Rings and modules

Based on lectures by Richard Borcherds
Written by Jack DeSerrano

Last updated on April 29, 2022

These notes are based on Richard Borcherds's YouTube series on rings and modules (at https://www.youtube.com/playlist?list=PL8yHsr3EFj52XDLrmvrFDgwcf6XOm2TEE).

Contents

1 Introduction

Examples of rings are $\mathbf{Z}$, $\mathbf{R}$, and $\mathbf{C}$. All of them have an addition, subtraction, and multiplication. We also have polynomial rings like $\mathbf{R}[x]$, matrix rings such as $\mathrm{M}_n(\mathbf{R})$, the Gaussian integers, and coordinate rings like $\mathbf{C}[x,y]/(y^2-x^3-x)$. The quaternions are another example of a ring.

Suppose R is a set with operations $+$ and $\times$. The set R is a **ring** if

1. it is an abelian group under $+$,
2. $\times$ is associative, and
3. $a \times (b+c) = (a \times b) + (a \times c)$, $(a+b) \times c = (a \times c) + (a \times b)$.

There are optional axioms:

4. $\times$ is commutative,
5. there exists $1 \in R$ such that $1 \times a = a \times 1 = a$ for all $a \in R$.

Algebraists tend to require the existence of a multiplicative identity.

Groups and rings are somewhat analogous. Groups act on sets while rings act on **modules**. A module over a ring is like a vector space over a field: if R is a field, then R-modules are vector spaces. If R is a ring and M is a module, then we have a distributive multiplication $r(m)$ satisfying the associative condition $(rs)m = r(sm)$, and if R has a multiplicative identity then $1m = m$.

Example 1.1 (Modules)**.** Vector spaces are modules of fields. The R-modules of $R = \mathbf{Z}$ are abelian groups. Ideals of a ring R are submodules.

Group actions on a set are left, right, or two-sided (where $(gs)h = g(sh)$). Modules are left, right, or two-sided, and if the ring is commutative then there is no difference between the three.

If G acts on sets S and T then G acts on the disjoint union of S and T. If M and N are R-modules then $M \oplus N$ is an R-module. The group G acts on the product $S \times T$, and the analogue here is the tensor product: if V and W are vector spaces with bases v_i and w_j, then the vector space whose basis consists of the expressions $v_i \otimes w_j$ is the tensor product $V \otimes W$. One can see that

$$\dim(V \otimes W) = (\dim V)(\dim W)$$

just as $|S \times T| = |S| \cdot |T|$. The general definition of the tensor product is harder to define, so we won't explore it yet.

If X and Y are sets then $|X \cup Y| = |X| + |Y| - |X \cap Y|$, and if U and V are subspaces of W then

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V).$$

If Z is a set then

$$|X \cup Y \cup Z| = |X| + |Y| + |Z| - |X \cap Y| - |X \cap Z| - |Y \cap Z| + |X \cap Y \cap Z|.$$

The same is not true for three vector spaces. Three one-dimensional vector spaces in two-dimensional space is a simple counterexample.

Recall Cayley's theorem: every group G is the symmetry of something. This something is the set G acted on the right by G and the set of symmetries is G acting on the left. Similarly, every ring is the set of endomorphisms of a linear object M. A linear object is an abelian group with extra structure and an endomorphism is a map from the object to itself preserving the linear structure.

Suppose R is a ring. If we take $M = R$ with a right action of R, the set of endomorphisms is R acting on the left. So rings are the endomorphisms of linear objects in the same way that groups are the symmetries of objects.

A homomorphism of rings preserves all ring operations. One should be careful, though. The map $\mathbf{Z}/6\mathbf{Z} \longrightarrow \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$ is an isomorphism of groups and rings; however, the ring $\mathbf{Z}/2\mathbf{Z}$ is not a subring of $\mathbf{Z}/6\mathbf{Z}$ (even though it is a subgroup): the map does not take $1_{\mathbf{Z}/2\mathbf{Z}}$ to $1_{\mathbf{Z}/6\mathbf{Z}}$.

If G acts on S and T, then we can consider maps from S to T preserving the action of G (maps of G-sets). Similarly, we have homomorphisms (or linear transformations) of modules.

If M and N are left modules, then homomorphisms $f : M \longrightarrow N$ should be on the right since we want $(rm)f = r(mf)$.

A subgroup H of G is a normal subgroup if it is the kernel of a group homomorphism. (Moreover, H is normal in G if $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$.) The ring-analogue of a normal subgroup is an **ideal**, and we have left, right, and two-sided ideals. A two-sided ideal I of R is the kernel of a ring homomorphism: we have

1. I is closed under addition,
2. $0 \in I$, and
3. I is closed under multiplication by elements in R.

Example 1.2. Though $\mathbf{R}[x^2]$ is a subring of $\mathbf{R}[x]$, it is not an ideal as it is not closed under multiplication by x.

If N is a left R-module, then we have R-submodules $M \subset N$. If $N = R$, then M is a left ideal: a left submodule of R.

Example 1.3. Suppose $R = \mathrm{M}_2(\mathbf{Z})$. Then

$$I_1 = \left\{ \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix} : \alpha, \beta \in \mathbf{Z} \right\}$$

is a right ideal of R,

$$I_2 = \left\{ \begin{pmatrix} \alpha & 0 \\ \beta & 0 \end{pmatrix} : \alpha, \beta \in \mathbf{Z} \right\}$$

is a left ideal of R, and

$$I_3 = \left\{ \begin{pmatrix} 2\alpha & 2\beta \\ 2\gamma & 2\zeta \end{pmatrix} : \alpha, \beta, \gamma, \zeta \in \mathbf{Z} \right\}$$

is a two-sided ideal of R. Furthermore, we have $R/I_1 = \mathbf{Z}^2$ and $R/I_2 = \mathbf{Z}^2$ (the difference is the side on which R acts). Also, $R/I_3 = \mathrm{M}_2(\mathbf{Z}/2\mathbf{Z})$.

The ring-analogue of the symmetric group S_n is the symmetry group of

$$R^n = \underbrace{R \oplus \cdots \oplus R}_{n \text{ times}}.$$

These sums of copies of R are called **free modules**. In particular, if R is a field, then all modules are free modules. The analogue of n points is a free module on n generators, so the analogue of S_n is the set of linear transformations of R^n, namely $\mathrm{M}_n(R)$. One can see that $\mathrm{M}_n(R)$ is a ring by noticing that the linear transformations of a module is automatically a ring. (This is a good exercise.)

2 Group rings

There is a functor from the category of groups to the category of rings taking a group G to its group ring and group homomorphisms to ring homomorphisms. The group ring of G is $\bigoplus_{g \in G} \mathbf{Z}g$ where $g_i g_j$ in the ring is defined by $g_i g_j$ in the group. We extend this definition to an arbitrary (typically commutative) ring R so that the group algebra of G is

$$R[G] = \bigoplus_{g \in G} Rg$$

where the multiplication is the same as described above.

Definition 2.1. An **R-algebra** is a ring S with a ring homomorphism $R \longrightarrow S$. (Often R is commutative and its image commutes with S.)

Suppose G is the Klein four group. The group ring $\mathbf{C}[G]$ splits as a product of 4 rings. (It certainly splits as a product of vector spaces:

$$\mathbf{C} \oplus \mathbf{C}g_1 \oplus \mathbf{C}g_2 \oplus \mathbf{C}g_3,$$

though this obvious construction is not a product of rings.)

What do products of rings look like? Suppose

$$R = S \times T = \{(s,t) : s \in S, t \in T\}$$

is a product of rings. We notice that $u = (1,0)$ commutes with R and that u is an **idempotent**: so $u^2 = u$. If a ring R splits as a product of rings then one can find an idempotent commuting with R. Conversely, given such a u one can reconstruct R as a product of rings by $R = Ru \times R(1-u) = uR \times (1-u)R$. (Showing that this product is well-defined is a good exercise.)

Consider the following idempotents:

$$u_1 = \frac{1+g_1+g_2+g_3}{4}, \quad u_2 = \frac{1+g_1-g_2-g_3}{4},$$
$$u_3 = \frac{1-g_1+g_2-g_3}{4}, \quad u_4 = \frac{1-g_1-g_2+g_3}{4}.$$

One can confirm that $u_i^2 = u_i$ and $u_i u_j = 0$ for $i \neq j$. Then the group ring splits as follows:

$$\mathbf{C}[G] = \mathbf{C}u_1 \times \mathbf{C}u_2 \times \mathbf{C}u_3 \times \mathbf{C}u_4.$$

In general, if G is a finite group, then the group algebra splits as a product of matrix rings over the complex numbers:

$$\mathbf{C}[G] = \prod_i \mathrm{M}_{n_i}(\mathbf{C}).$$

In the case of the four group, the n_is were 1. The n_is are the dimensions of the irreducible representations of G.

One notices that G does not need to be a group: it can be a monoid.

Example 2.2. Suppose $G = (\mathbf{Z}^+ \cup \{0\}, +)$. Then $R[G] = R[x]$. If $G = (\mathbf{Z}, +)$, then $R[G] = R[x][x^{-1}]$ (the Laurent polynomials over R).

Example 2.3. If $G = (\mathbf{Z}^+, \times)$, then $R[G]$ is the ring of formal Dirichlet polynomials over R: things of the form

$$\frac{r_1}{1^s} + \frac{r_2}{2^s} + \frac{r_3}{3^s} + \cdots$$

where formally $(1/m^s)(1/n^s) = 1/(mn)^s$. If $R = \mathbf{C}$ then one can think of formal Dirichlet polynomials as functions.

We have the ring of formal power series $R[\![x]\!]$ consisting of all elements of the form $\sum_n a_n x^n$ where the a_n are in R. We can do the same for formal Dirichlet polynomials. If $R = \mathbf{C}$, then one can think of these elements as (possibly convergent) Dirichlet series, such as the Riemann zeta function and its multiplicative inverse. One notices that

$$M(s) := \frac{1}{\zeta(s)} = \sum_n \frac{\mu(n)}{n^s}$$

where

$$\mu(n) := \begin{cases} 0 & p^2 \mid n \text{ for some prime } p \\ (-1)^{\omega(n)} & \text{otherwise} \end{cases}$$

and $\omega(n)$ counts the distinct prime factors of n. So $M\zeta = 1$ in the ring of formal Dirichlet series. This identity is equivalent to

$$\sum_{d \mid n} \mu(d) = \begin{cases} 1 & n = 1 \\ 0 & n > 1. \end{cases}$$

The group ring is an adjoint functor. There is an obvious functor from the category of R-algebras to the category of groups taking an R-algebra S to the invertible elements of S. One can think of this (right adjoint) functor as a forgetful functor. Indeed, we have a functor taking a group G to the group ring $R[G]$. This functor is a left adjoint functor: there is a correspondence between group homomorphisms from G to S^* and ring homomorphisms from $R[G]$ to S. (One can think of $R[G]$ being freely generated by G.)

One can construct the free R-algebra (where R is in the centre of the R-algebra) generated by a set S by taking the free monoid G on the elements of S: the monoid ring $R[G]$ is the free R-algebra generated by the elements of S. In the commutative case, one takes the free commutative monoid on S, which is just a polynomial ring.

A Laurent polynomial is a map $\mathbf{Z} \longrightarrow \mathbf{C}$ that is 0 for almost z: one can think of a Laurent polynomial $\sum a_n z^n$ as the map $n \longmapsto a_n$. Suppose f and g are two functions $\mathbf{Z} \longrightarrow \mathbf{C}$. Then the convolution

$$(f * g)(n) \longmapsto \sum_m f(m) g(n - m)$$

defines multiplication in the ring. Now let us consider continuous functions $f, g : \mathbf{R} \longrightarrow \mathbf{C}$ with compact support. Now we define multiplication by the convolution

$$(f * g)(x) \longmapsto \int_0^x f(y) g(x - y)\, dy.$$

One can verify that this well-defines a ring with no identity.

Recall that one defines the Fourier transform by

$$\widehat{f}(x) := \int_{-\infty}^{\infty} f(y)e^{-2\pi ixy}\,dy$$

up to normalization. Under suitable conditions one has $\widehat{fg} = \widehat{f} * \widehat{g}$: the Fourier transform is approximately a ring homomorphism from a ring of continuous functions under pointwise multiplication to a ring of continuous functions under convolution (one needs to add conditions to ensure integral convergence).

3 Burnside ring and rings of differential operators

Let G be a finite group. The set of all actions of G on finite sets almost forms a ring. The addition is the disjoint union, the multiplication is the product, the zero element is $\emptyset$, and the unit is the one point set. We have no "subtraction," though.

Let $G = \mathrm{S}_3$. Any action of G on S is a union of transitive actions. Transitive actions correspond to subgroups of G up to conjugacy. So one has an action of G on 1, 2, 3, or 6 elements. So the almost-ring consists of linear combinations of these actions by nonnegative integers.

One turns this "almost-ring" into a ring by taking formal linear combinations of these classes by any integers. This ring is called the **Burnside ring** of S_3. This construction is reminiscent of forming the integers from the natural numbers. Generally, one can force a semigroup X into a group G by the same process. (Take care, though: the map from the semigroup to the group does not need to be injective.) The group G is sometimes called a **Grothendieck group**.

There is a contravariant functor from the category of groups to the category of rings: if one has a map from a group G to a group H then one has a map $\mathrm{Burn}(G) \longleftarrow \mathrm{Burn}(H)$. If $G = \{1\}$ and $H = \mathrm{S}_3$, then one gets a map $\mathrm{Burn}(\{1\}) = \mathbf{Z} \longleftarrow \mathrm{Burn}(\mathrm{S}_3)$.

A variation on the Burnside ring is the **representation ring** of G: instead of considering the action of G on finite sets, one considers the action of G on complex vector spaces. If G is finite, any complex vector space can be written as a sum of irreducible representations of G in the same way that any set acted on by G can be written as a union of transitive orbits. One endows a ring structure on this object by defining addition as addition of vector spaces and multiplication by the tensor product.

We also have **rings of differential operators**. Consider differential operators on the real line with polynomial coefficients. This ring contains two obvious elements x and $D = d/dx$, so one gets differentiable operators by $\mathbf{R}[x, D]$. However, $Dx = xD + 1$ (the generators don't commute). So $\mathbf{R}[x, D]$ has a basis of things of the form $x^m D^n$ where $m, n \geq 0$. One can do this in arbitrarily many variables.

One can make differential equations modules over this ring. Suppose

$$E = x^2\frac{d^2}{dx^2} + x\frac{d}{dx} + (x^2 - \alpha^2)$$

is a homogenous linear differential operator. The operator E corresponds to Bessel's differential equation: the solutions to $Ey = 0$ are Bessel functions. Let $R = \mathbf{R}[x, d/dx]$ be the ring of differential operators over the real numbers. Then the quotient $R/R(E)$ is a left R-module. Let M be the smooth functions on $\mathbf{R}$. Then M is also a left R-module. If one defines a map $R \longrightarrow M : E \longmapsto Ef$, then maps $R/R(E) \longrightarrow M$ correspond to solutions to Bessel's equation. One can construct the Bernstein-Sato polynomial of a differential operator using modules over this ring.

4 Unique factorization

Theorem 4.1 (Fundamental theorem of arithmetic)**.** *Every positive integer is a product of prime numbers in a unique way up to order.*

One can generalize this notion to commutative rings with no zero divisors (these rings are called **integral domains**). It is useful to rephrase the fundamental theorem of arithmetic to say "every nonzero integer is a product of prime numbers and a unit uniquely up to order and units." (A **unit** in a ring is an element with a multiplicative inverse in the ring. The units in $\mathbf{Z}$ are ± 1.)

Definition 4.2. A nonzero element $p \in R$ is **prime** if p is not a unit and

$$p \mid ab \implies p \mid a \vee p \mid b$$

for all $a, b \in R$.

Definition 4.3. A nonzero element $p \in R$ is **irreducible** if p is not a unit and if $p = ab$ then a or b is a unit.

If p is prime, then p is irreducible.

Proposition 4.4. *The ring $\mathbf{Z}$ is a Euclidean domain. So $\mathbf{Z}$ is a principal ideal domain. Hence $\mathbf{Z}$ is a unique factorization domain.*

Definition 4.5. A ring R is a **Euclidean domain** if there is a norm map $R \longrightarrow \mathbf{Z}^+ : r \longmapsto |r|$ and $a, b \in R$ and $a \neq 0$ implies that one can write $b = aq + r$ with $|r| < |a|$.

Example 4.6 (Euclidean domains)**.** The ring $\mathbf{Z}$ is Euclidean with respect to $|\cdot|$. If k is a field, the ring $k[x]$ is Euclidean with respect to

$$f \longmapsto \begin{cases} 1 + \deg f & f \neq 0 \\ 0 & f = 0. \end{cases}$$

Proposition 4.7. *The ring of Gaussian integers $\mathbf{Z}[i]$ is a Euclidean domain.*

Proof.—Let $m+ni \in \mathbf{Z}[i]$. Define the norm of $m+ni$ by $\|m+ni\| = m^2+n^2$. If $b \in \mathbf{Z}[i]$, we want to show that $b = qa+r$ where $\|r\| < \|a\|$. So one has

$$\frac{b}{a} = q + \frac{r}{a}$$

where $\|r/a\| < 1$. Notice that $b/a \in \mathbf{C}$. We want to show that every complex number can be written as a sum of a Gaussian integer and a complex number with norm less than 1. Indeed, (and this fact is easily seen geometrically) $\mathbf{C}$ is covered by open unit disks centred at integer lattice points (Gaussian integers). So $\mathbf{Z}[i]$ is Euclidean. □

Definition 4.8. An ideal I of R is a **principal ideal** if I is generated by one element. If I is generated by $a \in R$, then one writes $I = Ra = (a)$.

Definition 4.9. A ring R is a **principal ideal domain** (or PID) if every ideal of R is a principal ideal.

Example 4.10 (Non–PIDs)**.** The ring $\mathbf{C}[x,y]$ has an ideal I of polynomials with 0 constant term that is not principal (though I is generated by x and y). The ring $\mathbf{Z}[x]$ has an ideal $I = (2,x)$ that is not principal.

Proposition 4.11. *Suppose R is a Euclidean domain. Then R is a PID.*

Proof.—Pick an ideal $I \neq (0)$. Pick a minimal nonzero element $a \in I$. Let $b \in I$. We can write $b = aq+r$ with $|r| < |a|$, so $r = 0$ since a is minimal. □

Example 4.12 (Non–Euclidean PID)**.** The ring

$$R = \mathbf{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$$

is a principal ideal domain but not a Euclidean domain.

Proof.—Go to an algebraic number theory course if you want to see how R is a principal ideal domain. If R is Euclidean but not a field, pick a nonzero non-unit minimal element $a \in R$. Every element in $R/(a)$ is represented by 0 or a unit. The units of R are ± 1. So the set of units with zero has three elements. If $a \in R$ is not ± 1 or 0, then $R/(a)$ has at least 4 elements (it turns out that $\#R/(a) = |a|^2 \geq 4$ unless $a = 0, \pm 1$). So $R/(a)$ cannot be represented by units and 0. So R is not Euclidean. □

One can replace -19 by -43, -67, or -163.

Proposition 4.13. *Suppose R is a PID. Then R is a unique factorization domain (a ring with unique factorization).*

Example 4.14 (Non–PID UFD)**.** The ring $\mathbf{C}[x, y]$ is a UFD but not a PID.

Proof.—(Existence.) Every nonzero non-unit element is divisible by an irreducible element. Suppose $a_0 = a_1 b_1$ is not irreducible, so a_1 and b_1 are not units. Then we write $a_1 = a_2 b_2$, $a_2 = a_3 b_3$, et cetera until an a_i is irreducible, and $a_i \mid a_0$ as desired. This process might not stop, so suppose it doesn't. Then we get an infinite sequence $\{a_n\}$ where $a_{i+1} \mid a_i$ and a_i/a_{i+1} is not a unit. The ideal $I = (a_0, a_1, \ldots)$ is principal, so $I = (b)$ for some b. The element b is a linear combination of finitely many a_is, say $a_0, \ldots, a_k$. So b is a multiple of a_k and divides a_{k+1}. So $a_{k+1} \in (b)$. Since $a_k \mid a_{k+1}$ and $a_{k+1} \mid a_k$, a_{k+1} is a_k times a unit, which is a contradiction. So the process above terminates, and every nonzero non-unit element is divisible by an irreducible element.

We want to write a_0 as a product of irreducible. Write $a_0 = a_1 r_1$ where r_1 is irreducible. We can do this unless a_0 is a unit, in which case we stop. We write $a_1 = a_2 r_2$ where r_2 is irreducible, and continue until a_n is a unit, then we stop, and a_0 is the product of a unit with irreducible. Suppose this process doesn't terminate. Then we get a sequence $\{a_n\}$ where $a_{i+1} \mid a_i$ and a_i/a_{i+1} is not a unit. This situation is impossible in a PID, so in a PID, every nonzero non-unit element is a product of irreducibles. (The same works for Noetherian rings.) □

Proof.—(Unicity.) Suppose $a = p_1 \cdots p_n = q_1 \cdots q_m$ where p_i, q_i are irreducible. We want to show that these factorizations are unique up to order and units. Assume that irreducibles are primes. So p_1 is prime. So $p_1 \mid q_1 \cdots q_m$, so $p_1 \mid q_1$. But q_1 is irreducible, so q_1 is p_1 times a unit. So $\{p_i\} = \{q_i\}$.

We need to show that irreducibles are prime in a PID. Suppose p is irreducible and $p \mid ab$. We want to show that $p \mid a$ or $p \mid b$. Consider the ideal $(p, a) = (c)$. If $p \mid a$, we're done. If not, then $(p, a) = R$ since $c \mid p$, so c is a unit or p. So $xp + ya = 1$ for some x, y (since $1 \in R$ and all elements in (p, a) are combinations of p and a). So $xpb + yab = b$, and since $p \mid xpb$ and $p \mid yab$, we must have $p \mid b$. So a PID is a UFD. □

5 Examples of unique factorizations

The rings $\mathbf{Z}$ and, if k is a field, $k[x]$ are obvious unique factorization domains. We will first consider the Gaussian integers, which we will denote R. In this ring, we define $|m + ni| = m^2 + n^2$. So $|ab| = |a| \cdot |b|$. What are the primes and units in R?

An element $r \in R$ is a unit if $|r| = 1$. So the units are $\mp 1, \mp i$. Suppose $x + iy$ is a Gaussian integer. Then $(x + iy) \mid (x^2 + y^2) = (x + iy)(x - iy)$, and $(x^2 + y^2) \in \mathbf{Z}$. So any prime in R must divide a prime $p \in \mathbf{Z}$. One has $|p| = p^2$, so either p is a prime in R, or p factors as a product of two primes $(x + iy)(x - iy)$ in R where $|x + iy| = p$. So $x^2 + y^2 = p$, so factoring p is the same as writing p as a sum of two squares.

If $p = 2$, then $p = 1^2 + 1^2$, so $p = (1 + i)(1 - i) = -i(1 + i)^2$. The prime $p = 3$ cannot be written as the sum of two squares, so $\pm 3, \pm 3i$ are primes in R. One sees $5 = 2^2 + 1^2$, so

$5 = (2+i)(2-i)$, and we get 8 primes in R, namely $2+i$, $-1+2i$, $-2-i$, $1-2i$, $2-i$, $-1-2i$, $-2+i$, and $1+2i$. The primes 7 and 11 are both prime in R, and $13 = 3^2+2^2 = (3+2i)(3-2i)$ splits like 5 did.

When is a prime a sum of two squares? A square is 0 or 1 mod 4, so a sum of two squares is $0, 1$, or 2 mod 4. So if $p \equiv 3 \pmod 4$, then p is a prime in R. If $p \equiv 1 \pmod 4$, p always splits as a product of two primes in R.

Proposition 5.1 (Fermat)**.** *An odd prime p can be written as a sum of two squares if and only if $p \equiv 1 \pmod 4$.*

Proof.—The group $(\mathbf{Z}/p\mathbf{Z})^*$ is cyclic of order $p-1$, which is divisible by 4, so there must be an element a of order 4. Further, a^2 has order 2, so $a^2 \equiv -1$. So $a^2 = np - 1$, and

$$(a+i)(a-i) = np.$$

So p divides $(a+i)(a-i)$ but does not divide $(a+i)$ or $(a-i)$. So p is not prime in the Gaussian integers. Since $\mathbf{Z}[i]$ is a UFD, p is not irreducible. So write $p = (x+iy)(\text{something})$ where $|x+iy| = p$, so $x^2+y^2 = p$. □

The ring $R = \mathbf{Z}\left[\sqrt{-2}\right]$ is also a UFD. (This fact can be seen by drawing open unit disks centred at each lattice point.) Instead of checking whether primes can be written as x^2+y^2, we check whether they can be written as x^2+2y^2.

Is $\mathbf{Z}\left[\sqrt{-3}\right]$ a UFD? When we draw open unit disks, we see that, for example, $1/2 + (\sqrt{3}/2)i$ is not in a disk. In fact, $\mathbf{Z}\left[\sqrt{-3}\right]$ is not a UFD:

$$4 = 2 \cdot 2 = (1+\sqrt{-3})(1-\sqrt{-3}),$$

and $2, 1 \pm \sqrt{-3}$ are irreducible and distinct considering unit multiplication. Notice that $1+\sqrt{-3}$ is not prime.

One can fix this, though. Notice that

$$\mathbf{Z}\left[\sqrt{-3}\right] \subset \mathbf{Z}\left[\frac{1+\sqrt{-3}}{2}\right] = R,$$

and R is a ring and a UFD. (One can show this by the unit disk method.)

One notices that $R = \mathbf{Z}\left[\sqrt{-5}\right]$ is not a UFD:

$$6 = 2 \cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5}).$$

However, one also sees that

$$\mathbf{Z}\left[\frac{1+\sqrt{-5}}{2}\right] \neq \left\{m + n\frac{1+\sqrt{-5}}{2} : m, n \in \mathbf{Z}\right\},$$

and $\mathbf{Z}\left[\left(1+\sqrt{-5}\right)/2\right]$ is not a lattice in $\mathbf{C}$. There is no way to extend this ring to a UFD nontrivially.

What do non-principal ideals look like in R? The principal ideals look like aR, which is a rectangular lattice like R (scaled and rotated according to $|a|$ and $\arg a$). The ideal $(2, 1+\sqrt{-5})$ is not principal, and its lattice is diamond-shaped. Non-principal ideals look different than principal ideals. For this particular ring, all non-principal ideals are of this shape. Depending on the ring, one can get many different shapes of ideals. The number of shapes of ideals is called the class number of the quadratic field.

6 Prime and maximal ideals

All rings are commutative for this section.

Definition 6.1. An ideal $\mathfrak{m} \subseteq R$ is a **maximal ideal** if and only if $R/\mathfrak{m}$ is a field. Also, $\mathfrak{m}$ is maximal if there are no ideals containing $\mathfrak{m}$ except for R.

Definition 6.2. An ideal $\mathfrak{p} \subseteq R$ is a **prime ideal** if $R/\mathfrak{p}$ is an integral domain. Also, $\mathfrak{p}$ is prime if $ab \in \mathfrak{p}$ implies $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$ and $\mathfrak{p} \neq R$.

If $p \in R$ is prime, then (p) is a prime ideal.

Example 6.3 (Prime principal ideal does not imply prime element)**.** The ideal (0) is a prime ideal in $\mathbf{Z}$, but 0 is not prime.

Suppose X is a compact Hausdorff space. Let $R = C(X)$ be the ring of continuous complex-valued functions on X. This ring is easy to visualize. Is every ring of this form? No. But, in some sense, the answer is yes.

What are the points and topology of X in terms of R? (Can you reconstruct X from R?) Suppose $x \in X$ is a point. Then we have a homomorphism $\psi : R \longrightarrow \mathbf{C} : f \longmapsto f(x)$. One sees that $\ker \psi$ is a maximal ideal. So the points of X correspond to maximal ideals of R.

Let $f \in R$. A basis of open sets is given by the set of points x where $f(x)$ is nonzero. This set corresponds to maximal ideals $\mathfrak{m}$ with $f \notin \mathfrak{m}$. One can do the same for any ring R, and we get a topological space called the **maximal spectrum** of R, denoted Spec max(R). This is not a functor from rings to topological spaces. Suppose $\psi : R \longrightarrow S$ is a map of rings and $\mathfrak{m}$ is maximal in S. Though $\psi^{-1}(\mathfrak{m})$ is an ideal in R, it doesn't need to be maximal. If $\psi : \mathbf{Z} \longrightarrow \mathbf{Q}$, the ideal (0) is maximal in $\mathbf{Q}$, and its inverse image $\psi^{-1}((0)) = (0)$ is not maximal in $\mathbf{Z}$.

Let's figure out why this goes wrong. Suppose $\psi : R \longrightarrow S$ is a map of rings and $\mathfrak{m}$ is maximal in S. Consider the diagram below.

$$\begin{tikzcd} R \arrow[r] \arrow[d] & S \arrow[d] \\ R/\psi^{-1}(\mathfrak{m}) \arrow[r, hook] & S/\mathfrak{m} \end{tikzcd}$$

Notice that $S/\mathfrak{m}$ is a field. The quotient $R/\psi^{-1}(\mathfrak{m})$ is a subring of a field: it is not necessarily a field, but it is an integral domain. Since a subring of an integral domain is an integral domain, the inverse of a prime ideal is prime. Now we define the **spectrum** of R, written $\operatorname{Spec}(R)$, to be the set of prime ideals of R. The topology is defined as before: take $f \in R$ and define

$$U_f := \{\mathfrak{p} \text{ prime in } R : f \notin \mathfrak{p}\}.$$

Then each U_f is an open subset of $\operatorname{Spec} R$ and

$$\{U_f : f \in R\}$$

is a basis for the (Zariski) topology. Now we have a contravariant functor

$$R \longrightarrow S : \operatorname{Spec} R \longleftarrow \operatorname{Spec} S.$$

This idea of using prime ideals instead of maximal ideals is due to Grothendieck as part of his reworking of algebraic geometry. It turns out that in the rings that algebraic geometers used to use—coordinate rings of varieties—the inverse of a maximal ideal is usually maximal.

Suppose $R = (0)$. Then $\operatorname{Spec} R = \emptyset$. This suggests a problem: to study a ring by looking at its spectrum, we better make sure the spectrum is nonempty. Given $R \neq (0)$, does R have maximal or prime ideals? Well, when does a set X with a partial order have a maximal element a? This occurs when the set is nonempty and any totally ordered subset Y has an upper bound in X. This is **Zorn's lemma**.

Proof.—Assume the axiom of choice. Pick $x_0 \in X$. If x_0 is not maximal, pick $x_1 > x_0$. If x_1 is not maximal, pick $x_2 > x_1$. If X is finite then this process terminates. If not, we get a chain $x_0 < x_1 < x_2 < \cdots$ that is totally ordered. Pick $x_\omega > x_0, x_1, \ldots$. If x_ω is not maximal, pick $x_{\omega+1} > x_\omega$, et cetera. Then we either find a maximal element or get a totally ordered subset $x_\omega < x_{\omega+1} < x_{\omega+2} < \cdots$. Then pick $x_{2\omega} > x_\omega, x_{\omega+1}, \ldots$. We pick x_α for an ordinal α, and the number of ordinals is larger than the cardinality of any set, so we eventually get a maximal element. □

Take X to be the set of ideals not equal to R, and X is nonempty if $R \neq (0)$. A totally ordered set of ideals is bounded, so we take their union which is an ideal (not equal to R): take $i \in I$ and $j \in J$. Then $I \subseteq J$ or $J \subseteq I$ since the set of ideals is totally ordered, so either $i+j \in J$ or $i+j \in I$, so we have closure of addition. So, by Zorn's lemma, we have a maximal ideal.

Given $I \neq R$, we can find a maximal ideal containing I.

Let S be a subset of R closed under multiplication. Suppose P is a maximal element of the set of ideals disjoint from S. Then P is a prime ideal.

Proof.—Suppose $ab \in P$. If $a \notin P$ and $b \notin P$, then $Pa \supset P$, so for some $s_1 \in S$, $s_1 = p_1 + r_1 a$ where $p_1 \in P$. Similarly, $s_2 = p_2 + r_2 b$, so $s_1 s_2 \in P$. This contradicts that P is disjoint from S. □

So $\operatorname{Spec} R \neq \emptyset$ if $R \neq (0)$.

Example 6.4. Suppose R is a field. Then $\operatorname{Spec} R$ is a point.

Example 6.5. Suppose $R = \mathbf{C}[x]$. The prime ideals of R are multiples of $(x-\alpha)$ for $\alpha \in \mathbf{C}$ and (0). So $\operatorname{Spec} \mathbf{C}[x]$ is the union of $\mathbf{C}$ with a generic point. Suppose $f = (x-\alpha_1)^{n_1}(x-\alpha_2)^{n_2}\cdots$. Then

$$U_f = (\{\text{generic point}\} \cup \mathbf{C}) \setminus \{\alpha_i\}.$$

So the only open sets apart from the empty set are the complements of finite sets of complex numbers. This is called the **finite complement topology**. It is not Hausdorff (any two nonempty sets have nonzero intersection).

Example 6.6. Let $R = \mathbf{Z}$. The prime ideals of R are (0) together with principal ideals generated by primes. Suppose $n = p_1^{n_1} p_2^{n_2} \cdots$, then

$$U_n = (\operatorname{Spec} \mathbf{Z}) \setminus \{p_i\}.$$

We almost get the finite complement topology (though (0) is in every open set), and this is not Hausdorff.

7 Localization

Let X be a compact Hausdorff space and $R = C(X)$. We can recover X as the maximal ideals of R where the topology is given by a basis of open sets: for each $f \in R$ we get an open set given by the maximal ideals not containing f.

One can endow extra structure on X: for every open set U, we map $U \longmapsto C(U)$. The conditions we put on these maps form a **sheaf**.

Suppose R is commutative. Recall the topology defined on $\operatorname{Spec} R$. We map each open set U_f to $R\left[f^{-1}\right]$ (with some work one can show that this defines a sheaf of rings on $\operatorname{Spec} R$). What is $R\left[f^{-1}\right]$?

Suppose $S \subseteq R$ as sets. Can one form a ring $R\left[S^{-1}\right]$ such that all elements in S are invertible? Suppose $t_1, t_2, \ldots \in S$. Then we take $R\left[S^{-1}\right] = R[t_1, t_2, \ldots]/(s_1t_1 - 1, s_2t_2 - 1, \ldots)$ to force t_i to be the inverse of s_i. This is universal. But we're losing control over R. Is $R \longrightarrow R\left[S^{-1}\right]$ injective? Is $R\left[S^{-1}\right] \neq \{0\}$?

We construct $\mathbf{Q}$ from $\mathbf{Z}$ by taking equivalence classes of pairs $(r, s) \in \mathbf{Z} \times (\mathbf{Z} \setminus \{0\})$ that we write r/s where $r_1/s_1 \sim r_2/s_2$ if $r_1s_2 = r_2s_1$. We try to copy this definition for any ring R and subset S. We define $R\left[S^{-1}\right]$ to be "equivalence classes" of pairs $(r, s) = r/s$ where we define

$$r_1/s_1 \sim r_2/s_2 \iff r_1s_2 = r_2s_1$$

if this defines an equivalence relation. One makes $(r_1/s_1)(r_2/s_2) := r_1r_2/s_1s_2$ and $r_1/s_1 + r_2/s_2 := (r_1s_2 + r_2s_1)/s_1s_2$, as well as $R \longrightarrow R\left[S^{-1}\right] : r \longmapsto r/1$. One needs S to be closed under multiplication and $1 \in S$, so S is a multiplicative subset.

We need to check if this well–defines an equivalence relation and satisfies ring axioms. The only problem is that $\sim$ does not need to be transitive. Suppose $r_1s_2 = r_2s_1$ and $r_2s_3 = r_3s_2$. Hence $r_1s_2s_3 = r_2s_1s_3 = s_1r_3s_2$, so we get $r_1s_3 = r_3s_1$ if we can cancel s_2, so we need s_2 not to be a zero divisor.

So if S is multiplicative and has no zero divisors, then one can construct $R\left[S^{-1}\right]$ as above, and $R \longrightarrow R\left[S^{-1}\right]$ is injective. When S has zero divisors, let I be the ideal of elements r where $rs = 0$ for some $s \in S$. (This is an ideal because S is multiplicative.) Then one takes

$$R \longrightarrow R/I \hookrightarrow (R/I)\left[S^{-1}\right] = R\left[S^{-1}\right]$$

where $(R/I)\left[S^{-1}\right]$ corresponds to the images of the elements of S in R/I. The second arrow corresponds to an injective map because the images of elements of S are not zero divisors. So the map $R \longrightarrow R\left[S^{-1}\right]$ has kernel I.

Example 7.1 (S has zero divisors)**.** Let R be continuous functions on $\mathbf{R}$ and let S be functions not vanishing at 0. Then $R \longrightarrow R\left[S^{-1}\right]$ is surjective and the kernel is the functions vanishing near 0. This map does not need to be injective.

Example 7.2. Let $R = \mathbf{Z}$. "Suppose you . . . want to kill the prime 2 and get rid of it." Then you localize by inverting 2: $\operatorname{Spec}(R[1/2]) = (\operatorname{Spec} R) \setminus (2)$. Suppose instead that you want to focus on 2: then you take

$$\operatorname{Spec}(R[1/3, 1/5, 1/7, 1/9, \ldots]) = \{(2), (0)\}$$

which can be written as $R_{(2)} = R\left[S^{-1}\right]$ where $S = \overline{(2)}$.

More generally, if $\mathfrak{p}$ is a prime ideal of any ring R, then the **localization** of R at $\mathfrak{p}$ is

$$R_{\mathfrak{p}} = R\left[S^{-1}\right], \quad S = \overline{\mathfrak{p}}.$$

The ring $R_{\mathfrak{p}}$ is an example of a **local ring**: it has a unique maximal ideal.

Suppose R is noncommutative. We can still invert all elements of S in a universal way: we take the ring $R[t_1, \ldots]$ of noncommutative polynomials in $t_1, \ldots$ and quotient out by the ideal $(t_1 s_1 - 1, s_1 t_1 - 1, \ldots)$.

Can all elements of $R\left[S^{-1}\right]$ be written rs^{-1}? Is $s^{-1}r = r_1 s_1^{-1}$ for some r_1 and s_1? Is $rs_1 = sr_1$? This is called the **right Ore condition**, and if it is satisfied things are much simpler. You also want $sr = 0 \implies \exists\, s_1 \in S$, $rs_1 = 0$.

8 Free modules

Suppose R is a ring. One defines the multiplication in the **opposite ring** ${}^{\text{op}}R$ of R as follows:

$$r_1 \times_{{}^{\text{op}}R} r_2 := r_2 \times_R r_1.$$

If M and N are left R-modules, then $\operatorname{Hom}_R(M, N)$ acts on the right: if $f \in \operatorname{Hom}_R(M, N)$, then we want $(rm)\, f = r\,(mf)$. In general, $\operatorname{Hom}_R(M, N)$ is not an R-module. If M and N are two-sided modules, then one can make $\operatorname{Hom}_R(M, N)$ a two-sided R-module.

The simplest R-modules are **free modules**, which are sums of copies of R. A free module M has a vector space–resemblant basis. Over a field, all modules are free modules.

Example 8.1 (Non–free module)**.** The $\mathbf{Z}$-module $\mathbf{Z}/2\mathbf{Z}$ is not free.

The essential invariant of a vector space is dimension: we define $\dim(k^n) := n$. We define the **rank** of a free R-module in a similar way: $\operatorname{rank}(R^n) := n$. Is this well defined?

If R is a field, it's well defined. If $R = \{0\}$, then it is not. If R is commutative and nonzero, then it is, since R has a maximal ideal $\mathfrak{m}$, so $R/\mathfrak{m}$ is a field; then we can take a free module R^n to $R^n/(\mathfrak{m}R)^n$, which is a vector space over $R/\mathfrak{m}$. We read off the rank of R^n by reducing mod $\mathfrak{m}$. The rank is also well defined if R is a finite–dimensional vector space over a field.

Example 8.2 (Rank of finite–dimensional vector space)**.** Let k be a field and $R = \mathrm{M}_n(k)$. One takes $\operatorname{rank}(R^n) := (\dim_k(R^n))/(\dim_k(R))$.

Example 8.3 (Nonzero ring with rank not well defined)**.** Pick any ring $S \neq \{0\}$. Pick any right S-module M with $M \cong M \oplus M$. (Take $M = \bigoplus S$, for instance.) Then $R = \operatorname{End}_S(M)$ acts on M on the left. If $M = \bigoplus S$, then R is "$\infty \times \infty$ matrices," though not quite. Take

$$\begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ \vdots \end{pmatrix} \in M$$

where almost all $s_i = 0$. Every element in R acting on elements in M must have almost all entries 0 in each of its columns. So M is acted on the right by S and on the left by R. Further, $M \cong M \oplus M$ as right S-modules. This isomorphism works as follows:

$$z \longmapsto \begin{pmatrix} az \\ cz \end{pmatrix},$$
$$bx + dy \longleftarrow \begin{pmatrix} x \\ y \end{pmatrix} \in M^2$$

with $a, b, c, d \in R$ where $\begin{pmatrix} b & d \end{pmatrix}$ acts on the vector in M^2 and $\begin{pmatrix} a \\ c \end{pmatrix}$ gives the linear transformation in the other direction. We must have

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

so $ba + dc = 1$, $ab = 1$, $ad = 0$, $cb = 0$, and $cd = 1$. If the ring is commutative, this implies that $1 = 0$.

In a ring with such a, b, c, d, if M is any R–bimodule—R acts on the left and the right—then $M \cong M \oplus M$. (This follows by using a, b, c, d to give an isomorphism in both directions.) This isomorphism is an isomorphism of right R–modules. Then, if $M = R$, we get $R \cong R \oplus R$ as right R–modules. In particular, there is a free right module of rank 2 isomorphic to a free right module of rank 1. So the rank of a module is not well defined. In this case, $\mathrm{M}_2(R) \cong R$ as right modules.

If one has a surjective map of modules from B to C and a free module R^n mapping to C, then one can lift this to a map from R^n to B. This is quite obvious: one takes each basis element's image in C and lifts it to B. More generally, any module (not necessarily free) with this lifting property is called a **projective module**.

9 Projective modules

Let R be a ring. Suppose we have an exact sequence of R–modules

$$\begin{tikzcd} A \arrow[r] & B \arrow[r] & C \arrow[r] & \{0\} \end{tikzcd}$$

and another R–module M. Then we get an exact sequence

$$\begin{tikzcd} \mathrm{Hom}_R(A, M) & \mathrm{Hom}_R(B, M) \arrow[l] & \mathrm{Hom}_R(C, M) \arrow[l] & \{0\}. \arrow[l] \end{tikzcd}$$

Why didn't we put $\{0\}$ at the beginning of the exact sequence? Consider

$$\begin{tikzcd}\{0\} \arrow[r] & \mathbf{Z} \arrow[r, "\times 2"] & \mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\}\end{tikzcd}$$

and let $M = \mathbf{Z}/2\mathbf{Z}$. By the process above, we get the sequence

$$\begin{tikzcd}\{0\} & \mathbf{Z}/2\mathbf{Z} \arrow[l] & \mathbf{Z}/2\mathbf{Z} \arrow[l, "\times 2"'] & \mathbf{Z}/2\mathbf{Z} \arrow[l] & \{0\} \arrow[l]\end{tikzcd}$$

which is not exact (the ×2 map is not onto).

Suppose that C is a free R–module and consider

$$\begin{tikzcd}\{0\} \arrow[r] & A \arrow[r] & B \arrow[r] & C \arrow[r] & \{0\}.\end{tikzcd}$$

This sequences splits, and if C is free, we have a homomorphism from C to B:

$$\begin{tikzcd}\{0\} \arrow[r] & A \arrow[r] & B \arrow[r] & C \arrow[r] \arrow[l, bend left, dashed, "\psi"] & \{0\}.\end{tikzcd}$$

Let $c_1, c_2, \ldots$ be a basis for C. Map c_i to an element in B with image c_i. This defines a map from C to B. The module B splits as $A \oplus \mathrm{im}_\psi(C)$. If B is a direct sum, it is straightforward to check that the resulting sequence is exact.

Definition 9.1. Suppose there is a surjective module homomorphism f from A to B and a module homomorphism g from C to B. The module C is a **projective module** if it has a lifting property: namely there exists a module homomorphism h from C to A such that $g = f \circ h$. The homomorphism h does not need to be unique.

Above, we let $A = C$ and $g = \mathrm{id}$.

Example 9.2 (Projective modules)**.** All free modules are projective modules (see the beginning of the section). If F is free and $F = P \oplus Q$, then P and Q are projective. (The module F does not have to be free: it could be projective. Submodules of free and projective modules do not need to be projective.) A module is projective if and only if it is a direct summand of a free module.

Example 9.3 (Non–free projective modules)**.** Take $R = \mathbf{Z}/6\mathbf{Z}$. We know $\mathbf{Z}/3\mathbf{Z} \oplus \mathbf{Z}/2\mathbf{Z}$ is free of rank 1, so $\mathbf{Z}/3\mathbf{Z}$ is projective. But R has zero divisors (if a ring splits as a direct sum, it inevitably has zero divisors).

Example 9.4 (Non–free projective modules over an integral domain)**.** Consider the Möbius strip. Project the Möbius strip onto the circle S^1 so that the preimage of a point is a line on the Möbius strip. We consider an open Möbius strip, so the fibre at each point is an open unit interval which we identify with a copy of $\mathbf{R}$ as a topological space. Let R be continuous functions on S^1 and let the module M be sections of the Möbius strip. One can see that M is an R–module by pointwise multiplication. One can think of R as periodic functions with $f(\tau) = f(\tau + 2\pi)$, so M is functions $f(\tau) = -f(\tau + 2\pi)$. The module M is not a rank-one free R–module: you cannot take a nonzero section of it: you cannot take a nonzero–everywhere map to M because of the sign change. However, $M \oplus M = R \oplus R$, so M is projective. There are two ways of seeing this: one maps $M \oplus M$ to $R \oplus R$ by

$$(f, g) \longmapsto \left(\cos\frac{\tau}{2} f(\tau) + \sin\frac{\tau}{2} g(\tau), -\sin\frac{\tau}{2} f(\tau) + \cos\frac{\tau}{2} g(\tau)\right),$$

and this takes periodic functions to antiperiodic functions and vice versa. This translation works both ways, so we get the isomorphism. There is also a geometric interpretation. Consider S^1 and the Möbius strip over it. Take the Möbius strip's zero section and take the normal bundle to this zero section in $\mathbf{R}^3$. The normal bundles form another copy of the Möbius strip. The direct sum of Möbius strips is the normal bundle of S^1 in the centre of the first Möbius strip. The normal bundle of S^1 is the trivial bundle so corresponds to $R \oplus R$.

Example 9.5. Let $R = \mathbf{Z}\left[\sqrt{-5}\right]$. Let M be the nonprincipal ideal $(2, 1 + \sqrt{-5})$. We want to show that M is projective (it is not free since it is not a principal ideal). We will show that $M \oplus M \cong R \oplus R$. Consider the exact sequence

$$\begin{tikzcd} R \oplus R \arrow[r, "\varphi"] & M \arrow[r] & \{0\} \end{tikzcd}$$

where $(a, b) \longmapsto (2a + \sqrt{-5}b)$. This splits as there is a map given by

$$\psi : \left(-x, \frac{1 - \sqrt{-5}}{2} x\right) \longleftarrow\!\shortmid x \in M$$

in the opposite direction. This map is well defined. So $R \oplus R$ is a direct sum of $\operatorname{im}_\psi(M)$ with $\ker\varphi$, namely

$$\left(\frac{\sqrt{-5} - 1}{2} y,\ y\right),\ y \in M.$$

So $R \oplus R \cong M \oplus M$, so M is projective.

In Example 9.4 and Example 9.5, we have $M \oplus M \cong R \oplus R$: they are **line bundles**. Roughly, a line bundle is a map from one space to another such that the fibre at each point is a one–dimensional vector space. Clearly, the Möbius strip is a line bundle over S^1, but what about the second example? We have seen that $\operatorname{Spec}\left(\mathbf{Z}\left[\sqrt{-5}\right]\right)$ is a topological space; in algebraic geometry, any module over $\mathbf{Z}\left[\sqrt{-5}\right]$ can be used to provide a sheaf over

Spec $\left(\mathbf{Z}\left[\sqrt{-5}\right]\right)$. The module $(2, 1+\sqrt{-5})$ gives a sheaf that is more-or-less a line bundle. Projective modules are geometrically reminiscent of vector bundles.

If P is projective then we can find another projective module Q such that $P \oplus Q$ is free. Can one choose Q to be free? Yes.

Proof (Eilenberg(–Mazur) swindle).—The module

$$(P \oplus Q) \oplus (P \oplus Q) \oplus (P \oplus Q) \oplus \cdots$$

is free. Also,

$$F = (Q \oplus P) \oplus (Q \oplus P) \oplus (Q \oplus P) \oplus \cdots$$

is free. So we have $P \oplus F \cong F$. □

Can we find Q free where Q is finite–dimensional and $P \oplus Q$ is free? Sometimes.

Suppose M is the Möbius strip. Every time you go around the circle you get a "factor of -1," and the same thing will happen if you add a finite number of free modules to M.

Suppose R is continuous functions on S^2 and M is sections of the tangent bundle: tangent vector fields on S^2. The module structure is given by pointwise multiplication. We want to show that $M \oplus R \cong R \oplus R \oplus R$ and M is not free. The module M is the tangent bundle of S^2 and R is the normal bundle of S^2. The normal bundle is isomorphic to a one-dimensional trivial bundle. The direct sum of the tangent space at x with the normal space at x is canonically isomorphic to three copies of the trivial bundle. (This works for any smooth hypersurface in any n–dimensional space.) The second part follows from the hairy ball theorem. (This applies to any compact manifold with nonzero Euler characteristic.)

10 Tensor products of abelian groups

The direct sum of abelian groups is a coproduct: if A, B, and C are abelian groups, there exists a unique map from $A \oplus B$ to C making the following diagram commute.

$$\begin{tikzcd}
 & C & \\
A \arrow[r] \arrow[ur] & A \oplus B \arrow[u, dashed] & B \arrow[l] \arrow[ul]
\end{tikzcd}$$

The product of abelian groups has the dual universal property:

$$\begin{tikzcd}
 & C \arrow[dl] \arrow[d, dashed] \arrow[dr] & \\
A & A \times B \arrow[l] \arrow[r] & B
\end{tikzcd}$$

In abelian groups, $A \oplus B \cong A \times B$. The idea behind the tensor product is that bilinear maps $A \times B \longrightarrow C$ should be the same as linear maps $A \otimes B \longrightarrow C$. There is a map from $A \otimes B$ to itself, so we have a universal bilinear map $A \times B \longrightarrow A \otimes B$:

$$\begin{tikzcd}
A \times B \arrow[r, "\text{bilinear}"] \arrow[dr, "\text{bilinear}"'] & A \otimes B \arrow[d, dashed, "\text{linear}"] \\
& C
\end{tikzcd}$$

Let's show that the tensor product exists. We have a map $A \times B \longrightarrow A \otimes B$ where we write $a \times b \longmapsto a \otimes b$. By the universal property, $A \otimes B$ is free generated by elements $a \otimes b$ for $a \in A$ and $b \in B$ modulo relations corresponding to the bilinearity of the map: namely $(a_1 + a_2) \otimes b = a_1 \otimes b + a_2 \otimes b$ and $a \otimes (b_1 + b_2) = a \otimes b_1 + a \otimes b_2$. This construction is universal (we have added the minimum number of relations to force the map to be bilinear):

$$\begin{tikzcd}
A \times B \arrow[r, "\text{bilinear}"] \arrow[dr, "f"'] & A \otimes B \arrow[d, dashed] \\
& C
\end{tikzcd}$$

We have $a \otimes b \longmapsto f(a \times b)$. So $\otimes$ exists.

This definition makes imagining the structure of $A \otimes B$ difficult. So let's forget about this definition and use the fact that bilinear maps $A \times B \longrightarrow C$ are the same as linear maps $A \otimes B \longrightarrow C$.

In particular, bilinear maps $\mathbf{Z} \times B \longrightarrow C$ are the same as linear maps $B \longrightarrow C$. So $\mathbf{Z} \otimes B \cong B$.

Also, bilinear maps $(A_1 \oplus A_2) \times B \longrightarrow C$ are the same as pairs of bilinear maps $A_1 \times B \longrightarrow C$ and $A_2 \times B \longrightarrow C$. So $(A_1 \oplus A_2) \otimes B \cong A_1 \otimes B \oplus A_2 \otimes B$. So, for instance,

$$\mathbf{Z}^2 \otimes \mathbf{Z}^3 = (\mathbf{Z} \oplus \mathbf{Z}) \otimes (\mathbf{Z} \oplus \mathbf{Z} \oplus \mathbf{Z}) = \mathbf{Z}^6.$$

Tensor products of free abelian groups are easy to compute.

Suppose we have an exact sequence

$$\{0\} \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow \{0\}$$

and we consider

$$\{0\} \longrightarrow A \otimes Y \longrightarrow B \otimes Y \longrightarrow C \otimes Y \longrightarrow \{0\}.$$

This is not an exact sequence, though

$$A \otimes Y \longrightarrow B \otimes Y \longrightarrow C \otimes Y \longrightarrow \{0\}$$

is exact.

Proof.—We know $\operatorname{Mor}(X \otimes Y, Z)$ can be identified with bilinear maps from $X \times Y$ to Z in abelian groups, which can be identified with $\operatorname{Mor}(X, \operatorname{Hom}(Y, Z))$ in abelian groups. So the functor taking X to $X \otimes Y$ is left adjoint to the functor taking Z to $\operatorname{Hom}(Y, Z)$. Left adjoints preserve colimits. (This is a good exercise.) Quotients are a special case of colimits: suppose we have

$$\begin{tikzcd} A \arrow[r, "f"] & B \arrow[r] & C \arrow[r] & \{0\} \end{tikzcd}$$

so $C = B/\operatorname{im} A$. One can think of C as the colimit of 0 and f:

$$\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "0"'] & B \arrow[r] & C \end{tikzcd}$$

(a coequalizer). So $\otimes Y$ preserves right exactness (left adjoints preserve right exactness and right adjoints preserve left exactness). □

Example 10.1 (Calculating tensor products)**.** Suppose we want to compute $\mathbf{Z}/2\mathbf{Z} \otimes \mathbf{Z}/2\mathbf{Z}$. Then we have the exact sequence

$$\begin{tikzcd} \mathbf{Z} \arrow[r, "\times 2"] & \mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\} \end{tikzcd}$$

so tensoring by $\mathbf{Z}/2\mathbf{Z}$

$$\begin{tikzcd} \mathbf{Z}/2\mathbf{Z} \arrow[r, "\times 2"] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \otimes \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\} \end{tikzcd}$$

is exact. We know that $\operatorname{im}_{\times 2}(\mathbf{Z}/2\mathbf{Z}) = \{0\}$, so

$$\mathbf{Z}/2\mathbf{Z} \otimes \mathbf{Z}/2\mathbf{Z} \cong \mathbf{Z}/2\mathbf{Z}.$$

Suppose we want to compute $\mathbf{Z}/2\mathbf{Z} \otimes \mathbf{Z}/3\mathbf{Z}$. Then we have

$$\begin{tikzcd} \mathbf{Z} \arrow[r, "\times 2"] & \mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\} \end{tikzcd}$$

so

$$\begin{tikzcd} \mathbf{Z}/3\mathbf{Z} \arrow[r, "\times 2"] & \mathbf{Z}/3\mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \otimes \mathbf{Z}/3\mathbf{Z} \arrow[r] & \{0\} \end{tikzcd}$$

The ×2 map is an isomorphism, so $\mathbf{Z}/2\mathbf{Z} \otimes \mathbf{Z}/3\mathbf{Z} \cong (\mathbf{Z}/3\mathbf{Z})/(\mathbf{Z}/3\mathbf{Z})$, so

$$\mathbf{Z}/2\mathbf{Z} \otimes \mathbf{Z}/3\mathbf{Z} \cong \{0\}.$$

Example 10.2 ($\mathbf{Z}/m\mathbf{Z} \otimes \mathbf{Z}/n\mathbf{Z}$)**.** We have the exact sequence

$$\begin{tikzcd} \mathbf{Z} \arrow[r, "\times m"] & \mathbf{Z} \arrow[r] & \mathbf{Z}/m\mathbf{Z} \arrow[r] & \{0\} \end{tikzcd}$$

and tensoring by $\mathbf{Z}/n\mathbf{Z}$ we get

$$\begin{tikzcd}\mathbf{Z}/n\mathbf{Z} \arrow[r, "\times m"] & \mathbf{Z}/n\mathbf{Z} \arrow[r] & \mathbf{Z}/m\mathbf{Z} \otimes \mathbf{Z}/n\mathbf{Z} \arrow[r] & \{0\}\end{tikzcd}$$

So $\mathbf{Z}/m\mathbf{Z} \otimes \mathbf{Z}/n\mathbf{Z}$ is $(\mathbf{Z}/n\mathbf{Z})/(\mathrm{im}_{\times m}(\mathbf{Z}/n\mathbf{Z}))$, so

$$\mathbf{Z}/m\mathbf{Z} \otimes \mathbf{Z}/n\mathbf{Z} \cong \mathbf{Z}/\gcd(n,m)\mathbf{Z}.$$

We have seen that $\otimes Y$ preserves right exactness, but it also preserves (even infinite) direct sums: one can show this by seeing that direct sums are colimits or directly. The functor $\mathrm{Hom}(Y, *)$ is right adjoint, so it preserves left exactness and products (these are limits). (See Section 9.)

Any finitely generated abelian group is a direct sum of cyclic groups, so working out their tensor products is straightforward:

$$(\mathbf{Z} \oplus \mathbf{Z} \oplus \mathbf{Z}/6\mathbf{Z}) \otimes (\mathbf{Z} \oplus \mathbf{Z}/15\mathbf{Z}) = \mathbf{Z} \oplus \mathbf{Z} \oplus \mathbf{Z}/6\mathbf{Z} \oplus \mathbf{Z}/15\mathbf{Z} \oplus \mathbf{Z}/15\mathbf{Z} \oplus \mathbf{Z}/3\mathbf{Z}.$$

Example 10.3 (Tensor products of not finitely generated abelian groups)**.** Suppose we want to compute $\mathbf{Z}/n\mathbf{Z} \otimes \mathbf{Q}$. Then

$$\begin{tikzcd}\mathbf{Z} \arrow[r, "\times n"] & \mathbf{Z} \arrow[r] & \mathbf{Z}/n\mathbf{Z} \arrow[r] & \{0\}\end{tikzcd}$$

is exact, so tensoring by $\mathbf{Q}$ we get

$$\begin{tikzcd}\mathbf{Q} \arrow[r, "\times n"] & \mathbf{Q} \arrow[r] & \mathbf{Z}/n\mathbf{Z} \otimes \mathbf{Q} \arrow[r] & \{0\}\end{tikzcd}$$

and the $\times n$ map is an isomorphism, so

$$\mathbf{Z}/n\mathbf{Z} \otimes \mathbf{Q} \cong \{0\}.$$

Example 10.4 (Tensor products where neither group is finitely generated)**.** Suppose we want to compute $\mathbf{Q} \otimes \mathbf{Q}$. Recall that $\otimes \mathbf{Q}$ commutes with colimits. We see that $\mathbf{Q}$ is a union of

$$\mathbf{Z} \subset \frac{1}{2}\mathbf{Z} \subset \frac{1}{6}\mathbf{Z} \subset \frac{1}{24}\mathbf{Z} \subset \cdots$$

so $\mathbf{Q}$ is a colimit

$$\begin{tikzcd}\mathbf{Z} \arrow[r, "\times 2"] & \mathbf{Z} \arrow[r, "\times 3"] & \mathbf{Z} \arrow[r, "\times 4"] & \cdots .\end{tikzcd}$$

Tensoring by $\mathbf{Q}$ gives

$$\begin{tikzcd}\mathbf{Q} \arrow[r, "\times 2"] & \mathbf{Q} \arrow[r, "\times 3"] & \mathbf{Q} \arrow[r, "\times 4"] & \cdots ,\end{tikzcd}$$

and all of these arrows are isomorphisms. So the colimit of this is $\mathbf{Q}$, so

$$\mathbf{Q} \otimes \mathbf{Q} \cong \mathbf{Q}.$$

Example 10.5 (Mistake)**.** Suppose we want to calculate $\mathbf{Q} \otimes (\mathbf{Z}/2\mathbf{Z})$. We write $\mathbf{Q}$ as a colimit as before, and tensor by $\mathbf{Z}/2\mathbf{Z}$ to obtain an exact sequence with $\mathbf{Z}/2\mathbf{Z}$ everywhere. Take its colimit, and obtain $\mathbf{Q} \otimes (\mathbf{Z}/2\mathbf{Z}) \cong \mathbf{Z}/2\mathbf{Z}$. This is wrong. When we get the sequence

$$\mathbf{Z}/2\mathbf{Z} \xrightarrow{\times 2} \mathbf{Z}/2\mathbf{Z} \xrightarrow{\times 3} \mathbf{Z}/2\mathbf{Z} \xrightarrow{\times 4} \cdots,$$

we cannot say that its colimit is $\mathbf{Z}/2\mathbf{Z}$ since these arrows are not isomorphisms. These arrows mean nothing more than the 0 map, so the colimit (and tensor product) is $\{0\}$. Also, once we tensor with $\mathbf{Z}/2\mathbf{Z}$, the arrows are not injective maps, so we do not have an increasing union of $\mathbf{Z}/2\mathbf{Z}$.

Exercise 10.6. Compute

1. $\mathbf{Q} \otimes \mathbf{Q}/\mathbf{Z}$;
2. $\mathbf{Z}_p \otimes \mathbf{Q}$.

11 Tensor products of modules

Suppose R is commutative and M and N are two R–modules. We want to define their tensor product $M \otimes N$ such that if A is another R–module we can identify bilinear maps from $M \times N$ to A with a unique linear map from $M \otimes N$ to A:

$$\begin{tikzcd}
M \times N \arrow[r, "\text{bilinear}"] \arrow[dr, "\text{bilinear}"'] & M \otimes N \arrow[d, dashed, "\text{linear}"] \\
& A
\end{tikzcd}$$

As when $R = \mathbf{Z}$, $M \otimes N$ is generated by elements $m \otimes n$ modulo relations $(m_1 + m_2) \otimes n = m_1 \otimes n + m_2 \otimes n$ and $m \otimes (n_1 + n_2) = m \otimes n_1 + m \otimes n_2$. We also need other relations to force the map to be R–linear: namely $(rm) \otimes n = r(m \otimes n)$ and $m \otimes (rn) = r(m \otimes n)$. Other than these extra relations, the tensor product over $\mathbf{Z}$–modules (abelian groups) is mostly similar to the tensor product over modules of arbitrary commutative rings.

Example 11.1. Suppose I is an ideal of R and we want to compute $R/I \otimes_R M$. Take the exact sequence

$$\{0\} \longrightarrow I \longrightarrow R \longrightarrow R/I \longrightarrow \{0\}.$$

The functor $\cdot \otimes_R M$ is right exact, so the sequence

$$I \otimes_R M \longrightarrow M \longrightarrow R/I \otimes_R M \longrightarrow \{0\}$$

is exact. But $I \otimes_R M \longrightarrow M$ is given by $a \otimes m \longmapsto am$, so the image of $I \otimes_R M$ under the first arrow is IM. Hence

$$R/I \otimes_R M \cong M/IM.$$

We will make some attempts at defining the tensor product over noncommutative rings. First, we will try to define it the same way we did for commutative rings (if M and N are R–modules, take the free group generated by elements $m \otimes n$ modulo relations forcing the map to be R–linear). Suppose M and N are left R–modules. Then the relations imply that $(rsm) \otimes n = (sm) \otimes (rn) = m \otimes (srn)$ and $(rsm) \otimes n = m \otimes (rsn)$, so this tensor product forces elements of R to act in a commutative way on these modules. So this definition doesn't really work.

Now let M be a right R–module and N a left R–module. Then we take the relation

$$(mr) \otimes n = m \otimes (rn).$$

One can define the tensor product by adding this relation instead; however, there's a problem: $M \otimes_R N$ is an abelian group, not an R–module. Suppose you want $M \otimes N$ to be an R–module. Let M and N be R–bimodules. Again, we use the relation above, and everything works out well. We define the left action of R on the bimodule $M \otimes N$ by $r(m \otimes n) := (rm) \otimes n$ and the right action by $(m \otimes n)r := m \otimes (nr)$.

Example 11.2 (Differential geometry)**.** Suppose V and W are vector spaces over a field k with bases $(v_1, v_2, \ldots)$ and $(w_1, w_2, \ldots)$. Then $V \otimes_k W$ is a vector space with basis $v_i \otimes w_j$. In particular,

$$\dim(V \otimes_k W) = (\dim V)(\dim W).$$

In differential geometry, one uses this idea with a (co)tangent space and its dual. Suppose V^* has a basis v_i^*. Then the elements of $V \otimes V \otimes V \otimes V^*$, for instance, might be

$$\sum R^a bcd\ (v_a^* \otimes v_b \otimes v_c \otimes v_d)$$

which a differential geometer would write $R^a bcd$. One thinks of this as a tensor field; it's a function on a manifold.

Example 11.3 (Coproducts of commutative rings)**.** Suppose R and S are commutative rings. The tensor product $R \otimes_{\mathbf{Z}} S$ is the coproduct:

$$\begin{tikzcd}
R \times S \arrow[r] \arrow[d] & S \arrow[d] \arrow[rdd] & \\
R \arrow[r] \arrow[rrd] & R \otimes_{\mathbf{Z}} S \arrow[rd, dashed] & \\
 & & A
\end{tikzcd}$$

We define $(r_1 \otimes s_1)(r_2 \otimes s_2) := r_1r_2 \otimes s_1s_2$. Using the universal property of the tensor product one can confirm that this is well defined; then one can check that $\otimes_{\mathbf{Z}}$ is indeed the coproduct.

Suppose R and S are T–algebras. We want a universal T–algebra generated by R and S, and this T–algebra is $R \otimes_T S$:

$$\begin{tikzcd}
T \arrow[r] \arrow[d] & S \arrow[d] \\
R \arrow[r] & R \otimes_T S
\end{tikzcd}$$

Rings correspond to affine schemes, so an algebraic geometer might consider the following diagram of spectra:

$$\begin{tikzcd}
\operatorname{Spec} T & \operatorname{Spec} S \arrow[l] \\
\operatorname{Spec} R \arrow[u] & \operatorname{Spec}(R \otimes_T S) \arrow[l] \arrow[u]
\end{tikzcd}$$

The object $\operatorname{Spec}(R \otimes_T S)$ is the fibered product or pullback of the affine schemes $\operatorname{Spec} R$ and $\operatorname{Spec} S$.

Suppose k is a field and we take the rings $k[x]$ and $k[y]$. One relates k to a point and two affine lines to $k[x]$ and $k[y]$ mapping to that point (as per the construction above). What's the product of these two spaces (affine lines)? It's the affine plane. This corresponds to the tensor product $k[x] \otimes k[y]$. By considering the basis of each ring, one determines that $k[x] \otimes k[y] = k[x, y]$, which indeed corresponds to the affine plane. So we get the diagram

$$\begin{tikzcd}
k \arrow[r] \arrow[d] & k[x] \arrow[d] \\
k[y] \arrow[r] & k[x] \otimes k[y] = k[x, y]
\end{tikzcd}$$

One should be careful: the product of the topological spaces corresponding to the affine line is not the topological space corresponding to the affine plane. Recall that $\operatorname{Spec} k[x]$ (if k is algebraically closed) is a copy of k corresponding to ideals of the form $(x - \alpha)$ union a generic point corresponding to (0). And $\operatorname{Spec} k[y]$ is similar. So the product of $\operatorname{Spec} k[x]$ and $\operatorname{Spec} k[y]$ corresponds to a subset of the affine plane as points (α, β) correspond to ideals $(x - \alpha, y - \beta)$, et cetera. The affine plane $k[x, y]$ has prime ideals other than this subset: for example, $(y^2 - x^3 - x)$.

Example 11.4 (Linear representations of finite groups)**.** Let V be a vector space over $\mathbf{C}$ with $\dim V < \infty$ and let G be a finite group acting on V. Suppose W is a vector space with the same described properties as V. The direct sum $V \oplus W$ is acted on by G via $g(v \oplus w) \longmapsto g(v) + g(w)$. We define a tensor product $V \otimes W$ on which G acts via $g(v \otimes w) \longmapsto g(v)g(w)$. So one can "add" and "multiply" two representations of G. These "additions" and "multiplications" behave like they should in a ring: $(U \oplus V) \otimes W = (U \otimes W) \oplus (V \oplus W)$ and $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$. Is the set of isomorphism classes of (these) representations of G a ring? Not quite. There is no "subtraction." But, as we did with the Burnside ring (as one does with a Grothendieck group), we force a subtraction and get the representation ring of G.

Example 11.5 (Representation rings)**.** Suppose $G = \mathbf{Z}/n\mathbf{Z} = \langle g \rangle$. Suppose V and W are vector spaces as described in Example 11.4 with dimension 1 and

$$g(v) = e^{2\pi i a/n}, v \in V,$$
$$g(w) = e^{2\pi i b/n}, w \in W.$$

The element g acts on elements of $V \otimes W$ by $e^{2\pi i(a+b)/n.}$

Suppose $G = \mathrm{S}_3$. The group G acts on $U = \mathbf{C}$ trivially, on $V = \mathbf{C}$ according to the sign representation, and on $W = \mathbf{C}^2$ as symmetries of a triangle embedded in $\mathbf{C}^2$. One finds that

$$V \otimes V = U,$$
$$W \otimes W = W \oplus U \oplus V.$$

So U, V, and W are a basis for the three-dimensional ring over $\mathbf{C}$.

Example 11.6 (Tensor products of fields)**.** (Exercise: Compute $\mathbf{Q}_2 \otimes_{\mathbf{Q}} \mathbf{Q}[i]$.) Suppose we want to compute $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C}$. It has a basis $(1 \otimes 1, 1 \otimes i, i \otimes 1, i \otimes i)$. This tensor product splits as a direct sum $\mathbf{C} \oplus \mathbf{C}$. One finds that

$$\kappa_1 = \frac{1 \otimes 1 + i \otimes i}{2} \quad \text{and} \quad \kappa_2 = \frac{1 \otimes 1 - i \otimes i}{2}$$

are idempotents. So the tensor product is isomorphic to $\mathbf{C}\kappa_1 \oplus \mathbf{C}\kappa_2$.

12 Duality and injective modules

Suppose V is a vector space over a field k. The dual space is given by

$$V^* := \mathrm{Hom}_k(V, k).$$

There is a natural map from V to its double dual. Elements of the double dual are maps from V^* to k, and if $v \in V$ and $f \in V^*$ where $f : v \longrightarrow k$, one maps f to $f(v)$. This is linear in f, so the map $f \longmapsto f(v)$ gives a linear map from V^* to k. This is the natural map.

If V is finite dimensional, then V is isomorphic to its dual. One chooses a basis for V which identifies V with an n-dimensional vector space, and the dual is also an n-dimensional vector space. So they are isomorphic, but this isomorphism is not natural. The vector space V is naturally isomorphic to its double dual. The map that was described above is an isomorphism.

If V is infinite dimensional, then V might not be isomorphic to its dual or double dual. Suppose V has a basis $(v_1, v_2, v_3, \ldots)$. Then elements of V are of the form $\sum a_i v_i$ where almost all of the a_i are 0. Elements of V^* are sequences b_i where we do not need almost all of the b_i to be 0. An element $v \in V$ is mapped to $\sum a_i b_i$, but the space of all such b_i has uncountable dimension. So V^* is bigger than V in general. One can endow V with a topology.

Example 12.1 (L^p spaces)**.** Suppose V is defined over $k = \mathbf{C}$. We define spaces L^p to be sequences a_i with $a_i \in k$ such that $\sum |a_i|^p < \infty$ where the norm in this vector space is the pth root of that quantity. Here we define L^p over the naturals, but, more generally, one might define it over a measure space. The dual of L^p is L^q where $1 < p < \infty$ and $1/p + 1/q = 1$. If $p = 2$, then L^p is a Hilbert space. Since $q = 2$, this says that the dual of a Hilbert space is a Hilbert space. In fact, we see

$$(L^p)^{**} \cong L^p.$$

So spaces can be isomorphic to their double dual with the right choice of topology.

Example 12.2. Let $V = C_0$ be sequences a_i tending to 0. Then $V^* = L^1$ is sequences b_i with $\sum |b_i| < \infty$. One can pair V with its dual by taking $\sum a_i b_i$ which will converge. But $V^{**} = L^\infty$ is bounded sequences c_i. Then $\sum b_i c_i$ is well defined, and V^{**} is clearly bigger than V. Even defining duals with a topology does not mean the double dual will not be bigger than the original space. The triple dual has to do with the contents on a measure space, and things start to become complicated.

Suppose R is a commutative ring. The dual of an R–module M is given by $M^* := \operatorname{Hom}(M, R)$. If M is free and finitely generated, then duality is similar to that of vector spaces. The result $\operatorname{rank} M = \operatorname{rank} M^*$ still holds, and M^{**} is naturally isomorphic to M.

If P is a finitely generated projective module, then there exists another projective module Q such that $P \oplus Q$ is free of finite rank. We have $(P \oplus Q)^* \cong P^* \oplus Q^*$, so the natural map $P \longrightarrow P^{**}$ is an isomorphism (as it is for free modules).

Suppose $R = \mathbf{Z}$ and $M = \mathbf{Z}/2\mathbf{Z}$. Then $M^* = \operatorname{Hom}_R(M, R) = \{0\}$. The theory seems to fall apart for nonprojective modules.

We define a different dual where

$$M^* := \operatorname{Hom}_{\mathbf{Z}}(M, \mathbf{Q}/\mathbf{Z}).$$

The module $\mathbf{Q}/\mathbf{Z}$ is a **dualizing module**. We have a natural map $M \longrightarrow M^{**}$, and this is an isomorphism if M is a finite abelian group. (For free abelian groups, the dualizing module is $\mathbf{Z}$.)

Proof.—Notice that $(A \oplus B)^* \cong A^* \oplus B^*$ and M is a direct sum of cyclic groups. So it is enough to check that M is isomorphic to its double dual when M is cyclic. We know that $(\mathbf{Z}/n\mathbf{Z})^* = \mathrm{Hom}_{\mathbf{Z}}(\mathbf{Z}/n\mathbf{Z}, \mathbf{Q}/\mathbf{Z})$. Suppose g is a generator of $\mathbf{Z}/n\mathbf{Z}$. Then the image of g is an element of order dividing n in $\mathbf{Q}/\mathbf{Z}$, and there are n such elements: 0, $1/n$, $2/n$, et cetera. So there are n homomorphisms from $\mathbf{Z}/n\mathbf{Z}$ to $\mathbf{Q}/\mathbf{Z}$ corresponding to these elements. One sees that $\mathrm{Hom}(\mathbf{Z}/n\mathbf{Z}, \mathbf{Q}/\mathbf{Z}) \cong \mathbf{Z}/n\mathbf{Z}$, but this isomorphism is not natural: it depends on the choice of generator. However, $\mathbf{Z}/n\mathbf{Z}$ has the same size as its dual, so we have a natural isomorphism $\mathbf{Z}/n\mathbf{Z} \longrightarrow (\mathbf{Z}/n\mathbf{Z})^{**}$. So any finite abelian group is isomorphic to its double dual. (That it is isomorphic to its dual is an "accident.") □

Why choose $\mathbf{Q}/\mathbf{Z}$ as the dualizing module? We needed the elements of order dividing n in the module to form a cyclic subgroup of order n. We might have used S^1, $\mathbf{R}/\mathbf{Z}$, or even $\mathbf{C}^\times$.

Example 12.3 (Dual of $(\mathbf{Z}/8\mathbf{Z})^\times$)**.** (Notice that the usual notation for the unit ring conflicts with notation for duals. If context isn't enough to clarify things we use $(\,\cdot\,)^\times$ for the unit ring.) Take $\mathbf{C}^\times$ to be the dualizing module. There are 4 homomorphisms from $(\mathbf{Z}/8\mathbf{Z})^\times$ to $\mathbf{C}^\times$:

	1	3	5	7
χ_0	1	1	1	1
χ_1	1	1	-1	-1
χ_2	1	-1	1	-1
χ_3	1	-1	-1	1

So

$$\left((\mathbf{Z}/8\mathbf{Z})^\times\right)^* = \{\chi_0, \chi_1, \chi_2, \chi_3\}.$$

These are examples of **Dirichlet characters**: a Dirichlet character is a map

$$\chi : (\mathbf{Z}/n\mathbf{Z})^\times \longrightarrow \mathbf{C}^\times.$$

A Dirichlet L–series, for example, is given by

$$L(s) = \sum_{n \geq 1} \frac{\chi(n)}{n^s}.$$

These were introduced by Dirichlet in his proof of Dirichlet's theorem[1].

One can do Fourier analysis on finite abelian groups M with Dirichlet characters. One notices that Dirichlet characters χ_i and χ_j are orthogonal if $i \neq j$. We define an inner

[1] See https://arxiv.org/abs/0808.1408.

product on complex functions on M by

$$(f,g) := \sum_{g\in M} f(m)\overline{g(m)}.$$

So $(\chi_i,\chi_j) = 0$ if $i \neq j$ (and $|M|$ if $i = j$). We have

$$\sum_{g\in M} \chi_i(g)\overline{\chi_j(g)} = \sum_{g\in M} (\chi_i/\chi_j)(g).$$

So we want to show that $\sum_{g\in M} \chi(g) = 0$ if χ is not the unit character (so $\chi(h) \neq 1$ for some h). We have

$$\begin{aligned}\sum_{g\in M} \chi(g) &= \sum_{g\in M} \chi(gh)\\ &= \sum_{g\in M} \chi(g)\chi(h)\\ &= \chi(h)\sum_{g\in M} \chi(g).\end{aligned}$$

So either $\chi(h) = 1$ or $\sum_{g\in M} \chi(g) = 0$ as desired. So the set of characters of a finite abelian group is an orthogonal basis for the inner product space of functions on M. So any function can be written as

$$f = \text{constant} \times \sum f(\chi_i)\chi_i,$$

and this is a “Fourier series.” Suppose $G = \mathbf{Z}$. Then $\mathrm{Hom}(G,S^1)$ is S^1. If $G = S^1$, then $\mathrm{Hom}(G,S^1)$ is $\mathbf{Z}$. It turns out that if G is locally compact it is isomorphic to its double dual in this sense. If $G = \mathbf{R}$ then its dual is G. The expansion above in the case when the group is S^1 turns out to be the Fourier series of f. If the group is $\mathbf{R}$, the expansion (where the summation is replaced by an integral) turns out to be the Fourier transform.

Definition 12.4. Suppose X and Y are left R–modules, $f : X \longrightarrow Y$ is an injective module homomorphism, and $g : X \longrightarrow I$ is a module homomorphism. The module I is an **injective module** if there exists a module homomorphism $h : Y \longrightarrow I$ such that $h \circ f = g$. So

$$\begin{tikzcd}
\{0\} \arrow[r] & X \arrow[r, "f"] \arrow[d, "g"'] & Y \arrow[dl, dashed, "h"] \\
 & I &
\end{tikzcd}$$

commutes. This is the category-theoretic dual of a projective module. So we can extend maps from a submodule X to I to the entire module Y to I.

Example 12.5 (Injective modules over $R = \mathbf{Z}$)**.** Is $I = \mathbf{Z}$ injective? There is an injective map from $\mathbf{Z}$ to itself given by multiplication by 2. Suppose we take the identity map on $\mathbf{Z}$. We cannot extend this map to a map from $2\mathbf{Z}$ to $\mathbf{Z}$. So $\mathbf{Z}$ is not injective. The module $I = \mathbf{Z}/2\mathbf{Z}$ also is not injective by the same reasoning. The only injective module over $\mathbf{Z}$ is the zero module.

Definition 12.6. Suppose n is a nonzero integer, M is a module, and $m \in M$. The module M is **divisible** if there exists an element $r \in M$ such that $rn = m$.

If M is an injective module, then M is divisible: consider

$$\begin{tikzcd}
 & & M \\
\{0\} \arrow[r] & \mathbf{Z} \arrow[r] \arrow[ur] & \mathbf{Z}/n\mathbf{Z} \arrow[u]
\end{tikzcd}$$

If a $\mathbf{Z}$–module is divisible, it is injective (this is true for PIDs, but not in general). So $\mathbf{Q}$ is injective.

Proof.—Suppose we have

$$\begin{tikzcd}
 & & \mathbf{Q} \\
\{0\} \arrow[r] & C \arrow[r] \arrow[ur, "f"] & B
\end{tikzcd}$$

and we want to extend f to B. Pick an element $b \in B \setminus C$. Suppose $nb \notin C$ for any $n \neq 0$. Then this extension can be given by $f(b) = 0$. Suppose $nb \in C$ for some $n > 0$. Pick the smallest such n. One extends f to b by writing $f(b) = r$ where $nr = f(nb)$. One uses the axiom of choice or Zorn's lemma so that we can choose a maximal extension of f that must encompass B. □

So $\mathbf{Q}$ is injective. The module $\mathbf{Q}/\mathbf{Z}$ is divisible, so it is also injective. One can write $\mathbf{Q}/\mathbf{Z} = \bigoplus M_p$ where M_p consists of elements of order a power of p. The modules M_p also are injective.

There are enough injective modules: any $\mathbf{Z}$–module is a sumbodule of an injective module. Suppose M is a module. We want to find an injective module containing M. Pick $m \in M$ nonzero and map m to an element $q \in \mathbf{Q}/\mathbf{Z}$ with $q \neq 0$ where q has the same order as m. We extend this map to M as $\mathbf{Q}/\mathbf{Z}$ is injective. Take a copy of $\mathbf{Q}/\mathbf{Z}$ for every nonzero $m \in M$. So we map M to $\prod(\mathbf{Q}/\mathbf{Z})$, and we map each element of M to something nonzero in the corresponding copy of $\mathbf{Q}/\mathbf{Z}$, so this map is injective and the module $\prod(\mathbf{Q}/\mathbf{Z})$ is injective. One can usually find smaller modules with this desired property.

Example 12.7 (Smaller injective modules)**.** One can embed $\mathbf{Z}$ in $\mathbf{Q}$. One can embed $\mathbf{Z}/p\mathbf{Z}$ in M_p. This module is an increasing union of

$$\mathbf{Z}/p\mathbf{Z} \subset \mathbf{Z}/p^2\mathbf{Z} \subset \mathbf{Z}/p^3\mathbf{Z} \subset \cdots$$

and is the smallest injective module containing $\mathbf{Z}/p\mathbf{Z}$. The same goes for $\mathbf{Z}/p^f\mathbf{Z}$. These are the injective envelopes of the corresponding modules.

Definition 12.8. The **injective envelope** or **injective hull** of a module is the smallest injective module containing that module and the largest essential extension of it.

There is no unique "minimal" free module mapping onto a module M in general. For instance, one can write $\mathbf{Z}/5\mathbf{Z}$ as a quotient of a free module in at least two different ways.

What about injective modules over arbitrary rings R? One can lift injective modules from $\mathbf{Z}$ to R. If I is an injective $\mathbf{Z}$–module, then $\mathrm{Hom}_{\mathbf{Z}}(R, I)$ is an injective R–module.

If we want to show something is injective, we need to study maps to it. One needs to understand $\mathrm{Hom}_R(M, \mathrm{Hom}_{\mathbf{Z}}(R, I))$ to show that $\mathrm{Hom}_{\mathbf{Z}}(R, I)$ is injective. We have a natural isomorphism $\mathrm{Hom}_R(M, \mathrm{Hom}_{\mathbf{Z}}(R, I)) \longrightarrow \mathrm{Hom}_{\mathbf{Z}}(M, I)$. The module I is $\mathbf{Z}$–injective, so we can lift homomorphisms from submodules to something to modules to something. So $\mathrm{Hom}_{\mathbf{Z}}(R, I)$ is an injective R–module.

Exercise 12.9. Show that there are enough injective R–modules.

One notices that $\mathrm{Hom}_{\mathbf{Z}}(R, I)$ in the case that $I = \mathbf{Q}/\mathbf{Z}$ is a definition of the dual for a finite abelian group.

13 Colimits and exactness

We will present some examples of limits and the corresponding colimits.

The limit of modules C_i in a discrete category is the universal thing that maps to all C_i: this is given by the direct product of the modules. The colimit of these modules is the universal thing to which all C_i map, and this is the direct sum of the modules. These are **products** and **coproducts** respectively.

Suppose modules M and N both map to a module T. The limit of M and N is $M \times_T N$, consisting of pairs $(m, n) \in M \times N$ such that m and n have the same image in T. This construction is universal: any module mapping to M and N factors through $M \times_T N$. This is called a **pullback**. The corresponding colimit is a **pushout**. Suppose a module T maps to modules M and N. The colimit is given by $(M \oplus N)/(\mathrm{im}\, T)$. This construction is universal in that if M and N map to a module then there is a unique map from the colimit to that module. The tensor product of commutative rings is an example of a pushout.

Suppose $f : M \longrightarrow N$ is a module homomorphism. Then the limit of

$$\begin{tikzcd}
M \arrow[r, shift left, "f"] \arrow[r, shift right, "0"'] & N
\end{tikzcd}$$

is the **kernel** of f: the equalizer of f and the zero map from M to N. It is a module $\ker f$ and a module homomorphism $k : \ker f \longrightarrow M$ such that $f \circ k = 0 \circ k$. It is universal:

$$\begin{tikzcd}
\ker f \arrow[r, "k"] & M \arrow[r, shift left, "f"] \arrow[r, shift right, "0"'] & N \\
T \arrow[u, dashed] \arrow[ur]
\end{tikzcd}$$

The corresponding colimit is the **cokernel** of f: the module $N/(\operatorname{im} M)$ and a homomorphism $g : N \longrightarrow N/(\operatorname{im} M)$ such that $g \circ f = g \circ 0$. It is universal in the same way.

Another example of a colimit is a **direct limit** or injective limit. Suppose we have modules M_i with homomorphisms $M_i \longrightarrow M_{i+1}$. The direct limit is a universal module to which every M_i maps:

$$\begin{tikzcd}
M_0 \arrow[r] \arrow[drr] & M_1 \arrow[r] \arrow[dr] & M_2 \arrow[r] \arrow[d] & M_3 \arrow[r] \arrow[dl] & \cdots \arrow[dll] \\
& & \varinjlim M_i & &
\end{tikzcd}$$

If these homomorphisms are injective then $\varinjlim M_i = \bigcup M_i$. We have seen that the limit of

$$\begin{tikzcd}
\mathbf{Z}/p\mathbf{Z} \arrow[r] & \mathbf{Z}/p^2\mathbf{Z} \arrow[r] & \mathbf{Z}/p^3\mathbf{Z} \arrow[r] & \cdots
\end{tikzcd}$$

is the injective module M_p. The dual of this is given by

$$\begin{tikzcd}
\mathbf{Z}/p\mathbf{Z} & \mathbf{Z}/p^2\mathbf{Z} \arrow[l] & \mathbf{Z}/p^3\mathbf{Z} \arrow[l] & \cdots \arrow[l]
\end{tikzcd},$$

and the limit of the opposite diagram—in this case the p-adic numbers—is called the **inverse limit** or projective limit:

$$\begin{tikzcd}
\mathbf{Z}/p\mathbf{Z} & \mathbf{Z}/p^2\mathbf{Z} \arrow[l] & \mathbf{Z}/p^3\mathbf{Z} \arrow[l] & \mathbf{Z}/p^4\mathbf{Z} \arrow[l] & \cdots \arrow[l] \\
& & \varprojlim \mathbf{Z}/p^i\mathbf{Z} = \mathbf{Z}_p \arrow[ull] \arrow[ul] \arrow[u] \arrow[ur] \arrow[urr] & &
\end{tikzcd}$$

The dual of a direct limit is an inverse limit. The dual of an inverse limit is not necessarily a direct limit. Another example of an inverse limit is the **profinite completion** of $\mathbf{Z}$.

Suppose $\mathbf{Z}/m\mathbf{Z}$ maps to $\mathbf{Z}/n\mathbf{Z}$ if m divides n. Draw all of these maps. The inverse limit of the resulting diagram is the profinite completion of $\mathbf{Z}$, denoted $\widehat{\mathbf{Z}}$. One sees that $\widehat{\mathbf{Z}}$ is given by the submodule of the product of all $\mathbf{Z}/n\mathbf{Z}$ whose components are compatible with the maps. This profinite completion can be described differently. Suppose $m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots$. Then $\mathbf{Z}/m\mathbf{Z} \cong \prod_i \mathbf{Z}/p_i^{\alpha_i}\mathbf{Z}$. One sees that this means that

$$\widehat{\mathbf{Z}} \cong \prod_{p \text{ prime}} \mathbf{Z}_p.$$

One constructs the finite adeles over the integers by $\mathbf{A}_{\mathbf{Z}} = \mathbf{R} \times \widehat{\mathbf{Z}}$. One gets the ring of adeles, then, by

$$\mathbf{A}_{\mathbf{Q}} = \mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{A}_{\mathbf{Z}}.$$

Do colimits or limits preserve exactness? That is, given various exact sequences

$$\begin{tikzcd} \{0\} \arrow[r] & A_i \arrow[r] & B_i \arrow[r] & C_i \arrow[r] & \{0\}, \end{tikzcd}$$

is

$$\begin{tikzcd} \{0\} \arrow[r] & \lim A_i \arrow[r] & \lim B_i \arrow[r] & \lim C_i \arrow[r] & \{0\} \end{tikzcd}$$

exact? Sometimes. It is true if the (co)limit is a product or a direct sum. Limits are left exact (and kernels are limits). Limits preserve limits. Right adjoints also preserve left exactness, and, in fact, limits are special cases of right adjoints. Similarly, colimits preserve colimits. So (co)limits preserve (co)kernels.

Do limits preserve colimits? Not in general: if $\mathcal{I}$ and $\mathcal{J}$ are categories with two objects and limits are taken in Set (the colimit in Set is disjoint union and the limit is product), one sees elementarily that $\operatorname{colim}\lim \neq \lim\operatorname{colim}$.

However, there is always a natural map

$$\operatorname*{colim}_{\mathcal{I}} \lim_{\mathcal{J}} F(I, J) \longrightarrow \lim_{\mathcal{J}} \operatorname*{colim}_{\mathcal{I}} F(I, J).$$

Clearly, there is a map from $F(I, J)$ to $\operatorname{colim}_{\mathcal{I}} F(I, J)$. Then we get a map $\lim_{\mathcal{J}} F(I, J) \longrightarrow \lim_{\mathcal{J}} \operatorname{colim}_{\mathcal{I}} F(I, J)$. Taking $\operatorname{colim}_{\mathcal{I}}$ on the left, we get $\operatorname{colim}_{\mathcal{I}} \lim_{\mathcal{J}} F(I, J)$, and by definition of the colimit, we have the desired map.

Example 13.1 (Pushouts do not preserve exactness of modules)**.** Consider the following three identical exact sequences.

$$\begin{tikzcd}
\{0\} \arrow[r] & \mathbf{Z} \arrow[r] & \mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\} \\
\{0\} \arrow[r] & \mathbf{Z} \arrow[r] \arrow[u, "\times 2"] \arrow[d, "\times 2"] & \mathbf{Z} \arrow[r] \arrow[u, "\times 2"] \arrow[d, "\times 2"] & \mathbf{Z}/2\mathbf{Z} \arrow[r] \arrow[u, "\times 2"] \arrow[d, "\times 2"] & \{0\} \\
\{0\} \arrow[r] & \mathbf{Z} \arrow[r] & \mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\}
\end{tikzcd}$$

This diagram has three pushouts, and the colimits of each (by methods described eariler) give

$$\mathbf{Z} \oplus \mathbf{Z}/2\mathbf{Z} \xrightarrow{\times 2} \mathbf{Z} \oplus \mathbf{Z}/2\mathbf{Z} \longrightarrow \mathbf{Z}/2\mathbf{Z} \oplus \mathbf{Z}/2\mathbf{Z} \longrightarrow \{0\}.$$

The first arrow is not injective, so pushouts do not preserve exact sequences.

Direct limits preserve exactness. The simplest direct limit is one over the positive integers. More generally, one takes a direct limit over a directed set. The key idea is that $\varinjlim M_i = \bigsqcup M_i / \sim$ where if $m \in M_i$ and $n \in M_j$ then $m \sim n$ if the images of m and n are the same in some M_k with $k \geq i$ and $k \geq j$. Any element in the direct limit is represented by some $m_i \in M_i$.

One can show that colimits over $\mathbf{Z}$ (in general, directed sets) are exact. We have exact sequences

$$\begin{tikzcd}
 & & \text{etc.} & & \\
\{0\} \arrow[r] & A_1 \arrow[r] \arrow[u] & B_1 \arrow[r] \arrow[u] & C_1 \arrow[r] \arrow[u] & \{0\} \\
\{0\} \arrow[r] & A_0 \arrow[r] \arrow[u] & B_0 \arrow[r] \arrow[u] & C_0 \arrow[r] \arrow[u] & \{0\}
\end{tikzcd}$$

We know that

$$\varinjlim A_i \longrightarrow \varinjlim B_i \longrightarrow \varinjlim C_i \longrightarrow \{0\}$$

is exact, but is the first arrow injective? Pick some $a \in \varinjlim A_i$ with image 0 in $\varinjlim B_i$. The element a is represented by some $a_i \in A_i$. The image of a_i is some $b_i \in B_i$, which must map to $0 \in B_j$ for some $j \geq i$. So the image of a_i in A_j maps to 0, so is 0. So the map is injective.

Directed sets can be generalized to a filtered category. Colimits over filtered categories preserve exactness.

Definition 13.2. A category $\mathcal{C}$ is a **filtered category** if

1. $\mathcal{C}$ is nonempty;
2. given objects $a, b \in \mathcal{C}$ there exists an object $c \in \mathcal{C}$ such that there are morphisms from a to c and from b to c;
3. given objects $a, b \in \mathcal{C}$ and two morphisms f and g from a to b there exists an object $c \in \mathcal{C}$ such that c is a coequalizer for f and g.

Example 13.3. Consider the following diagram.

$$\begin{tikzcd}
\{0\} \arrow[r] & \mathbf{Z} \arrow[r, "\times 2"] \arrow[d, shift right, "\times 2"'] \arrow[d, shift left, "\times 0"] & \mathbf{Z} \arrow[r] \arrow[d, shift right, "\times 2"'] \arrow[d, shift left, "\times 0"] & \mathbf{Z}/2\mathbf{Z} \arrow[r] \arrow[d, shift right, "\times 2"'] \arrow[d, shift left, "\times 0"] & \{0\} \\
\{0\} \arrow[r] & \mathbf{Z} \arrow[r, "\times 2"] & \mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\}
\end{tikzcd}$$

Taking colimits gives

$$\begin{tikzcd}
\mathbf{Z}/2\mathbf{Z} \arrow[r, "\times 2"] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\},
\end{tikzcd}$$

but the first arrow is not injective. So we need all conditions of a filtered category for colimits to preserve exactness.

14 Limits and exactness

Suppose we have exact sequences

$$\begin{tikzcd}
\{0\} \arrow[r] & A_i \arrow[r] & B_i \arrow[r] & C_i \arrow[r] & \{0\}
\end{tikzcd}$$

where $i \in \mathcal{I}$ is an element of an indexed category. We will take limits over $\mathcal{I}$. Is

$$\begin{tikzcd}
\{0\} \arrow[r] & \lim A_i \arrow[r] & \lim B_i \arrow[r] & \lim C_i \arrow[r] & \{0\}
\end{tikzcd}$$

exact? We have seen that it is always exact if the rightmost arrow is removed.

If the limit is the product, then limits preserve exactness: so

$$\begin{tikzcd}
\{0\} \arrow[r] & \prod A_i \arrow[r] & \prod B_i \arrow[r] & \prod C_i \arrow[r] & \{0\}
\end{tikzcd}$$

is exact. One proves this elementarily assuming the axiom of choice.

Pullbacks do not preserve exactness. Consider the following diagram.

$$\begin{tikzcd}
\{0\} \arrow[r] & \{0\} \arrow[r] \arrow[d] & M \arrow[r] \arrow[d, "\mathrm{id}"] & M \arrow[r] \arrow[d] & \{0\} \\
\{0\} \arrow[r] & M \arrow[r] & M \arrow[r] & \{0\} \arrow[r] & \{0\} \\
\{0\} \arrow[r] & \{0\} \arrow[r] \arrow[u] & M \arrow[r] \arrow[u, "\mathrm{id}"] & M \arrow[r] \arrow[u] & \{0\}
\end{tikzcd}$$

Taking limits, one gets

$$\begin{tikzcd}\{0\} \arrow[r] & \{0\} \arrow[r] & M \arrow[r] & M \times M.\end{tikzcd}$$

The rightmost arrow is not surjective if M is nonzero, so pullbacks do not preserve exactness.

Kernels do not preserve exactness. Consider the following diagram.

$$\begin{tikzcd}
\{0\} \arrow[r] & \mathbf{Z} \arrow[r] \arrow[d, "\times 2"] & \mathbf{Z} \arrow[r] \arrow[d, "\times 2"] & \mathbf{Z}/2\mathbf{Z} \arrow[r] \arrow[d, "\times 2"] & \{0\} \\
\{0\} \arrow[r] & \mathbf{Z} \arrow[r] & \mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\}
\end{tikzcd}$$

Taking kernels, one gets

$$\begin{tikzcd}\{0\} \arrow[r] & \{0\} \arrow[r] & \{0\} \arrow[r] & \mathbf{Z}/2\mathbf{Z}.\end{tikzcd}$$

The rightmost arrow is not onto, so kernels do not preserve exactness.

Inverse limits do not preserve exactness. Consider the following diagram.

$$\begin{tikzcd}
 & {} \arrow[d, "\times 3"] & \text{etc.} \arrow[d, "\times 3"] & {} \arrow[d, "\times 3"] & \\
\{0\} \arrow[r] & \mathbf{Z} \arrow[r, "\times 2"] \arrow[d, "\times 3"] & \mathbf{Z} \arrow[r] \arrow[d, "\times 3"] & \mathbf{Z}/2\mathbf{Z} \arrow[r] \arrow[d, "\times 3"] & \{0\} \\
\{0\} \arrow[r] & \mathbf{Z} \arrow[r, "\times 2"] \arrow[d, "\times 3"] & \mathbf{Z} \arrow[r] \arrow[d, "\times 3"] & \mathbf{Z}/2\mathbf{Z} \arrow[r] \arrow[d, "\times 3"] & \{0\} \\
\{0\} \arrow[r] & \mathbf{Z} \arrow[r, "\times 2"] & \mathbf{Z} \arrow[r] & \mathbf{Z}/2\mathbf{Z} \arrow[r] & \{0\}
\end{tikzcd}$$

One finds that the inverse limits of the second and third columns must be $\{0\}$ (we need the thing to be a multiple of every power of 3). The inverse limit of the fourth column is $\mathbf{Z}/2\mathbf{Z}$ since we have $1 \longmapsto 1$ at each arrow. So inverse limits do not preserve exactness (despite direct limits preserving exactness).

Mittag–Leffler came up with a condition on the A_i to guarantee exactness. First, suppose each $A_{i+1} \longrightarrow A_i$ is surjective. We have sequences

$$\begin{tikzcd}\{0\} \arrow[r] & A_i \arrow[r] & B_i \arrow[r] & C_i \arrow[r] & \{0\},\end{tikzcd}$$

and we choose an element in the inverse limit of the C_i: compatible elements c_i where the image of c_{i+1} is c_i. We need to find an element in the inverse limit of the B_i so it maps onto the c_i. Suppose we have found elements b_0 and b_1. We can find an element b_2 mapping onto c_2, but its image may not be b_1. We diagram chase[2] and determine that we can find such a b_2. So if the maps of A_i are surjective, we get exactness.

Now suppose $A_{i+1} \longrightarrow A_i$ is zero. Again, we have compatible elements c_i and elements b_i mapping onto the c_i. But the b_i do not have to be compatible. Notice that b_2 and b_1 have the same image in B_0 since the map from A_1 to A_0 is zero. So we replace b_i with $\operatorname{im}(b_{i+1})$, and these elements still map onto the c_i. Now the "b_i" are compatible. So if the maps of A_i are zero, we get exactnesss.

Now suppose that $\operatorname{im}(A_j)$ in A_i is zero for j large enough. Pick A_0 and pick A_{i_1} such that its image in A_0 is $\{0\}$, then pick A_{i_2} such that its image in A_{i_1} is $\{0\}$, et cetera. Taking such a subsequence does not affect the inverse limit, so this case reduces to the last one.

Now suppose that the image of A_j in A_i stabilizes. So we have

$$\operatorname{im}(A_i) \supseteq \operatorname{im}(A_{i+1}) \supseteq \cdots \supseteq \operatorname{im}(A_j) = \operatorname{im}(A_{j+1}) = \cdots .$$

The module A_i has a submodule S_i that is the stable limit of A_j for $j \gg i$. Then we get exact sequences

$$\begin{tikzcd}
\{0\} \arrow[r] & S_{i+1} \arrow[r] \arrow[d] & A_{i+1} \arrow[r] \arrow[d] & A_{i+1}/S_{i+1} \arrow[r] \arrow[d] & \{0\} \\
\{0\} \arrow[r] & S_i \arrow[r] & A_i \arrow[r] & A_i/S_i \arrow[r] & \{0\}
\end{tikzcd}$$

Our assumption implies that the arrows in the second column are surjective and the arrows in the fourth column satisfy the previous case. So exactness is preserved. This is the **Mittag–Leffler condition**.

Example 14.1 (Does not satisfy the Mittag–Leffler condition)**.** The sequence

$$\begin{tikzcd}
\mathbf{Z} & \mathbf{Z} \arrow[l, "\times 3"'] & \mathbf{Z} \arrow[l, "\times 3"'] & \mathbf{Z} \arrow[l, "\times 3"'] & \cdots \arrow[l, "\times 3"']
\end{tikzcd}$$

does not satisfy the Mittag–Leffler condition, and as we have seen, exactness is not prserved.

Example 14.2 (Satisfies the Mittag–Leffler condition)**.** Suppose the A_i are finite. Then the Mittag–Leffler condition is satisfied.

15 Polynomials

The polynomial ring over a field k is a Euclidean domain, so it is also a PID and a UFD. So any finitely generated $k[x]$–module is a direct sum of cyclic $k[x]$–modules.

There is an algorithm similar to the Sieve of Eratosthenes to determine the irreducibles over $k[x]$.

[2] See https://youtu.be/u8Pd7ozrXRo?t=637.

Recall that if a is a root of a polynomial, then that polynomial is divisible by $x - a$. So, over a field, a polynomial of degree n has at most n roots. It is important that k is a field, or, more specifically, that k is commutative and has no zero divisors.

Example 15.1 (k must be a field)**.** Over $\mathbf{Z}/8\mathbf{Z}[x]$, the polynomial $x^2 - 1$ has 4 roots. (Notice that $\mathbf{Z}/8\mathbf{Z}$ is not a field.) Over $\mathbf{H}[x]$, the polynomial $x^2 + 1$ has roots $\pm i$, $\pm j$, and $\pm k$ (and uncountably many more).

Proposition 15.2. *The group $(\mathbf{Z}/p\mathbf{Z})^*$ is cyclic. So it has a so-called **primitive root** that generates it.*

Proof.—Consider the field $\mathbf{Z}/p\mathbf{Z}$. Over this field, $x^n - 1$ has at most n roots. So $(\mathbf{Z}/p\mathbf{Z})^*$ has at most n elements of order dividing n, and hence it is cyclic. There is a counting argument due to Gauss[3] that proves this. One can also use the classification of finite abelian groups. □

Proposition 15.3. *Suppose F is a finite field. Then F^* is cyclic. Suppose k is a field. If G is a finite subgroup of k^*, then G is cyclic.*

Proof.—See the proof for Proposition 15.2. □

Example 15.4 (F must be a field)**.** The group $(\mathbf{Z}/8\mathbf{Z})^*$ is not cyclic (the elements 1, 3, 5, and 7 all satisfy $g^2 = 1$). The group of unit quaternions has a noncyclic subgroup $\{\pm 1, \pm i, \pm j, \pm k\}$.

Warning 15.5.—If a polynomial of degree n vanishes at more than n points, it is 0. However, a polynomial may vanish at all elements of the field without being 0.

Example 15.6. The polynomial $x^p - x$ vanishes at all points of $\mathbf{Z}/p\mathbf{Z}$.

Proposition 15.7. *The ring $\mathbf{Z}[x]$ is a UFD.*

Proof.—First note that $\mathbf{Z}[x]$ is neither Euclidean nor a PID. Recall that $\mathbf{Q}[x]$ is a UFD (since $\mathbf{Q}$ is a field). Write the content $c(f)$ of a polynomial f to be the greatest common divisor of the coefficients of f. We show that $c(f)c(g) = c(fg)$: replace f by $f/c(f)$ and g by $g/c(g)$. So $c(f) = c(g) = 1$, and we want to show that $c(fg) = 1$. Pick an integer prime p. Write

$$f(x) = a_m x^m + \cdots + a_i x^i + a_{i-1} x^{i-1} + \cdots + a_0$$

and

$$g(x) = b_n x^n + \cdots + b_j x^j + b_{j-1} x^{j-1} + \cdots + b_0.$$

[3] See Disquisitiones Arithmeticae arts. 52–56.

Assume that $a_{i-1}, \ldots, a_0$ and $b_{j-1}, \ldots, b_0$ are divisible by p and a_i and b_j are not divisible by p. Then the coefficient of the x^{i+j} term of $f(x)g(x)$ is $a_0 b_{i+j} + a_1 b_{i+j-1} + \cdots + a_i b_j + \cdots + a_{i+j-1} b_1 + a_{i+j} b_0$. The prime p does not divide this quantity. Since $\mathbf{Z}$ is a UFD, any irreducible is prime. So $c(fg) = 1$. The irreducibles of $\mathbf{Z}[x]$ are the irreducibles (primes) in $\mathbf{Z}$ together with polynomials of degree greater than 0 with content 1 and irreducible in $\mathbf{Q}[x]$. All elements of $\mathbf{Z}[x]$ are products of these irreducibles. We need to show that these are prime in $\mathbf{Z}[x]$. One shows this by using the fact that $\mathbf{Q}[x]$ is a UFD and $c(fg) = c(f)c(g)$. □

Similarly, if R is a UFD, then $R[x]$ is a UFD. So, if k is a field, one can prove that $k[x]$ is a UFD, then prove that $k[x, y]$ is a UFD, et cetera. So a ring of polynomials in n variables over a field is a UFD. Polynomials in an infinite number of variables over a field is still a UFD. Likewise, $\mathbf{Z}[x_1, \ldots, x_n]$ is a UFD.

16 Factorization of polynomials

Given a polynomial $f \in \mathbf{Z}[x]$, there are algorithms for writing f as a product of irreducible polynomials. Kronecker came up with the first such algorithm. If g_1 and g_2 are equal at more than $\min\{\deg g_1, \deg g_2\}$ points, then $g_1 = g_2$. Suppose $\deg f = n$. Pick $n+1$ points $a_1, \ldots, a_{n+1}$ and consider $f(a_1), \ldots, f(a_{n+1})$. If $f = gh$, then $g(a_1) \mid f(a_1)$, etc. There are a finite number of possibilites for these $n+1$ cases, so there are a finite number of possibilities for g. But this can be very slow. Kronecker's algorithm can be extended to polynomials in many variables.

There is no algorithm for determining whether $f \in \mathbf{Z}[x_1, \ldots, x_n]$ has a root in $\mathbf{Z}^n$. (This was Hilbert's tenth problem solved by Matiyasevich[4] along with Robinson, Davis, and Putnam. Roughly, they showed that one can encode any Turing machine in a Diophantine equation such that the Turing machine halts if and only if the equation has a solution, and since the halting problem is undecidable, Hilbert's tenth problem has no positive solution.)

There is almost a "fast" algorithm for factorizing polynomials over $\mathbf{Z}[x]$. One takes $f \in \mathbf{Z}[x]$ and writes it as $f = c(f)g$, and the LLL (Lenstra–Lenstra–Lovász) algorithm allows one to factor g in polynomial time. However, determining $c(f)$ comes down to factoring integers (and we don't know whether integers can be factored in polynomial time). Of course, Shor's algorithm would allow one to determine $c(f)$ in polynomial time.

One notices that almost all polynomials with degree greater than 0 and content 1 are irreducible over $\mathbf{Z}$. If a polynomial is irreducible over $\mathbf{Z}/p\mathbf{Z}$, then it is irreducible over $\mathbf{Z}$ (granted the degree over $\mathbf{Z}/p\mathbf{Z}$ is the same as that over $\mathbf{Z}$). Recall Eisenstein's criterion: if $x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbf{Z}[x]$ satisfies

1. p divides $a_0, \ldots, a_{n-1}$;

[4] See https://mitpress.mit.edu/books/hilberts-10th-problem.

2. p^2 does not divide a_0;

then f is irreducible.

Proof.—Write $f = gh$ where $g = x^m + \cdots + b_0$ and $h = x^{n-m} + \cdots + c_0$. Then p divides one of b_0 and c_0. So assume p divides b_0 and not c_0. Suppose $b_0, \ldots, b_k$ are divisible by p and b_{k+1} is not divisible by p. Then the coefficient of x^{k+1} in gh is $b_{k+1}c_0 + b_k c_1 + \cdots$. The prime p does not divide this quantity. Contradiction. So f is irreducible. □

Example 16.1. Claim: the polynomial

$$\frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + 1$$

is irreducible. Write $x + 1$ instead of x. Then we get the polynomial

$$\frac{(x+1)^p - 1}{x} = x^{p-1} + px^{p-2} + \binom{p}{2} x^{p-3} + \cdots + p.$$

By Eisenstein's criterion, this polynomial is irreducible, so the polynomial in question is also irreducible.

Such a change of variables works because p is totally ramified in $\mathbf{Z}[\zeta]$ (where ζ is a primitive pth root of unity). The field $\mathbf{Z}[\zeta]$ is a cyclotomic field, and p being totally ramified means that $p = \text{unit} \times (1-\zeta)^{p-1}$. Whenever one has a prime totally ramified in an extension field, one gets an analogue of Eisenstein's criterion that proves that the polynomial defining this field is irreducible.

In the nineteenth century, people were interested in factoring numbers of the form $p^n \pm 1$. For instance, after years of cumbersome computation, Landry (1869) determined that

$$2^{58} + 1 = 5 \cdot 107367629 \cdot 536903681.$$

Aurifeuille (1871) determined that this prime factorization does not require manual computation: one can use that

$$4a^4 + b^4 = (2a^2 + 2ab + b^2)(2a^2 - 2ab + b^2).$$

These factorizations are called Aurifeuillian factorizations. Take, for example,

$$x^4 + 4 = (x^2 + 2x + 2)(x^2 - 2x + 2).$$

Exercise 16.2. Show that $n^4 + 4^n$ is never prime if $n > 1$.

All linear factors $x-\alpha$ of $x^n+a_{n-1}x^{n-1}+\cdots+a_0 \in \mathbf{Z}[x]$ satisfy $\alpha \mid a_0$. If $n \leq 3$ and the polynomial is irreducible, it has a linear factor. For example, the polynomial x^3-3x+1 is irreducible since it has no linear factors (one checks $x \pm 1$). This polynomial has roots $2\cos(2\pi/9)$, $2\cos(4\pi/9)$, and $2\cos(8\pi/9)$. (You cannot trisect a 60° angle with ruler and compass.)

17 Noetherian rings

The ring $k[x]$ is a principal ideal domain. The ring $k[x,y]$ has a nonprincipal ideal (x,y) generated by 2 elements. Similarly, $k[x,y,z]$ has a nonprincipal ideal (x,y,z) generated by 3 elements. However, the ideal $(x^2,xy,y^2) \subset k[x,y]$ requires 3 generators, and, more generally, $(x^n,x^{n-1}y,\ldots,y^n) \subset k[x,y]$ requires more than n generators.

Nevertheless, Hilbert determined that every ideal of $k[x_1,\ldots,x_n]$ is finitely generated.

Definition 17.1. Let R be a ring. If all ideals $I \subset R$ are finitely generated, then R is a **Noetherian ring**.

This concept was first considered by Hilbert, but Noether determined that rings other than polynomial rings were Noetherian. She found that if R is Noetherian, then $R[x]$ is Noetherian.

Let R be a ring. The following are equivalent.

1. R is Noetherian;
2. every ideal of R is finitely generated;
3. every nonempty set of ideals has a maximal element;
4. every strictly increasing chain of ideals is finite.

The first and second things are equivalent by definition. The fact that the third and the fourth statements are equivalent has nothing to do with rings: it is true for all partially ordered sets.

If there is a strictly increasing infinite chain $I_0 \subset I_1 \subset \cdots$, then the set $\{I_0, I_1, \ldots\}$ has no maximal element. If a nonempty set S has no maximal elements, we pick $I_0 \in S$, then pick $I_1 \supset I_0$ since S has no maximal element. We continue and get an infinite chain $I_0 \subset I_1 \subset \cdots$. Thank you, axiom of choice. So the third and fourth statements are equivalent.

Suppose we have a chain $I_0 \subset I_1 \subset \cdots$. The union $I = \bigcup_j I_j$ is an ideal (a union of ideals doesn't need to be an ideal in general, but we have this chain). The ideal I is finitely generated by elements $i_1,\ldots,i_n$ all in I_m. Then $I_m = I_{m+1} = \cdots$, so this chain must be finite.

Suppose I is an ideal. Pick $i_0 \in I$, and if $I \neq (i_0)$, pick $i_1 \in I \setminus (i_0)$. If $I \neq (i_0, i_1)$, pick $i_2 \in I \setminus (i_0, i_1)$, et cetera. We get a chain $(i_0) \subset (i_0, i_1) \subset (i_0, i_1, i_2) \subset \cdots$, and by assumption all strictly increasing chains are finite, so this process stops.

Example 17.2 (Non–Noetherian ring)**.** Consider the ring $k[x_0, x_1, \ldots]$. The ideal $(x_0, x_1, \ldots)$ is not finitely generated. Also, the chain of ideals $(x_0) \subset (x_0, x_1) \subset \cdots$ is infinite. Further, the set of these ideals has no maximal element.

Example 17.3 ("Flipping" the conditions)**.** In $\mathbf{Z}$, there is a stricly decreasing chain of ideals $(2) \supset (4) \supset (8) \supset \cdots$. Indeed, there is a nonempty collection of ideals with no minimum element. But $\mathbf{Z}$ is Noetherian, so these conditions are stronger. If all nonempty sets of ideals have a minimum element, the ring is an **Artinian ring**.

Example 17.4 (Artinian ring)**.** The ring $k[x, y]/(x^n, y^n)$ is an n^2–dimensional vector space over k, and there cannot be infinite chain of decreasing vector spaces in a finite dimensional vector space, so this ring is Artinian.

To be Noetherian is somewhat subtle. Consider the following table of (mostly) inclusions of rings.

Ring	Noetherian?
$\mathbf{C}[x]$	Yes
Holomorphic functions in $\mathbf{Z}$	No
Holomorphic functions on closed unit disk	Yes
Holomorphic functions on open unit disk	No
Functions holomorphic in some neighbourhood of 0	Yes
Germs of smooth real functions at 0	No
Formal power series (not an inclusion)	Yes

The ring $\mathbf{C}[x]$ is a PID, so it is Noetherian; the only ideals of $\mathbf{C}[\![x]\!]$ are (x^n) and (0), so it is also Noetherian; et cetera. The ring of germs of smooth functions is non-Noetherian. The ideal of all germs of functions vanishing to infinite order at 0 not finitely generated. Also, all elements of a Noetherian ring are a product of irreducibles, and

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is not: $f = \left(\sqrt{f}\right)^2 = \left(\sqrt[4]{f}\right)^4 = \cdots$.

Noetherian rings tend to correspond to finite dimensional spaces, whereas non-Noetherian rings tend to correspond to infinite dimensional spaces (though this is not always true). Noetherian rings seem to turn up in algebra, and non-Noetherian rings seem to turn up in analysis. Noetherian rings are associated with "nice" roots (on $\mathbf{C}[x]$, one has a finite number of roots), and non-Noetherian rings are associated with "nasty" roots (consider holomorphic functions on $\mathbf{C}$).

If R is Noetherian, any quotient of R is Noetherian. However, a subalgebra of R does not need to be Noetherian.

18 Hilbert's theorems

Theorem 18.1 (Hilbert). *Let k be a field. Then the ring of polynomials $k[x_1, \ldots, x_n]$ in finitely many variables is Noetherian.*

Hilbert also showed that $\mathbf{Z}[x_1, \ldots, x_n]$ is Noetherian. Noether proved this by showing that if R is Noetherian, then $R[x]$ is Noetherian.

Suppose R is Noetherian and I is an ideal of $R[x]$. We want to show that I is finitely generated as an ideal. Let $I_0 \subset R$ be the ideal of leading coefficients of degree 0 elements in I. Let I_j be the ideal of leading coefficients of degree j elements in I. So $I_0 \subset I_1 \subset I_2 \subset \cdots$. This chain must stabilize, so $I_n = I_{n+1} = \cdots$ for some n. Then the set of generators of I consists of a finite set of degree j polynomials whose leading coefficients generate I_j for $0 \leq j \leq n$. We will show these generate I.

Pick $f \in I$ and by induction on $\deg f$ we show that f is generated by the set described above. We can find a polynomial of the form $\sum x^\alpha \cdot$ (element of generating set) with the same leading coefficient as f.

Theorem 18.2 (Hilbert again). *Rings of invariants are often finitely generated as algebras.*

Suppose a group G acts on a finite dimensional vector space V defined over k. Then G acts on the algebra of polynomials on V: if the dual of V has a basis $x_1, \ldots, x_n$, then the algebra of polynomials is $k[x_1, \ldots, x_n]$. The ring of invariants $k[x_1, \ldots, x_n]^G$ consists of all polynomials fixed by G. Is this invariant ring a finitely generated algebra?

Example 18.3. Suppose $G = \mathrm{S}_n$ and $V = \mathbf{C}^n$ is finite dimensional. The invariants consist of polynomials in $\mathbf{C}[x_1, \ldots, x_n]^{\mathrm{S}_n}$: these are polynomials that are invariant under permutation of the formal variables. Examples of elements in this invariant ring include $\sum_i x_i$, $\sum_{i<j} x_i x_j$, $\sum_{i<j<k} x_i x_j x_k$, etc. These elementary symmetric polynomials generate the ring of invariants (like $\sum_i x_i^2$).

Example 18.4. Let $G = \mathbf{Z}/4\mathbf{Z}$ and V be a 2–dimensional vector space. The group G with generator g acts on $k[x, y]$ by $g(x) = ix$ and $g(y) = iy$. The ring of invariants comprises 1 and the polynomials of degree dividing 4, and it is generated by x^4, x^3y, x^2y^2, xy^3, and y^4. (It is not normally the case that one can write down the generators of the invariant ring with ease.)

Example 18.5. Suppose $G = \mathrm{SL}_2(\mathbf{C})$ acts on binary forms (things of the form $a_n x^n + a_{n-1}x^{n-1}y + \cdots + a_0 y^n$ of arbitrary degree). The group G acts on x and y ($\langle a_0, \ldots, a_n\rangle$) by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}.$$

An invariant is a polynomial in $a_0, \ldots, a_n$ invariant under the action of G. Gordan showed[5] that the ring of invariants is a finitely generated algebra for more general groups and forms.

Hilbert's solution was simpler than Gordan's.

Example 18.6 (A non–finitely-generated subring of a polynomial ring)**.** The subalgebra of $k[x, y]$ consisting of linear combinations of elements in $k[x, y] \setminus \{x^\alpha : \alpha > 0\}$ is not finitely generated as an algebra.

These subalgebras have a Reynolds operator. Suppose k is a field of zero characteristic and G is a finite group. Then the **Reynolds operator** ρ is given by averaging over G:

$$\rho(f) = \frac{1}{|G|}\sum_{g\in G} g(f).$$

It has the following properties.

1. $\rho(1) = 1$;
2. $\rho(f) = f$ if f is fixed by G;
3. $\rho(fg) = f\rho(g)$ if f is fixed by G.

The group G acts on the ring of polynomials $k[x_1, \ldots, x_n]$. Let R be the ring of invariants. We want to show that R is finitely generated as an algebra. The ring R is graded: $R = R_0 \oplus R_1 \oplus \cdots$ where R_n is homogenous of degree n. Let I be the ideal of $k[x_1, \ldots, x_n]$ generated by R_1, R_2, et cetera (but not R_0 since $1 \in R_0$). The ideal I is finitely generated. Pick a finite set $\{i_1, \ldots, i_m\}$ of homogenous generators of I in R. We want to show that $i_1, \ldots, i_m$ generate R as an algebra (not an ideal).

Pick $f \in R$ homogenous with degree greater than 0. The element f is in the ideal I, so $f = a_1 i_1 + \cdots + a_m i_m$ for some $a_i \in k[x_1, \ldots, x_n]$. (The a_is do not need to be in R.) So

$$f = \rho(f) = \rho(a_1)i_1 + \cdots + \rho(a_m)i_m$$

since f and the i_js are invariants. Further, $\rho(a_1), \ldots, \rho(a_m) \in R$, so by induction on the degree, they are polynomials in $i_1, \ldots, i_m$. So f is a polynomial in $i_1, \ldots, i_m$, hence R is generated by $i_1, \ldots, i_m$. □

[5] See p. 85 of Mathematics of the 19th Century: Mathematical Logic, Algebra, Number Theory, Probability Theory.

The same proof works for compact groups: one defines ρ by integrating over G instead. But groups such as $\mathrm{SL}_2(\mathbf{C})$ are not compact. One notices that $\mathrm{SL}_2(\mathbf{C}) \supset \mathrm{SU}_2(\mathbf{R})$ essentially have the same invariants, and $\mathrm{SU}_2(\mathbf{R})$ is compact. This—Weyl's unitarian trick—makes Hilbert's finiteness theorem applicable to any reductive Lie group.

Nagata found[6] examples of groups G such that $k[x_1, \ldots, x_n]^G$ is not finitely generated. Noether—for finite groups—and Haboush—for infinite groups—extended Hilbert's theorem for fields of characteristic p.

19 Symmetric functions

The group S_n acts on n variables $\alpha_1, \ldots, \alpha_n$ by permuting them. The invariant polynomials under this action include

$$\begin{aligned}
e_1 &= \alpha_1 + \cdots + \alpha_n = \sum_i \alpha_i \\
e_2 &= \alpha_1\alpha_2 + \alpha_1\alpha_3 + \cdots + \alpha_{n-1}\alpha_n = \sum_{i<j} \alpha_i\alpha_j \\
e_3 &= \sum_{i<j<k} \alpha_i\alpha_j\alpha_k \\
&\ \ \vdots \\
e_n &= \alpha_1 \cdots \alpha_n = \prod_i \alpha_i.
\end{aligned}$$

These are elementary symmetric functions. Other such invariant polynomials include $\sum_i \alpha_i^2 = e_1^2 - 2e_2$. Any symmetric polynomial is a polynomial in the e_i.

The idea of the proof is to lexicographically order the monomials: $\alpha_1^{n_1}\alpha_2^{n_2}\cdots > \alpha_1^{m_1}\alpha_2^{m_2}\cdots$ if $n_1 > m_1$, or $n_1 = m_1$ and $n_2 > m_2$, etc.

The algorithm for writing a symmetric polynomial in terms of the elementary symmetric polynomials works as follows. Suppose f is a symmetric polynomial. Kill off the largest monomial in f by subtracting a monomial in the e_is. One can always do this.

Suppose the largest monomial is $\alpha_1^{n_1}\alpha_2^{n_2}\cdots$. Then subtract

$$\left(\sum_i \alpha_i\right)^{n_1-n_2} \left(\sum_{i<j} \alpha_i\alpha_j\right)^{n_2-n_3} \cdots.$$

This replaces $\alpha_1^{n_1}\alpha_2^{n_2}\cdots$ with a smaller monomial, and we repeat.

[6] See https://web.archive.org/web/20131102202816/http://www.mathunion.org/ICM/ICM1958/Main/icm1958.0459.0462.ocr.pdf.

One can guarantee that $n_i - n_{i+1} \geq 0$ only if one can arbitrarily permute the variables, so f must be symmetric for this to work.

How does one write $p_m = \sum_i \alpha_i^m$ in terms of elementary symmetric polynomials? The first few are straightforward:

$$\begin{aligned} p_0 &= \alpha_1^0 + \cdots + \alpha_n^0 = n, \\ p_1 &= \alpha_1 + \cdots + \alpha_n = e_1, \\ p_2 &= \alpha_1^2 + \cdots + \alpha_n^2 = e_1^2 - 2e_2. \end{aligned}$$

The polynomial p_3 seems somewhat tricky, and p_i for bigger i seems like it will be even more difficult. The following identity is useful:

$$\prod_i (x - \alpha_i) = \sum_i (-1)^i e_i x^{n-i} := f(x).$$

One can define the derivative of a polynomial over any field as follows:

$$\frac{d}{dx}(x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) := nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.$$

Identities such as the product rule hold. Applying $(d/dx)(\log\ \cdot\ )$ on both sides (to turn products to sums and to get rid of logarithms), we get

$$\sum_i \frac{1}{x - \alpha_i} = \frac{f'(x)}{f(x)}.$$

Writing the lefthand side as a formal power series, we get

$$\sum_i \left(\frac{1}{x} + \sum_j \frac{\alpha_j}{x^{j+1}} \right) = \frac{f'(x)}{f(x)},$$

so

$$\left(\sum_i \frac{p_i}{x^{i+1}} \right) (x^n - e_1 x^{n-1} + \cdots + (-1)^n e_n) = nx^{n-1} - (n-1)e_1 x^{n-2} + \cdots + (-1)^{n-1} e_{n-1}.$$

Comparing coefficients, we see that $p_0 = n$, $p_1 = e_1$, and $p_2 - e_1 p_1 = -2e_2$. Similarly, we find $p_3 - e_1 p_2 + e_2 p_1 = 3e_3$, $p_4 - e_1 p_2 + e_2 p_2 - e_3 p_1 = -4e_4$, et cetera.

Example 19.1. Consider the polynomial $z^3 + z + 1$. What is the sum of the fifth powers of the roots of this polynomial? Let its roots be α_1, α_2, and α_3. The sum of the fifth powers is p_5. We find that $p_0 = 3$, $p_1 = 0$, $p_2 + p_0 = 1$, $p_3 + p_1 + p_0 = 0$, $p_4 + p_2 + p_1 = 0$, and $p_5 + p_3 + p_2 = 0$, so $p_2 = -2$, $p_3 = -3$, $p_4 = -1$, and $p_5 = 5$.

Newton's identities also show up in Adams operations of K-theory. Let V be a vector space acted on by a group G. The tensor product $V \otimes V$ is acted on by G. This tensor product splits as a direct sum of two vector spaces acted on by G: one has $V \otimes V = \Sigma^2 V \oplus \Lambda^2 V$ where Σ^2 is the **symmetric square** and Λ^2 is the **alternating square**. The summand $\Sigma^2 V$ is spanned by things of the form $a \otimes b + b \otimes a$ and $\Lambda^2 V$ is spanned by things of the form $a \otimes b - b \otimes a$, so if $\dim V = n$, then $\dim(\Sigma^2 V) = n(n+1)/2$ and $\dim(\Lambda^2 V) = n(n-1)/2$. The **Adams operation** is given by $\Psi^2 V = \Sigma^2 V - \Lambda^2 V$. One "[takes] the semiring of all vector spaces under direct sum and tensor product and [forces] it to be a ring by decreeing that [one] can subtract vector spaces." This difference is called a **virtual vector space**.

The group G clearly cannot act on this non–vector space. Nevertheless, one can consider the trace of $g \in G$ on $\Psi^2(V)$. One sees that $\operatorname{Tr} g$ on $\Sigma^2 V$ is $\sum_{i \le j} \lambda_i \lambda_j$ where the λ_is are the eigenvalues of g on V. The trace of g on $\Lambda^2 V$ is $\sum_{i<j} \lambda_i \lambda_j$. So the trace of g on $\Psi^2 V$ is (by the additivity of the trace) $\sum_i \lambda_i^2 = p_2$. One may also consider the nth Adams operation $\Psi^n V$—where one considers $V^{\otimes n}$—in which the trace of an element is given by p_n.

Consider A_3 acting on x, y, and z. What are the invariant polynomials? Examples include e_1, e_2, and e_3. Suppose g swaps x and y and f is invariant under A_3. Then

$$f = \frac{f + gf}{2} + \frac{f - gf}{2},$$

and the first term is invariant under S_3 and the second term is antiinvariant under S_3. An example of an antiinvariant polynomial is $\Delta = (x-y)(y-z)(z-x)$. This polynomial Δ divides any antiinvariant (since any antiinvariant vanishes when any two variables are equal), so the invariants of A_3 are the same as the invariants of S_3 along with $\Delta \times$ the invariants of S_3. So the invariant ring is $\mathbf{C}[e_1, e_2, e_3, \Delta]$ modulo some ideal (since these polynomials are not algebraically independent). Notice that Δ^2 is invariant, so it must be equal to some polynomial in e_1, e_2, e_3. This polynomial turns out to be

$$\Delta^2 = 18 e_1 e_2 e_3 - 4 e_1^3 e_3 + e_1^2 e_2^2 - 4 e_2^3 - 27 e_3^2.$$

Can one determine this without manual computation? Suppose we have a polynomial $x^n - e_1 x^{n-1} + \cdots$ with roots α_i. If $\Delta := \prod_{i<j}(\alpha_i - \alpha_j)$, then $\Delta^2 = \prod_{i<j}(\alpha_i - \alpha_j)^2$ is a polynomial in the e_is. In the degree 2 case, we know that $\Delta^2 = b^2 - 4c$. One uses the resultant of two polynomials (the topic of the next section) to determine the discriminant of higher degree polynomials.

20 Resultants

Suppose $f(x) := x^n - e_1 x^{n-1} + e_2 x^{n-2} - \cdots$ has roots $\alpha_1, \ldots, \alpha_n$. The discriminant $\prod_{i<j}(\alpha_i - \alpha_j)^2$ is zero if f has a repeated root. The polynomial f has a repeated root if and only if f and f' have a root in common.

Warning 20.1.—Over fields of characteristic p, the derivative of $f(x) = x^p - \alpha$ is zero: a polynomial might divide its derivative.

When do f and f' have a common root? More generally, when do f and g have a common root? If $x - \alpha$ divides f and g, then $f(x)p(x) = g(x)q(x)$ where $\deg p < \deg g$ and $\deg q < \deg f$. This equation corresponds to a set of linear equations in the coefficients of p and q, where the coefficients of these equations relate to the coefficients of f and g. A set of linear equations has a nonzero solution if and only if some determinant is zero. What is this determinant?

Suppose $f(x) = a_m x^m + \cdots + a_0$ and $g(x) = b_n x^n + \cdots + b_0$. Then the matrix giving the linear equations is the Sylvester matrix:

$$\begin{pmatrix}
a_m & \cdots & a_0 & & & & \\
 & a_m & \cdots & a_0 & & & \\
 & & & \ddots & & & \\
 & & & a_m & \cdots & a_0 & \\
 & & & & a_m & \cdots & a_0 \\
b_n & \cdots & b_0 & & & & \\
 & b_n & \cdots & b_0 & & & \\
 & & & \ddots & & & \\
 & & & b_n & \cdots & b_0 & \\
 & & & & b_n & \cdots & b_0
\end{pmatrix}$$

The determinant of this matrix is the **resultant** of f and g (where $a_m \neq 0$ and $b_n \neq 0$). The polynomials f and g have a common root if and only if the resultant is 0.

So the discriminant of f is the resultant of f and f' up to a constant.

Example 20.2. Suppose $f(x) = x^3 + ax + b$. Then $f'(x) = 3x^2 + a$. In this case, the resultant is given by

$$\begin{vmatrix}
1 & 0 & a & b & 0 \\
0 & 1 & 0 & a & b \\
3 & 0 & a & 0 & 0 \\
0 & 3 & 0 & a & 0 \\
0 & 0 & 3 & 0 & a
\end{vmatrix} = \pm(4a^3 + 27b^2).$$

To get the sign, one evaluates both sides for an explicit polynomial. The roots of $x^3 - 1$ are the 3rd roots of unity, and the product $\prod_{i<j}(\alpha_i - \alpha_j)^2$ is less than 0. So $\operatorname{Disc} f = -4b^3 - 27c^2$.

Example 20.3 (Discriminants of degenerate cubics)**.** If α is a root of $x^3 + x + 1$, the discriminant of $\mathbf{Q}[\alpha]$ is -31. The curve $y^2 = x^3 + ax + b$ is elliptic if $-4b^3 - 27c^2 \neq 0$.

Suppose f and g are homonogenous polynomials in y and z with coefficients in $k[x_1, \ldots, x_n]$. Then f and g define hypersurfaces on n-dimensional affine space times a copy of the projective line. The roots of the resultant of f and g is (hypersurface defined by f) $\cap$ (hypersurface defined by g) projected on n-dimensional affine space, so this space is closed. (The projection from the affine plane to the affine line is not closed.) The projective line $\mathbf{P}^1$ is a complete variety, and $\mathbf{P}^n$, in general, is a complete variety.

We saw that the invariants of A_3 on $k[x, y, z]$ are $k[e_1, e_2, e_3, \Delta]/(\Delta - \mathrm{Disc}(e_1, e_2, e_3))$. The ideal $(\Delta - \mathrm{Disc}(e_1, e_2, e_3))$ is a **syzygy**.

Example 20.4 (Syzygies)**.** The ring $\mathbf{Z}/n\mathbf{Z}$ acts on $k[x, y]$ by $x \longmapsto x\zeta$ and $y \longmapsto y\zeta$ where ζ is a primitive nth root of unity. Then the ring of invariants is $k[x^n, x^{n-1}y, \ldots, y^n]$. Call these generators a_n, a_{n-1}, etc. We have relations like $a_{n-1}^2 = a_n a_{n-2}$, $a_{n-2}^2 = a_n a_{n-3}$, etc. These are syzygies.

21 Formal power series

Elements of $\mathbf{C}[\![x]\!]$ are of the form $\sum_i a_i x^i$. One does not worry about convergence.

Let R be a ring. Then $R[\![x]\!] = \varprojlim R[x]/(x^i)$. The ring $R[x]/(x)$ consists of elements of the form $a_0 + O(x)$, the ring $R[x]/(x^2)$ consists of elements of the form $a_0 + a_1 x + O(x^2)$, et cetera. To compute the inverse limit one picks elements of the rings in the limit that are compatible (with respect to natural homomorphisms between them). We have homomorphisms from $R[x]/(x^{i+1})$ to $R[x]/(x^i)$. So the constant terms, the linear terms, the quadratic terms, etc. must be the same, and this gives formal power series over R.

This concept generalizes. If R is a ring and I is an ideal of R, then $\widehat{R} = \varprojlim R/I^n$ is called the **completion** of R. Completions tend to behave like formal power series rings. An example of a complete ring is $\mathbf{Z}_p$. In this case, we have $R = \mathbf{Z}$ and $I = (p)$. Elements in $\mathbf{Z}/(p)$ are of the form a_0 in base p, elements in $\mathbf{Z}/(p^2)$ are of the form $a_1 a_0$ in base p, et cetera. The inverse limit results in elements of the form $\cdots a_3 a_2 a_1 a_0 = \cdots + a_3 p^3 + a_2 p^2 + a_1 p + a_0$ going on forever to the left. This is reminiscent of formal power series.

Let k be a field. The ring $k[\![x]\!]$ has one maximal ideal: if $\sum_i a_i x^i \in k[\![x]\!]$ where $a_0 \neq 0$, we can write it as $1 + xg(x)$, which has an inverse $\sum_i (-1)^i x^i g(x)^i \in k[\![x]\!]$. Since every power series with a nonzero constant term has an inverse, the only maximal ideal is $\mathfrak{m} = (x)$. Any ideal is either (0) or (x^n). The ring $k[\![x]\!]$ is a PID—further, it is a discrete valuation ring.

Definition 21.1. The ring R is a **discrete valuation ring** if it is a PID that has a unique nonzero prime ideal $\mathfrak{p}(R)$. The field $R/\mathfrak{p}(R)$ is the **residue field** of R.

The ring of formal power series $k[\![x, y]\!]$ is not a PID: the ideal (x, y) is nonprincipal. Recall that the ring $k[x, y]$ is Noetherian. We want to show that R is Noetherian means that $R[\![x]\!]$ is Noetherian. We copy the proof for polynomials—though it doesn't really work. Let $I \subset R[\![x]\!]$ be an ideal. We wrote ideals I_n generated by highest degree terms of polynomials $a_0 + \cdots + a_n x^n \in I$, but this doesn't work for power series. Instead, we let I_n be generated by a_n where $a_n x^n + a_{n+1}x^{n+1} + \cdots \in I$. We take powers of x less than n to be zero instead of taking powers of x greater than n to be zero. We get an increasing chain of ideals $I_0 \subset I_1 \subset I_2 \subset \cdots$. Then the proof is nearly identical—but why does this proof not work for polynomials?

We have also seen that if R is a UFD then $R[x]$ is a UFD. We want to show that $k[\![x, y]\!]$ is a UFD. One might attempt this by showing that if R is a UFD then $R[\![x]\!]$ is a UFD, but this is wrong.

Theorem 21.2 (Weierstrass preparation theorem)**.** *A power series $k[\![x_1, \ldots, x_n]\!]$ can be made to look like a polynomial in x_1.*

We want to prove this for $k[\![x, y]\!]$. Suppose $f \in k[\![x, y]\!]$. We can write $f = \text{unit} \times y^f \times (x^n + a_{n-1}x^{n-1} + \cdots + a_0)$ where the a_i are power series in y with zero constant term. The third term is called a Weierstrass polynomial. The idea is to kill off coefficients of f by multiplying it by suitable units (of the form $1 + \alpha x^m y^n$) so that we only have the x^n term. We multiply by some power of y such that a coefficient of some power of x is nonzero. Pick n to be the smallest power of x with a nonzero coefficient. We kill the coefficients in increasing order of powers of y, and in increasing order of powers of x therein. That is, we kill the coefficient of x^{n+1}, x^{n+2}, etc., then the coefficient of $x^{n+1}y$, $x^{n+2}y$, etc., then the coefficient of $x^{n+1}y^2$, $x^{n+2}y^2$, etc., etc.

Suppose we have killed all of the coefficients up to $x^{n+i}y^j$. Then we multiply f by $1 + \alpha x^i y^j$ where α is chosen such that the coefficient of x^n kills of the coefficient we wish to kill. This element is a unit. We need to check that an infinite product of things of the form $1 + \alpha x^m y^n$ converges, and it does.

Every power series is a product of irreducibles since the ring of power series is Noetherian. Then we need to show that if f is irreducible then f is prime. Suppose $f \mid gh$. We can assume that f, g, and h are Weierstrass polynomials (since we can harmlessly multiply by units and powers of y). So $fr = gh$ for some power series r. The key point is that r is a Weierstrass polynomial. Weierstrass polynomials are unique. (Exercise: Prove this.) If r is replaced by a Weierstrass polynomial, then fr is the Weierstrass polynomial of gh, so r must be a Weierstrass polynomial. We know that f, r, g, and h are in $k[\![y]\!][x]$, and f is irreducible in $k[\![y]\!][x]$, so f is prime in $k[\![y]\!][x]$ ($k[\![y]\!]$ is a UFD). So f divides g or h in $k[\![y]\!][x]$, so it must divide g or h in $k[\![x, y]\!]$.

Warning 21.3 (Traps)**.**—If $f \mid g$ in $k[\![x, y]\!]$, one cannot conclude that $f \mid g$ in $k[\![y]\!][x]$. (Consider $g = 1$ and $f = 1 + x$.) If f is irreducible in the polynomial ring, it does not need to be irreducible in the power series ring. (Consider $f = y^2 - x^2 - x^3$. We get $y^2 = x^2(1 + x)$, so $y^2 - (x\sqrt{1+x})^2 = (y + x\sqrt{1+x})(y - x\sqrt{1+x})$. The formal power series converges only at (an infinitesimally small neighbourhood of) 0, so the formal power series "sees" two branches.) Do not confuse $k[\![y]\!][x]$ and $k[x][\![y]\!]$ (or, for that matter, $k[\![y]\!][x]$ with $k[y][\![x]\!]$).

(Consider the element $1 + xy + x^2y^2 + \cdots$, or, for that matter, $1 + x + x^2 + \cdots$.) Also,

$$k[\![x, y]\!][x^{-1}, y^{-1}] \neq k[\![x]\!][x^{-1}][\![y]\!][y^{-1}] \neq k[\![y]\!][y^{-1}][\![x]\!][x^{-1}].$$

For instance, $(x - y)^{-1}$ is $1/x + y/x^2 + \cdots$ in one and $-1/y - x/y^2 + \cdots$ in another.

22 Hensel's lemma

Let $R = \mathbf{Z}_p$. Take $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Suppose α_0 is a root of f mod p: so $f(\alpha_0) \equiv 0 \pmod p$. Suppose $f'(\alpha_0) \not\equiv 0 \pmod p$. Then we can lift α_0 to a root α of f.

Example 22.1. Suppose we want to solve $x^2 = 7$ over $\mathbf{Z}_3$. First, we solve $\alpha_0^2 \equiv 7 \pmod 3$, so take $\alpha_0 = 1$. Then we lift to α_1 satisfying $\alpha_1^2 \equiv 7 \pmod{3^2}$. Then we write $\alpha_1 = \alpha_0 + 3b$. So

$$\begin{aligned} (\alpha_0 + 3b)^2 &\equiv 7 \pmod{3^2} \\ 1 + 6b + 3^2b^2 &\equiv 7 \pmod{3^2} \\ b &\equiv 1 \pmod{3^2}, \end{aligned}$$

so $b = 1$. Then we lift to α_2, so $\alpha_2^2 \equiv 7 \pmod{3^3}$, and $\alpha_2 = \alpha_1 + 3c$, so $2c \equiv * \pmod 3$. (One sees that $(d/dx)(x^2 - 7)(\alpha_0) = 2$, and we can only solve this at each step because $2 \not\equiv 0 \pmod 3$.)

Example 22.2. Suppose we want to solve $x^2 = 5$ in $\mathbf{Z}_2$. We see that $1^2 \equiv 5 \pmod 2$ and $1^2 \equiv 5 \pmod{2^2}$, but $\alpha^2 \equiv 5 \pmod{2^3}$ has no solutions. Notice that $(d/dx)(x^2 - 5) = 2x$, and this is $0 \pmod 2$ when $\alpha_0 = 1$.

One can prove Hensel's lemma by doing Example 22.1 with an arbitrary polynomial and showing that the coefficient is always $f'(\alpha_0)$, but we will prove it differently. We will show that Hensel's lemma is equivalent to Newton's method, which is given by the recursive relation

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Now we take x_0 to be a root of a function mod p and ensure that $f'(x_0) \not\equiv 0 \pmod p$. The relation above gives us

$$f(x_{n+1}) = f(x_n) - f'(x_n)\frac{f(x_n)}{f'(x_n)} + \frac{f''(x_n)}{2!}\left(\frac{f(x_n)}{f'(x_n)}\right)^2 - \cdots.$$

The first two terms cancel. Since $f'(x_n) \not\equiv 0 \pmod p$ and $f(x_n) \equiv 0 \pmod{p^f}$, the fraction $(f(x_n)/f'(x_n))^2$ is $0 \pmod{p^{2f}}$. (Don't worry about the 2!, 3!, etc. in the denominators:

$f^{(n)}(x_n)$ is always divisible by $n!$.) So if $f(x_n) \equiv 0 \pmod{p^k}$, then $f(x_{n+1}) \equiv 0 \pmod{p^{2k}}$. (So the number of correct p-adic digits doubles after each iteration; recall that something similar happens with Newton's method.)

Hensel's lemma applies to complete rings. Suppose we have the polynomial (that we will consider a power series) $y^2 - x^2 - x^3$. This factorizes as $(y+x)(y-x) - x^3$. By Hensel's lemmas for power series, that this factorizes means that we can lift it to a power series factorization: $(y - x\sqrt{1+x})(y + x\sqrt{1+x})$. What about $y^2 - x^3$? Again, the degree 2 terms factorize, giving $yy - x^3$. But there is no factorization. The condition that $f'(\alpha_0) \not\equiv 0$ means that α_0 is a simple root mod p, and this condition fails here.

Index