BORCHERDS LECTURES ON COMMUTATIVE ALGEBRA

LINA WEITZENBÖCK

ABSTRACT. These notes are a loose transcription of Richard Borcherds' online lecture series on Commutative Algebra, available on YouTube under https://youtube.com/playlist?list=PL8yHsr3EFj53rSexSz7vsYt-3rpHPR3HB&si=IjQa58LUzA7hKNQr.

CONTENTS

Part 1. Basic Notions, Invariants, and Noetherian Rings

1. Introduction

Commutative algebra serves as a framework for various fields in mathematics, sitting at the intersection of Algebraic Geometry, Algebraic Number Theory, and Invariant Theory. The level of this material assumes a first-year graduate (by American standards) background, particularly a working knowledge of basic commutative rings.

The philosophy adopted here is that commutative algebra is primarily a tool utilized in algebraic geometry and number theory, rather than a subject studied purely in isolation.

1.1. **Motivating Examples.** To illustrate the scope of commutative algebra, we consider typical examples of commutative rings from its three main areas of application.

1.1.1. *Number Theory.* In number theory, we study rings such as the integers $\mathbb{Z}$, and rings of integers in algebraic number fields. Examples include the Gaussian integers:

$$\mathbb{Z}[i] = \{m + ni \mid m, n \in \mathbb{Z}, i^2 = -1\}$$

and cyclotomic fields, such as the ring generated by a root of unity ζ, where $\zeta = e^{2\pi i/23}$.

A central question in commutative algebra regarding these rings is whether they form a *Unique Factorization Domain* (UFD), where every element factors uniquely into a product of primes (up to units and ordering). While $\mathbb{Z}$ and $\mathbb{Z}[i]$ are UFDs, Kummer demonstrated that the cyclotomic field generated by $e^{2\pi i/23}$ is not. To measure the obstruction to unique factorization, one studies the *Picard group*, denoted $\mathrm{Pic}(R)$, which classifies the invertible modules over the ring R.

1.1.2. *Algebraic Geometry.* A fundamental object of study is the *coordinate ring* of an algebraic variety, which is the ring generated by the coordinate functions of that variety. The simplest example is the affine plane over a field k, corresponding to the polynomial ring in two variables, $k[x, y]$.

For a more complex structure, consider an elliptic curve defined by the equation $y^2 = x^3 - x$. Its coordinate ring R is constructed as a quotient of the polynomial ring by the ideal generated by the defining equation:

$$R = \frac{k[x, y]}{(y^2 - x^3 + x)}$$

One may ask if R is a UFD. It is not, as it possesses a non-trivial Picard variety.

Commutative algebra establishes the precise relationship between the geometry of the curve and the algebraic structure of the ring. For instance, a geometric point on the elliptic curve corresponds to a homomorphism that evaluates the coordinate functions at that point:

$$\phi \colon R \to k$$

These homomorphisms naturally correspond to the maximal ideals of the coordinate ring R. Thus, points on the algebraic curve are intrinsically related to the ideals of the ring.

1.1.3. *Invariant Theory.* Historically, invariant theory provided much of the motivation for the development of commutative algebra. Consider an icosahedron embedded in $\mathbb{R}^3$. The symmetries of the icosahedron form a group G of order 120 (consisting of rotations and reflections). This group acts on the ambient space $\mathbb{R}^3$, and consequently on the polynomial ring $\mathbb{R}[x, y, z]$.

We are interested in the *invariant ring*, which consists of all polynomial functions fixed by the group action:

$$\mathbb{R}[x, y, z]^G = \{f \in \mathbb{R}[x, y, z] \mid g \cdot f = f \text{ for all } g \in G\}$$

An obvious invariant is the squared distance from the origin: $x^2 + y^2 + z^2$. Felix Klein demonstrated that the invariant ring for the icosahedron is generated by three polynomials a, b, and c, of degrees 2, 6, and 10, respectively.

A fundamental question in 19th-century invariant theory was whether the ring of invariants is always finitely generated; that is, whether there exists a finite set of invariant polynomials such that any other invariant can be expressed as a polynomial in them. We will address this by proving Hilbert's Basis Theorem, which establishes finite generation under broad conditions.

1.2. **Literature and References.** The primary textbook guiding this course is Eisenbud's "Commutative Algebra: with a View Toward Algebraic Geometry".

Other classic and advanced reference texts include:

- Atiyah and MacDonald, "Introduction to Commutative Algebra": A concise and standard introductory text.
- Zariski and Samuel, "Commutative Algebra" (Volumes 1 and 2): Classic foundational texts covering extensive technical details on valuation rings and power series.
- Serre, "Local Algebra": Highly recommended for its clarity.
- Nagata, "Local Rings": Contains a famous appendix of bad rings—counterexamples to properties one might naively hope hold for all commutative rings.
- Matsumura, "Commutative Ring Theory".

- Bourbaki, "Commutative Algebra": An encyclopedic reference covering the subject in maximum generality.
- Grothendieck, "Elements of Algebraic Geometry" (EGA): The monumental foundational text for modern algebraic geometry.
- Macaulay, "Algebraic Theory of Modular Systems": The earliest historical text on the subject (note that a *modular system* was Macaulay's terminology for an *ideal*).

2. Rings, Ideals, Modules

In this section, we review the basic definitions of rings, ideals, and modules. While this material is standard, it is important to clarify the exact conventions used in commutative algebra, as there are at least four inequivalent definitions of a ring circulating in the literature.

2.1. **Rings.**

Definition 2.1. A *ring* R is a set equipped with two binary operations, addition and multiplication, such that:

- R is an abelian group under addition with identity 0.
- R is associative under multiplication: $a(bc) = (ab)c$.
- R is distributive on the left and right: $a(b + c) = ab + ac$ and $(a + b)c = ac + bc$.

In the broader mathematical literature, there is disagreement on two further properties: whether multiplication must be commutative ($ab = ba$), and whether the ring must possess a multiplicative identity (an element 1 such that $1a = a1 = a$). Throughout this course, unless explicitly stated otherwise, all rings are assumed to be *commutative rings with an identity element*.

2.1.1. *Non-Commutative Rings.* Before restricting our attention entirely to commutative rings, it is instructive to recall several important examples of non-commutative rings:

- **Matrix Rings:** The ring $M_n(R)$ of $n \times n$ matrices with entries in a ring R, for instance, $M_2(\mathbb{Z})$.
- **Quaternions:** The ring $\mathbb{H}$ consisting of elements:

  $$a + bi + cj + dk$$

  where $a, b, c, d \in \mathbb{R}$, satisfying the relations:

  $$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, \quad \dots$$

- **Group Rings:** For a group G, the group ring $\mathbb{Z}[G]$ has a basis consisting of the elements $g \in G$. The ring multiplication is induced directly by the group product:

  $$(g_1)(g_2) = g_1 g_2$$

  This construction allows one to translate an enormous amount of group theory into the language of ring theory.

- **Rings of Differential Operators:** Consider operators acting on real functions on a line. An operator might take the form:

$$\sum a_{ij} x^i \left(\frac{d}{dx}\right)^j$$

where $a_{ij} \in \mathbb{R}$. These form a ring that is not commutative, as demonstrated by Leibniz's rule. If we let $A = \frac{d}{dx}$ and B be the operator representing multiplication by x, then applying them to a function f yields:

$$\frac{d}{dx}(xf) - x\left(\frac{d}{dx}f\right) = f$$
$$AB - BA = 1$$

Notice that the commutator $AB - BA = 1$ is, in some sense, "simpler" than A and B. In a commutative ring, $AB - BA = 0$, which is the ultimate simplification. Because the commutator in differential operator rings drops in degree or complexity, these rings are actually not too far removed from commutative rings, and techniques from commutative algebra often apply to them.
- **Non-Commutative Polynomial Rings:** Denoted $k\langle x, y\rangle$, where $xy \neq yx$. A basis consists of all arbitrary words in x and y (e.g., $1, x, y, x^2, xy, yx, y^2, \dots$).
- **Other Examples:** Universal enveloping algebras of Lie algebras, and Clifford algebras.

2.1.2. *Rings Without Identity.* Why do some authors study rings without an identity element? The motivation often stems from functional analysis and topology.

If R is a ring without an identity, one can always formally adjoin one by taking the direct sum $\mathbb{Z} \oplus R$ (or $\mathbb{R} \oplus R$ in analytical contexts) and defining multiplication appropriately, essentially forcing the element $(1, 0)$ to act as the identity.

However, in analysis, consider the convolution algebra of functions on a locally compact group, such as the reals $\mathbb{R}$. The convolution product is given by:

$$(f * g)(b) = \int_{\mathbb{R}} f(a)g(b-a)\,da$$

For the integral to converge generally, one restricts to continuous functions with compact support. This ring, however, lacks an identity element, because an identity under convolution would have to act like a Dirac delta "function" at the origin, which is not continuous.

Similarly, consider the ring $C_0(X)$ of continuous functions vanishing at infinity on a locally compact topological space X. This ring lacks an identity unless X is compact (since the constant function 1 does not vanish at infinity).

Adjoining an identity to $C_0(X)$ corresponds geometrically to taking the *one-point compactification* of X. Since drastically different locally compact spaces can share identical one-point compactifications, forcibly adjoining an identity to a ring can result in a severe loss of geometric and topological information.

Despite these analytical considerations, in commutative algebra, we demand that all rings have an identity 1. Consequently, a *ring homomorphism* $\phi\colon R \to S$

must explicitly preserve this identity:

$$\phi(1_R) = 1_S$$

Failure to enforce this condition leads to pathological homomorphisms, such as the trivial map $n \mapsto 0$ from $\mathbb{Z} \to \mathbb{Z}$, which is a valid homomorphism for rings without identity, but strictly forbidden when identities are enforced.

2.2. **Ideals.**

Definition 2.2. An *ideal* I of a commutative ring R (with identity) is defined as the kernel of a ring homomorphism from R to another ring S. Equivalently, a subset $I \subseteq R$ is an ideal if it satisfies the following two conditions:

$$a, b \in I \implies a \pm b \in I$$
$$a \in I,\ r \in R \implies ra \in I$$

It is crucial to note that I must be closed under multiplication by *arbitrary* elements $r \in R$, not merely elements within I. For instance, $\mathbb{Z}$ is a subring of $\mathbb{Q}$ and is closed under multiplication, but it is not an ideal of $\mathbb{Q}$, because multiplying an integer by an arbitrary rational number does not generally yield an integer.

(Incidentally, an ideal satisfies all axioms of a ring except for the possession of an identity element, which is why early algebraists sometimes allowed rings without identities).

Important examples of ideals and their corresponding quotient rings include:

- In $\mathbb{Z}$, the subset $n\mathbb{Z}$ of all multiples of n is an ideal. The quotient is the ring of integers modulo n.
- In the polynomial ring $\mathbb{R}[x]$, the ideal generated by $x^2 + 1$ yields the complex numbers:

$$\frac{\mathbb{R}[x]}{(x^2+1)} \cong \mathbb{C}$$

- In $k[x, y]$, the quotient by the principal ideal $(y^2 - x^3 + x)$ represents the ring of polynomial functions restricted to the corresponding elliptic curve. This operation of quotienting by ideals is the primary algebraic mechanism for studying functions on subsets defined by polynomial equations.

2.3. **Modules.**

Modules are the ring-theoretic analogue of vector spaces over fields.

Definition 2.3. A *module* M over a ring R is an abelian group equipped with a scalar multiplication map $R \times M \to M$, denoted $(r, m) \mapsto rm$, satisfying the following axioms for all $r, r_1, r_2 \in R$ and $m, m_1, m_2 \in M$:

$$r(m_1 + m_2) = rm_1 + rm_2$$
$$(r_1 + r_2)m = r_1m + r_2m$$
$$(r_1r_2)m = r_1(r_2m)$$
$$1m = m$$

The final axiom, $1m = m$, is absolutely crucial and explicitly relies on our assumption that the ring possesses an identity element.

A century ago, commutative algebra focused predominantly on ideals (then called "modular systems"). Today, modules are widely preferred because they offer significantly greater structural flexibility.

Fundamental examples of modules include:

- **$\mathbb{Z}$-modules:** Modules over the integers are precisely abelian groups. A major goal in commutative algebra is to generalize structure theorems for abelian groups to modules over more complex rings (e.g., the Lasker-Noether theorem serves as a generalization of the fundamental theorem of finitely generated abelian groups).
- **k-modules:** Where k is a field, a k-module is simply a standard vector space.
- **$k[x]$-modules:** A module over a polynomial ring $k[x]$ requires a vector space over k, equipped with an action by the element x. This action is simply a linear transformation. Thus, the entire theory of linear transformations is subsumed by the theory of modules over $k[x]$.
- **Submodules of R:** The ring R is a module over itself. Its submodules are exactly the ideals of R.
- **Quotient Modules:** If I is an ideal of R, the quotient ring R/I is an R-module (generated by a single element). Notice that you can reconstruct the ideal I from the module $M = R/I$ by considering its *annihilator*:

$$\mathrm{Ann}(M) = \{r \in R \mid rm = 0 \text{ for all } m \in M\} = I$$

The flexibility of modules is best illustrated by taking quotients. If M is a submodule of N, the quotient N/M is always a well-defined module. In contrast, if I and J are ideals in R with $I \subseteq J$, the quotient J/I is an R-module, but it is *not* an ideal of R.

2.3.1. *Analogy Between Groups and Rings.* To conclude, there is a profound structural analogy between group theory and commutative ring theory:

- Groups act on sets, whereas rings act on modules.
- Normal subgroups ($N \trianglelefteq G$) allow for the formation of quotient groups G/N. Analogously, ideals ($I \subseteq R$) allow for the formation of quotient rings R/I.
- If a set S is acted upon by a group G, one can form a $\mathbb{Z}$-module with basis S, which naturally becomes a module over the group ring $\mathbb{Z}[G]$.

In this light, many concepts in ring theory can be viewed as vast generalizations of corresponding concepts in group theory.

3. What is a Syzygy?

The word "syzygy" originates from a Greek term for a yoke, traditionally used to tie oxen together. In commutative algebra, a syzygy refers to an algebraic relation between generators of a module or an algebra, effectively "yoking" them together. We introduce this concept through the study of invariant rings, which historically motivated Hilbert's theorems on finite generation.

3.1. **Invariant Rings and Group Actions.** Suppose a group G acts on a vector space V over a field k. We are primarily interested in the *invariant ring* (or algebra of invariants), which consists of polynomial functions on V that remain constant under the action of G.

To formalize this, we must specify how G acts on the polynomial functions $f\colon V \to k$. For $g \in G$, the action is defined by:

$$(g \cdot f)(x) = f(g^{-1} \cdot x)$$

The inclusion of the inverse g^{-1} is necessary to satisfy the associativity axiom of a left group action, namely $(g_1 g_2) \cdot f = g_1 \cdot (g_2 \cdot f)$. Defining the action as $f(g \cdot x)$ would incorrectly yield a right group action, leading to contradictory evaluations.

3.2. **Examples of Invariants.**

3.2.1. *Rotations.* Consider the orthogonal group $O_3(\mathbb{R})$ acting on $\mathbb{R}^3$. Orthogonal transformations preserve Euclidean length. Consequently, the squared distance from the origin is an invariant polynomial:

$$f(x, y, z) = x^2 + y^2 + z^2$$

3.2.2. *Determinants.* Let the special linear group $SL_n(k)$ act on the direct sum of n copies of k^n:

$$V = k^n \oplus k^n \oplus \cdots \oplus k^n$$

By definition, elements of $SL_n(k)$ preserve the volume form. Thus, taking the determinant of the $n \times n$ matrix formed by n vectors in V yields a natural invariant polynomial.

3.2.3. *The Symmetric Group.* Let the symmetric group S_n act on $\mathbb{C}^n$ by permuting the coordinates $x_1, \ldots, x_n$. The invariant polynomials are precisely the symmetric polynomials. A standard algebraic basis of invariants is given by the elementary symmetric functions:

$$\begin{aligned} e_1 &= \sum_i x_i \\ e_2 &= \sum_{i<j} x_i x_j \\ &\cdots \\ e_n &= x_1 x_2 \cdots x_n \end{aligned}$$

The fundamental theorem of symmetric polynomials states that every S_n-invariant polynomial can be expressed as a unique polynomial in $e_1, \ldots, e_n$. Therefore, the invariant ring is itself a polynomial ring over $\mathbb{C}$, freely generated by $e_1, \ldots, e_n$ without any algebraic relations. This unconstrained structure is characteristic of reflection groups but is generally rare.

3.3. **First-Order Syzygies.**

3.3.1. *The Alternating Group.* Consider the alternating group $A_n \subset S_n$ acting on $\mathbb{C}^n$. Define the polynomial:

$$\Delta = \prod_{i<j}(x_i - x_j)$$

Elements of S_n permute the factors of Δ up to a sign. Specifically, A_n is the index-two subgroup of S_n that fixes Δ. Consequently, the invariant ring for A_n is finitely generated by $e_1, \ldots, e_n$ and Δ.

However, these generators are no longer algebraically independent. The square Δ^2 is manifestly symmetric and must therefore be expressible as a polynomial P in the elementary symmetric functions $e_1, \ldots, e_n$. This yields a non-trivial algebraic relation:

$$\Delta^2 - P(e_1, \ldots, e_n) = 0$$

For instance, when $n = 2$, $\Delta = x_1 - x_2$, and the relation takes the form $\Delta^2 = e_1^2 - 4e_2$. An algebraic relation between the generators of a ring is called a *first-order syzygy*. In the case of A_n, this is essentially the fundamental relation from which all other relations are generated.

3.4. **Higher-Order Syzygies.** To observe higher-order syzygies, consider the cyclic group C_3 of order 3, generated by an element σ such that $\sigma^3 = 1$. Let C_3 act on $\mathbb{C}^2$ by:

$$\sigma \cdot (x, y) = (\omega x, \omega y)$$

where $\omega = e^{2\pi i/3}$. A polynomial monomial $x^a y^b$ is invariant if and only if $a+b \equiv 0 \pmod 3$. The invariant algebra is generated by four basic monomials:

$$\begin{aligned} z_0 &= x^3 \\ z_1 &= x^2 y \\ z_2 &= xy^2 \\ z_3 &= y^3 \end{aligned}$$

These generators satisfy three immediate first-order syzygies:

$$\begin{aligned} a_1 &= z_2^2 - z_1 z_3 = 0 \\ a_2 &= z_0 z_3 - z_1 z_2 = 0 \\ a_3 &= z_1^2 - z_0 z_2 = 0 \end{aligned}$$

These syzygies, however, are not algebraically independent. We can construct a relation between the relations:

$$\begin{aligned} z_1 a_1 + z_2 a_2 + z_3 a_3 &= z_1(z_2^2 - z_1 z_3) + z_2(z_0 z_3 - z_1 z_2) + z_3(z_1^2 - z_0 z_2) \\ &= z_1 z_2^2 - z_1^2 z_3 + z_2 z_0 z_3 - z_1 z_2^2 + z_3 z_1^2 - z_3 z_0 z_2 \\ &= 0 \end{aligned}$$

This linear dependence between the first-order syzygies over the invariant ring is called a *second-order syzygy*.

3.5. **Free Resolutions and Finiteness.** We can frame the structure of syzygies in terms of module homomorphisms. Let $R = k[z_0, \dots, z_n]$ be a formal polynomial ring representing the generators. Mapping these formal variables to the actual generators of the invariant ring R^G gives a surjective k-algebra homomorphism:

$$R \twoheadrightarrow R^G$$

The kernel of this map is an ideal $I \subset R$ generated by the first-order syzygies (e.g., $a_1, \dots, a_m$). We can define a map from a free module R^m mapping basis elements to these generators:

$$R^m \to R$$

The image is I. The kernel of this second map consists precisely of the second-order syzygies. Iterating this process constructs a sequence of free R-modules, known as a free resolution:

$$\cdots \to R^p \to R^m \to R \to R^G \to 0$$

This exact sequence naturally raises three important finiteness questions in commutative algebra:

(1) Is the invariant ring R^G finitely generated as a k-algebra?
(2) Are the successive modules of syzygies (e.g., R^m, R^p) finitely generated as R-modules?
(3) Is this chain of free modules finite in length?

Hilbert established affirmative answers to these questions provided the base field k has characteristic zero and the group G is finite (or, more generally, reductive), laying the groundwork for modern commutative algebra.

4. Invariant Theory

In this section, we provide a historical summary of David Hilbert's work in invariant theory, contrasting it with the explicit, often arduous, computational methods of the 19th century.

4.1. **Binary Quantics.** A *binary quantic* is a homogeneous polynomial in two variables of arbitrary degree. (For three or four variables, the terms are *ternary* and *quaternary*, respectively; specific degrees are denoted as *quadric* for degree two, *cubic* for degree three, and so on). A general binary quantic of degree n takes the form:

$$f(x, y) = a_n x^n + a_{n-1} x^{n-1} y + \cdots + a_0 y^n$$

By convention, one might insert binomial coefficients $\binom{n}{k}$ before each term a_k.

The group $SL_2(\mathbb{C})$ acts on the $(n+1)$-dimensional basis space $\mathbb{C}^{n+1}$ of coefficients $(a_n, \dots, a_0)$. The action on the coordinates x and y is given by the standard linear transformation:

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

Substituting this into the binary quantic yields a highly complex expression for the transformed coefficients. The central problem of invariant theory is to find

polynomial functions in the coefficients $a_0, \ldots, a_n$ that remain invariant under this action.

4.1.1. *Examples of Invariants.*

- **Discriminant:** For a binary quadratic $a_2x^2+a_1xy+a_0y^2$, the discriminant is a familiar invariant:

$$\Delta = a_1^2 - 4a_2a_0$$

  More generally, the discriminant is proportional to the squared product of the differences of the roots of the polynomial.
- **Catalecticant:** For a degree four polynomial, an invariant introduced by Sylvester (who coined many such terms) is the determinant of a matrix formed by the coefficients, often written with binomial weights.

Paul Gordan, known as the "king of invariant theory", proved that the invariants of binary quantics are finitely generated.

4.2. **The Computational Complexity of Invariants.** While invariants can be proven to be finitely generated, explicitly computing them is exceptionally difficult. For example, consider *ternary cubics* (degree three, three variables x, y, z). There are 10 possible monomials:

$$x^3, x^2y, xy^2, y^3, x^2z, xyz, y^2z, xz^2, yz^2, z^3$$

This defines a 10-dimensional space of coefficients acted upon by $SL_3(\mathbb{C})$.

The invariant ring is generated by two invariants of degrees 4 and 6. The degree 4 invariant spans several lines of text, while the degree 6 invariant is extraordinarily dense, traditionally requiring the solution of hundreds of linear equations. Mathematicians in the 19th century, such as George Salmon, computed even larger invariants entirely by hand, engaging in informal contests to derive the most complex explicit expressions.

4.3. **Hilbert's Basis Theorem and Syzygies.** Given the immense complexity of computing invariants, establishing finite generation by explicit construction for arbitrary degrees and variables was impossible. The problem becomes even more severe when considering *syzygies*—the algebraic relations between invariants.

If the invariants are complicated, the first-order syzygies are far worse, and second-order syzygies (relations between relations) are worse still. Hilbert's spectacular conceptual breakthrough was to reverse the problem. He demonstrated that if invariants are finitely generated, then the syzygies of all orders are also finitely generated.

To prove this, Hilbert showed that for a polynomial ring $k[x_1, \ldots, x_n]$, any ideal is finitely generated. Rings satisfying this condition are now called *Noetherian rings*, named after Emmy Noether, who systematically abstracted and generalized Hilbert's polynomial ring theorems to arbitrary rings with this property.

4.4. **Three Meanings of "Finitely Generated".** When discussing finite generation, one must distinguish between three different contexts:

(1) **As a Module (or Ideal):** Generated by linear combinations over the base ring. For example, the ideal $(x, y) \subset k[x, y]$ is finitely generated as a module over $k[x, y]$.
(2) **As an Algebra (Ring):** Generated by polynomial expressions over the base ring. A polynomial ring $k[x]$ is finitely generated as a k-algebra by the element x, but it is *not* finitely generated as a k-module (it requires an infinite basis $1, x, x^2, \dots$).
(3) **As a Field:** Generated by rational functions. The field of rational functions $k(x)$ is finitely generated as a field over k (by x), but not as a k-algebra, because an infinite number of inverses (e.g., $1/(x-1)$) cannot be generated by finitely many polynomials.

4.5. **Extensions and Counterexamples.** Hilbert proved that invariant rings $k[x_1, \dots, x_n]^G$ are finitely generated as algebras when G is reductive (and the characteristic of k is zero).

Hilbert's 14th problem asked if this finite generation holds for *all* groups G. Decades later, Masayoshi Nagata found a counterexample: a specific group action for which the invariant ring is *not* finitely generated.

Furthermore, while all ideals in a polynomial ring with finitely many variables are finitely generated, this fails for infinitely many variables. For instance, in $k[x_1, x_2, \dots]$, the ideal generated by all polynomials with a constant term of zero has no finite generating set.

4.6. **Finite Free Resolutions.** Let R be the ring of invariants. We can express R as a quotient of a polynomial ring by mapping generators onto R. The kernel of this map represents the first-order syzygies, which form an ideal. By Hilbert's theorem, this ideal is finitely generated, leading to a map from a free module onto these syzygies.

Iterating this process yields an exact sequence of free modules, known as a *finite free resolution*:

$$\cdots \to F_2 \to F_1 \to F_0 \to M \to 0$$

Hilbert demonstrated that over a polynomial ring $k[x_1, \dots, x_n]$, every finitely generated module M possesses a finite free resolution.

However, this property does not hold for all Noetherian rings. Consider the ring $R = k[x]/(x^2)$ and the module $M = R/(x) \cong k$. Attempting to resolve M yields:

$$\dots \xrightarrow{x} R \xrightarrow{x} R \xrightarrow{x} R \to M \to 0$$

Because the kernel of multiplication by x in R is exactly the ideal (x), the resolution repeats infinitely and never terminates.

5. Noetherian Rings

In this section, we formally define Noetherian rings, establish several equivalent conditions, and explore a variety of examples and counterexamples. These concepts are essential for the subsequent proof of Hilbert's Basis Theorem.

5.1. **Equivalent Conditions for Noetherian Rings.** Recall from the previous lecture that a ring R is defined to be Noetherian if every ideal of R is finitely generated. This definition is equivalent to two other important structural conditions.

Theorem 5.1. *For a commutative ring R, the following conditions are equivalent:*

1. *R is a* Noetherian ring.
2. *Every ideal $I \subseteq R$ is finitely generated.*
3. *Every strictly increasing chain of ideals is finite (Ascending Chain Condition). That is, any chain*
$$I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \dots$$
must terminate.
4. *Every non-empty set of ideals in R has a maximal element (with respect to inclusion).*

Proof. The equivalence (1) $\iff$ (2) is true by definition. We proceed to prove (2) $\implies$ (3) $\implies$ (2) and (3) $\iff$ (4).

(2) $\implies$ (3)**:** Suppose all ideals are finitely generated. Given an ascending chain $I_1 \subseteq I_2 \subseteq \dots$, let
$$I = \bigcup_{n=1}^{\infty} I_n$$
Because the sequence is increasing, I is an ideal. By assumption, I is finitely generated by some elements $a_1, \dots, a_k$. Since all $a_i \in I$, there must exist some index N such that $a_i \in I_N$ for all $i = 1, \dots, k$. Consequently,
$$I = I_N$$
Thus, the chain stabilizes at I_N, meaning any strictly increasing chain must be finite.

(3) $\implies$ (2)**:** Assume every strictly increasing chain is finite. Suppose, for contradiction, there exists an ideal I that is not finitely generated. We can construct an infinite strictly increasing sequence of finitely generated ideals contained in I. Choose $a_1 \in I \setminus \{0\}$. Since $I \neq (a_1)$, choose $a_2 \in I \setminus (a_1)$. Continuing this process, if $I \neq (a_1, \dots, a_n)$, we can always choose $a_{n+1} \in I \setminus (a_1, \dots, a_n)$. This yields a strictly increasing chain:
$$(0) \subsetneq (a_1) \subsetneq (a_1, a_2) \subsetneq \dots$$
This contradicts the assumption that strictly increasing chains are finite. Thus, I must be finitely generated.

(3) $\iff$ (4)**:** This equivalence is a general property of partially ordered sets (posets). Let S be a non-empty set of ideals.

- (3) $\implies$ (4)**:** If S has no maximal element, then for any $I_1 \in S$, there exists $I_2 \in S$ such that $I_1 \subsetneq I_2$. Repeating this choice (invoking the Axiom of Choice) constructs an infinite strictly increasing chain, violating (3).
- (4) $\implies$ (3)**:** Given a chain $I_1 \subseteq I_2 \subseteq \dots$, the set of these ideals $\{I_n\}$ must have a maximal element I_N by (4). Therefore, for all $m \geq N$, $I_m = I_N$, and the chain is finite.

□

5.2. **A Non-Noetherian Example.** To see how these conditions fail, consider the polynomial ring in infinitely many variables $R = k[x_1, x_2, \dots]$.

- The ideal $(x_1, x_2, \dots)$ is not finitely generated.
- There exists an infinite strictly increasing chain:

$$(x_1) \subsetneq (x_1, x_2) \subsetneq (x_1, x_2, x_3) \subsetneq \dots$$

- The set $S = \{(x_1, \dots, x_n) \mid n \in \mathbb{N}\}$ has no maximal element.

Remark 5.2. The Noetherian property concerns *ascending* chains. The analog for *descending* chains—where every strictly decreasing chain of ideals terminates—defines an *Artinian ring*, which is a much stronger condition. For instance, the integers $\mathbb{Z}$ form a Noetherian ring, but it is not Artinian, as demonstrated by the infinite descending chain:

$$(2) \supsetneq (4) \supsetneq (8) \supsetneq (16) \supsetneq \dots$$

5.3. **Examples of Rings of Functions.** Determining whether a ring is Noetherian can be subtle. Consider the following sequence of rings of functions, which alternate between Noetherian and non-Noetherian:

(1) **Polynomials $\mathbb{R}[x]$:** *Noetherian.* This is a Principal Ideal Domain (PID); every ideal is generated by a single element.

(2) **Analytic functions on $\mathbb{R}$:** *Not Noetherian.* Let $Z = \{1, 2, 3, \dots\}$ be a set of points with no limit point. The ideal of functions vanishing at all but finitely many points of Z is not finitely generated.

(3) **Analytic functions on $[-1, 1]$:** *Noetherian.* An analytic function on a compact interval has only finitely many zeros. Any such function f can be written as:

$$f = P \cdot u$$

where P is a polynomial and u is a unit (an analytic function with no zeros). Because it is so closely related to $\mathbb{R}[x]$, all ideals are finitely generated.

(4) **Analytic functions on $(-1, 1)$:** *Not Noetherian.* We can choose an infinite sequence of zeros with no limit point inside the open interval, such as $1 - \frac{1}{n}$. The same construction as in (2) yields an ideal that is not finitely generated.

(5) **Analytic at 0 (Germs of holomorphic functions):** *Noetherian.* Any such function can be factored as $x^n \cdot u$, where $u(0) \neq 0$ (so u is a unit). The only non-zero ideals are of the form (x^n), making this a Discrete Valuation Ring (DVR).

(6) **Smooth functions near 0:** *Not Noetherian.* Consider a flat function with an infinite-order zero at the origin, such as:

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

We can form an infinite strictly increasing chain of ideals by taking fractional powers:

$$(f) \subsetneq (f^{1/2}) \subsetneq (f^{1/4}) \subsetneq \dots$$

(7) **Formal power series $\mathbb{R}[[x]]$:** *Noetherian.* Similar to analytic functions at 0, any formal power series can be written as $x^n \cdot u$ where u is invertible. It is a DVR.

5.4. **Properties of Noetherian Rings.**

Proposition 5.3. *If R is a Noetherian ring and I is an ideal, then the quotient ring R/I is Noetherian.*

Proof. By the correspondence theorem, ideals of R/I are in bijection with ideals of R containing I. Since any set of ideals in R has a maximal element, the same holds for the restricted set of ideals containing I. Thus, R/I satisfies the maximal condition and is Noetherian. □

A direct consequence is that if S is a Noetherian ring, any finitely generated algebra over S is also Noetherian, because it can be written as a quotient of a polynomial ring $S[x_1, \dots, x_n]$. This applies to coordinate rings of algebraic varieties and rings of integers in number fields.

Remark 5.4. A subring of a Noetherian ring is *not* necessarily Noetherian. For example, the non-Noetherian ring of analytic functions on $\mathbb{R}$ is a subring of the Noetherian formal power series ring. Another trivial example is any non-Noetherian integral domain embedded in its field of fractions (all fields are trivially Noetherian).

5.5. **Generators in Multivariable Polynomial Rings.** While every ideal in $k[x]$ is generated by 1 element, it is natural to ask if every ideal in $k[x, y]$ is generated by 2 elements. The answer is no. Consider the ideal $I \subset k[x, y]$ generated by all monomials of degree 3:

$$I = (y^3, y^2x, yx^2, x^3)$$

This ideal requires at least 4 generators. If we consider I modulo the degree 4 polynomials, it forms a 4-dimensional vector space, meaning any generating set must contain at least 4 elements. This generalizes: one can construct ideals in $k[x, y]$ requiring an arbitrarily large number of generators.

5.6. **Puiseux Series.** As a final example, consider the ring of *Puiseux series*, which extends formal power series by allowing fractional exponents with a common denominator. It is the union:

$$R = \bigcup_{n=1}^{\infty} k[[x^{1/n}]]$$

Despite its similarity to $k[[x]]$, this ring is *not* Noetherian. It contains an infinite strictly increasing chain of ideals:

$$(x) \subsetneq (x^{1/2}) \subsetneq (x^{1/4}) \subsetneq \dots$$

6. Proof of Hilbert's Basis Theorem

In this section, we present the proof of Hilbert's Basis Theorem, establishing that ideals in polynomial rings over a Noetherian ring are finitely generated. We begin with a modernized, streamlined proof before investigating an adaptation for formal power series, and finally discussing an alternative approach grounded in Dickson's Lemma and Gröbner bases.

6.1. **The Standard Proof.** The property that all ideals are finitely generated is equivalent to the ascending chain condition: all chains of ideals $I_1 \subseteq I_2 \subseteq I_3 \subseteq \dots$ are eventually constant. Hilbert's original proof focused directly on polynomials, but we can prove a more general result.

Theorem 6.1 (Hilbert's Basis Theorem)**.** *If R is a* Noetherian ring*, then the polynomial ring $R[x]$ is Noetherian.*

Since fields and the ring of integers $\mathbb{Z}$ are Noetherian, induction on the number of variables immediately yields that polynomial rings in finitely many variables over these rings are also Noetherian.

Proof. Suppose I is an ideal of $R[x]$. We will construct a sequence of ideals in the base ring R. Let I_k be the ideal consisting of 0 and the leading coefficients of all polynomials in I of degree exactly k.

It is clear that I_k is an ideal of R, as we can add polynomials of the same degree and multiply by elements of R without increasing the degree. Furthermore, we have an ascending chain of ideals:

$$I_0 \subseteq I_1 \subseteq I_2 \subseteq \dots$$

This inclusion holds because if we multiply a polynomial of degree k by x, we obtain a polynomial of degree $k+1$ with the exact same leading coefficient.

Because R is Noetherian, this chain of ideals must eventually stabilize. Thus, there exists some integer n such that:

$$I_n = I_{n+1} = I_{n+2} = \dots$$

Additionally, since R is Noetherian, each ideal I_k is finitely generated. We construct a finite set of generators for I as follows:

- For each $k \leq n$, let S_k be a finite set of polynomials of degree k in I whose leading coefficients generate I_k.

Let S be the union of these sets:

$$S = \bigcup_{k=0}^{n} S_k$$

We claim that the finite set S generates the ideal I. Let $f \in I$ be a polynomial of degree m with leading coefficient $a_m \in I_m$.

- If $m \leq n$, there exists a polynomial g in the ideal generated by S_m (and thus by S) with the same degree m and the same leading coefficient a_m.
- If $m > n$, we note that $a_m \in I_m = I_n$. Thus, we can find a polynomial h in the ideal generated by S_n with degree n and leading coefficient a_m. Multiplying h by x^{m-n} yields a polynomial $g = x^{m-n}h$ with degree m and leading coefficient a_m.

In either case, the difference $f - g$ lies in I and has a strictly smaller degree than f. By iteratively repeating this process, the degree strictly decreases at each step until we reach 0. This expresses f as a linear combination of the elements of S, completing the proof. □

While this proof establishes the existence of a finite generating set, it is largely non-constructive. Finding the integer n where the sequence of ideals stabilizes involves checking potentially infinitely many cases, a problem inherent not just to polynomial rings, but to finding generators of ideals even in $\mathbb{Z}$.

6.2. **Formal Power Series.** A variation of Hilbert's approach applies to formal power series.

Theorem 6.2. *If R is a Noetherian ring, then the ring of formal power series $R[[x]]$ is Noetherian.*

Proof. We adapt the proof for polynomial rings by turning it "upside down". Because a power series has no highest degree term, we define I_k based on the *lowest* degree terms.

Let I_k be the ideal in R generated by 0 and the coefficients of x^k for series in I whose lowest non-zero term has degree k. As before, $I_k \subseteq I_{k+1}$ (by multiplying a series by x), and the chain stabilizes at some integer n.

We select finite sets of power series S_k whose lowest degree coefficients generate I_k. To express an arbitrary $f \in I$ in terms of $S = \bigcup S_k$, we subtract multiples of elements in S to eliminate the lowest degree term of f, iteratively increasing the order of the lowest term.

The conceptual difference is that this process does not terminate in finitely many steps; it requires an infinite sum to kill off all terms. However, in the ring of formal power series $R[[x]]$, an infinite series of the form

$$\sum_{j=0}^{\infty} r_j x^j s_j$$

converges to a well-defined element, meaning the infinite sum of elements from the ideal generated by S combines into a single valid formal power series. □

6.3. **Dickson's Lemma and Gordan's Proof.** The original proofs by Hilbert and Gordan were constructed somewhat differently. Instead of proceeding by induction on the ring, they often worked directly with the monomials. This requires an essential combinatorial result.

Lemma 6.3 (Dickson's Lemma)**.** *Any set of monomials in a polynomial ring over finitely many variables has only a finite number of* minimal elements *under the partial order of divisibility.*

To sketch the proof for two variables, we can arrange the monomials $x^a y^b$ in a two-dimensional grid. Selecting a minimal element $x^{a_1} y^{b_1}$ eliminates a semi-infinite rectangular quadrant of multiples. Any subsequent minimal elements must lie in the remaining "L-shaped" region, which consists of strips of finite width. Each new minimal element effectively shrinks the width of these strips. Because the strips have finite width, they can only be shrunk a finite number of times before all minimal elements are exhausted. (Cf. the lecture recording for an illustration of this.)

We can use Dickson's Lemma to prove the finite generation of polynomial ideals directly.

Theorem 6.4. *Every ideal I in $k[x_1, \ldots, x_n]$ is finitely generated.*

Proof. First, we order the monomials using a total ordering, such as the lexicographic order. We define the relation such that:

$$x^a y^b > x^c y^d \quad \text{if } a > c, \text{ or if } a = c \text{ and } b > d$$

This is a well-ordering, meaning there are no infinite strictly decreasing chains of monomials.

For any polynomial, we define its *leading term* as its maximal monomial with respect to this lexicographic order. Let $L(I)$ be the set of all leading terms of all polynomials in the ideal I.

By Dickson's Lemma, $L(I)$ has only a finite number of minimal elements under divisibility. For each of these minimal leading terms, we select a corresponding polynomial in I. This yields a finite set of polynomials S.

We claim S generates I. If $f \in I$, its leading term must be divisible by the leading term of some $s \in S$ (since we selected all minimal elements). By subtracting an appropriate monomial multiple of s from f, we strictly decrease the leading term in the lexicographic order. Because the lexicographic order is a well-order, this process must terminate, reducing f to 0 in finitely many steps. Thus, f is a combination of elements in S. □

Remark 6.5. This procedure is inherently more constructive than the standard proof and forms the basis for the theory of *Gröbner bases*. While we utilized lexicographic ordering here, other total well-orderings (such as graded lexicographic order) are also permissible and often more computationally efficient.

7. Finite Generation of Invariants

In this section, we present the proof of Hilbert's Theorem on the finite generation of the algebra of invariants. We suppose that G is a finite group acting on a

finite-dimensional vector space V over a field k of characteristic zero. Our goal is to show that the corresponding algebra of invariants is finitely generated as a k-algebra.

7.1. **Setup and the Algebra of Invariants.** Let the dimension of our vector space V be n. The algebra of all polynomial functions on V can be identified with the polynomial ring R:

$$R = k[x_1, \dots, x_n]$$

The group G acts on R by linear substitutions. We are interested in the polynomials that are fixed by this action.

Definition 7.1. The *algebra of invariants*, denoted R^G (or I), is the sub-algebra of R consisting of all polynomials f such that $g \cdot f = f$ for all $g \in G$:

$$I = R^G = \{f \in R \mid g \cdot f = f \text{ for all } g \in G\}$$

A key structural property of the polynomial ring R is that it is graded by degree. We can assign each variable x_i a degree of 1, such that any monomial has a well-defined total degree:

$$\deg(x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}) = m_1 + m_2 + \cdots + m_n$$

Thus, we can decompose R into a direct sum of its homogeneous components R_d of degree d:

$$R = R_0 \oplus R_1 \oplus R_2 \oplus \dots$$

where $R_0 = k$ and R_1 is spanned by $x_1, \dots, x_n$.

Because the group action is linear, it preserves the degree of homogeneous polynomials. Consequently, the invariant sub-algebra I is also graded:

$$I = I_0 \oplus I_1 \oplus I_2 \oplus \dots$$

where $I_0 = k$. Graded rings are highly advantageous because they allow us to prove structural theorems using induction on the degree.

7.2. **Hilbert's Strategy and a Cautionary Example.** To prove that I is finitely generated as a k-algebra, we first construct an ideal in R. Let J be the ideal of R generated by all homogeneous invariants of strictly positive degree:

$$J = \langle I_1 \oplus I_2 \oplus I_3 \oplus \dots \rangle$$

Note that we exclude $I_0 = k$; otherwise, J would trivially be the entire ring R, as it would contain the identity element 1.

By Hilbert's Basis Theorem, R is a Noetherian ring, which immediately implies that the ideal J is finitely generated. We can choose a finite set of homogeneous invariant generators $a_1, a_2, \dots, a_k \in I$ for the ideal J.

We now wish to show that these elements $a_1, \dots, a_k$ generate I as a k-algebra. At this point, extreme care must be taken to distinguish between two concepts:

- **Generating J as an ideal over R:** Every element in J is a linear combination of the a_i with coefficients in the large ring R.
- **Generating I as an algebra over k:** Every element in I is a polynomial expression in the a_i with coefficients in the base field k.

In general, generators of an ideal do *not* translate to generators of a sub-algebra. Consider a generic sub-algebra generated by monomials such as y, y^2, y^3, xy, x^2y. If we form an ideal J from the positive-degree elements of this sub-algebra in $k[x, y]$, J might simply be generated by the single element x. However, the original sub-algebra is clearly not generated by x alone as an algebra, and may not even be finitely generated at all.

Therefore, to prove that I is finitely generated, we must use a special property unique to rings of invariants that arbitrary sub-algebras lack: the existence of a Reynolds operator.

7.3. **The Reynolds Operator.**

Definition 7.2. The *Reynolds operator* $\rho\colon R \to I$ is a projection map defined by averaging a polynomial over the entire group action:

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

This definition explicitly requires two assumptions:

(1) We must sum over G, necessitating that G is a finite group.
(2) We must divide by the order of G, meaning the characteristic of the field k must be zero (or at least, strictly coprime to $|G|$).

The Reynolds operator enjoys several critical properties:

$$\rho(1) = 1$$
$$\rho(f) = f \quad \text{if } f \in I$$

Most importantly, while ρ is not generally a ring homomorphism (i.e., $\rho(f_1 f_2) \neq \rho(f_1)\rho(f_2)$), it is a homomorphism of I-modules. If $f \in I$ is an invariant and $h \in R$ is arbitrary, then:

$$\rho(fh) = f\rho(h)$$

This implies that the short exact sequence of I-modules $0 \to \ker(\rho) \to R \to I \to 0$ splits, allowing us to write $R = I \oplus \ker(\rho)$. This splitting is the key to completing the proof.

7.4. **Proof of Hilbert's Theorem.**

Theorem 7.3 (Hilbert)**.** *Let G be a finite group acting linearly on a polynomial ring $R = k[x_1, \ldots, x_n]$, where k has characteristic zero. The algebra of invariants $I = R^G$ is finitely generated as a k-algebra.*

Proof. Let J be the ideal of R generated by homogeneous elements of I of degree > 0. By the Basis Theorem, J is finitely generated by homogeneous invariants $a_1, \ldots, a_k \in I$. We will prove by induction on the degree that any homogeneous invariant $f \in I$ lies in the k-algebra generated by $a_1, \ldots, a_k$, denoted $k[a_1, \ldots, a_k]$.

The base case is trivial: if $\deg(f) = 0$, then $f \in k$, which is trivially in the algebra.

Now suppose $f \in I$ is a homogeneous invariant with $\deg(f) > 0$. Because $f \in I$ and has positive degree, it resides in the ideal J. Thus, we can express f as an R-linear combination of our ideal generators:

$$f = \sum_{i=1}^{k} c_i a_i$$

where $c_i \in R$. Because the a_i are homogeneous, we can assume without loss of generality that the c_i are homogeneous and $\deg(c_i) = \deg(f) - \deg(a_i)$. Since $\deg(a_i) > 0$, it must be that $\deg(c_i) < \deg(f)$.

The issue is that the coefficients c_i are in R, not necessarily in I. We resolve this by applying the Reynolds operator to both sides of the equation. Since f is invariant, $\rho(f) = f$. Furthermore, since $a_i \in I$, we can pull them out of the operator:

$$\begin{aligned} f &= \rho(f) \\ &= \rho\left(\sum_{i=1}^{k} c_i a_i\right) \\ &= \sum_{i=1}^{k} a_i \rho(c_i) \end{aligned}$$

By the definition of the Reynolds operator, $\rho(c_i)$ is an invariant in I. Because ρ preserves degree, $\deg(\rho(c_i)) = \deg(c_i) < \deg(f)$.

By our inductive hypothesis, since the invariants $\rho(c_i)$ have strictly lower degree than f, they must lie in the algebra $k[a_1, \ldots, a_k]$. Therefore, f is expressed as a sum of products of elements that are all in $k[a_1, \ldots, a_k]$. Thus, $f \in k[a_1, \ldots, a_k]$, completing the proof. □

7.5. **Historical Context and Weyl's Unitary Trick.** The Reynolds operator is named after Osborne Reynolds, who did not study commutative algebra, but fluid dynamics. Reynolds developed a technique to manage the immense complexity of fluid flows varying over time by replacing the flow at each point with its time-averaged flow. Mathematically, he was applying an averaging operator over the continuous group of time translations. Algebraists later adopted "Reynolds operator" to mean an averaging operator over any group.

Hilbert originally proved this theorem not just for finite groups, but for continuous groups such as the Special Linear Group $SL_n(\mathbb{C})$. To adapt the proof above, one requires a continuous analog to the Reynolds operator.

For a *compact* group K, we can replace the finite sum with an integral utilizing the Haar measure, normalizing by the finite volume of the group:

$$\rho(f) = \frac{1}{\mathrm{Vol}(K)} \int_K g \cdot f \, dg$$

However, $SL_n(\mathbb{C})$ is not compact, making its volume infinite and rendering this direct integration ill-defined.

To bypass this, one employs *Weyl's Unitary Trick*. The compact Special Unitary Group $SU(n)$ is a subgroup of $SL_n(\mathbb{C})$. At the level of Lie algebras, we have the containment $\mathfrak{su}(n) \subset \mathfrak{sl}_n(\mathbb{C})$. If we complexify the real Lie algebra $\mathfrak{su}(n)$, we recover $\mathfrak{sl}_n(\mathbb{C})$:

$$\mathfrak{su}(n) \otimes_{\mathbb{R}} \mathbb{C} \cong \mathfrak{sl}_n(\mathbb{C})$$

Because they share the same complexification, their finite-dimensional complex representation theories are essentially identical. A finite-dimensional representation of $SL_n(\mathbb{C})$ can be restricted to a representation of $SU(n)$, averaged using the compact integration to find invariants, and these invariants remain valid for the entirety of $SL_n(\mathbb{C})$.

Remark 7.4. This elegant equivalence completely shatters when transitioning to infinite-dimensional representations, as the exponential map mapping the Lie algebra to the group generally fails to converge in infinite dimensions. Furthermore, if the base field k has characteristic $p > 0$, the Reynolds operator fails (division by zero occurs). Emmy Noether circumvented this by pioneering new, purely commutative algebraic techniques to establish finite generation in positive characteristic without relying on group averaging.

8. Noetherian Modules

In this section, we extend the concept of the Noetherian condition from rings to modules. We will discuss the essential properties of Noetherian modules and demonstrate their utility by proving a landmark theorem by Emmy Noether: the algebra of invariants of a finite group is finitely generated, even over fields of positive characteristic.

8.1. **Definitions and Equivalent Conditions.**

Definition 8.1. Let R be a commutative ring. An R-module M is called *Noetherian* if every submodule of M is finitely generated as an R-module.

In particular, a ring R is a Noetherian ring if and only if it is Noetherian when considered as a module over itself. This follows because the submodules of R are precisely the ideals of R.

Just as with Noetherian rings, the Noetherian condition for modules can be formulated in three equivalent ways. The proof of their equivalence is identical to the proof for rings and is therefore omitted.

Theorem 8.2. *For an R-module M, the following conditions are equivalent:*

1. *M is a* Noetherian module *(all submodules are finitely generated).*
2. *Every strictly increasing chain of submodules*

$$M_1 \subsetneq M_2 \subsetneq M_3 \subsetneq \cdots \subseteq M$$

   is finite (Ascending Chain Condition).
3. *Every non-empty set of submodules of M has a maximal element with respect to inclusion.*

8.2. **Properties of Noetherian Modules.** To construct and identify Noetherian modules, we rely on their behavior under exact sequences.

Proposition 8.3. *Suppose we have a short exact sequence of R-modules:*

$$0 \to A \to B \to C \to 0$$

Then B is a Noetherian module if and only if both A and C are Noetherian modules.

Proof. The implication that B being Noetherian forces A and C to be Noetherian is straightforward, as submodules of A and C naturally lift to submodules of B.

Conversely, assume A and C are Noetherian. Let $N \subseteq B$ be an arbitrary submodule. We must show that N is finitely generated. We construct a finite generating set for N in two steps:

- First, consider the image of N in C. Because C is Noetherian, this image is a finitely generated submodule of C. Choose a finite set of elements in N whose images generate the image of N in C.
- Second, consider the intersection $N \cap A$. Because A is Noetherian, this intersection is a finitely generated submodule of A. Choose a finite set of elements in N that generate $N \cap A$.

Taking the union of these two finite sets yields a finite set of elements in N. It is a standard exercise to verify that this combined set generates the entirety of N. Thus, N is finitely generated, and B is Noetherian. □

An immediate consequence of this proposition is that the direct sum of two Noetherian modules is Noetherian. If A and B are Noetherian, the exact sequence

$$0 \to A \to A \oplus B \to B \to 0$$

implies that $A \oplus B$ is Noetherian. By induction, any finite direct sum of Noetherian modules is Noetherian.

Corollary 8.4. *If R is a Noetherian ring, then any finitely generated R-module M is a Noetherian module.*

Proof. Because M is finitely generated, there exists a surjective module homomorphism from a free module of finite rank onto M:

$$R^n \twoheadrightarrow M$$

Since R is a Noetherian ring, it is a Noetherian module over itself. Thus, the finite direct sum R^n is a Noetherian module. Because M is isomorphic to a quotient of R^n, it corresponds to an exact sequence $0 \to K \to R^n \to M \to 0$. By Proposition 8.3, M must also be a Noetherian module. □

This result is ubiquitous in commutative algebra. A vast portion of the subject is dedicated to studying finitely generated modules over Noetherian rings precisely because the Noetherian property ensures that submodules (and hence kernels of homomorphisms) remain well-behaved and finitely generated.

8.3. **Application 1: Finiteness of Syzygies.** As a first application, we revisit the concept of syzygies. Suppose R^G is a finitely generated algebra of invariants. This means we can map a polynomial ring $R = k[x_1, \dots, x_n]$ onto R^G:

$$R \twoheadrightarrow R^G$$

The kernel of this map consists of the relations between the generators, known as the *first-order syzygies*.

Because R is a Noetherian ring, this kernel is finitely generated as an ideal (and thus as an R-module). Consequently, we can map a finite free module onto this kernel, yielding an exact sequence:

$$R^{n_1} \to R \to R^G \to 0$$

The kernel of the map $R^{n_1} \to R$ defines the *second-order syzygies*. Because R^{n_1} is a finitely generated module over a Noetherian ring, it is a Noetherian module. Thus, its submodules—including the kernel—are finitely generated. We can therefore map another finite free module R^{n_2} onto this kernel.

Iterating this process, we deduce that *all higher-order syzygies* are finitely generated. This allows us to construct a free resolution by finite free modules:

$$\cdots \to R^{n_3} \to R^{n_2} \to R^{n_1} \to R \to R^G \to 0$$

For arbitrary modules over arbitrary rings, the ranks n_i might be infinite, leading to a loss of control. The Noetherian property guarantees that these ranks remain finite at every step. (Hilbert's Syzygy Theorem further guarantees that for polynomial rings, this sequence eventually terminates in 0).

8.4. **Application 2: Noether's Theorem.** We now turn to a landmark theorem by Emmy Noether concerning invariant theory in arbitrary characteristic.

Let G be a finite group acting linearly on a finite-dimensional vector space V over a field k. The action of G extends to the polynomial ring $R = k[x_1, \dots, x_n]$, which represents functions on V. We want to prove that the ring of invariants R^G is finitely generated as a k-algebra.

Hilbert's original proof used the Reynolds operator ρ:

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

This approach breaks down if the characteristic of the field k divides the order of the group $|G|$, because division by $|G|$ is undefined (it is division by zero). Furthermore, a non-linear Reynolds operator generally does not exist in characteristic $p > 0$.

To observe this failure, consider the following characteristic p counterexample. Let k be the finite field of order p, and let $G \subset SL_2(k)$ be the group of order p defined by:

$$G = \left\{ \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \;\middle|\; a \in k \right\}$$

G acts on k^2, yielding a short exact sequence of G-modules:

$$0 \to k \to k^2 \to k \to 0$$

There is a one-dimensional subspace on which G acts trivially, and G also acts trivially on the quotient. However, G does *not* act trivially on the whole space k^2. The existence of a Reynolds operator would imply that this sequence splits (i.e., that the representation is a direct sum of trivial representations), which is false.

Emmy Noether bypassed the Reynolds operator entirely by using the theory of Noetherian modules.

Theorem 8.5 (Noether)**.** *Let G be a finite group acting linearly on $R = k[x_1, \dots, x_n]$. Then the invariant ring R^G is finitely generated as a k-algebra, regardless of the characteristic of k.*

Proof. For each generator x_i of R, construct the polynomial:

$$P_i(X) = \prod_{g \in G} (X - g \cdot x_i)$$

where X is a formal variable. By construction, x_i is a root of $P_i(X)$, meaning $P_i(x_i) = 0$.

If we expand $P_i(X)$, the coefficients are the elementary symmetric functions evaluated on the orbit $\{g \cdot x_i \mid g \in G\}$. Because the group G merely permutes the elements of this orbit, these elementary symmetric functions are manifestly invariant under G.

Let S be the k-subalgebra of R generated by the coefficients of $P_i(X)$ for all $i = 1, \dots, n$. We have a tower of rings:

$$S \subseteq R^G \subseteq R$$

We proceed in a sequence of steps, carefully distinguishing between generation as an algebra and generation as a module:

(1) S **is a finitely generated k-algebra.** For each i, the polynomial $P_i(X)$ has $|G|$ coefficients. Thus, S is generated by at most $n|G|$ invariant elements.

(2) S **is a Noetherian ring.** Because S is a finitely generated algebra over a field, it is Noetherian (by Hilbert's Basis Theorem).

(3) R **is a finitely generated S-module.** Every x_i is integral over S, because x_i is a root of the monic polynomial $P_i(X)$ whose coefficients lie in S. Using the relation $P_i(x_i) = 0$, any power x_i^m with $m \geq |G|$ can be reduced to a linear combination of lower powers $1, x_i, \dots, x_i^{|G|-1}$ with coefficients in S. Therefore, R is spanned as an S-module by the finite set of monomials $x_1^{m_1} \dots x_n^{m_n}$ where all $m_i < |G|$. There are at most $|G|^n$ such generators.

(4) R **is a Noetherian S-module.** Because S is a Noetherian ring and R is a finitely generated S-module, R is a Noetherian module over S.

(5) R^G **is a finitely generated S-module.** The invariant ring R^G is an S-submodule of R (since $S \subseteq R^G$). Because R is a Noetherian S-module, all of its submodules must be finitely generated. Thus, R^G is a finitely generated S-module.

(6) R^G **is a finitely generated k-algebra.** We construct a finite generating set for R^G as a k-algebra by taking the union of the generators of S (as a k-algebra) and the generators of R^G (as an S-module). This union is finite, completing the proof.

□

8.5. **Infinite Groups and Reductivity.** Hilbert's proof works for some infinite groups (such as special linear groups) in characteristic zero via the unitary trick. Noether's proof is limited to finite groups, because the polynomial $P_i(X)$ requires a product over all elements of G.

This leaves open the question: *If G is an infinite algebraic group acting on a k-algebra in characteristic $p > 0$, is the invariant ring finitely generated?*

This problem is substantially more difficult. The answer depends heavily on the structure of the group G. The fundamental result, proven by Nagata and Haboush, establishes the following equivalences for an algebraic group G over an algebraically closed field k:

Theorem 8.6 (Nagata, Haboush)**.** *The following conditions are equivalent:*

(1) For any finitely generated k-algebra on which G acts, the invariant ring is a finitely generated k-algebra.
(2) G is reductive*. This means G contains no non-trivial normal subgroups isomorphic to the additive group $\mathbb{G}_a^n$ (copies of k under addition).*
(3) If G acts linearly on a vector space k^n and fixes a non-zero vector v, there exists an invariant polynomial f such that $f(0) \neq f(v)$.

Haboush proved that condition (2) implies (3) (often called the geometric reductivity conjecture, which implies the existence of a non-linear Reynolds operator). Nagata proved that (3) implies (1), and also demonstrated that if G is the additive group of the field, one can construct pathological actions where the invariant ring is *not* finitely generated.

Paradoxically, while the additive group of a field might seem algebraically simpler than matrix groups like $SL_n(k)$, it is precisely the presence of additive subgroups that obstructs finite generation in invariant theory.

9. Euclidean Domains

This lecture focuses on visualizing rings by drawing a picture of them. There are generally three ways to visualize a ring: drawing a point for each element of the ring, drawing a point for each basis element (if the ring is a vector space), or drawing a point for each prime ideal. Here, we investigate the first method: drawing a point for each element of the ring. (The reader is reffered to the recordings for all visualizations throughout this lecture.)

9.1. **Visualizing Basic Rings.** Some familiar examples include:

- The ring of real numbers, $\mathbb{R}$, which we draw as a continuous line.
- The ring of complex numbers, $\mathbb{C}$, which forms a plane.
- The ring of integers, $\mathbb{Z}$, visualized as a subset of discrete points on the real line.

A more structured example is the ring of Gaussian integers:

$$\mathbb{Z}[i] = \{m + ni \mid m, n \in \mathbb{Z}\}$$

Since $\mathbb{Z}[i] \subset \mathbb{C}$, we can visualize it as a square lattice in the complex plane, with points at $0, 1, 2, i, 1 + i$, and so on.

As an application of this geometric perspective, we will show that $\mathbb{Z}[i]$ is a *Unique Factorization Domain* (UFD).

9.2. **Euclidean Domains and Factorization.** We begin by reviewing the definition and properties of a Euclidean domain.

Definition 9.1. A *Euclidean domain* is an integral domain R equipped with a division with remainder algorithm. That is, there exists a map from $R \setminus \{0\}$ to a well-ordered set (such as the non-negative integers), denoting the absolute value $|r|$, such that for any $a, b \in R$ with $b \neq 0$, we can write:

$$a = qb + r$$

where q is the quotient, and r is the remainder, with the strict condition that $|r| < |b|$ or $r = 0$.

For $\mathbb{Z}$, the absolute value is the standard one. For complex subrings, we can use the complex modulus (or its square, the norm, which explicitly maps to integers).

Theorem 9.2. *Every Euclidean domain is a* Principal Ideal Domain *(PID), and every PID is a* Unique Factorization Domain *(UFD).*

Proof Sketch. **Euclidean implies PID:** Let $I \subseteq R$ be a non-zero ideal. Choose an element $a \in I \setminus \{0\}$ with minimal absolute value. For any $b \in I$, division with remainder yields $b = qa + r$ with $|r| < |a|$. Since $r = b - qa \in I$ and a is minimal, r must be 0. Thus, $b = qa$, and $I = (a)$, making it principal.

PID implies UFD: The critical step is to show that in a PID, every *irreducible* element is *prime*. Recall that p is *irreducible* if it is not zero or a unit, and $p = ab$ implies a or b is a unit. An element p is *prime* if $p \mid ab$ implies $p \mid a$ or $p \mid b$.

Suppose p is irreducible and $p \mid ab$. If $p \nmid a$, consider the ideal (p, a). Since R is a PID, $(p, a) = (x)$ for some x. Thus, $x \mid p$, meaning x is either a unit or $x = up$ for some unit u. But x cannot be up, since $x \mid a$ would imply $p \mid a$, a contradiction. Therefore, x is a unit, and the ideal $(p, a) = (1)$.

This means there exist elements $u, v \in R$ such that:

$$up + va = 1$$

Multiplying by b yields:

$$upb + vab = b$$

Since p divides ab, p divides both terms on the left side, which implies $p \mid b$. With irreducible elements being prime, unique factorization into primes follows easily by induction. If $p_1 p_2 \cdots = q_1 q_2 \ldots$, p_1 must divide some q_i, meaning they are equal up to a unit, allowing us to cancel and repeat. □

9.3. **Geometric Proofs of Euclidean Domains.** We now return to $\mathbb{Z}[i]$ to demonstrate geometrically that it is Euclidean. Let $a, b \in \mathbb{Z}[i]$ with $b \neq 0$. We want to find $q, r \in \mathbb{Z}[i]$ such that $a = qb + r$ and $|r| < |b|$.

Dividing by b in the field of complex numbers, we get:

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

This requires finding an element $q \in \mathbb{Z}[i]$ such that the distance from a/b to q in the complex plane is strictly less than 1 (i.e., $|r/b| < 1$).

Geometrically, this is asking: if we draw an open disk of radius 1 around every point in the lattice $\mathbb{Z}[i]$, do these disks cover the entire complex plane?

The maximum distance from any point in $\mathbb{C}$ to the nearest Gaussian integer occurs at the center of the unit squares, which is $\sqrt{(1/2)^2 + (1/2)^2} = 1/\sqrt{2} < 1$. Because the open unit disks fully cover the plane, $\mathbb{Z}[i]$ is Euclidean.

This visual method extends to other rings:

- $\mathbb{Z}[\sqrt{-2}]$**:** The lattice points form rectangles. The maximum distance to a lattice point is strictly less than 1, so the open unit disks cover the plane. Thus, $\mathbb{Z}[\sqrt{-2}]$ is a Euclidean domain.
- $\mathbb{Z}[\sqrt{-3}]$**:** The lattice consists of points $m + n\sqrt{-3}$. If we draw open unit disks around each point, they fail to cover the plane. lattice points showing an uncovered gap] Specifically, the point $\frac{1+\sqrt{-3}}{2}$ is exactly at distance 1 from the lattice points $0, 1, \sqrt{-3}$, and $1 + \sqrt{-3}$. Because we require a strict inequality ($|r/b| < 1$), the argument fails.

In fact, $\mathbb{Z}[\sqrt{-3}]$ is not a Euclidean domain, nor a PID, nor a UFD. To see that unique factorization fails, observe that:

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

These represent two distinct factorizations into irreducibles that do not differ by merely units.

To visually confirm it is not a PID, consider the non-principal ideal $I = (2, 1 + \sqrt{-3})$.

If we plot the points of I, they form a triangular lattice consisting of equilateral triangles. However, any principal ideal (a) in this ring is merely a scaled and rotated version of the base ring itself, which forms a rectangular lattice. Since a triangular lattice cannot be structurally isomorphic to a rectangular lattice via scaling and rotation, the ideal I cannot be principal.

We can fix the defect in $\mathbb{Z}[\sqrt{-3}]$ by enlarging the ring to include the "missing" points. We define the ring of *Eisenstein integers*:

$$\mathbb{Z}\left[\frac{1 + \sqrt{-3}}{2}\right]$$

Let $\omega = \frac{-1+\sqrt{-3}}{2}$, an element satisfying $\omega^2 + \omega + 1 = 0$. This ring consists of all elements $m + n\omega$, which forms a perfect triangular lattice.

Drawing open unit disks around these points successfully covers the plane, proving that the Eisenstein integers form a Euclidean domain.

9.4. **Limits of Euclidean Domains.** It is important to note that Euclidean domains are relatively rare, even among UFDs. For example, the polynomial ring $k[x_1, \ldots, x_n]$ over a field is a UFD, but it is not a PID (e.g., the ideal (x, y) is not principal), and thus not Euclidean.

Among principal ideal domains, most encountered in basic contexts (e.g., $\mathbb{Z}$, $k[x]$, discrete valuation rings) happen to be Euclidean. However, there are PIDs that are not Euclidean domains. The classic minimal example is the ring of integers of $\mathbb{Q}(\sqrt{-19})$:

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

To prove R is not Euclidean, suppose for a contradiction that it were. We could choose a non-zero, non-unit element $a \in R$ with minimal absolute value. By the division algorithm, every element of the quotient ring $R/(a)$ would be represented by either 0 or a unit in R. The only units in R are ± 1. Therefore, the quotient $R/(a)$ could contain at most 3 elements. However, one can check that R has no quotients of size 2 or 3 (the smallest non-trivial quotient, $R/(2)$, has exactly 4 elements). Thus, no such division algorithm can exist, regardless of the absolute value function chosen.

Finally, the technique of drawing a point for every element works beautifully for rings whose additive groups embed neatly into finite-dimensional vector spaces (such as rings of integers in algebraic number fields, where it leads directly to Minkowski's theorem on the finiteness of the ideal class group). However, this geometric method cannot easily be applied to abstract rings or polynomial rings with many variables, motivating the need for alternative methods such as drawing pictures of prime ideals.

10. Weierstrass Preparation Theorem

In this lecture, we introduce a third method for drawing pictures of rings: visualizing a ring by drawing a point for each basis element. We assume the ring is a vector space over a field k, such as a polynomial ring, and we plot its basis elements. While this is trivial for polynomials in one variable, it becomes highly effective for rings of polynomials in two variables, which can be imagined as a two-dimensional array. (Again, the reader is reffered to the recordings for all visualizations throughout this lecture.)

10.1. **Visualizing Rings via Basis Elements.** Consider the polynomial ring $k[x, y]$. We can draw a point for each basis element $1, x, y, x^2, xy, y^2$, and so on. This forms a two-dimensional grid in the first quadrant.

This visualization allows us to easily identify and work with *monomial ideals*—ideals generated by monomials in x and y. By plotting the generators, the ideal consists of all points "above and to the right" of the generators. This visually demonstrates that $k[x, y]$ has a vast number of ideals of finite co-dimension. Because the number of such monomial ideals is already enormous, the classification of finite-dimensional algebras over k generated by two elements is prohibitively complicated.

10.2. **Examples of Visual Computations.**

Example 10.1. Find the dimension of the algebra:

$$A = \frac{k[x,y]}{(x^3, x^2y^2, y^5)}$$

This is incredibly easy to calculate by drawing a picture. We map out the basis elements of the ideal in a grid. The generators are at coordinates $(3,0), (2,2)$, and $(0,5)$. The ideal consists of all monomials divisible by these generators. The quotient ring is spanned by the monomials not in the ideal. By simply counting the remaining points under the "staircase" formed by the ideal, we find there are exactly 12 basis elements:

$$1, y, y^2, y^3, y^4$$
$$x, xy, xy^2, xy^3, xy^4$$
$$x^2, x^2y$$

Thus, the dimension of the algebra is 12.

Example 10.2. Let G be a group of order 2 acting on $R = k[x,y]$ by taking $x \mapsto -x$ and $y \mapsto -y$. We wish to show that the ring of invariants R^G is a two-dimensional free module over a polynomial ring.

First, we visualize $k[x,y]$ as a basis array. The polynomials invariant under this automorphism are exactly those spanned by monomials of even total degree.

We take our base polynomial ring to be generated by x^2 and y^2:

$$P = k[x^2, y^2]$$

Visually, P forms a sub-grid. It is completely obvious from the diagram that the invariant ring R^G is a free module of rank 2 over P, with basis elements 1 and xy. Any even-degree monomial is either in P (if the powers of x and y are both even) or it is xy times an element of P (if the powers of x and y are both odd).

Setting $u = x^2, v = xy$, and $w = y^2$, the ring of invariants is generated by u, v, w modulo the relation:

$$v^2 = uw$$

Thus, R^G is isomorphic to the coordinate ring of a cone.

10.3. **The Weierstraß Preparation Theorem.** A more complex application of visualizing rings is proving that the ring of formal power series in two variables, $k[[x,y]]$, is a Unique Factorization Domain (UFD).

For background:

- The formal power series ring in one variable, $k[[x]]$, is trivially a UFD because it is a Discrete Valuation Ring (DVR); its only non-zero ideals are of the form (x^n).
- By Gauss's Lemma, if R is a UFD, the polynomial ring $R[x]$ is a UFD. By induction, $k[x_1, \ldots, x_n]$ is a UFD.

- However, if R is a UFD, the formal power series ring $R[[x]]$ is *not* necessarily a UFD (counterexamples exist, though they are somewhat messy). Thus, we cannot use simple induction on the number of variables for power series.

To overcome this, we use Weierstraß polynomials.

Theorem 10.3 (Weierstraß Preparation Theorem)**.** *Any non-zero formal power series $f \in k[[x, y]]$ can be written uniquely in the form:*

$$f = x^k \cdot u \cdot P$$

where $k \geq 0$, u is a unit in $k[[x, y]]$, and P is a Weierstraß polynomial *in y.*

Definition 10.4. A *Weierstraß polynomial* in y over $k[[x]]$ is a polynomial of the form:

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

where the coefficients $a_i \in k[[x]]$ are formal power series in x that are divisible by x (i.e., they evaluate to 0 at $x = 0$).

It is crucial to note that the ring of polynomials in y with coefficients in $k[[x]]$ is strictly smaller than the ring of formal power series $k[[x, y]]$. For example, the series $\sum_{i=0}^{\infty} x^i y^i$ is a formal power series, but it is not a polynomial in y because it has infinitely many non-zero terms in y.

Proof Sketch. The theorem is easiest to prove by drawing a picture of the power series. We indicate the basis for the power series as a rectangular array of monomials $x^i y^j$.

Assume f is not divisible by x. This means f has a non-zero coefficient somewhere in the $x = 0$ column (the pure powers of y). Suppose the lowest such power is y^m. We want to "kill off" all other terms in the array to the right of y^m and above it.

We can eliminate the coefficient of y^{m+1} by multiplying f by a unit of the form $(1 + c \cdot y)$. By iteratively multiplying by units of the form $(1 + c \cdot x^i y^j)$, we can systematically eliminate all coefficients not belonging to the desired Weierstraß polynomial structure. Because we are in a formal power series ring, this infinite product of units converges to a well-defined unit u^{-1}. What remains is precisely the Weierstraß polynomial P. □

10.4. **Formal Power Series as a UFD.** We can now use the Weierstraß Preparation Theorem to prove our main result.

Theorem 10.5. *The ring of formal power series $k[[x, y]]$ is a Unique Factorization Domain.*

Proof. We already know that $k[[x, y]]$ is a Noetherian ring, which implies that every element can be factored into a product of irreducible elements. To establish unique factorization, the key step is to show that every irreducible element is prime. That is, if f is irreducible and f divides the product gh, then f divides g or f divides h.

By the Weierstraß Preparation Theorem, we can multiply f, g, and h by units to assume without loss of generality that they are Weierstraß polynomials. Units do not affect divisibility or primeness.

Suppose $f \mid gh$ in $k[[x, y]]$. Then there exists some formal power series $r \in k[[x, y]]$ such that:

$$fr = gh$$

Since f, g, and h are Weierstraß polynomials, gh is also a Weierstraß polynomial. We can apply the preparation theorem to r, writing $r = u \cdot P$, where u is a unit and P is a Weierstraß polynomial (assuming x does not divide r). Thus:

$$f \cdot (u \cdot P) = gh$$

Since the decomposition into a unit times a Weierstraß polynomial is unique, and gh is already a Weierstraß polynomial (meaning its unit factor is 1), we must have $u = 1$. Therefore, $r = P$, which means r is a polynomial in y.

This implies that the divisibility relation $fr = gh$ actually holds in the ring $k[[x]][y]$, which is the polynomial ring in one variable y over the UFD $k[[x]]$. Since polynomial rings over UFDs are UFDs, $k[[x]][y]$ is a UFD.

Within $k[[x]][y]$, the irreducible element f divides gh, so f must divide g or f must divide h. This divisibility in $k[[x]][y]$ naturally carries over to the larger ring $k[[x, y]]$. Thus, f is prime in $k[[x, y]]$, confirming it is a UFD. □

Remark 10.6. A strong warning must be issued regarding *convergent* power series. The ring of convergent power series is *not* a UFD. For example, consider a convergent polynomial $f(x)$ with an infinite number of roots $z_1, z_2, \ldots$ converging to 0. We might attempt to write:

$$f(x) = \prod_{i=1}^{\infty} (x - z_i)$$

Because there are infinitely many zeros, $f(x)$ does not have a factorization into a *finite* number of prime elements. (While Weierstraß formulated a method for writing convergent power series as infinite products, that is a theorem of complex analysis, not finite algebra). The algebraic theory of UFDs requires finite factorizations, making the distinction between formal and convergent power series critical.

Part 2. The Prime Spectrum, Topology, and Localization

11. Spectrum of a Ring

In the last couple of lectures, we covered two intuitive ways of drawing pictures of a ring: drawing a point for each element, or drawing a point for each basis element. The *spectrum* of a ring is a third approach, where we draw a point for each prime ideal. While the first two methods work well in specific simple cases, drawing a point for each prime ideal works universally well for all commutative rings.

11.1. **Motivation from Topology.** Suppose X is a compact Hausdorff space. Let R be the ring of continuous functions on X under pointwise multiplication. The idea is that X serves as a good picture of the ring R. The ring R in this case is a commutative C^*-algebra (an algebra over $\mathbb{R}$ equipped with a supremum norm).

We can recover the space X from the ring R:

- The points $x \in X$ correspond exactly to the maximal closed ideals of R. Specifically, a point x corresponds to the ideal of functions vanishing at x:
$$\mathfrak{m}_x = \{f \in R \mid f(x) = 0\}$$
It follows from the Stone-Weierstraß theorem that all closed maximal ideals of R are of this form.
- We can also reconstruct the topology on X. A basis of open sets is given by:
$$U_f = \{x \in X \mid f(x) \neq 0\}$$
Equivalently, in terms of maximal ideals, U_f consists of those ideals $\mathfrak{m}$ such that $f \notin \mathfrak{m}$.
- Alternatively, for any closed ideal $I \subseteq R$, we can define a closed set as the common zeros of all elements in I:
$$Z(I) = \{\mathfrak{m} \mid I \subseteq \mathfrak{m}\}$$

11.2. **The Maximal Spectrum and its Defect.** We can attempt to copy this construction for any commutative ring R by taking the set X to be the maximal ideals of R. This yields a space called the *maximal spectrum* of R, denoted $\mathrm{Spec}_m(R)$.

However, this approach has a critical defect regarding ring homomorphisms. Given a homomorphism of C^*-algebras $\phi\colon R \to S$, we naturally get a continuous map between their maximal spectra in the reverse direction, pulling back maximal ideals via $\phi^{-1}(\mathfrak{m})$.

For arbitrary rings, however, the preimage of a maximal ideal under a ring homomorphism might not be maximal.

Example 11.1. Consider the inclusion map $\phi\colon \mathbb{Z} \hookrightarrow \mathbb{Q}$. The zero ideal (0) is a maximal ideal in the field $\mathbb{Q}$. However, its preimage $\phi^{-1}((0)) = (0)$ is not a maximal ideal in $\mathbb{Z}$.

To fix this, we recall that a maximal ideal $\mathfrak{m} \subset S$ means $S/\mathfrak{m} = K$, a field. The pullback $\phi^{-1}(\mathfrak{m})$ is the kernel of the composition $R \to S \to K$. The image of R in K is a subring of a field, which is an integral domain, but not necessarily a field itself. Thus, the quotient of R by the preimage is an integral domain, which means that the preimage $\phi^{-1}(\mathfrak{m})$ is a *prime* ideal.

11.3. **The Prime Spectrum.** Because the preimage of a prime ideal under a ring homomorphism is always a prime ideal, we use prime ideals instead of maximal ideals to define the spectrum.

Definition 11.2. The *spectrum* of a commutative ring R, denoted $\mathrm{Spec}(R)$, is the set of all prime ideals of R.

We equip $\mathrm{Spec}(R)$ with a topology, known as the Zariski topology, defined in either of two equivalent ways:

- **Open sets:** A basis for the topology is given by the sets U_f, defined for each $f \in R$ as:

$$U_f = \{\mathfrak{p} \in \mathrm{Spec}(R) \mid f \notin \mathfrak{p}\}$$

- **Closed sets:** For any ideal $I \subseteq R$, we define a closed set:

$$Z(I) = \{\mathfrak{p} \in \mathrm{Spec}(R) \mid I \subseteq \mathfrak{p}\}$$

Lemma 11.3. *The sets $Z(I)$ satisfy the axioms for the closed sets of a topology.*

Proof. We must verify that these sets are closed under arbitrary intersections and finite unions:

- **Arbitrary intersections:**

$$\bigcap_\alpha Z(I_\alpha) = Z\left(\sum_\alpha I_\alpha\right)$$

- **Finite unions:**

$$Z(I) \cup Z(J) = Z(IJ)$$

This relies on the property of prime ideals: $IJ \subseteq \mathfrak{p}$ if and only if $I \subseteq \mathfrak{p}$ or $J \subseteq \mathfrak{p}$.

□

11.4. **Examples of** $\mathrm{Spec}(R)$**.**

Example 11.4. Let R be the *zero ring* (where $1 = 0$). By definition, a prime ideal $\mathfrak{p}$ must yield an integral domain quotient $R/\mathfrak{p}$, which requires $1 \neq 0$ in the quotient. Thus, the zero ring has no prime ideals:

$$\mathrm{Spec}(0) = \emptyset$$

Example 11.5. Let $R = k$, a *field*. A field has exactly one prime ideal, the zero ideal (0).

$$\mathrm{Spec}(k) = \{(0)\}$$

Geometrically, this corresponds to a single point.

Example 11.6. Let $R = \mathbb{C}[x]$. Since R is a Principal Ideal Domain (PID), its prime ideals are easy to classify:

- Maximal ideals of the form $(x - \alpha)$ for $\alpha \in \mathbb{C}$.
- The non-maximal prime ideal (0).

Thus, as a set, $\mathrm{Spec}(\mathbb{C}[x]) = \mathbb{C} \cup \{(0)\}$. The closed sets in this topology are the whole space and finite subsets of $\mathbb{C}$.

The point (0) is not a "point at infinity" like in the Riemann sphere. Instead, its closure is the entire space:

$$\overline{\{(0)\}} = \mathrm{Spec}(\mathbb{C}[x])$$

This makes (0) a *generic point*. Consequently, the topology is highly non-Hausdorff.

Example 11.7. Let $R = \mathbb{Z}$. The prime ideals are:

- Maximal ideals (p) for prime numbers $p = 2, 3, 5, \ldots$
- The zero ideal (0).

The closed sets are the whole space and finite sets of maximal ideals. The generic point (0) is dense. We can visualize (0) as a one-dimensional "fuzzy" point encompassing all the zero-dimensional closed points (p).

Example 11.8. Let $R = \mathbb{R}[x]$. The prime ideals correspond to irreducible polynomials:

- Linear polynomials $(x - \alpha)$ for $\alpha \in \mathbb{R}$.
- Irreducible quadratics $(x^2 + bx + c)$ with $b^2 - 4c < 0$, which correspond to conjugate pairs of complex roots $x \pm iy$ $(y \neq 0)$.
- The generic point (0).

Geometrically, $\mathrm{Spec}(\mathbb{R}[x])$ consists of the real line, the upper half of the complex plane (the complex plane folded in half, identifying each complex number with its conjugate), and the generic point (0).

In general, for $k[x]$ over a non-algebraically closed field k, the points of the spectrum correspond to the Galois orbits of the algebraic closure $\bar{k}$ over k, together with the generic point (0).

12. Examples of Spec(R)

In this section, we provide several concrete examples of the spectrum of a ring, $\mathrm{Spec}(R)$, to build geometric intuition for the prime ideal structure of commutative rings. We will see how $\mathrm{Spec}(R)$ bridges algebra, geometry, and number theory.

12.1. **Why is it called the "Spectrum"?** Before discussing complex rings, it is worth answering a basic question: why is the topological space of prime ideals called a "spectrum"? The answer lies in linear algebra and physics.

Consider a matrix $A \in M_n(\mathbb{C})$. Let R be the finite-dimensional commutative algebra over $\mathbb{C}$ generated by A. If the minimal polynomial of A factors as:

$$p(x) = (x - \alpha_1)^{n_1}(x - \alpha_2)^{n_2} \cdots (x - \alpha_k)^{n_k}$$

then the ring R is isomorphic to the quotient:

$$R \cong \frac{\mathbb{C}[x]}{(p(x))}$$

In this quotient ring, all prime ideals are maximal. They are simply generated by the linear factors corresponding to the roots of $p(x)$:

$$\mathfrak{m}_i = (x - \alpha_i)$$

Thus, as a set, $\mathrm{Spec}(R) = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}$. This is exactly the set of eigenvalues of the matrix A, which in linear algebra and quantum mechanics is called the *spectrum* of the operator. The ring-theoretic definition perfectly recovers the physical one.

12.2. **Decomposing Spec($\mathbb{Z}$) and Localization.** Recall that the spectrum of the integers, $\mathrm{Spec}(\mathbb{Z})$, consists of the closed points corresponding to the maximal ideals $(2), (3), (5), \ldots$, along with the dense generic point (0).

We can study specific regions of this space by altering the ring:

- **Inverting an element:** Consider $R_1 = \mathbb{Z}[1/2]$. By formally adjoining a half, the element 2 becomes a unit, so the ideal (2) is "killed off". The spectrum $\mathrm{Spec}(\mathbb{Z}[1/2])$ looks identical to $\mathrm{Spec}(\mathbb{Z})$, except the prime 2 has been removed. We do this when we want to eliminate the pathological behaviors often associated with the prime 2.
- **Localizing at a prime:** Consider $R_2 = \mathbb{Z}_{(2)}$, the ring of rational numbers with odd denominators. Here, we invert all odd primes. This kills off $(3), (5), (7)$, and so on. The only surviving prime ideal is the maximal ideal (2) (along with the generic point (0)). We use this ring when we wish to focus exclusively on the local behavior at the prime 2.

There is a natural map $\mathbb{Z}_{(2)} \to \mathbb{Z}/2\mathbb{Z}$. This induces a continuous map of spectra in the opposite direction:

$$\mathrm{Spec}(\mathbb{Z}/2\mathbb{Z}) \hookrightarrow \mathrm{Spec}(\mathbb{Z}_{(2)})$$

The spectrum of the field $\mathbb{Z}/2\mathbb{Z}$ is a single point, which geometrically embeds as the closed point (2) inside the localized spectrum.

12.3. **Ramification in the Gaussian Integers.** Consider the inclusion map $\mathbb{Z} \hookrightarrow \mathbb{Z}[i]$. This induces a map of spectra:

$$\mathrm{Spec}(\mathbb{Z}[i]) \to \mathrm{Spec}(\mathbb{Z})$$

Because $\mathbb{Z}[i]$ is a Principal Ideal Domain, its non-zero prime ideals are generated by prime elements. By studying the fibers of this topological map, we visually recover the number-theoretic decomposition of primes.

- The prime 2 ramifies: $2 = -i(1+i)^2$. Up to units, there is only one prime ideal in $\mathbb{Z}[i]$ above 2, but the square indicates a geometric "branch point".
- Primes $p \equiv 3 \pmod 4$ (like $3, 7, 11$) remain prime. The fiber over these points is a single point.
- Primes $p \equiv 1 \pmod 4$ (like $5, 13$) split into two distinct primes: $5 = (2+i)(2-i)$. The fiber over these points consists of two distinct points.

Thus, Spec provides a geometric visualization of the decomposition of primes in algebraic number fields.

12.4. **Polynomial Rings in Two Variables.** Let $k = \mathbb{C}$ and consider the polynomial ring in two variables $R = \mathbb{C}[x, y]$. This is the coordinate ring of the affine plane $\mathbb{C}^2$. The prime ideals of this ring categorize into three geometric dimensions:

(1) **Maximal Ideals (0-dimensional):** Ideals of the form $(x-\alpha, y-\beta)$ for $\alpha, \beta \in \mathbb{C}$. These are closed points, uniquely corresponding to the actual geometric points $(\alpha, \beta) \in \mathbb{C}^2$.
(2) **Prime Ideals of Curves (1-dimensional):** Ideals of the form (f), where f is an irreducible polynomial in x and y. These correspond to algebraic curves. In the Zariski topology, the closure of the point (f) consists of (f) itself, plus all the maximal ideals $(x-\alpha, y-\beta)$ that lie on the curve $f(\alpha, \beta) = 0$.
(3) **The Generic Point (2-dimensional):** The zero ideal (0). This is a massive, highly non-closed point. Its closure is the entire space $\mathrm{Spec}(\mathbb{C}[x, y])$.

12.5. **Formal Power Series Rings.** Now consider the formal power series ring $R = \mathbb{C}[[x, y]]$.

Because a formal power series is a local algebraic object, its spectrum represents a drastically zoomed-in geometric picture near the origin:

- There is only *one* closed point, corresponding to the unique maximal ideal (x, y).
- There are prime ideals generated by irreducible formal power series (f). Geometrically, these represent "infinitely small curves" or formal branches passing through the origin. These curves are merely "ghosts"; they have no global geometric extent.
- There is a generic 2-dimensional point (0).

The natural map $\mathbb{C}[x, y] \to \mathbb{C}[[x, y]]$ mapping polynomials to their Taylor series induces a map of spectra that formally embeds the local structure of $\mathrm{Spec}(\mathbb{C}[[x, y]])$ infinitesimally around the origin of the global plane $\mathbb{C}^2$.

12.6. **Hecke Algebras and Modular Forms.** As a final example illustrating the deep connections between commutative algebra and number theory, we examine the spectrum of a Hecke ring.

Consider the two-dimensional space of modular forms of level 1 and weight 12, spanned by the Eisenstein series E_{12} and the discriminant function Δ_{12}:

$$E_{12} = \frac{691}{65520} + \sum_{n \geq 1} \sigma_{11}(n) q^n$$

$$\Delta_{12} = q \prod_{n=1}^{\infty} (1 - q^n)^{24} = \sum_{n \geq 1} \tau(n) q^n$$

where $\sigma_{11}(n)$ is the sum of the 11^{th} powers of divisors of n, and $\tau(n)$ is the Ramanujan tau function.

The Hecke operators T_n act diagonally on this basis:

$$T_n(E_{12}) = \sigma_{11}(n) E_{12}$$
$$T_n(\Delta_{12}) = \tau(n) \Delta_{12}$$

We define the Hecke algebra R as the subring of $\mathbb{Z} \times \mathbb{Z}$ generated by the action of T_n on this basis, explicitly generated by the pairs $(\sigma_{11}(n), \tau(n))$.

To understand R, we look for congruences between the eigenvalues. Ramanujan famously discovered that:

$$\sigma_{11}(n) \equiv \tau(n) \pmod{691}$$

Because of this congruence, the Hecke ring R consists precisely of all pairs $(m, n) \in \mathbb{Z} \times \mathbb{Z}$ such that $m \equiv n \pmod{691}$.

To draw $\mathrm{Spec}(R)$, we analyze the natural projection $R \to \mathbb{Z} \times \mathbb{Z}$. The spectrum of $\mathbb{Z} \times \mathbb{Z}$ consists of two disjoint copies of $\mathrm{Spec}(\mathbb{Z})$. However, because of the condition $m \equiv n \pmod{691}$, the points modulo 691 on the first line must equal the points modulo 691 on the second line.

Consequently, $\mathrm{Spec}(R)$ looks geometrically like two separate lines (copies of $\mathrm{Spec}(\mathbb{Z})$) that physically intersect at exactly one point: the prime 691. Such intersections are called *Eisenstein primes*. Every modular form corresponds to a line in the spectrum of the full Hecke algebra, and the geometric intersection points of these lines correspond perfectly to arithmetic congruences between the modular forms.

(Remarkably, the coefficient 691 appearing here is also deeply related to the order of the Conway sporadic simple group, the automorphism group of the Leech lattice, highlighting the profound interconnectedness of algebraic structures).

13. Topology of $\mathrm{Spec}(R)$

In this lecture, we explore the basic topological properties of the spectrum of a ring, $\mathrm{Spec}(R)$, such as compactness, connectedness, and irreducibility.

13.1. **Recap of the Zariski Topology.** Recall that for a commutative ring R, the points of $\mathrm{Spec}(R)$ are its prime ideals. The *Zariski topology* on $\mathrm{Spec}(R)$ can be defined in two equivalent ways:

- **Basis of open sets:** The sets U_f, informally representing the loci where f is not zero, are defined as:
$$U_f = \{\mathfrak{p} \in \mathrm{Spec}(R) \mid f \notin \mathfrak{p}\}$$
- **Closed sets:** For any ideal $I \subseteq R$, the closed set consisting of primes containing I is:
$$Z(I) = \{\mathfrak{p} \in \mathrm{Spec}(R) \mid I \subseteq \mathfrak{p}\}$$

This topology is highly non-Hausdorff. For instance, $\mathrm{Spec}(\mathbb{Z})$ contains a non-closed generic point (0), which we visualize as one-dimensional, whose closure contains all the maximal ideals (p).

13.2. **Quasi-Compactness.**

Theorem 13.1. *The spectrum of a ring* $\mathrm{Spec}(R)$ *is* quasi-compact.

Remark 13.2. The term "quasi-compact" is synonymous with "compact" (every open cover has a finite subcover). The "quasi-" prefix is a historical artifact from when some mathematicians reserved "compact" to mean "compact and Hausdorff". Since $\mathrm{Spec}(R)$ is almost never Hausdorff, the term quasi-compact was invented, though today most use compact to mean not necessarily Hausdorff.

Proof. Suppose $\mathrm{Spec}(R)$ is covered by an arbitrary collection of open sets. Without loss of generality, we may assume these are basis open sets U_{f_i}. The condition that the entire spectrum is covered by the sets U_{f_i} means there is no prime ideal (and hence no maximal ideal) that fails to be in at least one U_{f_i}. Equivalently, no prime ideal contains all the elements f_i.

This implies that the ideal generated by the set of all f_i must be the entire ring R:

$$\langle \{f_i\} \rangle = R$$

Consequently, the identity element 1 can be expressed as a linear combination of the f_i:

$$1 = c_1 f_{i_1} + c_2 f_{i_2} + \cdots + c_k f_{i_k}$$

for some $c_j \in R$. Because this is an algebraic relation, it inherently involves only a finite number of terms. Working backwards, the finite subset $\{f_{i_1}, \ldots, f_{i_k}\}$ generates the unit ideal, meaning the finite collection $U_{f_{i_1}}, \ldots, U_{f_{i_k}}$ covers $\mathrm{Spec}(R)$.

Notice that this proof relies crucially on algebra prohibiting infinite sums. In analysis, where infinite series are permitted, this argument would fail, and the analogous topological spaces need not be compact. □

13.3. **Connectedness.**

Definition 13.3. A topological space is *connected* if it is non-empty and cannot be written as the disjoint union of two non-empty open sets.

The spectrum of a ring is sometimes connected and sometimes not.

Proposition 13.4. *If R is an integral domain, then* $\mathrm{Spec}(R)$ *is connected.*

Proof. If R is an integral domain, the zero ideal (0) is prime. The closure of the point (0) is exactly $Z((0)) = \mathrm{Spec}(R)$. Because the entire space is the closure of a single point, any non-empty open set must contain (0). Therefore, it is impossible to find two non-empty disjoint proper open sets, so the space is connected. □

13.3.1. *Disconnected Spectra.* A primary example of a disconnected spectrum arises when R is a direct product of rings, $R = A \times B$. If we denote the identities of A and B by 1_A and 1_B respectively, we have:

$$1_A \cdot 1_B = 0$$
$$1_A + 1_B = 1$$

Any prime ideal $\mathfrak{p} \subset R$ must contain 0, and thus contains the product $1_A 1_B$. Because $\mathfrak{p}$ is prime, it must contain either 1_A or 1_B. It follows easily that the prime ideals of R are exactly the disjoint union of the prime ideals of A and the prime ideals of B:

$$\mathrm{Spec}(A \times B) = \mathrm{Spec}(A) \coprod \mathrm{Spec}(B)$$

Example 13.5. By the Chinese Remainder Theorem, the ring $\mathbb{Z}/120\mathbb{Z}$ factors as:

$$\mathbb{Z}/120\mathbb{Z} \cong \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$$

Because it is a product of three rings, its spectrum is the disjoint union of their spectra. Since the spectrum of $\mathbb{Z}/p^k\mathbb{Z}$ is a single point, $\mathrm{Spec}(\mathbb{Z}/120\mathbb{Z})$ consists of exactly three discrete points corresponding to the prime ideals $(2), (3)$, and (5). We can view the Chinese Remainder Theorem visually as decomposing the spectrum into disjoint points.

Example 13.6. Let G be a finite abelian group, such as the Klein four-group $V_4 = \{1, a, b, c\}$ where $a^2 = b^2 = c^2 = 1$ and $abc = 1$. Consider the rational group ring $\mathbb{Q}[G]$. The group has four characters, which correspond to four orthogonal idempotents e_i such that:

$$\begin{aligned} e_i^2 &= e_i \\ e_i e_j &= 0 \quad (i \neq j) \\ e_0 + e_1 + e_2 + e_3 &= 1 \end{aligned}$$

For instance, the trivial character corresponds to $e_0 = \frac{1+a+b+c}{4}$. The ring splits as a product:

$$\mathbb{Q}[G] \cong e_0\mathbb{Q}[G] \times e_1\mathbb{Q}[G] \times e_2\mathbb{Q}[G] \times e_3\mathbb{Q}[G] \cong \mathbb{Q}^4$$

Thus, $\mathrm{Spec}(\mathbb{Q}[G])$ consists of four discrete, disconnected points. For an arbitrary finite abelian group over $\mathbb{C}$, the group ring similarly splits via idempotents $e_\chi = \frac{1}{|G|}\sum \chi(g)g$.

13.4. **Irreducibility.** In commutative algebra, a far more useful and stronger concept than connectedness is irreducibility.

Definition 13.7. A topological space X is *irreducible* if it is non-empty and cannot be expressed as the union of two proper closed subsets. Equivalently, any two non-empty open subsets of X must intersect.

For Hausdorff spaces, irreducibility is largely useless: the only irreducible Hausdorff spaces are single points (if a space has two distinct points, they can be separated by disjoint open sets, immediately violating the intersection property). However, in the highly non-Hausdorff Zariski topology, irreducible spaces are ubiquitous.

For example, $\mathrm{Spec}(\mathbb{Z})$ is irreducible. Generally, if X contains a dense point x (such that $\overline{\{x\}} = X$), then X is irreducible. The spectrum of any integral domain is irreducible for this exact reason.

13.4.1. *Irreducible Components.* A closed subset that is irreducible and maximal with respect to inclusion is called an *irreducible component*. Many spaces are connected but not irreducible, meaning they can be decomposed into a finite union of these irreducible components.

Example 13.8. Consider the integer group ring $\mathbb{Z}[G]$ where G is the Klein four-group. We have four character homomorphisms $\mathbb{Z}[G] \to \mathbb{Z}$, each inducing a continuous map of spectra:

$$\mathrm{Spec}(\mathbb{Z}) \to \mathrm{Spec}(\mathbb{Z}[G])$$

The spectrum of $\mathbb{Z}[G]$ is the union of four copies of $\mathrm{Spec}(\mathbb{Z})$. Unlike the rational group ring $\mathbb{Q}[G]$, these four components are not disjoint. If we tensor with $\mathbb{Z}/2\mathbb{Z}$, the resulting ring $(\mathbb{Z}/2\mathbb{Z})[G]$ has only one prime ideal. Geometrically, the four copies of $\mathrm{Spec}(\mathbb{Z})$ are glued together at the point corresponding to the prime 2.

Thus, $\mathrm{Spec}(\mathbb{Z}[G])$ is connected, but it is *reducible*, as it is the union of four proper closed subsets (its irreducible components). It acts as a connected refinement of the completely disconnected space $\mathrm{Spec}(\mathbb{Q}[G])$.

Example 13.9. Let $R = \mathbb{C}[x, y]/(xy)$. Any prime ideal must contain xy, so it must contain either x or y. Geometrically, the spectrum is the union of the x-axis and the y-axis. The space is connected (they meet at the origin) but reducible, as it is the union of two proper closed irreducible components (the two affine lines).

Remark 13.10. One must exercise caution when visualizing algebraic curves. For instance, the elliptic curve defined by $y^2 = x^3 + x$ may appear geometrically disconnected when plotted in the real Euclidean plane. However, the spectrum of its coordinate ring:

$$R = \frac{\mathbb{C}[x, y]}{(y^2 - x^3 - x)}$$

is not only connected but entirely irreducible in the Zariski topology. The Euclidean topology is vastly finer and can exhibit visual disconnections that do not exist algebraically.

14. Irreducible Subsets of Spec(R)

In this lecture, we discuss the irreducible subsets of the spectrum of a ring R. We will establish that every closed irreducible subset corresponds to the closure of a single point (a prime ideal). Following this, we will apply these concepts to the ring of continuous functions on a compact Hausdorff space to illustrate the strange behavior of non-closed prime ideals.

14.1. **Classifying Irreducible Closed Subsets.** Recall that if X is any topological space and $x \in X$, the closure $\overline{\{x\}}$ is always an irreducible closed subset. We aim to show that the converse holds in the Zariski topology of $\mathrm{Spec}(R)$.

Theorem 14.1. *In the spectrum of a commutative ring R, any closed irreducible subset is of the form $\overline{\{\mathfrak{p}\}}$ for some $\mathfrak{p} \in \mathrm{Spec}(R)$.*

Proof. Suppose we have an irreducible closed subset. Any closed subset in the Zariski topology is of the form $Z(I)$ for some ideal I. Let $Z(I)$ be irreducible. We want to show that there exists a prime ideal $\mathfrak{p}$ such that:

$$Z(I) = \overline{\{\mathfrak{p}\}}$$

A natural first guess is to set $\mathfrak{p} = I$. However, I need not be prime even if $Z(I)$ is irreducible. For instance, in the integers $\mathbb{Z}$, if $I = (4)$, then $Z(I) = \{(2)\}$, which is a single point and thus irreducible, but (4) is not a prime ideal.

To resolve this, we first replace I with its *radical*, denoted $\sqrt{I}$. The radical of I is defined as:

$$\sqrt{I} = \{r \in R \mid r^n \in I \text{ for some } n \geq 1\}$$

Let us briefly verify that $\sqrt{I}$ is indeed an ideal. It is clearly closed under multiplication by elements of R. To check closure under addition, suppose $r, s \in \sqrt{I}$. Then there exist integers $m, n \geq 1$ such that $r^m \in I$ and $s^n \in I$. Consider the binomial expansion of $(r+s)^{m+n}$:

$$(r+s)^{m+n} = r^{m+n} + \cdots + \binom{m+n}{k} r^{m+n-k} s^k + \cdots + s^{m+n}$$

For each term in the expansion, either the exponent of r is at least m (in which case the term is divisible by $r^m \in I$), or the exponent of s is at least n (divisible by $s^n \in I$). Thus, every term belongs to I, which implies:

$$(r+s)^{m+n} \in I$$

Therefore, $r + s \in \sqrt{I}$, proving it is an ideal.

Furthermore, any prime ideal $\mathfrak{p}$ containing I must also contain $\sqrt{I}$. If $r^n \in I \subseteq \mathfrak{p}$, then $r \in \mathfrak{p}$ by primeness. Thus, the closed sets coincide:

$$Z(I) = Z(\sqrt{I})$$

Consequently, we may assume without loss of generality that I is a radical ideal, meaning $I = \sqrt{I}$.

We now show that under this assumption, I must be prime. Suppose, for the sake of contradiction, that there exist elements $a, b \in R$ such that:

$$ab \in I$$

but $a \notin I$ and $b \notin I$. Consider the ideals $I + (a)$ and $I + (b)$. We have:

$$(I + (a))(I + (b)) \subseteq I$$

This implies that the prime ideals containing I are exactly those containing $I+(a)$ or $I + (b)$. Topologically, this means:

$$Z(I) = Z(I + (a)) \cup Z(I + (b))$$

Since $Z(I)$ is an irreducible closed set, it cannot be written as the union of two proper closed subsets. Therefore, one of these sets must equal $Z(I)$. Without loss of generality, assume:

$$Z(I) = Z(I + (a))$$

Because we assumed $a \notin I$ and I is a radical ideal, no power of a can be in I. Thus, the multiplicative subset $S = \{1, a, a^2, \dots\}$ is disjoint from I. By Lemma 14.2 (proven below), there exists a prime ideal $\mathfrak{p}$ such that:

$$I \subseteq \mathfrak{p}$$
$$\mathfrak{p} \cap S = \emptyset$$

Since $I \subseteq \mathfrak{p}$, we have $\mathfrak{p} \in Z(I)$. However, because $\mathfrak{p} \cap S = \emptyset$, the element a is not in $\mathfrak{p}$. This means $I + (a) \not\subseteq \mathfrak{p}$, so $\mathfrak{p} \notin Z(I + (a))$. This directly contradicts our deduction that $Z(I) = Z(I + (a))$.

We conclude that our assumption was false, so I must be a prime ideal. Finally, if I is prime, then $Z(I)$ is precisely the closure of the point I in the spectrum:

$$Z(I) = \overline{\{I\}}$$

This completes the classification of irreducible closed subsets. □

14.2. **A Useful Lemma on Prime Ideals.** In the proof above, we relied on an essential existence lemma for prime ideals.

Lemma 14.2. *Suppose S is a* multiplicative subset *of a ring R (meaning $1 \in S$, and $a, b \in S \implies ab \in S$). Let I be an ideal of R disjoint from S. Then there exists a prime ideal $\mathfrak{p}$ such that:*

$$I \subseteq \mathfrak{p}$$
$$\mathfrak{p} \cap S = \emptyset$$

Proof. We use Zorn's Lemma. Let Σ be the set of all ideals in R that contain I and are disjoint from S. Since $I \in \Sigma$, the set is non-empty. By Zorn's Lemma, Σ has a maximal element, which we will call $\mathfrak{p}$. We must show that $\mathfrak{p}$ is a prime ideal.

Suppose $a, b \in R$ such that:

$$ab \in \mathfrak{p}$$

We want to show that either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$. Suppose for contradiction that $a \notin \mathfrak{p}$ and $b \notin \mathfrak{p}$. Then the ideals $\mathfrak{p} + (a)$ and $\mathfrak{p} + (b)$ are strictly larger than $\mathfrak{p}$. By the maximality of $\mathfrak{p}$ in Σ, these larger ideals must intersect S. Therefore, there exist elements in S of the form:

$$s_1 \in \mathfrak{p} + (a)$$
$$s_2 \in \mathfrak{p} + (b)$$

Since S is a multiplicative subset, their product $s_1 s_2$ must also be in S. However, computing the product yields:

$$s_1 s_2 \in (\mathfrak{p} + (a))(\mathfrak{p} + (b)) \subseteq \mathfrak{p} + (ab)$$

Since $ab \in \mathfrak{p}$ by assumption, the entire product $s_1 s_2$ is in $\mathfrak{p}$. This implies that $\mathfrak{p}$ contains an element of S, contradicting the fact that $\mathfrak{p} \cap S = \emptyset$. Thus, either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$, proving that $\mathfrak{p}$ is prime. □

Remark 14.3. This argument is highly typical in commutative algebra: maximal elements of families of ideals subject to disjointness conditions have a very strong tendency to be prime ideals.

14.3. **Application: The Spectrum of $C(X)$ is "Weird".** As an application of Lemma 14.2, we will analyze the spectrum of the ring of continuous functions, highlighting pathological behavior when working with non-closed ideals.

Let X be a compact Hausdorff space, and let $R = C(X)$ be the ring of continuous real-valued functions on X under pointwise multiplication. We wish to understand $\mathrm{Spec}(C(X))$.

14.3.1. *Maximal Ideals.* The maximal ideals of $C(X)$ are completely determined by the points of X. For each $x \in X$, we define:

$$\mathfrak{m}_x = \{f \in C(X) \mid f(x) = 0\}$$

As established earlier, one can completely recover the space X from $C(X)$ via these maximal ideals.

14.3.2. *Prime Ideals.* The behavior of arbitrary prime ideals in $C(X)$ is considerably stranger. First, every prime ideal $\mathfrak{p}$ is contained in a unique maximal ideal $\mathfrak{m}_x$. To see this, suppose $\mathfrak{p}$ is contained in two distinct maximal ideals $\mathfrak{m}_x$ and $\mathfrak{m}_y$. Because X is compact Hausdorff, we can find continuous functions f and g such that:

$$\begin{aligned} f(x) &\neq 0 \\ g(y) &\neq 0 \\ f(t)g(t) &= 0 \quad \text{for all } t \in X \end{aligned}$$

Thus, $fg = 0 \in \mathfrak{p}$. Since $\mathfrak{p}$ is prime, either $f \in \mathfrak{p}$ or $g \in \mathfrak{p}$. If $f \in \mathfrak{p} \subseteq \mathfrak{m}_x$, then $f(x) = 0$, a contradiction. A similar contradiction arises if $g \in \mathfrak{p}$. Thus, $\mathfrak{p}$ belongs to a unique maximal ideal.

Furthermore, if a prime ideal $\mathfrak{p}$ is a *closed* set in the topological ring $C(X)$, and $\mathfrak{p} \subseteq \mathfrak{m}_x$, one can show via a short analysis argument that:

$$\mathfrak{p} = \mathfrak{m}_x$$

Therefore, the closed prime ideals are exactly the maximal ideals.

14.3.3. *Non-Closed Prime Ideals.* The true weirdness arises from non-closed prime ideals. We will use Lemma 14.2 to prove their existence without explicitly constructing one (which inherently requires the Axiom of Choice).

Assume X has a non-isolated point x. Let I be the ideal of functions vanishing on some neighborhood of x. Because x is not isolated, we can find a continuous function $f \in C(X)$ such that:

$$f(x) = 0$$

but f does not vanish completely on any neighborhood of x. This implies that $f \notin I$, and similarly, no power of f is in I.

We construct a multiplicative subset from the powers of f:

$$S = \{1, f, f^2, \dots\}$$

By our construction, $S \cap I = \emptyset$. Invoking Lemma 14.2, there exists a prime ideal $\mathfrak{p}$ such that:

$$I \subseteq \mathfrak{p}$$
$$\mathfrak{p} \cap S = \emptyset$$

Because I is contained in $\mathfrak{p}$, and I consists of functions vanishing near x, it is easy to see that $\mathfrak{p}$ must be contained in the maximal ideal $\mathfrak{m}_x$. However, $f \in \mathfrak{m}_x$ (since $f(x) = 0$), but $f \notin \mathfrak{p}$ (since $\mathfrak{p} \cap S = \emptyset$).

This yields a strict inclusion:

$$\mathfrak{p} \subsetneq \mathfrak{m}_x$$

We have proven the existence of a non-maximal, non-closed prime ideal $\mathfrak{p}$.

Remark 14.4. The moral of this example is that if a ring is equipped with a topology, one should generally restrict operations to closed ideals. Quotienting a topological ring by a non-closed ideal like $\mathfrak{p}$ results in a quotient space with a non-Hausdorff topology, which is analytically extremely poorly behaved.

15. Noetherian Topological Spaces

This lecture introduces *Noetherian topological spaces*, their properties, and their relation to the spectrum of Noetherian rings. We will use these concepts to decompose closed sets into irreducible components and investigate examples from algebraic geometry and representation theory.

15.1. **Definition and Equivalent Conditions.** We begin by defining what it means for a topological space to be Noetherian.

Theorem 15.1. *For a topological space X, the following conditions are equivalent:*

1. *Every non-empty set of closed subsets has a minimal element.*
2. *Every non-empty set of open subsets has a maximal element.*
3. *Every strictly increasing sequence of open subsets stabilizes. That is, any chain*
$$U_1 \subseteq U_2 \subseteq U_3 \subseteq \dots$$
eventually satisfies $U_n = U_{n+1} = \dots$ for some n.
4. *Every strictly decreasing sequence of closed subsets stabilizes. That is, any chain*
$$C_1 \supseteq C_2 \supseteq C_3 \supseteq \dots$$
eventually satisfies $C_n = C_{n+1} = \dots$ for some n.
5. *Every open subset of X is* quasi-compact *(meaning every open cover has a finite subcover).*

Proof Sketch. The equivalence (1) $\iff$ (2) and (3) $\iff$ (4) follows immediately by taking complements. The equivalence between (2) and (3) is a standard property of partially ordered sets.

To see (3) $\iff$ (5), suppose condition (5) holds. Given an increasing sequence of open sets $U_1 \subseteq U_2 \subseteq \dots$, let $U = \bigcup U_i$. Since U is an open subset, it is quasi-compact. The sets U_i form an open cover of U. By quasi-compactness, U is covered by a finite number of the U_i, and since they are nested, $U = U_n$ for some n. Thus, the sequence stabilizes, yielding condition (3). The reverse implication follows similarly. □

15.2. **Spectrum of a Noetherian Ring.**

Proposition 15.2. *If R is a Noetherian ring, then its spectrum* $\mathrm{Spec}(R)$ *is a Noetherian topological space.*

Proof. Closed sets in $\mathrm{Spec}(R)$ correspond to ideals of R. Because taking the zero locus $Z(I)$ reverses inclusions, an ascending chain of ideals in R corresponds to a descending chain of closed sets in $\mathrm{Spec}(R)$.

Since R is a Noetherian ring, every ascending chain of ideals stabilizes. Consequently, every descending chain of closed sets in $\mathrm{Spec}(R)$ stabilizes. Thus, $\mathrm{Spec}(R)$ is a Noetherian topological space. □

Remark 15.3. The Noetherian condition on a topological space is almost completely incompatible with the Hausdorff property. One can easily verify that if X is both Noetherian and Hausdorff, then X must be finite and discrete. Thus, apart from trivial examples, Hausdorff spaces are never Noetherian spaces.

We might ask if the converse holds: if $\mathrm{Spec}(R)$ is a Noetherian space, is R a Noetherian ring? The answer is no.

Example 15.4. Let k be a field, and consider the polynomial ring in infinitely many variables $k[x_1, x_2, \dots]$. We construct a quotient ring:

$$R = \frac{k[x_1, x_2, \dots]}{(x_1^2, x_2^2, \dots)}$$

In this ring, any prime ideal must contain each x_i^2, and therefore must contain each x_i. Thus, there is only one prime ideal in the entire ring:

$$\mathfrak{p} = (x_1, x_2, \dots)$$

Since $\mathrm{Spec}(R)$ consists of a single point, it has only two open sets ($\emptyset$ and the whole space), making it trivially a Noetherian topological space. However, R is *not* a Noetherian ring, because the maximal ideal $\mathfrak{p}$ is not finitely generated.

15.3. **Noetherian Induction.** When working with Noetherian spaces, a powerful proof technique is *Noetherian induction*.

Theorem 15.5 (Noetherian Induction)**.** *Let X be a Noetherian topological space, and let P be a property of closed subsets of X. Suppose that for any closed subset $C \subseteq X$, if all proper closed subsets of C possess property P, then C possesses property P. Then all closed subsets of X possess property P.*

Proof. Suppose, for the sake of contradiction, that there exist closed subsets of X that do not have property P. Because X is Noetherian, the set of all such closed

subsets must have a minimal element (with respect to inclusion). Let C be this minimal closed subset lacking property P.

By minimality, every proper closed subset of C must have property P. But by our assumption, this implies that C itself must have property P, which is a contradiction. Therefore, no such minimal element can exist, and all closed subsets must possess property P. □

15.4. **Decomposition into Irreducible Components.** We now apply Noetherian induction to prove a fundamental structural theorem for closed sets in Noetherian spaces.

Theorem 15.6. *In a Noetherian topological space X, every closed set can be written as the union of a finite number of closed irreducible subsets.*

Proof. We proceed by Noetherian induction. Let P be the property that a closed set C is a finite union of closed irreducible subsets. We assume that every proper closed subset of C has property P, and we must show that C has property P.

There are two cases:

(1) C **is irreducible:** In this case, C is the union of one irreducible subset (itself), so it trivially satisfies property P.
(2) C **is not irreducible:** By definition, C can be written as the union of two proper closed subsets, $C = D \cup E$. By our inductive hypothesis, since D and E are proper closed subsets of C, they both possess property P. That is, they are both finite unions of irreducible closed subsets. Consequently, their union $C = D \cup E$ is also a finite union of irreducible closed subsets, so C satisfies property P.

By Noetherian induction, every closed subset of X satisfies property P. □

Corollary 15.7. *If R is a Noetherian ring, every closed subset of* $\mathrm{Spec}(R)$ *is a finite union of irreducible closed sets. Since irreducible closed sets in* $\mathrm{Spec}(R)$ *are precisely the closures of prime ideals $\overline{\{\mathfrak{p}\}}$, knowing the prime ideals provides a complete geometric description of all closed subsets.*

15.5. **Examples and Applications.** We illustrate these concepts with examples ranging from algebraic geometry to representation theory.

Example 15.8 (Algebraic Sets)**.** Let $R = \mathbb{C}[x, y]$. This is a Noetherian ring, so $\mathrm{Spec}(\mathbb{C}[x, y])$ is a Noetherian space. An *algebraic set* in $\mathbb{C}^2$ corresponds to a closed subset $Z(I)$ of the spectrum.

By our theorem, any algebraic set can be uniquely decomposed into a finite number of irreducible components (which correspond to points and irreducible curves). In algebraic geometry, these irreducible components are called *varieties*.

Example 15.9 (Continuous Functions)**.** Let X be a compact Hausdorff space, and $R = C(X)$ be the ring of continuous real-valued functions on X. Because R is generally highly non-Noetherian, its spectrum does not enjoy the finite decomposition property.

Recall that every prime ideal is contained in a unique maximal ideal. Thus, an irreducible closed subset (which is the closure of a prime ideal) can contain

at most one maximal ideal. Since the maximal ideals form a copy of the compact Hausdorff space X, it contains a vast number of closed subsets that cannot possibly be expressed as a finite union of irreducibles.

Example 15.10 (Representation Theory of S_3)**.** Let us apply the spectrum to visualize the representation theory of a non-abelian group. We consider the symmetric group S_3 and its group ring over the integers, $\mathbb{Z}[S_3]$.

Because this is a course on commutative algebra, we restrict our attention to its center, $R = Z(\mathbb{Z}[S_3])$. A basis for the center is given by the conjugacy classes of S_3:

$$\begin{aligned} &1 \\ &a = (12) + (23) + (13) \quad \text{(order 2)} \\ &b = (123) + (132) \quad \text{(order 3)} \end{aligned}$$

As an abelian group, $R \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$. The multiplication is governed by the relations:

$$\begin{aligned} a^2 &= 3 + 3b \\ b^2 &= 2 + b \\ ab &= 2a \end{aligned}$$

We wish to determine the spectrum of R. We first find the maps from R to integral domains (like $\mathbb{Z}$) by taking the relations modulo a prime p.

In R/p, the relation $b^2 - b - 2 = 0$ implies $(b+1)(b-2) = 0$. Since R/p is an integral domain, we have two cases:

- $b = -1$: Then $a^2 = 3 + 3(-1) = 0$, which implies $a = 0$.
- $b = 2$: Then $a^2 = 3 + 3(2) = 9$, which implies $a = 3$ or $a = -3$.

This yields three distinct homomorphisms from R to $\mathbb{Z}$:

$$\begin{aligned} &\phi_1 \colon a \mapsto 3, b \mapsto 2 \quad \text{(Trivial representation)} \\ &\phi_2 \colon a \mapsto -3, b \mapsto 2 \quad \text{(Sign representation)} \\ &\phi_3 \colon a \mapsto 0, b \mapsto -1 \quad \text{(Standard 2D representation)} \end{aligned}$$

Each map gives a continuous embedding $\operatorname{Spec}(\mathbb{Z}) \hookrightarrow \operatorname{Spec}(R)$. Thus, $\operatorname{Spec}(R)$ has three irreducible components, corresponding to the three irreducible representations of S_3 over $\mathbb{C}$.

However, these components intersect at specific primes, reflecting modular representations over fields of positive characteristic:

- Over the prime (2), the evaluations $a = 3$ and $a = -3$ coincide ($3 \equiv -3 \pmod 2$). Thus, the trivial and sign representations become indistinguishable modulo 2.
- Over the prime (3), the evaluations $b = 2$ and $b = -1$ coincide ($2 \equiv -1 \pmod 3$). Furthermore, $a = 3 \equiv 0 \pmod 3$. Here, the trivial representation and the two-dimensional representation merge.

Geometrically, $\operatorname{Spec}(R)$ consists of three lines over the base line $\operatorname{Spec}(\mathbb{Z})$. The lines are separate for characteristic $p > 3$, but they pinch together at the primes

$p = 2$ and $p = 3$. This topological picture of the spectrum perfectly captures the block structure of the modular representation theory of S_3.

16. Localization

This lecture introduces the operation of *localization*. Given a commutative ring R and a subset $S \subset R$, the goal is to construct a new ring, denoted $S^{-1}R$, in which all elements of S are forced to have inverses.

16.1. **Motivating Examples.** We begin by examining the localization of familiar rings to see how this algebraic operation corresponds geometrically to restricting the spectrum of the ring.

16.1.1. *Localizing the Integers.* Let the base ring be the integers, $R = \mathbb{Z}$.

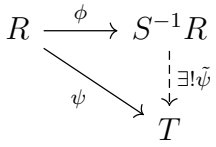
- Let $S = \mathbb{Z} \setminus \{0\}$ be the set of all non-zero integers. Inverting all non-zero elements yields the field of rational numbers:
$$S^{-1}R = \mathbb{Q}$$
The spectrum of $\mathbb{Q}$ consists of only the generic point (0).
- Let $S = \{2^n \mid n \geq 0\}$ be the powers of 2. The localized ring consists of rationals with only powers of 2 in the denominator:
$$S^{-1}R = \left\{ \frac{a}{2^n} \;\middle|\; a \in \mathbb{Z}, n \geq 0 \right\}$$
Geometrically, we have killed off the prime ideal (2) by making 2 invertible. The spectrum of this ring is $\operatorname{Spec}(\mathbb{Z})$ with the closed point (2) removed.
- Let S be the multiplicatively closed set generated by all odd primes (i.e., we invert $3, 5, 7, \dots$). The localized ring consists of rationals with odd denominators:
$$S^{-1}R = \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{Z}, b \text{ is odd} \right\}$$
Here, we have thrown away all odd primes from the spectrum, keeping only the prime (2) and the generic point (0). We use this ring to study the arithmetic behavior strictly locally at the prime 2.

16.1.2. *Localizing Polynomial Rings.* Let $R = \mathbb{C}[x]$, the coordinate ring of the affine line.

- If $S = \mathbb{C}[x] \setminus \{0\}$, the localization is the field of rational functions:
$$S^{-1}R = \mathbb{C}(x) = \left\{ \frac{p(x)}{q(x)} \;\middle|\; q(x) \neq 0 \right\}$$
The spectrum retains only the generic point.
- If $S = \{x^n \mid n \geq 0\}$, we invert only x. The result is the ring of Laurent polynomials:
$$S^{-1}R = \mathbb{C}[x, x^{-1}]$$
Geometrically, this restricts the spectrum to the affine line minus the origin.

- If S is the set generated by $(x-\alpha)$ for all $\alpha \neq 0$, we invert all points except the origin. The localized ring consists of rational functions that are defined at $x=0$:
$$S^{-1}R = \left\{ \frac{p(x)}{q(x)} \;\middle|\; q(0) \neq 0 \right\}$$
The spectrum of this ring contains only the origin (x) and the generic point (0). The term "localization" precisely refers to focusing locally on this neighborhood near zero.

16.2. **Universal Property and Construction.** We can define the localization via its universal property. Suppose we have a ring homomorphism $\phi\colon R \to T$ such that the image $\phi(s)$ is invertible in T for all $s \in S$. Then there exists a unique ring homomorphism from $S^{-1}R$ to T making the following diagram commute:



We can construct $S^{-1}R$ abstractly by taking R, adding a new variable t_s for each $s \in S$, and quotienting by the relations that force t_s to be the inverse of s:
$$S^{-1}R = \frac{R[\{t_s\}_{s\in S}]}{\langle st_s - 1\rangle_{s\in S}}$$
While this construction trivially satisfies the universal property, it is impractically huge and obscures the structure of the ring. For instance, determining the kernel of the natural map $R \to S^{-1}R$ is highly non-obvious from this quotient.

16.3. **Construction via Fractions.** To simplify the construction and notation, we assume from now on that S is a *multiplicative subset* of R. This means:
$$1 \in S$$
$$s_1, s_2 \in S \implies s_1 s_2 \in S$$
This assumption is harmless; if we start with an arbitrary set, we can replace it with its multiplicative closure without altering the resulting localized ring.

16.3.1. *Case 1: S has no zero divisors.* We mirror the construction of $\mathbb{Q}$ from $\mathbb{Z}$. We consider ordered pairs (r,s) with $r \in R$ and $s \in S$, which we write suggestively as r/s. We define an equivalence relation on these fractions:
$$\frac{r_1}{s_1} \sim \frac{r_2}{s_2} \iff r_1 s_2 = r_2 s_1$$
We define the ring operations in the standard way:
$$\frac{r_1}{s_1} + \frac{r_2}{s_2} = \frac{r_1 s_2 + r_2 s_1}{s_1 s_2}$$
$$\frac{r_1}{s_1} \cdot \frac{r_2}{s_2} = \frac{r_1 r_2}{s_1 s_2}$$

The only subtle point in verifying that this forms a well-defined ring is checking that the relation $\sim$ is transitive. Suppose:

$$r_1 s_2 = r_2 s_1$$
$$r_2 s_3 = r_3 s_2$$

Multiplying the first equation by s_3 and substituting the second yields:

$$r_1 s_2 s_3 = r_2 s_1 s_3 = r_3 s_2 s_1$$

Rearranging, we get:

$$s_2(r_1 s_3 - r_3 s_1) = 0$$

Because we assumed S contains no zero divisors, s_2 cannot be a zero divisor. Thus, we can cancel it to conclude:

$$r_1 s_3 = r_3 s_1$$

which means $r_1/s_1 \sim r_3/s_3$. In this case, the natural map $R \to S^{-1}R$ given by $r \mapsto r/1$ is strictly injective.

16.3.2. *Case 2: S contains zero divisors.* If S contains zero divisors, the transitivity argument above breaks down, and the naive equivalence relation fails. To correct this, we define I to be the set of elements in R annihilated by some element of S:

$$I = \{r \in R \mid rs = 0 \text{ for some } s \in S\}$$

We must check that I is an ideal. If $r_1, r_2 \in I$, there exist $s_1, s_2 \in S$ such that $r_1 s_1 = 0$ and $r_2 s_2 = 0$. Using the fact that S is multiplicative, $s_1 s_2 \in S$, and we have:

$$(r_1 + r_2)(s_1 s_2) = (r_1 s_1)s_2 + (r_2 s_2)s_1 = 0 + 0 = 0$$

Thus, $r_1 + r_2 \in I$, making it an ideal.

In the quotient ring R/I, the image of S has no zero divisors. We can safely form the localization of R/I by the image of S. Consequently, the kernel of the map $R \to S^{-1}R$ is precisely the ideal I.

Equivalently, one can define the localized ring in a single step by modifying the equivalence relation on fractions:

$$\frac{r_1}{s_1} \sim \frac{r_2}{s_2} \iff s(r_1 s_2 - r_2 s_1) = 0 \text{ for some } s \in S$$

This explicitly absorbs the zero-divisor issue into the relation, ensuring transitivity holds universally. Two elements of the localization are equal if and only if they are equivalent under this relation, giving us complete structural control over $S^{-1}R$.

16.4. **Important Special Cases.**

Example 16.1 (Total Quotient Ring)**.** Let R be any ring, and let S be the set of all non-zero divisors in R. Then S is multiplicative, and $S^{-1}R$ is called the *total quotient ring* of R. The map $R \to S^{-1}R$ is injective, making this the largest ring of fractions containing R as a subring. If R is an integral domain, this construction produces the standard field of fractions.

Example 16.2 (Killing Elements via Localization)**.** Let $R = \mathbb{C}[x, y]/(xy)$, whose spectrum consists of the x-axis and the y-axis. Let $S = \{1, x, x^2, \dots\}$. In the localization $S^{-1}R$, the relation $xy = 0$ implies:

$$x \cdot y = 0$$

Since $x \in S$, y is annihilated by an element of S. Therefore, y maps to 0 in the localized ring. The resulting ring is isomorphic to $\mathbb{C}[x, x^{-1}]$, geometrically corresponding to throwing away the y-axis and the origin.

Example 16.3 (Localization at a Prime Ideal)**.** Let $\mathfrak{p}$ be a prime ideal in R. We define the subset S as the complement of the prime ideal:

$$S = R \setminus \mathfrak{p}$$

By the very definition of a prime ideal ($a, b \notin \mathfrak{p} \implies ab \notin \mathfrak{p}$), the set S is a multiplicative subset. The localization $S^{-1}R$ is of paramount importance in commutative algebra; it is denoted by $R_{\mathfrak{p}}$ and is called the *localization of R at $\mathfrak{p}$*.

17. Spec$(S^{-1}R)$

In this section, we investigate the relationship between a ring R and its localization $S^{-1}R$ with respect to a multiplicative subset S. We will analyze the correspondence of ideals between these rings and describe the topological and geometric structure of the spectrum of a localized ring.

17.1. **Ideals in Localizations.** Let $f\colon R \to S^{-1}R$ be the natural localization map. We wish to relate the ideals of R to the ideals of $S^{-1}R$.

If I is an ideal of R, its direct image $f(I)$ is generally not an ideal in $S^{-1}R$. For example, under the natural map $\mathbb{Z} \to \mathbb{Q}$, the image of the ideal (2) is not an ideal in $\mathbb{Q}$. Therefore, we must map I to the ideal *generated by* its image:

$$\langle f(I) \rangle$$

Conversely, if J is an ideal in $S^{-1}R$, its inverse image $f^{-1}(J)$ is always an ideal in R, without requiring any modifications.

Remark 17.1. In the literature, $\langle f(I) \rangle$ is often called the *extension* of I, denoted I^e, while $f^{-1}(J)$ is called the *contraction* of J, denoted J^c.

These operations are not strictly inverses of each other, as there is no bijection between the complete sets of ideals of both rings. However, they are closely related in one direction.

Proposition 17.2. *If J is an ideal of $S^{-1}R$, then:*

$$J = \langle f(f^{-1}(J)) \rangle$$

Proof. The inclusion $\langle f(f^{-1}(J))\rangle \subseteq J$ is trivial. To prove the reverse inclusion, suppose $x \in J$. By the definition of the localization, we can write $x = r/s$ for some $r \in R$ and $s \in S$. Then:

$$\frac{r}{1} = \frac{r}{s} \cdot \frac{s}{1} \in J$$

This implies that $r \in f^{-1}(J)$. Consequently, $r/1 = f(r) \in f(f^{-1}(J))$. Since $1/s \in S^{-1}R$, we have:

$$\frac{r}{s} = \frac{r}{1} \cdot \frac{1}{s} \in \langle f(f^{-1}(J))\rangle$$

Thus, $J \subseteq \langle f(f^{-1}(J))\rangle$, completing the proof. □

An immediate consequence of this proposition is that the map $J \mapsto f^{-1}(J)$ is an injection from the set of ideals of $S^{-1}R$ into the set of ideals of R. We can therefore view the ideals of $S^{-1}R$ as a subset of the ideals of R.

Corollary 17.3. *If R is a Noetherian ring, then $S^{-1}R$ is also a Noetherian ring.*

Proof. A ring is Noetherian if and only if its ideals satisfy the ascending chain condition. Since the ideals of $S^{-1}R$ form a subset of the ordered set of ideals of R (preserving inclusions), any ascending chain of ideals in $S^{-1}R$ corresponds to an ascending chain of ideals in R. Since chains in R stabilize, chains in $S^{-1}R$ must also stabilize. □

Remark 17.4. Note that $S^{-1}R$ is typically not finitely generated as an R-algebra (unless S is generated by a finite set of elements). Therefore, one cannot generally invoke Hilbert's Basis Theorem to prove this corollary; the direct structural argument above is necessary. This highlights a fundamental principle: localization is an exceptionally well-behaved operation that preserves nearly all "nice" properties of commutative rings, unlike operations such as completion which frequently introduce technical pathologies.

17.2. **The Topology of Spec$(S^{-1}R)$.** We can now describe the topological relationship between the spectra of R and its localization.

Theorem 17.5. *The spectrum* $\mathrm{Spec}(S^{-1}R)$ *is homeomorphic to the subspace of* $\mathrm{Spec}(R)$ *consisting of all prime ideals $\mathfrak{p} \subset R$ such that $\mathfrak{p} \cap S = \emptyset$.*

Proof Sketch. We have seen that ideals of $S^{-1}R$ inject into ideals of R. It is a standard exercise to verify that this bijection restricts to prime ideals, precisely targeting those primes in R that do not intersect S.

To check that the topology matches the subspace topology, consider the basis of open sets for $\mathrm{Spec}(S^{-1}R)$, which are of the form $U_{r/s}$. Because s is invertible, this set is exactly the primes not containing r/s, which is equivalent to the primes not containing $r/1$. This directly corresponds to the open set U_r in $\mathrm{Spec}(R)$. Thus, the topology on $\mathrm{Spec}(S^{-1}R)$ is exactly the subspace topology inherited from $\mathrm{Spec}(R)$. □

Geometrically, one can imagine elements of S as functions on $\mathrm{Spec}(R)$. Passing to the localization $S^{-1}R$ corresponds to deleting the zero loci of all functions in S.

We are effectively throwing away a union of closed subsets (often of codimension 1) from the spectrum of R.

17.3. **Examples of Localized Spectra.** The most important special case is when $S = R \setminus \mathfrak{p}$ for some prime ideal $\mathfrak{p}$. The resulting localized ring is denoted $R_{\mathfrak{p}}$.

17.3.1. *Dimension One.*

- $\mathbb{Z}$ **localized at** (0)**:** Here, $\mathfrak{p} = (0)$. The localized ring is $\mathbb{Q}$. Its spectrum is a single point, corresponding to the generic point of $\mathrm{Spec}(\mathbb{Z})$.
- $\mathbb{Z}$ **localized at** (2)**:** Here, $\mathfrak{p} = (2)$. The localized ring $\mathbb{Z}_{(2)}$ consists of rational numbers with odd denominators. Its spectrum consists of the closed point (2) and the generic point (0).
- $\mathbb{C}[x]$ **localized at** (0)**:** The result is $\mathbb{C}(x)$, the field of rational functions, containing only the generic point.
- $\mathbb{C}[x]$ **localized at** (x)**:** The localized ring consists of rational functions defined at $x = 0$. Its spectrum contains the point (x) and the generic point (0).

17.3.2. *Dimension Two.* Consider the coordinate ring of the complex affine plane, $R = \mathbb{C}[x, y]$. Its prime ideals are the zero ideal (0), principal ideals generated by irreducible polynomials (f), and maximal ideals $(x - \alpha, y - \beta)$. Let us examine the localizations at each type of prime:

(1) **Localize at** (0)**:** The multiplicative set S is all non-zero elements. The localization is the field of rational functions $\mathbb{C}(x, y)$. The spectrum collapses to a single point.
(2) **Localize at** (f)**:** We invert all polynomials not divisible by the irreducible curve f. The resulting spectrum retains the generic point (0) and the prime (f). All other points (maximal ideals and other curves) are killed off.
(3) **Localize at** $\mathfrak{m} = (x - \alpha, y - \beta)$**:** We invert all polynomials that do not vanish at the point (α, β). The spectrum of this local ring $R_{\mathfrak{m}}$ retains the closed point $\mathfrak{m}$, the generic point (0), and a "ghost" of every 1-dimensional irreducible curve (f) that passes through (α, β).

17.4. **Quotients vs. Localizations.** It is instructive to contrast taking the quotient of a ring by a prime ideal with localizing at that prime ideal. Let $\mathfrak{p} \in \mathrm{Spec}(R)$.

- **Quotient** $R/\mathfrak{p}$**:** The spectrum $\mathrm{Spec}(R/\mathfrak{p})$ corresponds to all prime ideals of R *containing* $\mathfrak{p}$. Topologically, this is the closure of the point $\mathfrak{p}$. In the quotient ring, $\mathfrak{p}$ becomes the zero ideal, making it a *minimal* prime.
- **Localization** $R_{\mathfrak{p}}$**:** The spectrum $\mathrm{Spec}(R_{\mathfrak{p}})$ corresponds to all prime ideals of R *contained in* $\mathfrak{p}$. Topologically, this consists of the points outside $\mathfrak{p}$ that are "nearby". In the localized ring, $\mathfrak{p}$ becomes the unique maximal ideal, making it *maximal*.

Thus, quotients and localizations are dual operations: quotienting makes a prime minimal, while localizing makes a prime maximal.

17.5. **A Geometric Computation.** To demonstrate the power of visual geometric reasoning, consider the following problem. Let:

$$A = \frac{\mathbb{C}[x, y]}{(xy)}$$

We wish to determine the spectrum of the localization of A at the ideal generated by x and y, denoted $A_{(x,y)}$.

Instead of an algebraic calculation, we draw the spectrum. The spectrum of A consists of the x-axis and the y-axis intersecting at the origin. Its points are the origin (x, y), the generic point of the x-axis (y), the generic point of the y-axis (x), and all the closed points on both axes.

Localizing at the origin (x, y) means we discard all closed points except the origin itself, but we keep the generic points of any irreducible components passing through the origin. Therefore, we retain the origin, the generic point of the x-axis, and the generic point of the y-axis. The spectrum of the localized quotient $A_{(x,y)}$ consists of exactly three points.

18. Functions on Spec(R)

In this lecture, we discuss how to interpret the elements of an arbitrary commutative ring R as functions defined on its spectrum, Spec(R). This geometric viewpoint provides the foundation for the theory of schemes.

18.1. **Motivating Example: Continuous Functions.** Suppose we look at the ring of continuous functions on a topological space X, where X is compact Hausdorff. Let this ring be $C(X)$. The elements of $C(X)$ are literally continuous real-valued functions on X.

We saw previously that X is homeomorphic to the maximal ideal space of $C(X)$. A point $x \in X$ corresponds exactly to the maximal ideal of functions vanishing at that point:

$$\mathfrak{m}_x = \{f \in C(X) \mid f(x) = 0\}$$

When we evaluate a function f at the point x, it takes values in the real numbers. Algebraically, this value lives in the quotient field:

$$\frac{C(X)}{\mathfrak{m}_x} \cong \mathbb{R}$$

This makes it easy to visualize the ring $C(X)$: we just think of the space X, and we think of the ring elements as functions taking values in $\mathbb{R}$ at each point.

18.2. **Generalizing to Arbitrary Rings.** The problem is whether we can do something similar for any arbitrary ring R. We want the elements of R to be understood as functions on Spec(R).

The first question is: what should these functions take values in?

For an arbitrary ring R and a prime ideal $\mathfrak{p} \in \text{Spec}(R)$, the quotient $R/\mathfrak{p}$ is not generally a field; it is an integral domain. However, we can take its field of quotients, which we will denote by $k(\mathfrak{p})$. We can think of an element $f \in R$ as a

function from $\mathrm{Spec}(R)$ to these varying integral domains or fields. For each point $\mathfrak{p}$, the function f takes values in $R/\mathfrak{p}$ (or $k(\mathfrak{p})$).

Example 18.1. Suppose $R = \mathbb{C}[x]$. The maximal ideals are of the form $(x-\alpha)$, corresponding to the complex number α. If we have a function in R, such as $f(x) = x^2$, we can evaluate it at these points.

For the ideal $(x-\alpha)$, the quotient is:

$$\frac{\mathbb{C}[x]}{(x-\alpha)} \cong \mathbb{C}$$

Here, the function evaluates to the standard complex value α^2.

However, there is a slight exception at the generic point (0). Here, the quotient is $\mathbb{C}[x]/(0) \cong \mathbb{C}[x]$, and the value lies in its quotient field $\mathbb{C}(x)$. Thus, the target space where our function evaluates changes depending on the point.

Example 18.2. Let $R = \mathbb{Z}$. Things are slightly more complicated because the fields change dramatically. The prime ideals are $(2), (3), (5), \ldots$ and the generic point (0). The corresponding quotients are $\mathbb{F}_2, \mathbb{F}_3, \mathbb{F}_5$, and so on.

If we take a typical element like $f = 8 \in \mathbb{Z}$, we can evaluate this function across the spectrum:

$$\begin{aligned} f \pmod 2 &= 0 \in \mathbb{F}_2 \\ f \pmod 3 &= 2 \in \mathbb{F}_3 \\ f \pmod 5 &= 3 \in \mathbb{F}_5 \end{aligned}$$

At the generic prime (0), the quotient is $\mathbb{Z}$, and the field of fractions is $\mathbb{Q}$, so the function passes through the point 8 in $\mathbb{Q}$. This is a roundabout way of drawing the graph of the function $f = 8$ mapping from $\mathrm{Spec}(\mathbb{Z})$ to a varying space of fields.

18.3. **The Nilradical and Injectivity.** We have a map from R to the space of functions on $\mathrm{Spec}(R)$. We must ask: is this map injective? If we want to faithfully represent elements of R as functions, a non-zero element of R should not evaluate to zero everywhere.

Suppose an element $f \in R$ maps to zero in all the integral domains $R/\mathfrak{p}$. This is equivalent to saying:

$$f \in \mathfrak{p} \quad \text{for all } \mathfrak{p} \in \mathrm{Spec}(R)$$

Thus, f maps to zero everywhere if and only if it lies in the intersection of all prime ideals.

For the integers $\mathbb{Z}$, the intersection of all prime ideals is exactly (0), so the map is injective. But in general, the intersection of all prime ideals is the *nilradical* of the ring.

Theorem 18.3. *The nilradical of a ring R (the set of all nilpotent elements) is exactly equal to the intersection of all prime ideals of R.*

Proof. If a is nilpotent, then $a^n = 0$ for some $n \geq 1$. Since $0 \in \mathfrak{p}$ for every prime ideal, $a^n \in \mathfrak{p}$. By the definition of a prime ideal, this implies $a \in \mathfrak{p}$. Thus, the nilradical is contained in the intersection of all primes.

Conversely, suppose a is not nilpotent. We consider the multiplicative subset $S = \{1, a, a^2, \dots\}$. Since no power of a is 0, the subset S is disjoint from the zero ideal $I = (0)$.

By a previous lemma, if a multiplicative subset is disjoint from an ideal, we can find a prime ideal $\mathfrak{p}$ containing the ideal and disjoint from the multiplicative subset. Thus, there exists a prime $\mathfrak{p}$ such that:

$$\mathfrak{p} \cap S = \emptyset$$

This means $a \notin \mathfrak{p}$, so a is not in the intersection of all prime ideals. □

Therefore, if R has non-zero nilpotent elements, we cannot faithfully represent those elements as functions into the domains $R/\mathfrak{p}$.

18.4. **Evaluating in Local Rings.** To fix this problem and capture the nilpotent elements, we adjust our definition of the fiber. Instead of mapping to the integral domain $R/\mathfrak{p}$, we consider $f \in R$ as a function taking values in the local rings $R_\mathfrak{p}$ (the localization of R at $\mathfrak{p}$).

Proposition 18.4. *If an element $f \in R$ maps to zero in the local ring $R_\mathfrak{p}$ for every prime ideal $\mathfrak{p} \in \operatorname{Spec}(R)$, then $f = 0$.*

Proof. An element f maps to zero in $R_\mathfrak{p}$ if and only if there exists some element $s \notin \mathfrak{p}$ such that $sf = 0$.

Consider the annihilator ideal of f:

$$\operatorname{Ann}(f) = \{x \in R \mid xf = 0\}$$

The condition that $sf = 0$ for some $s \notin \mathfrak{p}$ implies that $\operatorname{Ann}(f)$ is not contained in $\mathfrak{p}$. Since this is true for *every* prime ideal $\mathfrak{p}$ in $\operatorname{Spec}(R)$, it means $\operatorname{Ann}(f)$ cannot be contained in any maximal ideal.

Therefore, the annihilator ideal must be the entire ring R. This implies that $1 \in \operatorname{Ann}(f)$, meaning:

$$1 \cdot f = 0$$

which forces $f = 0$. □

Local rings are not quite as easy to deal with as fields, but they are much simpler than general rings. By evaluating functions into local rings rather than fields, we obtain a fully faithful representation of R, allowing us to study the ring locally.

18.5. **The Structure Sheaf.** To formalize the idea of functions on $\operatorname{Spec}(R)$, we will construct a sheaf. Returning to the continuous functions $C(X)$, if $U \subseteq X$ is an open set, we define $\mathcal{O}(U)$ as the ring of continuous functions on U. This assignment has three fundamental properties:

(1) **Restriction:** If $U \subseteq V$, there is a restriction map $\mathcal{O}(V) \to \mathcal{O}(U)$.
(2) **Pre-sheaf property:** If U is covered by open sets U_i, and a function $f \in \mathcal{O}(U)$ restricts to the zero function on all U_i, then $f = 0$ on the whole of U.

(3) **Sheaf property:** If we have functions $f_i \in \mathcal{O}(U_i)$ that agree on all intersections $U_i \cap U_j$, we can uniquely glue them together to form a function $f \in \mathcal{O}(U)$.

We want to define analogous rings $\mathcal{O}(U)$ for open sets $U \subseteq \mathrm{Spec}(R)$ such that they behave like "nice functions" on U. Because functions on the spectrum take values in varying spaces, it is not immediately obvious how to define this.

We bypass the complication of arbitrary open sets by focusing on the distinguished open sets U_f, which form a basis for the topology. Recall that U_f is the complement of the vanishing locus of f:

$$U_f = \{\mathfrak{p} \in \mathrm{Spec}(R) \mid f \notin \mathfrak{p}\}$$

What should the "nice functions" on U_f be?

- The elements of the ring R itself ought to be nice functions.
- Since f does not vanish anywhere on U_f, it ought to be invertible as a function on U_f.

This strongly suggests that we define the nice functions on U_f by adjoining the inverse of f to R. Thus, we define:

$$\mathcal{O}(U_f) = R_f$$

where R_f is the localization of R at f.

By assigning the localized ring R_f to the distinguished open set U_f, we define a *sheaf of rings* on $\mathrm{Spec}(R)$. This construction allows us to formally and rigorously treat a commutative ring as the ring of global functions on a topological space.

19. Affine Schemes

This lecture continues our exploration of the spectrum of a commutative ring. Having previously defined the topological space $\mathrm{Spec}(R)$ and identified its basis of open sets, we now formalize the idea of assigning a "ring of functions" to each open set. This construction yields the *affine scheme* associated with R.

19.1. **Functions on Open Sets and Restriction.** Recall that for an element $f \in R$, we defined the basic open set U_f as the set of prime ideals not containing f:

$$U_f = \{\mathfrak{p} \in \mathrm{Spec}(R) \mid f \notin \mathfrak{p}\}$$

Informally, U_f is the set of points where the function f is non-zero. To each such open set, we assign a ring of functions $\mathcal{O}(U_f)$, which is defined to be the localization of R at f:

$$\mathcal{O}(U_f) = R_f = R[f^{-1}]$$

For this assignment to behave like a true space of functions, it must satisfy three properties: restriction, the presheaf property, and the sheaf property.

19.1.1. *Restriction Property.* If we have an inclusion of open sets, say $U_{f_1 f_2} \subseteq U_{f_1}$, we must have a corresponding restriction map of functions in the opposite direction:

$$\mathcal{O}(U_{f_1}) \to \mathcal{O}(U_{f_1 f_2})$$

Since there is a natural localization homomorphism mapping R_{f_1} to $R_{f_1 f_2}$ (by further inverting f_2), this restriction map clearly exists and satisfies the obvious composition conditions for nested open sets.

19.2. **The Presheaf Property.** The presheaf condition requires that if an open set U_f is covered by a collection of smaller open sets U_{f_i}, and a function $g \in \mathcal{O}(U_f)$ restricts to zero on every U_{f_i}, then g must have been zero on U_f to begin with.

Proposition 19.1. *The assignment $\mathcal{O}(U_f) = R_f$ satisfies the presheaf property.*

Proof. By replacing R with R_f, we can assume without loss of generality that $f = 1$. Thus, we are considering a collection of open sets U_{f_i} that cover the entire spectrum $\mathrm{Spec}(R)$.

The condition that the U_{f_i} cover $\mathrm{Spec}(R)$ means that no prime (or maximal) ideal contains all the f_i. Consequently, the ideal generated by the f_i is the unit ideal, so there exist elements $a_i \in R$ such that:

$$\sum a_i f_i = 1$$

Now, let $g \in R$ be a global function. Suppose g restricts to 0 in every localized ring $\mathcal{O}(U_{f_i}) = R_{f_i}$. By the definition of localization, this means that for each i, there exists some integer n_i such that:

$$g \cdot f_i^{n_i} = 0 \quad \text{in } R$$

Since the elements f_i generate the unit ideal, their powers $f_i^{n_i}$ also generate the unit ideal. (One can see this by raising the equation $\sum a_i f_i = 1$ to a sufficiently high power). Therefore, there exist elements $b_i \in R$ such that:

$$\sum b_i f_i^{n_i} = 1$$

Multiplying this entire equation by g yields:

$$g = \sum b_i (f_i^{n_i} g)$$

Since $f_i^{n_i} g = 0$ for all i, we conclude that $g = 0$. This confirms the presheaf property. □

19.3. **The Sheaf Property.** The sheaf condition is the gluing property: if we are given functions on each open set U_{f_i} of a cover, and these functions agree on all overlaps $U_{f_i} \cap U_{f_j}$, we must be able to uniquely glue them together into a single continuous function on the union.

Proposition 19.2. *The assignment $\mathcal{O}(U_f) = R_f$ satisfies the sheaf property.*

Proof (for Integral Domains). To simplify the proof significantly, we will assume R is an integral domain. (The theorem holds generally, but the zero-divisor conditions require tedious book-keeping).

As before, we assume $f = 1$, so the U_{f_i} cover $\mathrm{Spec}(R)$. We may also assume without loss of generality that the exponents of our given local fractions are 1, so we are given local functions:

$$\frac{r_i}{f_i} \in \mathcal{O}(U_{f_i})$$

The agreement on the intersections $U_{f_i} \cap U_{f_j} = U_{f_i f_j}$ means that:

$$\frac{r_i}{f_i} = \frac{r_j}{f_j} \quad \text{in } R_{f_i f_j}$$

Because R is an integral domain, this fractional equality strictly implies:

$$r_i f_j = r_j f_i \quad \text{in } R$$

Our goal is to find a global element $r \in R$ such that its restriction to U_{f_i} matches the given data:

$$r = \frac{r_i}{f_i} \quad \text{in } R_{f_i}$$

which is equivalent to showing $r f_i = r_i$ in R.

Because the U_{f_i} cover the spectrum, we have $\sum a_j f_j = 1$ for some $a_j \in R$. Motivated by this, we explicitly construct our candidate global function r as:

$$r = \sum a_j r_j$$

We must now verify that $r f_i = r_i$. We compute:

$$\begin{aligned} r f_i &= \left(\sum a_j r_j \right) f_i \\ &= \sum a_j (r_j f_i) \end{aligned}$$

Using the intersection agreement property $r_j f_i = r_i f_j$, we substitute to get:

$$\begin{aligned} r f_i &= \sum a_j (r_i f_j) \\ &= r_i \left(\sum a_j f_j \right) \end{aligned}$$

Since $\sum a_j f_j = 1$, we obtain $r f_i = r_i$. This demonstrates that the locally defined functions glue uniquely into a global function $r \in R$, satisfying the sheaf condition. □

What we have effectively constructed is the *affine scheme* associated with the ring R. An affine scheme pairs the topological space $\mathrm{Spec}(R)$ with the sheaf of regular functions defined by localizations $\mathcal{O}(U_f) = R_f$.

19.4. **The Algebra-Geometry Dictionary.** To use commutative algebra effectively in algebraic geometry, one must internalize the correspondence between algebraic structures in the ring and geometric structures in the affine scheme.

Algebra (Ring R)	**Geometry (Affine Scheme** $\mathrm{Spec}(R)$**)**
Ring R	Topological space $\mathrm{Spec}(R)$ with its structure sheaf
Prime ideal $\mathfrak{p}$	Point in the space
Maximal ideal $\mathfrak{m}$	Closed point
Element $f \in R$	Function on the space (taking values in $R_{\mathfrak{p}}$) *or* Hypersurface of zeros $V(f)$
Ideal I	Closed algebraic set $V(I)$
Local ring $R_{\mathfrak{p}}$	Germs of functions defined near the point $\mathfrak{p}$
Localization $R_f = R[f^{-1}]$	Functions on the distinguished open set U_f
Localization $S^{-1}R$	Functions on an intersection of open sets
Idempotent e ($e^2 = e$)	Clopen (closed and open) set
Module M	Sheaf over $\mathrm{Spec}(R)$

Remark 19.3. The correspondence for elements $f \in R$ and ideals I is not strictly one-to-one, because multiplying an element by a unit, or replacing an ideal with its radical, yields the exact same geometric zero-locus.

19.5. **Examples:** $\mathrm{Spec}(\mathbb{C}[x])$ **and** $\mathrm{Spec}(\mathbb{Z})$**.** To visualize how the sheaf of functions operates, we compare our standard continuous and discrete examples.

Example 19.4 (The Complex Line)**.** Let $R = \mathbb{C}[x]$, corresponding to the complex affine line. Suppose we take $f(x) = (x-1)(x-4)$.

The open set U_f is the complement of the closed points 1 and 4 in the complex plane. The ring of functions on this open set is:

$$\mathcal{O}(U_f) = \mathbb{C}[x]_{(x-1)(x-4)}$$

This ring consists of rational functions that are regular everywhere except possibly at $x = 1$ and $x = 4$, where they are permitted to have poles. A typical element of this ring is:

$$\frac{x^2}{(x-1)^2(x-4)}$$

This function has a pole of order 2 at $x = 1$ and a pole of order 1 at $x = 4$.

Example 19.5 (The Arithmetic Line)**.** We perform the exact same geometric operation on $R = \mathbb{Z}$. Let $f = 14$.

The open set U_{14} is the spectrum of $\mathbb{Z}$ with the closed prime ideals (2) and (7) removed. The ring of functions on this open set is the localization:

$$\mathcal{O}(U_{14}) = \mathbb{Z}[1/14]$$

This ring consists of rational numbers whose denominators only contain powers of 2 and 7. A typical element is:

$$\frac{9}{28} = \frac{3^2}{2^2 \cdot 7}$$

Under our geometric dictionary, we view the rational number 9/28 as a "function" defined on $\mathrm{Spec}(\mathbb{Z})$. It is regular everywhere except at the primes 2 and 7. Specifically, we can say it has a "pole of order 2" at the prime 2, and a "pole of

order 1" at the prime 7, perfectly mirroring the behavior of the complex rational function.

Part 3. Tensor Products, Artinian Rings, and Primary Decomposition

20. Review of Tensor Products

This lecture reviews the tensor product of modules over a commutative ring. Tensor products allow us to linearize bilinear maps and are a fundamental construction in commutative algebra.

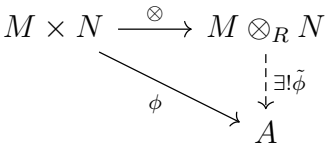
20.1. **Definition and Universal Property.** Let R be a commutative ring, and let M and N be R-modules. The tensor product of M and N over R is denoted by $M \otimes_R N$. When the base ring R is obvious from context, the subscript is frequently omitted, writing simply $M \otimes N$.

Definition 20.1. The *tensor product* $M \otimes_R N$ is the universal module for bilinear maps from $M \times N$. Explicitly, it is an R-module equipped with a bilinear map:

$$M \times N \to M \otimes_R N$$
$$(m, n) \mapsto m \otimes n$$

such that for any R-module A and any bilinear map $\phi\colon M \times N \to A$, there exists a unique linear map $\tilde{\phi}\colon M \otimes_R N \to A$ making the diagram commute. That is,

$$\tilde{\phi}(m \otimes n) = \phi(m, n)$$



Bilinearity means that the map is linear in each argument when the other is held fixed:

$$(r_1 m_1 + r_2 m_2) \otimes n = r_1(m_1 \otimes n) + r_2(m_2 \otimes n)$$
$$m \otimes (r_1 n_1 + r_2 n_2) = r_1(m \otimes n_1) + r_2(m \otimes n_2)$$

The universal property essentially states that bilinear maps from $M \times N$ to A are in natural bijection with linear maps from $M \otimes_R N$ to A. Thus, the tensor product is a mechanism for linearizing bilinear maps.

20.2. **Uniqueness and Existence.**

Proposition 20.2. *The tensor product $M \otimes_R N$ is unique up to canonical isomorphism.*

Proof. Suppose both $M \otimes_R N$ and $M \boxtimes_R N$ satisfy the universal property. The bilinear map $M \times N \to M \boxtimes_R N$ induces a unique linear map $M \otimes_R N \to M \boxtimes_R N$. Reversing their roles, the bilinear map $M \times N \to M \otimes_R N$ induces a unique linear map $M \boxtimes_R N \to M \otimes_R N$.

The composition of these two linear maps must be the identity on $M \otimes_R N$ (by the uniqueness of maps to itself making the trivial diagram commute), and similarly for $M \boxtimes_R N$. Thus, the two modules are canonically isomorphic. This canonical nature allows us to treat the tensor product as essentially unique. □

Proposition 20.3. *The tensor product $M \otimes_R N$ exists.*

Proof. We construct it by forcing existence. Let F be the free R-module generated by the symbols (m, n) for all $m \in M$ and $n \in N$. We then quotient F by the submodule generated by all relations necessary to enforce bilinearity:

$$\begin{aligned}(m_1 + m_2, n) - (m_1, n) - (m_2, n)\\ (m, n_1 + n_2) - (m, n_1) - (m, n_2)\\ (rm, n) - r(m, n)\\ (m, rn) - r(m, n)\end{aligned}$$

The resulting quotient module clearly satisfies the universal property by construction. □

While this construction guarantees existence, it is completely useless for practical calculations, as it involves quotienting an uncountably infinite-rank free module by a similarly massive submodule. We must rely on the universal property to compute tensor products.

20.3. **Warning: Non-Commutative Rings.**

Remark 20.4. The tensor product behaves much more subtly over non-commutative rings. If R is non-commutative, we must define $M \otimes_R N$ using a *right* R-module M and a *left* R-module N. The central relation becomes:

$$mr \otimes n = m \otimes rn$$

The critical problem is that the resulting tensor product $M \otimes_R N$ is generally *only an abelian group*, not an R-module. One cannot multiply by r on the left because M lacks a left module structure, and multiplying on the right interferes with the tensor relation.

To endow the tensor product with an R-module structure, one must use bimodules. If M is an (S, R)-bimodule and N is an (R, T)-bimodule, then $M \otimes_R N$ naturally becomes an (S, T)-bimodule. For commutative rings, any left module is canonically a bimodule, allowing us to safely treat the tensor product of two R-modules as an R-module.

20.4. **Calculating Tensor Products.** We compute tensor products using a series of natural isomorphisms derived from the universal property.

Proposition 20.5. *The tensor product distributes over direct sums and absorbs the base ring:*

$$\begin{aligned}(M_1 \oplus M_2) \otimes_R N &\cong (M_1 \otimes_R N) \oplus (M_2 \otimes_R N)\\ R \otimes_R M &\cong M\end{aligned}$$

Example 20.6 (Vector Spaces)**.** Let $V = k^m$ and $W = k^n$ be vector spaces over a field k. Using the additive and absorbing properties:

$$k^m \otimes_k k^n \cong k^{mn}$$

If V has a basis $v_1, \ldots, v_m$ and W has a basis $w_1, \ldots, w_n$, then $V \otimes_k W$ has a basis consisting of the $m \times n$ pure tensors $v_i \otimes w_j$.

It is critical not to confuse the tensor product with the direct product (or direct sum). The dimension of $V \times W$ is $m + n$, with basis $\{v_i\} \cup \{w_j\}$, whereas the dimension of $V \otimes_k W$ is mn. The natural map $V \times W \to V \otimes_k W$ is bilinear, not linear.

Example 20.7 (Finitely Generated Abelian Groups)**.** Any finitely generated abelian group ($\mathbb{Z}$-module) decomposes into a direct sum of copies of $\mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$. Because tensor products commute with direct sums, we only need to compute three basic cases:

$$\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z} \cong \mathbb{Z}$$
$$\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/n\mathbb{Z}$$
$$\mathbb{Z}/m\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/d\mathbb{Z}$$

where $d = \gcd(m, n)$.

To prove the last isomorphism, note that the module is generated by $1 \otimes 1$. We have:

$$m(1 \otimes 1) = (m \cdot 1) \otimes 1 = 0 \otimes 1 = 0$$
$$n(1 \otimes 1) = 1 \otimes (n \cdot 1) = 1 \otimes 0 = 0$$

Thus, the order of $1 \otimes 1$ divides both m and n, meaning the tensor product is bounded above by $\mathbb{Z}/d\mathbb{Z}$. The natural bilinear map $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/d\mathbb{Z}$ factoring through the tensor product proves equality.

In particular, if m and n are coprime, their tensor product is exactly 0. For example, $\mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/3\mathbb{Z} = 0$.

Generalizing the abelian group result to arbitrary commutative rings yields useful isomorphisms for ideals and quotients:

$$\frac{R}{I} \otimes_R \frac{R}{J} \cong \frac{R}{I + J}$$
$$\frac{R}{I} \otimes_R M \cong \frac{M}{IM}$$

20.5. **Tensor Products and Exactness.** A fundamental question in commutative algebra is how tensor products interact with exact sequences.

Problem 20.8. Suppose we have a short exact sequence of R-modules:

$$0 \to A \to B \to C \to 0$$

If we tensor the entire sequence with an R-module M, is the resulting sequence exact?

$$0 \to A \otimes_R M \to B \otimes_R M \to C \otimes_R M \to 0$$

If R is a field, the answer is yes. Every exact sequence of vector spaces splits, meaning $B \cong A \oplus C$. Since the tensor product preserves direct sums, the exactness is trivially preserved.

However, if R is not a field (e.g., $R = \mathbb{Z}$), the answer is a resounding *no*.

Example 20.9 (The Universal Counterexample)**.** Consider the short exact sequence of $\mathbb{Z}$-modules given by multiplication by 2:

$$0 \to \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

Let us tensor this sequence with $M = \mathbb{Z}/2\mathbb{Z}$. Using our calculation rules, $\mathbb{Z} \otimes \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z}$, and $\mathbb{Z}/2\mathbb{Z} \otimes \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z}$. The tensored sequence becomes:

$$0 \to \frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\times 2} \frac{\mathbb{Z}}{2\mathbb{Z}} \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

However, multiplication by 2 in the ring $\mathbb{Z}/2\mathbb{Z}$ is simply the zero map! Thus, the map is not injective, and the initial 0 must be crossed out. The sequence loses exactness on the left.

The failure of the tensor product to preserve exactness on the left (while preserving it on the right) is a major structural feature of module theory. The entire field of homological algebra was developed largely to measure and manage this precise defect.

21. Tensor Products and Exactness

This lecture investigates the exactness properties of tensor products and the Hom functor. We will demonstrate that these operations preserve exactness on one side but generally fail on the other. We will then establish the exactness properties rigorously and conclude by examining how tensor products commute with direct limits.

21.1. **Failures of Exactness.** In the previous lecture, we considered the short exact sequence of $\mathbb{Z}$-modules:

$$0 \to \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

21.1.1. *Tensoring with a Module.* If we tensor this sequence with $\mathbb{Z}/2\mathbb{Z}$, we obtain:

$$0 \to \mathbb{Z} \otimes \frac{\mathbb{Z}}{2\mathbb{Z}} \to \mathbb{Z} \otimes \frac{\mathbb{Z}}{2\mathbb{Z}} \to \frac{\mathbb{Z}}{2\mathbb{Z}} \otimes \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

Evaluating the tensor products, this simplifies to:

$$0 \to \frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\times 2} \frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\mathrm{id}} \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

The problem arises at the first non-trivial map. Multiplication by 2 on $\mathbb{Z}/2\mathbb{Z}$ is the zero map, which is not injective. Therefore, we must remove the initial 0 because the sequence is not exact on the left.

21.1.2. *The Hom Functor (Covariant).* Similarly, we can apply the functor $\mathrm{Hom}(\mathbb{Z}/2\mathbb{Z}, -)$ to our original exact sequence:

$$0 \to \mathrm{Hom}\left(\frac{\mathbb{Z}}{2\mathbb{Z}}, \mathbb{Z}\right) \to \mathrm{Hom}\left(\frac{\mathbb{Z}}{2\mathbb{Z}}, \mathbb{Z}\right) \to \mathrm{Hom}\left(\frac{\mathbb{Z}}{2\mathbb{Z}}, \frac{\mathbb{Z}}{2\mathbb{Z}}\right) \to 0$$

Since there are no non-trivial homomorphisms from a torsion module to a free module, we get:

$$0 \to 0 \to 0 \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

This is fine on the left, but the map from 0 to $\mathbb{Z}/2\mathbb{Z}$ is clearly not surjective. Thus, we must cross off the final 0; the functor is not exact on the right.

21.1.3. *The Hom Functor (Contravariant).* Alternatively, we can take homomorphisms *from* the exact sequence *to* $\mathbb{Z}/2\mathbb{Z}$. Recall that if we have a homomorphism $A \to B$, composition yields a map $\mathrm{Hom}(B, X) \to \mathrm{Hom}(A, X)$. Thus, the functor $\mathrm{Hom}(-, X)$ reverses the direction of all arrows.

Applying $\mathrm{Hom}(-, \mathbb{Z}/2\mathbb{Z})$ to $\mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$, we get:

$$0 \to \mathrm{Hom}\left(\frac{\mathbb{Z}}{2\mathbb{Z}}, \frac{\mathbb{Z}}{2\mathbb{Z}}\right) \to \mathrm{Hom}\left(\mathbb{Z}, \frac{\mathbb{Z}}{2\mathbb{Z}}\right) \to \mathrm{Hom}\left(\mathbb{Z}, \frac{\mathbb{Z}}{2\mathbb{Z}}\right) \to 0$$

Evaluating these yields:

$$0 \to \frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\mathrm{id}} \frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\times 2} \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

Again, multiplication by 2 is the zero map, which is not surjective onto $\mathbb{Z}/2\mathbb{Z}$. We must cross off the final 0.

Definition 21.1. Operations that preserve exactness on the right are called "right exact". Operations that preserve exactness on the left are called "left exact".

21.2. **Exactness of the Hom Functor.** We now formally verify the left exactness of the Hom functors.

Proposition 21.2. *Suppose we have an exact sequence:*

$$0 \to A \to B \to C$$

Then for any module M, the following sequence is exact:

$$0 \to \mathit{Hom}(M, A) \to \mathit{Hom}(M, B) \to \mathit{Hom}(M, C)$$

Proof. Let $f \in \mathrm{Hom}(M, B)$. We need to show that if f maps to 0 in $\mathrm{Hom}(M, C)$, then f is in the image of $\mathrm{Hom}(M, A)$.

The image of f in $\mathrm{Hom}(M, C)$ being zero means that the composition $M \xrightarrow{f} B \to C$ is zero. Thus, the image of f in B lies entirely within the kernel of $B \to C$. By the exactness of the original sequence, the kernel of $B \to C$ is exactly the image of A. Treating A as a submodule of B (since $A \to B$ is injective), the image of f lies in A. Therefore, f factors through A, meaning it is the image of some map in $\mathrm{Hom}(M, A)$. The other exactness conditions at A are completely routine. □

Proposition 21.3. *Suppose we have an exact sequence:*

$$A \to B \to C \to 0$$

Then for any module M, the following sequence is exact:

$$0 \to Hom(C, M) \to Hom(B, M) \to Hom(A, M)$$

Proof. Take $f \in \mathrm{Hom}(B, M)$ and suppose its image is 0 in $\mathrm{Hom}(A, M)$. This means f vanishes on the image of A in B. By exactness, the image of A is the kernel of the map to C. Thus, f vanishes on the kernel of $B \to C$, meaning f descends to a well-defined homomorphism on the quotient $B/\mathrm{im}(A) \cong C$. More precisely, f is the image of something in $\mathrm{Hom}(C, M)$. □

21.3. **Right Exactness of Tensor Products.** The proof of exactness for tensor products uses the exactness of Hom via the property of adjoint functors.

Theorem 21.4. *Suppose we have an exact sequence:*

$$A \to B \to C \to 0$$

Then for any module M, the tensored sequence is exact:

$$M \otimes A \to M \otimes B \to M \otimes C \to 0$$

Proof. We use the two previous results. First, pick an arbitrary module X. Applying $\mathrm{Hom}(-, X)$ to our exact sequence yields the exact sequence:

$$0 \to \mathrm{Hom}(C, X) \to \mathrm{Hom}(B, X) \to \mathrm{Hom}(A, X)$$

Now, apply $\mathrm{Hom}(M, -)$ to this new sequence. Since $\mathrm{Hom}(M, -)$ is left exact, we obtain another exact sequence:

$$0 \to \mathrm{Hom}(M, \mathrm{Hom}(C, X)) \to \mathrm{Hom}(M, \mathrm{Hom}(B, X)) \to \mathrm{Hom}(M, \mathrm{Hom}(A, X))$$

We now invoke *adjointness*. The space of linear maps from M to $\mathrm{Hom}(A, X)$ is canonically isomorphic to the space of bilinear maps from $M \times A \to X$, which by definition is isomorphic to the linear maps from $M \otimes A \to X$. Thus, we have the identity:

$$\mathrm{Hom}(M, \mathrm{Hom}(A, X)) \cong \mathrm{Hom}(M \otimes A, X)$$

Substituting this adjoint identity into our sequence, we find that the following sequence is exact for any module X:

$$0 \to \mathrm{Hom}(M \otimes C, X) \to \mathrm{Hom}(M \otimes B, X) \to \mathrm{Hom}(M \otimes A, X)$$

If we trace the definitions, this says that homomorphisms from $M \otimes C$ to X are the same as homomorphisms from $M \otimes B$ to X that vanish on the image of $M \otimes A$. This is exactly the universal property of a quotient module, which implies that $M \otimes C$ is precisely the quotient of $M \otimes B$ by the image of $M \otimes A$.

Therefore, the sequence:

$$M \otimes A \to M \otimes B \to M \otimes C \to 0$$

is right exact. □

Remark 21.5. In category theory, the functor taking the tensor product with M is the *left adjoint* to the functor $\mathrm{Hom}(M, -)$. A general theorem states that left adjoints preserve colimits, and the quotient is a special case of a colimit. What the argument here is really doing is giving the category theoretic arguments in this particular case.

21.4. **Calculating Tensor Products.** The right exactness of the tensor product makes it highly computable. Suppose we want to calculate $A \otimes M$. We can present A as a quotient of a free module by mapping another free module onto the relations:

$$R^m \to R^n \to A \to 0$$

By right exactness, tensoring this sequence with M yields:

$$R^m \otimes M \to R^n \otimes M \to A \otimes M \to 0$$

Since tensoring commutes with direct sums, $R^n \otimes M \cong M^n$. Thus, we have an exact sequence:

$$M^m \to M^n \to A \otimes M \to 0$$

This expresses $A \otimes M$ explicitly as a quotient of two modules. If we know enough about A and M, this presentation allows us to compute the tensor product directly.

21.5. **Tensor Products and Direct Limits.** Another useful property for computation is that tensor products commute with direct limits.

Suppose we have a direct system of modules and maps:

$$A_1 \to A_2 \to A_3 \to \dots$$

The *direct limit*, denoted $\varinjlim A_i$, can be thought of as the disjoint union of the A_i modulo the relation that $a_i \in A_i$ is equivalent to its image $a_j \in A_j$ under the system's maps. If the maps are injective, the direct limit is simply the union.

Proposition 21.6. *Tensor products commute with direct limits:*

$$\left(\varinjlim A_i\right) \otimes M \cong \varinjlim (A_i \otimes M)$$

21.5.1. *Example:* $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q}$. Let us calculate $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q}$ over the integers. We can express $\mathbb{Q}$ as a direct limit of copies of $\mathbb{Z}$, where the maps are multiplications by successive integers:

$$\mathbb{Q} \cong \varinjlim \left(\mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \xrightarrow{\times 3} \mathbb{Z} \xrightarrow{\times 4} \dots\right)$$

Now, tensor this direct system with $\mathbb{Q}$. Since tensoring commutes with direct limits:

$$\mathbb{Q} \otimes \mathbb{Q} \cong \varinjlim \left(\mathbb{Z} \otimes \mathbb{Q} \xrightarrow{\times 2} \mathbb{Z} \otimes \mathbb{Q} \xrightarrow{\times 3} \dots\right)$$

Since $\mathbb{Z} \otimes \mathbb{Q} \cong \mathbb{Q}$, the system becomes:

$$\varinjlim \left(\mathbb{Q} \xrightarrow{\times 2} \mathbb{Q} \xrightarrow{\times 3} \mathbb{Q} \xrightarrow{\times 4} \dots\right)$$

Every map here is an isomorphism. The direct limit of a sequence of isomorphisms is just the space itself. Thus:

$$\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q} \cong \mathbb{Q}$$

21.5.2. *A Trap:* $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z}$. Consider the tensor product $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z}$. Following the same procedure, we write $\mathbb{Q}$ as a direct limit of $\mathbb{Z}$ and tensor the system with $\mathbb{Z}/2\mathbb{Z}$:

$$\mathbb{Q} \otimes \frac{\mathbb{Z}}{2\mathbb{Z}} \cong \varinjlim \left(\mathbb{Z} \otimes \frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\times 2} \mathbb{Z} \otimes \frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\times 3} \dots \right)$$

This simplifies to:

$$\varinjlim \left(\frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\times 2} \frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\times 3} \frac{\mathbb{Z}}{2\mathbb{Z}} \xrightarrow{\times 4} \frac{\mathbb{Z}}{2\mathbb{Z}} \dots \right)$$

If we are sloppy and assume the limit is simply the union, we might incorrectly guess the answer is $\mathbb{Z}/2\mathbb{Z}$. However, taking tensor products does not necessarily preserve injectivity.

The maps in this direct limit are multiplication by $2, 3, 4, \dots$ evaluated in the ring $\mathbb{Z}/2\mathbb{Z}$. This sequence of maps is:

$$\times 0, \times 1, \times 0, \times 1, \dots$$

Since every second map is the zero map, any element introduced at stage i gets mapped to 0 at stage $i+1$ (or $i+2$). Because every element eventually becomes zero, the direct limit is actually zero:

$$\mathbb{Q} \otimes_{\mathbb{Z}} \frac{\mathbb{Z}}{2\mathbb{Z}} = 0$$

22. Flatness, Tensor Products, Localization

This lecture explores the relationship between tensor products and localization, introducing the fundamental concept of flatness.

22.1. **Flat Modules.** We begin by defining what it means for a module to be flat. As we have seen, the tensor product is right exact, but it generally fails to be left exact. Flat modules are precisely those for which the tensor product preserves full exactness.

Definition 22.1. An R-module M is called *flat* if taking the tensor product with M preserves exactness. That is, if the sequence

$$0 \to A \to B \to C \to 0$$

is exact, then the tensored sequence

$$0 \to M \otimes_R A \to M \otimes_R B \to M \otimes_R C \to 0$$

is also exact.

Since the right-hand side is always exact for any module M, the key consequence of flatness is that it preserves injectivity. Specifically, if A is a submodule of B, then $M \otimes_R A$ embeds naturally as a submodule of $M \otimes_R B$.

Example 22.2.
- Over the integers $\mathbb{Z}$, the module $\mathbb{Z}/2\mathbb{Z}$ is *not* exact (not flat), as tensoring with it destroys the injectivity of the map $\mathbb{Z} \xrightarrow{\times 2} \mathbb{Z}$.
- Over any ring R, the ring R itself is flat, because tensoring an R-module A with R yields $R \otimes_R A \cong A$, which trivially preserves exactness.

Flatness as a formal concept was introduced by Jean-Pierre Serre in the 1950s, and Alexander Grothendieck subsequently demonstrated that it is absolutely fundamental in algebraic geometry. Informally, if we view a module M as a family of localizations $M_{\mathfrak{p}}$ over the local rings of a spectrum, flatness dictates that this family varies "nicely" without jumping suddenly or behaving pathologically. Consequently, when studying families of modules or schemes, imposing a flatness condition forces the family to be well-behaved.

22.2. **Localization of Modules.** The main result of this lecture is that the localized ring $S^{-1}R$ is a flat R-module. To prove this, we first define the localization of an arbitrary module.

Definition 22.3. Let S be a multiplicative subset of R, and M an R-module. The *localization* $S^{-1}M$ is constructed similarly to $S^{-1}R$. The elements are equivalence classes of fractions m/s for $m \in M$ and $s \in S$, where:

$$\frac{m_1}{s_1} = \frac{m_2}{s_2} \iff s(m_1 s_2 - m_2 s_1) = 0 \quad \text{for some } s \in S$$

Addition and scalar multiplication follow the standard rules of arithmetic. The localization of M at a prime ideal $\mathfrak{p}$ is denoted $M_{\mathfrak{p}}$, where $S = R \backslash \mathfrak{p}$. Geometrically, $M_{\mathfrak{p}}$ acts as the *stalk* of M at the prime $\mathfrak{p}$. Just as localizing R assigns a ring of functions to open sets, localizing M constructs a quasi-coherent sheaf of modules over the spectrum of R.

The critical property of module localization is that it is an exact functor.

Proposition 22.4. *The map $M \mapsto S^{-1}M$ preserves exactness. If the sequence*

$$0 \to A \to B \to C \to 0$$

is exact, then the localized sequence

$$0 \to S^{-1}A \to S^{-1}B \to S^{-1}C \to 0$$

is also exact.

Proof. The main problem is to verify exactness at $S^{-1}B$. Suppose an element $b/s \in S^{-1}B$ maps to 0 in $S^{-1}C$. If the image of b in C is c, this means $c/s = 0$ in $S^{-1}C$. By the definition of fractions, there exists $s_1 \in S$ such that:

$$s_1 c = 0 \quad \text{in } C$$

This implies that the element $s_1 b \in B$ maps to 0 in C. Because the original sequence is exact at B, $s_1 b$ must lie in the image of A. Thus, there exists some $a \in A$ mapping to $s_1 b$. In the localized module $S^{-1}B$, we can write:

$$\frac{b}{s} = \frac{s_1 b}{s_1 s} = \frac{a}{s_1 s}$$

This shows that b/s is in the image of $S^{-1}A$, proving that localization preserves exactness. □

22.3. **Flatness of Localizations.** We can now prove our main theorem regarding localization and flatness.

Theorem 22.5. *For any multiplicative subset $S \subset R$, the localization $S^{-1}R$ is a flat R-module.*

Proof. We want to show that mapping a module A to the tensor product $A \otimes_R S^{-1}R$ preserves exactness. We have already proven that mapping A to $S^{-1}A$ preserves exactness. Therefore, it suffices to show that there is a canonical isomorphism:

$$S^{-1}A \cong A \otimes_R S^{-1}R$$

We construct mutually inverse maps. We can map:

$$\frac{a}{s} \mapsto a \otimes \frac{1}{s}$$
$$a \otimes \frac{r}{s} \mapsto \frac{ar}{s}$$

It is straightforward to check that these maps are well-defined and inverse to one another. Since tensoring with $S^{-1}R$ is canonically equivalent to localizing with respect to S, and localization preserves exactness, $S^{-1}R$ is a flat R-module. □

Remark 22.6. While localizing is a very well-behaved and flat operation, taking quotients is not. If I is an ideal, the quotient module A/IA is isomorphic to $A \otimes_R R/I$. Because R/I is generally not flat (e.g., $\mathbb{Z}/2\mathbb{Z}$), taking quotients does not preserve exactness.

22.4. **Flat Modules and Torsion.** What do flat modules look like? Over an integral domain, we can establish a basic necessary condition.

Proposition 22.7. *Let R be an integral domain. If M is a flat R-module, then M is* torsion-free.

Proof. Let $a \in R$ be a non-zero element. Because R is an integral domain, multiplication by a on R is injective, giving an exact sequence:

$$0 \to R \xrightarrow{\times a} R \to R/aR \to 0$$

Since M is flat, tensoring this sequence with M preserves exactness:

$$0 \to M \otimes_R R \xrightarrow{\times a} M \otimes_R R$$

Because $M \otimes_R R \cong M$, this implies that the multiplication by a map on M is injective:

$$0 \to M \xrightarrow{\times a} M$$

This means that no non-zero element of M is annihilated by a. Since this holds for all $a \neq 0$, M is torsion-free. □

For principal ideal domains (like $\mathbb{Z}$), the converse holds: a module is flat if and only if it is torsion-free. In particular, finitely generated flat modules over $\mathbb{Z}$ are exactly the free modules. However, over more general rings, torsion-free modules are not necessarily flat. Flatness is a much stronger and better-behaved property, which is why it is vastly preferred in commutative algebra.

22.5. **Local Properties.** We conclude with three extremely useful properties relating flatness, exactness, and localization. These properties are summarized by the catchphrase: "vanishing, exactness, and flatness are local properties".

Theorem 22.8. *Let M be an R-module.*

1. ***Vanishing is local:*** *$M = 0$ if and only if $M_{\mathfrak{m}} = 0$ for all maximal ideals $\mathfrak{m}$.*
2. ***Exactness is local:*** *A sequence $0 \to A \to B \to C \to 0$ is exact if and only if $0 \to A_{\mathfrak{m}} \to B_{\mathfrak{m}} \to C_{\mathfrak{m}} \to 0$ is exact for all maximal ideals $\mathfrak{m}$.*
3. ***Flatness is local:*** *M is flat if and only if $M_{\mathfrak{m}}$ is a flat $R_{\mathfrak{m}}$-module for all maximal ideals $\mathfrak{m}$.*

Proof. **(1) Vanishing is local:** If $M = 0$, its localizations are trivially zero. Conversely, assume $M_{\mathfrak{m}} = 0$ for all maximal ideals $\mathfrak{m}$. Pick any $x \in M$. Since x maps to 0 in $M_{\mathfrak{m}}$, its annihilator ideal must not be contained in $\mathfrak{m}$. This holds for every maximal ideal, meaning $\mathrm{Ann}(x)$ is not contained in any maximal ideal. Thus, $\mathrm{Ann}(x) = R$, which implies $1 \cdot x = 0$, so $x = 0$. Since every element is zero, $M = 0$.

(2) Exactness is local: One direction follows immediately because localization preserves exactness. For the other direction, observe that a sequence is exact at B if and only if the homology module:

$$H = \frac{\ker(B \to C)}{\mathrm{im}(A \to B)} = 0$$

Since localization is flat, localizing the homology module commutes with taking kernels and images:

$$H_{\mathfrak{m}} = \frac{\ker(B_{\mathfrak{m}} \to C_{\mathfrak{m}})}{\mathrm{im}(A_{\mathfrak{m}} \to B_{\mathfrak{m}})}$$

If the sequence is exact locally everywhere, then $H_{\mathfrak{m}} = 0$ for all maximal ideals. By part (1), this forces $H = 0$, meaning the global sequence is exact.

(3) Flatness is local: Suppose we want to check if M preserves the exactness of an arbitrary sequence $0 \to A \to B \to C \to 0$. Tensoring this sequence with M is exact if and only if the tensored sequence is exact locally at every maximal ideal $\mathfrak{m}$ (by part 2). The local sequence at $\mathfrak{m}$ is:

$$0 \to (A \otimes_R M)_{\mathfrak{m}} \to (B \otimes_R M)_{\mathfrak{m}} \to (C \otimes_R M)_{\mathfrak{m}} \to 0$$

By associativity of tensor products and localization, this sequence is canonically isomorphic to:

$$0 \to A_{\mathfrak{m}} \otimes_{R_{\mathfrak{m}}} M_{\mathfrak{m}} \to B_{\mathfrak{m}} \otimes_{R_{\mathfrak{m}}} M_{\mathfrak{m}} \to C_{\mathfrak{m}} \otimes_{R_{\mathfrak{m}}} M_{\mathfrak{m}} \to 0$$

This local sequence is exact for all exact A, B, C if and only if $M_{\mathfrak{m}}$ is a flat $R_{\mathfrak{m}}$-module for all maximal ideals $\mathfrak{m}$. □

23. Flat Extensions

This lecture discusses flat extensions of rings. Suppose we have two rings R and S, and a homomorphism between them, so we can consider S as an R-algebra (or an extension of R). The fundamental problem we want to discuss is how we relate R-modules to S-modules.

There are obvious ways to transition between these categories. We can convert an S-module into an R-module simply by restricting the ring action (restriction of scalars). On the other hand, if we have an R-module M, we can convert it into an S-module by taking the tensor product $S \otimes_R M$ (extension of scalars). This operation is, in fact, the left adjoint to the restriction functor.

23.1. **Comparing Spaces of Homomorphisms.** We want to know how the spaces of homomorphisms are related when we pass from R-modules to S-modules. Suppose we have two R-modules, M and N.

- We can construct the module of R-homomorphisms: $\mathrm{Hom}_R(M, N)$.
- We can extend scalars and construct the module of S-homomorphisms: $\mathrm{Hom}_S(S \otimes_R M, S \otimes_R N)$.

First, there is an obvious functorial map from one to the other. A homomorphism $\phi\colon M \to N$ naturally induces an S-module homomorphism $\mathrm{id}_S \otimes \phi\colon S \otimes_R M \to S \otimes_R N$. We can ask: is this induced map an isomorphism?

$$\mathrm{Hom}_R(M, N) \to \mathrm{Hom}_S(S \otimes_R M, S \otimes_R N)$$

In general, there is absolutely no reason why this should be an isomorphism. The target space is an S-module, but the domain $\mathrm{Hom}_R(M, N)$ is merely an R-module; there is no S-module structure anywhere in sight on the left-hand side.

To fix this, we must formally convert the left-hand side into an S-module by tensoring it with S. We thus consider the natural map:

$$S \otimes_R \mathrm{Hom}_R(M, N) \to \mathrm{Hom}_S(S \otimes_R M, S \otimes_R N) \tag{23.1}$$

Because we have applied the extension functor to both sides, there is now a reasonable chance that this map might be an isomorphism. However, in general, it is still not an isomorphism.

23.2. **A Counterexample.** We can demonstrate the failure of the isomorphism with a simple counterexample. Let $R = \mathbb{Z}$ and $S = \mathbb{Z}/2\mathbb{Z}$. Let our R-modules be $M = \mathbb{Z}/2\mathbb{Z}$ and $N = \mathbb{Z}$.

First, we compute the space of R-homomorphisms:

$$\mathrm{Hom}_R(M, N) = \mathrm{Hom}_{\mathbb{Z}}\left(\frac{\mathbb{Z}}{2\mathbb{Z}}, \mathbb{Z}\right) = 0$$

Because there are no non-zero homomorphisms from a torsion module to a free module, the left side vanishes. Consequently, extending scalars yields zero:

$$S \otimes_R \mathrm{Hom}_R(M, N) = \frac{\mathbb{Z}}{2\mathbb{Z}} \otimes_{\mathbb{Z}} 0 = 0$$

Now we evaluate the right-hand side. We tensor M and N with S:

$$S \otimes_R M = \frac{\mathbb{Z}}{2\mathbb{Z}} \otimes_{\mathbb{Z}} \frac{\mathbb{Z}}{2\mathbb{Z}} \cong \frac{\mathbb{Z}}{2\mathbb{Z}}$$
$$S \otimes_R N = \frac{\mathbb{Z}}{2\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbb{Z} \cong \frac{\mathbb{Z}}{2\mathbb{Z}}$$

The module of S-homomorphisms between these extended modules is:

$$\mathrm{Hom}_S(S \otimes_R M, S \otimes_R N) = \mathrm{Hom}_{\mathbb{Z}/2\mathbb{Z}}\left(\frac{\mathbb{Z}}{2\mathbb{Z}}, \frac{\mathbb{Z}}{2\mathbb{Z}}\right) \cong \frac{\mathbb{Z}}{2\mathbb{Z}}$$

Comparing the two sides, we see that $0 \ncong \mathbb{Z}/2\mathbb{Z}$. Thus, the map is not an isomorphism.

23.3. **The Isomorphism Theorem.** Despite the failure in the general case, there is one highly important case where the map is an isomorphism.

Theorem 23.1. *The natural map*

$$S \otimes_R \mathrm{Hom}_R(M, N) \to \mathrm{Hom}_S(S \otimes_R M, S \otimes_R N)$$

is an isomorphism provided that:

(1) S is a flat *R-module.*
(2) M is a finitely presented *R-module.*

Definition 23.2. An R-module S is *flat* if tensoring with S preserves exact sequences.

Definition 23.3. An R-module M is *finitely presented* if there exists an exact sequence:

$$R^m \to R^n \to M \to 0$$

where both m and n are finite. This means M is not only finitely generated (by n elements), but the kernel of that generating map (the module of relations) is also finitely generated (by m relations).

Remark 23.4. If R is a Noetherian ring, the condition of being finitely generated automatically implies being finitely presented. However, for non-Noetherian rings, finite presentation is the correct, stronger condition needed here.

Proof. We proceed in steps, proving the theorem for increasingly complex modules M.

Step 1: $M = R$. If $M = R$, the left side is:

$$S \otimes_R \mathrm{Hom}_R(R, N) \cong S \otimes_R N$$

The right side is:

$$\mathrm{Hom}_S(S \otimes_R R, S \otimes_R N) \cong \mathrm{Hom}_S(S, S \otimes_R N) \cong S \otimes_R N$$

Both sides are naturally isomorphic to $S \otimes_R N$, making the map an isomorphism.

Step 2: $M = R^n$ (where n is finite). The functors $\mathrm{Hom}(-,\cdot)$ and $S \otimes_R -$ are additive; they commute with finite direct sums. Since the theorem holds for $M = R$, it naturally extends to any finite direct sum $M = R \oplus \cdots \oplus R = R^n$. Note critically that n must be finite; infinite direct sums generally do not pull out of the first slot of the Hom functor as direct sums, but rather as direct products.

Step 3: Arbitrary finitely presented M. We start with the finite presentation exact sequence for M:

$$R^m \to R^n \to M \to 0$$

We apply the functor $\mathrm{Hom}_R(-, N)$. Since this functor is contravariant and left exact, it reverses the arrows and preserves exactness at the first two steps:

$$0 \to \mathrm{Hom}_R(M, N) \to \mathrm{Hom}_R(R^n, N) \to \mathrm{Hom}_R(R^m, N)$$

Next, we tensor this sequence with S. Here we critically use the hypothesis that S is *flat*. Flatness guarantees that tensoring with S preserves the exactness of the sequence (including the zero on the left):

$$0 \to S \otimes_R \mathrm{Hom}_R(M, N) \to S \otimes_R \mathrm{Hom}_R(R^n, N) \to S \otimes_R \mathrm{Hom}_R(R^m, N) \quad (23.2)$$

For the second part, we return to the presentation $R^m \to R^n \to M \to 0$ and first tensor it with S. Since the tensor product is right exact, we obtain an exact sequence of S-modules:

$$S \otimes_R R^m \to S \otimes_R R^n \to S \otimes_R M \to 0$$

Now we apply the functor $\mathrm{Hom}_S(-, S \otimes_R N)$ to this sequence. This is left exact, yielding:

$$0 \to \mathrm{Hom}_S(S \otimes M, S \otimes N) \to \mathrm{Hom}_S(S \otimes R^n, S \otimes N) \to \mathrm{Hom}_S(S \otimes R^m, S \otimes N) \quad (23.3)$$

We now place sequence (23.2) on top of sequence (23.3). The natural transformation between our functors provides vertical maps bridging the rows, forming a commutative diagram:

$$\begin{array}{ccccccc}
0 & \longrightarrow & S \otimes \mathrm{Hom}(M, N) & \longrightarrow & S \otimes \mathrm{Hom}(R^n, N) & \longrightarrow & S \otimes \mathrm{Hom}(R^m, N) \\
 & & \downarrow & & \downarrow \cong & & \downarrow \cong \\
0 & \longrightarrow & \mathrm{Hom}_S(S \otimes M, S \otimes N) & \longrightarrow & \mathrm{Hom}_S(S \otimes R^n, S \otimes N) & \longrightarrow & \mathrm{Hom}_S(S \otimes R^m, S \otimes N)
\end{array}$$

By Step 2, the vertical maps for the free modules R^n and R^m are isomorphisms. We want to conclude that the first vertical map is also an isomorphism. This setup begs for the Five Lemma. □

23.4. **The Five Lemma and Diagram Chasing.** The *Five Lemma* is an essential tool in homological algebra. It states that given a commutative diagram with exact rows of the following form:

$$\begin{array}{ccccccccc} A & \longrightarrow & B & \longrightarrow & C & \longrightarrow & D & \longrightarrow & E \\ \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \beta} & & \downarrow{\scriptstyle \gamma} & & \downarrow{\scriptstyle \delta} & & \downarrow{\scriptstyle \epsilon} \\ A' & \longrightarrow & B' & \longrightarrow & C' & \longrightarrow & D' & \longrightarrow & E' \end{array}$$

If α, β, δ, and ϵ are isomorphisms, then γ is also an isomorphism.

In our specific case, the diagram has a 0 on the left side. By padding the left with an extra column of zeros, our setup fits exactly into the Five Lemma (with α and β being the identity maps on 0, which are trivially isomorphisms).

The Five Lemma is proven using a technique called "diagram chasing". It generally splits into two sub-lemmas (one for injectivity, one for surjectivity). We will demonstrate the technique for surjectivity.

Lemma 23.5. *If the rows are exact, β and δ are surjective, and ϵ is injective, then γ is surjective.*

Proof. We pick an element $c' \in C'$ and we want to show there exists $c \in C$ such that $\gamma(c) = c'$. The proof traces a complex path through the diagram (cf. the lecture recording for an illustration):

(1) Let d' be the image of c' in D'.
(2) By exactness at C', the image of d' in E' is 0.
(3) Since δ is surjective, there exists $d \in D$ such that $\delta(d) = d'$.
(4) The image of d in E maps via ϵ to $0 \in E'$ (because the diagram commutes). Since ϵ is injective, the image of d in E must be 0.
(5) By exactness at D, since $d \mapsto 0 \in E$, there must exist some element $c \in C$ mapping to d.
(6) We compute $\gamma(c)$. By commutativity, the image of $\gamma(c)$ in D' is exactly d'.
(7) This means c' and $\gamma(c)$ both map to the same element $d' \in D'$. Their difference $c' - \gamma(c)$ maps to $0 \in D'$.
(8) By exactness at C', $c' - \gamma(c)$ must be the image of some element $b' \in B'$.
(9) Since β is surjective, there exists $b \in B$ such that $\beta(b) = b'$.
(10) Let the image of b in C be $\tilde{c}$. Then by commutativity, $\gamma(\tilde{c}) = b' = c' - \gamma(c)$.
(11) Rearranging, we get $\gamma(c + \tilde{c}) = c'$. Thus, $c + \tilde{c} \in C$ is our desired preimage, proving γ is surjective.

□

A similarly intricate diagram chase establishes injectivity (requiring the left-side maps). Combining them completes the Five Lemma, which in turn completes the proof of Theorem 23.1.

23.5. **Failure for Infinitely Generated Modules.** We finish with an example to demonstrate what happens if M is not finitely presented. If M is infinitely generated, the isomorphism usually fails, even with nice, flat rings over fields.

Let $R = \mathbb{Q}$, which is a field, meaning all modules are flat. Let $S = \mathbb{Q}[x]$, the polynomial ring, which is a flat R-algebra. Let M be an infinite-dimensional vector space over $\mathbb{Q}$, meaning M is not finitely generated. Let $N = \mathbb{Q}$, a one-dimensional vector space.

We compare the two sides of our mapped relationship:

$$\text{LHS} = \mathbb{Q}[x] \otimes_{\mathbb{Q}} \text{Hom}_{\mathbb{Q}}(M, N)$$
$$\text{RHS} = \text{Hom}_{\mathbb{Q}[x]}(\mathbb{Q}[x] \otimes_{\mathbb{Q}} M, \mathbb{Q}[x] \otimes_{\mathbb{Q}} N)$$

Because M is an infinite-dimensional vector space, it is a direct sum $\bigoplus_{i\in I} \mathbb{Q}$. The dual space $\text{Hom}_{\mathbb{Q}}(M, \mathbb{Q})$ is therefore an infinite *product* of copies of $\mathbb{Q}$, denoted $\prod_{i\in I} \mathbb{Q}$. Thus, the left side is:

$$\text{LHS} = \mathbb{Q}[x] \otimes_{\mathbb{Q}} \left(\prod_{i\in I} \mathbb{Q} \right)$$

On the right side, $S \otimes M$ is a free $\mathbb{Q}[x]$-module of infinite rank, meaning $\bigoplus_{i\in I} \mathbb{Q}[x]$. Homomorphisms from an infinite direct sum to $\mathbb{Q}[x]$ turn into an infinite product:

$$\text{RHS} = \text{Hom}_{\mathbb{Q}[x]} \left(\bigoplus_{i\in I} \mathbb{Q}[x], \mathbb{Q}[x] \right) \cong \prod_{i\in I} \mathbb{Q}[x]$$

We are essentially asking whether tensoring with $\mathbb{Q}[x]$ commutes with an infinite direct product:

$$\mathbb{Q}[x] \otimes_{\mathbb{Q}} \left(\prod \mathbb{Q} \right) \overset{?}{\cong} \prod \left(\mathbb{Q}[x] \otimes_{\mathbb{Q}} \mathbb{Q} \right) \cong \prod \mathbb{Q}[x]$$

While tensor products always commute with infinite *direct sums*, they do not generally commute with infinite *direct products*. We can see the difference explicitly.

A typical element on the left side is a finite sum of tensors, which means it corresponds to a sequence of polynomials where all polynomials have degrees bounded by some maximum integer D. In contrast, an element on the right side is an arbitrary sequence of polynomials, where the degrees can grow arbitrarily large (e.g., $1, x, x^2, x^3, \dots$).

The left space can be viewed as a dense subspace of the right space under a suitable topology, but they are not algebraically equal. This shows that the homomorphisms from M to N behave weirdly under extension if M is not finitely presented. (As a side note, if M has countable dimension, the LHS remains countable, while the infinite product on the RHS becomes uncountably infinite, highlighting how drastically they diverge).

24. Artinian Modules

In this lecture, we introduce Artinian rings and modules, investigating their properties and the profound symmetries—and asymmetries—between the Artinian and Noetherian conditions. We will also examine simple modules and modules of finite length.

24.1. **Definitions and Dual Conditions.** We begin by recalling the characterization of a Noetherian module. A module is Noetherian if it satisfies any of three equivalent conditions: the ascending chain condition (ACC), the property that every non-empty set of submodules has a maximal element, or the condition that every submodule is finitely generated.

We can formalize a "dual" concept by reversing the inclusions.

Definition 24.1. An R-module M is *Artinian* if it satisfies the descending chain condition (DCC) for submodules. That is, any decreasing chain of submodules

$$M_1 \supseteq M_2 \supseteq M_3 \supseteq \dots$$

must eventually stabilize, meaning $M_n = M_{n+1} = \dots$ for some n.

Just as with partially ordered sets for the Noetherian case, the DCC is immediately equivalent to a second condition:

- Every non-empty set of submodules of M has a *minimal* element.

There is, however, no obvious dual to the "finitely generated" condition.

24.2. **Examples of Artinian and Noetherian Modules.** To build intuition, let us classify various modules over the integers (or other rings) into the four possible combinations of these two properties.

24.2.1. *Modules that are both Artinian and Noetherian.*

- The zero module over any ring.
- Modules with a finite number of elements, such as $\mathbb{Z}/n\mathbb{Z}$ over $\mathbb{Z}$.
- Any finite-dimensional vector space over a field. (We will see that modules satisfying both conditions are the natural generalization of finite-dimensional vector spaces).

24.2.2. *Modules that are Noetherian but not Artinian.*

- The ring of integers $\mathbb{Z}$ viewed as a $\mathbb{Z}$-module. We have previously seen it is Noetherian, but it possesses an infinite descending chain:

$$\mathbb{Z} \supsetneq 2\mathbb{Z} \supsetneq 4\mathbb{Z} \supsetneq 8\mathbb{Z} \supsetneq \dots$$

- The localization $\mathbb{Z}_{(2)}$, which consists of rational numbers with odd denominators, is Noetherian but not Artinian over itself.

24.2.3. *Modules that are neither Artinian nor Noetherian.*

- The rational numbers $\mathbb{Q}$ as a $\mathbb{Z}$-module. We can form infinite chains extending in both directions:

$$\dots \supsetneq 8\mathbb{Z} \supsetneq 4\mathbb{Z} \supsetneq 2\mathbb{Z} \supsetneq \mathbb{Z} \supsetneq \frac{1}{2}\mathbb{Z} \supsetneq \frac{1}{4}\mathbb{Z} \supsetneq \dots$$

24.2.4. *Modules that are Artinian but not Noetherian.* These modules tend not to be finitely generated and are often structurally unusual.

Example 24.2. Let $M = \mathbb{Z}[1/2]/\mathbb{Z}$. This module consists of all rational numbers of the form $a/2^n$ modulo 1. We can view this as an infinite union of an ascending chain of cyclic subgroups:

$$\frac{\mathbb{Z}}{\mathbb{Z}} \subsetneq \frac{\frac{1}{2}\mathbb{Z}}{\mathbb{Z}} \subsetneq \frac{\frac{1}{4}\mathbb{Z}}{\mathbb{Z}} \subsetneq \frac{\frac{1}{8}\mathbb{Z}}{\mathbb{Z}} \subsetneq \dots$$

Because there is an infinite strictly increasing chain of submodules, M is explicitly *not* Noetherian. However, it is Artinian, because any strictly decreasing chain

must eventually terminate. This module is known as the injective envelope of $\mathbb{Z}/2\mathbb{Z}$.

24.3. **Artinian Rings.** The definition effortlessly extends to rings.

Definition 24.3. A commutative ring R is an *Artinian ring* if it is an Artinian module over itself; that is, it satisfies the descending chain condition for ideals.

Let us look at the four combinations for rings:

- **Noetherian but not Artinian:** $\mathbb{Z}$.
- **Neither:** The polynomial ring in infinitely many variables $k[x_1, x_2, \dots]$, which can be made arbitrarily "large".
- **Both:** $\mathbb{Z}/n\mathbb{Z}$, any PID modulo a non-zero ideal, and any finite-dimensional k-algebra (since its ideals must be k-vector subspaces, precluding infinite chains).
- **Artinian but not Noetherian:** Surprisingly, *there are no examples.*

Historically, early algebraists (like Emil Artin, after whom rings with the "minimum condition" are named) studied rings assuming both a maximum (Noetherian) and minimum (Artinian) condition. It was only later discovered that the minimum condition automatically enforces the maximum condition.

Theorem 24.4 (Akizuki-Hopkins-Levitzki)**.** *Every Artinian ring is automatically a Noetherian ring.*

We will prove this remarkable fact in the next lecture by classifying Artinian rings completely. To prepare for this, we must further study modules that are simultaneously Artinian and Noetherian.

24.4. **Simple Modules and Finite Length.**

Definition 24.5. A module M is *simple* if $M \neq 0$ and its only submodules are 0 and M.

Over a ring R, the simple modules are exactly those isomorphic to $R/\mathfrak{m}$ for some maximal ideal $\mathfrak{m}$. For instance, $\mathbb{Z}/p\mathbb{Z}$ is a simple $\mathbb{Z}$-module.

Definition 24.6. A module M has *finite length* if there exists a finite composition series of submodules:

$$0 = M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \cdots \subsetneq M_n = M$$

such that each successive quotient M_i/M_{i-1} is a simple module. The integer n is called the *length* of M.

Finite-dimensional vector spaces trivially have finite length. The module $\mathbb{Z}/p^n\mathbb{Z}$ also has finite length, with a chain of submodules whose successive quotients are all isomorphic to the simple module $\mathbb{Z}/p\mathbb{Z}$.

Theorem 24.7. *An R-module M has finite length if and only if it is both Noetherian and Artinian.*

Proof. ($\Rightarrow$) Any simple module is trivially both Noetherian and Artinian. If we have a short exact sequence $0 \to A \to B \to C \to 0$, B is Noetherian (resp. Artinian) if and only if both A and C are. By inducting on the length of the composition series, a module of finite length is built from a finite number of simple modules, and thus must be both Noetherian and Artinian.

($\Leftarrow$) Suppose M is Noetherian and Artinian. If $M = 0$, it has length 0. Otherwise, because M is Artinian, the set of non-zero submodules has a minimal element, say M_1. By definition, M_1 must be simple. If $M_1 \neq M$, we consider the submodules of M strictly containing M_1. Because M is Artinian, there is a minimal such module M_2, making M_2/M_1 simple. We continue this process, generating a strictly increasing chain:

$$0 \subsetneq M_1 \subsetneq M_2 \subsetneq \dots$$

Because M is Noetherian, this ascending chain must eventually terminate at $M_n = M$. Thus, M has finite length. □

24.5. **The Jordan-Hölder Theorem.** A natural worry is that a module might have multiple maximal chains of different lengths. The following theorem, directly analogous to the Jordan-Hölder theorem for finite groups, shows this is impossible.

Theorem 24.8. *If M has finite length, any two composition series of M have the same length and the same multiset of simple quotients (up to isomorphism).*

Proof Sketch. Suppose we have two composition series, M_i (of length m) and N_j (of length n). We can form a rectangular grid of submodules by taking intersections:

$$M_{i,j} = M_{i-1} + (M_i \cap N_j)$$

By analyzing the squares of this grid (evaluating the quotients $M_{i,j}/M_{i,j-1}$), one finds that for each square, either the horizontal quotients are isomorphic and the vertical quotients are zero, or vice-versa. Geometrically, we can trace a "taxicab route" along the edges of this grid from $(0,0)$ to (m,n). Passing from one composition series to the other corresponds to altering the taxicab route one square at a time. Each local adjustment may reorder the simple quotients, but it preserves both their total number and their isomorphism classes. Thus, the two chains must share the exact same simple quotients and the same length. □

Corollary 24.9. *The length of a module is well-defined. Furthermore, length is additive on exact sequences: if $0 \to A \to B \to C \to 0$ is an exact sequence of modules of finite length, then:*

$$\text{length}(B) = \text{length}(A) + \text{length}(C)$$

This additivity firmly establishes the length of a module as the correct ring-theoretic analogue to the dimension of a vector space.

25. Artinian Rings

In this lecture, we focus on Artinian rings. The main theorem we are going to prove is that all Artinian rings are Noetherian. In fact, we prove a slightly more precise version that establishes the equivalence of several structural conditions.

25.1. **Main Theorem and Equivalent Conditions.**

Theorem 25.1. *The following four conditions are equivalent for a commutative ring R:*

(1) R is Noetherian and the zero ideal (0) is a product of maximal ideals.
(2) R is Noetherian and all prime ideals are maximal. (In dimension theory, this condition precisely states that R is zero-dimensional. Thus, Artinian rings are exactly the zero-dimensional Noetherian rings).
(3) R has finite length as an R-module.
(4) R is Artinian.

Proof. We will first prove the three relatively straightforward implications (1) $\implies$ (2) $\implies$ (3) $\implies$ (4), and then tackle the more difficult implication (4) $\implies$ (1).

(1) $\implies$ (2): Suppose R is Noetherian and (0) is a product of maximal ideals. We can write:

$$(0) = \mathfrak{m}_1\mathfrak{m}_2\cdots\mathfrak{m}_n$$

Let $\mathfrak{p}$ be any prime ideal of R. Since $(0) \subseteq \mathfrak{p}$, we have:

$$\mathfrak{m}_1\mathfrak{m}_2\cdots\mathfrak{m}_n \subseteq \mathfrak{p}$$

If a prime ideal contains a product of ideals, it must contain at least one of the factors. Thus, $\mathfrak{p}$ contains some $\mathfrak{m}_i$. However, $\mathfrak{m}_i$ is a maximal ideal, so we must have $\mathfrak{p} = \mathfrak{m}_i$. Therefore, every prime ideal is maximal.

(2) $\implies$ (3): Suppose R is Noetherian and all primes are maximal. Assume for the sake of contradiction that R does not have finite length. Because R is Noetherian, we can choose an ideal I that is maximal among all ideals such that the quotient R/I is *not* of finite length.

We will show that I is a prime ideal. Let $S = R/I$. To prove S is an integral domain, suppose there exist elements $a, b \in S$ such that:

$$ab = 0$$

with $a \neq 0$ and $b \neq 0$. Because a and b are non-zero, the ideals (a) and (b) in S correspond to strictly larger ideals in R containing I. By our maximal choice of I, both quotients $S/(a)$ and $S/(b)$ must have finite length.

Consider the multiplication map by a:

$$S/(b) \twoheadrightarrow aS$$

Since $S/(b)$ has finite length, its surjective image aS also has finite length. We can now form the short exact sequence:

$$0 \to aS \to S \to S/aS \to 0$$

Since aS has finite length and S/aS (which is exactly $S/(a)$) has finite length, it follows that S must have finite length. This directly contradicts our assumption that R/I is not of finite length.

Thus, S must be an integral domain, meaning I is a prime ideal. By condition (2), all primes are maximal, so I must be maximal. But if I is maximal, $S = R/I$ is a field, which trivially has finite length (length 1). This is again a contradiction. Therefore, our initial assumption was false, and R must have finite length.

(3) $\implies$ (4): We have previously established that any module of finite length is necessarily both Artinian and Noetherian.

(4) $\implies$ (1): This is the most complex (or "hairy") part of the proof. Assume R is Artinian. We must show that (0) is a product of maximal ideals and that R is Noetherian.

First, let us choose an ideal J that is minimal among all ideals that can be written as a product of maximal ideals (this minimal element exists because R is Artinian). For any maximal ideal $\mathfrak{m}$, the product $J\mathfrak{m}$ is also a product of maximal ideals and $J\mathfrak{m} \subseteq J$. By the minimality of J, we must have:

$$J\mathfrak{m} = J \quad \text{for all maximal ideals } \mathfrak{m}$$

Since J itself is a product of maximal ideals, applying this repeatedly yields:

$$J^2 = J$$

We now aim to show that $J = 0$. Suppose for contradiction that $J \neq 0$. Because R is Artinian, we can choose an ideal I that is minimal among all ideals satisfying:

$$IJ \neq 0$$

Since $IJ \neq 0$, there must exist some element $i \in I$ such that $(i)J \neq 0$. The principal ideal (i) is contained in I, and by the minimality of I, we must have $I = (i)$.

Furthermore, observe that:

$$(IJ)J = I(J^2) = IJ \neq 0$$

Since $IJ \subseteq I$, the minimality of I forces the equality $IJ = I$. Thus, we have:

$$(i)J = (i)$$

This implies that there exists some element $j \in J$ such that $i = ij$, which rearranges to:

$$i(j-1) = 0$$

Now we consider the properties of j. The ideal J must be contained in all maximal ideals of R. If it were not contained in some maximal ideal $\mathfrak{m}$, then $J\mathfrak{m}$ would be a strictly smaller product of maximal ideals, contradicting the minimality of J. Since $j \in J$, it follows that j is in every maximal ideal. Consequently, $j-1$ cannot belong to any maximal ideal, which means $j-1$ must be a unit.

Since $i(j-1) = 0$ and $j-1$ is a unit, we are forced to conclude that $i = 0$. This contradicts the fact that $(i)J \neq 0$. The contradiction arose from assuming

$J \neq 0$. Therefore, we must have $J = 0$. Since J was defined as a product of maximal ideals, this proves that (0) is a product of maximal ideals.

Finally, we must show that R is Noetherian. We know that (0) is a product of maximal ideals, so we can write:

$$(0) = \mathfrak{m}_1\mathfrak{m}_2 \cdots \mathfrak{m}_n$$

Consider the descending filtration of ideals:

$$R \supset \mathfrak{m}_1 \supset \mathfrak{m}_1\mathfrak{m}_2 \supset \mathfrak{m}_1\mathfrak{m}_2\mathfrak{m}_3 \supset \cdots \supset \mathfrak{m}_1 \cdots \mathfrak{m}_n = (0)$$

Let us look at the successive quotients of this filtration. Each factor:

$$V_i = \frac{\mathfrak{m}_1 \cdots \mathfrak{m}_i}{\mathfrak{m}_1 \cdots \mathfrak{m}_{i+1}}$$

is annihilated by $\mathfrak{m}_{i+1}$, so it naturally forms a vector space over the field $R/\mathfrak{m}_{i+1}$. Because R is an Artinian ring, each of these factors V_i is an Artinian R-module, and hence an Artinian vector space. An Artinian vector space cannot have an infinite descending chain of subspaces, which means it must be finite-dimensional. A finite-dimensional vector space over a field inherently has finite length.

Since the ring R is built out of a finite number of extensions of modules of finite length, R itself must have finite length. As we established earlier, any module of finite length is Noetherian, completing the proof. □

Remark 25.2. It is tempting to assume that the factor V_i in the proof above is isomorphic to the field $R/\mathfrak{m}_{i+1}$. However, it is generally a vector space of dimension greater than one. For example, consider the ring:

$$R = \frac{k[x, y]}{(x^2, xy, y^2)}$$

This ring has a basis consisting of $1, x$, and y. The unique maximal ideal is $\mathfrak{m} = (x, y)$. We have the filtration $R \supset \mathfrak{m} \supset \mathfrak{m}^2 = (0)$. The first factor, $R/\mathfrak{m}$, is the field k (dimension 1). The second factor, $\mathfrak{m}/\mathfrak{m}^2$, has a basis consisting of x and y, making it a 2-dimensional vector space over k.

25.2. **Classification and Spectrum of Artinian Rings.** Having established this theorem, we can completely classify Artinian rings.

Corollary 25.3. *Any Artinian ring is isomorphic to a finite direct product of Artinian local rings.*

Proof. By the theorem, an Artinian ring has the property that the zero ideal is a product of maximal ideals. By grouping identical maximal ideals, we can write:

$$(0) = \mathfrak{m}_1^{k_1}\mathfrak{m}_2^{k_2} \cdots \mathfrak{m}_n^{k_n}$$

where the $\mathfrak{m}_i$ are distinct maximal ideals. Since distinct maximal ideals are coprime, their powers are also coprime. By the Chinese Remainder Theorem, we immediately obtain an isomorphism:

$$R \cong \frac{R}{\mathfrak{m}_1^{k_1}} \times \frac{R}{\mathfrak{m}_2^{k_2}} \times \cdots \times \frac{R}{\mathfrak{m}_n^{k_n}}$$

Each factor $R/\mathfrak{m}_i^{k_i}$ is an Artinian ring with exactly one prime ideal (which is necessarily maximal), making them Artinian local rings. □

This structural decomposition allows us to easily visualize the spectrum of an Artinian ring. Because every prime ideal is maximal, there are no inclusions among prime ideals, and the number of prime ideals is finite. Therefore, $\mathrm{Spec}(R)$ is simply a finite discrete set of points. This geometrically reinforces the intuition that Artinian rings are the ring-theoretic analogs of finite sets or finite-dimensional vector spaces.

However, one must be cautious, as the converse is not entirely true. A ring whose spectrum is a finite discrete set of points is not necessarily Artinian, unless we also assume it is Noetherian. For instance, consider the polynomial ring in infinitely many variables modulo the squares of all variables:

$$R = \frac{k[x_1, x_2, \dots]}{(x_1^2, x_2^2, \dots)}$$

This ring possesses a unique maximal ideal $\mathfrak{m} = (x_1, x_2, \dots)$. Thus, its spectrum consists of a single point. Yet R is neither Noetherian nor Artinian, demonstrating that the topological structure of the spectrum does not capture all the algebraic properties of the ring.

26. Examples of Artinian Rings

In the previous lecture, we established the rather heavy theorem that Artinian rings are necessarily Noetherian. We also demonstrated, as a consequence of the proof, that any Artinian ring is isomorphic to a product of local Artinian rings, which makes them particularly simple to study.

A typical motivating example of this decomposition is the Artinian ring $\mathbb{Z}/60\mathbb{Z}$. By the Chinese Remainder Theorem, this splits into a product of local Artinian rings:

$$\frac{\mathbb{Z}}{60\mathbb{Z}} \cong \frac{\mathbb{Z}}{4\mathbb{Z}} \times \frac{\mathbb{Z}}{3\mathbb{Z}} \times \frac{\mathbb{Z}}{5\mathbb{Z}}$$

Here we have an Artinian ring written as a product of three local Artinian rings. This is fairly typical of what happens in general, except it can become significantly more complicated. In this section, we will look at various examples of Artinian rings and attempt to classify them in terms of their *length*.

Recall that every Artinian ring has finite length. For simplicity, we will restrict our attention to local Artinian rings. If R is a local Artinian ring, it has a single maximal ideal $\mathfrak{m}$, and the quotient $R/\mathfrak{m}$ is some field k. If R has length n over k, it is in some sense "about the same size" as an n-dimensional vector space over k, although it need not actually be a vector space over k.

26.1. **Classification by Length.**

26.1.1. *Lengths 0 and 1.* The lowest lengths are trivial to classify:

- **Length 0:** This is trivially the zero ring, where $1 = 0$.
- **Length 1:** This yields exactly one example, which is the field k itself.

26.1.2. *Length 2.* At length two, the structure begins to get more interesting. One way to construct length two Artinian rings is to take a polynomial ring $k[x]$ and quotient it out by some polynomial $f(x)$ of degree two:

$$R = \frac{k[x]}{(x^2 + ax + b)}$$

This yields several slightly different cases depending on the factorization of $f(x)$:

(1) **Non-reduced ring:** We might get a ring with nilpotent elements, such as:

$$R = \frac{k[x]}{(x^2)}$$

Here, the element x is nilpotent (specifically, $x^2 = 0$).

(2) **Product of fields:** The polynomial might split into distinct linear factors, $f(x) = (x - \alpha)(x - \beta)$. By the Chinese Remainder Theorem, this splits:

$$\frac{k[x]}{(x - \alpha)(x - \beta)} \cong k \times k$$

Here we have an Artinian ring that is the product of two local Artinian rings.

(3) **Field extension:** The polynomial $f(x)$ might be irreducible. In this case, the quotient is simply a larger field containing k. A typical example is the Gaussian rationals:

$$\mathbb{Q}(i) \cong \frac{\mathbb{Q}[x]}{(x^2 + 1)}$$

These are all examples of rings that are vector spaces over k. However, we can also find examples that are *not* vector spaces over k. A simple example is the ring:

$$R = \frac{\mathbb{Z}}{p^2\mathbb{Z}}$$

This ring has length two, and its quotient field is $k = \mathbb{Z}/p\mathbb{Z}$. It is about the same size as a two-dimensional vector space over a field of p elements, but it is not a vector space because it is not annihilated by p.

26.1.3. *Lengths 3 and 4.* For higher lengths, rings that are not vector spaces become tricky to classify. In fact, even the ones that are vector spaces become quite complicated.

For **length 3**, we can get structures analogous to the length two cases:

- A product of three fields: $k \times k \times k$.
- A product of a non-reduced ring and a field: $k[x]/(x^2) \times k$.
- A field extension of degree 3, or a product of a field of degree 2 and a field of degree 1.

- Something slightly new: we can take a polynomial ring in two variables and quotient out by the ideal generated by all quadratic monomials:

$$R = \frac{k[x,y]}{(x^2, xy, y^2)}$$

This is a three-dimensional vector space with basis $1, x$, and y, where any product of x and y is strictly zero.

For **length 4**, things become even more intricate. Suppose R is an Artinian local ring with maximal ideal $\mathfrak{m}$, such that $\mathfrak{m}^3 = 0$. We have a natural mapping:

$$\frac{\mathfrak{m}}{\mathfrak{m}^2} \to \frac{\mathfrak{m}^2}{\mathfrak{m}^3}$$

If we take the dimension of $\mathfrak{m}^2/\mathfrak{m}^3$ to be 1 (so we can identify it with k), and the dimension of $\mathfrak{m}/\mathfrak{m}^2$ to be 2, then the map taking $x \mapsto x^2$ encodes a quadratic form on a two-dimensional vector space. Over various fields k (such as $\mathbb{Q}$), there are a large number of inequivalent quadratic forms in two variables, meaning there are many distinct Artinian rings of this specific type.

26.2. **The Complexity of High-Length Rings.** For lengths greater than or equal to 5 or 6, the classification starts becoming overwhelmingly complicated. In fact, it becomes so wild that you cannot reasonably describe a classification at all.

This is a universal phenomenon whenever you attempt to classify nilpotent objects. In our Artinian local ring, the maximal ideal $\mathfrak{m}$ is nilpotent, meaning $\mathfrak{m}^n = 0$ for some finite n. Nilpotent objects include Artinian rings, nilpotent Lie algebras, and finite p-groups.

If nilpotent objects are generated by only one or two elements, you can usually classify them. But as soon as they are generated by more elements and reach a sufficiently high dimension, the classification goes completely out of control. For example, the number of finite p-groups of order 2^{10} is exactly $49,487,365,422$. The number of nilpotent objects grows so rapidly with the dimension that it is hopeless to classify them beyond a very low bound.

26.3. **Hilbert Schemes and the Space of Ideals.** We can formalize this "wildness" by asking: what is the dimension of the space of Artinian rings over a field k that have codimension n and are generated by m elements?

We are essentially looking at ideals I in the polynomial ring $k[x_1, \ldots, x_m]$ such that the quotient has dimension n over k:

$$R = \frac{k[x_1, \ldots, x_m]}{I}$$

We can think of these ideals as forming a parameter space. What is the dimension of this space?

A naive geometric approach is to take n distinct points in affine space k^m, and let I be the ideal of all polynomials vanishing on these n points. Since we have n points, each with m coordinates, this configuration gives us an mn-dimensional space of Artinian rings. One might naturally guess that an arbitrary Artinian

ring is a limit of n points moving around in m-dimensional space, suggesting that the dimension of the space of such ideals is mn.

This plausible argument works perfectly for $m = 1$ and $m = 2$:

- For $m = 1$, an ideal is generated by a single polynomial of degree n, which obviously has dimension $n \times 1$.
- For $m = 2$, the dimension is indeed $2n$.

However, for $m \geq 3$, this intuition fails completely. There are ridiculously larger numbers of ideals of codimension n than geometry would suggest.

To see this, consider $m = 3$ with the ring $k[x_1, x_2, x_3]$. Let $\mathfrak{m} = (x_1, x_2, x_3)$ be the maximal ideal at the origin. We look for ideals I nested between powers of $\mathfrak{m}$:

$$\mathfrak{m}^{i+1} \subseteq I \subseteq \mathfrak{m}^i$$

for some integer $i > 0$. Notice that *any* vector subspace lying between $\mathfrak{m}^i$ and $\mathfrak{m}^{i+1}$ is automatically an ideal, because multiplying any element of $\mathfrak{m}^i$ by variables pushes it into $\mathfrak{m}^{i+1}$, which is already contained in I.

Let us estimate the dimensions:

- The codimension n of such an ideal is roughly bounded by a constant times i^3.
- The dimension of the vector space quotient $\mathfrak{m}^i/\mathfrak{m}^{i+1}$ is roughly a constant times i^2.

The space of all subspaces of a vector space is called a *Grassmannian*. The dimension of the Grassmannian of subspaces of an N-dimensional space is roughly proportional to N^2. Here, $N \sim i^2$, so the dimension of the space of such ideals is on the order of:

$$\dim(\text{Space of Ideals}) \sim (i^2)^2 = i^4$$

Comparing this to our geometric guess mn: since $m = 3$ and $n \sim i^3$, the geometric dimension is order i^3. But the actual space of ideals has dimension order i^4. For large i, $i^4 \gg i^3$.

This demonstrates that Artinian rings with at least three generators are deeply pathological. The space parameterizing them, called a *Hilbert scheme* (constructed by Grothendieck), becomes unmanageably wild. The general rule of thumb is that anything that can possibly go wrong *will* go wrong for a Hilbert scheme of dimension 3 or more.

26.4. **Tensor Products of Fields.** Another context where Artinian rings frequently arise is in the tensor product of two fields.

Example 26.1. Consider the tensor product of the complex numbers with themselves over the real numbers. This is a 4-dimensional algebra over $\mathbb{R}$, making it an Artinian ring. It splits as a product of local Artinian rings:

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C} \times \mathbb{C}$$

(Do not confuse the tensor product on the left with the ordinary product on the right; it is a numerical coincidence that $2 \times 2 = 2 + 2$). Because it splits as a

product, it must be generated by *idempotents*. We can write down these two idempotents explicitly:

$$e_1 = \frac{1 \otimes 1 + i \otimes i}{2}$$
$$e_2 = \frac{1 \otimes 1 - i \otimes i}{2}$$

One can easily check that $e_1^2 = e_1$ and $e_2^2 = e_2$, giving the required decomposition.

Example 26.2. If we tensor $\mathbb{Q}(\sqrt{2})$ with $\mathbb{Q}(\sqrt{3})$ over $\mathbb{Q}$, the result is actually a single field:

$$\mathbb{Q}(\sqrt{2}) \otimes_{\mathbb{Q}} \mathbb{Q}(\sqrt{3}) \cong \mathbb{Q}(\sqrt{2}, \sqrt{3})$$

When does a tensor product of fields yield a field, and when does it split into a product? If we take $K \otimes_k L$, where L is a finite separable extension of k, we can write $L \cong k[x]/(f(x))$ for some irreducible, separable polynomial f. Then:

$$K \otimes_k L \cong \frac{K[x]}{(f(x))}$$

While $f(x)$ is irreducible over k, it might factor over K as $f(x) = f_1(x) \cdots f_r(x)$. Because L is separable, f has no multiple roots, meaning these factors $f_i(x)$ are strictly coprime. By the Chinese Remainder Theorem:

$$\frac{K[x]}{(f(x))} \cong \prod_{i=1}^{r} \frac{K[x]}{(f_i(x))}$$

Thus, the tensor product decomposes into a product of fields.

26.4.1. *Inseparable Extensions.* The behavior is much stranger for inseparable extensions. Suppose k has characteristic $p > 0$, and we choose an element $a \in k$ that is not a p-th power. The polynomial $x^p - a$ is irreducible. Let $L = k[x]/(x^p - a)$ and let K be a field extension containing a root b such that $b^p = a$. Then:

$$K \otimes_k L \cong \frac{K[x]}{(x^p - a)}$$

Over the field K, we can factor $x^p - a$ as $(x - b)^p$. Thus:

$$K \otimes_k L \cong \frac{K[x]}{(x - b)^p}$$

This ring is clearly not a field, nor is it a product of fields. It is a local Artinian ring containing nilpotent elements (such as $x - b$).

Finally, we remark that the tensor product of two Artinian rings need not be Artinian at all. If we take the field of rational functions $k(x)$ and tensor it with $k(y)$ over k:

$$k(x) \otimes_k k(y)$$

The resulting ring is infinite-dimensional over k and is not an Artinian ring.

27. Associated Primes

This lecture introduces the concept of associated primes of a module M. The set of associated primes, denoted $\mathrm{Ass}(M)$, is a collection of prime ideals that geometrically control where the module "lives," a concept we will investigate geometrically in the subsequent lecture.

27.1. **Motivation: Finite Length vs. Finitely Generated.** To motivate this, let us first consider modules M of finite length. If a module has finite length, we can construct a composition series:

$$0 = M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \cdots \subsetneq M_n = M$$

where each successive quotient M_{i+1}/M_i is a simple module. A simple module contains no submodules other than 0 and itself, and is therefore isomorphic to the base ring modulo a maximal ideal:

$$\frac{M_{i+1}}{M_i} \cong \frac{R}{\mathfrak{m}_i}$$

This structural decomposition is perfect for modules of finite length. The trouble is that most modules we encounter in commutative algebra do not have finite length; rather, they are merely finitely generated.

Can we find an analogue for finitely generated modules? Over arbitrary rings, the answer is generally no. However, if we assume R is a *Noetherian ring*, a finitely generated module M over R is a Noetherian module. This finiteness condition allows us to recover a similar structural decomposition, replacing maximal ideals with prime ideals.

27.2. **Filtrations by Prime Quotients.** Let us examine some examples to see how we might build finitely generated modules.

Example 27.1. Let $R = \mathbb{Z}$. The finitely generated modules are precisely the finitely generated abelian groups. Such a group can be decomposed into a direct sum:

$$M = \bigoplus \mathbb{Z} \oplus \bigoplus_{p} \bigoplus_{n_i} \frac{\mathbb{Z}}{p^{n_i}\mathbb{Z}}$$

We can build M out of basic blocks. The simple modules are $\mathbb{Z}/p\mathbb{Z}$, corresponding to maximal ideals (p). However, we cannot build the free part $\mathbb{Z}$ out of simple modules. Instead, we must add $\mathbb{Z}$ itself to our list of building blocks, which we can write as $\mathbb{Z}/(0)$.

Notice that while (p) is maximal, (0) is merely a prime ideal in $\mathbb{Z}$. This suggests that we should look for prime ideals in the denominator rather than maximal ones.

This leads to a central question: For a finitely generated module M over a Noetherian ring R, can we find a sequence of submodules

$$0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_n = M$$

such that each successive quotient is isomorphic to $R/\mathfrak{p}_i$ for some prime ideal $\mathfrak{p}_i$?

$$\frac{M_{i+1}}{M_i} \cong \frac{R}{\mathfrak{p}_i}$$

The answer is yes.

Theorem 27.2. *Let M be a non-zero finitely generated module over a Noetherian ring R. Then M contains a submodule isomorphic to $R/\mathfrak{p}$ for some prime ideal $\mathfrak{p}$.*

Proof. We look for an ideal $\mathfrak{p}$ that is maximal among the set of annihilators of non-zero elements of M. Let this set of ideals be:

$$\Sigma = \{\mathrm{Ann}(x) \mid x \in M, x \neq 0\}$$

Since $M \neq 0$, Σ is non-empty. Because R is Noetherian, Σ has a maximal element, which we will call $\mathfrak{p}$. This means $\mathfrak{p} = \mathrm{Ann}(a)$ for some $a \in M \setminus \{0\}$, so $R/\mathfrak{p} \cong Ra \subseteq M$.

We must show that $\mathfrak{p}$ is a prime ideal. As is typical in commutative algebra, maximal elements of collections of annihilators have a strong tendency to be prime. Suppose $\mathfrak{p}$ is not prime. Then we can find elements $x, y \in R$ such that $x \notin \mathfrak{p}$ and $y \notin \mathfrak{p}$, but their product $xy \in \mathfrak{p}$.

Let $\bar{x}$ be the image of x under the natural mapping to the submodule $R/\mathfrak{p} \subseteq M$. Consider the annihilator of this element, $\mathrm{Ann}(\bar{x})$.

- It contains the ideal $\mathfrak{p}$ inherently.
- It contains y, because $xy \in \mathfrak{p} \implies y\bar{x} = 0$.

Because $y \notin \mathfrak{p}$, the ideal $\mathrm{Ann}(\bar{x})$ is strictly larger than $\mathfrak{p}$. Furthermore, because $x \notin \mathfrak{p}$, $\bar{x} \neq 0$, meaning $\mathrm{Ann}(\bar{x})$ is the annihilator of a non-zero element in M. This strictly larger annihilator contradicts the maximality of $\mathfrak{p}$ in Σ.

Thus, our assumption that $\mathfrak{p}$ is not prime must be false, so $\mathfrak{p}$ is indeed a prime ideal. □

Using this theorem, we can iteratively construct the desired filtration.

(1) If $M = 0$, we stop. Otherwise, M contains a submodule $M_1 \cong R/\mathfrak{p}_1$ with $\mathfrak{p}_1$ prime.
(2) Look at the quotient module M/M_1. By the theorem, it contains a submodule isomorphic to $R/\mathfrak{p}_2$.
(3) We define M_2 as the inverse image of this submodule under the projection $M \twoheadrightarrow M/M_1$. Then $M_2/M_1 \cong R/\mathfrak{p}_2$.
(4) Continue to find M_3, M_4, and so on.

This yields an increasing sequence of submodules $M_1 \subseteq M_2 \subseteq M_3 \subseteq \dots$. Because M is a Noetherian module, this chain must eventually terminate at some $M_n = M$. Thus, the module can be fully decomposed into a finite sequence of prime quotients.

27.3. **The Problem of Multiplicity.** Having found such a decomposition, we naturally ask: How often does a specific module $R/\mathfrak{p}$ occur in M? We might call this the *multiplicity* of $R/\mathfrak{p}$ in M.

For finite length modules, we previously saw that the multiplicity is well-defined (the Jordan-Hölder theorem) and additive on exact sequences. If $0 \to A \to B \to C \to 0$ is exact, then:

$$\text{mult}_{R/\mathfrak{m}}(B) = \text{mult}_{R/\mathfrak{m}}(A) + \text{mult}_{R/\mathfrak{m}}(C)$$

Does this additivity hold for finitely generated modules? Let us investigate.

Example 27.3. Let $R = \mathbb{Z}$ and $M = \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$. It seems obvious that the multiplicity of $\mathbb{Z}$ in M is 1, and the multiplicity of $\mathbb{Z}/2\mathbb{Z}$ is 1.

However, consider the universal counterexample:

$$0 \to \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

If multiplicity were additive, the multiplicity of $\mathbb{Z}/2\mathbb{Z}$ in the middle term ($\mathbb{Z}$) should equal its multiplicity in the left term plus its multiplicity in the right term. But $\mathbb{Z}/2\mathbb{Z}$ clearly does not occur as a submodule in $\mathbb{Z}$, yielding an equation $0 = 0 + 1$, which is false.

The root of the problem is that $\mathbb{Z}/2\mathbb{Z}$ *does* conceptually occur in $\mathbb{Z}$ as a quotient. In fact, we can pull out an infinite number of such quotients. There is no finite limit to how often $\mathbb{Z}/2\mathbb{Z}$ occurs "inside" $\mathbb{Z}$.

We can reliably define multiplicity when the prime ideal is (0). For $R = \mathbb{Z}$, the multiplicity of $\mathbb{Z}/(0)$ in M is given by the dimension of the rational vector space $M \otimes_{\mathbb{Z}} \mathbb{Q}$:

$$\dim_{\mathbb{Q}}(M \otimes_{\mathbb{Z}} \mathbb{Q})$$

This is just the rank of the abelian group, and it is perfectly additive because tensoring with $\mathbb{Q}$ (a flat module) preserves exactness. We cannot define the multiplicity of $\mathbb{Z}/2\mathbb{Z}$ analogously by tensoring with $\mathbb{Z}/2\mathbb{Z}$, precisely because $\mathbb{Z}/2\mathbb{Z}$ is not flat and does not preserve exactness.

27.4. **A Disturbing Geometric Example.** To show that the situation is even more dire than losing additivity, consider a case where the set of primes that occur is not even well-defined across different filtrations.

Example 27.4. Let $R = k[x, y]$ and let M be the ideal generated by x and y. We can view this as a submodule of R.

We can construct two completely different filtrations of M:

(1) We take a free submodule generated by x, giving an exact sequence:

$$0 \to R \to M \to \frac{R}{(y)} \to 0$$

Here, $M_1 \cong R$, and the quotient M/M_1 is isomorphic to $R/(y)$. The sequence of primes used is (0) and (y).

(2) We take a free submodule generated by y, giving another sequence:

$$0 \to R \to M \to \frac{R}{(x)} \to 0$$

Here, the quotient is isomorphic to $R/(x)$. The sequence of primes is (0) and (x).

The prime (y) occurs in the first filtration, but not the second. The prime (x) occurs in the second, but not the first. Neither of these quotients definitively "belongs" to M. We cannot build M exclusively out of copies of R (i.e., $0 \to R \to M \to R \to 0$ is impossible).

27.5. **Definition of Associated Primes.** To resolve the ambiguity of which primes actually belong to the module intrinsically, we restrict our focus to the submodules that literally sit inside M, rather than sub-quotients.

Definition 27.5. Let M be an R-module. The set of *associated primes* of M, denoted $\mathrm{Ass}(M)$, is the set of prime ideals $\mathfrak{p} \subset R$ that are annihilators of elements of M. Equivalently:

$$\begin{aligned}\mathrm{Ass}(M) &= \{\mathfrak{p} \in \mathrm{Spec}(R) \mid \mathfrak{p} = \mathrm{Ann}(m) \text{ for some } m \in M\} \\ &= \{\mathfrak{p} \in \mathrm{Spec}(R) \mid R/\mathfrak{p} \text{ is isomorphic to a submodule of } M\}\end{aligned}$$

Informally, $\mathrm{Ass}(M)$ is the set of prime ideals $\mathfrak{p}$ such that $R/\mathfrak{p}$ definitively occurs inside M as a direct sub-object. We proved earlier that if M is a non-zero finitely generated module over a Noetherian ring, it must contain at least one such submodule. Thus:

$$M \neq 0 \implies \mathrm{Ass}(M) \neq \emptyset$$

Let us revisit our previous examples using this rigorous definition.

Example 27.6.

- $\mathrm{Ass}(\mathbb{Z})$: The only prime ideal that is an annihilator of a non-zero integer is the zero ideal. $\mathrm{Ass}(\mathbb{Z}) = \{(0)\}$.
- $\mathrm{Ass}(\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z})$: The element $(1, 0)$ has annihilator (0). The element $(0, 1)$ has annihilator (2). Thus, $\mathrm{Ass}(\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}) = \{(0), (2)\}$.
- $\mathrm{Ass}((x, y))$ as an ideal in $k[x, y]$: The ring is an integral domain, so the only annihilator of any non-zero element is (0). Thus, $\mathrm{Ass}((x, y)) = \{(0)\}$. The "ghost" quotients $R/(x)$ and $R/(y)$ from our filtrations are correctly excluded.

27.6. **Properties of Associated Primes.** We now examine how the set of associated primes behaves with respect to short exact sequences.

Proposition 27.7. *Suppose we have a short exact sequence of R-modules:*

$$0 \to A \to B \to C \to 0$$

Then the associated primes satisfy:

$$\mathrm{Ass}(B) \subseteq \mathrm{Ass}(A) \cup \mathrm{Ass}(C)$$

Furthermore, we have a trivial lower bound:

$$\mathrm{Ass}(A) \subseteq \mathrm{Ass}(B)$$

Proof. The inclusion $\mathrm{Ass}(A) \subseteq \mathrm{Ass}(B)$ is completely obvious: if $R/\mathfrak{p}$ is a submodule of A, and A is a submodule of B, then $R/\mathfrak{p}$ is a submodule of B.

To prove $\mathrm{Ass}(B) \subseteq \mathrm{Ass}(A) \cup \mathrm{Ass}(C)$, let $\mathfrak{p} \in \mathrm{Ass}(B)$. This means there is a submodule $X \subseteq B$ such that $X \cong R/\mathfrak{p}$. We must consider the intersection of X with the submodule A:

- **Case 1:** $X \cap A = 0$**.** In this case, the projection map from B to C maps X injectively into C. Therefore, X is isomorphic to a submodule of C, which implies $\mathfrak{p} \in \mathrm{Ass}(C)$.
- **Case 2:** $X \cap A \neq 0$**.** Let a be a non-zero element in $X \cap A$. Because $a \in X \cong R/\mathfrak{p}$, and $R/\mathfrak{p}$ is an integral domain, the annihilator of *any* non-zero element in $R/\mathfrak{p}$ is precisely $\mathfrak{p}$. Thus, $\mathrm{Ann}(a) = \mathfrak{p}$. Since a is also an element of A, this proves that $\mathfrak{p} \in \mathrm{Ass}(A)$.

In both cases, $\mathfrak{p}$ must belong to either $\mathrm{Ass}(A)$ or $\mathrm{Ass}(C)$, establishing the subset inclusion. □

Remark 27.8. It is not generally true that $\mathrm{Ass}(B) = \mathrm{Ass}(A) \cup \mathrm{Ass}(C)$. Referring back to the exact sequence $0 \to \mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$, we have:

$$\mathrm{Ass}(\mathbb{Z}) = \{(0)\}$$
$$\mathrm{Ass}(\mathbb{Z}) \cup \mathrm{Ass}(\mathbb{Z}/2\mathbb{Z}) = \{(0)\} \cup \{(2)\} = \{(0), (2)\}$$

Clearly, $\{(0)\} \neq \{(0), (2)\}$.

We conclude with two structural consequences regarding the finiteness of $\mathrm{Ass}(M)$ for Noetherian modules.

Corollary 27.9. *Let M be a finitely generated module over a Noetherian ring R.*

(1) The set $\mathrm{Ass}(M)$ is finite.

(2) If we decompose M into a finite filtration $0 = M_0 \subsetneq M_1 \cdots \subsetneq M_n = M$ such that $M_{i+1}/M_i \cong R/\mathfrak{p}_i$ for prime ideals $\mathfrak{p}_i$, then:

$$\mathrm{Ass}(M) \subseteq \{\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_n\}$$

Proof. The proof relies immediately on induction via the exact sequence formula. By splitting the filtration into a sequence of short exact sequences $0 \to M_i \to M_{i+1} \to R/\mathfrak{p}_i \to 0$, we repeatedly apply the union bound:

$$\mathrm{Ass}(M_{i+1}) \subseteq \mathrm{Ass}(M_i) \cup \mathrm{Ass}(R/\mathfrak{p}_i)$$

Since $\mathrm{Ass}(R/\mathfrak{p}_i) = \{\mathfrak{p}_i\}$, iterating this from $M_0 = 0$ up to $M_n = M$ demonstrates that any prime in $\mathrm{Ass}(M)$ must be one of the $\mathfrak{p}_i$ appearing in the filtration. Because the filtration is finite, the set of associated primes is finite. □

28. Geometry of Associated Primes

In the previous lecture, we introduced the set of associated primes of a module. We now wish to investigate the geometric interpretation of this set and relate it to the support of a module. We will see that while the associated primes provide a refined picture of where a module "lives" geometrically, they also introduce complications, particularly in the form of embedded primes.

28.1. **Review and Limitations of Prime Filtrations.** Let R be a Noetherian commutative ring and M be a finitely generated R-module. Recall that the set of associated primes of M, denoted $\mathrm{Ass}(M)$, is defined as:

$$\begin{aligned}\mathrm{Ass}(M) &= \{\mathfrak{p} \in \mathrm{Spec}(R) \mid \mathfrak{p} = \mathrm{Ann}(m) \text{ for some } m \in M \setminus \{0\}\} \\ &= \{\mathfrak{p} \in \mathrm{Spec}(R) \mid R/\mathfrak{p} \hookrightarrow M\}\end{aligned}$$

We established that this set is finite. Furthermore, we can decompose M into a finite filtration:

$$0 = M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \cdots \subsetneq M_n = M$$

where each successive quotient M_{i+1}/M_i is isomorphic to $R/\mathfrak{p}_i$ for some $\mathfrak{p}_i \in \operatorname{Spec}(R)$.

We also showed that if $\mathfrak{p} \in \operatorname{Ass}(M)$, then $R/\mathfrak{p}$ must occur as a quotient in *every* such filtration of M. However, the converse question—whether every module can be built up purely by extensions of $R/\mathfrak{p}$ for $\mathfrak{p} \in \operatorname{Ass}(M)$—has a negative answer.

Example 28.1. Let $R = \mathbb{Z}[\sqrt{-5}]$, which is a standard example of a Dedekind domain that is not a Unique Factorization Domain. It contains non-principal ideals. Let M be the non-principal ideal:

$$M = (2, 1 + \sqrt{-5})$$

Because M is a torsion-free module over an integral domain, its only associated prime is the zero ideal:

$$\operatorname{Ass}(M) = \{(0)\}$$

However, we cannot build M simply by taking extensions of $R/(0) \cong R$, precisely because M is not a principal ideal and does not have a free structure. Thus, the exact relationship between the associated primes and the structural building blocks of M is somewhat messier than in the case of modules of finite length.

28.2. **The Support of a Module.** To better understand the geometry of M, we relate the associated primes to the spectrum of R. Since $\operatorname{Ass}(M) \subset \operatorname{Spec}(R)$, we can visualize $\operatorname{Ass}(M)$ as a specific subset of points. We introduce another closely related geometric invariant called the support of a module.

Definition 28.2. The *support* of an R-module M, denoted $\operatorname{Supp}(M)$, is the set of prime ideals $\mathfrak{p}$ such that the localization $M_{\mathfrak{p}}$ is non-zero:

$$\operatorname{Supp}(M) = \{\mathfrak{p} \in \operatorname{Spec}(R) \mid M_{\mathfrak{p}} \neq 0\}$$

Geometrically, you can think of $M_{\mathfrak{p}}$ as the *stalk* of the module M at the point $\mathfrak{p} \in \operatorname{Spec}(R)$. The support is simply the set of points where this stalk does not vanish.

For a finitely generated module M, the support is easy to determine algebraically. It coincides exactly with the closed subset defined by the annihilator of M:

$$\operatorname{Supp}(M) = \{\mathfrak{p} \in \operatorname{Spec}(R) \mid \operatorname{Ann}(M) \subseteq \mathfrak{p}\} = Z(\operatorname{Ann}(M))$$

Thus, $\operatorname{Supp}(M)$ is a closed subset in the Zariski topology.

28.3. **Comparing Support and Associated Primes.** Let us look at a few examples to contrast the support and the associated primes.

Example 28.3. Case 1: Modules of Finite Length. Let $R = \mathbb{Z}$ and $M = \mathbb{Z}/12\mathbb{Z}$.

$$\operatorname{Supp}(M) = \{(2), (3)\}$$
$$\operatorname{Ass}(M) = \{(2), (3)\}$$

For modules of finite length, the support and the associated primes completely coincide.

Case 2: The Base Ring. Let $R = \mathbb{Z}$ and $M = \mathbb{Z}$.

$$\operatorname{Supp}(M) = \operatorname{Spec}(\mathbb{Z})$$
$$\operatorname{Ass}(M) = \{(0)\}$$

Here, the support is the entire spectrum, including all maximal ideals (p). However, the only associated prime is the generic point (0).

Case 3: A Mixed Module. Let $R = \mathbb{Z}$ and $M = \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$.

$$\operatorname{Supp}(M) = \operatorname{Spec}(\mathbb{Z})$$
$$\operatorname{Ass}(M) = \{(0), (2)\}$$

The support is again the entire spectrum because the free part $\mathbb{Z}$ does not vanish anywhere. However, the associated primes detect both the free part (via (0)) and the torsion part (via (2)).

The support is a very crude invariant. Once a module contains a free copy of R, its support becomes the entire spectrum, and the support cannot detect any additional torsion components added to the module. The set of associated primes, however, is a more refined invariant that accurately records these extra components.

Despite their differences, the two sets are intimately related topologically.

Theorem 28.4. *For a finitely generated module M over a Noetherian ring R, the support of M is the Zariski closure of the set of associated primes of M:*

$$\operatorname{Supp}(M) = \overline{\operatorname{Ass}(M)}$$

Proof. First, we show $\operatorname{Ass}(M) \subseteq \operatorname{Supp}(M)$. Let $\mathfrak{p} \in \operatorname{Ass}(M)$. By definition, $\mathfrak{p} = \operatorname{Ann}(a)$ for some $a \in M$. This implies:

$$\operatorname{Ann}(M) \subseteq \operatorname{Ann}(a) = \mathfrak{p}$$

Because $\mathfrak{p}$ contains the annihilator of the entire module M, $\mathfrak{p} \in \operatorname{Supp}(M)$. Since the support is a closed set, it must contain the closure of any of its subsets. Thus:

$$\overline{\operatorname{Ass}(M)} \subseteq \operatorname{Supp}(M)$$

Conversely, suppose $\mathfrak{p} \in \operatorname{Supp}(M)$. By definition, the localized module $M_\mathfrak{p}$ is non-zero. Because $M_\mathfrak{p}$ is a non-zero finitely generated module over the Noetherian local ring $R_\mathfrak{p}$, it must have at least one associated prime. Choose some prime $\mathfrak{q} \in \operatorname{Ass}_{R_\mathfrak{p}}(M_\mathfrak{p})$.

The prime $\mathfrak{q}$ is an element of $\operatorname{Spec}(R_\mathfrak{p})$. The natural localization map $f\colon R \to R_\mathfrak{p}$ induces a map of spectra that pulls $\mathfrak{q}$ back to a prime ideal $f^{-1}(\mathfrak{q})$ in R. One

can verify that this pullback is an associated prime of the original module M, and by the properties of localization, it is strictly contained in $\mathfrak{p}$.

Therefore, there exists an associated prime of M contained in $\mathfrak{p}$. Topologically, this means $\mathfrak{p}$ is in the closure of that associated prime. Hence:

$$\mathrm{Supp}(M) \subseteq \overline{\mathrm{Ass}(M)}$$

This completes the proof. □

28.4. **Minimal and Embedded Primes.** Let us visualize these concepts using the polynomial ring $R = k[x, y]$. The spectrum $\mathrm{Spec}(k[x, y])$ is the affine plane.

Example 28.5. Let $M = k[x, y]/(y)$. The support of M is the x-axis, defined by $y = 0$. The associated prime is simply the ideal (y). Thus, the module "lives" exactly on the x-axis.

Example 28.6. Now, let us slightly modify the module. Let $M = k[x, y]/(y^2, xy)$.

The support of this module is again just the x-axis. The locus of points where both $y^2 = 0$ and $xy = 0$ is exactly the line $y = 0$. However, the associated primes tell a different story. If we write out a k-basis for M, we find it consists of:

$$1, x, x^2, x^3, \ldots \quad \text{and} \quad y$$

We have two distinct types of annihilators:

- The element x is annihilated by (y).
- The element y is annihilated by (x, y).

Thus, the associated primes are:

$$\mathrm{Ass}(M) = \{(y), (x, y)\}$$

Notice that we have a strict inclusion: $(y) \subsetneq (x, y)$.

Definition 28.7. Let M be a module. An associated prime $\mathfrak{p} \in \mathrm{Ass}(M)$ is called a *minimal* (or *isolated*) prime if it is minimal with respect to inclusion in $\mathrm{Ass}(M)$.

An associated prime that is not minimal (i.e., it strictly contains another associated prime) is called an *embedded prime*.

In Example 28.6, (y) is a minimal prime, and (x, y) is an embedded prime. The terminology "embedded" is geometrically intuitive: the maximal ideal (x, y) corresponds to the origin $(0, 0)$, which is literally a point *embedded* inside the closure of the minimal prime (y) (the x-axis).

Geometrically, one can visualize the module M as a function distribution that lives mostly on the x-axis, but has an extra "thick" point or extra structure specifically concentrated at the origin. Almost all pathologies and technical difficulties encountered when working with general modules are caused by the presence of these embedded primes.

28.5. **Towards Primary Decomposition.** We have seen that a module M can be visualized as living on its support, with extra structural information located at its embedded primes. However, $\mathrm{Ass}(M)$ is still a somewhat crude invariant. If we take $M \oplus M$, the associated primes remain exactly the same, yet the module has doubled in size. We desire a more refined structure theorem.

For a finitely generated abelian group A (a $\mathbb{Z}$-module), the fundamental theorem allows us to write it uniquely as a direct sum of its primary components and a free part:

$$A = A_0 \oplus A_2 \oplus A_3 \oplus \dots$$

where $\text{Ass}(A_0) = \{(0)\}$, $\text{Ass}(A_2) = \{(2)\}$, $\text{Ass}(A_3) = \{(3)\}$, and so on. We have completely decomposed the group into submodules, each possessing exactly one associated prime.

Can we achieve a similar direct sum decomposition for a finitely generated module M over an arbitrary Noetherian ring? That is, can we write $M = \bigoplus M_i$ where each M_i has exactly one associated prime?

The answer is generally no, and embedded primes are the culprits. Consider again $M = k[x, y]/(y^2, xy)$. The associated primes are (y) and (x, y). We can map M into a direct sum of modules corresponding to these primes:

$$M \hookrightarrow \frac{k[x, y]}{(y)} \oplus \frac{k[x, y]}{(x, y)}$$

While this mapping exists and is injective (meaning M is a *submodule* of a direct sum of pieces with unique associated primes), M itself does not split into a direct sum. This inclusion represents the best possible structural generalization: the *co-primary decomposition* of a module.

28.6. **Historical Note: The Lasker-Noether Theorem.** Historically, this decomposition theory was not developed for modules, but for ideals. The foundational result is the Lasker-Noether theorem.

Theorem 28.8 (Lasker-Noether)**.** *If I is an ideal of a Noetherian ring R, then I can be written as a finite intersection of primary ideals:*

$$I = Q_1 \cap Q_2 \cap \dots \cap Q_k$$

Definition 28.9. An ideal $J \subseteq R$ is called *primary* if it satisfies a slightly relaxed version of the prime ideal condition: if $xy \in J$, then either $x \in J$ or $y^n \in J$ for some integer $n \geq 1$.

Equivalently, J is primary if every zero divisor in the quotient ring R/J is nilpotent.

For example, in $\mathbb{Z}$, the prime ideals are (0) and (p). The primary ideals are (0) and (p^n). While primary ideals look suspiciously like powers of prime ideals, in more complicated rings (especially those with embedded primes), primary ideals are not necessarily just prime powers.

The theorem was originally proven by Emanuel Lasker. Lasker was a fascinating figure: he was the World Chess Champion and held the title longer than anyone else in history. He effectively took time off from his chess career to write a 100-page Ph.D. thesis proving this theorem for polynomial rings over fields and integers. His proof was computationally nightmarish.

Decades later, Emmy Noether condensed Lasker's 100-page computational proof into a short, elegant, abstract argument that worked for *all* rings satisfying the ascending chain condition, permanently attaching both their names to the

result. In the next lecture, we will generalize this further by defining co-primary modules and exploring how they relate to primary ideals.

29. The Lasker-Noether Theorem

In this section, we present the *Lasker-Noether theorem*. Historically, the terminology in this subject has become significantly convoluted. To understand why we rely on a concept called a *co-primary* module rather than a *primary* module, we will trace through three historical versions of the theorem. Throughout this section, let R be a Noetherian ring and M be a finitely generated R-module.

29.1. **Three Versions of the Theorem.** The original version of the theorem, proven by Emanuel Lasker for polynomial rings and generalized by Emmy Noether to all Noetherian rings, concerns ideals.

Theorem 29.1 (Version 1: Ideals)**.** *If I is an ideal of a Noetherian ring R, then I is a finite intersection of* primary ideals*. Recall that an ideal J is primary if:*

$$xy \in J \implies x \in J \text{ or } y^n \in J \text{ for some integer } n > 0$$

The second version attempts to generalize this from ideals to modules, introducing the notion of a *primary submodule*.

Theorem 29.2 (Version 2: Submodules)**.** *Suppose we have submodules $N \subseteq M$. Then N is a finite intersection of* primary submodules*. A submodule $X \subseteq M$ is called primary if for all $r \in R$ and $m \in M$:*

$$rm \in X \implies m \in X \text{ or } r^n M \subseteq X \text{ for some } n > 0$$

Notice that there is no such thing as a "primary module" in isolation; it is a property of how the submodule X is embedded inside M. In fact, it is intrinsically a property of the quotient module $Y = M/X$. The condition translates to saying that if $r \in R$ is a zero divisor on Y (meaning $ry = 0$ for some $y \neq 0$), then r is nilpotent on Y:

$$r^n Y = 0 \text{ for some } n > 0$$

A module Y possessing this property is called *co-primary*.

This leads to the third, most conceptually clean version of the theorem, which entirely discards the cumbersome language of primary submodules in favor of co-primary quotient modules.

Theorem 29.3 (Version 3: Modules)**.** *Any finitely generated module M over a Noetherian ring R is contained in a finite direct sum of modules $M_{\mathfrak{p}}$:*

$$M \hookrightarrow \bigoplus M_{\mathfrak{p}}$$

where each $M_{\mathfrak{p}}$ is co-primary *and has exactly one associated prime $\mathfrak{p}$.*

To see why Version 3 implies Version 2, taking $N = 0$ in Version 2 means we can write $0 = \bigcap J_{\mathfrak{p}}$, where each $J_{\mathfrak{p}}$ is a primary submodule. This immediately induces an injective map:

$$M \hookrightarrow \prod \frac{M}{J_{\mathfrak{p}}}$$

Since $J_{\mathfrak{p}}$ is a primary submodule, the quotient $M/J_{\mathfrak{p}}$ is a co-primary module. Thus, studying a single co-primary module cleanly replaces studying pairs of modules and submodules.

29.2. **Equivalence of Co-Primary Definitions.** Alert readers will notice we now have two seemingly distinct definitions of a *co-primary* module. We must prove they are equivalent for finitely generated modules over a Noetherian ring.

Definition 29.4. Let M be a finitely generated R-module. The two definitions of M being *co-primary* are:

(1) M has at most one associated prime (i.e., $|\operatorname{Ass}(M)| \leq 1$).
(2) If $rm = 0$ for some $m \neq 0$, then $r^n M = 0$ for some $n > 0$.

Proof of Equivalence. **(2) $\implies$ (1):** Suppose $M \neq 0$ satisfies condition (2). Let $m \in M$ be a non-zero element, and define $\mathfrak{p} = \operatorname{Ann}(m)$. By deciphering condition (2), any element $r \in \mathfrak{p}$ must satisfy $r^n M = 0$, meaning r belongs to the radical of the annihilator of M. Therefore:

$$\mathfrak{p} \subseteq \sqrt{\operatorname{Ann}(M)}$$

Now assume $\mathfrak{p}$ is prime, meaning $\mathfrak{p} \in \operatorname{Ass}(M)$. We always have the trivial inclusion:

$$\operatorname{Ann}(M) \subseteq \operatorname{Ann}(m) = \mathfrak{p}$$

Taking the radical of both sides (and noting the radical of a prime ideal is itself), we obtain:

$$\sqrt{\operatorname{Ann}(M)} \subseteq \mathfrak{p}$$

Combining the two inclusions, we find that any associated prime $\mathfrak{p}$ must equal $\sqrt{\operatorname{Ann}(M)}$. Since this value is uniquely determined by M, there can be at most one associated prime.

(1) $\implies$ (2): Assume M has exactly one associated prime, $\mathfrak{p}$. Without loss of generality, we can quotient out by $\operatorname{Ann}(M)$ and assume $\operatorname{Ann}(M) = 0$. Since $\mathfrak{p}$ is the only associated prime, it must contain all maximal annihilators of non-zero elements.

We need to show that $\mathfrak{p}$ is nilpotent, meaning $\mathfrak{p}^n = 0$ for some n. Suppose for contradiction that some $a \in \mathfrak{p}$ is not nilpotent. Because $\operatorname{Ann}(M) = 0$ and a is not nilpotent, the localized module M_a (where $S = \{1, a, a^2, \dots\}$) is non-zero.

Since $M_a \neq 0$, it has at least one associated prime $T \in \operatorname{Spec}(R_a)$. Let Q be the inverse image of T in R. Q is a prime ideal and is defined as the union of annihilators:

$$Q = \bigcup_{k=1}^{\infty} \operatorname{Ann}(Ma^k)$$

Because R is Noetherian, this ascending sequence of annihilators stabilizes, so $Q = \operatorname{Ann}(Ma^n)$ for some n. Thus, Q is the annihilator of some non-zero element in M, meaning $Q \in \operatorname{Ass}(M)$.

However, $a \in \mathfrak{p}$ by assumption, but $a \notin Q$ (otherwise $a/1$ would not be a unit in R_a, contradicting that T is a proper ideal). Therefore, $\mathfrak{p} \neq Q$. This contradicts the assumption that M has only one associated prime. Thus, $\mathfrak{p}$ must be nilpotent, proving condition (2). □

29.3. **Proof of the Lasker-Noether Theorem.** Using the clean definition of co-primary modules (at most one associated prime), the proof of the Lasker-Noether theorem becomes spectacularly brief, bypassing the hundred pages of computations required in Lasker's original manuscript.

Proof of Theorem 29.3. Suppose the theorem is false. Since M is a Noetherian module, we can pick a maximal submodule $N \subsetneq M$ such that the quotient M/N is *not* contained in a finite direct sum of co-primary modules. By quotienting out by N, we may assume without loss of generality that $N = 0$. Thus, 0 is the maximal submodule failing the condition, meaning M itself cannot be embedded into a finite sum of co-primary modules, but every proper quotient of M can be.

First, M cannot be co-primary (otherwise it would trivially embed into itself). By our equivalent definition, M must have at least two distinct associated primes, $\mathfrak{p}_1 \neq \mathfrak{p}_2$.

This implies M contains two submodules, $M_1 \cong R/\mathfrak{p}_1$ and $M_2 \cong R/\mathfrak{p}_2$. The annihilators of these submodules are precisely their corresponding primes:

$$\operatorname{Ann}(M_1) = \mathfrak{p}_1$$
$$\operatorname{Ann}(M_2) = \mathfrak{p}_2$$

Consider the intersection $M_1 \cap M_2$. Any element $x \in M_1 \cap M_2$ is annihilated by both $\mathfrak{p}_1$ and $\mathfrak{p}_2$. However, M_1 is an integral domain $R/\mathfrak{p}_1$, so the annihilator of any non-zero element in M_1 is exactly $\mathfrak{p}_1$. Since $\mathfrak{p}_1 \neq \mathfrak{p}_2$, the intersection must be trivial:

$$M_1 \cap M_2 = 0$$

We can now construct a natural map from M into the direct sum of its quotients:

$$M \to \frac{M}{M_1} \oplus \frac{M}{M_2}$$

The kernel of this map is exactly $M_1 \cap M_2 = 0$. Therefore, the map is injective:

$$M \hookrightarrow \frac{M}{M_1} \oplus \frac{M}{M_2}$$

Because we assumed 0 was the *maximal* submodule for which the theorem fails, and $M_1, M_2 \neq 0$, the proper quotients M/M_1 and M/M_2 both satisfy the theorem. That is, both can be embedded into a finite product of co-primary modules.

Since M embeds into the direct sum of these two quotients, M itself embeds into a finite product of co-primary modules. This is a direct contradiction of our starting assumption. Therefore, the theorem must be true for all finitely generated modules. □

29.4. **Recovering the Original Theorem.** We conclude by explicitly demonstrating that this modernized module-theoretic version implies Lasker's original version for ideals.

Let $I \subseteq R$ be an ideal. We apply Theorem 29.3 to the quotient module $M = R/I$. We obtain an injection:

$$\frac{R}{I} \hookrightarrow \prod_{j=1}^{k} M_j$$

where each M_j is a co-primary module. Without loss of generality, we can replace the M_j with the images of R under the composition of the injection and the canonical projections. Thus, we may assume $M_j \cong R/I_j$ for some ideals I_j.

Because R/I_j is a co-primary module, the ideal I_j is, by definition, a primary ideal in Lasker's sense. The injectivity of the map from R/I into the product $\prod R/I_j$ implies that the kernel is zero, yielding:

$$I = \bigcap_{j=1}^{k} I_j$$

Thus, I is successfully expressed as a finite intersection of primary ideals, recovering the classical Lasker-Noether theorem in a handful of lines.

30. Symbolic Powers

In the previous lectures, we discussed the Lasker-Noether theorem, which establishes that in a Noetherian ring, every ideal can be expressed as a finite intersection of primary ideals. In this section, we will construct explicit examples of primary ideals and discuss their relationship to prime ideals. In particular, we will introduce *symbolic powers*, a construction that elegantly remedies the failure of ordinary powers of prime ideals to remain primary.

30.1. **Primary Ideals versus Powers of Primes.** A natural first question to ask is: what do primary ideals look like?

For a principal ideal domain such as the integers $\mathbb{Z}$, the primary ideals are precisely (0) and the ideals generated by prime powers, (p^n). This simple classification suggests a naive hypothesis: perhaps primary ideals are always just powers of prime ideals.

This hypothesis is overwhelmingly false. In general commutative rings, the correspondence fails in both directions:

(1) There are primary ideals that are not powers of any prime ideal.
(2) There are powers of prime ideals that are not primary ideals.

We will construct explicit counterexamples for both cases.

30.1.1. *Primary Ideals that are Not Powers of Primes.* Constructing a primary ideal that is not a prime power is straightforward. We simply require a polynomial

ring with at least two variables. Let $R = k[x, y]$. We define our prime ideal to be the maximal ideal at the origin:

$$\mathfrak{p} = (x, y)$$

We can visualize R by drawing a 2D grid of its monomials. The ideal $\mathfrak{p}$ is generated by x and y. Its powers are generated by all monomials of a given total degree. For instance, the cube of $\mathfrak{p}$ is:

$$\mathfrak{p}^3 = (x^3, x^2y, xy^2, y^3)$$

Any power of $\mathfrak{p}$ must look geometrically like a perfect diagonal "staircase" cutting across the monomial grid.

However, the geometric space of primary ideals is far less rigid. Let us choose an arbitrary ideal generated by a somewhat irregular set of monomials, such as:

$$I = (x^6, x^3y, x^2y^2, y^3)$$

We claim that I is a $\mathfrak{p}$-primary ideal. To verify this, we look at the quotient module $M = R/I$. Because I contains pure powers of both x and y (namely x^6 and y^3), the quotient M is a finite-dimensional vector space over k. Consequently, M has a finite composition series of submodules:

$$0 = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_n = M$$

Because the only prime ideal containing I is the maximal ideal $\mathfrak{p} = (x, y)$, the only associated prime of the module M is (x, y). A module with exactly one associated prime is co-primary, which inherently means that the ideal I is primary.

Yet, I is manifestly not a power of $\mathfrak{p}$, because its generators do not possess a uniform total degree. Thus, there are vastly more primary ideals corresponding to a prime than there are powers of that prime.

Remark 30.1. It is crucial to distinguish between two distinct usages of the ideal $\mathfrak{p} = (x, y)$. If we view $\mathfrak{p}$ itself as an R-module, it is a co-primary module with associated prime (0). However, if we view the quotient $R/\mathfrak{p}$ as a module, it is a co-primary module with associated prime $\mathfrak{p} = (x, y)$.

30.1.2. *Powers of Primes that are Not Primary.* The failure in the opposite direction is far more surprising. One naturally expects the algebraic powers of a prime ideal to be algebraically well-behaved. However, the presence of geometric singularities severely disrupts this behavior.

Let R be the coordinate ring of a double cone over a field k:

$$R = \frac{k[x, y, z]}{(xy - z^2)}$$

Consider the ideal $\mathfrak{p} = (x, z)$. This ideal corresponds geometrically to the y-axis, which is a straight line lying entirely on the surface of the cone. We first verify that $\mathfrak{p}$ is prime by examining the quotient:

$$\frac{R}{\mathfrak{p}} = \frac{k[x, y, z]}{(x, z, xy - z^2)} \cong k[y]$$

Since $k[y]$ is an integral domain, $\mathfrak{p}$ is indeed a prime ideal.

Now, we compute the square of this prime ideal, $\mathfrak{p}^2$. In the polynomial ring, it is generated by x^2, xz, and z^2. However, inside our quotient ring R, we have the relation $z^2 = xy$. Substituting this into the generators yields:

$$\mathfrak{p}^2 = (x^2, xz, xy) = x(x, y, z)$$

We now test whether $\mathfrak{p}^2$ is primary. By definition, $\mathfrak{p}^2$ is primary if and only if every zero divisor in $R/\mathfrak{p}^2$ is nilpotent.

Let us look at the element y in the quotient $R/\mathfrak{p}^2$. From our generating set, we see that $xy \in \mathfrak{p}^2$. Therefore, in the quotient ring:

$$x \cdot y = 0$$

Since $x \notin \mathfrak{p}^2$, the element y is a zero divisor.

If $\mathfrak{p}^2$ were primary, y would have to be nilpotent. However, no power of y lies in $\mathfrak{p}^2$. The elements $1, y, y^2, \dots$ are all linearly independent modulo $\mathfrak{p}^2$ because $\mathfrak{p}^2$ is entirely contained within the ideal (x). Since y is a non-nilpotent zero divisor, $\mathfrak{p}^2$ is *not* a primary ideal.

Because $\mathfrak{p}^2$ is not primary, the Lasker-Noether theorem guarantees it must have a non-trivial primary decomposition. In this case, the decomposition isolates the singularity:

$$\mathfrak{p}^2 = (x) \cap (x, y, z)^2$$

Here, (x) is the $\mathfrak{p}$-primary component corresponding to the line itself (a "doubled" version of the y-axis), while $(x, y, z)^2$ is an embedded primary component corresponding to the singular point at the origin. The set of associated primes is $\mathrm{Ass}(R/\mathfrak{p}^2) = \{(x, z), (x, y, z)\}$, which explicitly reveals the embedded component.

30.2. **Powers of Maximal Ideals.** While powers of arbitrary prime ideals need not be primary, there is one critical exception: maximal ideals.

Theorem 30.2. *Let R be a Noetherian ring and let $\mathfrak{m}$ be a maximal ideal. If I is any ideal such that $\mathfrak{m}^n \subseteq I \subseteq \mathfrak{m}$ for some integer $n \geq 1$, then I is an $\mathfrak{m}$-primary ideal.*

Proof. Consider the quotient ring R/I. Because I contains $\mathfrak{m}^n$, the maximal ideal $\mathfrak{m}/I$ in the quotient ring is nilpotent.

Since $\mathfrak{m}$ is a maximal ideal in R, $\mathfrak{m}/I$ is the unique maximal ideal in R/I, meaning R/I is a local ring. In a local ring where the maximal ideal is nilpotent, every element is either a unit (if it lies outside the maximal ideal) or nilpotent (if it lies inside). Consequently, every zero divisor is necessarily nilpotent. This precisely means that I is a primary ideal. □

Corollary 30.3. *If $\mathfrak{m}$ is a maximal ideal, its n-th power $\mathfrak{m}^n$ is always a primary ideal.*

In our double cone example, the ideal (x, y, z) is maximal, so its square $(x, y, z)^2$ is automatically primary. However, $\mathfrak{p} = (x, z)$ is not maximal, which allowed its square to pick up pathological embedded components.

30.3. **Symbolic Powers.** We have seen that ordinary algebraic powers of prime ideals are defective because they can pick up embedded associated primes at singularities. To remedy this, we introduce a variation called the *symbolic power*, which forces the result to be primary by explicitly filtering out embedded components via localization.

Definition 30.4. Let $\mathfrak{p}$ be a prime ideal in a commutative ring R. The n-th *symbolic power* of $\mathfrak{p}$, denoted $\mathfrak{p}^{(n)}$, is defined as the contraction of the n-th power of the extended ideal in the local ring $R_{\mathfrak{p}}$:

$$\mathfrak{p}^{(n)} = \{x \in R \mid xy \in \mathfrak{p}^n \text{ for some } y \notin \mathfrak{p}\}$$

Let us break down why this definition works:

(1) We localize the ring at $\mathfrak{p}$, constructing the map $\phi\colon R \to R_{\mathfrak{p}}$.
(2) In the localized ring $R_{\mathfrak{p}}$, the extended ideal $\mathfrak{p}R_{\mathfrak{p}}$ is the unique maximal ideal.
(3) Because it is maximal, its algebraic power $\mathfrak{p}^n R_{\mathfrak{p}}$ is primary.
(4) We pull this ideal back to R. The inverse image of a primary ideal under a ring homomorphism is always a primary ideal.

Thus, $\mathfrak{p}^{(n)} = \phi^{-1}(\mathfrak{p}^n R_{\mathfrak{p}})$ is guaranteed to be a $\mathfrak{p}$-primary ideal.

30.3.1. *Symbolic Powers on the Cone.* Let us compute the symbolic square of $\mathfrak{p} = (x, z)$ in our singular cone ring $R = k[x, y, z]/(xy - z^2)$.

We previously computed the ordinary square:

$$\mathfrak{p}^2 = x(x, y, z)$$

This ideal contains the element xy. Notice that $y \notin \mathfrak{p}$ (since the y-axis is not contained in the x-axis).

According to the definition of the symbolic power, if a product xy lies in $\mathfrak{p}^2$ and $y \notin \mathfrak{p}$, then the element x must belong to the symbolic square $\mathfrak{p}^{(2)}$. Thus, the symbolic square contains x. Since the ideal (x) is already $\mathfrak{p}$-primary in this ring, we conclude:

$$\mathfrak{p}^{(2)} = (x)$$

Notice that $\mathfrak{p}^{(2)}$ is strictly larger than $\mathfrak{p}^2$. While $x \in \mathfrak{p}^{(2)}$, it is easily seen by degree arguments that $x \notin \mathfrak{p}^2$. The symbolic power has elegantly discarded the embedded singularity $(x, y, z)^2$ from the primary decomposition.

30.3.2. *Pathologies of Inclusion.* Symbolic powers resolve one pathology but introduce another regarding inclusions. In standard ideal arithmetic, if an ideal is contained in another, their powers respect this inclusion:

$$I \subseteq J \implies I^n \subseteq J^n$$

This fails spectacularly for symbolic powers.

In our cone example, consider the inclusion of prime ideals:

$$\mathfrak{p} = (x, z) \subseteq \mathfrak{m} = (x, y, z)$$

Let us compute their symbolic squares. For $\mathfrak{m}$, since it is maximal, its symbolic square is identical to its ordinary square:

$$\mathfrak{m}^{(2)} = \mathfrak{m}^2 = (x, y, z)^2$$

However, for $\mathfrak{p}$, we found that $\mathfrak{p}^{(2)} = (x)$.

Observe that $x \in \mathfrak{p}^{(2)}$, but $x \notin \mathfrak{m}^2$ because x is a degree 1 monomial and $\mathfrak{m}^2$ contains only elements of degree 2 and higher. Therefore:

$$\mathfrak{p}^{(2)} \not\subseteq \mathfrak{m}^{(2)}$$

Just because one prime is contained in another does *not* imply their symbolic powers are contained in one another.

30.4. **Geometric Meaning.** Finally, we can ask what the symbolic power represents geometrically.

Informally, the algebraic power $\mathfrak{p}^n$ consists of functions that can be written as linear combinations of products of n functions, each vanishing to order 1 along the variety $Z(\mathfrak{p})$.

The symbolic power $\mathfrak{p}^{(n)}$, on the other hand, consists of all functions that conceptually "vanish to order n" along $Z(\mathfrak{p})$.

On a smooth variety, these two concepts are identical. A function vanishing to order 2 along a smooth curve can always be expressed as a sum of products of pairs of functions vanishing to order 1. However, as our cone example demonstrates, at a geometric singularity, things go awry. The function x vanishes to order 2 along the y-axis, but because of the singularity at the origin, it cannot be decomposed into a product of functions vanishing to order 1. The symbolic power captures the true geometric notion of vanishing order, bypassing the algebraic defect at the singular point.

Part 4. Nullstellensatz, Integral Dependence, and Key Lemmas

31. Nullstellensatz

This lecture covers the famous Hilbert Nullstellensatz. The term "Nullstellensatz" translates from German as "zero-position-theorem" or simply "theorem of zeros". There are two primary forms of this theorem: a weak version and a strong version.

31.1. **The Weak Nullstellensatz.** We begin by considering the maximal ideals of a polynomial ring $k[x_1, \dots, x_n]$. For simplicity of notation, we often write this for two variables as $k[x, y]$.

There are some obvious maximal ideals, namely those generated by linear polynomials corresponding to points in the affine space:

$$(x - a, y - b, \dots)$$

These ideals naturally correspond to points $(a, b, \dots)$ in the affine space k^n. Note the slightly illogical standard notation: the parentheses in $(x - a, y - b)$ denote the ideal generated by these elements, whereas the parentheses in (a, b) denote the coordinates of a geometric point.

Theorem 31.1 (Weak Nullstellensatz)**.** *If k is an algebraically closed field, then all maximal ideals of the polynomial ring $k[x_1, \ldots, x_n]$ are of the form $(x_1 - a_1, \ldots, x_n - a_n)$ for some point $(a_1, \ldots, a_n) \in k^n$.*

If the field k is not algebraically closed, there can exist other maximal ideals that do not correspond to geometric points in k^n.

Example 31.2. Let $k = \mathbb{R}$, and consider the polynomial ring in one variable $\mathbb{R}[x]$. The ideal generated by $x^2 + 1$ is a maximal ideal:

$$I = (x^2 + 1)$$

However, this ideal does not correspond to any point on the real line, because $x^2 + 1 = 0$ has no solutions in $\mathbb{R}$. It geometrically relates to the point i in the complex numbers, which lies in the algebraic closure of $\mathbb{R}$.

31.2. **The Strong Nullstellensatz.** Suppose I is an arbitrary ideal in $k[x_1, \ldots, x_n]$. We define the *variety* of zeros of I, denoted V (or $Z(I)$), as the set of points where all elements of I vanish simultaneously.

Conversely, given this variety V, we can ask: what is the ideal of all polynomials vanishing on V? Let us call this ideal J (or $I(V)$). The most obvious guess in the early days of algebraic geometry was that J is simply equal to I.

However, while I is obviously contained in J, they are not necessarily equal.

Example 31.3. Consider the ring $k[x]$ and let the ideal be generated by x^2:

$$I = (x^2)$$

The variety of zeros V is simply the origin $\{0\}$. The set of all polynomials vanishing at the origin is the ideal generated by x:

$$J = (x)$$

Clearly, $J \neq I$. The polynomial $x^2 \in I$, but $x \notin I$.

More generally, if $f^n \in I$ for some integer $n > 0$, then f^n vanishes on the entirety of V. This implies that f itself must vanish on V, meaning $f \in J$. Therefore, the *radical* of I, denoted $\sqrt{I}$, is always contained in J.

Theorem 31.4 (Strong Nullstellensatz)**.** *If k is an algebraically closed field, then the ideal of polynomials vanishing on the zero locus of I is exactly the radical of I. That is:*

$$J = \sqrt{I}$$

In quotient ring terms, the ideal J/I forms the *nilradical* of the ring R/I, which consists precisely of the nilpotent elements of the quotient.

31.3. **The Difficulty of Finding the Radical.** One might assume that finding the nilpotent elements of a ring (and thus the radical of an ideal) is an easy task. If the square of an element is zero, it should seemingly be obvious what that element is. In practice, however, finding the radical of an ideal can be extraordinarily difficult.

31.3.1. *Nilpotent Matrices.* Consider the space of nilpotent $n \times n$ matrices over k. Let X be a matrix with generic entries x_{ij}. The condition that X is nilpotent implies that $X^n = 0$. We can define an ideal I in the polynomial ring $k[x_{11}, \ldots, x_{nn}]$ generated by the n^2 entries of the matrix X^n.

The zero set V of this ideal is precisely the set of nilpotent matrices. We want to find other functions (polynomials) that vanish on V.

- If X is nilpotent, all of its eigenvalues are zero.
- Therefore, the trace of X, which is the sum of the eigenvalues, must be zero.
- Thus, the trace polynomial $\sum x_{ii}$ vanishes on V.

The trace is a homogeneous polynomial of degree 1. However, the ideal I is generated by the entries of X^n, which are all homogeneous polynomials of degree n. Thus, the trace cannot possibly be in I, but it lies in the radical $\sqrt{I}$.

Furthermore, the characteristic polynomial of X is:

$$\det(\lambda I - X) = \lambda^n$$

for a nilpotent matrix. Therefore, all lower-degree coefficients of the characteristic polynomial also vanish on V and thus belong to the radical.

31.3.2. *Explicit* 2×2 *Example.* Let us do this explicitly for 2×2 matrices to see the hidden complexity. Let:

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

We compute X^2:

$$X^2 = \begin{pmatrix} a^2 + bc & ab + bd \\ ac + cd & d^2 + bc \end{pmatrix}$$

The ideal I is generated by these four entries:

$$I = (a^2 + bc, b(a + d), c(a + d), d^2 + bc)$$

We know from the Strong Nullstellensatz that the trace $a + d$ must be in the radical $\sqrt{I}$. This means some power $(a + d)^k$ must be in I.

A naive guess for 2×2 matrices might be $k = 2$. However, one can verify that $(a + d)^2 \notin I$. It turns out we must go up to the third power: $(a + d)^3 \in I$. To see this algebraically, note that I contains:

$$(a + d)b, \quad (a + d)c, \quad (a + d)(a - d)$$

where the last element comes from $(a^2 + bc) - (d^2 + bc) = a^2 - d^2$. It also contains a^2 and d^2 when multiplied appropriately. We can write $(a + d)^2$ as:

$$(a + d)^2 = 2a^2 + 2d^2 - (a - d)^2$$

By carefully manipulating these combinations, one can eventually deduce that $(a + d)^3 \in I$, but it requires significant algebraic work.

Alternatively, utilizing the determinant $\det(X) = ad - bc$, we know by the Cayley-Hamilton theorem that:

$$X^2 - \mathrm{Tr}(X)X + \det(X)I = 0$$

Since $X^2 \in I$, evaluating this carefully leads to the exact generators of the radical. The full radical is generated by the trace and determinant:

$$\sqrt{I} = (a + d, ad - bc)$$

If we quotient the polynomial ring $k[a, b, c, d]$ by this radical, substituting $d = -a$ yields:

$$\frac{k[a, b, c]}{(a^2 + bc)}$$

This quotient is an integral domain (the defining polynomial is an irreducible quadric cone), which confirms that we have found the complete radical, as it contains no further nilpotent elements.

31.3.3. *Commuting Matrices.* For an even harder example, let X and Y be two generic $n \times n$ matrices. Let I be the ideal generated by the entries of the commutator $XY - YX$. The vanishing set V is the space of pairs of commuting matrices.

What is the radical of I? Does $\sqrt{I} = I$? This remains an extraordinarily difficult open problem. Even though "commuting matrices" is one of the simplest conditions to state, determining the algebraic radical of this ideal is currently unsolved, demonstrating that finding radicals is generally a highly non-trivial task.

31.4. **Proof of the Weak Nullstellensatz over $\mathbb{C}$.** We can provide an incredibly short proof of the Weak Nullstellensatz if we "cheat" by restricting our field to the complex numbers $\mathbb{C}$ and using set-theoretic cardinality arguments rather than pure algebra.

Proof. Let $\mathfrak{m}$ be a maximal ideal in the polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$. We define the quotient field:

$$F = \frac{\mathbb{C}[x_1, \ldots, x_n]}{\mathfrak{m}}$$

This quotient F is a field (because $\mathfrak{m}$ is maximal) and it is a finitely generated $\mathbb{C}$-algebra.

Notice that F has at most countable dimension when viewed as a vector space over $\mathbb{C}$, because it is spanned by the countable set of all monomials in $x_1, \ldots, x_n$.

We claim that F must be an algebraic extension of $\mathbb{C}$. Suppose, for the sake of contradiction, that F is not algebraic. Then it must contain some element t that is transcendental over $\mathbb{C}$.

If t is transcendental, then for every complex number $\alpha \in \mathbb{C}$, the element $(t - \alpha)^{-1}$ exists in the field F. Consider the set of elements:

$$\left\{ \frac{1}{t - \alpha} \;\middle|\; \alpha \in \mathbb{C} \right\}$$

Because $\mathbb{C}$ is uncountably infinite, this is an uncountable set of elements in F. Furthermore, one can easily show via partial fractions that these elements are linearly independent over $\mathbb{C}$.

However, this implies F has an uncountable dimension as a $\mathbb{C}$-vector space, which contradicts our earlier observation that F has countable dimension. Thus, our assumption was false, and F contains no transcendental elements. F must be an algebraic extension of $\mathbb{C}$.

Since the complex numbers $\mathbb{C}$ are algebraically closed, the only algebraic extension of $\mathbb{C}$ is $\mathbb{C}$ itself. Therefore:

$$F = \mathbb{C}$$

This isomorphism means that the image of each variable x_i in the quotient field F is simply some complex number $a_i \in \mathbb{C}$. Consequently:

$$x_i \equiv a_i \pmod{\mathfrak{m}}$$

which implies that $(x_i - a_i) \in \mathfrak{m}$. Since the ideal generated by $(x_1 - a_1, \ldots, x_n - a_n)$ is already maximal, it must exactly equal $\mathfrak{m}$. □

This proof works over any uncountable algebraically closed field. For general fields, a more honest, purely algebraic proof (typically relying on Noether Normalization or Zariski's Lemma) is required.

31.5. **The Rabinowitsch Trick.** Having established the Weak Nullstellensatz, we can now elegantly prove that the Weak version implies the Strong version. This method is known as the *Rabinowitsch trick*.

Theorem 31.5. *The Weak Nullstellensatz implies the Strong Nullstellensatz.*

Proof. Let I be an ideal in $k[x_1, \ldots, x_n]$, let $V = Z(I)$ be its zero set, and let f be a polynomial that vanishes everywhere on V. We must show that $f^m \in I$ for some integer $m > 0$.

We introduce a new, auxiliary variable x_0. We define a new ideal J in the larger polynomial ring $k[x_0, x_1, \ldots, x_n]$, generated by the original ideal I and the polynomial $1 - x_0 f$:

$$J = \langle I, 1 - x_0 f \rangle$$

Let us find the common zeros of J in k^{n+1}. Any such zero must be a common root of all elements in I, meaning its coordinates $(x_1, \ldots, x_n)$ must lie in V. However, by our hypothesis, the polynomial f evaluates to 0 on any point in V.

Thus, evaluated at such a point, the polynomial $1 - x_0 f$ becomes:

$$1 - x_0(0) = 1 \neq 0$$

Therefore, the elements of J have absolutely no common roots. The zero set of J is empty.

By the Weak Nullstellensatz (applied in $n + 1$ variables), if an ideal has an empty zero set, it cannot be contained in any maximal ideal. The only such ideal is the entire ring. Thus:

$$J = (1)$$

Because $1 \in J$, we can express 1 as a finite polynomial linear combination of the generators of J:

$$1 = A_1 B_1 + \cdots + A_k B_k + A(1 - x_0 f)$$

where $B_i \in I$ and the coefficients $A_i, A \in k[x_0, x_1, \ldots, x_n]$.

This is an identity in the polynomial ring. We can evaluate it under the formal substitution $x_0 = 1/f$. This substitution takes place in the localized ring of rational functions $k[x_1, \ldots, x_n][1/f]$. Under this substitution, the term $1 - x_0 f$ vanishes completely:

$$1 = \tilde{A}_1 B_1 + \cdots + \tilde{A}_k B_k$$

where each $\tilde{A}_i$ is a rational function of the form C_i/f^{m_i} for some polynomial $C_i \in k[x_1, \ldots, x_n]$.

Let m be the highest power of f appearing in any denominator. We can clear all denominators by multiplying the entire equation by f^m:

$$f^m = (f^m \tilde{A}_1) B_1 + \cdots + (f^m \tilde{A}_k) B_k$$

Because $m \geq m_i$, each term $f^m \tilde{A}_i$ is now a standard polynomial in $k[x_1, \ldots, x_n]$. Since each $B_i \in I$, the right-hand side is a linear combination of elements of I with polynomial coefficients. Thus, the right-hand side is entirely contained in I, which forces:

$$f^m \in I$$

This concludes the proof of the Strong Nullstellensatz. □

32. Zariski's Lemma

This lecture focuses on providing a precise proof of the Nullstellensatz that applies to all algebraically closed fields. Previously, we proved the Weak Nullstellensatz over the complex numbers $\mathbb{C}$ by using its uncountability. While one could invoke the Lefschetz principle from model theory to extend this to all algebraically closed fields of a given characteristic, it is far more instructive (and standard) to rely on a purely commutative algebraic result known as *Zariski's Lemma*.

32.1. **Recap of the Nullstellensatz.** Recall the two forms of the Nullstellensatz for a polynomial ring over an algebraically closed field k:

- **Weak Nullstellensatz:** The maximal ideals of $k[x_1, \ldots, x_n]$ are precisely the obvious ones corresponding to the points of affine space, taking the form:

  $$\mathfrak{m} = (x_1 - a_1, \ldots, x_n - a_n)$$

  for $(a_1, \ldots, a_n) \in k^n$.
- **Strong Nullstellensatz:** If I is an ideal and $V(I)$ is its zero set, then any polynomial f vanishing entirely on $V(I)$ satisfies:

  $$f^k \in I$$

  for some integer $k \geq 1$. Equivalently, the ideal of functions vanishing on $V(I)$ is exactly the radical $\sqrt{I}$.

32.2. **Zariski's Lemma.** The key to proving the Nullstellensatz in full generality is Zariski's Lemma.

Lemma 32.1 (Zariski's Lemma)**.** *If a field K is finitely generated as an algebra over a field k, then K is finitely generated as a module over k. (In other words, K is a finite algebraic extension of k).*

Before diving into the formal proof, let us sketch the intuition. Suppose K is merely the field of rational functions $k(x_1, \ldots, x_m)$. We ask: can this field be finitely generated as an algebra over k?

If K were generated by a finite number of rational functions, all elements of K would be expressible as polynomials in these generators. The generators have the form:

$$\frac{f_i(x_1, \ldots, x_m)}{g_i(x_1, \ldots, x_m)}$$

Any polynomial combination of these generators will only ever have poles that are factors of the finitely many denominators g_i. Thus, the possible poles of all elements in the generated algebra are constrained to a finite set of irreducible polynomials. However, if $m > 0$, the polynomial ring $k[x_1, \ldots, x_m]$ has an infinite number of irreducible polynomials, meaning there are infinitely many places a rational function could have a pole. Thus, a finite number of generators simply do not provide "enough poles" to generate the entire field of rational functions.

We now formalize this argument to accommodate fields that are larger than purely rational function fields.

Proof of Zariski's Lemma. Assume K is a field that is finitely generated as a k-algebra by elements $x_1, \ldots, x_n$. We can sort and renumber these generators such that $x_1, \ldots, x_m$ form a maximal algebraically independent set over k. The remaining generators $x_{m+1}, \ldots, x_n$ are therefore algebraic over the field of fractions $k(x_1, \ldots, x_m)$.

If $m = 0$, all generators are algebraic over k, and K is a finite extension. The lemma is proven.

Thus, we assume $m > 0$ and look for a contradiction. The subring $k[x_1, \ldots, x_m]$ is a polynomial ring in m variables. Because $m > 0$, this ring has an infinite number of irreducibles up to units. (If k is infinite, one can simply take $x_1 - \alpha$ for all $\alpha \in k$. If k is finite, one can copy Euclid's proof for the infinitude of primes to construct infinitely many irreducibles).

Each of the remaining generators x_j (for $j = m + 1, \ldots, n$) is algebraic over $k(x_1, \ldots, x_m)$. Hence, x_j is the root of some non-zero polynomial with coefficients in $k(x_1, \ldots, x_m)$. By clearing denominators, we can find a polynomial equation for x_j with coefficients in $k[x_1, \ldots, x_m]$. Let $b_j \in k[x_1, \ldots, x_m]$ be the leading coefficient of this polynomial.

These b_j informally represent the "possible poles" introduced by the algebraic elements. We localize our polynomial ring by inverting these leading coefficients:

$$R = k[x_1, \ldots, x_m]\left[\frac{1}{b_{m+1}}, \ldots, \frac{1}{b_n}\right]$$

Over the ring R, the equations for x_j can be made monic (by dividing by b_j, which is now a unit). Thus, each x_j is *integral* over R. Because K is generated by elements integral over R, K is a finitely generated R-module.

Because there are infinitely many irreducibles in $k[x_1, \ldots, x_m]$ and only finitely many b_j, we can pick an irreducible polynomial $f \in k[x_1, \ldots, x_m]$ that does not divide any of the b_j.

Since K is a field, the inverse f^{-1} belongs to K. The field K is a finitely generated module over R, and R is Noetherian, so we can consider the ascending chain of R-submodules generated by powers of f^{-1}:

$$\langle 1 \rangle \subseteq \langle 1, f^{-1} \rangle \subseteq \langle 1, f^{-1}, f^{-2} \rangle \subseteq \ldots$$

This chain must eventually stabilize, meaning that for some N, f^{-N} can be expressed as an R-linear combination of strictly lower powers:

$$f^{-N} = a_0 + a_1 f^{-1} + \cdots + a_{N-1} f^{-(N-1)} \quad \text{for } a_i \in R$$

Multiplying both sides by f^{N-1} yields:

$$f^{-1} = a_0 f^{N-1} + a_1 f^{N-2} + \cdots + a_{N-1}$$

Since $f \in R$ and all $a_i \in R$, the right-hand side is an element of R. Therefore, $f^{-1} \in R$.

By the definition of R, any element in R can be written as a polynomial in $x_1, \ldots, x_m$ divided by some product of powers of the b_j. Thus:

$$\frac{1}{f} = \frac{P}{b_{m+1}^{k_{m+1}} \cdots b_n^{k_n}}$$

Cross-multiplying gives:

$$b_{m+1}^{k_{m+1}} \cdots b_n^{k_n} = f \cdot P$$

This implies that the irreducible polynomial f must divide one of the b_j, which directly contradicts our choice of f. This contradiction shows that m cannot be greater than 0, concluding the proof. □

32.3. **Proof of the Weak Nullstellensatz.** With Zariski's Lemma established, the Weak Nullstellensatz follows almost immediately.

Proof. Let $\mathfrak{m}$ be a maximal ideal in $k[x_1, \ldots, x_n]$, where k is algebraically closed. The quotient ring $K = k[x_1, \ldots, x_n]/\mathfrak{m}$ is a field. Furthermore, K is finitely generated as an algebra over k (generated by the images of the variables x_i).

By Zariski's Lemma, K is a finite algebraic extension of k. However, because k is algebraically closed, it has no non-trivial finite algebraic extensions. Therefore, the canonical embedding $k \hookrightarrow K$ is an isomorphism:

$$K = k$$

This means that the image of each variable x_i in K must be equal to some scalar $a_i \in k$. Consequently, in the quotient, we have:

$$x_i \equiv a_i \pmod{\mathfrak{m}}$$

This implies that the polynomial $(x_i - a_i)$ lies in $\mathfrak{m}$. Since this holds for all $i = 1, \ldots, n$, the maximal ideal $\mathfrak{m}$ must contain the ideal generated by these linear terms:

$$(x_1 - a_1, \ldots, x_n - a_n) \subseteq \mathfrak{m}$$

Because the ideal on the left is already maximal, the inclusion is an equality, completing the proof. □

32.4. **Pullbacks of Maximal Ideals.** A highly useful geometric corollary of Zariski's Lemma concerns homomorphisms of finitely generated algebras. In general ring theory, the inverse image of a maximal ideal under a homomorphism is prime, but not necessarily maximal. (For example, under $\mathbb{Z} \hookrightarrow \mathbb{Q}$, the inverse image of the maximal ideal (0) is (0), which is not maximal in $\mathbb{Z}$). However, for finitely generated algebras over fields, this pathology vanishes.

Corollary 32.2. *Let $f\colon A \to B$ be a homomorphism of finitely generated k-algebras. If $\mathfrak{m} \subset B$ is a maximal ideal, then the inverse image $f^{-1}(\mathfrak{m})$ is a maximal ideal in A.*

Proof. The homomorphism f induces an injective map on the quotients:

$$\frac{A}{f^{-1}(\mathfrak{m})} \hookrightarrow \frac{B}{\mathfrak{m}}$$

Since $\mathfrak{m}$ is maximal, $B/\mathfrak{m}$ is a field. Because B is a finitely generated k-algebra, $B/\mathfrak{m}$ is a field that is finitely generated as a k-algebra. By Zariski's Lemma, it is a finite algebraic extension of k.

Thus, $A/f^{-1}(\mathfrak{m})$ is a k-subalgebra of a finite algebraic extension of k. An elementary result from field theory states that any sub-domain of a finite field extension that contains the base field is itself a field. Because $A/f^{-1}(\mathfrak{m})$ is a field, the ideal $f^{-1}(\mathfrak{m})$ must be maximal. □

This result is why, in classical algebraic geometry, one could often work entirely with the maximal spectrum (the set of maximal ideals) without utilizing prime ideals; for finitely generated algebras over fields, pulling back functions geometrically properly maps points to points.

32.5. **Strong Nullstellensatz via Localization.** Finally, we present an alternate proof showing that the Weak Nullstellensatz implies the Strong Nullstellensatz, emphasizing localization rather than the explicit Rabinowitsch trick (though mathematically they are identical).

Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal, and let V be its zero set. Suppose f is a polynomial that vanishes on V. We want to show $f^k \in I$ for some k.

We consider the localization of the polynomial ring at the element f:

$$R_f = k[x_1, \ldots, x_n][f^{-1}]$$

By the properties of localization, the maximal ideals of R_f correspond precisely to the maximal ideals of $k[x_1, \ldots, x_n]$ that do not contain f.

Geometrically, the maximal ideals of $k[x_1, \ldots, x_n]$ take the form $(x_1 - a_1, \ldots, x_n - a_n)$. Such an ideal does not contain f if and only if $f(a_1, \ldots, a_n) \neq 0$.

Now, let $I \cdot R_f$ be the extension of I in this localized ring. We claim $I \cdot R_f$ is not contained in any maximal ideal of R_f. Suppose for contradiction that it were contained in some maximal ideal $\mathfrak{m}_{R_f}$. This would correspond to a maximal ideal $\mathfrak{m} \subset k[x_1, \ldots, x_n]$ such that $I \subseteq \mathfrak{m}$ and $f \notin \mathfrak{m}$.

- $I \subseteq \mathfrak{m}$ means the point defined by $\mathfrak{m}$ lies in the zero set V.
- $f \notin \mathfrak{m}$ means f does not vanish at this point.

This contradicts the hypothesis that f vanishes everywhere on V.

Therefore, the ideal $I \cdot R_f$ is not contained in any maximal ideal. In any commutative ring, the only ideal not contained in a maximal ideal is the entire ring itself. Thus:

$$I \cdot R_f = R_f$$

Because the extended ideal is the whole ring, it must contain the identity element 1:

$$1 \in I \cdot R_f$$

The elements of $I \cdot R_f$ are of the form a/f^k for $a \in I$. Thus, we have:

$$1 = \frac{a}{f^k}$$

Multiplying by the denominator yields $f^k = a \in I$, which completes the proof.

33. Integral Elements

This lecture introduces the concept of integral elements, which form the foundation for studying normal rings and the algebraic geometry of singularities. We begin by defining what it means for an element to be integral over a ring.

33.1. **Definitions and Basic Properties.**

Definition 33.1. Let S be an R-algebra. An element $s \in S$ is called *integral* over the ring R if it satisfies a monic polynomial equation with coefficients in R. That is, there exist elements $r_0, r_1, \ldots, r_{n-1} \in R$ such that:

$$s^n + r_{n-1}s^{n-1} + \cdots + r_1 s + r_0 = 0$$

Example 33.2. Let us find the integral elements of the rational numbers $\mathbb{Q}$ over the integers $\mathbb{Z}$. Suppose we have an element $s = a/b \in \mathbb{Q}$, where $a, b \in \mathbb{Z}$, and we assume that a and b are coprime. If s is integral over $\mathbb{Z}$, it satisfies a monic polynomial:

$$\left(\frac{a}{b}\right)^n + r_{n-1}\left(\frac{a}{b}\right)^{n-1} + \cdots + r_1\left(\frac{a}{b}\right) + r_0 = 0$$

where all $r_i \in \mathbb{Z}$. Multiplying this entire equation by b^n, we obtain:

$$a^n + r_{n-1}a^{n-1}b + \cdots + r_1 a b^{n-1} + r_0 b^n = 0$$

If any prime p divides b, then p must divide the entire sum. Since b divides every term except possibly the first, it follows that p must also divide a^n, and therefore p divides a. However, a and b were chosen to be coprime. Thus, no such prime p can exist, which forces b to be a unit ($b = \pm 1$).

Consequently, a/b must be an integer. The integral elements of the rational numbers over the integers are exactly the integers $\mathbb{Z}$. It is reassuring that the integral elements of $\mathbb{Q}$ turn out to be the integers, as otherwise the terminology "integral" would be deeply counterintuitive.

Notice that the only property we used in this proof is that $\mathbb{Z}$ is a Unique Factorization Domain (UFD) and $\mathbb{Q}$ is its field of fractions. This easily generalizes: if R is any UFD and K is its field of fractions, the integral elements of K over R are precisely the elements of R.

33.2. **Finiteness Conditions.** We now explore the relationship between integral elements and finite modules. In commutative algebra, one must carefully distinguish between an extension being finite as an *algebra* and finite as a *module*.

Proposition 33.3. *If $s \in S$ is integral over R, then the subring $R[s]$ is finitely generated as an R-module.*

Proof. Because s satisfies a monic polynomial of degree n, we have:

$$s^n = -r_{n-1}s^{n-1} - \cdots - r_1 s - r_0$$

This demonstrates that s^n can be expressed as an R-linear combination of strictly lower powers of s. By repeatedly substituting this relation, any higher power s^m for $m \geq n$ can be reduced to a linear combination of $1, s, s^2, \ldots, s^{n-1}$. Thus, the module $R[s]$ is generated by the finite set $\{1, s, s^2, \ldots, s^{n-1}\}$. □

The converse of this proposition is also true, and it yields an extremely powerful characterization of integral elements.

Theorem 33.4. *Let S be an R-algebra. If S is finitely generated as an R-module, then all elements of S are integral over R.*

We will provide two proofs of this theorem. The first proof is straightforward but requires the assumption that R is a Noetherian ring.

First Proof (Assuming R is Noetherian). Suppose S is a finite R-module, and let $s \in S$. We consider the ascending sequence of R-submodules generated by the powers of s:

$$\langle 1 \rangle \subseteq \langle 1, s \rangle \subseteq \langle 1, s, s^2 \rangle \subseteq \ldots$$

Because S is finitely generated as a module over a Noetherian ring R, S is a Noetherian module. Therefore, this strictly increasing chain of submodules must eventually stabilize.

This implies that for some n, the element s^n must already be contained in the module generated by the previous powers $\langle 1, s, \ldots, s^{n-1} \rangle$. Thus, we can write:

$$s^n = r_0 + r_1 s + \cdots + r_{n-1}s^{n-1}$$

Rearranging this yields a monic polynomial for s, proving it is integral. □

Usually, proofs relying on the Noetherian property fail for arbitrary rings. However, this is a remarkable exception where the theorem holds for all commutative rings. To prove it in complete generality, we employ a generalization of the Cayley-Hamilton Theorem.

Theorem 33.5 (Cayley-Hamilton Theorem for Modules)**.** *Let M be a finitely generated R-module. Any R-module endomorphism $\phi\colon M \to M$ satisfies a monic polynomial equation with coefficients in R.*

Proof. Suppose M is generated by a finite set of elements $m_1, \ldots, m_n$. These do not necessarily form a basis, and M might not even be free. The endomorphism ϕ can be represented by acting on the generators:

$$\phi(m_i) = \sum_{j=1}^{n} a_{ij} m_j$$

for some coefficients $a_{ij} \in R$. Notice that this representation matrix $A = (a_{ij})$ is not necessarily unique due to potential linear dependencies among the generators.

We can rewrite this system of equations using the Kronecker delta:

$$\sum_{j=1}^{n} (\phi\delta_{ij} - a_{ij}) m_j = 0$$

Let B be the $n \times n$ matrix with entries $B_{ij} = \phi\delta_{ij} - a_{ij}$, taking values in the commutative ring $R[\phi]$. The equation above states that the matrix B annihilates the column vector $\vec{m} = (m_1, \ldots, m_n)^T$.

Recall from linear algebra that for any matrix B, multiplying by its adjugate matrix $\operatorname{adj}(B)$ yields the determinant on the diagonal:

$$\operatorname{adj}(B)B = \det(B)I$$

Multiplying our equation $B\vec{m} = 0$ by the adjugate matrix gives:

$$\operatorname{adj}(B)B\vec{m} = \det(B)I\vec{m} = 0$$

This shows that the endomorphism $\det(B)$ acts as zero on all generators m_i, and therefore $\det(B) = 0$ as an endomorphism of M. Computing the determinant $\det(\phi I - A)$ expands to a monic polynomial in ϕ of degree n, proving the theorem. □

Second Proof of Theorem 33.4 (General Case). Assume S is finitely generated as an R-module. Let $s \in S$. We define an endomorphism $\phi_s\colon S \to S$ by left multiplication by s: $\phi_s(x) = sx$.

By the Cayley-Hamilton theorem for modules, the endomorphism ϕ_s satisfies a monic polynomial equation with coefficients in R:

$$\phi_s^n + r_{n-1}\phi_s^{n-1} + \cdots + r_1\phi_s + r_0 \operatorname{id}_S = 0$$

Applying this operator identity to the element $1 \in S$, and using the fact that $\phi_s^k(1) = s^k$, we obtain:

$$s^n + r_{n-1}s^{n-1} + \cdots + r_1 s + r_0 = 0$$

Thus, s is integral over R. □

33.3. **The Ring of Integral Elements.**

Corollary 33.6. *An R-algebra S is finitely generated as an R-module if and only if S is generated as an R-algebra by a finite number of integral elements.*

Proof. If S is finitely generated as a module, all its elements are integral (by Theorem 33.4), so we can just take its finite module generators as algebra generators.

Conversely, if S is generated as an algebra by integral elements $s_1, \ldots, s_n$, we proceed by induction. We showed that adjoining one integral element $R[s_1]$ forms a finite R-module. Adjoining a second integral element $R[s_1, s_2]$ is a finite module over $R[s_1]$, and the tower property of finite modules guarantees that $R[s_1, s_2]$ is finite over R. Extending this to all n generators proves S is a finite R-module. □

Corollary 33.7. *Let S be an R-algebra. The set of all elements in S that are integral over R forms a subring of S, called the* integral closure *of R in S.*

Proof. Let $a, b \in S$ be integral over R. By the previous corollary, the subalgebra $R[a, b]$ generated by these two elements is a finite R-module. Because the sum $a + b$ and the product ab belong to this finite R-module, they must themselves be integral over R (again by Theorem 33.4). Since the set of integral elements is closed under addition and multiplication, it forms a ring. □

Example 33.8. Let $R = \mathbb{Z}$ and $S = \mathbb{C}$. The integral elements of $\mathbb{C}$ over $\mathbb{Z}$ are known as the *algebraic integers*. By the corollary above, we immediately know that the algebraic integers form a ring.

It is worth emphasizing how difficult this is to prove using strictly polynomial manipulation. If α is a root of $x^3 - 2$ and β is a root of $x^5 - 3$, showing that $\alpha + \beta$ is the root of a monic integer polynomial is extremely painful (the minimal polynomial will have degree up to 15). The module-theoretic approach completely bypasses this combinatorial nightmare.

33.4. **Normalization.**

Definition 33.9. Let R be an integral domain with field of fractions K. The *normalization* of R is the integral closure of R in K.

The ring R is called *normal* if it equals its own normalization; that is, R is integrally closed in its field of fractions.

As noted earlier, any Unique Factorization Domain is a normal ring.

Example 33.10. Let us find the normalization of the ring $\mathbb{Z}[\sqrt{5}]$. Its field of fractions is $\mathbb{Q}(\sqrt{5})$, consisting of elements $a + b\sqrt{5}$ for $a, b \in \mathbb{Q}$.

If an element $\alpha = m + n\sqrt{5}$ is integral over $\mathbb{Z}$, its Galois conjugate $\bar{\alpha} = m - n\sqrt{5}$ is also a root of the same monic polynomial with integer coefficients, and is therefore also integral. Because the algebraic integers form a ring, their sum must be integral:

$$\alpha + \bar{\alpha} = 2m \in \mathbb{Z}$$

Since $2m \in \mathbb{Q}$ and the only rational algebraic integers are the standard integers, $2m$ must be an integer in $\mathbb{Z}$. Thus, m is either an integer or a half-integer.

If m is a half-integer, we check if such an element can be integral. Consider the element:

$$\phi = \frac{1+\sqrt{5}}{2}$$

which is the golden ratio. It is a root of the polynomial:

$$x^2 - x - 1 = 0$$

Because this polynomial is monic with integer coefficients, ϕ is indeed an integral element. Therefore, the normalization of $\mathbb{Z}[\sqrt{5}]$ is strictly larger than the original ring, explicitly given by $\mathbb{Z}[\frac{1+\sqrt{5}}{2}]$. Consequently, $\mathbb{Z}[\sqrt{5}]$ is not a normal ring.

More generally, for any square-free integer $d \equiv 1 \pmod 4$, the element $\frac{1+\sqrt{d}}{2}$ is integral, as it satisfies the monic polynomial:

$$x^2 - x - \frac{d-1}{4} = 0$$

Proposition 33.11. *The operation of taking the integral closure is idempotent. That is, the integral closure of an integral closure is just the integral closure itself.*

Proof. Suppose S is the integral closure of R in K, and let $s \in K$ be integral over S. Then s satisfies a monic polynomial:

$$s^n + c_{n-1}s^{n-1} + \cdots + c_1 s + c_0 = 0$$

where the coefficients $c_i \in S$. The subring $R[c_0, \ldots, c_{n-1}]$ is generated by finitely many elements that are integral over R, so it is a finite R-module. The element s is integral over this subring, meaning $R[c_0, \ldots, c_{n-1}, s]$ is a finite module over $R[c_0, \ldots, c_{n-1}]$.

By the transitivity of finite modules, $R[c_0, \ldots, c_{n-1}, s]$ is a finite module over R. Therefore, s is contained in a finite R-module, implying s is directly integral over R. Thus, it already belongs to S. □

We conclude with a preview of the geometric meaning of being a normal ring. In algebraic geometry, the coordinate ring of a variety being normal is a mild form of non-singularity. While a normal variety can still be singular, normality guarantees that the singular locus has codimension at least 2. Thus, normal curves must be completely smooth, and normal surfaces can only have isolated point singularities.

34. Geometry of Normalizations

In this lecture, we discuss the geometric meaning of finite extensions, integral elements, and normal rings. We will investigate how these algebraic concepts correspond to geometric mappings between spectra, focusing particularly on how normalization acts as a tool for resolving singularities.

34.1. **Recap of Definitions.** Let us recall the fundamental definitions. Suppose we have two rings $R \subseteq S$.

Definition 34.1. The ring S is *finite* over R if S is finitely generated as an R-module.

Definition 34.2. An element $s \in S$ is *integral* over R if the subring $R[s]$ is a finite R-module (equivalently, s satisfies a monic polynomial with coefficients in R).

Definition 34.3. An integral domain R is *normal* (or integrally closed) if every element in its field of quotients K that is integral over R already belongs to R.

Geometrically, the inclusion map $R \hookrightarrow S$ induces a continuous map of topological spaces in the opposite direction:

$$f\colon \operatorname{Spec}(S) \to \operatorname{Spec}(R)$$

We want to describe what our algebraic conditions mean in terms of the geometry of this map f.

34.2. **Geometric Meaning of Finiteness.** First, we consider what it means for S to be a finite extension of R. A natural geometric condition for the map f to be finite is that the preimage of any point is a finite set.

Definition 34.4. The map $f\colon \operatorname{Spec}(S) \to \operatorname{Spec}(R)$ is *quasi-finite* if for every prime ideal $\mathfrak{p} \in \operatorname{Spec}(R)$, the fiber $f^{-1}(\mathfrak{p})$ is a finite set.

While related, quasi-finiteness and finiteness are not exactly the same. There are plenty of extensions that are quasi-finite but not finite.

Example 34.5. Consider the extension of the polynomial ring into the ring of Laurent polynomials:

$$R[x] \subset R[x, x^{-1}]$$

Geometrically, $\operatorname{Spec}(R[x])$ represents the affine line over R, while $\operatorname{Spec}(R[x, x^{-1}])$ represents the affine line minus the origin. The induced map is simply the inclusion of the line minus the origin into the full line. The inverse image of any point is either a single point or empty, so the map is clearly quasi-finite. However, it is certainly not a finite extension, because $R[x, x^{-1}]$ is not finitely generated as a module over $R[x]$.

Despite this distinction, finiteness is a stronger condition.

Theorem 34.6. *If S is a finite extension of R, then the induced map $f\colon \operatorname{Spec}(S) \to \operatorname{Spec}(R)$ is quasi-finite.*

Proof. We pick a prime ideal $\mathfrak{p} \in \operatorname{Spec}(R)$. We want to show that $f^{-1}(\mathfrak{p})$, which corresponds to the prime ideals of S whose intersection with R is $\mathfrak{p}$, is a finite set. We proceed by simplifying the rings:

(1) **Quotient by $\mathfrak{p}$:** We quotient both rings by $\mathfrak{p}$ (or strictly speaking, quotient S by $\mathfrak{p}S$). We may assume without loss of generality that $\mathfrak{p} = (0)$. Thus, R is now an integral domain.

(2) **Localize at** (0)**:** We invert all non-zero elements of R, effectively replacing R with its field of quotients K, and replacing S with $S \otimes_R K$.

After these reductions, we are evaluating the case where the base ring is a field K. Since S was a finite R-module, it is now a finite-dimensional vector space over K.

A commutative ring that is a finite-dimensional vector space over a field is an Artinian ring. As previously established, Artinian rings have only a finite number of prime ideals, and all such ideals are maximal. Therefore, the fiber $f^{-1}(\mathfrak{p})$ is not just finite, but discrete as a topological space. □

The subtle relationship between quasi-finiteness and finiteness is formalized by Zariski's Main Theorem, which states that under general conditions, any quasi-finite morphism factors into an open immersion followed by a finite morphism.

34.3. **Geometric Meaning of Normalization.** The geometric interpretation of normalization is a mild form of *resolution of singularities*. While it does not resolve all singularities in general, it is often the critical first step.

Example 34.7 (A curve with a cusp)**.** Let k be a field and consider the ring:

$$R = k[t^2, t^3]$$

This ring is generated as a k-algebra by $x = t^2$ and $y = t^3$. The algebraic relation between these generators is $y^2 = x^3$. Thus, R is isomorphic to the coordinate ring of a curve with a cusp singularity at the origin:

$$R \cong \frac{k[x, y]}{(y^2 - x^3)}$$

The singularity occurs at the maximal ideal $\mathfrak{m} = (x, y)$.

We now normalize R. The field of quotients of R is $K = k(t)$, the field of all rational functions in t. The element t is clearly integral over R because it satisfies the monic polynomial relation:

$$t^2 - x = 0 \quad (\text{where } x \in R)$$

Because $k[t]$ is a polynomial ring (a Unique Factorization Domain), it is automatically normal. Thus, the integral closure of R in $k(t)$ is precisely $k[t]$.

Geometrically, $\mathrm{Spec}(k[t])$ is simply the smooth affine line. The inclusion $R \hookrightarrow k[t]$ induces a map from the smooth affine line onto the singular cubic curve. Normalizing the coordinate ring algebraically resolves the singularity geometrically.

Notice that $t = y/x$, which geometrically represents the slope of the line passing through the origin. The normalization effectively separates the branches of the curve by tracking the slope, an operation directly analogous to *blowing up* the singularity.

34.4. **Normalization of Quadratic Extensions.** Let us generalize this by examining the normalization of a quadratic extension. To simplify the analysis, suppose R is a Unique Factorization Domain and that the characteristic of the

base field is not 2 (so $1/2 \in R$). Let $r \in R$ be square-free, and consider the extension:

$$S = \frac{R[t]}{(t^2 - r)} = R[\sqrt{r}]$$

We want to find the integral closure of S in its field of fractions. An arbitrary element in the quotient field takes the form $p\sqrt{r} + q$, where p and q are in the field of fractions of R.

For this element to be integral over R, both its trace and norm must be integral over R.

- The trace is $2q$. Since $1/2 \in R$, $2q$ being integral over R (and R being normal) implies $q \in R$.
- We can subtract q to find that $p\sqrt{r}$ must be integral. Squaring this yields $p^2 r \in R$.

Example 34.8 (Singular vs. Smooth Cubics)**.** Consider the coordinate ring of a general cubic curve defined by $y^2 = f(x)$.

$$S = \frac{k[x, y]}{(y^2 - f(x))}$$

Suppose $f(x)$ has a repeated root: $f(x) = (x - \alpha)^2(x - \beta)$. We can write this as:

$$y^2 = (x - \alpha)^2(x - \beta)$$

Let $r = x - \beta$. We are adjoining the square root of r multiplied by $(x - \alpha)$. Based on our condition $p^2 r \in R$, we can choose $p = 1/(x - \alpha)$. Thus, the element:

$$t = \frac{y}{x - \alpha}$$

is in the integral closure. The normalization becomes $k[t]$, which represents the smooth affine line. Resolving the singularity again corresponds to taking the slope t of a line passing through the singular point $(\alpha, 0)$.

On the other hand, if $f(x)$ has three distinct roots, $f(x) = (x-\alpha)(x-\beta)(x-\gamma)$, the curve is a non-singular elliptic curve. In this case, $f(x)$ is square-free in $k[x]$. The condition $p^2 f(x) \in k[x]$ forces $p \in k[x]$. Thus, the coordinate ring is already its own integral closure; it is a normal ring, precisely because it has no singularities to resolve.

34.5. **Normal Rings with Singularities.** It is a common misconception that normal rings are always smooth. Normalization completely resolves singularities for *curves* (1-dimensional varieties), but for higher-dimensional varieties, normal rings can still possess singularities.

Example 34.9 (The Double Cone)**.** Consider the coordinate ring of a double cone:

$$R = \frac{k[x, y, z]}{(z^2 - x^2 - y^2)}$$

The double cone clearly has a geometric singularity at the origin $(0,0,0)$. We can view R as a quadratic extension of the polynomial ring $k[x,y]$ by adjoining the element $z = \sqrt{x^2+y^2}$.

Since $k[x,y]$ is a Unique Factorization Domain, and the polynomial x^2+y^2 is not divisible by the square of any irreducible element in $k[x,y]$, the condition $p^2(x^2+y^2) \in k[x,y]$ implies that p must be a polynomial. Thus, R is integrally closed in its field of fractions.

Therefore, R is a normal ring despite possessing a singularity.

The general behavior is summarized by Serre's criterion for normality. A Noetherian ring R is normal if and only if it satisfies two conditions:

(1) **Codimension condition (R_1):** All singular points of $\mathrm{Spec}(R)$ have codimension at least 2.
(2) **Depth condition (S_2):** The depth of localizations is at least $\min(2, \dim(R_{\mathfrak{p}}))$.

For reasonable rings (specifically, Cohen-Macaulay rings), the depth condition is automatically satisfied. Thus, normality is essentially equivalent to the statement that all singularities exist in codimension 2 or higher.

For a curve (dimension 1), a singularity of codimension 2 cannot exist, so normal curves are perfectly smooth. For a surface (dimension 2), codimension 2 singularities correspond to isolated points. Thus, taking the normalization of a singular surface removes all "lines" of singularities (codimension 1), leaving behind only isolated point singularities (like the tip of the cone), which must subsequently be resolved via successive blowing up.

35. Nakayama's Lemma

This lecture introduces Nakayama's Lemma, a crucial tool in commutative algebra. We begin with motivating examples concerning the intersection of powers of a maximal ideal.

35.1. **Motivation.** Suppose R is a local ring with unique maximal ideal $\mathfrak{m}$. We are often interested in whether the intersection of all powers of $\mathfrak{m}$ vanishes:

$$\bigcap_{n=1}^{\infty} \mathfrak{m}^n = 0$$

Example 35.1 (Analytic Functions)**.** Let R be the local ring of analytic functions near the origin $0 \in \mathbb{C}$. The maximal ideal $\mathfrak{m}$ consists of functions vanishing at 0. The power $\mathfrak{m}^n$ consists of functions vanishing to order n at the origin. Thus, the intersection $\bigcap_{n=1}^{\infty} \mathfrak{m}^n$ consists of analytic functions that vanish to infinite order. By the identity theorem for analytic functions, such a function must be identically zero. Here, the intersection is indeed 0.

Example 35.2 (Smooth Functions)**.** Let R be the local ring of smooth functions near $0 \in \mathbb{R}$. Again, $\mathfrak{m}$ consists of functions vanishing at 0. However, there exist non-zero smooth functions that vanish to infinite order at the origin, such as:

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

All derivatives of f evaluate to 0 at the origin, so $f \in \mathfrak{m}^n$ for all n. Thus, $\bigcap_{n=1}^{\infty} \mathfrak{m}^n \neq 0$.

Example 35.3 (Puiseux Series)**.** Let R be the union of formal power series rings $k[[x^{1/n}]]$ for all integers $n \geq 1$. The maximal ideal $\mathfrak{m}$ consists of series with a constant term of zero. Here, any element $x^{1/n} \in \mathfrak{m}$ can be written as the square of $x^{1/2n} \in \mathfrak{m}$. Therefore, $\mathfrak{m} = \mathfrak{m}^2$, which forces:

$$\bigcap_{n=1}^{\infty} \mathfrak{m}^n = \mathfrak{m} \neq 0$$

This fails to be zero in an even more drastic way than the smooth case.

In general, there is a natural ring homomorphism from R to its *completion* $\hat{R}$, defined as the inverse limit:

$$\hat{R} = \varprojlim_n \frac{R}{\mathfrak{m}^n}$$

The kernel of this map is exactly $\bigcap_{n=1}^{\infty} \mathfrak{m}^n$. Thus, the map to the completion is injective if and only if this intersection is zero. Our ultimate aim is to show that if R is a Noetherian local ring, this intersection is always zero. This will require Nakayama's Lemma and the Artin-Rees Lemma. Here, we focus solely on Nakayama's Lemma.

35.2. **Nakayama's Lemma.**

Lemma 35.4 (Nakayama's Lemma)**.** *Let R be a local ring with maximal ideal $\mathfrak{m}$, and let M be a finitely generated R-module. If $\mathfrak{m}M = M$, then $M = 0$.*

Remark 35.5. Historically, this result is originally due to Krull, and Nakayama reportedly disliked having it named after him, likely because the proof is essentially trivial. Nevertheless, the name has stuck. The lemma also holds more generally if $\mathfrak{m}$ is replaced by the Jacobson radical of R.

Proof. Suppose M is generated by a finite set of elements $a_1, \ldots, a_n$. We choose this generating set such that n is minimal.

If $n \geq 1$, the assumption $\mathfrak{m}M = M$ allows us to express a_1 as a linear combination of the generators with coefficients in $\mathfrak{m}$:

$$a_1 = \sum_{i=1}^{n} m_i a_i \quad \text{for some } m_i \in \mathfrak{m}$$

Rearranging this equation yields:

$$(1 - m_1)a_1 = \sum_{i=2}^{n} m_i a_i$$

Because R is a local ring and $m_1 \in \mathfrak{m}$, the element $1 - m_1$ is a unit in R. We can multiply by its inverse to express a_1 entirely in terms of $a_2, \ldots, a_n$:

$$a_1 = \sum_{i=2}^{n} (1 - m_1)^{-1} m_i a_i$$

This shows that a_1 can be dropped from the generating set, meaning M is generated by $n-1$ elements. This contradicts the minimality of n. Thus, we must have $n = 0$, which implies $M = 0$. □

Corollary 35.6. *Let R be a local ring with maximal ideal $\mathfrak{m}$, and let M be a finitely generated R-module. If elements $m_1, \ldots, m_n \in M$ generate the quotient module $M/\mathfrak{m}M$ over the field $R/\mathfrak{m}$, then they generate M as an R-module.*

Proof. Let N be the submodule of M generated by $m_1, \ldots, m_n$. We consider the quotient module M/N. Because M is finitely generated, M/N is also finitely generated.

By hypothesis, the images of $m_1, \ldots, m_n$ span $M/\mathfrak{m}M$. This implies that:

$$M = N + \mathfrak{m}M$$

Quotienting both sides by N, we obtain:

$$\frac{M}{N} = \mathfrak{m}\left(\frac{M}{N}\right)$$

Applying Nakayama's Lemma to the finitely generated module M/N, we conclude that $M/N = 0$, which means $M = N$. Thus, $m_1, \ldots, m_n$ generate the whole module M. □

Remark 35.7. A common mistake is omitting the hypothesis that M is *finitely generated*. If M is not finitely generated, the corollary fails. For example, let $R = \mathbb{Z}_{(2)}$, which is a local ring with maximal ideal $\mathfrak{m} = 2\mathbb{Z}_{(2)}$. Let $M = \mathbb{Q}$.

Since $\mathbb{Q}$ is divisible, $\mathfrak{m}M = 2\mathbb{Q} = \mathbb{Q}$. Thus, the quotient $M/\mathfrak{m}M = 0$. The empty set trivially generates $M/\mathfrak{m}M$, but it clearly does not generate $M = \mathbb{Q}$.

35.3. **Application: Integral Extensions and Surjectivity.** As an application of Nakayama's Lemma, we prove a geometric property of integral extensions.

Theorem 35.8. *Let $R \subseteq S$ be commutative rings such that S is integral over R. Then the induced map on spectra is surjective:*

$$\mathrm{Spec}(S) \twoheadrightarrow \mathrm{Spec}(R)$$

Proof. Let $\mathfrak{p} \in \mathrm{Spec}(R)$. We wish to find a prime ideal $\mathfrak{q} \in \mathrm{Spec}(S)$ such that $\mathfrak{q} \cap R = \mathfrak{p}$.

First, we localize both R and S at the prime $\mathfrak{p}$. This means we invert all elements in the multiplicative set $R \setminus \mathfrak{p}$. Localizing commutes with integral extensions, so $S_\mathfrak{p}$ remains integral over $R_\mathfrak{p}$. In the local ring $R_\mathfrak{p}$, the ideal generated by $\mathfrak{p}$ is the unique maximal ideal. By restricting to this localized case, we may assume without loss of generality that R is a local ring and $\mathfrak{p}$ is its maximal ideal.

We claim that the extended ideal $\mathfrak{p}S$ is strictly smaller than S. Suppose, for the sake of contradiction, that $\mathfrak{p}S = S$. Then we can express the identity element as a finite sum:

$$1 = \sum_{i=1}^{n} p_i s_i \quad \text{for } p_i \in \mathfrak{p}, s_i \in S$$

Let $M = R[s_1, \ldots, s_n]$ be the R-subalgebra of S generated by these specific elements. Because S is integral over R, each s_i is integral over R. Since M is generated by a finite number of integral elements, it is finitely generated as an R-module.

By our construction, $1 \in \mathfrak{p}M$. Because 1 generates M over R, this implies:

$$M = \mathfrak{p}M$$

Since M is a finitely generated R-module and $\mathfrak{p}$ is the maximal ideal of the local ring R, Nakayama's Lemma forces $M = 0$. However, $1 \in M$, so $1 = 0$, which is a contradiction.

Therefore, $\mathfrak{p}S \neq S$. Because $\mathfrak{p}S$ is a proper ideal of S, it is contained in some maximal ideal $\mathfrak{q}$ of S.

We must now verify that $\mathfrak{q} \cap R = \mathfrak{p}$. The intersection $\mathfrak{q} \cap R$ is an ideal in R that contains $\mathfrak{p}$ (since $\mathfrak{p} \subseteq \mathfrak{p}S \subseteq \mathfrak{q}$). Furthermore, $1 \notin \mathfrak{q}$, so $1 \notin \mathfrak{q} \cap R$, meaning it is a proper ideal. Because $\mathfrak{p}$ is a maximal ideal in R, we must have:

$$\mathfrak{q} \cap R = \mathfrak{p}$$

Thus, $\mathfrak{p}$ is the image of $\mathfrak{q}$ under the spectral map, proving surjectivity. □

36. The Artin-Rees Lemma

This lecture focuses on stating and proving the Artin-Rees Lemma. We will then apply it to prove Krull's Intersection Theorem, which asserts that the intersection of all powers of the maximal ideal in a Noetherian local ring is zero. Before stating the lemma, we must define filtrations and their associated topologies.

36.1. **Filtrations of Modules.**

Definition 36.1. A *decreasing filtration* on an R-module M is a descending sequence of submodules:

$$M = M_0 \supseteq M_1 \supseteq M_2 \supseteq \ldots$$

Definition 36.2. Let I be an ideal of R. The *I-adic filtration* on M is defined by setting:

$$M_n = I^n M$$

The term "I-adic" generalizes the classic p-adic filtration on the integers. If we take $R = \mathbb{Z}$, $M = \mathbb{Z}$, and $I = (p)$ for a prime p, the I-adic filtration is given by $M_n = p^n\mathbb{Z}$. This is the standard filtration used to construct the p-adic numbers (which are the completion of $\mathbb{Z}$ with respect to this filtration).

36.1.1. *Topologies Induced by Filtrations.* Any filtration on a module M naturally induces a topology on M. The topology is defined by declaring the submodules M_n to form a basis of open neighborhoods of $0 \in M$. By translation, a basis for the open neighborhoods around any element $x \in M$ is given by the sets:

$$x + M_n$$

A critical subtlety arises when relating submodules and topologies. Suppose M is a submodule of N. We can equip N with the I-adic topology (using I^nN). We can also equip M directly with its own I-adic topology (using I^nM).

If we restrict the I-adic topology of N to the subspace M, the basic open sets are of the form:

$$M \cap I^nN$$

We must ask: Does the subspace topology induced by N coincide with the intrinsic I-adic topology on M? In general, the answer is no.

Example 36.3. Let $R = \mathbb{Z}$, $N = \mathbb{Q}$, and $M = \mathbb{Z}$. Let $I = (2)$. Consider the 2-adic topology on $N = \mathbb{Q}$. Because multiplication by 2 is an automorphism of $\mathbb{Q}$, we have:

$$2^n\mathbb{Q} = \mathbb{Q}$$

Therefore, the only open sets in this topology are the empty set and $\mathbb{Q}$ itself, making it the *indiscrete topology*. Consequently, the subspace topology restricted to $\mathbb{Z}$ is also indiscrete.

On the other hand, the intrinsic 2-adic topology on $M = \mathbb{Z}$ has a neighborhood basis of zero given by $2^n\mathbb{Z}$. This topology is Hausdorff (since $\bigcap 2^n\mathbb{Z} = 0$), so it is vastly different from the indiscrete topology.

Even when the topologies coincide, the underlying filtrations might technically differ.

Example 36.4. Let $R = \mathbb{Z}$, $N = \mathbb{Z}$, and $M = 4\mathbb{Z}$, with $I = (2)$. The 2-adic filtration on N is:

$$\mathbb{Z} \supseteq 2\mathbb{Z} \supseteq 4\mathbb{Z} \supseteq 8\mathbb{Z} \supseteq \dots$$

Restricting these sets to $M = 4\mathbb{Z}$ via intersections yields the sequence:

$$4\mathbb{Z},\ 4\mathbb{Z},\ 4\mathbb{Z},\ 8\mathbb{Z},\ 16\mathbb{Z},\ \dots$$

Conversely, the intrinsic 2-adic filtration directly on M (I^nM) yields:

$$4\mathbb{Z},\ 8\mathbb{Z},\ 16\mathbb{Z},\ 32\mathbb{Z},\ \dots$$

These are two different filtrations, but because every term in one filtration eventually appears in the other (shifted by a finite index), they generate the exact same topology.

36.2. **Stable Filtrations.** To formalize this shifting behavior, we define the concept of a stable filtration.

Definition 36.5. An I-filtration (M_n) on M is called *stable* if:

$$IM_n \subseteq M_{n+1} \quad \text{for all } n$$
$$IM_n = M_{n+1} \quad \text{for all sufficiently large } n$$

The I-adic filtration $M_n = I^nM$ is clearly a stable filtration. Furthermore, it is the *smallest* stable filtration since the condition $IM_n \subseteq M_{n+1}$ forces $I^nM \subseteq M_n$.

Lemma 36.6. *Any two stable I-filtrations on M determine the same topology.*

Proof. Let (M_n) be a stable filtration. Since $I^n M \subseteq M_n$, the (M_n) topology is bounded by the I-adic topology. Conversely, because the filtration is stable, there exists an integer N such that $M_{N+k} = I^k M_N$ for all $k \geq 0$. This implies:

$$M_{N+k} = I^k M_N \subseteq I^k M$$

Thus, every neighborhood in the I-adic topology contains a neighborhood from the (M_n) filtration, proving the topologies are identical. □

36.3. **The Artin-Rees Lemma.** We can now state the Artin-Rees Lemma, which provides precise control over how filtrations behave on submodules.

Theorem 36.7 (Artin-Rees Lemma)**.** *Let R be a Noetherian ring, I an ideal of R, and $M \subseteq N$ be finitely generated R-modules.*

- ***Strong Form:*** *Any stable I-filtration on N restricts to a stable I-filtration on M.*
- ***Weak Form:*** *The I-adic topology on N restricts to the I-adic topology on M.*

The weak form follows immediately from the strong form, since restricting the I-adic filtration on N yields a stable filtration on M, which by our lemma induces the same topology as the intrinsic I-adic filtration on M.

To prove the strong form, we translate the algebraic property of stability into a finite generation property over a graded ring.

Definition 36.8. The *blow-up algebra* (or Rees algebra) of I is the graded ring:

$$\tilde{R} = R \oplus I \oplus I^2 \oplus I^3 \oplus \dots$$

Given an I-filtration (M_n) on an R-module M, we define an associated graded module:

$$\tilde{M} = M_0 \oplus M_1 \oplus M_2 \oplus M_3 \oplus \dots$$

Lemma 36.9. *The filtration (M_n) is stable if and only if $\tilde{M}$ is a finitely generated $\tilde{R}$-module.*

Proof. If (M_n) is stable, there exists some integer N such that for all $k > 0$, $M_{N+k} = I^k M_N$. Thus, the graded module $\tilde{M}$ is generated over $\tilde{R}$ entirely by the elements residing in the components $M_0 \oplus M_1 \oplus \dots \oplus M_N$. Since M is finitely generated and R is Noetherian, each M_i is finitely generated over R. Taking the union of these finite generating sets provides a finite set of generators for $\tilde{M}$ over $\tilde{R}$.

Conversely, if $\tilde{M}$ is finitely generated over $\tilde{R}$, we can choose a finite set of homogeneous generators. Let N be the maximum degree of these generators. Any element in M_{N+k} must be formed by multiplying generators of degree $\leq N$ by elements in $\tilde{R}$ of appropriate degree (which belong to I^k). Thus, $M_{N+k} = I^k M_N$, meaning the filtration is stable. □

Proof of the Artin-Rees Lemma. Let (N_n) be a stable I-filtration on N. By Lemma 36.9, the associated graded module $\tilde{N} = \bigoplus N_n$ is a finitely generated module over the blow-up algebra $\tilde{R}$.

Because R is a Noetherian ring, and $\tilde{R}$ is generated over R by a finite generating set of the ideal I, the blow-up algebra $\tilde{R}$ is a finitely generated R-algebra, and is therefore a Noetherian ring.

The restricted filtration on M is given by $M_n = M \cap N_n$. This naturally defines a graded submodule $\tilde{M} = \bigoplus M_n \subseteq \tilde{N}$.

Since $\tilde{N}$ is a finitely generated module over the Noetherian ring $\tilde{R}$, any submodule must also be finitely generated. Thus, $\tilde{M}$ is a finitely generated $\tilde{R}$-module. By Lemma 36.9, this immediately implies that the filtration (M_n) is stable. □

36.4. **Krull's Intersection Theorem.** We conclude by applying the Artin-Rees Lemma to prove a fundamental result in local commutative algebra.

Theorem 36.10 (Krull's Intersection Theorem)**.** *Let R be a Noetherian local ring with maximal ideal $\mathfrak{m}$. Then the intersection of all powers of $\mathfrak{m}$ is zero:*

$$\bigcap_{n=1}^{\infty} \mathfrak{m}^n = 0$$

Proof. Let $J = \bigcap_{n=1}^{\infty} \mathfrak{m}^n$. We wish to show $J = 0$.

Consider J as a submodule of R. Because J is the intersection of all powers of $\mathfrak{m}$, restricting the $\mathfrak{m}$-adic filtration of R to J yields:

$$J \cap \mathfrak{m}^n = J \quad \text{for all } n$$

Topologically, this means the restricted $\mathfrak{m}$-adic topology on J is the indiscrete topology (the only neighborhoods of zero are J itself).

By the weak form of the Artin-Rees Lemma, this restricted topology must coincide with the intrinsic $\mathfrak{m}$-adic topology on J. The basis for the intrinsic topology is $\mathfrak{m}^n J$. For the intrinsic topology to be indiscrete, we must have:

$$\mathfrak{m}J = J$$

Because R is Noetherian, the ideal J is a finitely generated R-module. We have a finitely generated module J over a local ring satisfying $\mathfrak{m}J = J$. By Nakayama's Lemma, this implies:

$$J = 0$$

This completes the proof. □

37. Blowup Algebras

This lecture surveys various algebras constructed from a filtration. While the title is "Blowup Algebras", we will discuss several closely related algebraic constructions derived from both decreasing and increasing filtrations.

37.1. **Filtrations.**

Definition 37.1. A *filtration* on a ring R is a sequence of additive subgroups R_i such that $R_i \cdot R_j \subseteq R_{i+j}$.

There are two primary types of filtrations we will consider:

- **Decreasing filtrations:** The subsets decrease, $R_0 \supseteq R_1 \supseteq R_2 \supseteq \dots$. The typical example is setting $R_n = I^n$ for some ideal $I \subseteq R$.
- **Increasing filtrations:** The subsets increase, $R_0 \subseteq R_1 \subseteq R_2 \subseteq \dots$. A standard example is fixing a subring R_0 and taking R_i to be the module spanned by products of at most i elements from a fixed set of generators $\{a_1, \dots, a_k\}$. In this case, R_i is the set of polynomials or monomials of degree at most i.

From these filtrations, we can construct three main types of algebras: a blowup algebra, an associated graded algebra, and a completion (an inverse limit).

37.2. **Algebras from Decreasing Filtrations.** Suppose we have a decreasing filtration, typically $R_n = I^n$ for some ideal I.

37.2.1. *The Blowup Algebra.*

Definition 37.2. The *blowup algebra* (often called the Rees algebra) associated with an ideal I is the graded algebra defined by the direct sum:

$$\mathrm{Bl}_I(R) = R \oplus I \oplus I^2 \oplus I^3 \oplus \dots$$

This algebra is naturally graded, where the degree n piece is I^n. We encountered a module-theoretic version of this algebra when proving the Artin-Rees Lemma (where it was used to show that stable filtrations yield finitely generated modules).

In algebraic geometry, this construction corresponds to a *blow-up*. Suppose W is a subvariety of a variety V. If R is the coordinate ring of V and I is the ideal defining W, the geometric blow-up of V along W replaces each point of W with the projective space of its normal bundle. Algebraically, this space is recovered by applying the Proj construction to the graded blowup algebra $R \oplus I \oplus I^2 \oplus \dots$. For instance, blowing up the origin in an affine plane replaces the origin with a full copy of $\mathbb{P}^1$.

37.2.2. *The Associated Graded Algebra.*

Definition 37.3. The *associated graded algebra* corresponding to a decreasing filtration $R_n = I^n$ is:

$$\mathrm{gr}_I(R) = \frac{R}{I} \oplus \frac{I}{I^2} \oplus \frac{I^2}{I^3} \oplus \dots$$

This construction turns the filtered ring R into a graded ring. In a sense, $\mathrm{gr}_I(R)$ is "about the same size" as R. Let us look at a few examples:

Example 37.4. Let R be a local ring with maximal ideal $\mathfrak{m}$, and let $k = R/\mathfrak{m}$ be its residue field. The associated graded algebra is:

$$\mathrm{gr}_{\mathfrak{m}}(R) = k \oplus \frac{\mathfrak{m}}{\mathfrak{m}^2} \oplus \frac{\mathfrak{m}^2}{\mathfrak{m}^3} \oplus \dots$$

Example 37.5. Let $R = k[x, y]$ localized at the origin (so we localize with respect to the ideal $\mathfrak{m} = (x, y)$). If we take the associated graded ring of this local ring with respect to its maximal ideal, we perfectly recover the original polynomial

ring $k[x, y]$. While not strictly an inverse to localization, it behaves similarly by recovering the global graded structure.

Example 37.6. Consider the localization of the integers at the prime 2, $R = \mathbb{Z}_{(2)}$, which consists of fractions a/b with b odd. Let $I = (2)$. The graded ring is:

$$\mathrm{gr}_{(2)}(R) = \frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{2\mathbb{Z}}{4\mathbb{Z}} \oplus \frac{4\mathbb{Z}}{8\mathbb{Z}} \oplus \dots$$

This is isomorphic to the polynomial ring $\mathbb{F}_2[x]$. We have transformed a local ring into a polynomial ring over the field with two elements.

37.2.3. *The Extended Rees Algebra.* To relate the ring R precisely to its associated graded algebra, we can use the *extended Rees algebra*.

Definition 37.7. The *extended Rees algebra* extends the standard blowup algebra to the left with copies of R:

$$\dots \oplus R \oplus R \oplus R \oplus I \oplus I^2 \oplus I^3 \oplus \dots$$

where the degree of the pieces are $\dots, -2, -1, 0, 1, 2, 3, \dots$.

If R is an algebra over a field k, the extended Rees algebra naturally forms an algebra over the polynomial ring $k[t]$, where t has degree -1.

Geometrically, this represents a flat deformation over the affine line. The fiber over any non-zero point of the affine line is isomorphic to the original ring R, while the special fiber over the origin ($t = 0$) is precisely the associated graded algebra $\mathrm{gr}_I(R)$. This demonstrates that the graded algebra is a flat deformation of R.

37.2.4. *Completion.* Finally, from a decreasing filtration I^n, we can form the completion:

$$\hat{R} = \varprojlim_n \frac{R}{I^n}$$

For instance, completing $k[x]$ with respect to the ideal (x) yields the formal power series ring $k[[x]]$. We will study completions in far more detail in subsequent sections.

37.3. **Algebras from Increasing Filtrations.** Now let us consider an increasing filtration $R_0 \subseteq R_1 \subseteq R_2 \subseteq \dots$.

37.3.1. *The Blowup Analog.* We can form a direct analogue to the blowup algebra:

$$\tilde{R} = R_0 \oplus R_1 \oplus R_2 \oplus \dots$$

Example 37.8. Let $R = k[x, y]$ and define R_n as the span of monomials of total degree $\leq n$. The resulting algebra $R_0 \oplus R_1 \oplus R_2 \oplus \dots$ turns out to be isomorphic to a polynomial ring in three variables $k[x, y, t]$, where t acts as an extra variable indicating the degree shifts (essentially corresponding to the element $1 \in R_1$). The grading matches perfectly. This is a somewhat convoluted way of defining a standard polynomial ring.

37.3.2. *Associated Graded Algebra for Non-Commutative Rings.* The graded algebra for an increasing filtration is given by the successive quotients:

$$\mathrm{gr}(R) = R_0 \oplus \frac{R_1}{R_0} \oplus \frac{R_2}{R_1} \oplus \dots$$

This construction is particularly powerful for turning certain non-commutative rings into commutative ones, allowing us to apply the tools of commutative algebra to them.

Example 37.9. Let R be the ring of differential operators (the Weyl algebra) over a field k of characteristic zero. It is generated as a k-algebra by the position variables $x_1, \dots, x_n$ and the differentiation operators $\partial_1, \dots, \partial_n$ (which we write as $d_1, \dots, d_n$).

This ring is non-commutative due to the Leibniz rule:

$$\begin{aligned} d_i x_j &= x_j d_i \quad \text{for } i \neq j \\ d_i x_i &= x_i d_i + 1 \end{aligned}$$

Notice that the obstruction to commutativity, the commutator $[d_i, x_i] = 1$, is structurally simpler (lower degree) than the elements themselves.

We define an increasing filtration where R_i is the vector space spanned by products of at most i generators from $\{x_1, \dots, x_n, d_1, \dots, d_n\}$. By induction on the Leibniz rule, one can show that for $a \in R_i$ and $b \in R_j$:

$$\begin{aligned} ab &\in R_{i+j} \\ ba &\in R_{i+j} \\ ab - ba &\in R_{i+j-1} \end{aligned}$$

Because the commutator $ab - ba$ drops in degree, it lies in R_{i+j-1}. When we form the associated graded algebra, we quotient by these lower-degree terms:

$$\mathrm{gr}(R) = R_0 \oplus \frac{R_1}{R_0} \oplus \frac{R_2}{R_1} \oplus \dots$$

In this quotient, the commutators vanish. Consequently, $\mathrm{gr}(R)$ is commutative. Specifically, it is isomorphic to a polynomial algebra in $2n$ variables:

$$\mathrm{gr}(R) \cong k[\bar{x}_1, \dots, \bar{x}_n, \bar{d}_1, \dots, \bar{d}_n]$$

where the bars denote the images of the generators in the graded pieces.

We can use this commutativity to easily prove that the non-commutative ring of differential operators R has no zero divisors.

Suppose $a, b \in R$ are non-zero elements. They have well-defined highest-degree parts, which correspond to non-zero images $\bar{a}$ and $\bar{b}$ in the graded ring $\mathrm{gr}(R)$. Because $\mathrm{gr}(R)$ is simply a polynomial ring in $2n$ variables over a field, it is an integral domain. Thus, the product $\bar{a}\bar{b} \neq 0$ in $\mathrm{gr}(R)$.

However, the top-degree part of the product ab in R is exactly equal to the product of the top-degree parts $\bar{a}\bar{b}$. Since this top-degree part is non-zero, the full product ab must be non-zero. Therefore, R has no zero divisors.

A nearly identical argument shows that the universal enveloping algebra of any Lie algebra also has no zero divisors, as its associated graded ring is similarly a commutative polynomial ring.

Part 5. Module Properties and Categories

38. Survey of Module Properties

This lecture provides a brief survey of different properties of modules, acting as a "cheat sheet" for concepts that will be explored in detail in subsequent lectures. We divide these properties into two main categories: finiteness conditions, and conditions that weaken the notion of a free module.

38.1. Finiteness Conditions.

Definition 38.1. The most common finiteness conditions for an R-module M are as follows:

- *Finitely generated*: There exists a surjective module homomorphism from a finite free module onto M:
$$R^n \twoheadrightarrow M$$
- *Finitely presented*: There exists an exact sequence of the form:
$$R^m \to R^n \to M \to 0$$
This means M is finitely generated, and the module of relations between the generators is also finitely generated.
- *Coherent*: M is finitely generated, and for any homomorphism from a finite free module $R^n \to M$, the kernel is finitely generated. (Note: the map $R^n \to M$ need not be surjective).
- *Finite length*: M has a finite composition series. This is equivalent to M being both Noetherian and Artinian.

Remark 38.2. Coherent modules do not appear frequently in basic commutative algebra, but they are crucial in sheaf theory over complex analytic manifolds, where the rings encountered are generally not Noetherian. In such settings, coherence provides a workable substitute for the Noetherian property.

Proposition 38.3. *Over a Noetherian ring R, the conditions of being finitely generated, finitely presented, and coherent are all equivalent.*

Because we frequently work over Noetherian rings, being finitely generated is the most common and useful finiteness condition.

38.2. Weakening the Free Condition.

Free modules are extremely well-behaved but relatively rare. Therefore, we study various properties that weaken the condition of being a free module.

Definition 38.4. Let M be an R-module. We define the following hierarchy of properties:

- *Free*: M is isomorphic to a direct sum of copies of the base ring:
$$M \cong \bigoplus R$$
- *Stably free*: There exists a finite integer n such that adding a finite free module yields a free module:
$$M \oplus R^n \cong R^m$$
- *Locally free*: There exist elements $f_1, \ldots, f_n \in R$ generating the unit ideal, meaning $(f_1, \ldots, f_n) = R$, such that the localization M_{f_i} is a free module over R_{f_i} for each i. Geometrically, this means the Zariski topology of $\mathrm{Spec}(R)$ is covered by a finite number of open sets over which M is free. These correspond to vector bundles.
- *Projective*: For any surjective module homomorphism $A \twoheadrightarrow B$ and any map $M \to B$, the map can be lifted to A:
$$\mathrm{Hom}_R(M, A) \twoheadrightarrow \mathrm{Hom}_R(M, B)$$
- *Stalks locally free*: The localization $M_{\mathfrak{p}}$ is a free $R_{\mathfrak{p}}$-module for all prime ideals $\mathfrak{p} \in \mathrm{Spec}(R)$.
- *Flat*: The functor $- \otimes_R M$ is exact. Equivalently, if $A \hookrightarrow B$ is an injective module homomorphism, then the induced map is also injective:
$$A \otimes_R M \hookrightarrow B \otimes_R M$$
- *Torsion-free*: If $a \in R$ is a regular element (not a zero divisor) and $m \in M$, then:
$$am = 0 \implies m = 0$$

38.3. **Relations Between the Properties.** The relationships between these properties depend heavily on the finiteness conditions imposed on the module.

Theorem 38.5. *For* finitely presented *modules, the following four conditions are equivalent:*

$$\textit{Locally free} \iff \textit{Projective} \iff \textit{Stalks locally free} \iff \textit{Flat}$$

These modules act as the algebraic analogue of finite-dimensional vector bundles.

If we weaken the finiteness condition from finitely presented to merely finitely generated, the perfect equivalence breaks down into a chain of specific implications.

Proposition 38.6. *We have the following implications under varying finiteness conditions:*

(1) Flat implies stalks locally free (for finitely generated *modules).*
(2) Stalks locally free implies projective (requires finitely presented *modules).*
(3) Projective implies locally free (for finitely generated *modules).*

Remark 38.7. Over a local ring or a Principal Ideal Domain (PID), all of the conditions from locally free down to torsion-free generally coincide.

38.4. **The Importance of Flatness.** One might ask: which of these properties is the most important? If one were sent to a desert island for a year and only allowed to take one module property, the clear choice is *flatness*.

This is historically and conceptually somewhat surprising. Flatness is the most technical and least intuitive property on the list; indeed, researchers studied commutative algebra for several decades before flatness was even formally defined.

However, for a mathematical property to be truly useful, it must satisfy two competing criteria:

(1) **Commonality:** There must be an abundance of modules possessing the property. (A property possessed by almost no modules is useless).
(2) **Structure:** The property must enforce good, well-behaved structure on the modules that possess it.

Torsion-free modules are extremely common, but they do not behave particularly nicely. On the other extreme, free and projective modules are incredibly well-behaved, but they are relatively rare.

Flat modules sit exactly in the "sweet spot": they are exceptionally well-behaved, yet they are still remarkably common. This optimal balance makes flatness a critically useful property in modern commutative algebra and algebraic geometry.

39. Stably Free Modules

In the previous lectures, we discussed various properties that weaken the condition of being a free module. We now investigate *stably free* modules, which are structurally the closest to free modules without necessarily being free themselves.

39.1. **Definition and the Tangent Space of S^2.**

Definition 39.1. Let R be a commutative ring. A finitely generated R-module M is *stably free* if there exists a finite integer n such that:

$$M \oplus R^n \cong R^m$$

for some integer m. The key point here is that n must be finite.

Trivially, any free module is stably free (by taking $n = 0$). We must first construct an example of a stably free module that is not free.

Example 39.2. Let S^2 be the two-dimensional sphere. We define our ring and module topologically:

- Let R be the ring of continuous real-valued functions on S^2.
- Let M be the space of continuous tangent vector fields on S^2.

M is an R-module under pointwise multiplication. Geometrically, M corresponds to the tangent bundle of S^2, and R corresponds to the normal bundle of S^2 embedded in $\mathbb{R}^3$. The direct sum of the tangent bundle and the normal bundle yields the trivial tangent bundle of the ambient space $\mathbb{R}^3$:

$$M \oplus R \cong R^3$$

This establishes that M is stably free. However, by the Hairy Ball Theorem, every continuous tangent vector field on S^2 must vanish at some point. Consequently, the tangent bundle cannot be trivial, which implies M is not a free R-module.

We can also express this example purely algebraically. Let R be the coordinate ring of the sphere S^2:

$$R = \frac{\mathbb{R}[x, y, z]}{(x^2 + y^2 + z^2 - 1)}$$

We define M as the submodule of R^3 representing tangent vectors:

$$M = \left\{ (a, b, c) \in R^3 \;\middle|\; ax + by + cz = 0 \right\}$$

We can establish the decomposition $R^3 \cong M \oplus R$ explicitly via the map:

$$(a, b, c) \mapsto (a - dx, b - dy, c - dz) \oplus d$$

where $d = ax + by + cz$. This confirms $M \oplus R \cong R^3$, meaning M is stably free.

39.2. **Vector Fields on Spheres.** This topological construction generalizes to the $(n-1)$-dimensional sphere S^{n-1} embedded in $\mathbb{R}^n$. The module M of continuous vector fields on S^{n-1} over the ring R of continuous functions satisfies:

$$M \oplus R \cong R^n$$

We can ask: when is M actually a free module? This corresponds to the topological question of when S^{n-1} possesses a trivial tangent bundle. It is free only when $n = 1, 2, 4$, or 8 (corresponding to S^0, S^1, S^3, and S^7).

For S^0, S^1, and S^3, these correspond to the elements of norm 1 in $\mathbb{R}, \mathbb{C}$, and $\mathbb{H}$ (quaternions), which naturally form groups. The tangent bundle of any Lie group is trivial via left multiplication by group elements. S^7 corresponds to the unit octonions (or Cayley numbers); while non-associative, the multiplication still admits inverses and yields a trivial tangent bundle.

For a long time, it was an open question whether any other spheres admitted trivial tangent bundles—the famous "vector fields on spheres" problem. J.F. Adams finally solved this, proving that no other dimensions work. Thus, for all other n, M provides an example of a stably free module that is not free.

39.3. **Serre's Conjecture.** Stably free modules also arise crucially in Serre's Conjecture, which was eventually proven by Quillen and Suslin.

Theorem 39.3 (Serre's Conjecture / Quillen-Suslin Theorem)**.** *Every finitely generated projective module over a polynomial ring $k[x_1, \ldots, x_n]$ over a field is free.*

Serre's motivation stemmed from algebraic topology: affine space corresponds to a real vector space, over which all finite-dimensional vector bundles are topologically trivial.

We can easily show that such a projective module is stably free. By Hilbert's Syzygy Theorem, any finitely generated projective module P_0 over a polynomial ring has a finite free resolution:

$$0 \to F_n \to \cdots \to F_1 \to F_0 \to P_0 \to 0$$

Because P_0 is projective, the surjection from F_0 splits:

$$F_0 \cong P_1 \oplus P_0$$

Iterating this splitting backward through the finite resolution yields:

$$P_0 \oplus F_1 \oplus F_3 \oplus \ldots \cong F_0 \oplus F_2 \oplus F_4 \oplus \ldots$$

Since the sum of finite free modules is a finite free module, P_0 added to a free module is free. Thus, P_0 is stably free.

Serre's Conjecture effectively reduces to demonstrating that stably free modules over $k[x_1, \ldots, x_n]$ are free. This can be phrased elementarily: if $f_1, \ldots, f_m$ are polynomials that generate the unit ideal, can the vector $(f_1, \ldots, f_m)$ be extended to an invertible $m \times m$ matrix over the polynomial ring? The Quillen-Suslin theorem proves that it always can.

39.4. **Stably Free Modules of Small Rank.** We can analyze stably free modules M over an arbitrary ring R by investigating their rank. Let M be a stably free module of rank r, satisfying:

$$M \oplus R^n \cong R^{n+r}$$

39.4.1. *Rank 0.* If $r = 0$, then $M \oplus R^n \cong R^n$. Taking the n-th exterior power of both sides yields:

$$\Lambda^n(R^n) \cong R$$
$$\Lambda^n(M \oplus R^n) \cong \bigoplus_{i=0}^{n} \left(\Lambda^i(M) \otimes \Lambda^{n-i}(R^n)\right)$$

Expanding the right-hand side, the leading terms are:

$$\begin{aligned}\Lambda^n(M \oplus R^n) &\cong \left(\Lambda^0(M) \otimes \Lambda^n(R^n)\right) \oplus \left(\Lambda^1(M) \otimes \Lambda^{n-1}(R^n)\right) \oplus \ldots \\ &\cong R \oplus \left(M \otimes R^{\binom{n}{n-1}}\right) \oplus \ldots\end{aligned}$$

Thus, we obtain an isomorphism $R \cong R \oplus M^{\oplus n} \oplus \ldots$. This algebraic relation forces $M = 0$.

39.4.2. *Rank 1.* If $r = 1$, then $M \oplus R^n \cong R^{n+1}$. We take the $(n+1)$-th exterior power of both sides:

$$\Lambda^{n+1}(R^{n+1}) \cong R$$

On the other side, because $\Lambda^{n+1}(R^n) = 0$, the expansion simplifies to its highest non-vanishing term:

$$\begin{aligned}\Lambda^{n+1}(M \oplus R^n) &\cong \big(\Lambda^1(M) \otimes \Lambda^n(R^n)\big) \oplus \big(\Lambda^2(M) \otimes \Lambda^{n-1}(R^n)\big) \oplus \dots \\ &\cong M \otimes R \oplus \dots \\ &\cong M \oplus \dots\end{aligned}$$

Equating the two results, we find that $M \cong R$.

Therefore, stably free modules of rank 0 and 1 are necessarily free. As we observed with the tangent space of S^2, stably free modules of rank 2 and higher are considerably more complicated and can fail to be free.

40. The Eilenberg-Mazur Swindle

This section introduces a technique known as the *Eilenberg-Mazur swindle* (or simply the Mazur swindle in topology) and explores its applications to the structure of free and stably free modules.

40.1. **The Algebraic Mechanism.** To illustrate the core idea behind the swindle, consider the classic fallacious proof that $1 = 0$. By examining the infinite series $1 - 1 + 1 - 1 + \dots$, we can bracket the terms in two different ways:

$$(1-1) + (1-1) + (1-1) + \dots = 0 + 0 + 0 + \dots = 0$$
$$1 + (-1+1) + (-1+1) + \dots = 1 + 0 + 0 + \dots = 1$$

This implies $1 = 0$. The obvious flaw in this argument is that infinite sums of integers are not well-defined. However, the Eilenberg-Mazur swindle applies exactly this logic to contexts where infinite sums *are* well-defined, associative, and commutative.

Proposition 40.1. *Suppose an operation permits well-defined, associative, and commutative infinite sums. If $A + B = 0$ (where 0 acts as the identity), then $A = 0$ and $B = 0$.*

Proof. Consider the infinite sum $A + B + A + B + \dots$. We can bracket the terms in two ways:

$$(A+B) + (A+B) + \dots = 0 + 0 + \dots = 0$$
$$A + (B+A) + (B+A) + \dots = A + 0 + 0 + \dots = A$$

Equating the two bracketings yields $A = 0$. By symmetry, $B = 0$. □

40.2. **Topological Example: Knots.** Before applying this to commutative algebra, we consider an example from topology introduced by Barry Mazur. We can define the sum of two knots A and B by cutting each and joining the loose ends. This operation is commutative and, in a natural sense, associative.

Furthermore, infinite sums of knots are well-defined for *continuous* knots (by scaling each subsequent knot to be half the size of the previous one and connecting them in an infinite sequence converging to a limit point). Note that this construction yields a continuous, but not smooth, limit point, which highlights a fundamental difference between topological and smooth manifolds.

Theorem 40.2 (Mazur Swindle for Knots)**.** *Knots cannot cancel. That is, if the sum of two knots A and B is the trivial knot, then both A and B must be the trivial knot.*

Proof. Let 0 denote the trivial knot. Suppose $A+B=0$. Taking the infinite sum $A+B+A+B+\dots$ and applying the swindle bracketing yields:

$$(A+B)+(A+B)+\dots=0$$
$$A+(B+A)+(B+A)+\dots=A$$

Thus, $A=0$, meaning A is the trivial knot (and similarly for B). □

40.3. **Application to Projective Modules.** We now transition to the algebraic version of the swindle, often attributed to Eilenberg and Mac Lane. Recall the definition of a stably free module:

Definition 40.3. A module M is *stably free* if there exists a *finite* integer n such that $M\oplus R^n$ is a free module.

A natural question is why the integer n must be finite. If we allow n to be infinite, the concept degenerates into the definition of a projective module.

Theorem 40.4. *For any module M, $M\oplus R^\infty\cong R^\infty$ if and only if M is projective.*

Proof. ($\Rightarrow$) Suppose $M\oplus R^\infty\cong R^\infty$. This implies that M is a direct summand of a free module, which inherently means M is a projective module.

($\Leftarrow$) Suppose M is projective. Then there exists some module N such that $M\oplus N\cong F$, where F is a free module. We now apply the Eilenberg-Mazur swindle by taking the infinite direct sum:

$$M\oplus N\oplus M\oplus N\oplus\dots$$

We can bracket this infinite direct sum in two different ways. First, grouping the M and N pairs:

$$(M\oplus N)\oplus(M\oplus N)\oplus\dots\cong F\oplus F\oplus\dots\cong R^\infty$$

Second, isolating the first M and grouping the remaining terms:

$$M\oplus(N\oplus M)\oplus(N\oplus M)\oplus\dots\cong M\oplus F\oplus F\oplus\dots\cong M\oplus R^\infty$$

Equating the two bracketings yields $M\oplus R^\infty\cong R^\infty$. □

This demonstrates exactly why n is restricted to be finite in the definition of stably free modules; without the finiteness condition, the definition captures all projective modules.

40.4. **Stably Free Modules of Infinite Rank.** We can also use the swindle to prove that stably free modules of infinite rank are necessarily free. While we will not rigorously define the "rank" of a stably free module here, we will take it to mean that $M\oplus R^n\cong R^\infty$ for some finite n.

Theorem 40.5. *If M is a stably free module of infinite rank, meaning $M\oplus R^n\cong R^\infty$ for a finite integer n, then M is a free module.*

Proof. Given the isomorphism $R^\infty \cong M \oplus R^n$, we analyze the natural map:

$$R^\infty \twoheadrightarrow M \oplus R^n$$

We can find a finite number of basis elements in R^∞ (spanning a submodule isomorphic to R^m) that map surjectively onto the finite module R^n. By subtracting suitable basis elements to ensure the image of the remaining free module lies entirely inside M, we can reduce the problem to a surjective map:

$$R^\infty \oplus R^m \twoheadrightarrow M \oplus R^n$$

where the map from R^m to R^n is surjective, and the map from R^∞ is injective into M.

Because R^n is free (and therefore projective), the surjective map $R^m \twoheadrightarrow R^n$ splits. Thus, we can write:

$$R^m \cong X \oplus R^n$$

for some submodule X that maps into M. By definition, since $X \oplus R^n \cong R^m$, the module X is stably free. Substituting this back, we find a decomposition for M:

$$M \cong R^\infty \oplus X$$

We now apply the Eilenberg-Mazur swindle to this expression. We construct the infinite direct sum:

$$X \oplus R^n \oplus X \oplus R^n \oplus X \oplus R^n \oplus \dots$$

and we add R^∞ to it. On the one hand, pairing the alternating terms yields:

$$R^\infty \oplus (X \oplus R^n) \oplus (X \oplus R^n) \oplus \dots \cong R^\infty \oplus R^m \oplus R^m \oplus \dots$$

which is an infinite direct sum of free modules, and therefore is itself a free module.

On the other hand, re-pairing the terms by absorbing $R^n \oplus X$ factors into R^∞ yields:

$$X \oplus (R^n \oplus X) \oplus (R^n \oplus X) \oplus \dots \oplus R^\infty \cong X \oplus R^\infty \oplus R^\infty \dots \cong X \oplus R^\infty$$

Since $M \cong X \oplus R^\infty$, we conclude that M is isomorphic to the free module derived from the first bracketing. Thus, M is a free module. □

41. Locally Free Modules

In this lecture, we introduce the concept of locally free modules and contrast them with stably free modules. We will investigate the geometric analogy between locally free modules and vector bundles, and demonstrate precisely why a locally free module is not necessarily stably free.

41.1. **Vector Bundles and Locally Free Modules.** We begin by recalling the geometric notion of a vector bundle. A vector bundle over a space X is a map from a space V to X such that the fibers are vector spaces. To be a well-behaved vector bundle, it must look locally trivial: there must exist an open covering $\{U_i\}$ of X such that locally, the bundle looks like a simple product:

$$V|_{U_i} \cong \mathbb{R}^n \times U_i \to U_i$$

where $\mathbb{R}^n$ is an n-dimensional vector space. For example, if X is a smooth manifold, the tangent bundle mapping the tangent space of X to X is a classic vector bundle.

This geometric structure translates directly into algebra. Let R be the ring of continuous functions on X. The sections of the vector bundle V naturally form a module M over R. Because we can perform pointwise addition and scalar multiplication, M inherits a robust R-module structure.

We generalize this algebraic behavior to arbitrary rings. The topological space X is replaced by the prime spectrum $\mathrm{Spec}(R)$. The open sets U_i correspond to the distinguished basis of open sets $U_{f_i} = \mathrm{Spec}(R_{f_i})$ for elements $f_i \in R$. For these open sets to cover $\mathrm{Spec}(R)$, the elements $\{f_i\}$ must generate the unit ideal.

Definition 41.1. An R-module M is *locally free* if there exists a finite set of elements $f_1, \ldots, f_k \in R$ generating the unit ideal such that for each i, the localized module M_{f_i} is a free module over the localized ring R_{f_i}.

41.2. **Locally Free vs. Stably Free.** Our primary question is to determine the exact relationship between locally free modules and stably free modules.

Proposition 41.2. *If a module M is stably free, then M is locally free.*

Proof. If M is stably free and not finitely generated, we know from previous results that M is inherently free. A free module is trivially locally free (by simply taking the covering set to be $\{1\}$).

If M is stably free and finitely generated, then M is a direct summand of a free module, which means M is a projective module. We will demonstrate later that finitely generated projective modules are always locally free. □

We must ask the converse: does locally free imply stably free? The answer is generally **no**. There are two distinct reasons for this failure—one trivial, and one profound.

41.2.1. *The Trivial Reason: Disconnected Spectra.* For the trivial case, consider a vector bundle over a disconnected space $X = X_1 \cup X_2$. We can easily construct a bundle whose fibers are 1-dimensional over X_1 and 2-dimensional over X_2. This bundle is locally free, but it cannot possibly be stably free. A stably free module M satisfies:

$$M \oplus R^n \cong R^m$$

which implies that the fibers of M must have a constant dimension of $m - n$ everywhere across the space.

Algebraically, we can construct this by taking a disconnected ring:

$$R = \frac{\mathbb{Z}}{6\mathbb{Z}} \cong \frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{3\mathbb{Z}}$$

The spectrum consists of two discrete points. Let $M = \mathbb{Z}/3\mathbb{Z}$. This module has dimension 0 over the first point and dimension 1 over the second. It is locally free (being free over the localized components), but it is not stably free due to its inconsistent rank.

41.2.2. *The Profound Reason: Twisted Modules.* The much more interesting failure occurs when the spectrum is connected, but the module is "twisted."

Geometrically, consider the Möbius band mapping to the circle S^1. This is a 1-dimensional vector bundle over a connected space. It is locally free because we can cover S^1 with two overlapping open intervals, and over each interval, the Möbius band trivializes to a simple product.

However, it is not stably free. As you trace the fiber around the entire circle, its orientation reverses. Adding a finite-dimensional trivial vector bundle to the Möbius band does not fix this orientability defect, meaning it can never be isomorphic to a free (trivial) bundle. Interestingly, if you take the direct sum of two Möbius bands, the orientations cancel out, and the resulting bundle is trivial:

$$M \oplus M \cong R \oplus R$$

41.3. **An Algebraic Example in a Dedekind Domain.** We can precisely model this twisted behavior using algebraic number theory. Let R be the ring of integers of the number field $\mathbb{Q}(\sqrt{-5})$:

$$R = \mathbb{Z}[\sqrt{-5}]$$

This is a standard Dedekind domain that is not a Principal Ideal Domain. We use the classic non-principal ideal M:

$$M = (2, 1 + \sqrt{-5})$$

Because M is not principal, it is not a free R-module. We will prove that M is locally free.

Proof. We must find elements that generate the unit ideal and over which M becomes free. Let us choose $f_1 = 2$ and $f_2 = 3$. Since $3 - 2 = 1$, they clearly generate the unit ideal, so their corresponding open sets U_2 and U_3 cover $\operatorname{Spec}(R)$.

First, we localize at 2. In the ring $R_{1/2}$, the element 2 is invertible. Since $2 \in M$, multiplying by its inverse means $1 \in M_{1/2}$. Therefore, the localized ideal is the entire ring:

$$M_{1/2} = R_{1/2}$$

This is trivially a free module of rank 1 over $R_{1/2}$.

Second, we localize at 3. In $R_{1/3}$, the element 3 is a unit. Recall the factorization in R:

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

Dividing by 3, we find:

$$2 = \frac{(1+\sqrt{-5})(1-\sqrt{-5})}{3}$$

This equation shows that in the localized ring $R_{1/3}$, the generator 2 is merely a multiple of the other generator $(1+\sqrt{-5})$. Consequently, $M_{1/3}$ is generated by the single element $(1+\sqrt{-5})$. Because this generator is not a zero divisor, $M_{1/3}$ is a free module of rank 1 over $R_{1/3}$.

Since M is free over both localizations, M is a locally free module. □

41.4. **Direct Sums of Twisted Modules.** Just as the direct sum of two Möbius bands yields a trivial bundle, the direct sum of our twisted module M with itself yields a free module. We claim that:

$$M \oplus M \cong R \oplus R$$

To see this elegantly, let $\alpha = 1+\sqrt{-5}$ and $\bar{\alpha} = 1-\sqrt{-5}$. We define a surjective module homomorphism from the free module R^2 to M:

$$\phi\colon R^2 \twoheadrightarrow M$$
$$(x, y) \mapsto 2x + \alpha y$$

The kernel of ϕ consists of all pairs (x, y) such that $2x + \alpha y = 0$. Since $\alpha\bar{\alpha} = 6$, we can easily find an element in the kernel. Let $x = 3$ and $y = -\bar{\alpha}$. Then:

$$2(3) + \alpha(-\bar{\alpha}) = 6 - 6 = 0$$

Thus, the kernel is generated by $(3, -\bar{\alpha})$. It turns out this kernel is itself isomorphic to M. Because M is a projective module, the short exact sequence

$$0 \to K \to R^2 \to M \to 0$$

splits, meaning the middle term is isomorphic to the direct sum of the ends. Thus:

$$R^2 \cong M \oplus K \cong M \oplus M$$

This explicit calculation demonstrates that while M is locally free but not free, adding it to itself "untwists" the algebraic structure to form a free module.

This behavior generalizes beautifully. For the ring of integers of any algebraic number field, the non-zero fractional ideals are always locally free modules of rank 1. They are free if and only if they are principal ideals. Taking the sum of a finite number of such ideals yields a module isomorphic to a sum of free modules, up to the ideal class group. Geometrically, these locally free modules of rank 1 correspond to line bundles, and the ideal class group is the arithmetic equivalent of the Picard group in algebraic geometry.

42. Projective Modules

In this lecture, we discuss projective modules. Specifically, we will examine the relationship between projective modules and locally free modules, analyzing when these two concepts coincide and when they diverge.

42.1. **Definitions and Basic Properties.** Let us first recall the definitions of locally free and projective modules.

Definition 42.1. An R-module M is *locally free* if we can cover the spectrum $\mathrm{Spec}(R)$ with a set of distinguished open sets $U_{f_i} = \mathrm{Spec}(R_{f_i})$ such that the localized module M_{f_i} is a free module over the localized ring R_{f_i} for each i. Equivalently, the elements f_i must generate the unit ideal:

$$(f_1, \ldots, f_n) = R$$

This concept directly corresponds to the geometric notion of vector bundles, which are locally trivial families of vector spaces parameterized by a base space.

Definition 42.2. An R-module P is *projective* if it satisfies the following lifting property: for any surjective module homomorphism $A \twoheadrightarrow B$ and any homomorphism $P \to B$, there exists a lifted homomorphism $P \to A$ making the diagram commute.

Several basic properties of projective modules are easily verified:

- Free modules are trivially projective. One can simply take a basis for the free module, map it into B, and lift those basis elements to preimages in A.
- Any direct summand of a projective module is projective. If $P = X \oplus Y$ is projective, then both X and Y are projective.
- Any direct sum of projective modules is projective.
- A submodule of a projective module is *not* necessarily projective.

42.2. **Are Projective and Locally Free Modules the Same?** We are faced with a fundamental question: are projective modules the same as locally free modules? Equivalently, are projective objects exactly the same as vector bundles?

The answer depends entirely on the specific field of mathematics you ask:

- **Commutative Algebra / Differential Geometry:** Yes, locally free objects are generally equivalent to projective objects.
- **Algebraic Geometry / Complex Analytic Geometry:** No, there are many examples of locally free objects that are not projective.

This discrepancy often causes significant confusion for mathematicians who transition between these fields. The underlying reason for this difference relies heavily on cohomology.

In algebraic geometry, one studies the obstruction to lifting using exact sequences. For modules or sheaves A and B, we encounter an exact sequence involving the Ext functor (or a first cohomology group):

$$\cdots \to \mathrm{Hom}(P, A) \to \mathrm{Hom}(P, B) \to \mathrm{Ext}^1(P, A) \to \ldots$$

(Note: This Ext^1 term acts as a piece of the Grothendieck spectral sequence of a composed functor). The critical point is that locally free implies projective if and only if this first cohomology group vanishes.

Does this group vanish?

- Over commutative rings or smooth manifolds, we possess *partitions of unity*. A partition of unity allows us to write 1 as a sum of functions supported on local open sets. This analytic/algebraic tool forces the cohomology group to identically vanish, meaning locally free objects are exactly projective.
- Over algebraic varieties or complex analytic manifolds, partitions of unity do not exist. Consequently, the cohomology group does not necessarily vanish, and we can find locally free objects (vector bundles) that are not projective. In fact, in algebraic geometry, even a free 1-dimensional sheaf over a 1-dimensional projective space might fail to act projectively on global sections, a fact that is highly disconcerting if one approaches it purely from the perspective of commutative algebra.

42.3. **The Commutative Ring Case.** Because this is a course in commutative algebra, we restrict our focus to the ring theory case, where the necessary algebraic partitions of unity (elements generating the unit ideal) are available.

Over a commutative ring, **locally free always implies projective**. We must then ask: does the converse hold? Does projective imply locally free?

The answer is **yes for finitely generated modules**, but **no in general**.

Let us briefly sketch why the finitely generated case holds:

(1) If P is a finitely generated projective module, it is relatively easy to show that it is finitely presented.
(2) With somewhat more effort, one can show that the stalks of P are free. The stalk of P at a prime ideal $\mathfrak{p}$ is the localization $P_{\mathfrak{p}}$ over the local ring $R_{\mathfrak{p}}$. A theorem by Kaplansky dictates that projective modules over local rings are always free.
(3) If a module is finitely presented and its stalks are free, it implies that the module is locally free.

In practice, locally free modules encountered in commutative algebra are almost always obviously projective. For instance, they are often stably free, meaning they are direct summands of finite free modules, making them trivially projective. Therefore, while the precise theoretical equivalence is satisfying to know, the distinction rarely causes practical issues for finitely generated modules.

42.4. **A Non-Finitely Generated Counterexample.** We now construct an explicit counterexample of a projective module that is *not* locally free. To circumvent the previous theorem, we must construct a module that is not finitely generated.

Let X be an infinite set (countable or uncountable). We define our ring R as the ring of functions from X to the finite field with two elements, $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$:

$$R = \prod_{x \in X} \frac{\mathbb{Z}}{2\mathbb{Z}}$$

This ring is an example of a complete Boolean algebra. Such structures are relatively rare in commutative algebra but are used heavily in set theory, specifically

within Paul Cohen's notion of forcing to construct models of Zermelo-Fraenkel set theory with various exotic properties.

We define our module I to be the ideal consisting of functions with *finite support* (functions that evaluate to non-zero on only a finite number of elements of X). We can view elements of R as infinite binary sequences, while elements of I are binary sequences that eventually stabilize to all zeros.

First, we prove that I is projective. Notice that I decomposes into a direct sum over the points of X:

$$I = \bigoplus_{x \in X} I_x$$

where I_x is the submodule of functions supported exclusively at the single point x. For each x, the ambient ring R naturally splits as a direct sum:

$$R = I_x \oplus J_x$$

where J_x consists of all functions whose support is disjoint from x. Since R is trivially a free module of rank 1, I_x is a direct summand of a free module. This establishes that I_x is projective. Because any direct sum of projective modules is projective, the infinite direct sum $I = \bigoplus I_x$ is inherently a projective module.

Next, we prove that I is *not* locally free. To test local freeness, we localize R at an element $f \in R$.

- If f has finite support, the localized module $I[f^{-1}]$ is indeed a free module over $R[f^{-1}]$ (it is directly isomorphic to $R[f^{-1}]$).
- If f has infinite support, it is easy to check that $I[f^{-1}]$ fails to be free over $R[f^{-1}]$.

For I to be locally free, we would need to cover the spectrum $\operatorname{Spec}(R)$ with distinguished open sets U_{f_i} where every f_i has finite support. Algebraically, this means the elements $f_1, \ldots, f_n$ must generate the unit ideal R.

However, a finite collection of elements $f_1, \ldots, f_n$ possessing finite support cannot possibly generate the unit ideal. The union of their supports is still a finite set, and generating R requires covering the entire infinite set X. (As a technical aside, if X is infinite, the spectrum $\operatorname{Spec}(R)$ contains highly pathological extra elements called ultrafilters, which cannot be covered by functions of finite support).

Thus, I provides an example of a module that is projective but undeniably not locally free.

Finally, this construction serves as a counterexample to another common intuition. Over smooth manifolds, an arbitrary direct sum of vector bundles remains a well-defined vector bundle. However, the analog for rings is false. In our example, the ideal $I = \bigoplus I_x$ is an infinite direct sum of the modules I_x. Each I_x is individually locally free, yet their infinite direct sum I fails to be locally free. When dealing with modules that are not finitely generated, one must be exceedingly careful, as the direct geometric analogy between vector bundles and locally free modules breaks down.

43. Stalkwise Locally Free Modules

In this lecture, we continue our investigation of module properties by focusing on *stalkwise locally free* modules. This is a somewhat technical topic, but it serves to bridge the gap between locally free modules and projective modules.

43.1. **Relation to Other Module Classes.** Recall that we have previously discussed two related classes of modules:

- *Locally free modules*
- *Projective modules*

Stalkwise locally free modules are conceptually bracketed between locally free modules and projective modules. Crucially, all three of these properties are completely equivalent for *finitely presented* modules.

In a previous lecture, we established the relationship between locally free modules and projective modules. Our focus here will be on the precise relationship between projective modules and stalkwise locally free modules. Specifically, we will demonstrate that while they coincide under finite presentation, they diverge otherwise.

43.2. **Counterexamples for Infinite and Finitely Generated Modules.** We begin by providing examples of modules that are stalkwise locally free but fail to be projective.

43.2.1. *A Non-Finitely Generated Counterexample.*

Example 43.1. Let the base ring be the integers, $R = \mathbb{Z}$. We define our module M to be the set of all rational numbers with a square-free denominator:

$$M = \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{Z}, b \text{ is square-free} \right\}$$

For any prime ideal $\mathfrak{p} = (p)$ in the spectrum $\operatorname{Spec}(\mathbb{Z})$, we can localize M at $\mathfrak{p}$. Because the denominator of any element in M can only contain a single factor of p, localizing (which allows us to invert any integer not divisible by p) "absorbs" this single factor. It is straightforward to check that:

$$M_{\mathfrak{p}} \cong \mathbb{Z}_{\mathfrak{p}}$$

Geometrically, one can think of all the stalks of M as being one-dimensional free modules over the local rings. Thus, M is stalkwise locally free.

However, M is obviously not a free module over $\mathbb{Z}$. Since projective modules over $\mathbb{Z}$ are exactly the free modules, M is not projective.

This example serves as a counterexample to naive attempts to define invertible modules. We define a module M as *invertible* if M is locally free of rank 1. This example shows that you cannot replace "locally free" with "stalkwise locally free" in this definition, because M is stalkwise locally free of rank 1, but it is neither free nor locally free of rank 1.

43.2.2. *A Finitely Generated Counterexample.* One might suspect that the previous failure occurred because M was not finitely generated. We now construct a finitely generated module that is stalkwise locally free but not projective.

Example 43.2. Recall the Boolean ring R constructed from an infinite set X:

$$R = \{f \colon X \to \mathbb{Z}/2\mathbb{Z}\}$$

We define the ideal I as the set of functions with finite support:

$$I = \{f \in R \mid f(x) \neq 0 \text{ for only finitely many } x \in X\}$$

Instead of looking at I, we define our module M as the quotient:

$$M = \frac{R}{I}$$

M is finitely generated because it is generated by the single identity element $1 \in R/I$. However, M is not finitely presented because the kernel I is not finitely generated.

To see that all stalks of M are free, we analyze the localizations of R. Because R is a Boolean ring, it satisfies the identity:

$$x^2 = x \quad \text{for all } x \in R$$

If we localize a Boolean ring at any prime ideal $\mathfrak{p}$, the resulting local ring $R_\mathfrak{p}$ must still be Boolean. In a local Boolean ring, any element not in the maximal ideal is a unit. Since $x^2 = x$, the only unit is 1. Thus, the only element not in the maximal ideal is 1, which forces the local ring to be exactly the field with two elements:

$$R_\mathfrak{p} \cong \frac{\mathbb{Z}}{2\mathbb{Z}}$$

Since every localization is a field, every module over it is a vector space, and thus automatically a free module. Therefore, all stalks of M are free, meaning M is stalkwise locally free.

Despite this, M is not projective. Consider the short exact sequence:

$$0 \to I \to R \to M \to 0$$

If M were projective, this exact sequence would split, giving an isomorphism:

$$R \cong I \oplus M$$

Since R is generated by 1 element, this splitting would imply that I is a finitely generated module (as it would be a direct summand of a finitely generated module). But we know I is not finitely generated, so the sequence cannot split, and M cannot be projective.

43.3. **Equivalence under Finite Presentation.** The counterexamples above demonstrate that for stalkwise locally free modules to be projective, they must be finitely presented. We now prove this implication.

Theorem 43.3. *If a module A is finitely presented and stalkwise locally free, then A is projective.*

Proof. Because A is finitely presented, we can write it as the quotient of a finite free module F by a finitely generated kernel K. This gives a short exact sequence:

$$0 \to K \to F \to A \to 0$$

We are given that A is stalkwise locally free. A key property of stalkwise locally free modules is that they are always *flat*. (This is because flatness is a local property, and free modules are flat; if all stalks are free, all stalks are flat, making the module flat globally).

We proceed in four steps.

Step 1: For any ideal $I \subseteq R$, we claim that:

$$K \cap FI = KI$$

This follows directly from the flatness of A. Because A is flat, tensoring the short exact sequence with the ideal I preserves exactness. By comparing the tensored sequence:

$$0 \to K \otimes_R I \to F \otimes_R I \to A \otimes_R I \to 0$$

with the standard submodule sequence:

$$0 \to K \cap FI \to FI \to AI \to 0$$

the natural isomorphism between the free module terms forces $K \cap FI = KI$.

Step 2: If $u \in K$, we can find a homomorphism $f\colon F \to K$ that fixes u (i.e., $f(u) = u$). Because F is a free module, we can write u in terms of a finite basis f_i of F:

$$u = \sum r_i f_i$$

Let I be the ideal generated by the coefficients r_i. By definition, $u \in FI$. Since u is also in K, we have:

$$u \in K \cap FI$$

By Step 1, $u \in KI$. Therefore, we can express u as a linear combination:

$$u = \sum k_i r_i \quad \text{for some } k_i \in K$$

We can now define the desired homomorphism $f\colon F \to K$ by specifying its action on the basis elements:

$$f(f_i) = k_i$$

Evaluating f on u, we obtain $f(u) = \sum r_i f(f_i) = \sum r_i k_i = u$. Thus, f fixes u.

Step 3: If $u_1, \ldots, u_n \in K$, we can find a single homomorphism $f\colon F \to K$ fixing all the u_i. This follows from Step 2 via a standard induction argument on the number of elements.

Step 4: Because A is finitely presented, the kernel K is finitely generated. Let $u_1, \dots, u_n$ be a finite generating set for K. By Step 3, there exists a homomorphism $f\colon F \to K$ such that $f(u_i) = u_i$ for all generators. This implies that the restriction of f to K is the identity map on K. Therefore, the short exact sequence:

$$0 \to K \to F \to A \to 0$$

splits. Since the sequence splits, A is isomorphic to a direct summand of the free module F. Any direct summand of a free module is, by definition, projective. Thus, A is projective. □

Remark 43.4. This highlights a subtle but profound difference between finitely generated and finitely presented modules. Over a Noetherian ring, any finitely generated module is automatically finitely presented, making this distinction irrelevant. However, over non-Noetherian rings, the gap between these finiteness conditions allows for modules that are stalkwise locally free but fail to be projective.

This concludes our technical discussion on stalkwise locally free modules. In the next lecture, we will move on to flat modules, which represent perhaps the single most important class of modules in commutative algebra.

44. Flat Modules

This lecture introduces flat modules. We recall the definition, summarize how flatness relates to other module properties, examine examples and counterexamples, and finally prove a powerful structure theorem for finitely presented flat modules over local rings.

44.1. **Definition and the Hierarchy of Modules.** We begin by recalling the definition of a flat module.

Definition 44.1. An R-module M is *flat* if tensoring with M preserves exact sequences. Specifically, M is flat if for every exact sequence of R-modules

$$0 \to A \to B$$

the tensored sequence remains exact:

$$0 \to M \otimes_R A \to M \otimes_R B$$

In other words, tensoring with a flat module preserves injectiveness.

As we have seen in previous lectures, we have established a strict hierarchy of module properties, each successively weakening the condition of being a free module. We can summarize the chain of implications as follows:

$$\text{Free} \implies \text{Stably Free} \implies \text{Locally Free} \implies \text{Projective} \implies \text{Stalkwise Free} \implies \text{Flat}$$

Out of all these conditions, flatness is arguably the most practically useful. This utility stems from two main advantages:

(1) It is a local property and thus relatively easy to verify. A module M is flat if and only if the localization $M_{\mathfrak{p}}$ is flat over $R_{\mathfrak{p}}$ for all prime ideals $\mathfrak{p} \in \mathrm{Spec}(R)$ (or equivalently, all maximal ideals).
(2) There is an enormous abundance of flat modules, whereas free, stably free, or projective modules are comparatively rare.

To highlight the geometric perspective: when transitioning to algebraic geometry and working with schemes, projective sheaves are almost non-existent over non-affine schemes. In contrast, flat sheaves are ubiquitous and serve as the essential tool for defining nice, continuous families of spaces.

44.2. **Examples and Counterexamples.** Because of the hierarchy established above, any free, stably free, locally free, projective, or stalkwise free module is automatically flat. However, flatness captures many more modules.

Example 44.2. Any localization $S^{-1}R$ of a ring R with respect to a multiplicative subset S is a flat R-module. In particular, the local ring $R_{\mathfrak{p}}$ at any prime ideal $\mathfrak{p}$ is flat. Generally, localizations are neither projective nor free, illustrating that flatness is a much broader condition.

Example 44.3 (A Non-Flat Module)**.** The standard counterexample of a module that is *not* flat is the module $\mathbb{Z}/2\mathbb{Z}$ over the ring of integers $R = \mathbb{Z}$. As we have seen previously, tensoring the injective map $\mathbb{Z} \xrightarrow{\times 2} \mathbb{Z}$ with $\mathbb{Z}/2\mathbb{Z}$ destroys exactness.

We must also exercise caution when performing standard module operations, as flatness is not universally preserved by taking submodules or quotients.

Proposition 44.4. *Quotients and submodules of flat modules are generally not flat.*

Proof. Consider the polynomial ring in two variables, $R = k[x, y]$. The ring R itself is a free module, and therefore flat. Consider the ideal $I = (x, y)$ as a submodule, yielding the exact sequence:

$$0 \to (x, y) \to k[x, y] \to k \to 0$$

Here, the quotient module is the field k. The middle term $k[x, y]$ is flat. However, the quotient k is not a flat R-module (it is annihilated by x and y, creating torsion). Since the quotient of a flat module by a submodule need not be flat, the property fails for quotients.

Similarly, the submodule $I = (x, y)$ is not flat, establishing that submodules of flat modules are also not generally flat. (We will rigorously prove that (x, y) is not flat in later lectures using homological algebra). □

Despite these failures, flatness behaves predictably within exact sequences under specific configurations.

Proposition 44.5. *Given a short exact sequence of R-modules:*

$$0 \to A \to B \to C \to 0$$

The following logical deductions hold:

- *If B and C are flat, then A is flat (the kernel of a surjection between flat modules is flat).*
- *If A and C are flat, then B is flat.*
- *If A and B are flat, it does* not *necessarily imply that C is flat.*

44.3. **Flat Modules over Local Rings.** One of the most powerful features of flat modules is that, under the assumption of finite presentation, flatness over a local ring forces the module to be entirely free.

Theorem 44.6. *Suppose R is a local ring with maximal ideal $\mathfrak{m}$, and let M be a finitely presented R-module. The following four conditions are equivalent:*

(1) M is free.
(2) M is projective.
(3) M is flat.
(4) For any exact sequence $0 \to A \to B \to M \to 0$, the sequence obtained by tensoring with $R/\mathfrak{m}$ remains exact:

$$0 \to \frac{A}{\mathfrak{m}A} \to \frac{B}{\mathfrak{m}B} \to \frac{M}{\mathfrak{m}M} \to 0$$

Remark 44.7. Once we develop the basics of homological algebra, we will recognize that Condition 4 is equivalent to stating that the torsion group vanishes:

$$\operatorname{Tor}_1^R\left(M, \frac{R}{\mathfrak{m}}\right) = 0$$

For now, we use Condition 4 as an ad hoc, elementary definition representing this vanishing condition. Additionally, finite generation is technically sufficient for this theorem, but the proof is substantially more involved. Since we primarily concern ourselves with Noetherian rings where finite generation implies finite presentation, we restrict our proof to finitely presented modules.

Proof. The implications (1) $\implies$ (2) $\implies$ (3) are universally true and have been established previously. The implication (3) $\implies$ (4) is trivial by the definition of flatness: if M is flat, then tensoring any exact sequence ending in M with any module (such as $R/\mathfrak{m}$) preserves exactness.

The crux of the theorem is proving (4) $\implies$ (1). We proceed by heavily utilizing Nakayama's Lemma.

Because M is finitely generated, the quotient $M/\mathfrak{m}M$ is a finite-dimensional vector space over the residue field $k = R/\mathfrak{m}$. Let n be the dimension of this vector space. We can choose a set of n basis elements in $M/\mathfrak{m}M$ and lift them to elements in M. This defines a module homomorphism from the free module R^n to M:

$$f\colon R^n \to M$$

By our construction, tensoring f with $R/\mathfrak{m}$ yields an isomorphism of vector spaces:

$$\bar{f}\colon \left(\frac{R}{\mathfrak{m}}\right)^n \xrightarrow{\sim} \frac{M}{\mathfrak{m}M}$$

Let N be the cokernel of f. We have $R^n \to M \to N \to 0$. Tensoring with $R/\mathfrak{m}$ is right-exact, so $(R/\mathfrak{m})^n \to M/\mathfrak{m}M \to N/\mathfrak{m}N \to 0$ is exact. Since $\bar{f}$ is an isomorphism (and thus surjective), we must have $N/\mathfrak{m}N = 0$. By Nakayama's Lemma, since N is finitely generated, $N = 0$. Thus, f is surjective.

We now have a short exact sequence, where K is the kernel of f:

$$0 \to K \to R^n \xrightarrow{f} M \to 0$$

Because M is a *finitely presented* module and R^n is a finite free module, it is a standard algebraic fact that the kernel K must be finitely generated.

Now we invoke Condition 4. Since the sequence ending in M is exact, tensoring with $R/\mathfrak{m}$ yields an exact sequence:

$$0 \to \frac{K}{\mathfrak{m}K} \to \left(\frac{R}{\mathfrak{m}}\right)^n \xrightarrow{\bar{f}} \frac{M}{\mathfrak{m}M} \to 0$$

We already established that the map $\bar{f}$ is an isomorphism. Therefore, its kernel must be strictly zero:

$$\frac{K}{\mathfrak{m}K} = 0$$

Since K is finitely generated, we can apply Nakayama's Lemma to K, which forces:

$$K = 0$$

Since the kernel is trivial, the surjective map $f\colon R^n \to M$ is an isomorphism. Therefore, $M \cong R^n$, proving that M is a free module. □

44.4. **Finitely Presented Modules over Arbitrary Rings.** We can seamlessly elevate the local ring theorem to arbitrary commutative rings by passing through stalks.

Corollary 44.8. *Let R be any commutative ring, and let M be a finitely presented R-module. The following conditions are equivalent:*

(1) M is projective.
(2) M is locally free.
(3) M is stalkwise free.
(4) M is flat.

Proof. We have previously established the chain (1) $\implies$ (2) $\implies$ (3) $\implies$ (4). It suffices to prove that (4) $\implies$ (1).

Suppose M is a flat, finitely presented R-module. Because flatness is a local property, for every prime ideal $\mathfrak{p} \in \operatorname{Spec}(R)$, the localized module $M_\mathfrak{p}$ is a flat module over the local ring $R_\mathfrak{p}$.

Furthermore, localizing a finitely presented module yields a finitely presented module over the local ring. By Theorem 44.6, since $M_\mathfrak{p}$ is a finitely presented flat module over a local ring, $M_\mathfrak{p}$ must be a free $R_\mathfrak{p}$-module.

Thus, the stalks of M are free. As we proved in a prior lecture, if a module is finitely presented and stalkwise free, it is inherently a projective module. This closes the loop of equivalences. □

45. Torsion Free Modules

This lecture concludes the sequence on the different types of modules. We will focus on torsion-free modules and their properties. While there is relatively little to say about them compared to flat modules, it is necessary to place them within the broader hierarchy we have established.

45.1. **The Hierarchy of Modules.** To summarize our progress, we have defined a chain of implications for modules:

$$\begin{aligned} \text{Free} &\implies \text{Stably Free} \\ &\implies \text{Locally Free} \\ &\implies \text{Projective} \\ &\implies \text{Stalkwise Free} \\ &\implies \text{Flat} \\ &\implies \text{Torsion Free} \\ &\implies \text{Co-primary (over integral domains)} \end{aligned}$$

The most critical property in this sequence is flatness. In this section, we examine the final two properties: torsion-free and co-primary.

45.2. **Definition of Torsion Free.**

Definition 45.1. Let R be an integral domain. An R-module M is *torsion free* if for any $x \in R$ and $m \in M$:

$$xm = 0 \implies x = 0 \text{ or } m = 0$$

Over a general commutative ring with zero divisors, there are several competing definitions in the literature. A common generalization states that $xm = 0$ implies $m = 0$ or x is a zero divisor. However, if R has zero divisors, this definition renders almost nothing torsion-free. Consequently, torsion-free modules are rarely used over general rings; their utility is almost entirely restricted to integral domains.

45.3. **Flatness Implies Torsion Free.**

Proposition 45.2. *If M is a flat module over an integral domain R, then M is torsion free.*

Proof. Let $x \in R$ be a non-zero element. Because R is an integral domain, x is not a zero divisor. Thus, multiplication by x is an injective map, yielding the exact sequence:

$$0 \to R \xrightarrow{\times x} R \to \frac{R}{xR} \to 0$$

Because M is flat, tensoring this exact sequence with M preserves exactness:

$$0 \to M \otimes_R R \xrightarrow{\times x} M \otimes_R R$$

Since $M \otimes_R R \cong M$, this forces the multiplication by x map on M to be injective:

$$0 \to M \xrightarrow{\times x} M$$

This injectivity precisely means that x is not a zero divisor on M. Since this holds for all non-zero $x \in R$, M is torsion free. □

45.4. **Torsion Free Modules over PIDs.** The primary reason torsion-free modules are frequently utilized is that over sufficiently nice rings, the property coincides exactly with flatness, while being analytically much simpler to verify.

Proposition 45.3. *Over a Principal Ideal Domain (PID), an R-module is torsion free if and only if it is flat.*

Proof Sketch. A general theorem states that an R-module M is flat if and only if the natural map $M \otimes_R I \to M$ is injective for all ideals $I \subseteq R$. In a PID, every ideal I is principal, generated by a single element. Testing flatness thus reduces to testing injectivity of multiplication by single elements, which is exactly the definition of being torsion free. □

Remark 45.4. This equivalence also holds over Dedekind domains. Over such rings, one may as well forget about torsion freeness and use flatness instead, as flatness behaves universally well over all rings.

45.5. **Counterexamples: Torsion Free but Not Flat.** While the properties coincide over PIDs, torsion freeness does not imply flatness in general, not even for local rings.

Example 45.5 (Polynomial Ring)**.** Let $R = k[x, y]$ be the polynomial ring over a field, and let $I = (x, y)$ be the maximal ideal at the origin. As a submodule of the free module R, I is trivially torsion free. However, I is not flat. One can demonstrate this by noting that the torsion group $\mathrm{Tor}_1(I, k) \neq 0$. Alternatively, tensoring the exact sequence $0 \to I \to k[x, y]$ with I reveals that the map $I \otimes_R I \to I$ is not injective.

Example 45.6 (Local Ring)**.** Recall from previous lectures that over local rings, many properties like flatness and projectiveness coincide. However, a torsion-free module over a local ring is *not* necessarily flat. We construct a counterexample analogously to the polynomial case: let $R = k[[x, y]]$ be the local ring of formal power series, and let $I = (x, y)$. The ideal I is torsion free, but it is demonstrably not flat. Thus, torsion freeness is a weaker condition, even locally.

45.6. **Relation to Co-primary Modules.** We conclude by comparing torsion freeness with the co-primary condition. Recall that a module is *co-primary* if it possesses exactly one associated prime.

Proposition 45.7. *Over an integral domain R, any torsion-free module M is co-primary.*

Proof. If M is torsion free, the annihilator of any non-zero element $m \in M$ is the zero ideal:

$$\mathrm{Ann}(m) = (0)$$

The associated primes of M are precisely the prime ideals that arise as annihilators of non-zero elements. Therefore, the only associated prime of M is (0):

$$\mathrm{Ass}(M) = \{(0)\}$$

Because M has exactly one associated prime, it is co-primary. □

The converse, however, is trivially false.

Example 45.8. Let $R = \mathbb{Z}$ and $M = \mathbb{Z}/2\mathbb{Z}$. The module M is clearly not torsion free, as $2 \cdot 1 = 0$ in M. However, the only associated prime of M is (2):

$$\operatorname{Ass}(M) = \{(2)\}$$

Since it has exactly one associated prime, M is perfectly co-primary. Thus, co-primary implies absolutely nothing about being torsion free.

In the upcoming lectures, having concluded our survey of module types, we will transition into an introduction to homological algebra, which will provide the necessary framework to formally calculate torsion groups and prove the various assertions we have postponed. (Note to the reader: This set of lectures is typed up in a separate document.)

Part 6. Limits, Colimits, and Completions

46. Limits and Colimits of Modules

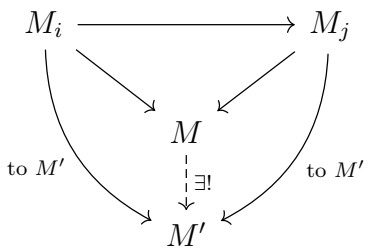
This section introduces limits and colimits, primarily focusing on modules over a commutative ring. We will begin with the abstract categorical definitions and proceed through numerous concrete examples to build intuition.

46.1. **Abstract Definitions.** To define a colimit or a limit, we start with a general category $\mathcal{C}$, which consists of a collection of objects and morphisms (arrows) between them. We then take a functor $F\colon \mathcal{C} \to \mathrm{Mod}_R$, where Mod_R is the category of modules over a ring R.

This means that for each object in the category $\mathcal{C}$, we assign a module M_i, and for each morphism in $\mathcal{C}$, we assign a corresponding module homomorphism between these modules.

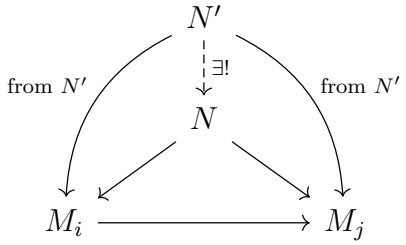
Definition 46.1. A *colimit* of the functor F is a universal object M such that there exist homomorphisms from all the individual modules M_i to M, making all relevant diagrams commute.

Furthermore, M is universal with this property: if M' is any other module equipped with compatible maps from all the M_i, then there exists a unique homomorphism from M to M' making the entire diagram commute.



Definition 46.2. A *limit* of the functor F is defined dually. It is a universal object N equipped with homomorphisms *to* all the modules M_i, making the diagrams commute.

If N' is any other object with compatible maps to all the M_i, there exists a unique map from N' to N.



These abstract definitions can be difficult to grasp initially, so we will examine several concrete examples of categories $\mathcal{C}$ to see what limits and colimits look like in practice.

46.2. **Examples of Limits and Colimits.**

Example 46.3 (Discrete Categories)**.** Let the category $\mathcal{C}$ consist of exactly two points with no morphisms between them (other than the identity maps). The functor F simply selects two modules, M_1 and M_2.

The colimit requires a universal module M with maps from M_1 and M_2. This is simply the direct sum:

$$\text{Colimit} = M_1 \oplus M_2$$

Any other module M' with maps from M_1 and M_2 uniquely factors through $M_1 \oplus M_2$.

Somewhat confusingly, the limit in this specific finite case is also the direct sum (or direct product, which coincides for finite sets in abelian categories).

If we instead let $\mathcal{C}$ consist of an infinite number of discrete points, assigning modules $M_1, M_2, M_3, \ldots$, the colimit and limit diverge:

$$\text{Colimit} = \bigoplus_i M_i$$

$$\text{Limit} = \prod_i M_i$$

The direct sum is the colimit because it satisfies the universal property for maps *from* the M_i. The direct product is the limit because it satisfies the universal property for maps *to* the M_i. The direct sum is a proper submodule of the direct product, consisting only of sequences where all but finitely many entries are zero.

Example 46.4 (Parallel Arrows)**.** Let $\mathcal{C}$ be a category with two objects and two morphisms between them. The functor F maps this to two modules, M_1 and M_2, with two homomorphisms between them. Let us assume one homomorphism is the zero map, and the other is a random homomorphism f:

$$M_1 \xrightarrow{0,f} M_2$$

For the colimit, we seek a module M with maps from M_1 and M_2. The commutativity of the diagram forces the composition $M_1 \xrightarrow{f} M_2 \to M$ to equal the

composition $M_1 \xrightarrow{0} M_2 \to M$. This means the image of f must map to zero in M. Thus, the colimit is a quotient:

$$\text{Colimit} = \frac{M_2}{\text{im}(f)} = \text{coker}(f)$$

For the limit, we seek a module N with maps to M_1 and M_2. Commutativity requires the map $N \to M_2$ to simultaneously equal zero and the composition $N \to M_1 \xrightarrow{f} M_2$. This forces N to map exactly into the kernel of f:

$$\text{Limit} = \ker(f)$$

Thus, kernels and cokernels are special cases of limits and colimits.

Example 46.5 (Sequential Categories)**.** Let $\mathcal{C}$ correspond to the natural numbers, with a morphism $i \to j$ if and only if $i \leq j$. This yields a sequence:

$$M_0 \to M_1 \to M_2 \to \dots$$

Suppose each M_i is a submodule of M_{i+1} and the maps are strict inclusions. The colimit is the module M containing all the M_i, which is simply their union:

$$\text{Colimit} = \bigcup_{i=0}^{\infty} M_i$$

The limit for this diagram is trivial. A module N mapping to all M_i is completely determined by its map to M_0, making the limit simply M_0.

If the maps are not inclusions, colimits can collapse unexpectedly. Consider:

$$\mathbb{Z} \xrightarrow{0} \mathbb{Z} \xrightarrow{0} \mathbb{Z} \xrightarrow{0} \dots$$

A colimit M must receive maps from all these $\mathbb{Z}$'s. However, because every map factors through a zero morphism, every element must map to zero. Thus:

$$\text{Colimit} = 0$$

even though all the individual modules are non-zero.

Example 46.6 (Inverse Sequences)**.** Suppose we reverse the arrows of the previous sequence:

$$\dots \to M_2 \to M_1 \to M_0$$

Here, the colimit is trivially M_0. The limit, however, is much more interesting. It is known as the *projective limit* (or inverse limit).

It is constructed as the submodule of the direct product consisting of compatible sequences:

$$\text{Limit} = \varprojlim M_i = \left\{ (m_0, m_1, m_2, \dots) \in \prod M_i \;\middle|\; f_{ij}(m_i) = m_j \text{ for } i > j \right\}$$

This construction is heavily utilized in algebra, such as forming the p-adic integers from the sequence $\mathbb{Z}/p\mathbb{Z} \leftarrow \mathbb{Z}/p^2\mathbb{Z} \leftarrow \mathbb{Z}/p^3\mathbb{Z} \leftarrow \dots$.

For a highly complicated, random category with many objects and interlocking arrows, the colimit can be understood via generators and relations. We assign a generator for each element in each module M_i, and every morphism $M_i \to M_j$ dictates a relation identifying the source element with its target in the colimit.

46.3. **Filtered Categories and Directed Posets.** Certain types of categories yield exceptionally well-behaved limits and colimits.

Definition 46.7. A *directed poset* is a partially ordered set where for any two elements A and B, there exists an element C such that $C \geq A$ and $C \geq B$. A poset can be viewed as a category with at most one morphism between any two objects.

Colimits taken over directed posets are often called *direct limits*. The generalization of a directed poset to an arbitrary category is a filtered category.

Definition 46.8. A category $\mathcal{C}$ is a *filtered category* if it satisfies two conditions:

(1) For any two objects A and B, there exists an object C with morphisms $A \to C$ and $B \to C$.
(2) For any two objects A and B and any two parallel morphisms $f, g\colon A \to B$, there exists an object C and a morphism $h\colon B \to C$ such that the compositions become equal:

$$hf = hg$$

Notice that the second condition is trivially satisfied in a poset since there is at most one morphism between objects.

Many fundamental constructions are filtered colimits.

Example 46.9 (Localization)**.** Let S be a multiplicative subset of a ring R. The localization $S^{-1}R$ is a filtered colimit of the rings $R[s^{-1}]$ for $s \in S$. The category has objects corresponding to elements of S, with a morphism $s \to t$ whenever s divides t. Because S is multiplicatively closed, given s and t, the product $st \in S$ provides the required object for the filtered condition.

Example 46.10 (The Rationals)**.** The field of rational numbers $\mathbb{Q}$ is a filtered colimit of copies of the integers $\mathbb{Z}$. We can construct the sequence:

$$\mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \xrightarrow{\times 3} \mathbb{Z} \xrightarrow{\times 4} \dots$$

The colimit of this filtered system is $\mathbb{Q}$.

Remark 46.11 (Warning on Functors vs. Subcategories). It is crucial to distinguish between a functor from a category to Mod_R, and a subcategory of Mod_R. This distinction occasionally leads to profound confusion.

Consider the diagram:

$$\mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \xrightarrow{\times 2} \dots$$

If we view this as a *functor* from the category $\mathbb{N}$ to $\mathbb{Z}$-modules, the colimit is the ring $\mathbb{Z}[1/2]$.

However, suppose we mistakenly attempt to treat this purely as a *subcategory* of $\mathbb{Z}$-modules consisting of the single object $\mathbb{Z}$ and the morphism of multiplication

by 2. In this erroneous interpretation, the identity map and the multiplication-by-2 map must eventually equalize in the colimit. This forces 1 and 2 to map to the same element, collapsing the colimit to 0.

This exact technical confusion regarding non-standard definitions of categories versus functors was at the heart of the initial controversy surrounding Mochizuki's proposed proof of the ABC conjecture.

46.4. **Exactness.** A fundamental question regarding limits and colimits is how they interact with exact sequences. Suppose we have a functor assigning a short exact sequence to each object in our category:

$$0 \to A_i \to B_i \to C_i \to 0$$

If we take the limit or colimit of this entire system, is the resulting sequence

$$0 \to \varinjlim A_i \to \varinjlim B_i \to \varinjlim C_i \to 0$$

still exact? The answer is that it is sometimes exact and sometimes not, a topic we will thoroughly investigate in the next section.

47. Colimits and Exactness

This lecture focuses on the problem of exactness preservation for colimits. Suppose we have a collection of exact sequences of modules over a ring, indexed by elements i in some category J, equipped with suitable morphisms. In other words, we have a functor from J to the category of modules. We ask: does taking the colimit over J preserve exactness?

47.1. Right Exactness of Colimits.

Theorem 47.1. *The colimit functor over any category J is* right exact.

Suppose we have a right exact sequence of functors from J to modules:

$$A_i \to B_i \to C_i \to 0$$

Taking the colimit yields a right exact sequence:

$$\varinjlim_{i \in J} A_i \to \varinjlim_{i \in J} B_i \to \varinjlim_{i \in J} C_i \to 0$$

There are two primary methods to see this.

47.1.1. *Method 1: Adjoint Functors.* Taking colimits is left adjoint to the diagonal functor. Given a functor from the category of modules to the category of functors from J to modules, which simply assigns a module M to the constant functor mapping everything in J to M, the corresponding left adjoint functor is precisely the colimit. Because left adjoints are universally right exact, colimits inherently preserve right exactness.

47.1.2. *Method 2: Commuting Colimits.* The second method relies on the observation that colimits commute with colimits. A quotient is merely a special case of a colimit (specifically, the cokernel of a map), meaning colimits commute with quotients.

To understand why colimits commute with colimits, consider the product of two categories. We can take colimits over the first category (horizontally) and then over the second (vertically).

Alternatively, we can take colimits vertically, then horizontally. Or, we can compute a "super colimit" over the entire product category at once. Under suitable conditions, all three approaches yield the same isomorphic object. This commutativity acts as an algebraic analogue to Fubini's Theorem in integration theory:

$$\int \int f(x,y)\,dx\,dy = \int \left(\int f(x,y)\,dx \right) dy = \int \left(\int f(x,y)\,dy \right) dx$$

Because the colimit functor is right exact, it possesses left derived functors, which we can denote by colim_1. However, we will not delve into these derived functors here, as they are rarely used in standard commutative algebra.

47.2. **Failure of Left Exactness.** We naturally ask about left exactness: is a colimit of injective maps necessarily injective? The answer is generally no.

Example 47.2. Let J be a category consisting of two objects and two parallel morphisms. Consider the functor mapping these to the maps $\times 2$ and 0.

$$\mathbb{Z} \rightrightarrows \mathbb{Z}$$
$$\mathbb{Q} \rightrightarrows \mathbb{Q}$$

Because $\mathbb{Z} \subset \mathbb{Q}$, the natural inclusion maps are injective.

The colimit over this category is the cokernel of the difference of the two maps.

$$\varinjlim(\mathbb{Z} \rightrightarrows \mathbb{Z}) = \frac{\mathbb{Z}}{2\mathbb{Z}}$$
$$\varinjlim(\mathbb{Q} \rightrightarrows \mathbb{Q}) = \frac{\mathbb{Q}}{2\mathbb{Q}} = 0$$

While the components inject into each other, the induced map on the colimits is the map $\mathbb{Z}/2\mathbb{Z} \to 0$, which is manifestly not injective. Thus, colimits do not preserve exactness in general.

47.3. **Filtered Colimits.** While general colimits fail to be left exact, a crucial exception exists: taking colimits over a *filtered category*.

Theorem 47.3. *A colimit over a* filtered *category J preserves exactness.*

To see this, we utilize a special explicit construction for filtered colimits of modules. If M_i are modules over a filtered category J, their colimit can be constructed as the disjoint union of all the M_i modulo an equivalence relation:

$$\varinjlim_J M_i = \frac{\coprod_{i \in J} M_i}{\sim}$$

We define $m_i \in M_i$ and $m_j \in M_j$ to be equivalent ($m_i \sim m_j$) if there exists an object $k \in J$ and morphisms $i \to k$ and $j \to k$ such that m_i and m_j map to the exact same element in M_k.

The proof that $\sim$ is a valid equivalence relation heavily relies on J being filtered. For instance, to prove transitivity, if $m_i \sim m_j$ (meeting in M_k) and $m_j \sim m_l$ (meeting in M_n), the filtered property guarantees the existence of a further object M_p that receives maps from M_k and M_n. Pushing the elements forward into M_p demonstrates that $m_i \sim m_l$. Note that if J is not filtered (e.g., a discrete category with two objects), the colimit is simply a direct sum $M_1 \oplus M_2$, which cannot generally be expressed as a quotient of a disjoint union.

Proof of Exactness. Suppose we have injective maps $A_i \hookrightarrow B_i$ for all $i \in J$. Let $a \in \varinjlim A_i$ be an element whose image in $\varinjlim B_i$ is 0.

The element a is represented by some $a_i \in A_i$. Its image in B_i is some $b_i \in B_i$. Because b_i becomes 0 in the colimit $\varinjlim B_i$, the definition of our equivalence relation forces b_i to map to 0 in some specific module B_k further along the directed system.

By the functoriality of the system, we have a commutative square mapping $A_i \to B_i \to B_k$ and $A_i \to A_k \to B_k$.

$$\begin{array}{ccc} A_i & \longrightarrow & B_i \\ \downarrow & & \downarrow \\ A_k & \longrightarrow & B_k \end{array}$$

The element a_i maps to some $a_k \in A_k$, which maps to $0 \in B_k$. However, by assumption, the map $A_k \hookrightarrow B_k$ is injective. This forces $a_k = 0$ inside A_k.

Since a_i maps to 0 in A_k, its equivalence class in the colimit $\varinjlim A_i$ is 0. Thus, $a = 0$, proving that the map on colimits is injective. Consequently, filtered colimits are exact, and their derived functors (such as colim_1) identically vanish. □

Remark 47.4. An awful warning must be given regarding limits. While colimits are right exact and become completely exact over filtered categories, limits (inverse limits) are always left exact, but they generally fail to be right exact *even if the category is filtered*. We will study pathological examples of this failure in upcoming lectures.

47.4. **Colimits of Flat Modules.** We conclude with an essential application to flat modules.

Theorem 47.5. *A filtered colimit of flat modules is flat.*

Proof. Suppose $M = \varinjlim M_i$, where each M_i is a flat module, and the colimit is filtered. To show M is flat, we must show that tensoring M with any exact sequence $0 \to A \to B \to C \to 0$ preserves exactness.

Because each M_i is flat, the tensored sequence is exact for every i:

$$0 \to A \otimes M_i \to B \otimes M_i \to C \otimes M_i \to 0$$

We now take the filtered colimit of this entire system of exact sequences. As proven above, filtered colimits preserve exactness, so the resulting sequence remains exact:

$$0 \to \varinjlim(A \otimes M_i) \to \varinjlim(B \otimes M_i) \to \varinjlim(C \otimes M_i) \to 0$$

A foundational property of tensor products is that they commute with arbitrary colimits. Applying this commutativity yields:

$$0 \to A \otimes (\varinjlim M_i) \to B \otimes (\varinjlim M_i) \to C \otimes (\varinjlim M_i) \to 0$$

Substituting $M = \varinjlim M_i$, we obtain:

$$0 \to A \otimes M \to B \otimes M \to C \otimes M \to 0$$

Since this tensored sequence is exact, M is a flat module. □

Remark 47.6. It is crucial that the colimit is filtered; an unfiltered colimit of flat modules is generally not flat. For instance, any module N can be expressed as the cokernel of a map between free modules:

$$F_1 \to F_0 \to N \to 0$$

Free modules are inherently flat, so N is an unfiltered colimit (a cokernel) of flat modules. Yet, N need not be flat.

Incidentally, there is a powerful converse to our theorem, known as Lazard's Theorem. It states that a module is flat *if and only if* it is a filtered colimit of finitely generated free (and thus flat) modules. The proof of this theorem in Eisenbud's textbook contains a well-known minor gap, so interested readers are advised to consult Lazard's original paper for a complete and rigorous treatment.

48. Limits and Exactness

This lecture focuses on the limits and exactness of modules over a commutative ring. We analyze the conditions under which the inverse limit functor preserves exact sequences, mirroring our previous discussion on colimits. The primary application of this theory will be the construction and study of module completions.

48.1. **Left Exactness of Limits.** Suppose we are given an exact sequence of modules over a ring R indexed by a category J:

$$0 \to A_i \to B_i \to C_i \to 0$$

We wish to know if the inverse limit of this sequence remains exact.

First, the limit functor $\varprojlim$ is *left exact*. This means that taking the inverse limit automatically yields the exact sequence:

$$0 \to \varprojlim A_i \to \varprojlim B_i \to \varprojlim C_i$$

The proof of this is identical to the proof that the colimit functor is right exact, simply by reversing the direction of all arrows.

Consequently, the limit functor has a right derived functor, typically denoted by $\varprojlim^1$. If $\varprojlim^1 A_i = 0$, then the map $\varprojlim B_i \to \varprojlim C_i$ is surjective, rendering the entire sequence exact. (While higher derived functors of the limit exist, they

frequently encounter set-theoretic complications and are rarely used in ordinary commutative algebra).

48.2. **Failure of Right Exactness.** Just as colimits need not be left exact, limits need not be right exact.

Example 48.1. Consider the short exact sequence of $\mathbb{Z}$-modules:

$$0 \to \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

We construct an inverse system by taking an infinite number of copies of this sequence, where the transition maps between successive sequences are multiplication by 3:

$$\begin{aligned} A_{i+1} &\to A_i \quad \text{(multiplication by 3)} \\ B_{i+1} &\to B_i \quad \text{(multiplication by 3)} \\ C_{i+1} &\to C_i \quad \text{(multiplication by 3)} \end{aligned}$$

Let us compute the inverse limits of these sequences. For $\varprojlim A_i$, we seek a sequence of integers x_i such that $x_i = 3x_{i+1}$. This requires x_0 to be divisible by 3^n for every n, which is only possible if $x_0 = 0$. Thus, $\varprojlim A_i = 0$. By the exact same logic, $\varprojlim B_i = 0$.

However, for $C_i = \mathbb{Z}/2\mathbb{Z}$, multiplication by 3 is an isomorphism (it is equivalent to multiplication by 1). Thus, the inverse limit of isomorphisms yields $\varprojlim C_i \cong \mathbb{Z}/2\mathbb{Z}$.

Applying the limit to our sequence gives:

$$0 \to 0 \to 0 \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to 0$$

This sequence is clearly not exact on the right, demonstrating that the limit functor generally fails to be right exact. This failure highlights the necessity of finding conditions under which $\varprojlim^1 A_i$ vanishes.

48.3. **The Mittag-Leffler Condition.** The derived functor $\varprojlim^1 A_i$ vanishes if the system satisfies the *Mittag-Leffler condition*.

Definition 48.2. An inverse system of modules $A_0 \leftarrow A_1 \leftarrow A_2 \leftarrow \dots$ satisfies the *Mittag-Leffler condition* if the image of A_{i+j} in A_i stabilizes for sufficiently large j. That is, for each i, there exists an integer N such that for all $k \geq 0$:

$$\operatorname{im}(A_{i+N} \to A_i) = \operatorname{im}(A_{i+N+k} \to A_i)$$

Remark 48.3. Historically, this condition traces its algebraic origins back to Bourbaki's work on general topology for inverse systems of Hausdorff uniform spaces, though the name is borrowed from Mittag-Leffler's earlier, vaguely related work in complex analysis.

Our previous counterexample fails this condition: the image of $A_{i+j} \to A_i$ is $3^j\mathbb{Z}$, which decreases as j grows and never stabilizes.

48.4. **Proof of Exactness.**

Theorem 48.4. *If the inverse system A_i satisfies the Mittag-Leffler condition, then $\varprojlim^1 A_i = 0$. Consequently, if $0 \to A_i \to B_i \to C_i \to 0$ is a short exact sequence of inverse systems, the sequence of inverse limits is exact:*

$$0 \to \varprojlim A_i \to \varprojlim B_i \to \varprojlim C_i \to 0$$

Proof. We prove this in three stages. We assume we are given a coherent sequence $(c_i) \in \varprojlim C_i$, and our goal is to construct a sequence $(b_i) \in \varprojlim B_i$ mapping onto it.

Case 1: The maps $A_{i+1} \to A_i$ are surjective for all i. In this case, the Mittag-Leffler condition is trivially satisfied because the image is always A_i. Suppose we have successfully chosen an element $b_i \in B_i$ mapping to c_i. We must lift c_{i+1} to an element $b_{i+1} \in B_{i+1}$ that maps to b_i.

Pick any element $x \in B_{i+1}$ mapping onto c_{i+1}. Let $y \in B_i$ be the image of x. Because both y and b_i map to $c_i \in C_i$, their difference $b_i - y$ lies in the kernel of $B_i \to C_i$. By exactness, there exists $z \in A_i$ mapping to $b_i - y$.

Since the map $A_{i+1} \to A_i$ is surjective, we can lift z to some $w \in A_{i+1}$. We define our corrected lift as:

$$b_{i+1} = x + w$$

Then b_{i+1} correctly maps to c_{i+1} (since $w \in A_{i+1}$ maps to 0 in C_{i+1}), and its image in B_i is exactly $y + (b_i - y) = b_i$. By induction (and invoking the Axiom of Choice), we can construct the infinite sequence (b_i). Thus, the limit map is surjective.

Case 2: The maps $A_{i+j} \to A_i$ are zero for large j. By replacing the sequence with an appropriate subsequence, we may assume without loss of generality that $A_{i+1} \to A_i$ is the zero map for all i.

Given $(c_i) \in \varprojlim C_i$, we lift c_{i+1} to an arbitrary element $x \in B_{i+1}$. Let b_i be the image of x in B_i. We claim b_i is uniquely determined and independent of the choice of x. If x' is another lift of c_{i+1}, the difference $x - x'$ maps to 0 in C_{i+1}, and thus is the image of some $y \in A_{i+1}$. Because the transition map $A_{i+1} \to A_i$ is zero, the image of y in B_i is zero. Therefore, x and x' map to the identical element b_i.

This constructs a unique, canonical sequence $(b_i) \in \varprojlim B_i$, proving surjectivity in this case.

Case 3: The general Mittag-Leffler case. Let $A'_i \subseteq A_i$ denote the stable image of A_{i+j} for large j. We form a short exact sequence of inverse systems:

$$0 \to A'_i \to A_i \to \frac{A_i}{A'_i} \to 0$$

By construction, the system (A'_i) has surjective transition maps, satisfying Case 1. The quotient system (A_i/A'_i) has transition maps that eventually become zero, satisfying Case 2.

Applying the long exact sequence for derived limit functors, we obtain:

$$\varprojlim{}^1 A_i' \to \varprojlim{}^1 A_i \to \varprojlim{}^1 \left(\frac{A_i}{A_i'}\right)$$

From Cases 1 and 2, the first and third terms vanish. Therefore, the middle term $\varprojlim^1 A_i$ must also equal 0, completing the proof. □

48.5. **Examples and Applications.**

Example 48.5. Suppose the modules A_i are all finite sets. Any decreasing sequence of submodules must stabilize simply because the sets are finite. Thus, any inverse system of finite modules inherently satisfies the Mittag-Leffler condition, guaranteeing that taking their limit preserves exactness. This similarly applies to systems of Artinian modules, as they satisfy the descending chain condition.

The primary application of this theorem is in the study of module completions. The completion of a module M with respect to an ideal I is defined as the inverse limit:

$$\hat{M} = \varprojlim_n \frac{M}{I^n M}$$

If we have a short exact sequence $0 \to A \to B \to C \to 0$, taking completions yields an exact sequence $0 \to \hat{A} \to \hat{B} \to \hat{C} \to 0$ provided the associated systems satisfy the requisite Mittag-Leffler condition, a property we will rely upon extensively in the next lecture.

49. Completions

This section introduces the concept of the completion of a ring, studying its definition, fundamental examples, and its geometric relationship to localization.

49.1. **Definition of Completion.** Suppose we have a commutative ring R and an ideal $I \subseteq R$.

Definition 49.1. The *completion* of R with respect to the ideal I, denoted $\hat{R}$, is the inverse limit (or projective limit) of the quotient rings R/I^n:

$$\hat{R} = \varprojlim_n \frac{R}{I^n}$$

This means we consider the sequence of natural quotient maps:

$$\cdots \to \frac{R}{I^3} \to \frac{R}{I^2} \to \frac{R}{I}$$

An element of the completion $\hat{R}$ is a sequence $(a_0, a_1, a_2, \dots)$ where $a_n \in R/I^{n+1}$, such that each a_n maps to a_{n-1} under the natural projection.

49.2. **Basic Examples.**

Example 49.2 (Formal Power Series)**.** Let $R = k[x]$ be the polynomial ring over a field, and let $I = (x)$. The quotients are:

$$\frac{R}{I} = k$$
$$\frac{R}{I^2} = \frac{k[x]}{(x^2)}$$

An element in R/I^2 can be uniquely represented as $a_0 + a_1 x$. Moving up the sequence, an element in the inverse limit is an infinitely long polynomial, meaning the completion is the ring of formal power series:

$$\hat{R} = k[[x]]$$

Similarly, completing the multivariable polynomial ring $k[x, y]$ with respect to the ideal (x, y) yields the formal power series ring $k[[x, y]]$.

Example 49.3 (p-adic and 10-adic Integers)**.** Let $R = \mathbb{Z}$ and $I = (10)$. The completion is the inverse limit of $\mathbb{Z}/10^n\mathbb{Z}$, which forms the ring of 10-*adic integers*, denoted $\hat{\mathbb{Z}}_{10}$. Its elements can be thought of as infinite decimal sequences extending infinitely to the left of the decimal point.

By the Chinese Remainder Theorem, the quotients split:

$$\frac{\mathbb{Z}}{10^n\mathbb{Z}} \cong \frac{\mathbb{Z}}{2^n\mathbb{Z}} \times \frac{\mathbb{Z}}{5^n\mathbb{Z}}$$

Taking the inverse limit preserves this product:

$$\hat{\mathbb{Z}}_{10} \cong \hat{\mathbb{Z}}_2 \times \hat{\mathbb{Z}}_5$$

Because $\hat{\mathbb{Z}}_{10}$ decomposes into a direct product, it inherently possesses zero divisors. Thus, one typically focuses on completing $\mathbb{Z}$ at prime ideals (p) to obtain the p-adic integers $\hat{\mathbb{Z}}_p$, which remain integral domains. The ring of p-adic integers serves as the number-theoretic analogue to the formal power series ring $k[[x]]$.

49.3. **Metric Construction of Completions.** Alternatively, completion can be constructed analytically via a metric. We define an absolute value on R with respect to the ideal I:

$$|x| = \begin{cases} c^{-n} & \text{if } x \in I^n \setminus I^{n+1} \text{ for a fixed } c > 1 \\ 0 & \text{if } x \in \bigcap_n I^n \text{ (e.g., } x = 0) \end{cases}$$

This defines a distance metric $d(x, y) = |x - y|$. The completion of R as a Cauchy metric space exactly coincides with the inverse limit definition. Furthermore, this distance function satisfies the strong *ultrametric inequality*:

$$|x + y| \leq \max(|x|, |y|)$$
$$d(x, y) \leq \max(d(x, z), d(z, y))$$

49.4. **Three Themes of Completions.** The study of completions generally falls into three recurring themes:

(1) **Analytic similarity:** Completions allow for analytic techniques. In the p-adic integers, one can define generalized operations such as exponentiation, Gamma functions, and Bessel functions.
(2) **Equation solving:** It is remarkably easy to lift approximate solutions of equations to exact solutions in complete rings, a property governed by Hensel's Lemma.
(3) **A stronger localization:** Completion acts as a more severe, localized focus on a point than standard localization.

49.5. **Completion as a Stronger Localization.** We expand on the third theme. Localizing focuses on a point by making elements outside the corresponding prime ideal invertible. Completion does this and more.

Suppose an element $r \in R$ acts as a unit modulo a prime power P^n. It automatically remains a unit modulo P^{2n}. This can be seen algebraically: if $1 - rs \in P^n$, then:

$$\begin{aligned} 1 - (1 - rs)^2 &\in P^{2n} \\ rs(2 - rs) &\equiv 1 \pmod{P^{2n}} \end{aligned}$$

Thus, r is invertible in R/P^{2n}. Because these inverses lift compatibly, r has a unique inverse in the completion $\hat{R}$.

Consequently, if $\mathfrak{m}$ is a maximal ideal, the completion $\hat{R}$ at $\mathfrak{m}$ is automatically a local ring. Elements not in $\mathfrak{m}$ are units in $R/\mathfrak{m}$ (a field), and thus lift to units in $\hat{R}$. This creates a natural chain of homomorphisms:

$$R \to R_{\mathfrak{m}} \to \hat{R} \to \frac{R}{\mathfrak{m}^k} \cdots \to \frac{R}{\mathfrak{m}}$$

49.5.1. *Pathologies.* This behavior relies heavily on completing at a *maximal* ideal in a *Noetherian* ring.

- **Non-maximal ideals:** If we complete $\mathbb{Z}$ at the prime ideal (0), we get $\hat{\mathbb{Z}} = \mathbb{Z}$, which is not a local ring. The localization $\mathbb{Q}$ does not naturally map into this completion.
- **Non-Noetherian rings:** Let $R = k[x^{1/n} \mid n \geq 1]$ be the ring of Puiseux polynomials, and $\mathfrak{m}$ be the ideal generated by all positive powers of x. Because $\mathfrak{m} = \mathfrak{m}^2$, all higher quotients collapse: $\mathfrak{m}^n = \mathfrak{m}$. The completion reduces to the base field k. The localization $R_{\mathfrak{m}}$ does not embed into k.

49.6. **Geometric Interpretation and Spectra.** Geometrically, the spectrum $\operatorname{Spec}(R)$ represents the global space. Localizing at $\mathfrak{m}$ creates $\operatorname{Spec}(R_{\mathfrak{m}})$, which geometrically restricts our view to a neighborhood of $\mathfrak{m}$ by throwing away unrelated points.

Passing to the completion $\operatorname{Spec}(\hat{R})$ acts as an "infinitesimal neighborhood" that zooms in so closely that global connectivity can shatter.

Example 49.4 (Reducibility of Completions)**.** Let R be the coordinate ring of a nodal curve over a field of characteristic zero:

$$R = \frac{k[x,y]}{(y^2 - x^3 - x^2)}$$

The localization at the origin $\mathfrak{m} = (x,y)$ remains an integral domain because the curve is globally irreducible. However, completing at the origin yields:

$$\hat{R} = \frac{k[[x,y]]}{(y^2 - x^2(1+x))}$$

In the formal power series ring $k[[x]]$, the term $\sqrt{1+x}$ has a valid Taylor expansion. Thus, the defining equation factors:

$$y^2 - x^2(1+x) = \left(y - x\sqrt{1+x}\right)\left(y + x\sqrt{1+x}\right)$$

Because the equation factors, the completed ring $\hat{R}$ contains zero divisors. Geometrically, taking the completion has zoomed in so deeply on the intersection point that the single nodal curve visibly splits into two disconnected branches. This proves that the completion of an integral domain need not be an integral domain.

49.7. **Visualizing the 2-adic Integers.** To physically visualize a completion like the 2-adic integers $\hat{\mathbb{Z}}_2 = \varprojlim \mathbb{Z}/2^n\mathbb{Z}$, we construct a branching tree.

- The last digit (mod 2) is either 0 or 1.
- If it is 0, the previous digit limits the number to either 00 or 10 (mod 4).
- Each subsequent digit choice branches the possibilities into two disjoint disks.

Taking the inverse limit equates to taking the infinite intersection of these nested topological disks. The resulting space is perfectly homeomorphic to the Cantor set, or equivalently, the topological product 2^∞.

50. Hensel's Lemma

This section focuses on Hensel's Lemma, a powerful tool for finding solutions to equations over completions of rings.

50.1. **Introduction.** Recall that the completion of a ring R with respect to an ideal I is defined as the inverse limit:

$$\hat{R} = \varprojlim_n \frac{R}{I^n}$$

A typical example is the ring of p-adic integers, $\mathbb{Z}_p$, which is the inverse limit of $\mathbb{Z}/p^n\mathbb{Z}$.

A central problem in algebra is determining whether a polynomial equation has roots. Consider the polynomial $f(x) = x^3 - 5x - 2 = 0$.

- Does it have a root in the rationals $\mathbb{Q}$? This requires some algebraic effort to answer.

- Does it have a root in the reals $\mathbb{R}$? The answer is immediately obvious: $f(x)$ is positive for large positive x and negative for large negative x. By the intermediate value theorem, it must have a real root.

It is often much easier to show that polynomials have roots over complete fields like $\mathbb{R}$ than over $\mathbb{Q}$. Because completions of rings are also complete in a topological sense, they possess a property analogous to the intermediate value theorem. Hensel's Lemma formalizes this: if we can solve an equation modulo I^n for a suitable n, we can lift this to a solution in the full completion $\hat{R}$.

50.2. **Hensel's Lemma: Version 1.**

Theorem 50.1 (Hensel's Lemma, Version 1)**.** *Suppose I is an ideal of a ring R with completion $\hat{R}$. Let $f(x)$ be a polynomial with coefficients in $\hat{R}$. Suppose there exists a root $a \in R/I$ such that:*

$$f(a) \equiv 0 \pmod{I}$$

If the formal derivative $f'(a)$ is invertible in R/I, then a can be lifted to a root in $\hat{R}$.

In many applications, I is a maximal ideal, so R/I is a field. In this case, the invertibility condition simply means $f'(a) \not\equiv 0 \pmod{I}$, which geometrically states that a is a *simple root* of f, rather than a double root or a root of higher order.

Proof. It suffices to show that any root in R/I^n can be lifted to a root in R/I^{n+1}. By starting with a root in R/I and iterating this process, we obtain a coherent sequence of roots in $R/I^2, R/I^3, \ldots$, which defines a unique root in the inverse limit $\hat{R}$.

Suppose a is a root modulo I^n, meaning $f(a) \in I^n$. We wish to find a perturbed root $a + \epsilon$ such that:

$$f(a + \epsilon) \equiv 0 \pmod{I^{n+1}}$$

Expanding via Taylor series, we have:

$$f(a + \epsilon) = f(a) + \epsilon f'(a) + \epsilon^2 \frac{f''(a)}{2} + \ldots$$

Since we want this to vanish modulo I^{n+1}, and ϵ will be drawn from I^n, any terms involving ϵ^2 or higher will inherently belong to $I^{2n} \subseteq I^{n+1}$. Thus, they vanish modulo I^{n+1}, leaving the linear approximation:

$$f(a) + \epsilon f'(a) \equiv 0 \pmod{I^{n+1}}$$

Solving for ϵ suggests:

$$\epsilon = -f(a)f'(a)^{-1}$$

Since $f'(a)$ is a unit in R/I, there exists some element $f'(a)^{-1}$ such that $f'(a)f'(a)^{-1} \equiv 1 \pmod{I}$. Because $f(a) \in I^n$, multiplying it by this approximate inverse yields an element ϵ that is well-defined modulo I^{n+1} and lies in I^n. Substituting this ϵ back into the expansion successfully lifts the root to R/I^{n+1}. □

50.3. **Examples in p-adic Integers.**

Example 50.2. Does 7 have a square root in the 3-adic integers $\mathbb{Z}_3$? Let $f(x) = x^2 - 7$, so $f'(x) = 2x$.

We first look for a root modulo 3. The equation $x^2 \equiv 7 \equiv 1 \pmod 3$ has a solution $a = 1$. We check the derivative condition:

$$f'(1) = 2(1) = 2 \not\equiv 0 \pmod 3$$

Because the derivative is invertible modulo 3, the root $a = 1$ lifts to a square root of 7 in $\mathbb{Z}_3$. Following the iterative proof construction, we can compute the base-3 digits of the root step-by-step: 1, then $1 + 1 \cdot 3 = 4 \pmod 9$, then $4 + 1 \cdot 9 = 13 \pmod{27}$, and so on.

Example 50.3 (Squares in $\mathbb{Z}_p^\times$ for odd p)**.** We can determine the structure of the group of units in $\mathbb{Z}_p$ modulo the squares. For an element $b \in \mathbb{Z}_p$ to be a square, it must satisfy two conditions:

(1) The number of zeros at the end of its p-adic expansion must be even (otherwise the p-adic valuation would be odd, which cannot happen for a square). We can factor out these even powers of p to assume b is a unit.
(2) The first non-zero digit must be a quadratic residue modulo p. That is, the equation $x^2 \equiv b \pmod p$ must have a solution.

If these conditions hold, we have $f(x) = x^2 - b \equiv 0 \pmod p$. The derivative is $f'(x) = 2x$. Because p is odd, $2 \not\equiv 0 \pmod p$, meaning the derivative evaluated at the root is invertible modulo p. By Hensel's Lemma, the root lifts to $\mathbb{Z}_p$.

Consequently, the group structure of the units modulo squares is:

$$\frac{\mathbb{Z}_p^\times}{(\mathbb{Z}_p^\times)^2} \cong \frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}}$$

where one factor accounts for the parity of the valuation, and the other accounts for the Legendre symbol modulo p.

50.4. **Hensel's Lemma for $p = 2$ and Newton's Method.** The previous argument breaks down completely if $p = 2$. If we attempt to solve $x^2 - a = 0$ in the 2-adic integers $\mathbb{Z}_2$, the derivative is $f'(x) = 2x$. Evaluated at any integer, this derivative is $\equiv 0 \pmod 2$. Thus, $f'(x)$ is never a unit in $\mathbb{Z}_2/2\mathbb{Z}_2$, and the standard version of Hensel's Lemma cannot be applied.

For instance, $a = 1$ is a root of $x^2 \equiv 5 \pmod 4$, but there is no solution to $x^2 \equiv 5 \pmod 8$. We need a refinement of Hensel's Lemma that can handle cases where the root is not perfectly simple modulo I.

The refinement relies on Newton's method. Recall from real analysis that to find a root of $f(x)$, we can iterate:

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

Over the reals, Newton's method can be unstable. Over the p-adics, however, it behaves remarkably well.

Theorem 50.4 (Refined Hensel's Lemma)**.** *Suppose we have a polynomial $f(x)$ with coefficients in $\mathbb{Z}_p$. If there exists an approximate root $a \in \mathbb{Z}_p$ and an integer $d \geq 0$ such that:*

$$f(a) \equiv 0 \pmod{p^{2d+1}}$$
$$f'(a) \not\equiv 0 \pmod{p^{d+1}}$$

then a can be lifted to a true root in $\mathbb{Z}_p$.

Notice that if $d = 0$, this reduces perfectly to the first version of Hensel's lemma (the function vanishes mod p, and the derivative does not vanish mod p).

Proof. We apply a single step of Newton's method. Let $x_n = a$ be our approximate root. We expand $f(x_{n+1})$ using the Taylor series around x_n:

$$\begin{aligned} f(x_{n+1}) &= f\left(x_n - \frac{f(x_n)}{f'(x_n)}\right) \\ &= f(x_n) - f'(x_n)\frac{f(x_n)}{f'(x_n)} + \frac{f''(x_n)}{2!}\left(\frac{f(x_n)}{f'(x_n)}\right)^2 + \dots \end{aligned}$$

The first two terms exactly cancel: $f(x_n) - f(x_n) = 0$. The remaining terms are bounded by the behavior of the quadratic term.

We are given that $f(x_n)$ is divisible by p^{2d+1}, and $f'(x_n)$ is divisible by at most p^d. Therefore, the fraction $f(x_n)/f'(x_n)$ is divisible by at least p^{d+1}. Squaring this fraction means it is divisible by p^{2d+2}.

(Note: The term $f''(x_n)/2!$ does not introduce problematic denominators because the formal second derivative of x^n divided by 2 is $\binom{n}{2}x^{n-2}$, which naturally has integer coefficients).

Consequently, all remaining terms in the Taylor expansion are divisible by p^{2d+2}. Thus, iterating Newton's method yields a sequence of approximations where the number of correct digits essentially doubles at each step:

$$f(x_n) \equiv 0 \pmod{p^{2d+k}} \implies f(x_{n+1}) \equiv 0 \pmod{p^{2d+2k}}$$

This rapid convergence guarantees a unique root in the limit. □

50.5. **Applications to $\mathbb{Z}_2$.**

Example 50.5 (Squares in $\mathbb{Z}_2$)**.** When does an odd integer $b \in \mathbb{Z}_2$ have a square root? Let $f(x) = x^2 - b$. The derivative is $f'(x) = 2x$.

For an odd x, $f'(x) = 2x$ is exactly divisible by 2^1. Thus, we must set $d = 1$ so that $f'(x) \not\equiv 0 \pmod{2^{d+1}} = 4$.

According to the refined Hensel's Lemma, to guarantee a lift to $\mathbb{Z}_2$, we require the function to vanish modulo 2^{2d+1}. Since $2(1) + 1 = 3$, we must find a root modulo $2^3 = 8$:

$$x^2 - b \equiv 0 \pmod{8}$$

Therefore, an odd 2-adic integer b is a square if and only if $b \equiv 1 \pmod{8}$.

Example 50.6 (Fourth roots in $\mathbb{Z}_2$)**.** When does an odd integer $b \in \mathbb{Z}_2$ have a fourth root? Let $f(x) = x^4 - b$. The derivative is $f'(x) = 4x^3$.

For an odd x, $f'(x) = 4x^3$ is exactly divisible by $2^2 = 4$. Thus, we must set $d = 2$ so that $f'(x) \not\equiv 0 \pmod{2^{d+1}} = 8$.

Hensel's Lemma requires a root modulo 2^{2d+1}. Since $2(2) + 1 = 5$, we require a root modulo $2^5 = 32$:

$$x^4 - b \equiv 0 \pmod{32}$$

Checking the odd fourth powers modulo 32, we find they are exactly 1 and 17. Notice that both 1 and 17 are congruent to 1 (mod 16). Thus, by numerical coincidence, an odd 2-adic integer has a fourth root if and only if $b \equiv 1 \pmod{16}$.

51. Hensel's Lemma Continued

This lecture serves as a continuation of our discussion on Hensel's lemma. We will apply the lemma to determine the structure of the group of units of the p-adic integers, introduce a generalized version of Hensel's lemma for lifting factorizations, and conclude with a discussion on Henselian local rings.

51.1. **The Group of Units of the p-adic Integers.** Recall that Hensel's lemma states that if R is a local ring with maximal ideal I such that R/I is a field, and $f(x)$ is a monic polynomial over R, then any simple root a of $f(x)$ in R/I can be lifted to a root in the completion $\hat{R}$.

We will use this to analyze the structure of the group of units of the p-adic integers, $\mathbb{Z}_p^\times$. The ring of p-adic integers $\mathbb{Z}_p$ is the completion of the integers $\mathbb{Z}$ at the prime ideal (p). It can be viewed as the inverse limit:

$$\mathbb{Z}_p = \varprojlim_n \frac{\mathbb{Z}}{p^n\mathbb{Z}}$$

Elements of $\mathbb{Z}_p$ can be thought of as numbers written in base p that extend infinitely to the left.

To find the structure of the unit group $\mathbb{Z}_p^\times$, we first identify the roots of unity within it. The finite field $\mathbb{Z}/p\mathbb{Z}$ has $p-1$ non-zero elements, which form a cyclic group of order $p-1$. Consequently, every non-zero element satisfies the polynomial equation:

$$x^{p-1} - 1 = 0 \pmod{p}$$

Because the roots of this polynomial are precisely the $p-1$ distinct non-zero elements of $\mathbb{Z}/p\mathbb{Z}$, they are all simple roots modulo p. By Hensel's lemma, each of these $p-1$ roots lifts uniquely to a root of $x^{p-1} - 1 = 0$ in $\mathbb{Z}_p$.

Thus, $\mathbb{Z}_p^\times$ contains exactly $p-1$ roots of unity, which are congruent to $1, 2, \ldots, p-1$ modulo p. This provides us with a split direct factor:

$$\mathbb{Z}_p^\times \cong \frac{\mathbb{Z}}{(p-1)\mathbb{Z}} \times (1 + p\mathbb{Z}_p)$$

The first factor represents the roots of unity, and the second factor consists of elements congruent to 1 modulo p. We are now reduced to determining the

structural group of $1 + p\mathbb{Z}_p$ under multiplication. We must distinguish between two cases: p odd and $p = 2$.

51.1.1. *The Case p is Odd.* If p is an odd prime, the multiplicative group $1+p\mathbb{Z}_p$ is isomorphic to the additive group $p\mathbb{Z}_p$. Because $p\mathbb{Z}_p$ is just a scaled copy of the p-adic integers, it is isomorphic to $\mathbb{Z}_p$ under addition.

This isomorphism is established via the formal exponential and logarithm maps:

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$
$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots$$

We must be careful regarding the convergence of these series in the p-adic topology. In $\mathbb{R}$, the factorial in the denominator grows rapidly, ensuring convergence. In $\mathbb{Z}_p$, however, dividing by p makes a number *larger* in the p-adic metric. We must ensure that the powers of x in the numerator accumulate factors of p faster than the factorials in the denominator strip them away.

Let us count the number of factors of p dividing $n!$, denoted $v_p(n!)$. Using Legendre's formula:

$$\begin{aligned} v_p(n!) &= \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \dots \\ &\leq \frac{n}{p} + \frac{n}{p^2} + \frac{n}{p^3} + \dots \\ &= \frac{n}{p-1} \end{aligned}$$

Suppose $x \in p\mathbb{Z}_p$, meaning p divides x. Then the power of p dividing x^n is at least n. The p-adic valuation of the general term in the exponential series is:

$$\begin{aligned} v_p\left(\frac{x^n}{n!}\right) &= v_p(x^n) - v_p(n!) \\ &\geq n - \frac{n}{p-1} \\ &= n\left(1 - \frac{1}{p-1}\right) \end{aligned}$$

Because $p > 2$, we have $p - 1 \geq 2$, which implies $1 - \frac{1}{p-1} > 0$. Therefore, as $n \to \infty$, the p-adic valuation of the general term goes to infinity, meaning the terms converge to 0. Thus, the exponential series converges for any $x \in p\mathbb{Z}_p$. The logarithm series is even easier to verify and converges under the same conditions.

Because these maps are mutually inverse isomorphisms between $(p\mathbb{Z}_p, +)$ and $(1 + p\mathbb{Z}_p, \times)$, we obtain the final structure for odd p:

$$\mathbb{Z}_p^\times \cong \frac{\mathbb{Z}}{(p-1)\mathbb{Z}} \times \mathbb{Z}_p$$

Notice that the first factor is the torsion subgroup (finite order), while the second factor $\mathbb{Z}_p$ is torsion-free. This strongly parallels the structure of the real units $\mathbb{R}^\times \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{R}_{>0}$.

51.1.2. *The Case $p = 2$.* If $p = 2$, the convergence argument fails for $x \in 2\mathbb{Z}_2$. The bound $v_2(n!) \leq n/(2-1) = n$ means the denominator can accumulate factors of 2 at the same rate as x^n. To force convergence, we must require x to be divisible by $p^2 = 4$.

Thus, for $p = 2$, the exponential map only establishes an isomorphism between the additive group $4\mathbb{Z}_2$ and the multiplicative group $1 + 4\mathbb{Z}_2$. The elements congruent to 1 (mod 2) split further:

$$\mathbb{Z}_2^\times \cong \{\pm 1\} \times (1 + 4\mathbb{Z}_2)$$

Using the isomorphism $1 + 4\mathbb{Z}_2 \cong 4\mathbb{Z}_2 \cong \mathbb{Z}_2$, we get the structure for $p = 2$:

$$\mathbb{Z}_2^\times \cong \frac{\mathbb{Z}}{2\mathbb{Z}} \times \mathbb{Z}_2$$

The prime 2 behaves slightly differently, yielding more roots of unity (a factor of 4 rather than 2 in the internal filtration) than the general formula for odd primes would suggest.

51.2. **Lifting Factorizations (Hensel's Lemma, Version 2).** There are many variations of Hensel's lemma in the literature. The version we have used lifts a simple root. A powerful generalization lifts coprime factorizations of polynomials.

Theorem 51.1. *Let R be a complete local ring with maximal ideal I. Let $f(x)$ be a monic polynomial in $R[x]$. Suppose that modulo I, $f(x)$ factorizes as:*

$$f(x) \equiv g_0(x)h_0(x) \pmod{I}$$

where $g_0(x)$ and $h_0(x)$ are monic and coprime *in $(R/I)[x]$. Then this factorization lifts to a global factorization in $R[x]$:*

$$f(x) = g(x)h(x)$$

such that $g(x) \equiv g_0(x) \pmod{I}$ and $h(x) \equiv h_0(x) \pmod{I}$.

If we take $g_0(x)$ to be a linear polynomial $x - a$, this theorem immediately recovers our previous version of Hensel's lemma. Saying $x - a$ and $h_0(x)$ are coprime simply means that a is not a root of $h_0(x)$, which is precisely the statement that a is a simple root of $f(x)$.

Proof Sketch. The proof proceeds by induction, iteratively lifting the factorization modulo higher powers of I.

Suppose we have successfully found polynomials $g_n(x)$ and $h_n(x)$ such that:

$$f(x) \equiv g_n(x)h_n(x) \pmod{I^n}$$

We want to find correction terms $a(x)$ and $b(x)$ in $I^n[x]$ to form the next lift:

$$g_{n+1}(x) = g_n(x) + a(x)$$
$$h_{n+1}(x) = h_n(x) + b(x)$$

such that $f(x) \equiv g_{n+1}(x)h_{n+1}(x) \pmod{I^{n+1}}$. Expanding the product yields:

$$\begin{aligned} g_{n+1}(x)h_{n+1}(x) &= (g_n + a)(h_n + b) \\ &= g_n h_n + g_n b + h_n a + ab \end{aligned}$$

Because a and b are in I^n, their product ab is in $I^{2n} \subseteq I^{n+1}$. We can discard it modulo I^{n+1}. Setting this equal to $f(x)$ gives the required linear equation for the corrections:

$$g_n(x)b(x) + h_n(x)a(x) \equiv f(x) - g_n(x)h_n(x) \pmod{I^{n+1}} \tag{51.1}$$

The right-hand side is some known polynomial with coefficients in I^n, which we can refer to as a "complicated mess". We must solve for $a(x)$ and $b(x)$.

Because g_0 and h_0 are coprime over the field R/I, the polynomial ring $(R/I)[x]$ is a Principal Ideal Domain. Thus, by Bezout's identity, there exist polynomials $c(x)$ and $d(x)$ such that:

$$g_0(x)c(x) + h_0(x)d(x) \equiv 1 \pmod{I}$$

To solve equation 51.1, we simply multiply our Bezout identity by the "mess" $(f - g_n h_n)$. We assign:

$$b(x) = c(x) \cdot \text{mess} \pmod{I^{n+1}}$$
$$a(x) = d(x) \cdot \text{mess} \pmod{I^{n+1}}$$

This explicitly constructs the higher lifts. Because R is complete, this infinite sequence of polynomial approximations converges to exact polynomials $g(x)$ and $h(x)$ in $R[x]$. □

51.3. **Henselian Local Rings.** We conclude by formalizing the environment in which these theorems operate.

Definition 51.2. A *Henselian local ring* is a local ring $(R, \mathfrak{m})$ in which Hensel's lemma holds. Specifically, it is a ring where any monic polynomial $f \in R[x]$ that factors modulo $\mathfrak{m}$ into coprime monic factors can be factored in $R[x]$ lifting the residual factorization.

As we have seen, any complete local ring (such as $k[[x]]$ or $\mathbb{Z}_p$) is automatically a Henselian ring. However, completions are often geometrically massive. If R is countable, its completion $\hat{R}$ is usually uncountable, containing formal limits that obscure the purely algebraic geometry.

Nagata demonstrated that for any local ring R, there exists a *Henselization*, which is roughly the smallest Henselian local ring containing R. You can construct a tower of inclusions:

$$R \subseteq R^h \subseteq \hat{R}$$

The Henselization R^h can be conceptually understood as the "algebraic part" of the completion. For instance, if R is the localization of a polynomial ring $k[x_1, \dots, x_n]$, its completion is the formal power series ring $k[[x_1, \dots, x_n]]$, but its Henselization is the ring of *algebraic power series* (power series that satisfy polynomial equations over R).

This is deeply analogous to the relationship between the algebraic numbers and the complex numbers. The rationals $\mathbb{Q}$ sit inside the algebraically closed field $\mathbb{C}$, but there is a much smaller algebraic closure $\bar{\mathbb{Q}}$ nested between them:

$$\mathbb{Q} \subseteq \bar{\mathbb{Q}} \subseteq \mathbb{C}$$

Definition 51.3. A *strictly Henselian ring* is a Henselian local ring whose residue field $R/\mathfrak{m}$ is separably closed.

Strictly Henselian rings are fundamental objects in modern algebraic geometry. In the standard Zariski topology, the local ring at a point $\mathcal{O}_{X,p}$ is formed by taking the direct limit of the coordinate rings of all Zariski-open neighborhoods. However, the Zariski topology is often too coarse. When working with the *étale topology* (a Grothendieck topology that allows for finite unramified coverings to act as "open neighborhoods"), the appropriate analogue of the local ring at a point is precisely a strictly Henselian local ring.

52. Flatness of Completions

In this lecture, we will investigate the flatness of completions. To motivate this, recall that for a commutative ring R and a prime ideal $\mathfrak{p}$, we can construct the localization $R_\mathfrak{p}$. We previously proved that $R_\mathfrak{p}$ is a flat R-module using two main steps:

(1) The operation of localization $M \mapsto M_\mathfrak{p}$ preserves exactness.
(2) There is a natural isomorphism between the localization and the tensor product:

$$M_\mathfrak{p} \cong M \otimes_R R_\mathfrak{p}$$

Combining these two facts showed that tensoring with $R_\mathfrak{p}$ preserves exactness, meaning $R_\mathfrak{p}$ is a *flat* module.

We can perform a similar construction by taking the completion of R at an ideal I, defined as the inverse limit:

$$\hat{R} = \varprojlim_n \frac{R}{I^n}$$

Completion is, in many ways, a "stronger version" of localization. Our goal in this lecture is to show that if R is a Noetherian ring, then $\hat{R}$ is a flat R-module. We will attempt to imitate the proof used for localization, but we will encounter several complications because taking the completion of a module does not always preserve exactness, and it does not always commute with the tensor product.

52.1. Exactness of Completion for Finitely Generated Modules.

The first step is to establish the conditions under which taking the completion of a module preserves exactness. This is the analogue of the corresponding result for localization.

Lemma 52.1. *Let R be a Noetherian ring and let*

$$0 \to A \to B \to C \to 0$$

be an exact sequence of finitely generated R-modules. Then the completed sequence

$$0 \to \hat{A} \to \hat{B} \to \hat{C} \to 0$$

is also exact.

Before proving this lemma, let us demonstrate that it fails if the modules are not finitely generated.

Example 52.2. Consider the following exact sequence of $\mathbb{Z}$-modules:

$$0 \to \mathbb{Z} \to \mathbb{Q} \to \frac{\mathbb{Q}}{\mathbb{Z}} \to 0$$

Let us take the completion with respect to the prime ideal (2).

- The completion of $\mathbb{Z}$ at (2) is the ring of 2-adic integers, $\hat{\mathbb{Z}} = \mathbb{Z}_2$.
- If we complete the rational numbers $\mathbb{Q}$ at (2), we get 0, because $\mathbb{Q} = 2^n\mathbb{Q}$ for all n, meaning the quotients $\mathbb{Q}/2^n\mathbb{Q}$ are all 0. Thus, $\hat{\mathbb{Q}} = 0$.
- Similarly, the completion of $\mathbb{Q}/\mathbb{Z}$ at (2) is 0.

The completed sequence becomes:

$$0 \to \mathbb{Z}_2 \to 0 \to 0 \to 0$$

This sequence is manifestly not exact because the map from $\mathbb{Z}_2$ to 0 is not injective. Thus, completion does not generally preserve exactness for modules that are not finitely generated.

Proof of Lemma 52.1. We assume the sequence $0 \to A \to B \to C \to 0$ is exact. For any $n \geq 1$, we can construct the following sequence:

$$0 \to \frac{A}{I^nB \cap A} \to \frac{B}{I^nB} \to \frac{C}{I^nC} \to 0$$

This sequence is exact. However, the first term looks somewhat unusual. One might expect it to be A/I^nA. If we replace the first term with A/I^nA, the sequence is not necessarily exact. For instance, consider the exact sequence $0 \to \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$ over $R = \mathbb{Z}$, and let $I = (2)$. Quotienting by I^1 naively yields $0 \to \mathbb{Z}/2\mathbb{Z} \xrightarrow{0} \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$, which loses injectivity on the left. Thus, we must use the expression $A/(I^nB \cap A)$.

We now take the inverse limit of these exact sequences. Recall that the inverse limit preserves exactness if the Mittag-Leffler condition holds. For the sequence of modules $A/(I^nB \cap A)$, the transition maps:

$$\frac{A}{I^{n+1}B \cap A} \to \frac{A}{I^nB \cap A}$$

are clearly surjective. Systems with surjective transition maps automatically satisfy the Mittag-Leffler condition. Therefore, taking the inverse limit yields an exact sequence:

$$0 \to \varprojlim \left(\frac{A}{I^nB \cap A}\right) \to \varprojlim \left(\frac{B}{I^nB}\right) \to \varprojlim \left(\frac{C}{I^nC}\right) \to 0$$

The middle term is precisely the completion $\hat{B}$, and the rightmost term is the completion $\hat{C}$.

We are left with identifying the leftmost term. Is the inverse limit of $A/(I^nB \cap A)$ equal to the completion $\hat{A}$? The completion of A is defined as:

$$\hat{A} = \varprojlim \left(\frac{A}{I^n A} \right)$$

By the Artin-Rees Lemma, since A is a submodule of a finitely generated module B over a Noetherian ring R, the filtration $I^nB \cap A$ is a *stable I-filtration*. This stability condition implies that the topology induced by the filtration $I^nB \cap A$ is exactly the same as the intrinsic I-adic topology induced by I^nA. Because the two topologies are identical, their inverse limits coincide:

$$\varprojlim \left(\frac{A}{I^nB \cap A} \right) = \varprojlim \left(\frac{A}{I^n A} \right) = \hat{A}$$

Therefore, substituting this into our sequence confirms that:

$$0 \to \hat{A} \to \hat{B} \to \hat{C} \to 0$$

is exact, completing the proof. Note that the application of the Artin-Rees Lemma required A, B, and C to be finitely generated modules over a Noetherian ring. □

52.2. **Tensor Product versus Completion.** The second step in our journey is to relate the completion of a module to the tensor product with the completed ring.

Lemma 52.3. *If M is a finitely generated module over a Noetherian ring R, then there is a natural isomorphism:*

$$\hat{R} \otimes_R M \xrightarrow{\sim} \hat{M}$$

Before diving into the proof, we must acknowledge that this isomorphism fails if M is not finitely generated. It can fail to be surjective, and it can fail to be injective:

- **Failure of Injectivity:** Let $M = \mathbb{Q}$ and $R = \mathbb{Z}$, with completion at the ideal $I = (2)$. The completion $\hat{M}$ is 0, but the tensor product $\hat{\mathbb{Z}}_2 \otimes_{\mathbb{Z}} \mathbb{Q}$ is the field of 2-adic numbers $\mathbb{Q}_2$, which is highly non-zero. The canonical map $\hat{\mathbb{Z}}_2 \otimes_{\mathbb{Z}} \mathbb{Q} \to \hat{\mathbb{Q}}$ is the zero map $\mathbb{Q}_2 \to 0$. Because this map has a massive kernel, the natural map is not injective.
- **Failure of Surjectivity:** Let $M = \bigoplus_{i=1}^{\infty} \mathbb{Z}$ be an infinite direct sum of copies of $\mathbb{Z}$. The tensor product commutes with direct sums:

  $$\hat{\mathbb{Z}}_2 \otimes_{\mathbb{Z}} \left(\bigoplus_{i=1}^{\infty} \mathbb{Z} \right) \cong \bigoplus_{i=1}^{\infty} \hat{\mathbb{Z}}_2$$

  However, the completion involves an inverse limit, which behaves like a direct product, yielding elements with an infinite number of non-zero components that converge to zero in the 2-adic topology. Thus, $\hat{M}$ is strictly larger than the tensor product, and the natural map is not surjective.

Proof of Lemma 52.3. Since M is a finitely generated module over a Noetherian ring R, we can present it as a quotient of a finite free module. Let F be a finitely generated free module mapping surjectively onto M, and let N be the kernel. This provides an exact sequence:

$$0 \to N \to F \to M \to 0$$

Because R is Noetherian, the submodule N of the finitely generated module F is also finitely generated.

We analyze this sequence in two different ways. First, we tensor the sequence with the completed ring $\hat{R}$. The tensor product is right-exact, yielding the exact sequence:

$$\hat{R} \otimes_R N \to \hat{R} \otimes_R F \to \hat{R} \otimes_R M \to 0$$

Second, because N, F, and M are all finitely generated, we can apply Lemma 52.1 to obtain the exact sequence of completions:

$$0 \to \hat{N} \to \hat{F} \to \hat{M} \to 0$$

We combine these into a commutative diagram:

$$\begin{array}{ccccccc} \hat{R} \otimes_R N & \longrightarrow & \hat{R} \otimes_R F & \longrightarrow & \hat{R} \otimes_R M & \longrightarrow & 0 \\ \downarrow & & \downarrow \cong & & \downarrow & & \\ \hat{N} & \longrightarrow & \hat{F} & \longrightarrow & \hat{M} & \longrightarrow & 0 \end{array}$$

Because F is a finitely generated free module, say $F \cong R^n$, both operations respect the finite direct sum perfectly. Therefore, the middle vertical map $\hat{R} \otimes_R F \to \hat{F}$ is trivially an isomorphism.

A basic diagram chase (a variant of the Five Lemma) shows that because the middle map is an isomorphism and the right horizontal maps are surjective, the rightmost vertical map $\hat{R} \otimes_R M \to \hat{M}$ must be surjective.

Since this surjectivity holds for *any* finitely generated module, it also holds for N. Thus, the leftmost vertical map $\hat{R} \otimes_R N \to \hat{N}$ is surjective. Applying the Five Lemma logic again with a surjective left map and an isomorphic middle map forces the rightmost vertical map to be an isomorphism:

$$\hat{R} \otimes_R M \cong \hat{M}$$

This completes the proof. □

52.3. **Flatness of Completion.** With these two lemmas established, we can now achieve our primary goal: demonstrating the flatness of the completion.

Theorem 52.4. *If R is a Noetherian ring and I is an ideal of R, then the completion $\hat{R}$ is a* flat *R-module.*

Proof. To prove that $\hat{R}$ is flat, we must show that tensoring with $\hat{R}$ preserves exactness. By standard homological algebra, this is equivalent to showing that the first Tor group vanishes for all R-modules M:

$$\mathrm{Tor}_1^R(\hat{R}, M) = 0 \quad \text{for all } M$$

We first establish this for finitely generated modules. Suppose M is finitely generated. By Lemma 52.3, we have the isomorphism $\hat{R} \otimes_R M \cong \hat{M}$. By Lemma 52.1, the functor $M \mapsto \hat{M}$ preserves exactness for finitely generated modules. Because this functor preserves exactness, its corresponding Tor group must vanish:

$$\operatorname{Tor}_1^R(\hat{R}, M) = 0 \quad \text{for } M \text{ finitely generated.}$$

To extend this to arbitrary modules, we utilize the property that the Tor functor commutes with filtered direct limits. Any arbitrary module M can be written as the direct limit of its finitely generated submodules $\{M_\alpha\}$:

$$M = \varinjlim M_\alpha$$

We compute the Tor group for the limit:

$$\begin{aligned} \operatorname{Tor}_1^R(\hat{R}, M) &= \operatorname{Tor}_1^R\left(\hat{R}, \varinjlim M_\alpha\right) \\ &= \varinjlim \operatorname{Tor}_1^R(\hat{R}, M_\alpha) \end{aligned}$$

Because each M_α is finitely generated, $\operatorname{Tor}_1^R(\hat{R}, M_\alpha) = 0$ for all α. The direct limit of zeros is zero:

$$\varinjlim 0 = 0$$

Therefore, $\operatorname{Tor}_1^R(\hat{R}, M) = 0$ for all R-modules M. This confirms that $\hat{R}$ is a flat R-module. □

52.4. **Summary: How to Complete Modules.** The flatness of completion yields a critical lesson for practical calculations. We frequently need to move modules from a ring R to its completion $\hat{R}$. There are two obvious methods to attempt this:

(1) **The "bad way":** Take the completion of the module itself, $M \mapsto \hat{M}$.
(2) **The "good way":** Tensor the module with the completed ring, $M \mapsto M \otimes_R \hat{R}$.

For finitely generated modules over Noetherian rings, these two methods produce the exact same result (by Lemma 52.3). However, for modules that are not finitely generated, the first method produces severe pathologies.

Example 52.5. Consider the sequence over the integers $R = \mathbb{Z}$:

$$0 \to \mathbb{Z} \to \mathbb{Q} \to \frac{\mathbb{Q}}{\mathbb{Z}} \to 0$$

We wish to turn this into a sequence over the 2-adic integers, $\hat{\mathbb{Z}}_2$.

- If we use the "bad way" (taking completions), the sequence devolves into:

  $$0 \to \mathbb{Z}_2 \to 0 \to 0 \to 0$$

  This sequence is an unmitigated mess: it is completely broken and fails to be exact.

- If we use the "good way" (tensoring with $\hat{\mathbb{Z}}_2$), we obtain:

$$0 \to \mathbb{Z}_2 \to \mathbb{Q}_2 \to \frac{\mathbb{Q}_2}{\mathbb{Z}_2} \to 0$$

  This sequence preserves the precise structure of the original sequence and remains perfectly exact, directly demonstrating why tensoring with the flat module $\hat{R}$ behaves a lot better.

So the take-home-message should be: taking completions behaves pathologically for modules that are not finitely generated. When handling such modules, one should avoid the formal completion operation and instead tensor with the completed base ring.

Part 7. Dimension Theory

53. Dimension Introductory Survey

This lecture serves as a broad survey of the various attempts to rigorously define the concept of *dimension* in mathematics, leading to the definitions most useful in commutative algebra. Our ultimate goal is to define the dimension of a commutative ring R such that it aligns with geometric intuition.

For instance, consider the polynomial ring in three variables over a field k:

$$R = k[x, y, z]$$

Geometrically, this is the coordinate ring of regular functions on 3-dimensional affine space. Therefore, any sensible definition of dimension should ensure that:

$$\dim(k[x, y, z]) = 3$$

While dimension is usually intuitively obvious, formulating a rigorous, universally applicable definition has historically been quite difficult.

53.1. **Topological and Historical Attempts.** Early mathematicians assumed the dimension of a space was simply the number of real parameters required to uniquely define a point within it.

53.1.1. *Cardinality and Continuous Parameters.* This naive parameter-counting approach fell apart when Georg Cantor discovered that there exists a bijection between the real line $\mathbb{R}$ and the plane $\mathbb{R}^2$ (and more generally, $\mathbb{R}^n$). Because these sets have the same cardinality, one actually only needs a single real number to parameterize points in any $\mathbb{R}^n$.

Cantor's bijections, however, were highly discontinuous. One might hope that restricting to *continuous* maps would save the parameter definition. This hope was dashed by Giuseppe Peano, who constructed continuous, surjective maps from $\mathbb{R}^1$ onto $\mathbb{R}^2$, now known as space-filling curves. Even if we insist on continuous maps, $\mathbb{R}^2$ can still be parameterized by a single real number.

(As an amusing historical aside, Norbert Wiener once suggested using Peano's space-filling curve to define integration over $\mathbb{R}^2$ by pulling it back to a standard integral over $\mathbb{R}^1$. For fairly obvious practical reasons, this attempt to define multidimensional integration never gained traction).

If we restrict ourselves to *differentiable* maps, the parameter definition finally works: there is no differentiable surjection from $\mathbb{R}^1$ onto $\mathbb{R}^2$. However, defining dimension via differentiable manifolds is extraordinarily difficult to work with purely algebraically. For instance, proving that a 2-dimensional real vector space is not homeomorphic to a 3-dimensional real vector space requires non-trivial tools from algebraic topology.

53.1.2. *Lebesgue Covering Dimension.* One of the earliest rigorous definitions that succeeded in general topology is the Lebesgue covering dimension.

Definition 53.1. A topological space X has *Lebesgue covering dimension* $\leq n$ if every open cover of X has a refinement such that every point $x \in X$ is contained in at most $n+1$ of the refining open sets.

For example, if we cover $\mathbb{R}^2$ with overlapping open disks, we can always shrink and arrange these disks so that no more than 3 of them intersect at any given point. Informally, this intersection property confirms that $\mathbb{R}^2$ has Lebesgue covering dimension 2. It turns out that $\mathbb{R}^n$ has covering dimension n, though proving this is notoriously difficult.

While successful in general topology, this definition fails catastrophically in algebraic geometry. In the Zariski topology of the spectrum of a ring, $\mathrm{Spec}(R)$, almost all non-empty open sets intersect one another. Consequently, the Lebesgue covering dimension of even the 1-dimensional affine line is infinite, rendering this definition useless for commutative algebra.

53.1.3. *Brouwer-Menger-Urysohn (BMU) Dimension.* The classical topological definition of dimension was developed by Brouwer, Menger, and Urysohn. It operates on the inductive principle that the boundary of a space should have a dimension one less than the space itself.

Definition 53.2. A topological space X has *BMU dimension* $\leq n$ (where $n \in \{-1, 0, 1, 2, \dots\}$) if every point $x \in X$ has arbitrarily small open neighborhoods U whose topological boundaries satisfy:

$$\dim(\partial U) \leq n - 1$$

The empty set is defined to be the unique space with dimension -1.

Classically, this was only utilized for separable metric spaces. However, it is a highly surprising fact that the BMU dimension works remarkably well for *Noetherian topological spaces*, which are about as far from separable metric spaces as one can get.

53.2. **Commutative Algebra Definitions.** We now examine the definitions of dimension specifically tailored for rings and algebraic varieties.

53.2.1. *Krull Dimension.* Developed by Wolfgang Krull, this is the standard definition of dimension used throughout commutative algebra.

Definition 53.3. The *Krull dimension* of a topological space X (or a ring R) is the supremum of the lengths n of all strictly increasing chains of irreducible closed subsets Z_i (or prime ideals $\mathfrak{p}_i$):

$$Z_0 \subsetneq Z_1 \subsetneq Z_2 \subsetneq \cdots \subsetneq Z_n$$
$$\mathfrak{p}_0 \supsetneq \mathfrak{p}_1 \supsetneq \mathfrak{p}_2 \supsetneq \cdots \supsetneq \mathfrak{p}_n$$

For example, consider the spectrum of the polynomial ring in two variables, $\mathrm{Spec}(k[x, y])$. We can construct a chain consisting of a closed point (maximal ideal), an irreducible curve passing through it, and the entire affine plane (the zero ideal):

$$(x, y) \supsetneq (x) \supsetneq (0)$$

This chain has length 2 (consisting of 3 distinct ideals), which proves that the Krull dimension is at least 2. Proving that it is *at most* 2 requires structurally bounding all possible chains of prime ideals.

The Krull dimension is extremely well-behaved for Noetherian topological spaces. However, it is dreadful for non-Noetherian spaces. For instance, the Krull dimension of $\mathbb{R}^n$ with the standard Euclidean topology is 0. This occurs because in any Hausdorff space, the only irreducible closed subsets are individual points, precluding any strictly increasing chains.

53.2.2. *Deviation of a Poset.* A beautiful, highly abstract definition is the deviation of a partially ordered set (poset).

Definition 53.4. The *deviation* of a poset P is an ordinal number α. We say the deviation is $\leq \alpha$ if, for any infinite descending chain:

$$a_0 > a_1 > a_2 > \dots$$

all but finitely many of the intervals $[a_{i+1}, a_i]$ have deviation strictly less than α.

If R is a Noetherian ring, we can look at the poset of *all* its ideals (not just prime ideals). The dimension of R turns out to be the deviation of this poset of ideals. While we will not use this definition frequently, it is extremely powerful because it naturally generalizes to *non-commutative* rings and arbitrary modules, where defining dimension via prime ideals (Krull dimension) breaks down or becomes highly pathological. Note that the dimension of a module in this sense is completely different from the dimension of a vector space; all vector spaces over fields have a poset deviation dimension of 0.

53.3. **Analytic and Algebraic Variants.**

53.3.1. *Hausdorff Dimension.* The Hausdorff dimension utilizes a metric structure. One covers a metric space X with balls of radius ϵ and analyzes the asymptotic rate of growth of the number of balls required as $\epsilon \to 0$:

$$N(\epsilon) \sim \epsilon^{-d}$$

The exponent d defines the dimension. This dimension is highly desirable in analysis because it can evaluate to real, non-integral numbers, making it the standard tool for studying fractals.

Currently, there are no serious applications of non-integral Hausdorff dimension in mainstream algebraic geometry, primarily because algebraic varieties are rarely equipped with analytic metrics in their raw algebraic form. This presents a rather wild open research question: can one construct natural examples of fractional-dimensional algebraic varieties?

53.3.2. *Transcendence Degree.* An older algebraic definition relies on field theory.

Definition 53.5. Let B be an integral domain that is a finitely generated algebra over a field k. The dimension of B is the *transcendence degree* of its field of fractions $K(B)$ over k.

For example, if $B = k[x, y, z]$, its field of fractions is the rational function field $k(x, y, z)$. The maximal number of algebraically independent elements is 3, so:

$$\dim(B) = \text{tr.deg}_k(k(x, y, z)) = 3$$

This perfectly matches our geometric expectation. However, this definition fails spectacularly if we step outside the realm of algebras over fields. For the ring of integers $\mathbb{Z}$, the field of fractions is $\mathbb{Q}$. The transcendence degree of $\mathbb{Q}$ over its prime subfield is 0, but the Krull dimension of $\mathbb{Z}$ is 1. Furthermore, the definition requires B to be an integral domain, making it useless for rings with zero divisors.

53.3.3. *Gelfand-Kirillov Dimension.* For a finitely generated algebra R over a field k (which may be non-commutative), we can choose a finite set of generators. Let R_n be the finite-dimensional k-vector space spanned by all monomials in the generators of length at most n. The *Gelfand-Kirillov dimension* is defined by the asymptotic growth rate:

$$\text{GKdim}(R) = \limsup_{n \to \infty} \frac{\log(\dim_k R_n)}{\log n}$$

Informally, this measures the polynomial growth rate of the algebra, asserting that $\dim_k R_n \approx n^d$. For commutative rings, this coincides precisely with the Krull dimension. For non-commutative rings, it is a robust invariant that can sometimes evaluate to non-integral real numbers ≥ 2.

53.4. **Dimension of Local Rings.** Because dimension is intrinsically a *local* property—a space can be 1-dimensional at one point and 2-dimensional at another—it is often best to define the dimension of an algebraic variety at a specific point. Algebraically, this corresponds to defining the dimension of a *local ring* $(R, \mathfrak{m})$.

53.4.1. *Hilbert Polynomials.* The most computable and arguably "best" definition of dimension in commutative algebra relies on the Hilbert polynomial.

For a local ring $(R, \mathfrak{m})$, we analyze the length of the quotient module $R/\mathfrak{m}^k$. For sufficiently large integers k, this length is given by a polynomial in k:

$$\text{length}_R\left(\frac{R}{\mathfrak{m}^k}\right) = P(k)$$

If R is Noetherian, $P(k)$ is guaranteed to be a polynomial for $k \gg 0$. The *dimension* of the local ring R is defined to be the degree of this Hilbert polynomial $P(k)$.

For example, in the power series ring $k[[x]]$, the length of $k[[x]]/(x^k)$ is exactly k. This is a polynomial of degree 1, confirming that the affine line is 1-dimensional.

53.4.2. *Tangent Space Dimension and Singularities.* A manifold is typically assigned dimension by looking at its tangent space. For a local ring $(R, \mathfrak{m})$ over a field $k = R/\mathfrak{m}$, the Zariski tangent space is defined as the dual of the vector space $\mathfrak{m}/\mathfrak{m}^2$. One might naively define the dimension of the ring as:

$$\dim(R) \stackrel{?}{=} \dim_k \left(\frac{\mathfrak{m}}{\mathfrak{m}^2}\right)$$

While this gives the correct answer for smooth spaces, it fails completely at singularities.

Consider the coordinate ring of the cusp curve $y^2 = x^3$. This is a 1-dimensional object. However, at the origin (the maximal ideal $\mathfrak{m} = (x, y)$), the space $\mathfrak{m}/\mathfrak{m}^2$ requires both x and y as basis elements, evaluating to a 2-dimensional tangent space.

This failure is turned into a feature: a local ring is defined to be *regular* (or non-singular) precisely when the dimension of its tangent space exactly equals its Krull dimension.

53.4.3. *System of Parameters.* To fix the tangent space definition, we generalize the concept of generators.

Definition 53.6. In a local ring $(R, \mathfrak{m})$, a *system of parameters* is a set of elements $x_1, \ldots, x_d \in \mathfrak{m}$ that generate an ideal containing some power of the maximal ideal:

$$\mathfrak{m}^n \subseteq (x_1, \ldots, x_d) \subseteq \mathfrak{m}$$

The dimension of R is exactly the minimal number d of elements required to form a system of parameters. If $n = 1$, we recover the generators of the maximal ideal (the tangent space), which explains why the minimal number of elements is smaller than the tangent space dimension at a singularity.

53.5. **Conclusion.** In the following lectures, we will focus primarily on three definitions of dimension for commutative rings:

(1) The Krull Dimension.
(2) The degree of the Hilbert Polynomial.
(3) The minimal size of a System of Parameters.

We will prove the highly non-trivial fact that for any Noetherian local ring, these three seemingly completely disjoint definitions yield the exact same integer.

54. Hilbert Polynomials

This lecture introduces the Hilbert polynomial, which serves as a tool for defining the dimension of a local Noetherian ring. We will construct the Hilbert polynomial abstractly by analyzing the growth rate of graded modules.

54.1. **Graded Rings and Modules.** Let us establish the basic setting. Suppose we have a graded ring R, which decomposes as a direct sum over the integers n:

$$R = \bigoplus_{n \in \mathbb{Z}} R_n$$

We impose two finiteness conditions on R:

(1) The degree-zero component R_0 is a Noetherian ring.
(2) R is finitely generated as an algebra over R_0.

By Hilbert's Basis Theorem, these two conditions immediately imply that R itself is a Noetherian ring.

Now, let M be a graded module over R, decomposing as:

$$M = \bigoplus_{n \in \mathbb{Z}} M_n$$

We assume that M is finitely generated *as a module* over R. (It is crucial to distinguish between being finitely generated as a module versus as an algebra).

The fundamental problem we wish to solve is: *How fast does the size of the graded component M_n grow as $n \to \infty$?*

54.2. **Measuring the Size of a Module.** To quantify the growth of M_n, we need a formal way to measure the "size" of a module. We define a size function $\lambda(M_n)$, which assigns an integer to each graded piece. The exact definition of λ depends on the base ring R_0:

- **Case 1 (R_0 is a field):** In this case, each M_n is a finite-dimensional vector space over the field R_0. We simply define $\lambda(M_n)$ to be the vector space dimension:

$$\lambda(M_n) = \dim_{R_0}(M_n)$$

- **Case 2 (R_0 is an Artinian ring):** Over an Artinian ring, all finitely generated modules have finite length. We can define $\lambda(M_n)$ to be the module length:

$$\lambda(M_n) = \text{length}_{R_0}(M_n)$$

The most critical requirement for our size function λ is that it must be additive on short exact sequences. If we have an exact sequence of R_0-modules:

$$0 \to A \to B \to C \to 0$$

then the size function must satisfy:

$$\lambda(B) = \lambda(A) + \lambda(C)$$

(In a more advanced K-theoretic context, this additivity means that λ induces a homomorphism from the Grothendieck group $K_0(R_0)$ to the integers $\mathbb{Z}$).

54.3. **The Poincaré Series.** We package the sequence of sizes $\lambda(M_n)$ into a formal power series, known as the *Poincaré series* (or Hilbert series) of the module M.

Definition 54.1. The *Poincaré series* of M is defined as the formal power series in a dummy variable t:

$$P_M(t) = \sum_{n \geq 0} \lambda(M_n) t^n$$

This is a formal power series with integer coefficients. The central theorem regarding the Poincaré series is that it is always a rational function.

Theorem 54.2. *The Poincaré series $P_M(t)$ is a rational function of the form:*

$$P_M(t) = \frac{f(t)}{(1-t^{k_1})(1-t^{k_2})\cdots(1-t^{k_s})}$$

where $f(t)$ is a polynomial with integer coefficients, and $k_1, k_2, \ldots, k_s$ are the degrees of the s generators of R as an algebra over R_0.

Proof. Because R is a finitely generated R_0-algebra, we can pick a finite set of homogeneous generators $x_1, \ldots, x_s$ for R over R_0. Let the degree of x_i be denoted by k_i. We proceed by induction on s, the number of generators.

Base Case ($s = 0$): If $s = 0$, then $R = R_0$. Because M is a finitely generated R_0-module, the graded pieces M_n must be zero for all sufficiently large n. Consequently, the infinite sum truncates, and the Poincaré series $P_M(t)$ is a polynomial. The theorem holds trivially.

Inductive Step ($s > 0$): Assume the theorem holds for graded modules over algebras generated by $s-1$ elements. We isolate the action of the last generator, x_s. Multiplication by x_s yields an R_0-module homomorphism from M_n to M_{n+k_s}. We can form an exact sequence by defining K_n as the kernel and L_{n+k_s} as the cokernel:

$$0 \to K_n \to M_n \xrightarrow{\cdot x_s} M_{n+k_s} \to L_{n+k_s} \to 0$$

Let $K = \bigoplus K_n$ and $L = \bigoplus L_n$. Because K and L are respectively a submodule and a quotient module of a finitely generated Noetherian module, they are both finitely generated R-modules. Crucially, the element x_s annihilates both K and L. Therefore, K and L can be viewed as modules over the smaller ring $R/(x_s)$, which is generated over R_0 by only $s-1$ elements $(x_1, \ldots, x_{s-1})$. By our inductive hypothesis, their Poincaré series $P_K(t)$ and $P_L(t)$ are rational functions with the desired denominator structure.

We can split our 4-term exact sequence into two short exact sequences. Utilizing the additivity of the size function λ, the alternating sum of the sizes must be zero:

$$\lambda(K_n) - \lambda(M_n) + \lambda(M_{n+k_s}) - \lambda(L_{n+k_s}) = 0$$

Multiplying this equation by t^{n+k_s} and summing over all $n \geq 0$, we convert this relation into an equation of Poincaré series. The terms involving index shifts naturally absorb factors of t:

$$t^{k_s} P_K(t) - t^{k_s} P_M(t) + P_M(t) - P_L(t) = g(t)$$

where $g(t)$ is some correction polynomial accounting for the lowest degree terms skipped by the shift. Factoring out $P_M(t)$ yields:

$$(1 - t^{k_s}) P_M(t) = P_L(t) - t^{k_s} P_K(t) + g(t)$$

Dividing by $(1 - t^{k_s})$ provides an explicit formula for $P_M(t)$:

$$P_M(t) = \frac{P_L(t) - t^{k_s} P_K(t) + g(t)}{1 - t^{k_s}}$$

Since $P_L(t)$ and $P_K(t)$ are rational functions with denominators containing factors of the form $(1 - t^{k_i})$ for $i < s$, $P_M(t)$ is a rational function with the extra factor $(1 - t^{k_s})$ in the denominator. This completes the induction. □

54.4. **The Hilbert Polynomial.** We now specialize to the most common geometric case: all generators x_i have degree 1 (i.e., $k_i = 1$ for all i). Theorem 54.2 then states that the Poincaré series simplifies to:

$$P_M(t) = \frac{f(t)}{(1-t)^s}$$

where s is the number of degree-1 generators.

We can expand the denominator using the generalized binomial theorem:

$$\frac{1}{(1-t)^s} = \sum_{n \geq 0} \binom{n+s-1}{s-1} t^n$$

For any fixed integer $s \geq 1$, the binomial coefficient $\binom{n+s-1}{s-1}$ acts as a polynomial in the variable n of degree $s-1$. Since $P_M(t)$ is the product of the polynomial $f(t)$ and this expansion, the coefficient of t^n in $P_M(t)$ (which is precisely our size measure $\lambda(M_n)$) will eventually stabilize into a polynomial in n.

Definition 54.3. For sufficiently large n, the value $\lambda(M_n)$ agrees identically with a polynomial in n. This unique polynomial is called the *Hilbert polynomial* of the module M, denoted $H_M(n)$.

54.5. **Integer-Valued Polynomials.** Because $\lambda(M_n)$ inherently represents an integer (a dimension or a length), the Hilbert polynomial $H_M(n)$ must evaluate to an integer for all sufficiently large integers n. Consequently, by standard algebraic continuation, $H_M(n)$ takes integer values for *all* integers n. Such polynomials are called *integer-valued polynomials*.

An integer-valued polynomial does not necessarily have integer coefficients. For instance, the polynomial:

$$f(n) = \frac{n(n-1)}{2} = \frac{1}{2}n^2 - \frac{1}{2}n$$

has rational coefficients, but yields an integer for every integer n.

Theorem 54.4. *Any integer-valued polynomial $f(n)$ of degree d can be expressed as a finite $\mathbb{Z}$-linear combination of binomial polynomials:*

$$f(n) = a_0\binom{n}{0} + a_1\binom{n}{1} + a_2\binom{n}{2} + \cdots + a_d\binom{n}{d}$$

where $a_i \in \mathbb{Z}$ and $\binom{n}{i} = \frac{n(n-1)\cdots(n-i+1)}{i!}$.

Proof. We proceed by analyzing the values of the binomial polynomials at consecutive integers $n = 0, 1, 2, \ldots$. Observe that the polynomial $\binom{n}{i}$ vanishes for all integers n such that $0 \leq n < i$, and evaluates exactly to 1 when $n = i$.

This triangular evaluation matrix allows us to recursively determine integer coefficients a_i. We set $a_0 = f(0)$. Then $f(1) - a_0\binom{1}{0}$ determines a_1, and so forth. Because we can iteratively force our linear combination to exactly match the values of $f(n)$ at the $d+1$ points $n = 0, 1, \ldots, d$, the two polynomials of degree d must be identical everywhere. □

A direct corollary of this structural decomposition is that the leading term of an integer-valued polynomial of degree d must take the form:

$$\frac{a_d}{d!}n^d$$

for some integer a_d. While the leading coefficient itself may be a fraction, multiplying it by $d!$ definitively clears the denominator.

54.6. **Examples in Projective Geometry.** The Hilbert polynomial provides a robust method for computing geometric invariants such as dimension and degree.

Suppose V is a projective variety defined by a homogeneous ideal I inside the polynomial ring $S = k[x_0, \ldots, x_N]$. We consider the homogeneous coordinate ring $R = S/I$. The components R_k are finite-dimensional vector spaces over the field k. For sufficiently large k, the dimension of R_k is given by the Hilbert polynomial:

$$\dim_k(R_k) = H_R(k)$$

The *dimension* of the projective variety V is defined as exactly the degree d of its Hilbert polynomial $H_R(k)$. The *degree* of the variety is defined as $d!$ times the leading coefficient of $H_R(k)$, which we proved above must be an integer.

Example 54.5. Consider a single hypersurface $V \subset \mathbb{P}^N$ defined by an irreducible homogeneous polynomial f of degree m. The ideal is simply $I = (f)$.

The dimension of the graded piece of the full polynomial ring S_k is the number of monomials of degree k in $N+1$ variables, which is exactly $\binom{k+N}{N}$. The ideal I in degree k is formed by multiplying f by all polynomials of degree $k-m$. Thus, the dimension of the quotient R_k is the difference:

$$\begin{aligned} H_R(k) &= \binom{k+N}{N} - \binom{k+N-m}{N} \\ &= \frac{(k+N)\cdots(k+1)}{N!} - \frac{(k+N-m)\cdots(k-m+1)}{N!} \end{aligned}$$

If we expand these terms for large k, the leading terms of $k^N/N!$ cancel each other out. The next highest term dictating the polynomial growth is k^{N-1}. Thus, the

degree of the Hilbert polynomial is $N-1$, perfectly confirming that a single hypersurface in $\mathbb{P}^N$ has geometric dimension $N-1$.

To find the degree of the variety, we compute the leading coefficient of k^{N-1}. Careful algebraic expansion of the binomial difference yields that the coefficient of k^{N-1} is exactly:

$$\frac{m}{(N-1)!}$$

Multiplying by the factorial of the dimension (which is $(N-1)!$) isolates the numerator, demonstrating that the degree of the variety is exactly m, precisely the degree of the defining polynomial f.

55. Dimension of Local Rings

In this lecture, we will focus on the dimension of local rings, providing four distinct definitions and establishing the groundwork to prove their equivalence. We will conclude by explicitly proving that the Krull dimension matches the topological Brouwer-Menger-Urysohn dimension for Noetherian spaces.

55.1. **Four Definitions of Dimension.** Let R be a local ring with a unique maximal ideal $\mathfrak{m}$. We define the dimension of R in the following four ways:

Definition 55.1. The *Brouwer-Menger-Urysohn (BMU) dimension* is a classic topological definition. A topological space X has BMU dimension $\leq n$ if every point $x \in X$ has arbitrarily small open neighborhoods whose boundary has BMU dimension strictly less than n. We say a local ring R has BMU dimension n if its prime spectrum $\mathrm{Spec}(R)$ has BMU dimension n.

Definition 55.2. The *Krull dimension* is defined via chains of prime ideals. The dimension of a topological space is the supremum of the lengths of all strictly increasing chains of irreducible closed subsets:

$$Z_0 \subsetneq Z_1 \subsetneq Z_2 \subsetneq \cdots \subsetneq Z_n$$

The Krull dimension of R is the Krull dimension of $\mathrm{Spec}(R)$, which corresponds to the supremum of the lengths of strictly increasing chains of prime ideals in R.

Definition 55.3. The *Hilbert polynomial dimension* uses the graded ring formed by the quotients of powers of the maximal ideal. Assuming R is a Noetherian local ring, the dimension of the vector space $R/\mathfrak{m}^n$ over the residue field $R/\mathfrak{m}$ is a polynomial in n for sufficiently large n:

$$\dim_{R/\mathfrak{m}} \left(\frac{R}{\mathfrak{m}^n} \right) = P(n)$$

The dimension of R is defined as the degree of this Hilbert polynomial $P(n)$.

Definition 55.4. A *system of parameters* is a set of generators $x_1, \ldots, x_d$ for an ideal I that is sandwiched between the maximal ideal and some power of the maximal ideal:

$$\mathfrak{m}^r \subseteq I \subseteq \mathfrak{m}$$

for some integer $r > 0$. The dimension of R is defined as the minimum cardinality of any valid system of parameters.

55.2. **An Example: The Ghost of a Cusp.**

Example 55.5. Let R be the ring of formal power series in two variables quotiented by the relation for a cusp:

$$R = \frac{k[[x, y]]}{(y^2 - x^3)}$$

Geometrically, $\mathrm{Spec}(R)$ represents the neighborhood of a cusp singularity. Let us evaluate the dimension of R using our definitions.

Krull Dimension: The ring R has exactly two prime ideals: the zero ideal (0) and the maximal ideal (x, y). This yields a maximal chain of length one:

$$(0) \subsetneq (x, y)$$

Thus, the Krull dimension is 1.

Hilbert Polynomial: We can parameterize R using formal power series in t by substituting $x = t^2$ and $y = t^3$. The maximal ideal $\mathfrak{m}$ essentially consists of all formal power series in t with no constant or linear term. The vector space $R/\mathfrak{m}^n$ has dimension $n - 1$ for all $n > 0$. The polynomial $P(n) = n - 1$ has degree 1, meaning the Hilbert polynomial dimension is 1.

System of Parameters: The maximal ideal $\mathfrak{m}$ inherently requires two generators, x and y. However, if we consider the ideal I generated solely by x, we observe that $I = (t^2)$. This ideal is clearly contained in $\mathfrak{m}$, and it contains $\mathfrak{m}^2$ (since $\mathfrak{m}^2$ consists of powers of t of degree 4 and higher). Thus:

$$\mathfrak{m}^2 \subseteq (x) \subseteq \mathfrak{m}$$

Therefore, $\{x\}$ constitutes a system of parameters. Because the minimum cardinality of this system is 1, the dimension by this definition is 1.

55.3. **Equivalence Strategy.** Our overarching goal is to prove that all four definitions of dimension are equivalent for Noetherian local rings. Over the course of the next several lectures, we will establish the following cyclic inequalities:

$$\dim_{\mathrm{Krull}}(R) \leq \dim_{\mathrm{Hilbert}}(R)$$
$$\dim_{\mathrm{Hilbert}}(R) \leq \dim_{\mathrm{Param}}(R)$$
$$\dim_{\mathrm{Param}}(R) \leq \dim_{\mathrm{Krull}}(R)$$

Proving this cycle will demonstrate that these three algebraic definitions are identical. In the remainder of this lecture, we will establish the equivalence of the Krull dimension and the topological BMU dimension.

Remark 55.6. It is important to note that dimension theory is highly pathological for non-Noetherian rings. For non-Noetherian spaces, the dimensions defined by the Hilbert polynomial or systems of parameters are often simply infinite. Furthermore, the Krull dimension of a polynomial ring $R[x]$ can be strictly greater than $1 + \dim(R)$, contradicting basic geometric expectations.

If we wish to define the dimension of a general, non-local ring R using the Hilbert or parameters definitions, we take the supremum of the dimensions of the local rings $R_{\mathfrak{p}}$ at all prime ideals $\mathfrak{p}$. This naturally equates to the Krull dimension of the global spectrum $\mathrm{Spec}(R)$.

55.4. **Krull Dimension equals BMU Dimension.** We now prove the geometric equivalence between the Krull and BMU dimensions.

Theorem 55.7. *For any Noetherian topological space X, the Krull dimension equals the Brouwer-Menger-Urysohn dimension.*

Proof. We prove this by establishing the inequality in both directions.

Part 1: $\dim_{\mathbf{Krull}}(X) \leq \dim_{\mathbf{BMU}}(X)$**.** Suppose we have a strictly increasing chain of irreducible closed subsets of length n:

$$Z_0 \subsetneq Z_1 \subsetneq Z_2 \subsetneq \cdots \subsetneq Z_n$$

We must show that the BMU dimension of Z_i is at least i. We proceed by induction. Pick a point $p \in Z_i$ such that $p \notin Z_{i-1}$. Any open neighborhood U of p within the subspace Z_i intersects Z_{i-1} trivially. Because Z_i is irreducible, Z_{i-1} must be fully contained in the boundary of U.

By our inductive hypothesis, the BMU dimension of Z_{i-1} is at least $i-1$. Therefore, the boundary of any arbitrarily small neighborhood of p has a BMU dimension of at least $i-1$. By the inductive definition of BMU dimension, this forces the BMU dimension of Z_i to be greater than $i-1$, meaning it is at least i. Since this holds for all valid chains, the inequality follows. (Notice this direction does not require the space to be Noetherian; it holds for all topological spaces).

Part 2: $\dim_{\mathbf{BMU}}(X) \leq \dim_{\mathbf{Krull}}(X)$**.** This direction fails for general spaces. For instance, the real line $\mathbb{R}$ with the standard Euclidean topology has a BMU dimension of 1, but its Krull dimension is 0 because it is a Hausdorff space (where irreducible closed subsets are merely single points). We must use the Noetherian property.

We proceed by Noetherian induction. Assume the inequality is false. Because X is a Noetherian space, the collection of closed subsets violating the inequality must have a minimal element. Let Y be this minimal closed subset. We may assume Y is irreducible; if Y were the union of two proper closed subsets, at least one of them would also violate the inequality, which contradicts the minimality of Y.

Suppose $\dim_{\mathrm{BMU}}(Y) \geq n$. By definition, there exists a point $p \in Y$ and an open neighborhood U such that the topological boundary ∂U has a BMU dimension of at least $n-1$. Because ∂U is a proper closed subset of the irreducible space Y, our minimal choice of Y guarantees that ∂U satisfies the inequality. Therefore, the Krull dimension of ∂U is at least $n-1$.

This implies there exists a strictly increasing chain of irreducible closed subsets in ∂U of length $n-1$:

$$Z_0 \subsetneq Z_1 \subsetneq \cdots \subsetneq Z_{n-1} \subseteq \partial U$$

Because ∂U is a proper closed subset of Y, we can extend this chain by appending the entire irreducible space Y to the end:

$$Z_0 \subsetneq Z_1 \subsetneq \cdots \subsetneq Z_{n-1} \subsetneq Y$$

This newly constructed chain has length n, demonstrating that the Krull dimension of Y is at least n. This directly contradicts our initial assumption that Y violated the inequality $\dim_{\text{BMU}} \leq \dim_{\text{Krull}}$. Thus, no such minimal counterexample can exist, and the inequality must hold for all closed subsets of the Noetherian space X. □

56. Hilbert Polynomial versus System of Parameters

Previously, we introduced several different definitions for the dimension of a local ring. Again, our overarching goal is to show that they are all equivalent.

In this section, we will demonstrate that the dimension defined using Hilbert polynomials is less than or equal to the dimension defined using a system of parameters.

56.1. **Two Different Hilbert Polynomials.** Before proceeding with the proof, we will clarify a potential source of confusion regarding the Hilbert polynomial of a local ring $(R, \mathfrak{m})$. There are two distinct graded rings one can construct, leading to two different lengths (and thus two different polynomials for sufficiently large n):

$$\lambda\left(\frac{R}{\mathfrak{m}^{n+1}}\right)$$
$$\lambda\left(\frac{\mathfrak{m}^{n}}{\mathfrak{m}^{n+1}}\right)$$

Both of these expressions become polynomials in n for sufficiently large n. However, their degrees differ. The degree of the polynomial corresponding to the quotient ring on the left is exactly one greater than the degree of the polynomial corresponding to the graded pieces on the right:

$$\deg \lambda\left(\frac{R}{\mathfrak{m}^{n+1}}\right) = 1 + \deg \lambda\left(\frac{\mathfrak{m}^{n}}{\mathfrak{m}^{n+1}}\right)$$

Consequently, when referring to the "degree of the Hilbert polynomial" of a local ring, there is a slight ambiguity. To be precise, the *dimension* of the local ring R is defined as the degree of the polynomial corresponding to the entire quotient:

$$\dim(R) = \deg \lambda\left(\frac{R}{\mathfrak{m}^{n+1}}\right)$$

56.2. **Systems of Parameters and Graded Rings.** We now recall the definition of a system of parameters.

Definition 56.1. A *system of parameters* for a local ring $(R, \mathfrak{m})$ is a set of generators of an ideal q such that:

$$\mathfrak{m}^r \subseteq q \subseteq \mathfrak{m}$$

for some positive integer r.

Instead of using the maximal ideal $\mathfrak{m}$, let us examine the graded ring constructed using the ideal q:

$$\bigoplus_{n=0}^{\infty} \frac{R}{q^{n+1}} = \frac{R}{q} \oplus \frac{q}{q^2} \oplus \dots$$

Because q contains a power of the maximal ideal ($\mathfrak{m}^r \subseteq q$), the quotient R/q is an Artinian ring. Therefore, it has finite length over R. This provides an example where we use the Hilbert polynomial for a graded ring whose degree-zero component is no longer a field, but an Artinian ring.

For sufficiently large n, the length $\lambda(R/q^{n+1})$ is a polynomial in n. We wish to compare the degree of this polynomial with the number of generators of the ideal q.

56.3. **Bounding the Degree.** To achieve this comparison, we look at the other associated graded algebra:

$$\bigoplus_{n=0}^{\infty} \frac{q^n}{q^{n+1}} = \frac{R}{q} \oplus \frac{q}{q^2} \oplus \dots$$

The number of generators of the ideal q is equal to the number of generators of this graded algebra over its degree-zero piece R/q.

A fundamental result for graded algebras generated by degree-one elements states that the degree of the Hilbert polynomial $\lambda(q^n/q^{n+1})$ is strictly less than the number of generators of the algebra $\bigoplus q^n/q^{n+1}$ over R/q (provided the number of generators of this algebra is strictly positive).

Therefore, using the degree shift relationship:

$$\deg \lambda\left(\frac{R}{q^{n+1}}\right) = 1 + \deg \lambda\left(\frac{q^n}{q^{n+1}}\right)$$

we can deduce that the degree of $\lambda(R/q^{n+1})$ is at most the number of generators of the graded algebra. Thus, it is less than or equal to the number of generators of the ideal q:

$$\deg \lambda\left(\frac{R}{q^{n+1}}\right) \leq \text{number of generators of } q$$

56.4. **Comparing the Polynomials.** We now have a bound on the degree of the polynomial defined by q, but to relate this back to the dimension of R, we must compare it to the Hilbert polynomial defined by the maximal ideal $\mathfrak{m}$.

Recall the inclusion $\mathfrak{m}^r \subseteq q \subseteq \mathfrak{m}$. Taking the n-th power yields:

$$\mathfrak{m}^{nr} \subseteq q^n \subseteq \mathfrak{m}^n$$

This chain of inclusions induces a sequence of natural surjective maps on the quotient rings:

$$\frac{R}{\mathfrak{m}^{nr}} \twoheadrightarrow \frac{R}{q^n} \twoheadrightarrow \frac{R}{\mathfrak{m}^n}$$

Because these maps are surjective, the lengths of the modules must satisfy the corresponding inequalities:

$$\lambda\left(\frac{R}{\mathfrak{m}^{nr}}\right) \geq \lambda\left(\frac{R}{q^n}\right) \geq \lambda\left(\frac{R}{\mathfrak{m}^n}\right)$$

Let $f(n)$ be the Hilbert polynomial generated by the ideal $\mathfrak{m}$, and let $g(n)$ be the polynomial generated by the ideal q:

$$f(n) = \lambda\left(\frac{R}{\mathfrak{m}^n}\right)$$
$$g(n) = \lambda\left(\frac{R}{q^n}\right)$$

The length inequalities directly translate to inequalities between these polynomials for large n:

$$f(nr) \geq g(n) \geq f(n)$$

From these inequalities, we can definitively determine the relationship between their degrees. If $f(n)$ and $g(n)$ are polynomials that bound each other in this manner (specifically, $g(n)$ is bounded between $f(n)$ and a linearly rescaled argument version of $f(n)$), they must possess the exact same degree. While they may be totally different polynomials, they grow at the same asymptotic rate with respect to the exponent.

$$\deg f(n) = \deg g(n)$$

56.5. **Conclusion.** This equivalence of degrees completes the proof of our first inequality. We substitute the dimension definition:

$$\dim(R) = \deg \lambda\left(\frac{R}{\mathfrak{m}^{n+1}}\right)$$

By our polynomial comparison, this is equal to the degree of the polynomial defined by q:

$$\dim(R) = \deg \lambda\left(\frac{R}{q^{n+1}}\right)$$

Finally, using the bound we established via the graded algebra, this degree is less than or equal to the number of generators of q, for any system of parameters q:

$$\dim(R) \leq \text{number of generators of } q$$

Thus, the dimension defined by the Hilbert polynomial is less than or equal to the cardinality of any system of parameters.

57. Krull versus Hilbert

In this lecture, we will prove that the Krull dimension is bounded by the Hilbert dimension.

57.1. **Dimension Reduction Lemma.** Before proving the main inequality between the Krull dimension and the Hilbert dimension, we establish a lemma stating that quotienting out by an element that is not a zero divisor will reduce the Hilbert dimension.

Lemma 57.1. *Let R be a Noetherian local ring with maximal ideal $\mathfrak{m}$, and let $x \in \mathfrak{m}$ be an element that is not a zero divisor. Then the Hilbert dimension of the quotient ring decreases:*

$$\dim_{Hilbert}\left(\frac{R}{xR}\right) \leq \dim_{Hilbert}(R) - 1$$

Remark 57.2. In fact, equality holds, but for our current purposes, the inequality is sufficient. The condition that x is not a zero divisor is absolutely essential.

To see why the non-zero divisor condition is necessary, consider the following example.

Example 57.3. Let R be the local ring at the origin of the coordinate axes:

$$R = \left(\frac{k[x,y]}{(xy)}\right)_{(x,y)}$$

Geometrically, this is the local ring at the intersection point of the two lines $x = 0$ and $y = 0$. Because the variety is reducible, the ring contains zero divisors (e.g., x and y).

If we quotient out by the zero divisor x, we are effectively setting $x = 0$, which restricts our view to the y-axis:

$$\frac{R}{(x)} \cong k[y]_{(y)}$$

The original ring R (two intersecting lines) has dimension 1. The quotient ring $k[y]_{(y)}$ (a single line) also has dimension 1. Thus, quotienting by a zero divisor did not drop the dimension.

Proof. Assume x is not a zero divisor. We start with the short exact sequence given by multiplication by x:

$$0 \to R \xrightarrow{\times x} R \to \frac{R}{xR} \to 0$$

This sequence is exact precisely because x is not a zero divisor. We want to relate this to the Hilbert polynomial, so we quotient the sequence by $\mathfrak{m}^n$. Applying the tensor product $- \otimes_R R/\mathfrak{m}^n$ yields:

$$\frac{R}{\mathfrak{m}^n} \xrightarrow{\times x} \frac{R}{\mathfrak{m}^n} \to \frac{R}{xR + \mathfrak{m}^n} \to 0$$

However, we lose exactness on the left. To maintain a short exact sequence, we must replace the first term with the appropriate intersection:

$$0 \to \frac{R}{xR \cap \mathfrak{m}^n} \to \frac{R}{\mathfrak{m}^n} \to \frac{R}{xR + \mathfrak{m}^n} \to 0$$

The third term corresponds to the lengths that define the Hilbert polynomial of R/xR. Because the sequence is exact, the lengths are additive:

$$\lambda\left(\frac{R}{xR+\mathfrak{m}^n}\right) = \lambda\left(\frac{R}{\mathfrak{m}^n}\right) - \lambda\left(\frac{R}{xR\cap\mathfrak{m}^n}\right)$$

We must now compare the two terms on the right. By the strong Artin-Rees lemma, the filtration given by $xR\cap\mathfrak{m}^n$ is a *stable* $\mathfrak{m}$-filtration. This means that, up to a finite shift k, it is essentially equivalent to the standard $\mathfrak{m}$-adic filtration on the submodule xR. Specifically, there exists a fixed integer k such that:

$$\mathfrak{m}^n x \subseteq xR\cap\mathfrak{m}^n \subseteq \mathfrak{m}^{n-k}x$$

Because multiplication by x is an isomorphism from R to xR, the term on the right is isomorphic to $R/\mathfrak{m}^{n-k}$. Thus, the length of $R/(xR\cap\mathfrak{m}^n)$ is bounded between the lengths of $R/\mathfrak{m}^n$ and $R/\mathfrak{m}^{n-k}$.

Consequently, the Hilbert polynomials of $R/\mathfrak{m}^n$ and $R/(xR\cap\mathfrak{m}^n)$ must have the exact same degree and the exact same leading coefficient. When we subtract them, the leading terms perfectly cancel out. Thus, the difference is a polynomial of strictly smaller degree:

$$\deg\lambda\left(\frac{R}{xR+\mathfrak{m}^n}\right) < \deg\lambda\left(\frac{R}{\mathfrak{m}^n}\right)$$

By definition, this means:

$$\dim_{\text{Hilbert}}\left(\frac{R}{xR}\right) \leq \dim_{\text{Hilbert}}(R) - 1$$

□

57.2. **Krull Dimension vs. Hilbert Dimension.** Using this lemma, we can now prove the main theorem of this section.

Theorem 57.4. *For any Noetherian local ring R, the Krull dimension is less than or equal to the Hilbert dimension:*

$$\dim_{\mathit{Krull}}(R) \leq \dim_{\mathit{Hilbert}}(R)$$

Proof. Suppose that $\dim_{\text{Krull}}(R) \geq n$. We must show that $\dim_{\text{Hilbert}}(R) \geq n$. We proceed by induction on n.

Because the Krull dimension is at least n, there exists a strictly increasing chain of prime ideals of length n:

$$\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \cdots \subsetneq \mathfrak{p}_n$$

We may without loss of generality pass to the quotient ring $R/\mathfrak{p}_0$. Since $\mathfrak{p}_0$ is a prime ideal, the quotient ring is an integral domain. Passing to a quotient ring can only potentially decrease the dimension, so if we prove the Hilbert dimension of the quotient is at least n, it must have been at least n for the original ring. Thus, we may assume R is an integral domain and $\mathfrak{p}_0 = (0)$.

Now, pick a non-zero element $x \in \mathfrak{p}_1$. Because R is an integral domain and $x \neq 0$, x is not a zero divisor. We can therefore apply our previous lemma to find:

$$\dim_{\text{Hilbert}}(R) \geq \dim_{\text{Hilbert}} \left(\frac{R}{xR} \right) + 1$$

In the quotient ring R/xR, the chain of prime ideals becomes:

$$\frac{\mathfrak{p}_1}{xR} \subsetneq \frac{\mathfrak{p}_2}{xR} \subsetneq \cdots \subsetneq \frac{\mathfrak{p}_n}{xR}$$

This is a strictly increasing chain of prime ideals of length $n-1$. Therefore, the Krull dimension of the quotient ring R/xR is at least $n-1$.

By our inductive hypothesis, the theorem holds for $n-1$, which implies:

$$\dim_{\text{Hilbert}} \left(\frac{R}{xR} \right) \geq \dim_{\text{Krull}} \left(\frac{R}{xR} \right) \geq n-1$$

Substituting this back into our inequality yields:

$$\dim_{\text{Hilbert}}(R) \geq (n-1) + 1 = n$$

Since this holds for any $n \leq \dim_{\text{Krull}}(R)$, it follows that $\dim_{\text{Krull}}(R) \leq \dim_{\text{Hilbert}}(R)$. □

In the next lecture, we will complete the cycle of inequalities by showing that the dimension defined by a system of parameters is at most the Krull dimension.

58. System of Parameters versus Krull

In this lecture, we will prove that the dimension defined using a system of parameters is at most the Krull dimension. Once this is established, we will have proven that these three different definitions of dimension are all, in fact, perfectly equivalent.

Before proving the main theorem, we require a fundamental result known as the Prime Avoidance Lemma.

58.1. **The Prime Avoidance Lemma.** The problem is the following: suppose we are given a finite number of ideals $P_1, \ldots, P_n$ contained in a ring R, and suppose we are given another ideal I of R. We wish to find some element $x \in I$ such that x is not in any of the ideals P_i.

There is an obvious necessary condition: the ideal I cannot be contained in any individual P_i. If it were, we obviously could not satisfy the condition.

Geometrically, we are trying to cover the ideal I by a finite number of other ideals such that none of them individually contain I.

Without further assumptions, this covering can sometimes happen.

Example 58.1. Let R be an algebra over the field with two elements, specifically:

$$R = \frac{\mathbb{F}_2[x, y]}{(x^2, xy, y^2)}$$

This is an Artinian ring and a 3-dimensional algebra over $\mathbb{F}_2$ with basis $1, x$, and y. Let I be the ideal generated by x and y. Because the field has only two elements, I contains exactly four elements:

$$I = \{0, x, y, x + y\}$$

Now we define three ideals:

$$P_1 = (x)$$
$$P_2 = (y)$$
$$P_3 = (x + y)$$

Each of these ideals contains exactly two elements (zero and its generator). We can clearly see that:

$$I \subseteq P_1 \cup P_2 \cup P_3$$

Thus, we cannot find an element in I that avoids all three ideals. However, notice that none of these three ideals P_i are prime. For instance, $y^2 = 0 \in P_1$, but $y \notin P_1$.

If we require the covering ideals to be prime, we can always find such an element. This is why the result is called *Prime Avoidance*.

Lemma 58.2 (Prime Avoidance)**.** *Let I be an ideal of a commutative ring R, and let $P_1, \ldots, P_n$ be prime ideals. If I is not contained in any P_i, then there exists an element $x \in I$ such that:*

$$x \notin P_i \quad \textit{for all } i = 1, \ldots, n$$

Proof. We may assume without loss of generality that $P_i \not\subseteq P_j$ for $i \neq j$. If one ideal were contained in another, it would be redundant in the union, and we could simply throw it out.

We proceed by induction on n, the number of prime ideals.

Base Case ($n = 1$): This is trivial. Because we assumed $I \not\subseteq P_1$, we can simply pick an element $x \in I$ such that $x \notin P_1$.

Inductive Step: Assume the lemma holds for $n - 1$ prime ideals. By the inductive hypothesis, we can pick an element x such that:

$$x \in I$$
$$x \notin P_1 \cup \cdots \cup P_{n-1}$$

If $x \notin P_n$, we are entirely done. We have found our element.

Thus, we may assume $x \in P_n$. For each $i < n$, because $P_i \not\subseteq P_n$, we can pick an element y_i such that:

$$y_i \in P_i$$
$$y_i \notin P_n$$

Additionally, because $I \not\subseteq P_n$, we can pick an element $y \in I$ such that $y \notin P_n$.

We now construct a new element:

$$z = x + y \cdot y_1 y_2 \cdots y_{n-1}$$

We claim that z satisfies all our required conditions.

- First, $z \in I$. This is because $x \in I$ and $y \in I$, so their sum (with any multiplier) is in I.
- Second, $z \notin P_n$. We know $x \in P_n$. The product $y \cdot y_1 \cdots y_{n-1}$ cannot be in P_n because P_n is a prime ideal and none of the individual factors $y, y_1, \ldots, y_{n-1}$ are in P_n. The sum of an element in P_n and an element not in P_n is not in P_n.
- Third, $z \notin P_i$ for any $i < n$. The product $y \cdot y_1 \cdots y_{n-1}$ is in P_i because it contains the factor $y_i \in P_i$. However, we originally chose x such that $x \notin P_i$. The sum of an element not in P_i and an element in P_i is not in P_i.

Therefore, z is an element of I that avoids all n prime ideals. □

Remark 58.3. If one is careful, the assumption that *all* P_i are prime is slightly stronger than necessary. The proof above never actually utilizes the primeness of $P_1, \ldots, P_{n-1}$, only the primeness of P_n. Thus, the lemma holds even if up to two of the ideals are not prime.

58.2. **Bounding the System of Parameters.** We now proceed to prove the main theorem of this lecture. Let d be the Krull dimension of our local ring R, which is the maximum length of a chain of prime ideals.

Theorem 58.4. *Let $(R, \mathfrak{m})$ be a local Noetherian ring of Krull dimension d. The minimal size of a system of parameters for R is less than or equal to d.*

Proof. Given the Krull dimension d, we need to find a system of parameters $x_1, \ldots, x_d$ of size exactly d. We will construct this sequence of elements inductively.

We demand that the sequence $x_1, \ldots, x_d$ satisfies the following geometric property: for every $i \leq d$, any prime ideal P containing the elements $x_1, \ldots, x_i$ must have a co-dimension greater than or equal to i.

Recall that the co-dimension (or height) of a prime P is the largest length of a chain of proper inclusions $P_0 \subsetneq P_1 \subsetneq \cdots \subsetneq P_i = P$.

It is easier to visualize this algebraically geometric condition: we think of the elements x_i as defining hypersurfaces. Setting $x_1 = 0$ defines a hypersurface. The condition requires that every prime containing x_1 and x_2 (which corresponds to an irreducible component of the intersection of the two hypersurfaces $x_1 = 0$ and $x_2 = 0$) has co-dimension at least 2. We are systematically choosing functions to properly slice the variety, ensuring we eliminate degenerate cases where an intersection of k hypersurfaces drops the co-dimension by less than k.

Suppose we have successfully constructed the sequence up to x_{i-1}. We now need to find an appropriate x_i.

Consider all prime ideals P that contain the sequence $(x_1, \ldots, x_{i-1})$. By our inductive hypothesis, all such primes have co-dimension at least $i-1$. We isolate the *minimal* prime elements among them that have co-dimension exactly $i-1$. Let these be $P_1, P_2, \ldots, P_k$.

There are only a finite number of such minimal primes. This is because they correspond exactly to the minimal associated primes of the finitely generated

quotient module:

$$\frac{R}{(x_1, \ldots, x_{i-1})}$$

Crucially, none of these primes P_j can contain the maximal ideal $\mathfrak{m}$. The co-dimension of $\mathfrak{m}$ is exactly the Krull dimension d. Since we have not yet reached the end of our construction, we know that $d \geq i > i-1$. Because the co-dimension of P_j is exactly $i-1$, it strictly cannot equal $\mathfrak{m}$. Since $\mathfrak{m}$ is the unique maximal ideal, this forces the condition:

$$\mathfrak{m} \not\subseteq P_j \quad \text{for all } j$$

This is precisely the setup for the Prime Avoidance Lemma. We treat $\mathfrak{m}$ as our ideal I. Because $\mathfrak{m}$ is not contained in any of the finite number of prime ideals $P_1, \ldots, P_k$, we can find an element x_i such that:

$$x_i \in \mathfrak{m}$$
$$x_i \notin P_j \quad \text{for all } j$$

We now verify the inductive condition for x_i. Consider any prime ideal P containing $(x_1, \ldots, x_i)$. Because P contains $(x_1, \ldots, x_{i-1})$, its co-dimension is at least $i-1$. Could its co-dimension be exactly $i-1$? If it were, P would have to be one of the minimal primes P_j. But we explicitly chose $x_i \notin P_j$, whereas $x_i \in P$. Thus, P cannot be any of the P_j. This forces the co-dimension of P to be strictly greater than $i-1$, meaning it has co-dimension at least i.

We iterate this construction until we reach x_d. We have now constructed elements $x_1, \ldots, x_d \in \mathfrak{m}$ such that any prime ideal containing all of them has co-dimension at least d.

What prime ideals have co-dimension at least d? Since d is the Krull dimension of the entire ring R, the only prime ideal with co-dimension d is the maximal ideal $\mathfrak{m}$ itself. Therefore, the only prime ideal in R containing the ideal $(x_1, \ldots, x_d)$ is $\mathfrak{m}$.

In a Noetherian ring, if the only prime ideal containing an ideal I is the maximal ideal, then I is an $\mathfrak{m}$-primary ideal. This means that I must contain some power of the maximal ideal:

$$\mathfrak{m}^n \subseteq (x_1, \ldots, x_d) \subseteq \mathfrak{m}$$

This is exactly the definition of a system of parameters. We have successfully found a system of parameters consisting of exactly d elements.

Therefore, the minimal size of a system of parameters is less than or equal to d, the Krull dimension of R. □

Putting this theorem together with the results from our previous two lectures, we obtain the full cycle of inequalities:

$$\text{Krull Dimension} \leq \text{Hilbert Dimension}$$
$$\text{Hilbert Dimension} \leq \text{Parameters Dimension}$$
$$\text{Parameters Dimension} \leq \text{Krull Dimension}$$

This cycle proves that all three definitions of dimension are entirely equivalent for Noetherian local rings. The equivalence of these three distinct perspectives—topological (Krull), algebraic/combinatorial (Hilbert), and geometric (Parameters)—is one of the most beautiful and technical foundational results in commutative algebra.

59. Krull's Principal Ideal Theorem

This lecture focuses on some applications of the dimension theory we have established in the previous lectures. In particular, we will build towards proving Krull's Principal Ideal Theorem.

59.1. **Review of Dimension Definitions.** Recall that we have three distinct but equivalent definitions for the dimension of a Noetherian local ring $(R, \mathfrak{m})$:

- **Krull dimension:** The supremum of the lengths of strictly increasing chains of prime ideals.
- **Hilbert polynomial:** The degree of the polynomial $P(n) = \text{length}(R/\mathfrak{m}^n)$ for large n.
- **System of parameters:** The minimal cardinality of a set of elements generating an $\mathfrak{m}$-primary ideal.

Having these three equivalent definitions is exceptionally useful because we can switch between them depending on which is easiest to apply. For instance:

- The *Krull dimension* provides an easy lower bound for the dimension. To show $\dim(R) \geq n$, one simply needs to find a chain of prime ideals of length n.
- The *system of parameters* provides an easy upper bound. If one can find a system of parameters $x_1, \ldots, x_d$, then $\dim(R) \leq d$.

It is generally quite difficult to use Krull's definition to establish an upper bound, and equally difficult to use a system of parameters to establish a lower bound. Because the definitions are equal, we can combine them to determine the exact dimension. The *Hilbert polynomial* definition technically yields both bounds simultaneously, but it requires substantial effort to compute the actual polynomial.

59.2. **Dimension of the Completion.** Our first application of these equivalent definitions is straightforward but theoretically important.

Theorem 59.1. *If R is a Noetherian local ring, it has the same dimension as its completion $\hat{R}$.*

$$\dim(R) = \dim(\hat{R})$$

Proof. Proving this using Krull's definition or a system of parameters is highly non-trivial. However, it becomes completely trivial if we use the Hilbert polynomial definition.

Recall that the completion $\hat{R}$ is defined as the inverse limit of the quotients $R/\mathfrak{m}^n$. By the properties of completion, we have a natural isomorphism of the

quotient rings:

$$\frac{R}{\mathfrak{m}^n} \cong \frac{\hat{R}}{\hat{\mathfrak{m}}^n}$$

where $\hat{\mathfrak{m}}$ is the maximal ideal of the completion. Because the Hilbert definition of dimension depends only on the lengths of these quotient modules, and the quotient modules are isomorphic, their lengths are identical for all n. Therefore, they share the exact same Hilbert polynomial, implying they have the exact same dimension. □

59.3. **The Dimension of the Zeros of a Function.** The remaining applications in this lecture revolve around the geometric principle that the zeros of a single function should have codimension one. Let us formalize a first version of this principle.

Theorem 59.2. *Suppose x is an element of a Noetherian local ring R, and x is neither a unit nor a zero divisor. Then:*

$$\dim\left(\frac{R}{(x)}\right) = \dim(R) - 1$$

Before proving this, we must explain the necessity of the exceptions. Why do we exclude units and zero divisors? Let us analyze cases where the zeros of a function do *not* have codimension one:

(1) **The function has no zeros:** If x is a unit in R, then the ideal (x) is the entire ring R. The quotient $R/(x)$ is the zero ring, which has dimension -1 (or is undefined). The zero set is empty. Thus, we exclude units.
(2) **The function vanishes on an entire component:** Suppose the geometric space consists of two disconnected components (or intersecting components). If x is zero on one component but non-zero on the other, the zero locus includes an entire irreducible component of the original space, meaning its codimension is 0, not 1. Algebraically, we could find a function y that vanishes on the first component but not the second. Then $xy = 0$ everywhere, meaning x is a zero divisor. Thus, we exclude zero divisors.
(3) **Pathologies over non-algebraically closed fields:** Consider the function x^2+y^2 on the real affine plane $\mathbb{R}^2$. Its only real zero is the origin $(0, 0)$, which has codimension 2. At first glance, this seems like a counterexample. However, in ring theory, we look at the entire spectrum $\mathrm{Spec}(\mathbb{R}[x, y])$. This spectrum contains a 1-dimensional prime ideal generated by (x^2+y^2), and the function vanishes on this entire 1-dimensional ideal. It also vanishes at complex points like $(1, \pm i)$, which correspond to perfectly valid maximal ideals in the spectrum. Thus, algebraically, the zeros do indeed have codimension 1; the apparent counterexample only arises because the real Euclidean topology is blind to the generic and complex points of the spectrum.

Proof of Theorem 59.2. We previously established that quotienting by a non-zero divisor drops the dimension by at least 1:

$$\dim\left(\frac{R}{(x)}\right) \leq \dim(R) - 1$$

We must now prove the reverse inequality:

$$\dim(R) \leq \dim\left(\frac{R}{(x)}\right) + 1$$

Let $d = \dim(R/(x))$. By the system of parameters definition of dimension, we can pick elements $x_1, \ldots, x_d \in R$ whose images form a system of parameters in the quotient ring $R/(x)$.

This means that in $R/(x)$, the ideal generated by the images of $x_1, \ldots, x_d$ is primary for the maximal ideal $\mathfrak{m}/(x)$. Lifting this back to the original ring R, the ideal generated by $x_1, \ldots, x_d$ together with x must be $\mathfrak{m}$-primary.

Thus, the elements $\{x_1, \ldots, x_d, x\}$ form a system of parameters for the local ring R. The cardinality of this system is $d+1$. Since the dimension of R is defined as the minimal size of any system of parameters, we immediately obtain:

$$\dim(R) \leq d + 1 = \dim\left(\frac{R}{(x)}\right) + 1$$

Combining the two inequalities yields exact equality. □

59.4. **Krull's Principal Ideal Theorem.** We can now state a much more precise and general version of this principle, which applies to rings that are not necessarily local. This is the celebrated Krull's Principal Ideal Theorem.

Theorem 59.3 (Krull's Principal Ideal Theorem)**.** *Suppose R is a Noetherian ring, and $x \in R$ is an element that is not a zero divisor. Then all* minimal *prime ideals containing x have codimension exactly* 1*.*

Remark 59.4. Recall that the codimension of a prime ideal $\mathfrak{p}$ is the maximum length of a chain of prime ideals ending at $\mathfrak{p}$. This is sometimes referred to as the "height" of the prime ideal, but "codimension" preserves the geometric intuition.

Geometrically, we can interpret this theorem as follows: think of R as the ring of functions on some space $\mathrm{Spec}(R)$, and x as a specific function. The zeros of x form a closed subset $Z(x)$. The prime ideals containing x correspond to the irreducible subsets of $Z(x)$. The *minimal* primes containing x correspond exactly to the maximal irreducible components of $Z(x)$. Krull's theorem asserts that all of these irreducible components of the zero locus have a codimension of exactly 1. This formalizes the informal principle that the zeros of a well-behaved function possess codimension 1.

Proof. Let $\mathfrak{p}$ be a minimal prime ideal among those containing x. We want to show that the codimension of $\mathfrak{p}$ is 1.

The codimension of $\mathfrak{p}$ in R is exactly equal to the Krull dimension of the localized ring $R_{\mathfrak{p}}$. Thus, we localize our entire setup at $\mathfrak{p}$. In the localized ring $R_{\mathfrak{p}}$, the extended ideal $\mathfrak{p}R_{\mathfrak{p}}$ is the unique maximal ideal.

Because $\mathfrak{p}$ was chosen to be minimal among the primes containing x, there are no other prime ideals in $R_{\mathfrak{p}}$ containing x besides the maximal ideal itself. In a Noetherian ring, if the only prime ideal containing an element is the maximal ideal, then the ideal generated by that element is primary for the maximal ideal.

Therefore, the ideal (x) is $\mathfrak{p}R_{\mathfrak{p}}$-primary. This implies that the single element x constitutes a valid system of parameters for the local ring $R_{\mathfrak{p}}$.

Because the dimension of a local ring is bounded above by the size of any system of parameters, we have:

$$\dim(R_{\mathfrak{p}}) \leq 1$$

Consequently, the codimension of $\mathfrak{p}$ is either 0 or 1.

We must now rule out codimension 0. If the codimension of $\mathfrak{p}$ were 0, it would mean $\mathfrak{p}$ is a minimal prime ideal of the entire ring R. A fundamental property of Noetherian rings is that all elements belonging to a minimal prime ideal are zero divisors.

However, our initial hypothesis explicitly stated that $x \in \mathfrak{p}$ and x is *not* a zero divisor. This provides a direct contradiction. Thus, $\mathfrak{p}$ cannot be a minimal prime of R, meaning its codimension cannot be 0.

We are left with the only remaining possibility: the codimension of $\mathfrak{p}$ is exactly 1. This completes the proof. □

Part 8. Regular, Cohen-Macaulay, and Gorenstein Rings

60. Regular Local Rings

Let R be a Noetherian local ring with unique maximal ideal $\mathfrak{m}$. There is a nested family of properties for local rings, ordered from most specialized to most general:

- Regular local rings
- Local complete intersection rings
- Gorenstein rings
- Cohen-Macaulay rings
- General local rings

Very informally, regular local rings correspond to the non-singular points of manifolds or algebraic varieties. Local complete intersections correspond roughly to varieties defined by the minimum possible number of equations. Gorenstein rings are those possessing nice duality properties. Cohen-Macaulay rings have to do with not mixing components of different dimensions (for instance, the union of a plane and a line is a typical example of a ring that is not Cohen-Macaulay). Finally, we have general local rings. This section focuses on the "best" possible local rings: regular local rings.

60.1. **Definition and Basic Properties.** Recall that for any Noetherian local ring $(R, \mathfrak{m})$, the dimension of the cotangent space $\mathfrak{m}/\mathfrak{m}^2$ as a vector space over the residue field $R/\mathfrak{m}$ satisfies a fundamental inequality with the Krull dimension of R:

$$\dim_{R/\mathfrak{m}} \left(\frac{\mathfrak{m}}{\mathfrak{m}^2} \right) \geq \dim(R)$$

The reason this inequality holds is that a minimal set of generators for $\mathfrak{m}$ forms a system of parameters, and any system of parameters has cardinality at least the Krull dimension of the ring R.

Definition 60.1. A Noetherian local ring $(R, \mathfrak{m})$ is called a *regular local ring* if equality holds:

$$\dim_{R/\mathfrak{m}} \left(\frac{\mathfrak{m}}{\mathfrak{m}^2} \right) = \dim(R)$$

60.2. **Examples of Regular and Non-Regular Local Rings.**

Example 60.2 (A Non-Regular Local Ring)**.** Let us consider the coordinate ring of a cusp curve, defined by $y^2 = x^3$. We take the polynomial ring and localize at the origin, which corresponds to the maximal ideal (x, y):

$$R = \left(\frac{k[x, y]}{(y^2 - x^3)} \right)_{(x,y)}$$

Geometrically, the curve is singular at the origin. We verify this algebraically. The maximal ideal is $\mathfrak{m} = (x, y)$. The dimension of the cotangent space $\mathfrak{m}/\mathfrak{m}^2$ is 2, because neither x nor y is killed off by the relation $y^2 - x^3 \in \mathfrak{m}^2$.

On the other hand, the Krull dimension of R is 1 (it is a curve). Since:

$$\dim_k \left(\frac{\mathfrak{m}}{\mathfrak{m}^2} \right) = 2 > \dim(R) = 1$$

the ring R is not regular. This formally captures the fact that the origin is a singular point.

Example 60.3 (Power Series Rings)**.** Let R be a power series ring in n variables over a field k:

$$R = k[[x_1, \ldots, x_n]]$$

This is a local ring with maximal ideal $\mathfrak{m} = (x_1, \ldots, x_n)$. The cotangent space $\mathfrak{m}/\mathfrak{m}^2$ clearly has dimension n, as it possesses the basis $x_1, \ldots, x_n$.

We must figure out the Krull dimension of R. The easy way to determine this is to remember that quotienting a local ring by an element that is neither a zero divisor nor a unit decreases the dimension by 1. If we quotient by x_n, we obtain:

$$\frac{R}{(x_n)} \cong k[[x_1, \ldots, x_{n-1}]]$$

Every time we remove a variable in this manner, the dimension decreases by 1. By induction, it is obvious that $\dim(R) = n$. Because the dimension of the ring equals the dimension of its cotangent space, R is regular.

Example 60.4 (Polynomial Rings)**.** Suppose we take the polynomial ring $k[x_1, \ldots, x_n]$ and localize at the ideal $(x_1, \ldots, x_n)$. This represents the local ring at the origin of affine space. This local ring is also regular. This follows because its completion is the formal power series ring $k[[x_1, \ldots, x_n]]$. A local ring and its completion have the exact same Krull dimension and the exact same cotangent space. Since the completion is regular, the original local ring must also be regular.

This allows us to extend the definition to global rings.

Definition 60.5. A commutative ring S (not necessarily local) is called a *regular ring* if its localization $S_{\mathfrak{p}}$ is a regular local ring for all prime ideals $\mathfrak{p} \in \operatorname{Spec}(S)$.

In fact, a theorem by Serre states that the localization of any regular local ring is again regular. Because of this, it is sufficient to check that $S_{\mathfrak{m}}$ is regular only for all *maximal* ideals $\mathfrak{m}$. Thus, the polynomial ring $k[x_1, \ldots, x_n]$ is a globally regular ring, which perfectly aligns with our intuition that n-dimensional affine space should be non-singular everywhere under any reasonable definition.

60.3. **Structure of Complete Regular Local Rings.** While we cannot definitively classify all regular rings, we can describe the structure of *complete* regular local rings very precisely.

Suppose R is a complete regular local ring. Furthermore, suppose that R contains a field that maps isomorphically onto the residue field $R/\mathfrak{m} = k$. (We must explicitly state this condition because it is not universally true. For example, the p-adic integers $\mathbb{Z}_p$ form a complete regular local ring of dimension 1, but $\mathbb{Z}_p$ does not contain any field mapping onto its residue field $\mathbb{F}_p$).

If R does contain such a field, then R is isomorphic to a formal power series ring:

$$R \cong k[[x_1, \ldots, x_n]]$$

for some integer n.

Proof. We pick elements $x_1, \ldots, x_n \in \mathfrak{m}$ that form a basis for the cotangent space $\mathfrak{m}/\mathfrak{m}^2$. Because R is a complete Noetherian local ring, this choice naturally induces a surjective ring homomorphism from the formal power series ring onto R:

$$\phi \colon k[[x_1, \ldots, x_n]] \twoheadrightarrow R$$

We must show that ϕ is injective. Suppose there is some non-zero element a in the kernel of ϕ. Because a is neither a unit nor a zero divisor in the formal power series ring, the quotient ring:

$$\frac{k[[x_1, \ldots, x_n]]}{(a)}$$

has Krull dimension exactly $n - 1$. Since R is a further quotient of this ring, the dimension of R can be at most $n - 1$:

$$\dim(R) \leq n - 1$$

However, this gives a direct contradiction. By assumption, R is a regular local ring, meaning its Krull dimension must equal the dimension of its cotangent space, which is n. Therefore, the kernel must be trivial, and ϕ is an isomorphism. □

This is a special case of the Cohen Structure Theorem. When R does not contain a field, the conclusion is more complicated and requires the presence of a complete discrete valuation ring (such as the p-adic integers).

60.4. **Singularities of Hypersurfaces.** Let us apply this theory to ask which points on a hypersurface are singular.

Let $f(x_1, \ldots, x_n)$ be a polynomial, and consider the hypersurface defined by $f = 0$ in n-dimensional affine space. We want to evaluate whether the local ring at the origin $(0, \ldots, 0)$ is regular. To simplify notation, we assume the point of interest is the origin, which can always be achieved by translating coordinates.

The local ring at the origin is given by localizing the quotient ring:

$$R = \left(\frac{k[x_1, \ldots, x_n]}{(f)} \right)_{(x_1, \ldots, x_n)}$$

Assuming f is not a unit (meaning the zero locus is not empty) and not a zero divisor (since $k[x_1, \ldots, x_n]$ is an integral domain), the Krull dimension drops by exactly 1:

$$\dim(R) = n - 1$$

Next, we compute the dimension of the cotangent space $\mathfrak{m}/\mathfrak{m}^2$. For the ambient localized polynomial ring, the cotangent space is simply the n-dimensional vector space spanned by $x_1, \ldots, x_n$. For the hypersurface quotient, we must further quotient this vector space by the *linear part* of the polynomial f.

The dimension of this new cotangent space depends entirely on f:

- It is n if f has no linear part.
- It is $n - 1$ if f has a non-zero linear part.

By Taylor expansion, f has a linear part if and only if at least one of its first partial derivatives evaluates to a non-zero value at the origin:

$$\frac{\partial f}{\partial x_i}(0) \neq 0 \quad \text{for some } i$$

Since the ring is regular if and only if the dimension of the cotangent space (n or $n - 1$) equals the Krull dimension ($n - 1$), we see that the local ring is regular if and only if some partial derivative of f is non-zero.

Conversely, the singular points of the hypersurface $f = 0$ are exactly the points satisfying the simultaneous system of equations:

$$\begin{aligned} f &= 0 \\ \frac{\partial f}{\partial x_1} &= 0 \\ \frac{\partial f}{\partial x_2} &= 0 \\ &\cdots \\ \frac{\partial f}{\partial x_n} &= 0 \end{aligned}$$

For example, if $f(x, y) = y^2 - x^3$, the singular points must satisfy:

$$\begin{aligned} y^2 - x^3 &= 0 \\ -3x^2 &= 0 \\ 2y &= 0 \end{aligned}$$

Assuming the characteristic of the field is not 2 or 3, the partial derivatives force $x = 0$ and $y = 0$. Thus, the origin is the unique singular point, perfectly confirming our geometric intuition that the cusp is the only "funny" point on the curve.

61. Examples of Regular Local Rings

This lecture focuses on examples of regular local rings. In previous lectures, we examined regular local rings primarily over algebraically closed fields. However, when working over non-algebraically closed fields in characteristic $p > 0$, some rather counterintuitive phenomena can occur. The first couple of examples will illustrate these pathologies.

61.1. **Pathologies in Positive Characteristic.** Suppose we take a variety V over a general field k. Its coordinate ring is the ring of polynomials over k modulo some ideal I:

$$R = \frac{k[x_1, \ldots, x_n]}{I}$$

We can localize this ring at some point. If we are working over an algebraically closed field, a point simply corresponds to coordinates $(a_1, \ldots, a_n) \in k^n$. If the field is not algebraically closed, a point corresponds more generally to a maximal ideal $\mathfrak{m}$, and we can form the local ring $R_{\mathfrak{m}}$.

We ask the following question: *If the local rings at all points over k are regular, is the same true over the algebraic closure $\bar{k}$?*

In characteristic $p > 0$, this is false. Zariski discovered a famous counterexample that initially caused significant confusion.

Example 61.1 (Zariski's Counterexample)**.** Let k be a field of characteristic $p > 0$, and let $a \in k$ be an element that is not a p-th power in k. Consider the curve defined by:

$$x^p + y^p = a$$

The coordinate ring is:

$$R = \frac{k[x, y]}{(x^p + y^p - a)}$$

Consider the maximal ideal generated by y. If we quotient R by (y), we get:

$$\frac{R}{(y)} = \frac{k[x]}{(x^p - a)}$$

Because a is not a p-th power in k, the polynomial $x^p - a$ is irreducible. Therefore, this quotient is a field, which proves that the ideal (y) is indeed a maximal ideal in R.

If we localize at this maximal ideal, the local ring has dimension 1. Furthermore, the maximal ideal is generated by a single element, y. Because the number of generators matches the dimension, the local ring is *regular*, and its cotangent space is 1-dimensional. Thus, over the non-algebraically closed field k, we have a regular local ring.

Now, let us examine what happens when we move to the algebraic closure $\bar{k}$. Over $\bar{k}$, there exists some element b such that $b^p = a$. In characteristic p, the algebraic relation becomes:

$$x^p + y^p - a = x^p + y^p - b^p = (x + y - b)^p$$

Thus, the coordinate ring over the algebraic closure is:

$$\bar{R} = \frac{\bar{k}[x, y]}{((x + y - b)^p)}$$

This ring is completely riddled with nilpotent elements. The element $(x + y - b)$ is non-zero but its p-th power is zero. Consequently, this ring is non-regular at *every* maximal ideal. We started with a variety that was non-singular (regular at every point) over k, but passing to the algebraically closed field caused it to become highly singular everywhere.

61.2. **Zero-Dimensional Case and Tensor Products.** To better understand the algebraic mechanism behind this failure, it is instructive to simplify the problem to dimension 0. Instead of the curve, consider the zero-dimensional variety defined by:

$$x^p - a = 0$$

Over k, the ring is:

$$L = \frac{k[x]}{(x^p - a)}$$

Since a is not a p-th power, L is a field. It has only one maximal ideal (the zero ideal), the local ring is the field itself, and it is trivially regular.

However, over the algebraic closure $\bar{k}$, the ring splits:

$$L \otimes_k \bar{k} = \frac{\bar{k}[x]}{(x^p - a)} = \frac{\bar{k}[x]}{(x - b)^p}$$

This ring is not a field and contains nilpotent elements, meaning it is not a regular local ring.

This phenomenon can be understood through the lens of tensor products. We are effectively examining the tensor product of two fields, L and M, over a base field k.

Theorem 61.2. *Let L and M be fields containing k, with L being a finite extension of k.*

- *In characteristic 0, $L \otimes_k M$ is always a finite product of fields. Thus, every localization is a field, and it is regular at every point.*

- *In characteristic $p > 0$, $L \otimes_k M$ can contain nilpotent elements and fail to be a product of fields, resulting in non-regular rings.*

In geometric terms, this implies a highly disconcerting fact: in characteristic $p > 0$, the product of two non-singular varieties can be singular. To rectify this, Zariski introduced a stricter condition.

Definition 61.3. A local ring R over a field k is called *geometrically regular* if the tensor product $R \otimes_k \bar{k}$ is a regular ring, where $\bar{k}$ is the algebraic closure of k.

In our previous example, the field $k[x]/(x^p - a)$ is regular, but it is not geometrically regular. Grothendieck later generalized this via the concept of *smoothness* for morphisms of rings, which reduces to geometric regularity when the base ring is a field.

61.3. **An Example from Algebraic Number Theory.** We conclude with a classic example of regular local rings in algebraic number theory, focusing on an order in a quadratic field.

Consider the ring $R = \mathbb{Z}[\sqrt{-3}]$. As is well known, this ring is not a Unique Factorization Domain (UFD). We have the non-unique factorization:

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

To remedy this, one takes its *normalization* (the integral closure in its quotient field), which yields the ring of Eisenstein integers:

$$\tilde{R} = \mathbb{Z}\left[\frac{1 + \sqrt{-3}}{2}\right]$$

This normalization is a UFD. We will examine the local rings of both R and $\tilde{R}$ at a specific prime.

Let us define a maximal ideal $\mathfrak{m}$ in $R = \mathbb{Z}[\sqrt{-3}]$:

$$\mathfrak{m} = (2, 1 + \sqrt{-3})$$

The quotient $R/\mathfrak{m}$ is simply the field with two elements. We want to check if the localization $R_{\mathfrak{m}}$ is a regular local ring.

We analyze the cotangent space by calculating the length of $R/\mathfrak{m}^2$. The square of the ideal is generated by $4, 2(1 + \sqrt{-3})$, and $(1 + \sqrt{-3})^2 = -2 + 2\sqrt{-3}$. Simplifying this in R, one can check that the quotient $R/\mathfrak{m}^2$ has length 3. Since the length of $R/\mathfrak{m}$ is 1, the cotangent space $\mathfrak{m}/\mathfrak{m}^2$ must have dimension 2 as a vector space over the residue field.

However, the Krull dimension of the ring R is 1. Because the dimension of the cotangent space (2) is strictly greater than the dimension of the ring (1), the local ring $R_{\mathfrak{m}}$ is *not regular*.

Conversely, if we look at the normalization $\tilde{R}$, it is a Dedekind domain. Every localization at a maximal ideal is a Discrete Valuation Ring (DVR). Since DVRs are 1-dimensional and their maximal ideals are generated by a single element (the uniformizer), they are inherently regular local rings. Geometrically, taking the normalization corresponds to resolving the singularities of the spectrum, ensuring every point becomes regular.

61.3.1. *A Subtle Pitfall with Prime Preimages.* There is a subtle trap when comparing prime ideals between these two rings. Consider the ideal (2) in the normalized ring $\tilde{R}$. One can check that:

$$\frac{\tilde{R}}{(2)} \cong \mathbb{F}_4$$

Since the quotient is a field (the finite field of 4 elements), the ideal (2) is maximal and therefore prime in $\tilde{R}$.

However, if we look at the ideal (2) in the original ring $R = \mathbb{Z}[\sqrt{-3}]$, it is *not prime*. The quotient $R/(2)$ has zero divisors and is isomorphic to a sum of two fields of order 2.

This presents an apparent paradox: the inverse image of a prime ideal under a ring homomorphism must be prime. The resolution is that the inverse image of the prime ideal $(2) \subset \tilde{R}$ is *not* the ideal generated by 2 in R. If one takes a set of generators for a prime ideal in an extension ring, the inverse image of the ideal is not necessarily generated by the inverse images of those generators. The actual inverse image of the ideal $(2) \subset \tilde{R}$ in R is the maximal ideal $\mathfrak{m} = (2, 1 + \sqrt{-3})$ that we analyzed earlier.

62. Cohen-Macaulay Local Rings

In the previous lectures, we discussed regular rings and regular local rings, noting that they sit at the top of a hierarchy of favorable ring conditions. There is a sequence of progressively broader classes of local rings: regular rings, local complete intersection rings, Gorenstein rings, Cohen-Macaulay rings, and finally, arbitrary Noetherian local rings. In this lecture, we turn our attention to the Cohen-Macaulay local rings.

Historically, these rings are named after Irvin Cohen and Francis Sowerby Macaulay. Macaulay was an English mathematician working in the early 20th century. During that period, university mathematics positions in England were scarce, so highly capable mathematicians often became either high school teachers or Church of England vicars; Macaulay himself was a high school teacher.

62.1. **Regular Sequences and Depth.** To motivate the definition of a Cohen-Macaulay ring, we must first define the depth of a ring via regular sequences. Throughout this section, let R be a Noetherian local ring with unique maximal ideal $\mathfrak{m}$.

Recall from our discussion on dimension theory that if an element $x \in \mathfrak{m}$ is neither a zero divisor nor a unit, quotienting by the ideal generated by x drops the Krull dimension of the ring by exactly one:

$$\dim\left(\frac{R}{(x)}\right) = \dim(R) - 1$$

We can iterate this process by choosing a sequence of elements that are consecutively non-zero divisors.

Definition 62.1. A sequence of elements $x_1, x_2, \ldots, x_n \in \mathfrak{m}$ is called a *regular sequence* for R if:

(1) x_1 is not a zero divisor in R.
(2) x_2 is not a zero divisor in the quotient ring $R/(x_1)$.
(3) Generally, x_i is not a zero divisor in the quotient ring $R/(x_1, \ldots, x_{i-1})$ for each $1 \leq i \leq n$.

This is yet another instance where the word "regular" is heavily overloaded in commutative algebra.

Notice that each time we extend the regular sequence by quotienting out by the next element, the Krull dimension of the resulting ring drops by one. Consequently, the length of any regular sequence $x_1, \ldots, x_n$ is bounded above by the dimension of the local ring:

$$n \leq \dim(R)$$

Definition 62.2. The *depth* of a local ring R is defined as the maximum length of a regular sequence in $\mathfrak{m}$. By the inequality above, we always have $\operatorname{depth}(R) \leq \dim(R)$.

Definition 62.3. A Noetherian local ring R is a *Cohen-Macaulay ring* if its depth is equal to its Krull dimension:

$$\operatorname{depth}(R) = \dim(R)$$

62.2. **Examples of Cohen-Macaulay Rings.**

Example 62.4 (Dimension Zero Rings)**.** Any Noetherian local ring of dimension zero is trivially Cohen-Macaulay. Its Krull dimension is 0, and the maximum length of a regular sequence is 0 (the empty sequence).

Example 62.5 (Regular Local Rings)**.** Every regular local ring is Cohen-Macaulay. To prove this, recall that a regular local ring R is necessarily an integral domain. The proof of this relies on examining its associated graded ring:

$$\operatorname{gr}_{\mathfrak{m}}(R) = \frac{R}{\mathfrak{m}} \oplus \frac{\mathfrak{m}}{\mathfrak{m}^2} \oplus \frac{\mathfrak{m}^2}{\mathfrak{m}^3} \oplus \ldots$$

For a regular local ring, this graded ring is isomorphic to a polynomial algebra over the residue field, which is clearly an integral domain. If the associated graded ring of a filtered ring is an integral domain, the original ring must also be an integral domain. To see this, suppose $a, b \in R$ are non-zero such that $ab = 0$. If $a \in \mathfrak{m}^i \setminus \mathfrak{m}^{i+1}$ and $b \in \mathfrak{m}^j \setminus \mathfrak{m}^{j+1}$, their product in the graded piece $\mathfrak{m}^{i+j}/\mathfrak{m}^{i+j+1}$ must be non-zero (since the graded ring is an integral domain), contradicting $ab = 0$.

Because R is an integral domain, any non-zero element $x_1 \in \mathfrak{m}$ is not a zero divisor. Furthermore, if we choose $x_1 \in \mathfrak{m} \setminus \mathfrak{m}^2$, the quotient ring $R/(x_1)$ is again a regular local ring (of dimension $\dim(R) - 1$). We can iteratively pick non-zero divisors until the dimension drops to zero, yielding a regular sequence of length $\dim(R)$.

62.3. **Non-Examples and Equidimensionality.** It is highly instructive to analyze local rings that fail to be Cohen-Macaulay.

Example 62.6 (Embedded Point)**.** Let R be the formal power series ring in two variables modulo the ideal (y^2, xy):

$$R = \frac{k[[x, y]]}{(y^2, xy)}$$

Geometrically, this represents a 1-dimensional line (the x-axis, where $y = 0$) with an embedded nilpotent "fuzz" or generic point sticking out at the origin.

The maximal ideal is $\mathfrak{m} = (x, y)$. Notice that the element $y \in R$ is non-zero, but it annihilates every generator of the maximal ideal: $x \cdot y = 0$ and $y \cdot y = 0$. Therefore, *all* elements in the maximal ideal $\mathfrak{m}$ are zero divisors.

Since we cannot even pick a first element for a regular sequence, we have:

$$\operatorname{depth}(R) = 0$$
$$\dim(R) = 1$$

Thus, R is not Cohen-Macaulay. This pathology generally arises when a ring has irreducible components of different dimensions or embedded primes.

Example 62.7 (Non-Equidimensional Ring)**.** Let us construct a ring with no nilpotent elements that is not Cohen-Macaulay. Consider the union of a 2-dimensional plane ($z = 0$) and a 1-dimensional line ($x = y = 0$) intersecting at the origin. We define the local ring via completion:

$$R = \frac{k[[x, y, z]]}{(xz, yz)}$$

The Krull dimension of R is 2 (due to the plane). Let us attempt to find a regular sequence. We can pick $x_1 = x + z$ as our first non-zero divisor. Taking the quotient yields:

$$\frac{R}{(x + z)} \cong \frac{k[[x, y, z]]}{(xz, yz, x + z)}$$

Setting $z = -x$, we can eliminate z, resulting in the ring:

$$\frac{k[[x, y]]}{(-x^2, -xy)} = \frac{k[[x, y]]}{(x^2, xy)}$$

This is isomorphic to the ring from our previous example. In this quotient, all elements of the maximal ideal are zero divisors, so the regular sequence cannot be extended further.

Thus, $\operatorname{depth}(R) = 1$, but $\dim(R) = 2$. R is not Cohen-Macaulay. This ring fails to be *equidimensional* because it consists of a 2-dimensional component and a 1-dimensional component. It is a general fact that Cohen-Macaulay rings are always equidimensional.

Example 62.8 (Equidimensional but not Cohen-Macaulay)**.** Is the converse true? Does equidimensionality guarantee the Cohen-Macaulay property? No. We can construct an equidimensional local ring without nilpotent elements that is still not Cohen-Macaulay.

Consider two 2-dimensional planes intersecting at a single point in 4-dimensional space. (In 3D, two planes must intersect in a line, but in 4D they can intersect

in a point). Let the planes be $w = x = 0$ and $y = z = 0$. The corresponding completed local ring is:

$$R = \frac{k[[w, x, y, z]]}{(wy, wz, xy, xz)}$$

This ring is equidimensional of dimension 2. We pick our first regular element, $x_1 = w - y$. Quotienting by x_1 corresponds to setting $w = y$:

$$\frac{R}{(w-y)} \cong \frac{k[[x, y, z]]}{(y^2, yz, xy, xz)}$$

In this quotient ring, the element y is non-zero, but it kills the entire maximal ideal (x, y, z). Thus, every element in the maximal ideal is a zero divisor. We find $\operatorname{depth}(R) = 1$ while $\dim(R) = 2$. Being Cohen-Macaulay is a strictly stronger condition than merely being equidimensional.

62.4. **Depth for Modules and Homological Characterization.** To show that the depth of a ring does not depend on the specific choice of the elements in the regular sequence, we generalize the concept of depth to modules and utilize homological algebra.

Definition 62.9. Let M be a finitely generated R-module. A sequence $x_1, \ldots, x_n \in \mathfrak{m}$ is a *regular sequence for* M if x_1 is not a zero divisor on M, and x_i is not a zero divisor on the quotient $M/(x_1, \ldots, x_{i-1})M$ for each i.

The depth of a module M is the length of the longest regular sequence for M. The following theorem relates depth to the vanishing of Ext functors, proving that depth is a homological invariant.

Theorem 62.10. *For a finitely generated module M over a Noetherian local ring $(R, \mathfrak{m})$, the following statements are equivalent:*

1. *There exists a regular sequence for M of length n.*
2. $\operatorname{Ext}^i_R(R/\mathfrak{m}, M) = 0$ *for all $i < n$.*
3. *Any regular sequence for M can be extended to a regular sequence of length n.*

Proof. We prove that (1) implies (2), and (2) implies (3). The implication (3) implies (1) is completely trivial.

(1) $\implies$ (2): We proceed by induction on n. Let $x_1, \ldots, x_n$ be a regular sequence for M. We have a short exact sequence given by multiplication by x_1:

$$0 \to M \xrightarrow{\times x_1} M \to \frac{M}{x_1 M} \to 0$$

This sequence is exact precisely because x_1 is not a zero divisor on M. Applying the functor $\operatorname{Hom}_R(R/\mathfrak{m}, -)$ yields a long exact sequence of Ext modules. The relevant portion is:

$$\cdots \to \operatorname{Ext}^{i-1}_R\left(\frac{R}{\mathfrak{m}}, \frac{M}{x_1 M}\right) \to \operatorname{Ext}^i_R\left(\frac{R}{\mathfrak{m}}, M\right) \xrightarrow{\times x_1} \operatorname{Ext}^i_R\left(\frac{R}{\mathfrak{m}}, M\right) \to \ldots$$

By our inductive hypothesis, the regular sequence $x_2, \ldots, x_n$ of length $n-1$ for the module M/x_1M guarantees that the leftmost term $\operatorname{Ext}_R^{i-1}(R/\mathfrak{m}, M/x_1M)$ vanishes for $i-1 < n-1$, which corresponds to $i < n$.

Because this term is zero, the map induced by multiplication by x_1 is injective. However, the R-module structure on $\operatorname{Ext}_R^i(R/\mathfrak{m}, M)$ is induced by its first argument, $R/\mathfrak{m}$. Since $x_1 \in \mathfrak{m}$, multiplication by x_1 must act as the zero map.

An injective map that is also the zero map implies that the module itself must be the zero module. Therefore:

$$\operatorname{Ext}_R^i\left(\frac{R}{\mathfrak{m}}, M\right) = 0 \quad \text{for } i < n$$

(2) $\implies$ (3): We assume $\operatorname{Ext}_R^i(R/\mathfrak{m}, M) = 0$ for $i < n$ and wish to show that any regular sequence can be extended to length n.

Consider the base case $n = 1$. The condition translates to $\operatorname{Ext}_R^0(R/\mathfrak{m}, M) = \operatorname{Hom}_R(R/\mathfrak{m}, M) = 0$. If this Hom group were non-zero, it would mean that $R/\mathfrak{m}$ embeds into M, which implies that the maximal ideal $\mathfrak{m}$ is an associated prime of M. Since it is zero, $\mathfrak{m} \notin \operatorname{Ass}(M)$.

The set of zero divisors of M is exactly the union of its associated primes. Because M is finitely generated, this is a finite union of prime ideals. By the Prime Avoidance Lemma, the maximal ideal $\mathfrak{m}$ cannot be entirely contained in a finite union of proper prime ideals. Thus, there exists an element $x_1 \in \mathfrak{m}$ that avoids all associated primes, meaning x_1 is not a zero divisor on M. This establishes a regular sequence of length 1.

For $n > 1$, after selecting the non-zero divisor x_1, we simply apply the inductive hypothesis to the quotient module M/x_1M, carefully utilizing the long exact sequence of Ext to confirm that $\operatorname{Ext}_R^i(R/\mathfrak{m}, M/x_1M) = 0$ for $i < n-1$. The details are standard homological bookkeeping left as an exercise. □

62.5. **Corollaries.** Because condition (2) in the theorem above depends on the module M and the maximal ideal $\mathfrak{m}$, but not on the elements x_i, we obtain the following powerful corollary:

Corollary 62.11. *Any two maximal regular sequences for a module M over a Noetherian local ring have the exact same length.*

This assures us that we do not need to worry about making a "bad choice" for our first regular element; any valid choice can eventually be extended to the global depth of the module.

Corollary 62.12. *Any quotient of a regular local ring R by a regular sequence (possibly non-maximal) is a Cohen-Macaulay ring.*

Proof. Let the regular local ring be R, and let $x_1, \ldots, x_i$ be a regular sequence. The Krull dimension of the quotient ring $R/(x_1, \ldots, x_i)$ is $\dim(R) - i$.

Because all maximal regular sequences have the same length, we can extend $x_1, \ldots, x_i$ to a maximal regular sequence of R of length $\dim(R)$. The remaining elements $x_{i+1}, \ldots, x_{\dim(R)}$ inherently form a regular sequence of length $\dim(R) - i$ in the quotient ring $R/(x_1, \ldots, x_i)$. Since the depth is at least $\dim(R) - i$ and the dimension is $\dim(R) - i$, the quotient ring is Cohen-Macaulay. □

As an immediate application, any hypersurface singularity is defined by quotienting a regular local ring by a single non-zero divisor. Thus, hypersurface singularities are always Cohen-Macaulay. This covers the vast majority of singularities one encounters in basic algebraic geometry (such as any singularity on a plane curve), demonstrating that the Cohen-Macaulay condition is quite prevalent.

63. Koszul Complex

This section introduces a gadget called the *Koszul complex*. We will first recall the definition of regular sequences, explain the construction of the Koszul complex, and demonstrate how to use it to solve problems concerning regular sequences.

63.1. **Regular Sequences and Permutations.** Recall that if we have a ring R, a sequence of elements $x_1, x_2, \ldots, x_n$ is called a *regular sequence* if it satisfies two conditions:

(1) For each $i \geq 1$, x_i is not a zero divisor in the quotient ring $R/(x_1, \ldots, x_{i-1})$. (For $i = 1$, x_1 is not a zero divisor in R).
(2) The quotient ring is non-trivial, meaning:

$$\frac{R}{(x_1, \ldots, x_n)} \neq 0$$

The non-triviality condition is strictly necessary; without it, the zero ring would allow for many pathological and silly examples of regular sequences.

Because the definition inherently depends on the order of the quotienting process, we must ask: is a permutation of a regular sequence also a regular sequence? If true, it would greatly simplify matters. However, the answer is generally no.

Example 63.1. Let R be the polynomial ring in three variables quotiented by xz:

$$R = \frac{k[x, y, z]}{(xz)}$$

This is the coordinate ring of two planes meeting along a line. Consider the sequence of elements $x - 1$ and xy.

This sequence is regular. The element $x - 1$ is not a zero divisor in R. Quotienting by $(x - 1)$ corresponds to setting $x = 1$, yielding:

$$\frac{R}{(x-1)} \cong \frac{k[y, z]}{(z)} \cong k[y]$$

In the integral domain $k[y]$, the element $xy \equiv y$ is clearly not a zero divisor. Thus, $x - 1, xy$ is a regular sequence.

Now, consider the permuted sequence $xy, x - 1$. In the original ring R, xy is immediately a zero divisor, because:

$$(xy)z = y(xz) = 0$$

Thus, xy fails the first condition, meaning the permuted sequence is not regular.

Despite this failure in general rings, it turns out that a permutation of a regular sequence remains regular if R is a local ring. For a local ring with maximal ideal $\mathfrak{m}$, the non-triviality condition $R/(x_1, \dots, x_n) \neq 0$ implies that all elements x_i of the regular sequence must reside in $\mathfrak{m}$. The proof relies heavily on the Koszul complex.

63.2. **Definition of the Koszul Complex.** We can define the *Koszul complex* for any sequence of elements $x_1, \dots, x_n \in R$, though its exactness properties will depend on the sequence being regular.

63.2.1. *The Base Cases.* For $n = 1$, the Koszul complex for an element x_1 is the simple sequence:

$$0 \to R \xrightarrow{\times x_1} R \to \frac{R}{(x_1)} \to 0$$

This sequence is obviously exact if and only if x_1 is not a zero divisor (which is the main condition for x_1 to be a regular sequence of length 1).

For $n = 2$, the Koszul complex for x_1, x_2 is:

$$0 \to R \xrightarrow{d_2} R^2 \xrightarrow{d_1} R \to \frac{R}{(x_1, x_2)} \to 0$$

The differentials are defined as follows. The map d_2 takes $1 \in R$ to $(x_1, x_2) \in R^2$. The map d_1 takes an element $(a, b) \in R^2$ to:

$$d_1(a, b) = x_2 a - x_1 b$$

The minus sign is absolutely critical to ensure that the composition $d_1 \circ d_2 = 0$. The final map is the standard canonical projection.

63.2.2. *The General Construction.* In general, the terms of the Koszul complex are constructed using the exterior powers of the free module R^n:

$$0 \to \bigwedge^n R^n \to \cdots \to \bigwedge^2 R^n \to \bigwedge^1 R^n \to R \to \frac{R}{(x_1, \dots, x_n)} \to 0$$

This yields a sequence of free modules with binomial ranks:

$$0 \to R \to R^n \to R^{\binom{n}{2}} \to \cdots \to R^{\binom{n}{n-2}} \to R^n \to R \to \frac{R}{(x_1, \dots, x_n)} \to 0$$

An elegant way to inductively define the complex and its differentials is by splicing together two copies of the Koszul complex for $n - 1$ elements. Suppose we have the complex for $x_1, \dots, x_{n-1}$:

$$0 \to R \to R^{n-1} \to \cdots \to R \to \frac{R}{(x_1, \dots, x_{n-1})} \to 0$$

We write this complex twice, forming two rows. We define vertical maps between the identical rows given entirely by multiplication by the element x_n. By introducing alternating minus signs into the differentials, we can take the direct sum of the shifted terms to form a new complex. We truncate and adjust the final quotient term by replacing it with $R/(x_1, \dots, x_n)$. This recursive mapping cone construction yields the full Koszul complex for n elements.

By explicitly writing out the differentials, one observes that up to isomorphism, the Koszul complex does not depend on the order of the sequence $x_1, \ldots, x_n$.

63.3. **Exactness and Free Resolutions.**

Theorem 63.2. *If $x_1, \ldots, x_n$ is a regular sequence in R, then the associated Koszul complex is exact.*

Proof Sketch. We proceed by induction on n. The base cases are easy to verify. For the inductive step, exactness relies on the mapping cone construction described above. The only non-trivial part of verifying exactness occurs at the very end of the spliced complex:

$$R^{n-1} \xrightarrow{d} R \to \frac{R}{(x_1, \ldots, x_{n-1})}$$

We require that if an element $b \in R$ maps to 0 in $R/(x_1, \ldots, x_{n-1})$, it must be in the image of the differential. By standard diagram chasing on the spliced rows, this exactness holds provided that the multiplication map by x_n:

$$\frac{R}{(x_1, \ldots, x_{n-1})} \xrightarrow{\times x_n} \frac{R}{(x_1, \ldots, x_{n-1})}$$

is injective. However, the injectivity of x_n on this quotient is exactly the definition of x_n extending the regular sequence $x_1, \ldots, x_{n-1}$. Thus, the regular sequence condition guarantees the exactness of the complex. □

This result is profoundly important for homological algebra. It provides a finite free resolution of the module $M = R/(x_1, \ldots, x_n)$:

$$0 \to \bigwedge^n R^n \to \cdots \to \bigwedge^1 R^n \to R \to M \to 0$$

Having a finite free resolution allows one to easily calculate Tor and Ext groups for M. Furthermore, it demonstrates that modules generated by regular sequences exhibit heavily restricted homological dimensions.

63.4. **The Converse and Nakayama's Lemma.** We now ask the converse: if the Koszul complex on $x_1, \ldots, x_n$ is exact, is the sequence necessarily regular?

In general rings, the answer is no. Reusing our previous counterexample, the sequence $x_1 = xy$ and $x_2 = x - 1$ is not regular. However, because the Koszul complex is independent of order up to isomorphism, the complex for $(xy, x-1)$ is isomorphic to the complex for $(x-1, xy)$. Since $(x-1, xy)$ is a regular sequence, its complex is exact. Thus, we have an exact Koszul complex for a non-regular sequence.

However, the converse *is* true if R is a local ring.

Theorem 63.3. *Let R be a local Noetherian ring. If the Koszul complex for elements $x_1, \ldots, x_n \in \mathfrak{m}$ is exact, then $x_1, \ldots, x_n$ is a regular sequence. Consequently, any permutation of a regular sequence in a local ring remains regular.*

Proof. We rely on the first homology group of the Koszul complex, denoted $H_1(x_1, \ldots, x_n)$. The recursive splicing construction of the complex yields a long exact sequence in homology. The relevant segment is:

$$H_1(x_1, \ldots, x_{n-1}) \xrightarrow{\times x_n} H_1(x_1, \ldots, x_{n-1}) \to H_1(x_1, \ldots, x_n) \to \\ \to \frac{R}{(x_1, \ldots, x_{n-1})} \xrightarrow{\times x_n} \frac{R}{(x_1, \ldots, x_{n-1})}$$

We assume the Koszul complex for n elements is exact, which implies its homology vanishes: $H_1(x_1, \ldots, x_n) = 0$. Substituting 0 into the sequence gives two immediate consequences:

(1) The rightmost map is injective. This means x_n is not a zero divisor on the quotient $R/(x_1, \ldots, x_{n-1})$. If we assume inductively that $x_1, \ldots, x_{n-1}$ is a regular sequence, this proves $x_1, \ldots, x_n$ is regular.
(2) The leftmost map is surjective. This means multiplication by x_n maps the module $H_1(x_1, \ldots, x_{n-1})$ surjectively onto itself:

$$x_n \cdot H_1(x_1, \ldots, x_{n-1}) = H_1(x_1, \ldots, x_{n-1})$$

We now invoke Nakayama's Lemma. Let $M = H_1(x_1, \ldots, x_{n-1})$. Because R is local, the elements x_i belong to the maximal ideal $\mathfrak{m}$, so $x_n \in \mathfrak{m}$. The surjectivity condition implies $\mathfrak{m}M = M$. By Nakayama's Lemma (applicable because homology modules of finite free complexes over Noetherian rings are finitely generated), we must have:

$$M = H_1(x_1, \ldots, x_{n-1}) = 0$$

This vanishing homology indicates that the Koszul complex for $n - 1$ elements is also exact. By induction down to $n = 1$, we sequentially prove that each x_i is a non-zero divisor on the prior quotient. Thus, the sequence is proven to be regular.

Because the exactness of the Koszul complex is invariant under permutations, and exactness implies regularity in local rings, any permutation of a regular sequence in a local ring is regular. □

64. Gorenstein Rings

This lecture will be about Gorenstein local rings. Throughout this section, we assume all rings are Noetherian and local. Gorenstein rings possess a profound duality property.

64.1. **The Zero-Dimensional Case.** In order to define Gorenstein rings, we first examine the zero-dimensional case, which is somewhat simpler than the general definition.

Definition 64.1. Let R be a zero-dimensional, Noetherian, local ring with maximal ideal $\mathfrak{m}$ and residue field $k = R/\mathfrak{m}$. The ring R is *Gorenstein* if the vector space dimension of homomorphisms from k to R is exactly one:

$$\dim_k \operatorname{Hom}_R(k, R) = 1$$

More generally, we can define the dual of an R-module M as $\mathrm{Hom}_R(M, R)$. If R is Gorenstein, this duality behaves exceptionally well. For instance, for any finitely generated R-module M, the double dual is isomorphic to M:

$$\mathrm{Hom}_R(\mathrm{Hom}_R(M, R), R) \cong M$$

For general zero-dimensional rings, this standard dual is inadequate. Instead, there exists an alternative definition utilizing a special module called the *dualizing module*, denoted ω. The dual of M is defined as $\mathrm{Hom}_R(M, \omega)$. A zero-dimensional ring is Gorenstein precisely when its dualizing module is isomorphic to the ring itself:

$$\omega \cong R$$

64.1.1. *Examples of Zero-Dimensional Rings.* Visualizing zero-dimensional rings as stacked blocks representing basis elements over k provides excellent geometric intuition for the Gorenstein property. The dimension of $\mathrm{Hom}_R(k, R)$ corresponds to the number of "blocks at the bottom" of the tower (i.e., elements annihilated by the maximal ideal).

Example 64.2. Let $R = k[[x]]/(x^5)$. The maximal ideal is (x), and the elements killed by x are multiples of x^4. Thus:

$$\mathrm{Hom}_R(k, R) = \mathrm{span}_k\{x^4\}$$

This is a 1-dimensional space. Visually, R is a single tower of five blocks $(1, x, x^2, x^3, x^4)$, with exactly one block at the bottom. Thus, R is Gorenstein.

Example 64.3. Let $R = k[x, y]/(x^2, xy, y^2)$. The basis is $1, x, y$. The elements killed by $\mathfrak{m} = (x, y)$ are exactly the linear span of x and y. Thus:

$$\dim_k \mathrm{Hom}_R(k, R) = 2$$

Visually, this ring has one block on top (1) and two blocks on the bottom (x, y). Because the dimension is 2, it is not Gorenstein.

Example 64.4. Let $R = k[x, y]/(x^2, y^2)$. The basis is $1, x, y, xy$. The only element annihilated by both x and y is xy. Thus:

$$\dim_k \mathrm{Hom}_R(k, R) = 1$$

Visually, this is a diamond shape: 1 on top, x and y in the middle, and xy at the bottom. There is only one block at the bottom, so R is Gorenstein.

An informal geometric interpretation is that a zero-dimensional Gorenstein ring "looks the same upside down", reflecting its self-duality. If you invert the tower of blocks for $k[[x]]/(x^5)$ or $k[x, y]/(x^2, y^2)$, the structural shape is preserved. In contrast, inverting $k[x, y]/(x^2, xy, y^2)$ yields two blocks on top and one on the bottom, which breaks the structural symmetry.

64.2. **Higher Dimensions and Grothendieck's Definition.** In higher dimensions, the definition becomes significantly more sophisticated, making use of homological algebra.

Definition 64.5. Let R be a local Noetherian ring of Krull dimension d, with residue field k. The ring R is *Gorenstein* if the Ext groups satisfy:

$$\operatorname{Ext}^i_R(k, R) = 0 \quad \text{for } i \neq d$$
$$\dim_k \operatorname{Ext}^d_R(k, R) = 1$$

For $d = 0$, Ext^0_R is exactly Hom_R, completely recovering our zero-dimensional definition.

This definition was established by Alexander Grothendieck. He named these rings after Daniel Gorenstein, a mathematician best known for directing the classification of finite simple groups. Before working in group theory, Gorenstein studied algebraic geometry and proved a theorem regarding plane curve singularities, identifying a property that later turned out to be equivalent to the Gorenstein condition. This inspired Grothendieck's naming. (Mathematical folklore humorously suggests that Gorenstein himself claimed he did not understand the definition of a Gorenstein ring, though this is likely an exaggeration). The terminology was heavily popularized by Hyman Bass in his seminal paper "On the Ubiquity of Gorenstein Rings".

64.3. **Testing the Gorenstein Property.** Grothendieck's definition is theoretically powerful but exceptionally heavy for practical calculations. Fortunately, testing whether a higher-dimensional ring is Gorenstein can be reduced to the zero-dimensional case.

Theorem 64.6. *If R has dimension $d > 0$, then R is Gorenstein if and only if there exists a non-zero divisor $x \in \mathfrak{m}$ such that the quotient ring $R/(x)$ is Gorenstein.*

This exact reductive property also holds for Cohen-Macaulay rings. Any non-zero divisor in $\mathfrak{m}$ suffices; the choice of x does not matter. The proof relies on analyzing the short exact sequence:

$$0 \to R \xrightarrow{\times x} R \to \frac{R}{(x)} \to 0$$

Applying the long exact sequence of Ext groups to this sequence establishes the equivalence between the Ext condition for R and the Gorenstein condition for $R/(x)$.

64.4. **Subtle Examples in Higher Dimensions.** Being Gorenstein is a remarkably subtle arithmetic property. A seemingly harmless change to the structure of a ring can drastically alter whether or not it is Gorenstein.

Consider the power series ring in two variables, $k[[x, y]]$, acted upon by a cyclic group C_3 of order 3. We assume the characteristic of k is not 3. Let ω be a primitive cube root of unity ($\omega^3 = 1$). We analyze the fixed subring R^{C_3} under two very similar group actions.

Example 64.7 (Action 1: Not Gorenstein)**.** Let the group act by mapping $x \mapsto \omega x$ and $y \mapsto \omega y$. The invariant subring R consists of all formal power series where the total degree of every monomial is divisible by 3.

To test if R is Gorenstein, we quotient by a regular sequence of length 2 to reduce it to dimension 0. We pick two non-zero divisors, say x^3 and y^3, and quotient R by the ideal (x^3, y^3). The remaining basis elements in the quotient are exactly those invariant monomials whose individual x and y degrees are strictly less than 3. These are $1, xy^2$, and x^2y.

$$\frac{R}{(x^3, y^3)} = \operatorname{span}_k\{1, xy^2, x^2y\}$$

In this quotient ring, any product of the non-trivial elements is zero. For example, $(xy^2)(x^2y) = x^3y^3 = 0$. Thus, both xy^2 and x^2y are annihilated by the maximal ideal.

$$\dim_k \operatorname{Hom}\left(k, \frac{R}{(x^3, y^3)}\right) = 2$$

Because the dimension is 2, the ring is *not* Gorenstein.

Example 64.8 (Action 2: Gorenstein)**.** Now consider a slightly different action: $x \mapsto \omega x$ and $y \mapsto \omega^2 y$. The invariant monomials are those where the degree of x plus twice the degree of y is divisible by 3. The fundamental invariants are x^3, y^3, and xy.

Again, we quotient the invariant subring R by the non-zero divisors x^3 and y^3. The remaining basis elements in the quotient are:

$$\frac{R}{(x^3, y^3)} = \operatorname{span}_k\{1, xy, (xy)^2\}$$

Here, multiplying xy by the maximal ideal does not necessarily give zero, as $(xy)(xy) = (xy)^2 \neq 0$. The only element completely annihilated by the maximal ideal is the bottom block $(xy)^2$.

$$\dim_k \operatorname{Hom}\left(k, \frac{R}{(x^3, y^3)}\right) = 1$$

Thus, this subtly altered invariant ring *is* Gorenstein.

64.5. **Curve Singularities in Four Dimensions.** The visual subtlety is even more pronounced for curve singularities. Consider three curves defined by mappings from the affine line $\mathbb{A}^1$ into $\mathbb{A}^4$. We investigate the local ring at the origin (their completion).

(1) **Curve 1:** $t \mapsto (t^4, t^5, t^6, t^7)$. The local ring is spanned by $1, t^4, t^5, t^6, t^7, t^8, \ldots$. (Note that $t^8 = (t^4)^2$, so it contains all powers ≥ 4). We quotient by the non-zero divisor t^4. The remaining basis elements are $1, t^5, t^6, t^7$. Because the product of any two non-trivial elements has degree ≥ 10, which is 0 modulo (t^4) (since $t^{10} = t^4 \cdot t^6$), they are all annihilated by the maximal ideal.

$$\dim_k \operatorname{Hom}\left(k, \frac{R}{(t^4)}\right) = 3$$

This ring is *not* Gorenstein.

(2) **Curve 2:** $t \mapsto (t^5, t^6, t^7, t^8)$. The local ring contains 1 and all powers of t from t^5 upwards. We quotient by the non-zero divisor t^5. The quotient basis is $1, t^6, t^7, t^8$, and t^{14}. (We exclude $t^{10}, t^{11}, t^{12}, t^{13}$ because they are multiples of t^5). Notice that $t^6 \cdot t^8 = t^{14}$ and $t^7 \cdot t^7 = t^{14}$. The only element entirely annihilated by the maximal ideal is t^{14}.

$$\dim_k \operatorname{Hom}\left(k, \frac{R}{(t^5)}\right) = 1$$

This ring *is* Gorenstein. The structure explicitly displays a bilinear duality mapping the 3-dimensional space spanned by $\{t^6, t^7, t^8\}$ against itself into the 1-dimensional space spanned by t^{14}.

(3) **Curve 3:** $t \mapsto (t^6, t^7, t^8, t^9)$. The local ring contains 1 and all powers of t from t^6 upwards. Quotienting by t^6 leaves the basis $1, t^7, t^8, t^9, t^{16}, t^{17}$. Here, the bottom layer consists of the two elements t^{16} and t^{17}, which are both annihilated by the maximal ideal.

$$\dim_k \operatorname{Hom}\left(k, \frac{R}{(t^6)}\right) = 2$$

This ring is *not* Gorenstein.

Geometrically, these three curves look nearly identical in 4-dimensional space, yet their arithmetic duality properties fluctuate completely.

64.6. **Regular Rings are Gorenstein.** We conclude by formally proving that all regular local rings are Gorenstein, utilizing the Koszul complex.

Theorem 64.9. *If R is a regular local ring, then R is Gorenstein.*

Proof. Let R be a regular local ring of dimension d with maximal ideal $\mathfrak{m}$ and residue field k. By definition, the maximal ideal is generated by a regular sequence $x_1, \ldots, x_d$. We can construct the Koszul complex for this sequence, which resolves the residue field k:

$$0 \to R \to R^d \to \cdots \to R^d \to R \to \frac{R}{(x_1, \ldots, x_d)} \to 0$$

Since $(x_1, \ldots, x_d) = \mathfrak{m}$, the final quotient is exactly k.

To compute the Ext groups $\operatorname{Ext}^i_R(k, R)$, we apply the functor $\operatorname{Hom}_R(-, R)$ to this free resolution and calculate the cohomology of the resulting dual complex. Because the Koszul complex is self-dual (up to shifts and isomorphism), the dual complex is isomorphic to the original complex read in reverse. The cohomology of this reversed sequence will therefore perfectly match the homology of the original sequence, just shifted to the end.

Since the original Koszul complex is exact everywhere except at the final quotient k, the dual complex will be exact everywhere except at the d-th position. This forces the vanishing of all lower Ext groups:

$$\operatorname{Ext}^i_R(k, R) = 0 \quad \text{for } i \neq d$$

At the d-th position, the cohomology is exactly the quotient of R by the ideal generated by the dual differentials, which is precisely $(x_1, \ldots, x_d) = \mathfrak{m}$. Therefore:

$$\operatorname{Ext}^d_R(k, R) \cong \frac{R}{\mathfrak{m}} = k$$

Thus, $\dim_k \operatorname{Ext}^d_R(k, R) = 1$. By Grothendieck's definition, this perfectly satisfies the conditions for R to be a Gorenstein ring. □

65. Fitting Ideals

This lecture introduces *Fitting ideals*, an important sequence of invariants for finitely generated modules over arbitrary commutative rings. The concept is named after Hans Fitting, who developed them in the 1930s.

65.1. **Definition and Motivation.** Suppose R is a commutative ring (not necessarily Noetherian) and M is a finitely generated R-module. We can define a sequence of ideals:

$$\operatorname{Fitt}_0(M) \subseteq \operatorname{Fitt}_1(M) \subseteq \operatorname{Fitt}_2(M) \subseteq \cdots \subseteq R$$

The i-th Fitting ideal, $\operatorname{Fitt}_i(M)$, acts as a sort of obstruction to generating the module M by i elements. Strictly speaking, it is more sensible to view the quotient ring $R/\operatorname{Fitt}_i(M)$ as the actual obstruction. This is a cyclic module over R, and if M can be generated by i elements, this quotient vanishes, which is equivalent to saying:

$$\operatorname{Fitt}_i(M) = R$$

To formally construct these ideals, we must express M via a presentation. Because M is finitely generated, we can find a surjective module homomorphism from a finite free module onto M. Suppose M is generated by n elements, $g_1, \ldots, g_n$. This gives an exact sequence:

$$R^J \xrightarrow{A} R^n \twoheadrightarrow M \to 0$$

where R^n is the free module with basis $e_1, \ldots, e_n$ mapping to $g_1, \ldots, g_n$. The kernel of this map is the module of relations. Because we do not assume R is Noetherian, this kernel might not be finitely generated, so the index set J could be infinite.

The map A can be represented by a matrix with n columns (corresponding to the generators) and possibly infinitely many rows (corresponding to the relations). Let the rows be indexed by J, where each row represents a relation of the form:

$$r_{j,1}g_1 + r_{j,2}g_2 + \cdots + r_{j,n}g_n = 0$$

Definition 65.1. Given a presentation matrix A for a module M generated by n elements, the *i-th Fitting ideal* of M, denoted $\operatorname{Fitt}_i(M)$, is the ideal generated by the determinants of all $(n-i) \times (n-i)$ minors of A.

In particular, the *zeroth Fitting ideal*, $\operatorname{Fitt}_0(M)$, is generated by the determinants of all $n \times n$ submatrices of A. If $n - i \leq 0$, we define $\operatorname{Fitt}_i(M) = R$.

65.2. **Independence of Presentation.** An obvious problem with this definition is that it appears to depend heavily on the chosen presentation. A module M has many different presentations; changing the generators or the relations completely alters the matrix A. However, the Fitting ideals are an intrinsic invariant of the module M.

Theorem 65.2. *The i-th Fitting ideal $Fitt_i(M)$ depends only on the module M, and not on the choice of presentation.*

Proof Sketch. Any two presentations of a module M can be related by a sequence of elementary operations and their inverses:

(1) **Adding a new generator:** We can add a new generator g_{n+1} to our generating set, provided we also add a trivial relation stating that g_{n+1} is a linear combination of the old generators:
$$g_{n+1} - \sum_{k=1}^{n} c_k g_k = 0$$
(2) **Adding a redundant relation:** We can add a new relation that is a linear combination of the existing known relations. Since R is not necessarily Noetherian, this might involve adding an infinite number of such redundant relations.

We must verify that these operations do not alter the ideal generated by the minors.

Consider the first operation. The new presentation matrix $\tilde{A}$ has one extra column and one extra row. By performing elementary row and column operations (which do not change the ideal generated by the minors), we can use the 1 in the new relation to clear out the rest of its row and column. The new matrix essentially looks like a block diagonal matrix with the old matrix A in one block and a 1×1 identity block. Thus, taking an $(n+1-i) \times (n+1-i)$ minor of $\tilde{A}$ that includes the new row and column simply yields the determinant of an $(n-i) \times (n-i)$ minor of A (up to sign). If a minor does not include the new row, it contains a column of zeros, and its determinant is zero. Thus, the generated ideal remains invariant.

For the second operation, adding a row that is a linear combination of other rows does not change the ideal of minors. Any minor involving the new row can be expanded using the multilinearity of the determinant into a linear combination of determinants of minors from the original matrix. Hence, the sequence of Fitting ideals is a well-defined invariant of M. □

65.3. **Examples over the Integers.** To illustrate Fitting ideals, let us compute them for finitely generated abelian groups, which are simply finitely generated modules over $R = \mathbb{Z}$.

Example 65.3. Let M be a finite abelian group. By the structure theorem, M decomposes into a direct sum of cyclic groups:
$$M \cong \frac{\mathbb{Z}}{n_1\mathbb{Z}} \oplus \frac{\mathbb{Z}}{n_2\mathbb{Z}} \oplus \cdots \oplus \frac{\mathbb{Z}}{n_k\mathbb{Z}}$$

The obvious presentation for M has k generators and k relations. The relation matrix A is a $k \times k$ diagonal matrix with entries $n_1, n_2, \ldots, n_k$ along the diagonal.

The zeroth Fitting ideal $\mathrm{Fitt}_0(M)$ is generated by the determinants of all $k \times k$ minors. There is only one such minor, the full matrix A. Thus:

$$\mathrm{Fitt}_0(M) = (n_1 n_2 \cdots n_k)$$

Notice that the product $n_1 n_2 \cdots n_k$ is precisely the order of the group $|M|$. Therefore, the zeroth Fitting ideal is generated by the order of the group. The obstruction to generating M with 0 elements vanishes if and only if $\mathrm{Fitt}_0(M) = (1)$, which occurs if and only if M is the trivial group.

Example 65.4. Now suppose M is a general finitely generated abelian group, including a possible free part:

$$M \cong \mathbb{Z}^j \oplus \frac{\mathbb{Z}}{n_1\mathbb{Z}} \oplus \cdots \oplus \frac{\mathbb{Z}}{n_k\mathbb{Z}}$$

We can assume without loss of generality that the invariants divide each other: $n_1 \mid n_2 \mid \cdots \mid n_k$. The number of generators is $n = j + k$. The relation matrix is a diagonal matrix with j zeros and the k integers $n_1, \ldots, n_k$ on the diagonal.

Let us evaluate the sequence of Fitting ideals:

- To find $\mathrm{Fitt}_i(M)$ for $i < j$, we must take determinants of $(n-i) \times (n-i)$ submatrices. Because there are j rows of zeros, any submatrix of this size must include at least one row of zeros. Thus, all such determinants vanish:

$$\mathrm{Fitt}_i(M) = (0) \quad \text{for } i < j$$

- For $i = j$, the largest non-zero minor is obtained by deleting the j rows and columns of zeros. The remaining $k \times k$ matrix gives:

$$\mathrm{Fitt}_j(M) = (n_1 n_2 \cdots n_k)$$

- For $i = j+1$, we delete one more row and column. To generate the largest possible ideal (smallest determinant), we delete the largest entry n_k:

$$\mathrm{Fitt}_{j+1}(M) = (n_1 n_2 \cdots n_{k-1})$$

- This pattern continues until we have deleted all relation rows:

$$\mathrm{Fitt}_{j+k}(M) = (1) = \mathbb{Z}$$

This shows that M can be generated by $j + k$ elements, as the obstruction $\mathrm{Fitt}_{j+k}(M)$ is the whole ring. Remarkably, for finitely generated abelian groups, the sequence of Fitting ideals completely and uniquely determines the isomorphism class of the module.

65.4. **Summary of Properties.** We can summarize the key properties of Fitting ideals as follows:

(1) The Fitting ideals form an ascending chain:

$$\mathrm{Fitt}_0(M) \subseteq \mathrm{Fitt}_1(M) \subseteq \mathrm{Fitt}_2(M) \subseteq \ldots$$

(2) If a module M can be generated by i elements, then the i-th Fitting ideal is the entire ring:

$$\mathrm{Fitt}_i(M) = R$$

(3) The converse to the above statement is generally false for arbitrary rings. However, if M is a finitely generated module over a *local ring* (or more generally, over a local complete intersection ring), the converse does hold.
(4) The Fitting ideal is deeply related to the annihilator of the module. Specifically, one can show that the annihilator of M raised to the power of the number of generators n is contained in the zeroth Fitting ideal, which is in turn contained in the annihilator:

$$\mathrm{Ann}(M)^n \subseteq \mathrm{Fitt}_0(M) \subseteq \mathrm{Ann}(M)$$

We will utilize Fitting ideals in subsequent lectures when discussing local complete intersection rings, where measuring the exact number of relations relative to the number of generators becomes critical.

66. Local Complete Intersection Rings

In this lecture, we conclude our classification of local rings by introducing a class of rings with the rather cumbersome name of *local complete intersection rings*, often abbreviated as *LCI rings*.

To contextualize, we have discussed several classes of local rings that fit into a strict hierarchy. We have regular local rings (corresponding to non-singular points of varieties), which are a special case of local complete intersection rings. These, in turn, are a special case of Gorenstein rings, which are a special case of Cohen-Macaulay rings, which are finally a special case of general Noetherian local rings. Having covered the other classes in previous lectures, we now focus on the LCI case.

66.1. **Definition and Geometric Meaning.**

Definition 66.1. Let S be a regular local ring. A *local complete intersection ring* R is defined as the quotient of S by an ideal generated by a regular sequence $x_1, \ldots, x_i$:

$$R = \frac{S}{(x_1, \ldots, x_i)}$$

While this definition may initially seem artificial or overly technical, it captures a fundamental and natural geometric concept. Suppose you have a variety of codimension k in an n-dimensional affine space. Globally, it requires at least k equations to define it. If we can define the local ring at a point using exactly k equations (the minimal possible number), we obtain a local complete intersection ring. Thus, the LCI condition means the space can be defined locally using the minimal possible number of equations.

66.1.1. *Basic Examples.*

- **Regular Rings:** All regular rings are trivially local complete intersections, as one can simply quotient a regular ring by the empty set (which trivially forms a regular sequence).
- **Hypersurface Singularities:** Any hypersurface singularity is an LCI ring. If we localize the polynomial ring and quotient by a single non-zero element f, we obtain:

$$R = \left(\frac{k[y_1, \ldots, y_n]}{(f(y_1, \ldots, y_n))} \right)_{\mathfrak{m}}$$

  Because the ambient ring is an integral domain, f is automatically a non-zero divisor, forming a regular sequence of length 1. In particular, all singularities of plane algebraic curves are LCI singularities.

For an example of something that is *not* a local complete intersection, one might consider the union of a plane and a point in 3D space. At the isolated point, the variety has codimension 1 overall, but its 0-dimensional component cannot be defined by a single equation due to the codimension 2 line involved. However, this is a rather clumsy example because it fails to even be Cohen-Macaulay. We desire a more refined counterexample.

66.2. **Gorenstein vs. Local Complete Intersection.** Every local complete intersection ring is inherently Gorenstein. The proof involves taking a regular sequence and utilizing long exact sequences of Ext groups to explicitly calculate the necessary homological vanishing conditions.

A highly non-trivial problem is to find an example of a ring that is Gorenstein but *not* a local complete intersection. Most obvious examples of Gorenstein rings automatically turn out to be LCI rings. Historically, the distinction between these two classes gained massive prominence when Andrew Wiles proved Fermat's Last Theorem in the 1990s; a critical step in his proof involved demonstrating that a specific Gorenstein local ring was, in fact, a local complete intersection.

We will construct a zero-dimensional ring that is Gorenstein but not LCI.

Example 66.2. Consider the ring of formal power series (or polynomials) in three variables, quotiented by the following specific ideal:

$$R = \frac{k[[x, y, z]]}{(x^2, xy, yz, z^2, y^2 - xz)}$$

First, we establish that R is a Gorenstein ring. The dimension of R is 0, and it has length 5. We can view this ring as being constructed from a vector space. Let V be a finite-dimensional vector space over k, equipped with a symmetric bilinear form $V \otimes V \to k$. We can form a ring:

$$R = k \oplus V \oplus k$$

where the first k spans the identity, the vector space V corresponds to the maximal ideal modulo its square $\mathfrak{m}/\mathfrak{m}^2$, and the final k corresponds to $\mathfrak{m}^2$. The ring multiplication between elements of V is governed precisely by the bilinear form mapping into the bottom k.

If the bilinear form on V is non-degenerate, the resulting zero-dimensional local ring is strictly Gorenstein. In our specific example, V is 3-dimensional (spanned by x, y, z), and the relations ensure the form is non-degenerate, confirming R is Gorenstein. (Note that if V were 2-dimensional, the resulting ring would actually be an LCI; 3 dimensions are required to break the LCI property).

66.2.1. *Why is it not an LCI?.* Informally, we can write R as a quotient of $k[[x, y, z]]$ by an ideal $I \subset \mathfrak{m}^2$. The vector space $\mathfrak{m}^2/\mathfrak{m}^3$ has dimension 6 (spanned by $x^2, xy, xz, y^2, yz, z^2$). The ring R requires us to kill off a 5-dimensional subspace of this 6-dimensional space, leaving a 1-dimensional bottom layer.

To kill a 5-dimensional subspace, we inherently need at least 5 defining relations. However, the ambient regular ring $k[[x, y, z]]$ has dimension 3, and our target ring R has dimension 0. If R were a local complete intersection, we should be able to define it using exactly $3 - 0 = 3$ relations. Since $5 > 3$, this strongly suggests the ring is not LCI. While intuitively compelling, making this proof rigorous requires evaluating the Fitting ideals.

66.3. **Fitting Ideal Criterion.** To rigorously prove that a ring is not an LCI, we make use of the following theorem concerning Fitting ideals (as introduced in the previous lecture).

Theorem 66.3. *A 0-dimensional Noetherian local ring $(R, \mathfrak{m})$ is a local complete intersection if and only if the zeroth Fitting ideal of the maximal ideal is not zero:*

$$\mathrm{Fitt}_0(\mathfrak{m}) \neq 0$$

For rings of positive dimension, one can recursively reduce to the zero-dimensional case: a local ring R of dimension > 0 is an LCI if and only if $R/(x_1)$ is an LCI, where x_1 is a regular element (a non-zero divisor).

Let us apply this criterion to Example 66.2. The maximal ideal $\mathfrak{m}$ is generated by 3 elements: x, y, z. We must write down the matrix of relations defining $\mathfrak{m}$ as an R-module. The rows of this matrix correspond to linear combinations of the generators that evaluate to zero.

$$\begin{aligned}
x^2 = 0 &\implies x \cdot x + 0 \cdot y + 0 \cdot z = 0 \implies (x, 0, 0) \\
yx = 0 &\implies y \cdot x + 0 \cdot y + 0 \cdot z = 0 \implies (y, 0, 0) \\
xy = 0 &\implies 0 \cdot x + x \cdot y + 0 \cdot z = 0 \implies (0, x, 0) \\
z^2 = 0 &\implies 0 \cdot x + 0 \cdot y + z \cdot z = 0 \implies (0, 0, z)
\end{aligned}$$

Similarly, $yz = 0$ gives $(0, z, 0)$ and $(0, 0, y)$. The relation $y^2 = xz$ implies $z \cdot x - y \cdot y = 0$, yielding $(-z, y, 0)$, and so forth.

We construct a large matrix A where every single entry is an element of the maximal ideal $\mathfrak{m}$. The zeroth Fitting ideal $\mathrm{Fitt}_0(\mathfrak{m})$ is generated by the determinants of all 3×3 minors of A. Because every entry of A resides in $\mathfrak{m}$, the determinant of any 3×3 minor will be a sum of products of 3 elements from $\mathfrak{m}$. Thus, every such determinant belongs to $\mathfrak{m}^3$.

However, in our ring R, the ideal $\mathfrak{m}^3$ is strictly zero. Therefore:

$$\mathrm{Fitt}_0(\mathfrak{m}) = 0$$

By our theorem, this rigorously proves that R is not a local complete intersection ring.

While we are computing Fitting ideals, we can evaluate the higher ones for practice:

- $\mathrm{Fitt}_1(\mathfrak{m})$ is generated by the 2×2 minors. These determinants lie in $\mathfrak{m}^2$. Since $\mathfrak{m}^2$ is a 1-dimensional vector space spanned by xz, and we can easily find a non-zero 2×2 minor, we have $\mathrm{Fitt}_1(\mathfrak{m}) = \mathfrak{m}^2 = (xz)$.
- $\mathrm{Fitt}_2(\mathfrak{m})$ is generated by the 1×1 minors, which are simply the entries of the matrix. These generate the entire maximal ideal, so $\mathrm{Fitt}_2(\mathfrak{m}) = \mathfrak{m}$.
- $\mathrm{Fitt}_3(\mathfrak{m})$ is generated by the 0×0 minors. By convention, the determinant of a 0×0 matrix is 1, so $\mathrm{Fitt}_3(\mathfrak{m}) = R$.

66.4. **A Geometric Example Without Nilpotents.** The previous example was a zero-dimensional ring heavily populated with nilpotent elements. We can construct a one-dimensional, reduced example (no nilpotent elements) that corresponds to a classical geometric object.

Consider the coordinate ring of the space curve parametrized by $t \mapsto (t^5, t^6, t^7, t^8)$. The local ring at the origin is:

$$S = k[[t^5, t^6, t^7, t^8]]$$

We previously demonstrated that this ring is Gorenstein. We now show it is not a local complete intersection. We reduce to the zero-dimensional case by quotienting by the non-zero divisor t^5:

$$\frac{S}{(t^5)}$$

This quotient has length 5, with a basis consisting of $1, t^6, t^7, t^8$, and t^{14}. Let us map this to our previous example by assigning $x = t^6$, $y = t^7$, and $z = t^8$.

$$xz = t^6 \cdot t^8 = t^{14}$$
$$y^2 = t^7 \cdot t^7 = t^{14}$$

Thus, $xz = y^2$. Furthermore, higher products vanish modulo t^5. For instance, $x^2 = t^{12} = t^5 \cdot t^7 \equiv 0 \pmod{t^5}$. Evaluating the remaining products ($xy = t^{13} \equiv 0$, $z^2 = t^{16} \equiv 0$, $yz = t^{15} \equiv 0$), we recover the exact same defining relations as Example 66.2.

Because the quotient ring $S/(t^5)$ is identically isomorphic to our zero-dimensional non-LCI ring, the original one-dimensional geometric ring S is also not a local complete intersection ring.

66.5. **Summary Flowchart for Local Rings.** To conclude, we can summarize the diagnostic procedure for classifying a given Noetherian local ring R into a flowchart.

(1) **Reduce Dimension:** If $\dim(R) > 0$, seek a non-zero divisor $x \in \mathfrak{m}$.
   - If no such element exists, R is **not Cohen-Macaulay** (and thus not Gorenstein, LCI, or Regular).
   - If x exists, replace R with $R/(x)$ and repeat until $\dim(R) = 0$.

(2) **Test Gorenstein:** For the 0-dimensional reduction, compute the dimension of the socle: $\dim_k \operatorname{Hom}_R(k, R)$.
   - If > 1, the ring is **not Gorenstein**.
   - If $= 1$, the ring is **Gorenstein**. Proceed to the next test.

(3) **Test Local Complete Intersection:** Compute the zeroth Fitting ideal of the maximal ideal, $\operatorname{Fitt}_0(\mathfrak{m})$.
   - If $= 0$, the ring is **not LCI**.
   - If $\neq 0$, the ring is **LCI**. Proceed to the next test.

(4) **Test Regularity:** Compute the dimension of the original ring R as a vector space over k, specifically evaluating if $\dim_k(\mathfrak{m}/\mathfrak{m}^2) = \dim(R)$. If they match, the ring is **Regular**. If not, one must verify that the sequence of non-zero divisors chosen in Step 1 did not inadvertently reside in $\mathfrak{m}^2$. If they were chosen properly (in $\mathfrak{m} \setminus \mathfrak{m}^2$) and the dimensions fail to match, the ring is not regular.

This structured approach reliably unwinds the homological properties of local rings. The upcoming lectures will move toward applications of commutative algebra, specifically concerning the Bernstein-Sato polynomial.

Part 9. D-Modules and Dedekind Domains

67. Introduction to the Bernstein Sato Polynomial

This lecture introduces the Bernstein-Sato polynomial. Unlike previous topics, this lecture covers slightly non-commutative algebra. A significant portion of commutative algebra theory can be extended to rings that are "almost" commutative.

67.1. **Almost Commutative Algebras.** Several important algebras fail to be strictly commutative but still retain a manageable structure where the commutator is "simpler" than the elements themselves.

- **Exterior Algebras:** For a vector space, the exterior algebra satisfies:
$$x \wedge y = -y \wedge x$$
- **Clifford Algebras:** Arising frequently in the study of Lie groups, if we have an inner product $\langle \cdot, \cdot \rangle$ on a vector space, the Clifford algebra relations are:
$$ab + ba = \langle a, b \rangle$$
Here, ab is closely related to $-ba$, differing only by a scalar, which is inherently simpler.
- **Rings of Differential Operators:** Consider the Weyl algebra generated over $\mathbb{C}$ by the position operator x and the differentiation operator $\frac{d}{dx}$. Applying Leibniz's rule to a test function yields the commutation relation:
$$\frac{d}{dx}(x) = x\frac{d}{dx} + 1$$
These operators do not commute, but their commutator $\left[\frac{d}{dx}, x\right] = 1$ is a constant, which is definitively simpler than either x or $\frac{d}{dx}$.

When an algebra is close to being commutative in this sense, techniques from commutative algebra can often be successfully applied. We will use the ring of differential operators to demonstrate the existence of the *Bernstein-Sato polynomial*. (Note: This is named after Joe Bernstein and Mikio Sato, and is entirely unrelated to the Bernstein polynomials used in numerical analysis).

67.2. **Motivation: The Gamma Function.** To understand the mechanism behind the Bernstein-Sato polynomial, we first review the analytic continuation of the Gamma function.

The Gamma function $\Gamma(s)$ is defined by the integral:

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1}\, dt$$

This integral converges absolutely for the real part of $s > 0$. If $\mathrm{Re}(s) \leq 0$, the integral diverges near $t = 0$. However, we can analytically continue the function by using integration by parts.

Observe the action of the derivative on the polynomial term:

$$\frac{d}{dt} t^{s+1} = (s+1)t^s$$

By differentiating e^{-t} and integrating t^{s-1}, integration by parts yields the functional equation:

$$\Gamma(s) = \frac{1}{s}\Gamma(s+1)$$

Because $\Gamma(s+1)$ is well-defined for $\mathrm{Re}(s) > -1$, dividing by s extends the domain of $\Gamma(s)$ to $\mathrm{Re}(s) > -1$, picking up a simple pole at $s = 0$. Repeating this process extends $\Gamma(s)$ to a meromorphic function on the entire complex plane, with poles at the non-positive integers. The critical mechanism here was trading a power of t for a derivative.

67.3. **Definition of the Bernstein-Sato Polynomial.** The Bernstein-Sato polynomial generalizes this exact mechanism to multivariate polynomials.

Definition 67.1. Let f be a non-zero polynomial in several variables $x_1, \ldots, x_n$ over $\mathbb{C}$. The *Bernstein-Sato polynomial* $b(s)$ is the monic polynomial of minimal degree such that there exists a differential operator $P(s)$ satisfying the functional equation:

$$P(s) f(x_1, \ldots, x_n)^{s+1} = b(s) f(x_1, \ldots, x_n)^s$$

where $P(s)$ is a polynomial in the variables x_i, the partial derivatives $\frac{\partial}{\partial x_i}$, and the complex variable s.

If such an operator $P(s)$ and polynomial $b(s)$ exist, they allow us to analytically continue integrals of the form:

$$\int_{\mathbb{R}^n} \phi(x) f(x)^s\, dx$$

where $f(x) \geq 0$ and $\phi(x)$ is a smooth test function of compact support. Just as with the Gamma function, we can use the differential operator $P(s)$ to integrate by parts, extending the integral to all complex values of s, except for poles corresponding to the roots of $b(s)$.

67.4. **Examples and the Main Theorem.**

Example 67.2. Let $f(t) = t$. The differential operator is simply $P(s) = \frac{d}{dt}$. The functional equation is:

$$\frac{d}{dt} t^{s+1} = (s+1)t^s$$

Here, $b(s) = s + 1$.

Example 67.3. Let $f(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2$. We can apply the standard Laplace operator:

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$$

Applying the Laplacian to f^{s+1} yields:

$$\Delta f^{s+1} = 4(s+1)\left(s + \frac{n}{2}\right) f^s$$

To normalize the Bernstein-Sato polynomial so its leading coefficient is 1, we set $P(s) = \frac{1}{4}\Delta$. The corresponding Bernstein-Sato polynomial is:

$$b(s) = (s+1)\left(s + \frac{n}{2}\right)$$

The set of all polynomials $b(s)$ satisfying the functional equation forms an ideal in $\mathbb{C}[s]$. The Bernstein-Sato polynomial is uniquely defined as the monic generator of this ideal. Proving the existence of such a polynomial is non-trivial, but calculating it is even harder.

Example 67.4. Let $f(x, y) = x^2 + y^3$. Despite being only slightly more complicated, finding the differential operator involves mountains of linear algebra. The resulting polynomial is:

$$b(s) = (s+1)\left(s + \frac{5}{6}\right)\left(s + \frac{7}{6}\right)$$

The foundational theorem of this subject guarantees that this structure is not a coincidence.

Theorem 67.5 (Bernstein-Sato)**.** *Every non-zero complex polynomial $f \in \mathbb{C}[x_1, \ldots, x_n]$ possesses a non-zero Bernstein-Sato polynomial.*

The proof of this theorem heavily uses commutative algebra and the theory of D-modules, specifically employing the Hilbert polynomial, which we will detail in subsequent lectures.

67.5. **Application: The Malgrange-Ehrenpreis Theorem.** To demonstrate the immense power of the Bernstein-Sato polynomial, we use it to trivially prove the Malgrange-Ehrenpreis theorem. For a long time, it was a major open question whether every linear differential equation had a fundamental solution.

Theorem 67.6 (Malgrange-Ehrenpreis)**.** *Every non-zero linear differential operator D with constant coefficients has a fundamental solution. That is, there exists a distribution F such that:*

$$DF = \delta$$

where δ is the Dirac delta distribution.

If a fundamental solution exists, one can solve $Df = g$ for any reasonable function g by taking the convolution $f = F * g$.

Proof. We take the Fourier transform of the equation $DF = \delta$. The Fourier transform converts differentiation into multiplication by coordinate variables, meaning the differential operator D becomes multiplication by a polynomial $q(x)$. The unknown distribution F becomes a distribution $\hat{F}$, and the delta function δ transforms into the constant 1.

$$q(x)\hat{F} = 1$$

The naive solution is to simply set $\hat{F} = \frac{1}{q(x)}$.

If $q(x)$ has no zeros, this is a perfectly well-behaved function. However, if $q(x)$ has zeros, $1/q(x)$ is generally not locally integrable. If the zeros form a simple non-singular variety or have normal crossings, defining the distribution via principal values is straightforward. If the zero locus contains highly complicated singularities, constructing the distribution is incredibly difficult. One classical method uses Hironaka's theorem on the resolution of singularities, which is a massive, highly technical sledgehammer.

Instead, we use the Bernstein-Sato polynomial. Without loss of generality, we can assume $q(x) \geq 0$ everywhere, because:

$$\frac{1}{q} = \frac{\bar{q}}{q\bar{q}}$$

where $q\bar{q} \geq 0$.

We consider the function $q(x)^s$. For $\mathrm{Re}(s) \geq 0$, this is a perfectly well-defined, locally integrable function, and thus a valid holomorphic distribution. By the Bernstein-Sato theorem, there exists a differential operator $P(s)$ and a polynomial $b(s)$ such that:

$$P(s)q^{s+1} = b(s)q^s$$

Rearranging this gives:

$$q^s = \frac{P(s)q^{s+1}}{b(s)}$$

We use this identity to analytically continue q^s to a meromorphic function of s taking values in the space of distributions. This means that for any smooth test

function ϕ of compact support, the integral $\int \phi(x)q(x)^s\,dx$ extends to a meromorphic complex-valued function. The poles of this distribution are related to the zeros of $b(s+n)$.

We wish to evaluate this near $s=0$ to construct the inverse q^{-1}. We expand q^s as a Laurent series around $s=0$:

$$q^s = q_{-m}s^{-m} + q_{1-m}s^{1-m} + \cdots + q_0 s^0 + \dots$$

where the coefficients q_i are distributions. We multiply this entire series by the smooth function $q(x)$:

$$q \cdot q^s = q \cdot q_{-m}s^{-m} + \cdots + q \cdot q_{-1}s^{-1} + q \cdot q_0 s^0 + \dots$$

On the other hand, the distribution q^s evaluated analytically at $s=0$ is simply the constant function 1. Because 1 is holomorphic at $s=0$, all the singular terms in its Laurent series must vanish:

$$q \cdot q^s = 1$$

Matching coefficients of the powers of s on both sides, we deduce that the coefficients of the negative powers must be annihilated by q:

$$\begin{aligned} q \cdot q_{-m} &= 0 \\ &\dots \\ q \cdot q_{-1} &= 0 \end{aligned}$$

And crucially, matching the constant s^0 term yields:

$$q \cdot q_0 = 1$$

This establishes that the distribution q_0 acts exactly as the inverse of q. Taking the inverse Fourier transform of q_0 provides the fundamental solution F, effortlessly proving the theorem. □

67.6. **A Note on the Associativity of Distributions.** An observant reader might spot an apparent algebraic contradiction in the proof above. We found that $q \cdot q_0 = 1$ and $q \cdot q_{-1} = 0$. This implies that we can form infinitely many distinct inverses for q:

$$q \cdot (q_0 + cq_{-1}) = 1 + 0 = 1$$

In standard ring theory, an element has at most one inverse. If $qa = 1$ and $qb = 1$, then:

$$b = b(qa) = (bq)a = 1 \cdot a = a$$

How can q have multiple inverses?

The first issue is that distributions do not form a ring; the product of two arbitrary distributions is not always well-defined. However, in our equations, q is a smooth polynomial, and the product of a smooth function and a distribution is always well-defined.

The true resolution to the paradox is that *even when the product of distributions is defined, it is not strictly associative.*

Example 67.7. Let us work in one dimension. Consider the smooth function x, the principal value distribution $\frac{1}{x}$, and the Dirac delta distribution $\delta(x)$. We evaluate the triple product in two different bracketings:

$$\left(\frac{1}{x} \cdot x\right) \cdot \delta(x) = 1 \cdot \delta(x) = \delta(x)$$
$$\frac{1}{x} \cdot (x \cdot \delta(x)) = \frac{1}{x} \cdot 0 = 0$$

Because $\delta(x) \neq 0$, the associative law catastrophically fails. This failure of associativity explicitly breaks the algebraic proof of unique inverses, allowing distributions like $q(x)$ to legitimately possess an infinite family of valid inverses.

68. Bernstein's Inequality

This lecture will be a continuation of the previous lecture on the Bernstein-Sato polynomial. What we will be doing in this lecture and the next is showing how to prove the existence of this polynomial. In fact, in the rest of this lecture, we will hardly mention the Bernstein-Sato polynomial. Instead, we will be proving an inequality called *Bernstein's inequality*, which is an inequality involving modules over the Weyl algebra.

68.1. **The Weyl Algebra.**

Definition 68.1. The *Weyl algebra*, denoted A, is the ring of differential operators with polynomial coefficients over the complex numbers $\mathbb{C}$. It is generated by the variables $x_1, \ldots, x_n$ and the partial derivatives $\partial_1, \ldots, \partial_n$.

Here, ∂_i is written as an abbreviation for differentiation with respect to x_i. One has to be a little bit careful because this ring is non-commutative. However, it is pretty close to being commutative, as the following relations hold:

$$x_i x_j = x_j x_i$$
$$\partial_i \partial_j = \partial_j \partial_i$$
$$\partial_i x_j = x_j \partial_i \quad \text{if } i \neq j$$

The only relation that stops it from being a commutative algebra is the Leibniz rule:

$$\partial_i x_i = x_i \partial_i + 1$$

Because it is so close to being a commutative ring, we can apply many techniques of commutative algebra to it.

If M is a module over the Weyl algebra A, we can think of it as a system of differential equations. Since A is non-commutative, we must specify left or right ideals, but suppose we have a cyclic module $M = A/I$ for some left ideal I. Each element of I represents a differential equation. A homomorphism of A-modules from A/I to a ring of smooth functions yields solutions to this system. The image of 1 under this homomorphism will be a smooth function that is annihilated by all differential operators in I. Thus, the Weyl algebra turns the problem of solving differential equations into a problem of linear algebra over a non-commutative ring.

68.2. **The Center of the Weyl Algebra.**

Proposition 68.2. *The center of the Weyl algebra A over $\mathbb{C}$ is exactly the field of complex numbers $\mathbb{C}$.*

Proof. The algebra A certainly contains $\mathbb{C}$ as a subring. We must show it contains nothing else. If we have a differential operator $d \in A$, we can map it via the commutator with x_i:

$$d \mapsto x_i d - d x_i$$

This is a $\mathbb{C}$-linear map from A to A. Working in the standard basis consisting of monomials in x multiplied by monomials in ∂, it is easy to compute the kernel. The kernel of the maps $d \mapsto x_i d - d x_i$ for all i consists precisely of the polynomials in $x_1, \ldots, x_n$.

Similarly, there is a symmetric map using the partial derivatives:

$$d \mapsto \partial_i d - d \partial_i$$

The kernel of these maps for all i consists of polynomials in $\partial_1, \ldots, \partial_n$. Anything commuting with all x_i and all ∂_i must be a polynomial in x and simultaneously a polynomial in ∂. The intersection of these two sets is simply $\mathbb{C}$. □

Remark 68.3. We implicitly used the fact that $\mathbb{C}$ has characteristic zero. In characteristic $p > 0$, the center of the Weyl algebra is no longer trivial. For example, it contains ∂_i^p and x_i^p. The commutator $[\partial_i, x_i^p]$ turns out to be a multiple of p, which is zero in characteristic p. Furthermore, in characteristic p, one encounters additional complications with divided power differential operators (informally resembling $\partial^p/p!$), so we will strictly restrict our attention to the complex numbers.

68.3. **The Bernstein Filtration.** We wish to convert the non-commutative Weyl algebra into a commutative algebra.

Definition 68.4. The *Bernstein filtration* on the Weyl algebra A is an ascending chain of complex vector spaces:

$$A_0 \subseteq A_1 \subseteq A_2 \subseteq \ldots$$

where $A_0 = \mathbb{C}$, and A_i is spanned by all monomials in $x_1, \ldots, x_n, \partial_1, \ldots, \partial_n$ of total degree less than or equal to i.

This filtration satisfies the standard multiplicative property:

$$A_i A_j \subseteq A_{i+j}$$

Crucially, because the non-commutative cross terms drop in degree (e.g., $\partial_i x_i - x_i \partial_i = 1 \in A_0$), the commutator of any two elements drops by one degree. If $P \in A_i$ and $Q \in A_j$, then:

$$PQ - QP \in A_{i+j-1}$$

We can form the associated graded ring:

$$\operatorname{gr} A = A_0 \oplus \frac{A_1}{A_0} \oplus \frac{A_2}{A_1} \oplus \ldots$$

Because the commutators vanish in the successive quotients, this graded ring is strictly commutative. In fact, it is isomorphic to a polynomial ring in $2n$ variables (the images of x_i and ∂_i).

We can do the same for modules. Suppose M is an A-module generated by a finite-dimensional complex vector space M_0. We define a filtration on M by:

$$M_i = A_i M_0$$

This immediately satisfies:

$$A_i M_j \subseteq M_{i+j}$$

The associated graded module $\operatorname{gr} M = \bigoplus M_i/M_{i-1}$ is a finitely generated module over the commutative Noetherian polynomial ring $\operatorname{gr} A$.

Because $\operatorname{gr} M$ is finitely generated over a polynomial ring, the dimension of the complex vector space M_i is a polynomial in i for sufficiently large i, dictated by the theory of Hilbert polynomials.

Definition 68.5. The *dimension* of M, denoted $\dim(M)$, is defined as the degree of the polynomial giving $\dim_{\mathbb{C}}(M_i)$ for large i.

The *multiplicity* of M is defined as $\dim(M)!$ times the leading coefficient of this Hilbert polynomial.

Remark 68.6. One should not confuse the dimension of M as a vector space (which is infinite) with this ring-theoretic dimension (which is the degree of the polynomial growth). The dimension and multiplicity are independent of the initial generating subspace M_0. Changing M_0 only shifts the filtration by a finite bounded amount, which preserves the degree and leading coefficient of the Hilbert polynomial, altering only the lower-order terms.

68.4. **Bernstein's Inequality.** In the commutative case, the dimension of a module over a polynomial ring in $2n$ variables can be any integer from 0 to $2n$. However, over the non-commutative Weyl algebra, there is a dramatic restriction.

Theorem 68.7 (Bernstein's Inequality)**.** *If M is a non-zero finitely generated module over the Weyl algebra A, then its dimension is bounded below by n:*

$$\dim(M) \geq n$$

Proof. The key point is to show that the natural map from A_i to the endomorphism space $\operatorname{Hom}_{\mathbb{C}}(M_i, M_{2i})$ is injective.

Suppose there exists an operator $a \in A_i$ such that $aM_i = 0$. We will show by induction on i that $a = 0$.

The base case $i < 0$ is trivial since $A_i = 0$. Assume the statement is true for $i - 1$. We have $a \in A_i$ and $aM_i = 0$. Consider the commutator of a with the partial derivative ∂_j. From our filtration properties:

$$a\partial_j - \partial_j a \in A_{i-1}$$

We evaluate how this commutator acts on M_{i-1}. Since $\partial_j \in A_1$, the term $\partial_j M_{i-1}$ lands inside M_i. By our assumption, a annihilates everything in M_i, so:

$$a(\partial_j M_{i-1}) = 0$$

Therefore, the action of the commutator simplifies:

$$(a\partial_j - \partial_j a)M_{i-1} = -\partial_j(aM_{i-1})$$

But $M_{i-1} \subseteq M_i$, so $aM_{i-1} = 0$. Thus, the entire commutator annihilates M_{i-1}.

By our inductive hypothesis, any element in A_{i-1} that annihilates M_{i-1} must be identically zero. Therefore, $a\partial_j - \partial_j a = 0$ for all j. By a completely symmetric argument, $ax_j - x_j a = 0$ for all j.

Since a commutes with all x_j and ∂_j, a must lie in the center of the Weyl algebra. As we proved earlier, the center is exactly $\mathbb{C}$, meaning a is simply a complex scalar. Since $aM_i = 0$ and M is a non-zero module (so $M_i \neq 0$ for large i), the scalar a must be 0. This proves injectivity.

Now we compare dimensions. Because $A_i \hookrightarrow \mathrm{Hom}_{\mathbb{C}}(M_i, M_{2i})$ is injective, we have:

$$\dim_{\mathbb{C}}(A_i) \leq \dim_{\mathbb{C}}(\mathrm{Hom}_{\mathbb{C}}(M_i, M_{2i}))$$

The dimension of A_i corresponds to the number of monomials of degree $\leq i$ in $2n$ variables, which is a polynomial in i of degree $2n$.

The dimension of the Hom space is bounded by the product of the dimensions of M_i and M_{2i}:

$$\dim_{\mathbb{C}}(\mathrm{Hom}_{\mathbb{C}}(M_i, M_{2i})) = \dim_{\mathbb{C}}(M_i) \times \dim_{\mathbb{C}}(M_{2i})$$

By the definition of the dimension of M, both $\dim_{\mathbb{C}}(M_i)$ and $\dim_{\mathbb{C}}(M_{2i})$ are bounded by polynomials in i of degree $\dim(M)$. Their product is therefore bounded by a polynomial of degree $2\dim(M)$.

Comparing the degrees of the bounding polynomials yields:

$$2n \leq 2\dim(M)$$

Dividing by two gives Bernstein's inequality:

$$\dim(M) \geq n$$

□

Definition 68.8. If a finitely generated Weyl algebra module M satisfies $\dim(M) = n$ (the minimum possible dimension) or $M = 0$, then M is called a *holonomic module*.

This term is related to the notion of a holonomic system of differential equations. In the next lecture, we will demonstrate how to use Bernstein's inequality and the properties of holonomic modules to prove the existence of the Bernstein-Sato polynomial.

69. Holonomic Modules

In this section, we study *holonomic modules* over the Weyl algebra and use their structural properties to prove the existence of the Bernstein-Sato polynomial. In previous lectures, we defined the Bernstein-Sato polynomial and established Bernstein's inequality; now, we bring these concepts together.

69.1. **Recap: The Weyl Algebra and Bernstein's Inequality.** Recall that the Weyl algebra A is the ring of differential operators with polynomial coefficients over the complex numbers:

$$A = \mathbb{C}[x_1, \ldots, x_n, \partial_1, \ldots, \partial_n]$$

where ∂_i is shorthand for the partial derivative $\frac{\partial}{\partial x_i}$.

If M is a non-zero finitely generated module over the Weyl algebra A, Bernstein's inequality states that the dimension of M is bounded below by the number of variables n:

$$\dim(M) \geq n$$

Here, the dimension of M does not refer to its dimension as a vector space over $\mathbb{C}$ (which is typically infinite). Instead, it is defined as the degree of the Hilbert polynomial associated with a good filtration of M.

In commutative algebra, modules over a polynomial ring can have dimension 0, and these zero-dimensional modules are precisely the ones of finite length. In our non-commutative setting over the Weyl algebra, the lowest possible dimension is n. We will show that a similar structural property holds: modules achieving this minimal dimension possess finite length.

69.2. **Definition and Finite Length of Holonomic Modules.**

Definition 69.1. An A-module M is called *holonomic* if it is finitely generated and either $M = 0$ or:

$$\dim(M) = n$$

Theorem 69.2. *Every holonomic module M over the Weyl algebra A has finite length.*

Proof. The intuition here is that if a module has dimension k, any filtration can only contain a finite number of subquotients of dimension k, though it could potentially contain infinitely many subquotients of dimension strictly less than k. Because Bernstein's inequality guarantees that there are absolutely no non-zero modules of dimension strictly less than n, a holonomic module cannot have infinitely many subquotients.

To formalize this, recall that the Hilbert polynomial is additive on exact sequences. If we have a short exact sequence of holonomic modules:

$$0 \to M' \to M \to M'' \to 0$$

then their Hilbert polynomials satisfy $P_M(t) = P_{M'}(t) + P_{M''}(t)$.

Because all non-zero submodules and quotient modules must have dimension exactly n (they cannot drop below n due to Bernstein's inequality), the polynomials $P_{M'}(t)$ and $P_{M''}(t)$ all have degree exactly n. Consequently, the leading coefficients of these Hilbert polynomials are additive.

We define the *multiplicity* of a holonomic module M, denoted $e(M)$, as:

$$e(M) = n! \times (\text{leading coefficient of } P_M(t))$$

Because the leading coefficients are additive, the multiplicity is additive on exact sequences:

$$e(M) = e(M') + e(M'')$$

Furthermore, the multiplicity is always an integer $e(M) \geq 0$, and $e(M) = 0$ if and only if $M = 0$.

If M has a composition series of length ℓ, say $0 = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_\ell = M$, then each successive quotient M_i/M_{i-1} is non-zero and therefore has multiplicity at least 1. By additivity:

$$e(M) = \sum_{i=1}^{\ell} e\left(\frac{M_i}{M_{i-1}}\right) \geq \sum_{i=1}^{\ell} 1 = \ell$$

Thus, the length ℓ of any chain of submodules is strictly bounded above by the multiplicity $e(M)$, which is finite. Therefore, M has finite length. □

69.3. **Modules with Polynomial Growth.** In practice, we often encounter modules where we do not yet know if they are finitely generated. The following lemma provides a powerful criterion for proving finite generation and holonomicity simultaneously.

Lemma 69.3. *Let M be an arbitrary A-module (not necessarily known to be finitely generated). Suppose M admits an ascending filtration of finite-dimensional $\mathbb{C}$-vector spaces $M_0 \subseteq M_1 \subseteq M_2 \subseteq \dots$ such that $M = \bigcup M_i$ and $A_1 M_i \subseteq M_{i+1}$.*

If there exists a polynomial $P(i)$ of degree at most n such that:

$$\dim_{\mathbb{C}}(M_i) \leq P(i) \quad \text{for all } i$$

then M is finitely generated, and thus M is holonomic.

Proof. Any finitely generated submodule $N \subseteq M$ inherits a filtration $N_i = N \cap M_i$, and its dimension is bounded by the same polynomial $P(i)$. Thus, every finitely generated submodule of M is holonomic and has a multiplicity $e(N)$ bounded by:

$$e(N) \leq n! \times (\text{leading coefficient of } P)$$

Suppose, for the sake of contradiction, that M is not finitely generated. We could iteratively pick elements to form an infinite, strictly increasing chain of finitely generated submodules:

$$N^{(0)} \subsetneq N^{(1)} \subsetneq N^{(2)} \subsetneq \cdots \subseteq M$$

Because each $N^{(j)}$ is holonomic, the sequence of multiplicities $e(N^{(j)})$ must form a strictly increasing sequence of non-negative integers. However, this sequence is universally bounded above by the multiplicity extracted from $P(i)$. An infinite strictly increasing sequence of integers cannot be bounded above. This contradiction forces M to be finitely generated, which immediately implies it is holonomic. □

Corollary 69.4. *Let $p(x_1, \dots, x_n)$ be a non-zero polynomial, and consider the localization of the polynomial ring at p:*

$$M = \mathbb{C}[x_1, \dots, x_n, p^{-1}]$$

Then M is a holonomic A-module.

Proof. We can define a filtration on M by bounding the degree of the numerator relative to the power of the denominator. Let $m = \deg(p)$. We set:

$$M_k = \left\{ \frac{f}{p^k} \;\middle|\; \deg(f) \leq (m+1)k \right\}$$

It is easy to verify that $A_1 M_k \subseteq M_{k+1}$, giving a valid filtration. The vector space dimension of M_k is bounded by the number of polynomials of degree at most $(m+1)k$, which is:

$$\dim_{\mathbb{C}}(M_k) \leq \binom{(m+1)k+n}{n}$$

This binomial coefficient is a polynomial in k of degree exactly n. By the preceding lemma, M does not grow faster than a polynomial of degree n, meaning M is finitely generated and holonomic. □

69.4. **Existence of the Bernstein-Sato Polynomial.** We now possess the necessary tools to prove the existence of the Bernstein-Sato polynomial.

Recall that for a given non-zero polynomial $p \in \mathbb{C}[x_1, \ldots, x_n]$, we seek a differential operator $Q(s)$ and a polynomial $b(s)$ such that the functional equation holds:

$$b(s)p^s = Q(s) \cdot p^{s+1}$$

Theorem 69.5. *For any non-zero polynomial $p \in \mathbb{C}[x_1, \ldots, x_n]$, the Bernstein-Sato polynomial exists.*

Proof. We introduce the formal parameter s. To make use of our holonomic machinery, we extend our base field from $\mathbb{C}$ to the field of rational functions in s, denoted $K = \mathbb{C}(s)$.

We define the Weyl algebra over K:

$$A_K = K[x_1, \ldots, x_n, \partial_1, \ldots, \partial_n]$$

We also define the module $M_K = K[x_1, \ldots, x_n, p^{-1}] \cdot p^s$. Note that p^s is treated as a formal basis element upon which the differentiation operators act via the standard chain rule: $\partial_i(p^s) = sp^{s-1}\partial_i(p)$.

By the exact same proof as the corollary above (merely extending the scalars to K), the module M_K is a holonomic module over A_K. Because M_K is holonomic, it possesses *finite length*.

Consider the following decreasing sequence of A_K-submodules generated by increasing powers of p:

$$A_K p^s \supseteq A_K p^{s+1} \supseteq A_K p^{s+2} \supseteq \ldots$$

Because this is a descending chain of submodules within a module of finite length, the Ascending Chain Condition's dual (the Descending Chain Condition for Artinian modules) dictates that the sequence must eventually stabilize.

Therefore, there exists some integer $k \geq 0$ such that:

$$A_K p^{s+k} = A_K p^{s+k+1}$$

This equality of modules implies that the generator p^{s+k} can be written as an A_K-linear combination of the elements in the smaller module. Thus, there exists an operator $Q \in A_K$ such that:

$$p^{s+k} = Q(x, \partial, s) \cdot p^{s+k+1}$$

Since this is a formal identity in the variable s, we can perform a safe translation of the variable, substituting $s \mapsto s - k$. This yields:

$$p^s = Q(x, \partial, s-k) \cdot p^{s+1}$$

Let $\tilde{Q}(s) = Q(x, \partial, s-k)$. The coefficients of the operator $\tilde{Q}(s)$ are rational functions in $\mathbb{C}(s)$. We can extract the finite set of denominators appearing in these coefficients and let $b(s)$ be their least common multiple.

Multiplying the entire equation by the polynomial $b(s)$ clears all denominators, resulting in:

$$b(s)p^s = \left(b(s)\tilde{Q}(s)\right) \cdot p^{s+1}$$

Since $b(s)\tilde{Q}(s)$ is now an operator strictly with polynomial coefficients in s, we have successfully found a differential operator $P(s) = b(s)\tilde{Q}(s)$ and a polynomial $b(s) \in \mathbb{C}[s]$ satisfying the Bernstein-Sato functional equation. This completes the proof of existence. □

70. Introduction to Dedekind Domains

This lecture introduces Dedekind domains. We begin with a motivating example of a ring that lacks unique factorization and investigate how ideals restore this property. We then establish the formal algebraic definition, its geometric interpretation, and examine counterexamples that fail exactly one of the defining conditions.

70.1. **Failure of Unique Factorization.** The standard example of an integral domain that is not a Unique Factorization Domain (UFD) is the ring $R = \mathbb{Z}[\sqrt{-5}]$. This consists of all numbers of the form $m + n\sqrt{-5}$ where $m, n \in \mathbb{Z}$.

In this ring, we can factor the number 6 in two completely different ways:

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

The elements $2, 3, 1 + \sqrt{-5}$, and $1 - \sqrt{-5}$ are all irreducible, and they are not simply unit multiples of each other. Thus, unique factorization of elements fails.

Kummer discovered a way to restore unique factorization by introducing "ideal elements" (which later formalized into the modern concept of ideals). Instead of factoring elements, we factor the ideals generated by those elements. We can factor the principal ideals (2) and (3) into products of prime ideals:

$$(2) = \mathfrak{ab}$$
$$(3) = \mathfrak{cd}$$

Similarly, the principal ideals generated by the other factors decompose as:

$$(1+\sqrt{-5}) = \mathfrak{ac}$$
$$(1-\sqrt{-5}) = \mathfrak{bd}$$

When we multiply these ideals together, the two different element factorizations correspond to the exact same ideal factorization, simply rearranged:

$$(6) = (\mathfrak{ab})(\mathfrak{cd})$$
$$(6) = (\mathfrak{ac})(\mathfrak{bd})$$

Although the factorization of elements is not unique, the factorization into ideals is unique up to the order of the factors.

70.2. **Definition of a Dedekind Domain.** This property of ideals motivates the fundamental definition.

Definition 70.1. An integral domain R is a *Dedekind domain* if every non-zero proper ideal factors into a product of prime ideals in a unique way (up to order).

Because this definition is practically difficult to verify for a given ring, people universally use an equivalent structural definition.

Theorem 70.2. *An integral domain R is a* Dedekind domain *if and only if it satisfies the following three conditions:*

(1) R is a Noetherian ring.
(2) Every non-zero prime ideal of R is maximal.
(3) R is integrally closed in its field of fractions.

Recall that being integrally closed means that if we embed R into its field of fractions K, any element $x \in K$ that satisfies a monic polynomial equation with coefficients in R:

$$x^n + r_{n-1}x^{n-1} + \cdots + r_0 = 0$$

must already belong to R.

These three conditions are entirely independent; dropping any single one yields a ring that is no longer a Dedekind domain.

There are two primary sources of Dedekind domains:

- **Number Theory:** The ring of integers of any algebraic number field. For example, $\mathbb{Z}$ itself, or $\mathbb{Z}[\sqrt{-5}]$, which is the integral closure of $\mathbb{Z}$ in the field $\mathbb{Q}(\sqrt{-5})$.
- **Algebraic Geometry:** The coordinate rings of non-singular affine algebraic curves. For example, the polynomial ring $k[x]$ (the affine line), or the coordinate ring of an elliptic curve.

70.3. **Geometric Interpretation.** It is easier to understand these three arbitrary-seeming algebraic conditions if we interpret them geometrically using the coordinate rings of curves.

(1) **Noetherian:** In algebraic geometry, the Noetherian condition roughly translates to the space being "not weird". It ensures finite generation of defining equations.

(2) **Non-zero primes are maximal:** The Krull dimension is the length of the longest chain of prime ideals. Since (0) is prime, a chain looks like $(0) \subsetneq \mathfrak{m}$. This guarantees that the geometric dimension of the space is exactly 1.
(3) **Integrally closed:** Geometrically, being integrally closed corresponds to being a *normal* variety. A normal variety is smooth in codimension 1. Since our space has dimension 1, the singular locus must have dimension less than 0, meaning there are absolutely no singularities.

Together, these three conditions dictate that a geometric Dedekind domain corresponds exactly to a *non-singular algebraic curve*.

Unique factorization of ideals also possesses a natural geometric dictionary. For a curve over an algebraically closed field:

- **Prime ideals** correspond to individual points P_i on the curve.
- **Ideals** correspond to *divisors*, which are formal linear combinations of points with positive integer multiplicities: $\sum n_i P_i$.
- **Elements** correspond to rational functions on the curve.
- **The ideal generated by an element** corresponds to the zeros of that function (counted with multiplicity).

From this viewpoint, unique factorization of ideals is geometrically trivial: it merely states that any effective divisor (a collection of points with multiplicities) can be uniquely expressed as a sum of individual points.

70.4. **Counterexamples.** We can better understand the necessity of the three conditions by providing counterexamples where exactly one condition is dropped.

70.4.1. *Not Noetherian.* Consider the ring of Puiseux series over a field k:

$$R = \bigcup_{n=1}^{\infty} k[[x^{1/n}]]$$

This ring is 1-dimensional and integrally closed, but it is vastly non-Noetherian. We can form an ideal generated by all positive rational powers of x:

$$I = (x, x^{1/2}, x^{1/6}, x^{1/24}, \dots)$$

Notice that $I^2 = I$. Consequently, we completely lose any notion of unique factorization into primes.

An analytic example is the ring of holomorphic functions on $\mathbb{C}$. Using Weierstraß factors, a holomorphic function can be factored by its roots, but a function might have infinitely many roots, requiring an infinite product of prime elements. This violates the algebraic requirement of finite generation and finite factorization.

70.4.2. *Not Integrally Closed.* Let $R = \mathbb{Z}[\sqrt{-3}]$. This is a 1-dimensional Noetherian ring, but it is not integrally closed because the element $(1 + \sqrt{-3})/2$ is integral over R but not contained in it.

Geometrically, this corresponds to a curve with a singularity. Consider the coordinate ring of a cusp:

$$S = \frac{k[x,y]}{(y^2 - x^3)}$$

If we parameterize this by setting $t = y/x$, the ring S embeds as $k[t^2, t^3] \subset k[t]$. It is not integrally closed. In both cases, ideals centered at the singularity fail to factor uniquely into prime ideals. For instance, we may find two distinct ideals $\mathfrak{a}$ and $\mathfrak{b}$ such that $\mathfrak{a}^2 = \mathfrak{a}\mathfrak{b}$.

70.4.3. *Dimension Greater Than One.* If we drop the condition that non-zero primes are maximal, we permit geometries of dimension 2 or higher.

- An arithmetic example is the polynomial ring $\mathbb{Z}[x]$.
- A geometric example is the affine plane $k[x, y]$.

In these rings, ideals can be extraordinarily complicated. For example, the ideal (x^3, x^2y, y^2) cannot be factored into a product of prime ideals. Instead, these rings satisfy the weaker *Lasker-Noether theorem*, which states that every ideal is a finite *intersection* of primary ideals, but not necessarily a product.

70.5. **Relation to Unique Factorization Domains.** It is important to contrast Dedekind domains with Unique Factorization Domains (UFDs).

- **PID implies both:** If R is a Principal Ideal Domain (PID), then R is both a UFD and a Dedekind domain.
- **UFD does not imply Dedekind:** The polynomial ring $k[x, y]$ is a UFD (elements factor uniquely), but it has dimension 2, so it is not a Dedekind domain.
- **Dedekind does not imply UFD:** The ring $\mathbb{Z}[\sqrt{-5}]$ is a Dedekind domain, but as we saw, elements do not factor uniquely.

Thus, neither class contains the other. While there exists a broader class known as *Krull domains* that generalizes and contains both UFDs and Dedekind domains, they represent fundamentally different algebraic concepts: one repairs the factorization of elements, while the other abstracts the smooth geometry of curves.

References

Modern and Standard Texts

1. Atiyah, Michael F., and Macdonald, Ian G. *Introduction to Commutative Algebra*. Addison-Wesley Publishing Co., Reading, Mass., 1969.
2. Eisenbud, David. *Commutative Algebra: with a View Toward Algebraic Geometry*. Graduate Texts in Mathematics, 150. Springer-Verlag, New York, 1995.
3. Matsumura, Hideyuki. *Commutative Ring Theory*. Translated from the Japanese by M. Reid. Cambridge Studies in Advanced Mathematics, 8. Cambridge University Press, Cambridge, 1986.
4. Serre, Jean-Pierre. *Local Algebra*. Translated from the French by CheeWhye Chin and revised by the author. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2000.

Classic and Foundational Texts

5. Bourbaki, Nicolas. *Commutative Algebra. Chapters 1–7*. Elements of Mathematics. Springer-Verlag, Berlin, 1989.
6. Grothendieck, Alexander, and Dieudonné, Jean. *Éléments de géométrie algébrique* (EGA). Publications Mathématiques de l'IHÉS, Bures-sur-Yvette, 1960–1967.
7. Nagata, Masayoshi. *Local Rings*. Interscience Tracts in Pure and Applied Mathematics, No. 13. Interscience Publishers, New York-London, 1962.
8. Zariski, Oscar, and Samuel, Pierre. *Commutative Algebra. Vol. I, II*. Graduate Texts in Mathematics, Vol. 28, 29. Springer-Verlag, New York-Heidelberg, 1975.

Historical Texts

9. Macaulay, Francis S. *The Algebraic Theory of Modular Systems*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 19. Cambridge University Press, Cambridge, 1916.