Noetherian.

Lemma 9.2. *Every quotient ring over a Noetherian ring is Noetherian.*

Proof. Every ideal in the quotient ring corresponds to an ideal in the covering ring, which is finitely generated, and these generators project onto generators for the original idea. □

Thus, provided we can prove that $A[x_1, \ldots, x_n]$ is Noetherian if A is Noetherian, we can prove that all finitely generated algebras over Noetherian rings are Noetherian. This is provided by Hilbert's basis theorem.

Theorem 9.3. *If A is Noetherian, then $A[x]$ is Noetherian.*

Proof. Let $\mathfrak{a}$ be an ideal of $A[x]$. The set $\mathfrak{b}$ of $a \in A$ which are the leading coefficients of polynomials in $\mathfrak{a}$ is an ideal, hence finitely generated. It follows that if we pick $f_1 \in \mathfrak{a}$ of minimal degree, then $f_2 \in \mathfrak{a} - (f_1)$ of minimal degree, then $f_3 \in \mathfrak{a} - (f_1, f_2)$, and so on. Eventually, the leading coefficients of $f_1, \ldots, f_n$ must generate $\mathfrak{b}$. Then $f_1, \ldots, f_n$ actually generate $\mathfrak{a}$, because if g has degree greater than f_n, then it's leading ceofficient is equal to $\sum b_n a_n$, and $g - \sum b_n X^{k_n} f_n$ has degree strictly less than g. Continuing this process, we find that g is equivalent to a polynomial of degree less than the degree of f_n in $\mathfrak{a}/(f_1, \ldots, f_n)$, but by assumption this polynomial is in $(f_1, \ldots, f_n)$, so $g \in (f_1, \ldots, f_n)$. □

A more general definition is useful in commutative algebra. A module M is *Noetherian* if every submodule is finitely generated. This is equivalent to an ascending chain condition for submodules. A quotient of a Noetherian module is Noetherian, as is the direct sum $M \oplus N$ of two Noetherian modules. To see the latter property, we note that if we have an exact sequence

$$0 \to K \to M \to N \to 0$$

Then M is Noetherian if and only if N and K are. If M is Noetherian, K can be identified as a submodule, and N as a quotient, so this is trivial. Conversely, if K and N are Noetherian, and M_0 is a submodule of M, then $M_0 \cap K$ is finitely generated by $x_1, \ldots, x_n$ and $M_0/N \cap M_0$, viewed as a submodule of N, is finitely generated by $y_1 + N, \ldots, y_m + N \in M_0$, with $y_n \in M_0$. Thus given any $x \in M_0$, there is a_n such that $x - \sum a_n y_n \in N = K$, so $x - \sum a_n y_n = \sum b_n x_n$, showing that $x \in (x_1, \ldots, x_n, y_1, \ldots, y_m)$.

Theorem 9.4. *Every finitely generated module over a Noetherian ring is Noetherian.*

Proof. Let A be a Noetherian ring, and M a finitely generated module over A. We let M be generated by $x_1, \ldots, x_n$, and we prove the theorem by induction on n. For $n = 1$, M is isomorphic to an ideal in A, which is finitely generated by assumption. In general, we have a commutative diagram

$0 \to Ax_n \to M \to M/Ax_n \to 0$ induced by multiplication, and Ax_n and M/Ax_n are Noetherian by induction, so M is also Noetherian. $\square$

Chapter 10

Graded Modules

A *graded ring* is a ring A which has a direct sum decomposition as

$$A = A_0 \oplus A_1 \oplus \ldots,$$

such that for each n and m, $A_n \cdot A_m \subset A_{n+m}$. Elements of A_n are known as *homogenous elements* of degree n. An ideal $\mathfrak{a}$ in a graded ring is *homogenous* if it is generated by homogenous elements, or, alternatively, if whenever $a \in \mathfrak{a}$, with $a = a_0 + a_1 + \ldots$ for $a_i \in A_i$, then $a_i \in \mathfrak{a}$. If $\mathfrak{a}_n = \mathfrak{a} \cap A_n$, then $A/\mathfrak{a}$ has a natural gradation as

$$A/\mathfrak{a} = A_0/\mathfrak{a}_0 \oplus A_1/\mathfrak{a}_1 \oplus \cdots.$$

Thus the quotient of a graded ring by a homogenous ideal also is naturally a graded ring. For any graded ring A, the set $A_+ = A_1 \oplus A_1 \oplus \ldots$ known as the *irrelevant ideal*, and the quotient A/A_+ is isomorphic to A_0, which is trivially graded. An obvious, though incredibly important fact about homogenous ideals is the following.

Example. *The fundamental example of a graded ring is the polynomial ring $k[x_1, \ldots, x_n]$, which is graded by degree. If $S = k[x_1, \ldots, x_n]$, then we can write*

$$S = S_1 \oplus S_2 \oplus \cdots$$

where for each k, S_k consists of homogenous polynomials of degree k. More generally, if V is a projective variety in $\mathbf{P}^n$ then it's homogenous coordinate ring has a natural gradation, since $I(V)$ is a homogenous ideal in $k[x_1, \ldots, x_n]$.

The natural modules to study over a graded rings A are the family of *graded modules*, an A module M with a decomposition as

$$M = M_0 \oplus M_1 \oplus \cdots,$$

where for each n and m, $A_n M_m \subset M_{n+m}$. A graded, or homogenous submodule N of a graded module M is then just a submodule that can be decomposed as a direct sum over the graded submodules, i.e. we can write

$$N = N_0 \oplus N_1 \oplus \cdots$$

where $N_i \subset M_i$ for each i. The quotient module M/N is then naturally graded as

$$M/N = (M_0/N_0) \oplus (M_1/N_1) \oplus \cdots.$$

Graded modules naturally occur when studying vector bundles over geometric spaces.

A simple, but powerful remark is useful in the analysis of graded modules. If M Is a finitely generated graded module, then it is certainly generated by finitely many homogenous elements of the module. If M is generated by homogenous elements $a_1, \ldots, a_N \in M$, with $a_i \in M_{n_i}$ for each i, then for each k and $a \in M_k$, we can write

$$a = \sum_{m=1}^{N} b_m a_m,$$

where $b_m \in A_{k-n_m}$ for each m. This is because we can certainly write this decomposition with $b_m \in A$ for each m, and then decompose b_m into a direct sum of gradations; any part of b_m that isn't in A_{k-n_m} will be cancelled out in the overall sum anyway.

Theorem 10.1. *If A is a graded ring with identity, then the following properties are equivalent to one another:*

1. *A is Noetherian.*
2. *A_0 is Noetherian, and A_+ is a finitely generated ideal of A.*
3. *A_0 is Noetherian, and A is a finitely generated A_0 algebra.*

Proof. Let us prove each implication separately:

- $(1 \Rightarrow 2)$ If A is Noetherian, then A_0 is Noetherian as a subring of A. Moreover, A_+ is an ideal of A, hence a finitely generated ideal of A since A is Noetherian.

- $(2 \Rightarrow 3)$ If A_0 is Noetherian, and A_+ is a finitely generated ideal of A with generators $(a_1, \ldots, a_N)$, with $a_i \in A_{n_i}$ for each i. We then claim that A is generated as an A_0 algebra by $\{1, a_1, \ldots, a_N\}$. We prove $A_n \subset A_0[1, a_1, \ldots, a_N]$ for each n by induction on n. The case $n = 0$ is trivial. If $n > 0$ and $a \in A_n$, then we can write

$$a = \sum_{i=1}^{N} c_i a_i$$

  where $a_0 \in A_0$, and $c_i \in A_{n-n_i}$ for each i. Since $n - n_i < n$, the inductive hypothesis implies $c_i \in A_0[1, a_1, \ldots, a_N]$ for each i, which implies that $a \in A_0[1, a_1, \ldots, a_N]$. This completes the inductive case and shows A is a finitely generated A_0 algebra.

- $(3 \Rightarrow 1)$ If A_0 is Noetherian, and A is a finitely generated A_0 algebra, then A is isomorphic to a quotient of the polynomial ring $A_0[x_1, \ldots, x_n]$; the ring $A_0[x_1, \ldots, x_n]$ is Noetherian, and so A is then Noetherian as well, as the quotient of a Noetherian ring. $\square$

Remark. Any finitely generated algebra over a field is Noetherian. This Lemma shows that if we work over the family of all graded rings A with A_0 a field, then such rings are Noetherian if and only if they are finitely generated over A_0.

The ideas behind graded modules were first developed by Hilbert, in the contexts of invariant theory. Here one studies an algebraic group G acting rationally on a variety V, inducing an action on the coordinate ring $k[V]$. We wish to study the ring $k[V]^G$ of invariants of G. Since constant functions are fixed by G, $k[V]^G$ is a k subalgebra of $k[V]$. Moreover, the group structure gives an 'averaging map' $\varphi : k[V] \to k[V]^G$, which is a graded $k[V]^G$ module homomorphism fixing $k[V]^G$. For finite groups G, we can choose

$$(\varphi f)(x) = \frac{1}{\#(G)} \sum_{a \in G} f(ax).$$

For larger groups, we must use integration against the Haar measure in G in a careful manner. If we define

$$k[V]k[V]_+^G = \{\sum_i b_i a_i : b_i \in k[V], a_i \in k[V]_+^G\},$$

then $k[V]k[V]_+^G$ is a homogenous ideal in $k[V]$. Since $k[V]$ is Noetherian, $k[V]k[V]_+^G$ is generated by finitely many homogenous elements $a_1, \ldots, a_n \in A_+$. If $a \in k[V]_+^G$, we may thus write

$$a = \sum_i b_i a_i,$$

and then $a = \varphi(a) = \sum_i \varphi(b_i)a_i$. Since $\varphi(b_i) \in k[V]^G$ for each i, this shows that $k[V]_+^G$ is a finitely generated ideal of $k[V]^G$. Since $k[V]_0^G = k$ is a field, and hence trivially Noetherian, we have seen this implies that $k[V]^G$ is a finitely generated algebra over k.

10.1 The Hilbert Function

Chapter 11

K Theory

Theorem 11.1 (Steinitz)**.** *Let R be a Dedekind domain. Let $P = \mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_r$, and let $Q = \mathfrak{b}_1 \oplus \cdots \oplus \mathfrak{b}_s$ be finitely generated projective modules, where the ideals is the direct sum are nonzero. Then P is isomorphic to Q if and only if $r = s$ and $\mathfrak{a}_1 \dots \mathfrak{a}_r = \mathfrak{b}_1 \dots \mathfrak{b}_s$.*

Proof. The last result implies P is isomorphic to $\mathbf{R}^{r-1} \oplus \mathfrak{a}_1 \dots \mathfrak{a}_r$, and Q is isomorphic to $\mathfrak{r}^{s-1} \oplus \mathfrak{b}_1 \dots \mathfrak{b}_s$. Just because P is isomorphic to Q does not imply that $\mathfrak{a}_1 \dots \mathfrak{a}_r$ is isomorphic to $\mathfrak{b}_1 \dots \mathfrak{b}_s$ over *general rings*, but we find that such is true when working over Dedekind domains. First, we remark that for an R linear map $\phi : \mathfrak{a} \to \mathfrak{b}$, there is an element q in the fraction field k such that $\phi(a) = qa$ for all $a \in \mathfrak{a}$. To see this take any nonzero $a_0 \in \mathfrak{a}$. Then, in k,

$$\phi(a) = \frac{a_0 \phi(a)}{a_0} = \frac{\phi(a_0 a)}{a_0} = a \frac{\phi(a_0)}{a_0}$$

Therefore, associated to any R linear map ϕ there is an $r \times s$ matrix M with entries in K. If ϕ is an isomorphism, then M^{-1} exists, and so $r = s$. We now claim that $\det(M)\, \mathfrak{a}_1 \dots \mathfrak{a}_r$ □

Corollary 11.2. *If R is a Dedekind domain, then $K_0(R) \cong \mathbf{Z} \oplus \widetilde{K_0}(R)$, and as a group, $\widetilde{K_0}(R)$ is isomorphic to the class group of R. Moreover, the product of any two elements of the reduced group is zero.*

Proof. The group $\widetilde{K_0}(R)$ is the kernel of the map from $K_0(R)$ to $\mathbf{Z}$. There is a correspondence $[\mathfrak{a}_1 \oplus \dots \mathfrak{a}_r] \mapsto r$ (taking the rank of the module) which extends to an isomorphism from $K_0(R)$ to $\mathbf{Z} \oplus \mathrm{Cl}(R)$. To prove the product

of any two elements of $\tilde{K_0}(R)$ is zero, we consider $[\mathfrak{a}]-1$ and $[\mathfrak{b}]-1$ in $\tilde{K_0}(R)$. Then

$$([\mathfrak{a}-1])([\mathfrak{b}]-1) = [\mathfrak{a}\otimes\mathfrak{b}]-[\mathfrak{a}]-[\mathfrak{b}]+1$$

Since $\mathfrak{a}\oplus\mathfrak{b} \cong R\oplus\mathfrak{a}\,\mathfrak{b} \cong R\oplus(\mathfrak{a}\otimes\mathfrak{b})$, this is zero. Because the elements $[\alpha]-1$ generate $\tilde{K_0}(R)$. □

11.1 Invertible Modules

A finitely generated module M over an commutative ring is invertible if there exists some module N such that $M\otimes N$ is isomorphic to R. We have a canonical homomorphism from $M\otimes M^*$ to R. Whenever M is invertible, this is precisely an isomorphism.

Lemma 11.3. *If P is a finitely generated projective module then P^* is finitely generated and projective, and $(P^*)^* \cong P$.*

Proof. If P is finitely generated, we have an exact diagram

$$0 \to Q \to R^n \to P \to 0$$

which induces an exact diagram

$$0 \to P^* \to (R^n)^* \to Q^* \to 0$$

and if the first diagram splits, the second one splits. Thus $R^n = P\oplus Q$, and $(R^n)^* \cong P^*\oplus Q^*$, and $(R^n)^*$ is isomorphic to R^n, showing P^* is projective. The double dual map $\nu : M \to M^{**}$ is an isomorphism if $M = R^n$. If P is finitely generated and projective, then $\nu : P \to P^{**}$. If $\mathfrak{p}$ is a prime ideal of R, then $R_\mathfrak{p}\otimes P \to R_\mathfrak{p}\otimes P^{**}$. □

M is an invertible R module if and only if M is finitely generated and projective of rank 1. TODO: ADD PROOF.

The Picard group of a commutative ring R is the group of isomorphism classes of invertible R modules, with the operation being the tensor product. We have an inclusion from the Picard group of R to $K_0(R)$, which is a morphism of multiplicative monoids. It is an inclusion of groups, but not necessarily an isomorphism.