MATH 210A NOTES

ARUN DEBRAY
DECEMBER 16, 2013

These notes were taken in Stanford's Math 210A class in Fall 2013, taught by Professor Zhiwei Yun. I TEXed them up using `vim`, and as such there may be typos; please send questions, comments, complaints, and corrections to `adebray@stanford.edu`.

Contents

Part 1. Basic Theory of Rings and Modules

1. Introduction to Rings and Modules: 9/23/13

This class doesn't require a lot of group-theoretical background, just the basic definitions and examples. But the linear-algebraic prerequisite is very important, because in this class it will be generalized to statements about

modules and rings. We care about groups because they act on sets: $G \circlearrowright X$. One can study groups abstractly, but the examples end up being about actions. Similarly, in linear algebra one has linear transformations acting on a vector space. In this class, this will be generalized to rings acting on modules: $R \circlearrowright M$.

Definition. A ring $(R, 0, 1, +, \cdot)$ is a set R with two distinguished elements $0, 1 \in R$ and two binary operations $+, \cdot : R \times R \to R$ that satisfy the following axioms:

- $(R, 0, +)$ is an abelian group.
- $(R, 1, \cdot)$ is a semigroup; that is, it satisfies every axiom of a group except that not all elements must be invertible.[1]
- The two operations are compatible: for any $a, b, c \in R$, $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$.
- $0 \neq 1$.[2]

Example 1.1.

- The ring of integers, $\mathbb{Z}$, is the One Ring and the most frequent example.
- Some larger rings: $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$. In these rings, every nonzero element is invertible; they are fields.
- $\mathbb{H} = \mathbb{R} \cdot 1 \oplus \mathbb{R} \cdot i \oplus \mathbb{R} \cdot j \oplus \mathbb{R} \cdot k$, the quaternions, a four-dimensional vector space over $\mathbb{R}$ with relations $ij = -ji = k$, $jk = -kj = i$, and $ki = -ik = j$, and $i^2 = j^2 = k^2 = -1$. This is a ring in which multiplication is not commutative, but every nonzero element is still invertible.

Given a ring R, one can adjoin some formal variables to get the polynomial ring $R[x_1, \ldots, x_n]$; if R is commutative, then this is also. However, one can also introduce noncommutative formal variables:

Example 1.2. The ring $\mathbb{Z}\langle x, y\rangle = \mathrm{Span}_{\mathbb{Z}}\{x^{a_1}y^{b_1}x^{a_2}y^{b_2}\cdots\}$ is the free abelian group on the set of strings over x and y. Addition is componentwise, and multiplication is given by concatenation, so $xy \neq yx$. Here, 1 is the empty word.

Another noncommutative ring is $\mathrm{Mat}_n(k)$, the ring of $n \times n$ matrices over a field k.

In addition to making rings larger, it is possible to obtain smaller rings out of rings, using quotients, such as $\mathbb{Z}/n\mathbb{Z}$. This works with polynomial rings, too: if k is a field, then $k[x]/(x^3 - 1)$ implies that one only cares about residue classes modulo $x^3 - 1$. More generally, given $f_1, \ldots, f_r \in R[x_1, \ldots, x_n]$, one can take the ideal generated by them, $(f_1, \ldots, f_r)$, and then quotient out by it: $R[x_1, \ldots, x_n]/(f_1, \ldots, f_r)$.

Definition. An ideal is a subset of a ring $I \subset R$ such that:

- I is a subgroup of $(R, 0, +)$, and
- for any $a \in R$ and $b \in I$, $ab, ba \in I$.

Thus, an ideal $I \subset R$ is closed under left and right multiplication by R, not just I. For example, $k[x^2] \subset k[x]$ is a subring, but not an ideal, because $1 \in k[x^2]$, but $x \cdot 1 \notin k[x^2]$. In fact, this illustrates that if $I \subset R$ is an ideal and $1 \in I$, then $I = R$.

The notation $(x_1, \ldots, x_n) \subset R$ means the ideal generated by the x_i, which is the set of linear combinations of the x_i weighted by elements of R:

$$(x_1, \ldots, x_n) = \left\{ y \mid y = \sum a_i x_i b_i, a_i, b_i \in R \right\},$$

though if R is commutative this simplifies to

$$= \left\{ y \mid y = \sum a_i x_i, a_i \in R \right\}.$$

Why do we care about ideals?

Lemma 1.1. *If $I \subset R$ is an ideal, then R/I taken as a quotient group under addition[3] has a natural ring structure.*

Proof. The multiplication is, of course, given as follows: for $a, b \in R/I$, choose their representatives $\tilde{a}, \tilde{b} \in R$, and define $ab = \tilde{a}\tilde{b} \bmod I$. It is necessary to check that $\tilde{a}\tilde{b}$ is well-defined in R for any possible choices of representatives, but this is where I is required to be an ideal: if $\tilde{a}_1, \tilde{b}_1$ and $\tilde{a}_2, \tilde{b}_2$ are two choices of representatives of a and b in R, then $\tilde{a}_1 = \tilde{a}_2 + i$ for an $i \in I$, and similarly $\tilde{b}_1 = \tilde{b}_2 + j$ with $j \in I$. Thus,

$$\tilde{a}_1\tilde{b}_1 = (\tilde{a}_2 + i)(\tilde{b}_2 + j) = \tilde{a}_2\tilde{b}_2 + j\tilde{a}_2 + i\tilde{b}_2 + ij \equiv \tilde{a}_2\tilde{b}_2 \bmod I. \qquad \boxtimes$$

Definition. If $(R, 0, 1, +, \cdot)$ and $(R', 0', 1', +', \cdot')$ are rings, then a set map $f : R \to R'$ is called a ring homomorphism if:

[1] The standard terminology for this is a monoid; a semigroup is usually not required to have an identity.

[2] This is required to eliminate the pathological example $R = \{0\}$, where $0 = 1$, since there's no good reason to call this a ring.

[3] The lemma statement brushes over the fact that R/I is a quotient group at all, but since I is an ideal, then it is a subgroup of R additively, and since R is an abelian group, then $I \trianglelefteq R$, and R/I is in fact a group.

- f is a group homomorphism under $+$ and $+'$.
- $f(ab) = f(a)f(b)$ for any $a, b \in R$.
- $f(1) = 1'$.

From the first requirement, one also can see that $f(0) = 0'$. In the homework you are asked to find a nontrivial group homomorphism that satisfies the second condition but not the third, so that it is in some sense almost a ring homomorphism. The trivial $f : x \mapsto 0$ is a sillier example. In this class, though, all rings will have 1 and all ring homomorphisms will preserve it.

Definition. The kernel of a ring homomorphism f is the kernel of the underlying group homomorphism: $\ker(f) = f^{-1}(0)$.

Kernels and ideals behave well in an analogue to the isomorphism theorems from group theory:

Theorem 1.2. *If $f : R \to R'$ is a ring homomorphism, then:*

(1) $\ker(f)$ *is an ideal of* R*, and*

(2) $R/\ker(f) \xrightarrow{\sim} \mathrm{Im}(f)$ *as rings.*[4]

These are the basic notions; now for some more examples. Rings appear everywhere in mathematics, such as these examples from analysis.

Example 1.3. Let X be a manifold, such as $\mathbb{R}^n$, and let $C(X, \mathbb{C})$ be the set of continuous complex-valued functions on X.[5] Then, $C(X, \mathbb{C})$ is a ring under pointwise addition and multiplication: if $f, g : X \to \mathbb{C}$, then $(f+g)(x) = f(x)+g(x)$ and $(fg)(x) = f(x)g(x)$. This is a commutative ring because $\mathbb{C}$ is.

Example 1.4. Let $C^\infty(\mathbb{R})$ denote the set of smooth functions $\mathbb{R} \to \mathbb{R}$.

Definition. A differential operator

$$D = a_0(x) + a_1(x)\frac{\mathrm{d}}{\mathrm{d}x} + a_2(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \cdots + a_n(x)\frac{\mathrm{d}^n}{\mathrm{d}x^n}$$

is a finite linear combination of taking derivatives and multiplying by smooth functions $a_0, \ldots, a_n \in C^\infty(\mathbb{R})$.

Let $\mathcal{D}$ be the set of differential operators. Then, $\mathcal{D}$ is closed under addition and function composition, and forms a ring under these operations. However, because of the Leibniz Rule, it isn't commutative: $\frac{\mathrm{d}}{\mathrm{d}x}(xf) = f + x\frac{\mathrm{d}f}{\mathrm{d}x}$.

Rings also appear in topology:

Example 1.5. If this example doesn't make any sense, don't worry. But if X is a "reasonable" topological space, then the cohomology ring $H^*(X, \mathbb{Z}) = \bigoplus H^i(X, \mathbb{Z})$ is a graded ring (i.e. it has extra structure) with the multiplication given by the cup product $\smile : H^i(X, \mathbb{Z}) \times H^j(X, \mathbb{Z}) \to H^{i+j}(X, \mathbb{Z})$.

Connections between rings and groups. Given a group G, one can enlarge it to form a ring. Given a ring R, the group ring $R[G]$ is given as

$$R[G] = \left\{ \sum_{g \in G} a_g \cdot [g] \mid a_g \in R, \text{ and only finitely many are nonzero} \right\}.$$

Here, the brackets are used to denote a formal expression representing a group element. The set of all such linear combinations is denoted R^G, the G-fold Cartesian product,[6] and the finitely supported set is more commonly written as $\bigoplus_{g \in G} R$, but multiplication is defined first on elements of the form $1 \cdot [g]$: $(1 \cdot [g])(1 \cdot [h]) = 1 \cdot [gh]$, which is the reasonable thing to do. This can be extended to be R-linear; for example, $(a_g[g] + a_{g'}[g']) \cdot [h] = a_g[gh] + a_{g'}[g'h]$. The requirement that there be only finitely many nonzero terms ensures this sort of multiplication is possible, as there is no *a priori* definition of an infinite sum in R.

To confuse you a bit more, there is another definition of multiplication on $\bigoplus_{g \in G} R$, viewed as the set of finitely supported functions $G \xrightarrow{f} R$ (i.e. the set of functions f for which $f(g) \neq 0$ for only finitely many $g \in G$). Here, multiplication is defined pointwise, but be warned! This is *not* a ring if G is infinite, because then the identity function $e : x \mapsto 1$, doesn't have finite support. However, since $1 \in R^G$, then R^G is a ring under pointwise multiplication, even if G is infinite.

[4] This holds even though $\mathrm{Im}(f)$ is in general not an ideal of R', though it is always a subring.

[5] This works identically as well if the functions are chosen to be smooth, or C^r, or differentiable, etc.

[6] This is also known as the Cartesian power, and is just copies of R indexed by G.

If G is finite, this gives two ring structures on $R[G] = \bigoplus_{g \in G} R$. They are in general different; if G is noncommutative, the former is noncommutative as well, but the latter is commutative. In fact, the latter structure is less interesting in general, since it doesn't use any of the group structure. Group rings, however, are used to study the representations of finite groups, which leads to the notion of modules.

Definition. If R is a ring, a left R-module is an abelian group $(M, 0, +)$ together with an action of R, $R \times M \to M$ denoted $(a, m) \mapsto a \cdot m$, which satisfies the following natural conditions:

- $\cdot$ is bilinear: $(a + b) \cdot m = a \cdot m + b \cdot m$, and $a \cdot (m + n) = a \cdot m + a \cdot n$.
- $\cdot$ is associative: $(a \cdot b) \cdot m = a \cdot (b \cdot m)$. This axiom is tricky; which dots mean which kind of multiplication?
- $1 \cdot m = m$.

Sometimes, one hears "R acts on M" [from the left] to mean that M is a left R-module, and by default, modules are left modules.

There is an obvious notion of homomorphism between modules, so the definition wasn't given in class.

Example 1.6. The ring of $n \times n$ matrices over a field k, $\mathrm{Mat}_n(k)$, acts on n-dimensional vector space over k, $V = k^n$, with the action given by matrix multiplication. Thus, V becomes a $\mathrm{Mat}_n(k)$-module. Similarly, $\mathrm{Mat}_{n \times m}(k)$ is acted on by $\mathrm{Mat}_n(k)$, again by left multiplication.

Example 1.7. The set $C(X, \mathbb{C})$ of continuous complex-valued functions on a topological space X act on $C(Y, \mathbb{C})$, where $Y \subset X$ is a topological subspace. The action is given by restriction: $f \in C(X), g \in C(Y) \mapsto fg \in C(Y)$, with multiplication defined pointwise. This can also be seen as multiplying $f|_Y$ and g in $C(Y)$.

Alternatively, let E be a vector bundle over X and $C(E)$ be the set of its continuous sections.[7] It turns out that one can multiply sections by functions, and obtain an action.

Example 1.8. If $I \subset R$ is an ideal, then R/I is an R-module by left multiplication. R/I is also a ring, as seen above, and this has a special name.

Definition. If R is a ring, then an R-algebra is a ring A together with a ring homomorphism $R \to A$.

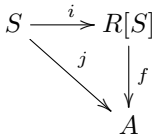
This induces a natural R-module structure on A, given by "multiplication" through the homomorphism. In Example 1.8, $A = R/I$, and the homomorphism is the quotient map.

Lemma 1.3. *If R and A are rings, then to have a ring homomorphism $R \to A$ is the same as defining an R-module structure on A.*

Proof. If $f : R \to A$, then the module structure is given by $(r, a) \mapsto f(r) \cdot a$, and it can be checked that this satisfies the axioms. In the other direction, just restrict the module action to 1. ⊠

As an example, it is possible to define polynomial rings more generally, and in a different way. Let R be a polynomial ring and S a set. Then, in an abuse of notation not to be confused with group rings, the polynomial ring generated by the set S, written $R[S]$, is characterized by the following universal property:

- There is a map $i : S \to R[S]$ (the map that sends each element to its corresponding formal variable).
- For any commutative ring A together with a map $j : S \to A$, there exists a unique ring homomorphism $f : R[S] \to A$ such that the following diagram commutes:



That is, $f \circ i = j$.

This property characterizes $R[S]$ uniquely, a fact that requires proof, though a similar characterization for the noncommutative case $R\langle S \rangle$ will be presented in the homework.

2. More Kinds of Rings and Modules: 9/25/13

For this lecture, all rings will be commutative.

Definition. If A is a ring and M and N are A-modules, then an A-linear map $f : M \to N$ is an abelian group homomorphism such that $f(ax) = af(x)$ for all $a \in A$ and $x \in M$.

[7] One example, but not the only example, is $E = X \times \mathbb{R}^n$, with sections as functions $f : X \to X \times \mathbb{R}^n$ such that $f(x) = (x, \dots)$. Generally, sections are more interesting.

Following this definition, $\ker(f)$, $\mathrm{Im}(f)$, and $\mathrm{coker}(f)$, defined identically to the abelian group case, are also A-modules.

Definition. A free A-module is a direct sum of copies of A: $M = \bigoplus A$. If this is a finite sum, it becomes $M = A^n$, and if it is infinite, the direct sum means that only finitely many terms of each coordinate can be nonzero.

Not all modules are free, which forms a crucial difference between modules and vector spaces.

Exercise 2.1. Suppose that A is a ring and $I \neq (0)$ is an ideal of A. Show that A/I is not a free A-module.

Additionally, submodules of free modules aren't always free. For example, if $A = k[x^2, x^3] = \mathrm{Span}\{1, x^2, x^3, x^4, x^5, \dots\}$, then A is the ring of all polynomials without linear terms, and $A \subset k[x]$ is a subring of codimension 1 (viewing $k[x]$ as a vector space). Let $M \subseteq A$ be given by $M = \mathrm{Span}\{x^2, x^3, x^4, \dots\} = x^2 k[x]$. Then, M is an ideal in A, because multiplying an $a \in A$ by an $m \in M$ can only increase its minimum degree (and therefore eliminate any constant or linear terms), but M is not itself free. To see why, notice that at least two elements are necessary to generate M; if $f \in M$, then $A \cdot f \subsetneq M$ because if the lowest term of f has degree d, then $A \cdot f$ is missing polynomials with terms of degree $d+1$, since A has no terms of degree 1. However, any two elements $f, g \in M$ satisfy some relations such as $fg = gf$. This seems silly, but imagine the first factor on each side to be in A and the second in M: $fg = gf$. This means that they aren't linearly independent in M, since they satisfy this linear combination with weights in A, but no single element generates M, and thus M cannot be free. There are plenty of other examples: if $A = k[x, y]$, define $M \subset A$ as $M = (x, y)$, which is similarly not generated by a single element, yet is not free. There are also subtleties as to whether a module is finitely generated or not.

Definition. An A-module M is finitely generated if there is a surjective homomorphism $A^n \twoheadrightarrow M$ (where n is finite).

Here, the coordinate elements of A^n are mapped to the generators of M.

Definition. An A-module M is torsion-free if there are no $m \in M \setminus 0$ and $a \in A \setminus 0$ such that $am = 0$.

Example 2.1. If $A = \mathbb{Z}$, then $\mathbb{Z}$-modules are just abelian groups, so there is no extra structure. It is a fact[8] that submodules of free $\mathbb{Z}$-modules, or equivalently free abelian groups, $M \subset \mathbb{Z}^{\oplus n}$ are free, so long as n is finite. Another fact is that if M is a finitely generated, torsion-free $\mathbb{Z}$-module (which is the same notion as that of a torsion-free abelian group), then M is free.

Here, being finitely generated is important: if M isn't finitely generated, but is torsion-free, it isn't necessarily free. The simplest example is $M = \mathbb{Q}$. This is a field, so it is torsion-free, but since any two elements are linearly dependent, then it isn't free.

Theorem 2.1 (Chinese Remainder Theorem)**.** *Let A be a commutative[9] ring and $I_1, \dots, I_n \subset A$ be ideals such that $I_i + I_j = A$ whenever $i \neq j$.[10] Then, for any $b_1 \in A/i_1, \dots, b_n \in A/I_n$, the system of congruence equations*

$$x_i \equiv b_i \pmod{I_i}, i = 1, \dots, n \tag{1}$$

has a solution in A which is unique modulo $\bigcap_{i=1}^n I_i$.

This is a slightly generalized setting of the classical Chinese Remainder Theorem, in which $A = \mathbb{Z}$, $I_i = (m_i)$, and $m_1, \dots, m_n$ are pairwise coprime, so that $(m_i) + (m_j) = \mathbb{Z}$ if $i \neq j$. Then, one can always solve a system of equations

$$x \equiv b_i \pmod{m_i}, i = 1, \dots, n$$

in the integers: some $x \in \mathbb{Z}$ solves them all simultaneously.

Proof of Theorem 2.1. The proof itself is going to be fairly simple, but first it is necessary to restate the theorem slightly. We have the quotient homomorphisms $A \to A/I_i$, so there is a ring homomorphism $f : A \to \prod_{j=1}^n A/I_j$ (defined componentwise). Then, the statement that (1) has a solution is equivalent to f being surjective, and its uniqueness is equivalent to the existence of a ring isomorphism

$$A/\bigcap_{j=1}^n I_j \xrightarrow{\sim} \prod_{j=1}^n A/I_j,$$

or that $\ker(f) = \bigcap I_j$.

Now, we can completely ignore the equations, and focus on the structure of the rings. In fact, it makes the second part of the theorem much clearer, assuming the first part: if $f(x) = 0$, then $x \in I_i$ for all I_i, so $\ker(f) = \bigcap I_i$.

[8] This will be proven in greater generality next lecture.

[9] This isn't strictly necessary for the hypothesis, but it makes the argument easier.

[10] The sum of two ideals I and J of a ring R is $I + J = \{i + j \mid i \in I, j \in J\}$, and it is also an ideal of R.

The first part is harder. Notice that it suffices to show that the coordinate elements (that is, the basis elements $e_i = (0, \dots, 1, \dots, 0)$) are in $\mathrm{Im}(f)$, since if $a_i \mapsto e_i$, then $\sum b_i a_i \mapsto (b_1, \dots, b_n)$. In fact, since they're all the same, just pick one; the argument for the rest is the same. What is the preimage of $(1, 0, \dots, 0)$? In other words, what is the element of A that is congruent to 1 mod I_1, but is in $I_2, \dots, I_n$? We haven't yet used the fact that $I_1 + I_j = A$ for $j \neq 1$, so there exist $x_j \in I_1$ and $y_j \in I_j$ such that $x_j + y_j = 1$ for each $j = 2, \dots, n$. Then, $x = y_2 y_3 \cdots y_n$ will be the solution: certainly, it is in each of $I_2, \dots, I_n$, and modulo I_1, each relation becomes $y_i \equiv 1 \bmod I_j$, so their product must be as well. ⊠

Special Kinds of Commutative Rings. Interesting things can be said by imposing conditions on commutative rings.

Definition. A is a domain if whenever $a, b \in A$ and $ab = 0$, then $a = 0$ or $b = 0$.

In other words, A cannot have any zero divisors. In domains, one has cancellation: $ab = ac$ implies $b = c$.

Definition. An ideal $I \subset A$ is prime if whenever $a, b \in A$ and $ab \in I$, then $a \in I$ or $b \in I$.

It is an easy consequence of these two definitions that A is a domain iff (0) is a prime ideal. Additionally (almost the same statement), I is prime if A/I is a domain. Generally, prime ideals are denoted with the German $\mathfrak{p}$.

Definition. An ideal $I \subset A$ is maximal if:

- $I \neq A$, and
- If $J \supsetneq A$ is an ideal of A, then $J = A$.

In the partial order of ideals, (0) is minimal and A is maximal, but the maximal ideals are those just below I. In some sense, we don't want to call A an ideal, because A/A is not a useful ring. But it meets the technical definition.

Proposition 2.2. *I is a maximal ideal of A iff A/I is a field.*

Proof. $\Longleftarrow$ is clear: if $J \supsetneq I$, then J/I is an ideal of A/I, but A/I is a field, so either $J/I = A/I$ or is zero.
$\Longrightarrow$ is harder: if I is maximal, then A/I has only (0) and A/I as ideals. If $\overline{a}$ is not invertible, then $(\overline{a}) \neq A/I$, so $(\overline{a}) = (0)$, or $\overline{a} = 0$. Thus, every nonzero element of A/I is invertible, so it is a field. ⊠

See how these quickly fail if A isn't required to be commutative.

Corollary 2.3. *Since fields are integral domains, then I is a maximal ideal implies I is a prime ideal.*

The converse is not true: $(x, y) \subset k[x, y]$ is maximal, because $k[x, y]/(x, y) \cong k$ is a field, but $(x) \subset (x, y)$, so it isn't maximal, and $k[x, y]/(x) \cong k[x]$, which is an integral domain, but not a field. Thus, (x) is prime, but not maximal.

If A and B are commutative rings and $f : A \to B$ is a ring homomorphism, and $\mathfrak{p} \subset B$ is a prime ideal, then $f^{-1}(\mathfrak{p})$ is a prime ideal of A. Thus, prime ideals are closed under preimage, but not under image. Neither direction works for maximal ideals: with $\mathbb{Z} \hookrightarrow \mathbb{Q}$, (0) is maximal in $\mathbb{Q}$, but not $\mathbb{Z}$.

Definition. If A is a domain and $x \in A \setminus 0$, then x is irreducible if whenever $x = yz$ for $y, z \in A$, then either y or z is invertible.

This is a natural generalization of the prime numbers over $\mathbb{Z}$, or the irreducible polynomials over a field (i.e. $A = k[x]$).

Definition. A is a unique factorization domain (UFD) if every nonzero $x \in A$ can be written as a finite product of irreducible elements, and this expression is unique up to reordering and multiplication by units.

For example, if u_1 is a unit and $y_1, \dots$ are irreducibles, then $x = y_1 y_2 y_3 = (u_1 y_1) y_2 (u_1^{-1} y_3)$ are considered equivalent.

Both $\mathbb{Z}$ and $k[x]$ are UFDs, as is $\mathbb{Z}[i]$ (where $i^2 = -1$). This latter ring is a free $\mathbb{Z}$-module of rank 2, spanned by 1 and i. As a ring (the ring of Gaussian integers), this is a UFD, which is a little less obvious than in the previous cases.

Example 2.2. A famous non-example is $\mathbb{Z}[\sqrt{-5}]$, because $(1+\sqrt{-5})(1-\sqrt{-5}) = (2)(3) = 6$, and $2, 3$, and $1 \pm \sqrt{-5}$ are irreducibles that can't be transformed into each other with units. Every element factors, but not necessarily in a unique way. Similarly, in $k[x^2, x^3]$, $x^6 = x^2 \cdot x^2 \cdot x^2 = x^3 \cdot x^3$, which is a non-unique factorization (since $x \notin k[x^2, x^3]$, so each of these is irreducible).

Example 2.3. A more disturbing non-example is given by $\mathbb{C}[\sqrt{x}, \sqrt[3]{x}, \sqrt[4]{x}, \dots]$. Any expression in x in this ring has a nonterminating factorization. Worse still, this is a perfectly natural object that arises in algebra, along with its cousin, its ring of power series: $\mathbb{C}[[x, \sqrt{x}, \sqrt[3]{x}, \sqrt[4]{x}, \dots]]$. The field of Laurent series $\mathbb{C}((x))$ (power series with a finite number of negative-degree terms) has algebraic closure $\overline{\mathbb{C}((x))} = \mathbb{C}((x, \sqrt{x}, \sqrt[3]{x}, \sqrt[4]{x}, \dots))$, which is a polynomial

indexed over $\mathbb{Q}$ with the finiteness conditions that there are only finitely many negative terms and the sizes of the denominators is bounded. You can check that this is a field, and $\mathbb{C}[x, \sqrt{x}, \sqrt[3]{x}, \dots]$ shows up when one wants to solve polynomial equations in $\mathbb{C}((x))$.

Theorem 2.4. *If A is a UFD, then so is $A[x_1, \dots, x_n]$.*

Proof sketch. The problem will be reduced in several steps:

1. By induction, it's clearly only necessary to show that $A[x]$ is a unique factorization domain if A is.
2. When A is a field, $A[x]$ admits the Euclidean algorithm for calculating the greatest common divisor of two elements, which implies that $A[x]$ is a UFD. This is considerably easier than the general case.

Let's try to list the irreducibles in $A[x]$. If $a \in A$ is irreducible, then $a \in A[x]$ is too. There are less trivial irreducibles, too: the polynomials of positive degree that are irreducibles in $K[x]$. Here, K is the field of fractions of A.[11]

The coefficients of f could have common divisors, so force the greatest common divisor of the coefficients of f to be 1 (i.e. if a divides all of the coefficients of f, then $a = 0$ or a is a unit).

Now, let's try to factorize a general polynomial $f \in A[x]$. First, let a be the greatest common divisor of coefficients in f, and write $f(x) = af_1(x)$. This makes sense because A is a UFD, so a is well-defined up to units.

Definition. If $f \in A[x]$ and the greatest common divisor of the coefficients of f is 1, then f is called primitive.

So the f_1 given above is primitive. Since $K[x]$ is a UFD, then f_1 can be factorized there. Considering the case where $f_1 = gh$ for irreducible g and h, it is possible to clear the denominators, by multiplying by the right units (scalars) in $K[x]$, so one has $cf_1 = g_1h_1$, where $c \in A$ is the thing that it was necessary to multiply by.

Lemma 2.5 (Gauss)**.** *c is a unit in A.*

Proof. If not, then take an irreducible $a \mid c$ and consider $cf_1 = g_1h_1 \bmod a$, giving $0 = \overline{g_1}(x)\overline{h_1}(x)$ in $A/(a)[x]$. But! $A/(a)$ is an integral domain because (a) is a prime ideal (since A is a UFD). Thus, $A/(a)[x]$ is a domain, leading to a contradiction. ⊠

This means that f factors into irreducibles. Uniqueness is not hard once existence is shown. ⊠

To understand this more concretely, consider $A = \mathbb{Z}$.

3. Tema con Variazioni, or Modules Over Principal Ideal Domains: 9/27/13

Definition. If A is a domain, then A is called a principal ideal domain (PID) if every ideal of A is generated by one element.

Typically, this would be written $I = (a)$ for an ideal $I \subset A$ and an $a \in A$. Two good examples are $\mathbb{Z}$ and $k[x]$ for a field k. However, $k[x, y]$ is not a PID, since (x, y) is not principal (as discussed last time).

In some sense, PIDs cannot be very big. There's a notion of a dimension of a ring, in which $k[x]$ has dimension 1 and $k[x, y]$ has dimension 2. In this schema, PIDs are required to have dimension 1. Though the definition won't formally be given in this class, it is occasionally useful for intuition.

Exercise 3.1. Show that if A is a PID, then A is a UFD.

Since every ideal is generated by one element, one can ask questions about which elements generate which ideals. Here are a couple facts:

- (a) is a prime ideal iff a is irreducible or $a = 0$ (since (0) is a prime ideal).
- (a) is a maximal ideal iff a is an irreducible element.

This is because if $(a) \subsetneq I = (b)$, then $a = b \cdot c$ because $a \in (b)$, and so on.

Corollary 3.1. *Almost every prime ideal in a PID (all except the zero ideal) is a maximal ideal.*

This is another consequence of the fact that PIDs have dimension 1.

[11] This construction can be made for any domain A, analogous to constructing $\mathbb{Q}$ from $\mathbb{Z}$ by defining quotients of elements with a notion of equivalence, etc.

The Structure Theorems. Today, several variations of the Structure Theorem for Finitely Generated Modules over a Principal Ideal Domain will be presented. Each such module can be built out of the following building blocks: the ring of scalars A, and quotients $A/I \cong A/(a)$ for some $a \in A$.

Theorem 3.2 (Structure Theorem for Finitely Generated Modules Over a PID, Version A)**.** *Let A be a PID, and m be a finitely-generated A-module. Then,*

$$M \cong A^r \oplus A/(a_1) \oplus A/(a_2) \oplus \cdots \oplus A/(a_n),$$

where $r \in \mathbb{Z}_{\geq 0}$, $a_1, \ldots, a_n \in A \setminus 0$, and $a_1 \mid a_2 \mid \cdots \mid a_n$. Moreover, r is unique for this choice of M, and $(a_1, \ldots, a_n)$ are as well, up to multiplication by units.

The finite generation condition is very important: there are infinitely generated modules that can't be written in this way, such as $\mathbb{Q}$ or $\mathbb{R}$ as $\mathbb{Z}$-modules.

The first obvious step is to consider consequences and special cases.

Definition. An A-module M is called torsion if for every $m \in M$, there exists a nonzero $a \in A$ which kills it (i.e. so that $am = 0$).

This is the other extreme to the idea of a torsion-free module discussed in the previous lecture, which is analogous to the concept of an integral domain.

Corollary 3.3. *If M is a finitely generated module over a PID, then M is torsion-free iff M is free.*

The reverse direction of the equivalence is true for modules over any domain, and the forward direction requires the Structure Theorem 3.2.

Corollary 3.4. *If A is a PID and M is a finitely-generated torsion A-module, then*

$$M \cong \bigoplus_{i=1}^{n} A/(a_i)$$

such that $a_1 \mid a_2 \mid \cdots \mid a_n$, and subject to the same uniqueness condition as in Theorem 3.2.

Definition. The annihilator of an A-module M is the ideal of A $\operatorname{Ann}(M) = \{a \in A \mid a \cdot M = 0\}$.

When studying the structure of objects it is useful to have some invariants, so in the context of Corollary 3.4, the ideal $\chi_M = (a_1 \cdots a_n)$ is called the *characteristic*, by analogy with the characteristic polynomial of a matrix.[12] Another invariant is the annihilator of (a_n), denoted m_M. This ideal kills all of M, and is the analogue of the matrix-theoretic minimal polynomial. It is the largest such ideal that kills M; if b kills M, then $b \mid a_n$.

Theorem 3.5 (Structure Theorem for Finitely Generated Modules Over a PID, Version B)**.** *If A is a PID and $M \subset A^{\oplus n}$ as modules, then there exists a basis $e_1, \ldots, e_n$ for $A^{\oplus n}$ and nonzero elements $a_1 \mid a_2 \mid \cdots \mid a_m$ for some $m \leq n$, such that $M = \operatorname{Span}\{a_1 e_1, \ldots, a_m e_m\}$. Moreover, the a_i are unique up to multiplication by units.*

This theorem says that submodules of free modules over a PID are free, and that by multiplying the basis elements for $A^{\oplus n}$, one obtains a basis for M. However, the uniqueness isn't obvious, even with the condition, so consider an example.

Example 3.1. Consider $2\mathbb{Z} \oplus 3\mathbb{Z} \subset \mathbb{Z} \oplus \mathbb{Z}$. The first one has as a basis $\{(2,0),(0,3)\}$, or the standard basis multiplied by 2 and 3, respectively. This does satisfy the basis condition in Theorem 3.5, but $2 \nmid 3$, so it doesn't satisfy the divisibility condition. Thus, a different basis is necessary for $\mathbb{Z} \oplus \mathbb{Z}$: choose $e_1 = (2,3)$ and $e_2 = (1,2)$.[13] Now, $M = \mathbb{Z}e_1 \oplus \mathbb{Z} \cdot 6e_2$, so it works.

This works because every element in M looks like $(2x, 3y)$ in $\mathbb{Z} \oplus \mathbb{Z}$, so solving

$$\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2x \\ 3y \end{pmatrix}$$

for $a, b \in \mathbb{Z}$ shows that b always happens to be a multiple of 6:

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 2x \\ 3y \end{pmatrix} = \begin{pmatrix} 2x - 3y \\ -6x + 6y \end{pmatrix}.$$

Thus, it works, and $(2x, 3y) = (2x - 3y) \cdot e_1 + (-x + y) \cdot 6e_2$.

[12] The $a_1, \ldots, a_n$ are only well-defined up to units, but the ideal is thus completely well-defined.

[13] To ensure that this is a basis, take the determinant: $\det(e_1\ e_2) = \pm 1$ is the requirement.

Theorem 3.6 (Structure Theorem for Finitely Generated Modules Over a PID, Version C)**.** *Let X be a matrix with entries in a PID A. Then, there exist invertible matrices U and V[14] over A such that*

$$UXV = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_m & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

for $a_1, \ldots, a_m \neq 0$ satisfying $a_1 \mid \cdots \mid a_m$ and the uniqueness condition from Theorem 3.5.

This theorem was implicitly referred to in Example 3.1: $\begin{pmatrix} 2 & \\ & 3 \end{pmatrix}$ represents M, but it doesn't satisfy the divisibility condition, and was thus fixed into $\begin{pmatrix} 1 & \\ & 6 \end{pmatrix}$. Here, $U = \begin{pmatrix} 2 & -3 \\ -1 & 1 \end{pmatrix}$, and $V = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$.

Of course, we will see eventually that all of these theorems are equivalent to each other. The last such theorem is a refinement of Theorem 3.2.

Theorem 3.7 (Structure Theorem for Finitely Generated Modules Over a PID, Version D)**.** *Let A be a PID and M be a finitely generated A-module. Then, there is a finite set $P \subset A$ of irreducible elements such that for each $p \in P$, there is a sequence of natural numbers $1 \leq e_1(p) \leq e_2(p) \leq \cdots \leq e_{n_p}(p)$ such that*

$$M \cong A^r \oplus \bigoplus_{p \in P} \left(\bigoplus_{i=1}^{n_p} A / \left(p^{e_i(p)} \right) \right),$$

where $r \in \mathbb{Z}$ is as in Theorem 3.2, and subject to the condition that all of the mentioned values are unique.

Notice that the divisibility condition is hidden into $e_i(p) \leq e_{i+1}(p)$.

Equivalence of the Structure Theorems. To show that Theorem 3.2 implies Theorem 3.7 is as easy as splitting $A/(a)$ into a sum of quotients by irreducible elements, but thanks to the Chinese Remainder Theorem (Theorem 2.1), it works, since A is a UFD. If $a = p_1^{t_1} p_2^{t_2} \cdots p_r^{t_r}$ for irreducibles $p_1, \ldots, p_r$ and $t_1, \ldots, t_r \in \mathbb{N}$, up to units and reordering, then we also want $p_1, \ldots, p_r$ to be distinct, which is easy to ensure. In particular, applying the CRT to $I_i = \left(p_i^{t_i}\right)$, they are pairwise coprime, so $I_i + I_j = A$ whenever $i \neq j$. Thus,

$$A/(a) = A/\bigcap_{i=1}^{r} I_i \cong \prod_{i=1}^{r} A/I_i = \bigoplus_{i=1}^{r} A/I_i.^{15}$$

Doing this for every factor $a_1, \ldots, a_n$ shows that Theorem 3.2 implies Theorem 3.7.

The reverse direction is also reasonable, but when assembling the prime powers,[16] it's necessary to meet the divisibility condition. This can be done by collecting them in the following manner:

$$\begin{array}{llll} p^{e_1} & q^{e_1'} & *^{e_1''} & = a_m \\ p^{e_2} & q^{e_2'} & *^{e_2''} & = a_{m-1} \\ p^{e_3} & q^{e_3'} & *^{e_3''} & = a_{m-2} \\ \vdots & \vdots & \vdots & \vdots \\ p^{e_\ell} & q^{e_\ell'} & *^{e_\ell''} & = a_3 \\ p^{e_m} & q^{e_m'} & & = a_2 \\ p^{e_n} & & & = a_1 \end{array}$$

Theorem 3.5 can be rewritten as

$$M = \bigoplus_{i=1}^{n} A \cdot a_i \cdot e_i \subset A^n = \bigoplus_{i=1}^{n} A \cdot e_i.$$

To show that Theorem 3.2 implies Theorem 3.5, apply the former to A^n/M. Certainly, this is finitely generated, so

$$A^n \twoheadrightarrow A^n/m \simeq A^r \oplus A/(a_1) \oplus \cdots \oplus A/(a_t).$$

---

[14] Since X is not necessarily a square matrix, then U and V might not be the same size. But their sizes are chosen to make the theorem work.

[15] $\times$ and $\oplus$ are the same in the finite case, but $\oplus$ is the preferred notation for modules.

[16] "[A]ssembling the prime powers" sounds like part of an evil supervillain's plans for a doomsday device, doesn't it?

The direct sum induces coordinate elements $x_1, \dots, x_r$ of the free part and $y_1, \dots, y_r$ of the torsion part; let $\tilde{x}_1, \dots, \tilde{x}_r$ and $\tilde{y}_1, \dots, \tilde{y}_r$ be their preimages in A^n. This might not yet be a basis of A^n. Ignoring the $\tilde{x}_i$ for now, $\tilde{y}_i \cdot a_i \to 0$ in A^n/M, so $\tilde{y}_i \cdot a_i \in M$ and thus are good candidates for the basis elements.

It's easy to show that the $\tilde{x}_i$ give a direct-sum decomposition for $A^n \simeq A^r \oplus A^{n-r}$, where the $\tilde{x}_i$ span A^r and A^{n-r} surjects onto the torsion part of the module. Since there's no interesting kernel in A^r, this means that $M \subset A^{n-r}$.

Now,

$$A^{n-r} \twoheadrightarrow A^{n-r}/M \cong \bigoplus_{i=1}^{t} A/(a_i).$$

The coordinate elements satisfy

$$\begin{array}{ccccccccc} & & A^t & = & A & \oplus & \cdots & \oplus & A \\ & \overset{\tilde{\varphi}}{\nearrow} & & & \downarrow & & & & \downarrow \\ A^{n-r} & \underset{\varphi}{\twoheadrightarrow} & A^{n-r}/M & \cong & A/(a_1) & \oplus & \cdots & \oplus & A/(a_n) \end{array}$$

where $\tilde{\varphi}$ is a surjective homomorphism chosen to make this work. With a little bit of head-scratching, this should lead to the implication.

To show that Theorem 3.5 implies Theorem 3.6, it will be demonstrated that one can diagonalize a matrix with the former theorem.

A matrix is just a map $A^M \xrightarrow{f} A^n$. Since one's allowed to left- and right-multiply by invertible matrices, one can choose any basis of A^m or A^n to work in. Applying Theorem 3.5 to $\mathrm{Im}(f) \subset A^n$, one has a basis $\{e_1, \dots, e_n\}$ of A^n and $a_1 \mid \cdots \mid a_r$ such that $\{a_1 e_1, \dots, a_r e_r\}$ is a basis for $A^r \cong \mathrm{Im}(f)$. In this sense, any section of the short exact sequence $0 \to A^{m-r} \cong \ker(f) \to A^n \to A^r \cong \mathrm{Im}(f) \to 0$ gives a decomposition of $A^m = A^r \oplus A^{m-r}$.

The other direction is mostly clear; though one must show that the submodule of a free module in this context is free, so that the diagonal matrix is an injective map. Once M is known to be free, Theorem 3.6 can be directly applied to prove Theorem 3.5.

4. Proof of the Structure Theorem for Modules over a Principal Ideal Domain: 10/2/13

"No modules were harmed in the making of this lecture."

Today, one of the structure theorem from the last lecture will be proven, and as indicated, this implies the result for the remaining three.

Proof of Theorem 3.7. The first step is to, of course, reduce the problem. Let $M_{\mathrm{tor}} = \{m \in M \mid a \cdot m = 0 \text{ for some } a \in A \setminus 0\}$.

Exercise 4.1.

(1) Show that $M_{\mathrm{tor}} \subset M$ is an A-submodule.
(2) Show that M/M_{tor} is torsion-free.

Assuming the above exercise, denote $N = M_{\mathrm{tf}} = M/M_{\mathrm{tor}}$. Then, the first major step in the proof is to show that N is of finite rank, or equivalently that every torsion-free module of finite rank is free.[17]

Proceed by induction on the minimal number of generators r of N as an A-module. When $r = 0$, the result is obvious, For $r > 0$, choose a generating set $\{x_1, \dots, x_r\}$ for N. Since r is minimal, then every one of these x_i is nonzero. Then, look at $A \cdot x_i \subset N$.

Definition. The saturation of an A-submodule N of a module M is $\mathrm{Sat}(N) = \{\frac{a}{b} n \mid n \in N, a, b \in A \text{ are in lowest terms}\}$.[18]

To saturate $A \cdot x_1$ is equivalent to taking $y \in N$ such that $by = ax_1$, since N is torsion-free, so we have cancellation. Let $N_1 = \mathrm{Sat}(N)$. Then, by definition, $A \cdot x_1 \subset N_1$.

Exercise 4.2. Use the fact that A is a PID to show that N_1 is of the form $N_1 = A \cdot y_1$ for some y_1. Intuitively, the goal is to find the $y \in N_1$ with the largest denominator (which can be made rigorous by using the sense of divisibility), so that a can be taken to be 1.

Now, N_1 is a free A-module of rank 1, and it is saturated. Thus, because y_1 has maximal denominator, every multiple of it is an integral multiple, not fractional.

[17] With this wording, this part of the proof is necessary and sufficient to imply Corollary 3.3.

[18] The concept of "in lowest terms" makes sense whenever the greatest common divisor of two elements is defined, and will mean that $a, b \in A$, $b \neq 0$, and $\gcd(a, b) = 1$.

Exercise 4.3. Using this and the definition of saturation, show that $N/(A \cdot y_1)$ is torsion-free.

In any case, $N/(A \cdot y_1)$ is generated by $\{x_2, \dots, x_r\}$, so by the inductive hypothesis, $N/(A \cdot y_1)$ is free with some basis $e_1, \dots, e_s$ of A^s. Then, this basis can be lifted to $\tilde{e}_1, \dots, \tilde{e}_s \in V$, which means that N is a free A-module with basis $y_1, \tilde{e}_1, \dots, \tilde{e}_s$. In summary, get the first element of the basis, and then apply induction.

Now that M_{tf} is known to be free, consider the short exact sequence $0 \to M_{\text{tor}} \to M \to M_{\text{tf}} \to 0$, and lift the basis $\{x_1, \dots, x_r\}$ of M_{tf} to $\{\tilde{x}_1, \dots, \tilde{x}_r\}$, a basis for the free part of M. Thus, $M = M_{\text{tor}} \oplus A \cdot \tilde{x}_1 \oplus \cdots \oplus A \cdot \tilde{x}_n$. However, this lift is not canonical; the quotient is, but choosing the realization involves a choice. For example, in $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}$, one could make the free part $(0,1)\mathbb{Z}$ or $(1,1)\mathbb{Z}$.

Now, the proof has been reduced to the case in which $M = M_{\text{tor}}$ is a finitely generated torsion module, so that M is killed by some nonzero $a \in A$ (which can be found as the product of things that are killed by the generators). This means that M is an $A/(a)$-module, since (a) doesn't really act on M other than killing it. Now, it's possible to use the Chinese Remainder Theorem: let $P = \{p_1, \dots, p_t\}$ be the prime factors of a, so that $a = up_1^{e_1} \cdots p_t^{e_t}$ for some unit $u \in A$. Let $I_i = (p_i^{e_i})$, so that I_i and I_j are coprime when $i \neq j$, implying that $I_i + I_j = A$ (otherwise, $I_i + I_j = (c)$ for some $c \mid p_i^{e_i}$ and $c \mid p_j^{e_j}$, so c must be a unit). By the Chinese Remainder Theorem, there exists an x_i such that $x_i \equiv 1 \bmod p_i$ and $x_i \equiv 0 \bmod p_j$ when $j \neq i$, for all $i = 1, \dots, t$.

Then, use these x_i as projectors to decompose M as $x_i \cdot M \oplus \cdots \oplus x_t \cdot M$. Each of these is clearly an A-submodule of M, and since the ideals are coprime, one can indeed write this direct sum, so $M \supseteq \bigoplus x_i \cdot M$. This is actually an equality, though, because every $m \in M$ is of the form $m = (x_1 + \cdots + x_t)m$, but $x_1 + \cdots + x_t \equiv 1 \bmod a$, but a acts by 0 in M, so $m = x_1 \cdot m + \cdots + x_t \cdot m$, so every element of M is a linear combination of these x_i.

Define $M_i = x_i \cdot m$. This is simpler than M itself; it's killed by $p_i^{e_i}$, since x_i is already a multiple of all of the $p_j^{e_j}$ for $j \neq i$, so it's a module killed by a prime power.

Now that the problem has been reduced even further, let M be a module killed by a prime power p^e for a prime p and $e \in \mathbb{N}$. Then, it will be possible to show that

$$M \cong A/(p^{\varepsilon_1}) \oplus \cdots \oplus A/(p^{\varepsilon_n}),$$

which implies Theorem 3.7. This can be done fairly easy with localization, but since that hasn't been taught yet, this slightly messier proof will be presented. It does provide an explicit algorithm for solving the problem, however.

It's known that M is the quotient space of some module: $A^n \twoheadrightarrow M$. Since M is killed by p^e, then consider $M \supset pM \supset p^2M \supset \cdots \supset p^eM = 0$, which is a filtration of M by submodules. Thus, one has the sequence

$$M/pM \xrightarrow{\cdot p} pM/p^2M \xrightarrow{\cdot p} p^2M/p^3M \xrightarrow{\cdot p} \cdots \xrightarrow{\cdot p} p^{e-1}M/p^eM,$$

where each map is multiplication by p. Thus, each map is surjective, and since (p) is a maximal ideal of A, then $A/(p)$ is a field, and each quotient is a finite-dimensional vector space over it. One can obtain a basis for the largest vector space as follows: pick a basis for the kernel of each map. Since the image of each map can be identified with the next $p^iM/p^{i+1}M$ in the sequence and each is a vector space (so that the kernel and the image make up the whole thing), one obtains a basis $e_1, e_2, \dots$ for M/pM, which surjects onto a basis for pM/p^2M, and so on.

Then, it is possible to lift this basis to a basis $\tilde{e}_1, \dots, \tilde{e}_n$ to M such that $p \cdot \tilde{e}_i = 0$. In general, an arbitrary lift is only zero mod p, so for the general case in which $p \cdot \hat{e}_i = p^2(*)$, take $\tilde{e}_i = \hat{e}_i - p*$, so that $p \cdot \tilde{e}_i = p(\hat{e}_i - p*) = 0$. Thus, such a lift is always possible.

Similarly, if $f_1, f_2, \dots$ is a basis of $\ker(pM/p^2M \xrightarrow{p} p^2M/p^3M)$, then lift them to $\tilde{f}_1, \tilde{f}_2, \dots$ such that $\tilde{f}_i = p\tilde{g}_i$ with the $\tilde{g}_i$ killed by p^2. Such $\tilde{g}_i$ can be found by the same reasoning as above.

Continuing, the goal is to lift stuff by a certain power of p such that it's killed by the next power. Thus, one obtains the necessary basis. Each $\tilde{e}_i$ generates a copy of $A/(p)$, each $\tilde{g}_i$ a copy of $A/(p^2)$, and so on, yielding the desired decomposition of M.

However, there's a little more work to do: a bit of care is needed to check that one has $\oplus$ rather than $+$, and so on. This is not difficult, but will require several lines of math.

Now we have existence; how about uniqueness? Well, it falls out of the existence part of the proof, because the number of copies of each $A/(p^k)$ depends on the size of the kernel of each map $p^{i-1}M/p^iM \to p^iM/p^{i+1}M$. This is invariant, so uniqueness follows. ⊠

Corollary 4.1. *If $\lambda_i = \dim_k p^iM/p^{i+1}M$, then $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{e-1}$ is a partition, and $\sum \lambda_i$ is the length of M.*

The length, a notion which will be developed further in the homework, is analogous to the dimension, except that M isn't a vector space, so strict dimension isn't available. In this scenario, the length serves as a reasonable analogue, because M has a graded filtration akin to a vector space, even though isn't one.

By Theorem 3.7, we obtain $M = A/(p^{\mu_0}) \oplus A/(p^{\mu_1}) \oplus \cdots$ such that $\mu_0 \geq \mu_1 \geq \cdots$. Then, the λ_i form the transpose partition to μ_i, such as if $\lambda = (5,3,1)$, then $\mu = (3,2,2,1,1)$. This is how the invariant factors are recovered from the invariant factors λ_i.

This relation is similar to the relation of Jordan blocks to the kernel of a nilpotent matrix, which is readable from the Jordan block sizes. This is no coincidence.

Applications to Linear Algebra. Let k be a field and V be a finite-dimensional k-vector space; let T be a k-linear transformation. Recall that the Jordan canonical form operates if k is algebraically closed, so that all eigenvalues exist in k. Then, it says that there exists a basis of V under which T has Jordan block form:

$$T = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_s \end{pmatrix}, \text{ where } B_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix},$$

up to the rearrangement of the Jordan blocks $B_1, \dots, B_s$.[19]

A stronger, more general version of this result holds for all fields k. Let V and $T : V \to V$ be as before. Then, define a $k[x]$-module structure on V by declaring how x acts on T; then, $f(x) = \sum a_i x^i$ acts on V by $\sum a_i T^i$. Since V is finite-dimensional, then it's finitely generated as a $k[x]$-module, and since $k[x]$ is a PID, then by Theorem 3.7,

$$V \simeq \bigoplus_{p \text{ monic, irreducible}} k[x]/(p^{e_1(p)} \oplus \cdots \oplus p^{e_{n_p}(p)}),$$

since we know the prime elements in $k[x]$. But T acts on V as multiplication by x acts on all of the factors, and this gives a nice basis for V. First, look at $k[x]/(p^e)$, which should be T-stable. Then, let $d = \deg(p)$, so this quotient space has a basis

$$\{1, x, x^2, \dots, x^{d-1}, p, xp, x^2p, \dots, x^{d-1}p, p^2, xp^2, \dots, p^{e-1}, xp^{e-1}, \dots, x^{d-1}p^{e-1}\},$$

so $\dim(k[x]/(p^e)) = de$, which is why that's a basis. It looks complicated, but multiplication by x, or equivalently action by T, has a fairly simple form: let

$$A = \begin{pmatrix} 0 & & & -a_0 \\ 1 & 0 & & -a_1 \\ & 1 & \ddots & \vdots \\ & & \ddots & 0 & -a_{d-2} \\ & & & 1 & -a_{d-1} \end{pmatrix} \text{ and } B = \begin{pmatrix} & 1 \end{pmatrix},$$

where the a_i are the coefficients of p; then T has block form

$$T \sim \begin{pmatrix} A & & & & \\ B & A & & & \\ & B & A & & \\ & & \ddots & \ddots & \\ & & & B & A \end{pmatrix}.$$

There are e copies of each of B and A in this matrix, so it is a $de \times de$ matrix.

From this point of view, one can read off many invariants of a linear transformation, such as the characteristic polynomial, which is just the product of the invariant factors of the $k[x]$-module V.

Part 2. Linear Algebra

5. Bilinear and Quadratic Forms: 10/4/13

Definition. Let A be a ring and M be a free A-module of rank n. Then:

- A *bilinear form* is a function $\beta : M \times M \to A$ that is A-linear in each argument.
- Such a bilinear form β is *symmetric* if $\beta(x, y) = \beta(y, x)$ for all $x, y \in M$, and is *skew-symmetric* if $\beta(x, y) = -\beta(y, x)$ for all $x, y \in M$.
- A bilinear form is *alternating* (also *symplectic*) if $\beta(x, x) = 0$ for all $x \in M$.

Generally, skew-symmetric and alternating forms are the same, except sometimes in characteristic 2. Every alternating form is skew-symmetric, because

$$\beta(x, y) + \beta(y, x) = \beta(x + y, x + y) - \beta(x, x) - \beta(y, y) = 0.$$

[19] This is also useful to demonstrate the non-obvious fact that a matrix A is similar to its transpose ${}^tA$.

Definition. A quadratic form is a map $q : M \to A$ satisfying:

(1) $q(a \cdot m) = a^2 \cdot q(m)$ (i.e. q is homogeneous of degree 2), and
(2) $(x, y) \mapsto q(x + y) - q(x) - q(y)$ is a bilinear form.

Obviously, the bilinear form above is symmetric (if this is not obvious, switch x and y and see what happens). In fact, given some symmetric bilinear form β, one can obtain a quadratic form $q(x) = \beta(x, x)$. But these are not inverse operators: $\beta \mapsto q \mapsto 2\beta$, and similarly $q \mapsto \beta \mapsto 2q$.

Corollary 5.1. *If 2 is invertible in A, then there is a bijection between symmetric bilinear forms and quadratic forms given by either of the above arrows. There is also a bijection between skew-symmetric quadratic forms and alternating forms.*

Again, they aren't each others' inverses, but multiplying one of them by $1/2$ solves that problem.

What happens if $\operatorname{char}(k) = 2$? Let $A = k$ be a field such that $\operatorname{char}(k) = 2$. In this case, symmetric and skew-symmetric bilinear forms are the same thing, because $-1 = 1$. Quadratic forms and alternating forms give these forms, but the implication in the other direction given in Corollary 5.1 doesn't apply here.

Example 5.1. If $q(x) = \sum_{i=1}^{n} x_i^2$, then $q(x+y) - q(x) - q(y) = 0$, because of the Frobenius automorphism. However, if $q(x) = x_1 x_2$, then $q(x + y) - q(x) - q(y) = x_1 y_2 + x_2 y_1$. This is nondegenerate.

What you get depends on what kind of quadratic form you start with; square terms drop out, but mixed ones live.

Definition. If (M, β) is a free A-module equipped with a bilinear form, β is called nondegenerate if the map $\beta^\sharp : M \to M^\vee = \operatorname{Hom}_A(M, A)$[20] given by $x \mapsto (y \mapsto \beta(x, y))$[21] is an isomorphism of A-modules. In this case, (M, β) is called an inner product space.

$\beta^\sharp$ is A-linear already, so it only remains to show in any given case that it is an isomorphism. If $A = k$ is a field, then there's an additional notion of nondegeneracy: that there is no $x \in M \setminus 0$ such that $\beta(x, y) = 0$ for all $y \in M$ (i.e. nothing is perpendicular to everything, except 0), or that $\beta^\sharp$ is injective. In general, this isn't equivalent to it being bijective, such as if $M = \mathbb{Z}$. Then, if β is given by $\beta(1, 1) = 2$, then β is injective but not surjective, and the stronger definition is needed.

From here on in the lecture, consider commutative rings A such that $2 \in A^\times$, so that symmetric bilinear forms and quadratic forms can be identified. Thus, one obtains the following invariant:

Definition. If $e_1, \ldots, e_n$ is an A-basis of M, then the determinant of a nondegenerate symmetric bilinear form is given as follows: define the matrix $G_{\beta,\{e_i\}} = (\beta(e_i, e_j))_{i,j}$ (that is, the $(i, j)^{\text{th}}$ entry is $\beta(e_i, e_j)$). This is a symmetric matrix, and in some other basis $\{e_i'\}$, it has the form $G_{\beta,\{e_i'\}} = S G_{\beta,\{e_i\}} S^{\mathrm{T}}$. Then, the determinant of β is defined as $\det \beta = \det G_{\beta,\{e_i\}}$, so that it is well-defined up to squares.

The way that the well-definedness is typically written is $\det \beta \in A^\times / (A^\times)^2$. Though the matrix G can be defined for any bilinear form, β is nondegenerate iff $G_{\beta,\{e_i\}}$ is invertible iff $\det(G)$ is invertible.

Example 5.2. If $A = \mathbb{R}$, then $\mathbb{R}^\times / (\mathbb{R}^\times)^2 \cong \{\pm 1\}$, so each quadratic form over $\mathbb{R}$ is sent to ± 1, which corresponds to positive or negative definiteness.

From earlier in linear algebra, recall the following result:

Theorem 5.2 (Classification of Nondegenerate Real-Valued Quadratic Forms)**.** *A real-valued quadratic form Q is equivalent to one of the form $x_1^2 + x_2^2 + \cdots + x_p^2 - x_{p+1}^2 - x_{p+1}^2 - \cdots - x_{p+1}^2$.*

The pair (p, q) is called the signature of Q, and the corresponding invariant is $\det(Q) = (-1)^q$.[22]

Example 5.3. If $A = k = \mathbb{F}_q$ for q odd, then $\det(Q) \in \mathbb{F}_q^\times / (\mathbb{F}_q^\times)^2 \cong \mathbb{Z}/2$, which can be seen in the following diagram, which is just the correspondence between multiplicative and additive notation:

$$\begin{array}{ccc} \mathbb{F}_q^\times & \xrightarrow{\sim} & \mathbb{Z}/(q-1) \\ \downarrow{\scriptstyle x \mapsto x^2} & & \downarrow{\scriptstyle \times 2} \\ \mathbb{F}_q^\times & \xrightarrow{\sim} & \mathbb{Z}/(q-1) \end{array}$$

Since $q - 1$ is known to be even, then $\ker(\times 2) = \{0, (q-1)/2\}$, and thus there are exactly two elements in the cokernel. This means that exactly have of $\mathbb{F}_q^\times$ are squares, which is already known in the context of quadratic resides. Here, the determinant takes values in $\mathbb{Z}/2$ again.

[20] This notation was introduced in the homework, where it was shown that since M is free, then $M^\vee$ is as well.

[21] Sometimes written $x \mapsto \beta(x, -)$.

[22] Apparently, this stuff is useful in Morse theory, understanding negative eigenvalues geometrically.

Theorem 5.3. *This is the only invariant for quadratic forms for $A = \mathbb{F}_q$, where q is odd: two nondegenerate forms Q and Q' on $\mathbb{F}_q^n$ are equivalent iff $\det(Q) = \det(Q')$ in $\mathbb{F}_q^\times/(\mathbb{F}_q^\times)^2$.*

This means that there are exactly two equivalence classes of nondegenerate quadratic forms on $\mathbb{F}_q^n$: one given by the identity matrix, $\sum x_i^2$, and one given by the matrix with a diagonal of $(\varepsilon, 1, \dots, 1)$, where ε isn't a square. This induces the form $\varepsilon x_1^2 + \sum_2^n x_i^2$.

Proof of Theorem 5.3. Clearly, the two cited forms aren't equivalent, because they have different determinants.

First, one an always bring a matrix for a quadratic form Q into diagonal form, because there exists an $e_1 \in V$ such that $\langle e_1, e_1\rangle \neq 0$, since Q is nondegenerate. Then, take the orthogonal decomposition $V = k \cdot e_1 \oplus (k \cdot e_1)^\perp$, and repeat, because $Q|_{(k\cdot e_1)^\perp}$ is still nondegenerate.

Now, there's an orthogonal basis for V, which is the bit that requires $\operatorname{char}(k) \neq 2$: $\langle e_i, e_j\rangle = 0$ if $i \neq j$, and $\langle e_i, e_i\rangle = u_i \neq 0$ for some unit u_i. In this basis, the matrix for Q has the form $\begin{pmatrix} u_1 & & \\ & \ddots & \\ & & u_n \end{pmatrix}$; multiplying e_i by some z changes u_i by z^2, but there are only two equivalence classes modulo squares, so the matrix can be put in the form

$$\begin{pmatrix} \varepsilon & & & & & & \\ & \varepsilon & & & & & \\ & & \ddots & & & & \\ & & & \varepsilon & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

for some $\varepsilon \notin (\mathbb{F}_q^\times)^2$ (the same one can be used, since all of the non-residues are equivalent).

Now, if there's more than one ε, they can also be reduced: two εs on the diagonal will go to two 1s. Consider the form $\begin{pmatrix} \varepsilon & \\ & \varepsilon \end{pmatrix} \sim \varepsilon x^2 + \varepsilon y^2$. If we can show this is equivalent to $x^2 + y^2$, then the whole theorem is proven, so the goal is to solve $\varepsilon x^2 + \varepsilon y^2 = 1$ in $\mathbb{F}_q$. This can actually be done naïvely: rearrange to $x^2 = \varepsilon^{-1} - y^2$, and count the possibilities. On the right-hand side, y can be zero or any quadratic residue, giving $(q+1)/2$ possibilities for the right-hand side. For x^2, there are also $(q+1)/2$ possible values, since x can also be zero, so by the pigeonhole principle, the two must meet somewhere, so a solution exists.

Thus, there is an orthonormal basis in which $\begin{pmatrix} \varepsilon & \\ & \varepsilon \end{pmatrix} \sim \begin{pmatrix} 1 & \\ & * \end{pmatrix}$. Since the determinants must be the same, then $*$ must be a square, so it might as well be 1, and the rest follows. ⊠

Notice that in the case where $k = \mathbb{R}$, one could take $\varepsilon = -1$, but $x^2 + y^2 = -1$ has no solutions, so there invariant takes a different form.

Definition. If (M, q) is a nondegenerate quadratic form, then it is split if there exists a submodule $N \subset M$ such that $\operatorname{rank} N = (\operatorname{rank} M)/2$ and $q|_N = 0$.

Clearly, this only makes sense if M is even-dimensional.

Example 5.4. Let $A = \mathbb{C}$ and $q = x^2 + y^2$. A one-dimensional subspace where it vanishes always exists, so every quadratic form is split. But since $\mathbb{C}^\times = (\mathbb{C}^\times)^2$, all quadratic forms of a given rank are equivalent.[23]

Proposition 5.4. *A quadratic form over $\mathbb{R}$ is split iff its signature is of the form $(n/2, n/2)$.*

Proof. If M is split, then $N = N \oplus L$ for some other L. Note that $N^\perp \neq L$, because $q|_N = 0$, so in fact $N = N^\perp$! Thus, this isn't an orthogonal decomposition, so it has a nontrivial inner product. With a suitable choice of L, $q|_L = 0$, so there is a basis $(e_1, \dots, e_n)$ of N and a dual basis $(f_1, \dots, f_n)$ of L such that the matrix has form $[f_n, f_{n-1}, \dots, f_1, e_n, \dots, e_1]$, creating a pairing $e_1 \leftrightarrow f_n$, $e_n \leftrightarrow f_{n-1}$, etc. This looks like a symplectic basis, except that it is symmetric. Then, each pair can be transformed into the desired signature, because $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \sim \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$: both of them are indefinite as quadratic forms, and every real qusdratic form is either positive definite, negative definite, or indefinite. When this is repeated with each pairing, the desired signature is obtained. ⊠

[23] This is also true in every algebraically closed field or even any field closed under taking square roots.

There are various ways to product split quadratic forms: for example, over any ring, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ works, since it forces the first basis vector to be orthogonal to itself. More generally, a matrix given in block form as $\begin{pmatrix} 0 & B \\ C & * \end{pmatrix}$, where B and C are invertible (and therefore inverse transposes to each other), is split. However, a split form can be made out of any quadratic form:

Claim. If (M, q) is a quadratic form, then $(M, q) \oplus (M, -q)$ is split.

Proof. Let $\Delta(M)$ denote the diagonal of M in $(M, q) \oplus (M, -q)$, which is a half-dimensional space on which the definition works right, showing that it's split. ⊠

Definition. If A is a commutative ring, the Witt group $W(A)$ is the set of nondegenerate quadratic forms on free A-modules modded out by an equivalence relation $\dot{\sim}$, called weak equivalence, given by $(M, q)\dot{\sim}(M', q')$ if there exist split forms (S_1, q_1) and (S_2, q_2) such that $(M, q) \oplus (S_1, q_1) = (M', q') \oplus (S_2, q_2)$.

For example, under this definition, every split form is weakly equivalent to the zero form, so one can't talk about dimension in this equivalence, only parity of dimension.

Exercise 5.1. Show that this set of equivalence classes is a group under direct sum.

The trick is to demonstrate that inverses exist: the rest of the operations are reasonably clear. Using $(M, -q) = (M, q)^{-1}$, it is possible to concoct a group structure, rather than just a monoid.

These seem complicated, but it happens that a ring homomorphism $A \to B$ induces a homomorphism of Witt groups by tensoring them, so it's possible to make interesting statements such as the following.

Theorem 5.5. $W(\mathbb{Z}) \xrightarrow{\sim} W(\mathbb{R}) \xrightarrow{\sim} \mathbb{Z}$.

Though the proof won't be presented in this class, the last step isn't too hard, and is given by the signature: if (M, Q) has signature (p, q), send it to the integer $q - p$ (or, equivalently, $p - q$), because the split form $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ goes to zero in the Witt group, so it can be cancelled, and only the difference in the signature terms is meaningful. The result that $W(\mathbb{Z}) \cong W(\mathbb{R})$ is more surprising: two quadratic forms over $\mathbb{Z}$ are weakly equivalent iff they have the same signature. For a more detailed discussion and a proof, consult [2].

Part 3. Category Theory

6. Categories, Functors, Limits, and Adjoints: 10/9/13

"Before functoriality, people lived in caves." – Brian Conrad

In order to understand localization of rings, it will be necessary to study some category theory.

The basic idea is that a category $\mathcal{C}$ is a pair of objects $\mathrm{Obj}\,\mathcal{C}$[24] and a set of morphisms for any two objects in $\mathrm{Obj}\,\mathcal{C}$, given by $\mathrm{Mor} : \mathrm{Obj}\,\mathcal{C} \times \mathrm{Obj}\,\mathcal{C} \to \underline{\mathrm{Set}}$. In some sense, $(X, Y) \to \mathrm{Mor}_{\mathcal{C}}(X, Y)$. These morphisms are required to obey composition, and to have an identity.

Definition. More formally, a category $\mathcal{C}$ is a collection $\mathrm{Obj}\,\mathcal{C}$ of objects and a set of morphisms $\mathrm{Mor}_{\mathcal{C}}(X, Y)$ for every $X, Y \in \mathrm{Obj}\,\mathcal{C}$ with a composition operation $\circ : \mathrm{Mor}_{\mathcal{C}}(X, Y) \times \mathrm{Mor}_{\mathcal{C}}(Y, Z) \to \mathrm{Mor}_{\mathcal{C}}(X, Z)$ such that:

- Composition is associative: if $f \in \mathrm{Mor}_{\mathcal{C}}(X, Y)$, $g \in \mathrm{Mor}_{\mathcal{C}}(Y, Z)$, and $h \in \mathrm{Mor}_{\mathcal{C}}(Z, W)$, then $f \circ (g \circ h) = (f \circ g) \circ h$.
- For every $X \in \mathrm{Obj}\,\mathcal{C}$, there exists an identity morphism $\mathrm{id}_X \in \mathrm{Mor}_{\mathcal{C}}(X, X)$ such that for any $Y \in \mathrm{Obj}\,\mathcal{C}$ and $f \in \mathrm{Mor}_{\mathcal{C}}(X, Y)$, $\mathrm{id}_X \circ f = f$ (and similarly, $f \circ 1_Y = f$).

One could view a set as just a collection of points, a zero-dimensional object, in which case a category is akin to a one-dimensional object, where the points (the objects) are connected by arrows (the morphisms).

Definition. A covariant functor is a map between categories $F : \mathcal{C} \to \mathcal{D}$ such that $F : \mathrm{Obj}\,\mathcal{C} \to \mathrm{Obj}\,\mathcal{D}$, and $f \in \mathrm{Mor}_{\mathcal{C}}(X, Y) \mapsto F(f) \in \mathrm{Mor}_{\mathcal{D}}(F(X), F(Y))$; additionally, F is required to preserve the identity morphism and composition.

Equivalently, one could specify $F_{X,Y} : \mathrm{Mor}_{\mathcal{C}}(X, Y) \to \mathrm{Mor}_{\mathcal{D}}(F(X), F(Y))$ such that identity and composition are preserved.

[24] This is not necessarily a set; for example, in the category of sets, the set of all sets would be a bad thing to have and has to be worked around.

Definition. A natural transformation between functors F and G, written $F \overset{\alpha}{\Longrightarrow} G$, is a collection of $\mathcal{D}$-morphisms $\alpha_X : F(X) \to G(X)$ such that for any $X, Y \in \operatorname{Obj}\mathcal{C}$ and $f \in \operatorname{Mor}_{\mathcal{C}}(X, Y)$, the following diagram commutes:

$$\begin{array}{ccc} F(X) & \xrightarrow{\alpha_X} & G(X) \\ \downarrow{\scriptstyle F(f)} & & \downarrow{\scriptstyle G(f)} \\ F(Y) & \xrightarrow{\alpha_Y} & G(Y) \end{array}$$

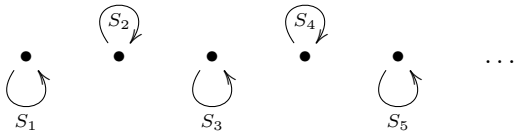
In some sense, the data that one needs to supply is a map from each $F(X) \to G(X)$, with the commutativity condition.

If $\mathcal{C}$ and $\mathcal{D}$ are categories, then $\operatorname{Fun}(\mathcal{C}, \mathcal{D})$ denotes the category of functors from $\mathcal{C}$ to $\mathcal{D}$: the objects are functors $F : \mathcal{C} \to \mathcal{D}$, and the morphisms are natural transformations $F \overset{\alpha}{\Longrightarrow} G$,

Natural transformations can be thought of as homotopies between maps: $\mathcal{C}$ and $\mathcal{D}$ look like graphs, and each map $f \in \operatorname{Mor}_{\mathcal{C}}(X, Y)$ can be thought of as a path in $\mathcal{C}$. Then, α_X and α_Y are paths from F to G, and the diagram being commutative is akin to these paths being homotopic.

In addition to covariant functors, one also has contravariant functors, which are exactly the same, except that in this case $F_{X,Y} : \operatorname{Mor}_{\mathcal{C}}(X, Y) \to \operatorname{Mor}_{\mathcal{D}}(F(Y), F(X))$. A contravariant functor is equivalent to a covariant functor $\mathcal{C}^{\mathrm{op}} \to \mathcal{D}$, where $\mathcal{C}^{\mathrm{op}}$ is the opposite category: defined with the same objects as $\mathcal{C}$ and $\operatorname{Mor}_{\mathcal{C}^{\mathrm{op}}}(X, Y) = \operatorname{Mor}_{\mathcal{C}}(Y, X)$. This is another potentially useful way of talking about things.

Example 6.1. The category $\underline{\mathrm{Set}}$ is created by taking the objects to be sets and the morphisms are any functions. One also has the category $\underline{\mathrm{Set}}^{\mathrm{bij}}$, where the objects are sets and the morphisms bijections between sets. This category looks like this:



where the first dot corresponds to $\{1\}$, the second to $\{1, 2\}$, and so on. Each set A is acted on by the permutation group S_A.

Lots of algebraic structures form categories, such as $\underline{\mathrm{Groups}}$, $\underline{\mathrm{Rings}}$, $\underline{k\text{-Vect}}$, where k is a field, and so on. These are all enrichments of the category of sets. Another enrichment that isn't really an algebraic structure is $\underline{\mathrm{Top}}$, the category of topological spaces. These enriched sets mean there is a forgetful functor $\underline{\mathrm{Top}} \to \underline{\mathrm{Set}}$ (and similarly for the other enriched categories) that takes a topological space and returns the underlying set. Some more interesting functors on these categories include $\underline{\mathrm{Set}} \to \underline{k\text{-Vect}}$ given by $S \mapsto \bigoplus_{s \in S} k$ or $S \mapsto k^S$.

Let $\mathcal{C}$ be an arbitrary category and $X \in \operatorname{Obj}\mathcal{C}$. Then, one has a functor $h_X : \mathcal{C} \to \underline{\mathrm{Set}}$ that sends $Y \mapsto \operatorname{Mor}_{\mathcal{C}}(X, Y)$. Letting X vary, one obtains a functor $h : \mathcal{C} \to \operatorname{Fun}(\mathcal{C}, \underline{\mathrm{Set}})$ given by $X \mapsto h_X$. This is a contravariant functor. An equivalent restatement is that built into the structure of any category is a pairing $\mathcal{C}^{\mathrm{op}} \times \mathcal{C} \to \underline{\mathrm{Set}}$ given by $(X, Y) \mapsto \operatorname{Mor}_{\mathcal{C}}(X, Y)$.

Definition. A functor $F : \mathcal{C} \to \mathcal{D}$ is faithful if $F_{X,Y} : \operatorname{Mor}_{\mathcal{C}}(X, Y) \to \operatorname{Mor}_{\mathcal{D}}(F(X), F(Y))$ is always injective. It is called fully faithful if all such maps are bijective.

Lemma 6.1 (Yoneda)**.** *The functor h given above is fully faithful.*

When a functor $F : \mathcal{C} \to \mathcal{D}$ is fully faithful, it is possible to view $\mathcal{C}$ as a subategory of $\mathcal{D}$: a fully faithful functor gives an embedding of categories. The actual number of objects is pretty unimportant, though, as in the following example:

Example 6.2. Let $\mathcal{C} = \underline{\mathrm{FinSet}}$, the category of finite sets, and $\mathcal{D} = \mathbb{Z}_{\geq 0}$ with morphisms $n \mapsto m$ given by maps $\{1, \ldots, n\} \to \{1, \ldots, m\}$. Then, consider the functor $S \mapsto \#S$. Since all of the objects of $\mathcal{C}$ with the same cardinality are isomorphic in $\mathcal{D}$, so this is a fully faithful functor.

Definition. An object $I \in \operatorname{Obj}\mathcal{C}$[25] is called initial if for every $X \in \operatorname{Obj}\mathcal{C}$, $\operatorname{Mor}_{\mathcal{C}}(I, X)$ is a singleton set. Similarly, $F \in \operatorname{Obj}\mathcal{C}$ is called final if $|\operatorname{Mor}_{\mathcal{C}}(X, F)| = 1$ for every $X \in \operatorname{Obj}\mathcal{C}$.

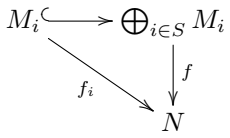
Example 6.3. In the category of sets, the empty set is minimal, because by convention there is one function $\emptyset \to S$ for each S. Every singleton set $\{e\}$ is final, because any set S maps to it in a unique way: $s \mapsto e$ for all $s \in S$.

In $\underline{A\text{-Mod}}$, 0 is both initial and final, since A-linear maps have more rigid requirements.

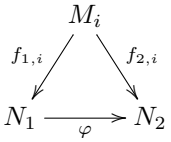
[25]By this point, one might start seeing $I \in \mathcal{C}$ to denote I as an object. This is reasonable notation, but not completely precise.

Sometimes, initial or final objects don't exist: let $\mathcal{C}$ be the category of nonempty sets, so that $\mathcal{C}$ has no initial object and $\mathcal{C}^{\text{op}}$ has no final object (since initial objects in $\mathcal{C}$ correspond to final objects in $\mathcal{C}^{\text{op}}$, and vice versa). Similarly, the category of nontrivial A-modules has neither initial nor final objects (in general, one can always throw stuff out of a category and it still works, though whether this is actually useful is a different question).

The concepts of initial objects lead to a very important motivation of category theory: *a universal property tends to be equivalent to saying that something is initial or final in some category.* For example, given a set S and A-modules $\{M_i\}_{i\in S}$, then the direct sum $\bigoplus_{i\in S} M_i$ is characterized by the following universal property: for any A-module N together with A-linear maps $f_i : M_i \to N$ (one for each i), then there exists a unique A-linear map $f : \bigoplus_{i\in S} M_i \to N$ such that the following diagram commutes for each $i \in S$:



This can be reformulated: let $\mathcal{C}$ be the category whose objects are data $(N, M_i \xrightarrow{f_i} N)$, where N is an A-module and f_i is as above, with the "obvious" morphisms, i.e the A-linear maps $N_1 \to N_2$ that are compatible with the f_i (i.e. φ such that the diagram



(2)

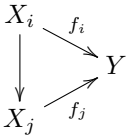
commutes). Then, the universal property is equivalent to $\bigoplus_{i\in S}$ being the initial object in $\mathcal{C}$.[26] This is convenient, because initial and final objects are unique up to isomorphism, which is why one occasionally hears of "the" initial or final object. In the same vein, the direct product is the final object in the category $\mathcal{D}$ with objects given by A-modules N along with A-linear homomorphisms $N \xrightarrow{g_i} M_i$, with the analogous morphisms.[27]

The direct sum and direct product are special cases of direct and inverse limits, respectively, in some sense.

Definition. A directed set S is a set equipped with some arrows $i \to j$ for some $i, j \in S$, subject to the condition that for any $i, j \in S$, there exists a $k \in S$ such that k is reachable from both i and j.

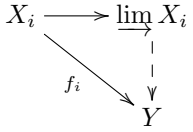
There's more than one way of thinking about S; it could be a certain type of directed graph, a poset, or even a category if $i \to i$ is added everywhere. Directed sets can be used to construct limits: for example, if $S = \mathbb{N}$ and $i \to j$ iff $i \le j$, then one will obtain the usual notion of the limit of a sequence.

Definition. Fix a category $\mathcal{C}$. For a given directed set S, associate an object X_i to each $i \in S$ and a morphism $X_i \to X_j$ for each arrow $i \to j$.[28] Build another category $\mathcal{C}_{X_i}$ whose objects are given by $Y \in \mathcal{C}$ along with maps $X_i \xrightarrow{f_i} Y$ such that



commutes, and the morphisms are those between the objects that are compatible with the f_i in the same sense as (2). If an initial object exists in $\mathcal{C}_{X_i}$, it is called the colimit (sometimes also the direct limit, or the inductive limit) of $\{X_i\}_{i\in S}$, denoted $\varinjlim_{i\in S} X_i$.

The equivalent formulation in terms of universal properties is that for every $i \in S$, there exists a map $X_i \to \varinjlim X_j$ such that whenever $f_i : X \to Y$, there is a unique way to fill in the third arrow as below:



[26] Technically, it has to be realized as an object of $\mathcal{C}$ for this to be true, rather than just in $\underline{A}$-Mod, but in this case the f_i are the canonical embeddings $M_i \hookrightarrow M$.

[27] Notice that the final object in $\mathcal{C}$ above is 0 again, though this time with the zero maps.

[28] If S is taken to be a category, this association is exactly a covariant functor $S \to \mathcal{C}$.

The direct sum $\bigoplus_{i\in S} X_i$ can be thought of as a special case of this construction: given the ordinary set S, form a directed set $\widehat{S}$ by adding an extra point t to S and then adding arrows $s \to t$ for every $s \in S$.[29] Then, $\bigoplus_{i\in S} X_i = \varinjlim_{i\in\widehat{S}} X_i$. In some sense, this is not an exact characterization according to the definition, but it's a helpful illustration of the concept nonetheless. S is an undirected set, and it's helpful to view the direct sum as the direct limit over the undirected set S. Everything in the definition still works, though.

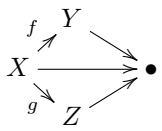
The direct product can be constructed in a similar way: instead of a directed set, take a set that looks like

$$S = \left(\bullet \nearrow \bullet \quad \bullet \searrow \bullet \right).$$

Then, the direct limit becomes

$$\varinjlim \left(\begin{array}{ccc} & \overset{f}{\nearrow} & Y \\ X & & \\ & \underset{g}{\searrow} & Z \end{array} \right) = (Y \times Z)/\operatorname{Im}(x \mapsto (f(x), -g(x))),$$

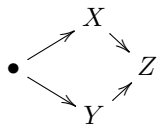
or the direct product modulo this "antidiagonal." If there is an object $\bullet$ such that



commutes, then the antidiagonal vanishes, and the direct product is in fact a direct limit. This typically arises in the category $\underline{A\text{-Mod}}$, but it can be taken elsewhere too.

It's actually possible to take limits of any random diagram: for example, if one has $\bullet \overset{\varphi}{\circlearrowright}$ (which would correspond to an A-module X with an endomorphism φ), then its limit is the set of coinvariants of $\varphi \circlearrowleft X$, given by $X/\{x - \varphi(x) \mid x \in X\}$, which is the set that just identifies x and $\varphi(x)$.

If the definitions and work above are made in the opposite category, one obtains limits (also inverse limits), in which all of the arrows are reversed: the objects are Y with maps $Y \to X_i$, and the quest is for a final object. Thus, the direct product is a special case of a limit in the same sense that the direct sum is a special case of a colimit. Specifically, the limit of a diagram $X \xrightarrow{f} Z \leftarrow Y$ is some initial



that makes the diagram commute, giving $\{(x, y) \in X \oplus Y \mid f(x) = g(y)\}$, which is a submodule of $X \oplus Y$. This is called the fiber product of X and Y, written $X \times_Z Y$. It also exists in more general categories. Additionally, $\varprojlim (\varphi \circlearrowright X) = X^{\varphi}$, or the invariants under the endomorphism φ.

Definition. Let $\mathcal{C}$ and $\mathcal{D}$ be categories. Then, a left adjoint to a functor $F : \mathcal{C} \to \mathcal{D}$ is a functor $G : \mathcal{D} \to \mathcal{C}$ together with bijections $\alpha_{X,Y} : \operatorname{Mor}_{\mathcal{C}}(G(X), Y) \xrightarrow{\sim} \operatorname{Mor}_{\mathcal{D}}(X, F(Y))$ such that the $\alpha_{X,Y}$ are natural transformations. Then, F is called a right adjoint to G.

The notation $\mathcal{C} \underset{F}{\overset{G}{\rightleftarrows}} \mathcal{D}$ is occasionally used, though it is nonstandard. Also, to remember which of the adjoint functors is left and which is right, the left adjoint appears in the first (left) argument of the bijection, and the right adjoint appears in the second (right-hand) argument.

Example 6.4. Consider the forgetful functor $\underline{\text{Top}} \to \underline{\text{Set}}$, which takes a topological space and returns the underlying set. It has a left adjoint given by taking a set and returning that set with the discrete topology, since all maps are continuous in that topology. A right adjoint to this functor would be a bijection between morphisms $Y \to X$ as topological spaces and maps $Y \to X$ as sets. This can be done by assigning X the trivial topology $\{\emptyset, X\}$.

The forgetful functor $\underline{\text{Rings}} \to \underline{\text{Set}}$ has left adjoint $S \mapsto \mathbb{Z}\langle S \rangle$.

[29] This requires the Axiom of Choice in the general case.

Part 4. Localization

7. Introduction to Localization: 10/11/13

Before jumping into localization, a clearer definition of direct and inverse limits will be provided. Specifically, revise the notion of a directed set to be a partially ordered set $(I, \leq)$ such that for any $i, j \in I$ there exists a $k \in I$ such that $i \leq k$ and $j \leq k$. This definition does *not* allow for self-loops.

The advantage of this approach is that it allows for an explicit construction: for example, if $X_i \in \underline{A\text{-Mod}}$, one can construct

$$\varinjlim_{i \in I} X_i = \coprod_{i \in I} X_i / \{x_i \in X_i \sim x_j \in X_j \text{ if their image is the same in some } X_k \text{ for some } k \geq i, j\}.$$

One can check that this is an equivalence relation. This gives the quotient set an abelian group operation in which $x_i + x_j$ is defined by taking some $k \geq i, j$ and considering the maps $X_i, X_j \to X_k$. Then, the images $x_i \mapsto y$ and $x_j \mapsto z$ give $x_i + x_j = y + z$ in $\varinjlim X_i$. This is well-defined because the equivalence relation removes any dependence of this on the k chosen. Finally, the action of A is inherited from that on the X_i, making the limit an A-module.

The inverse limit $\varprojlim X_i \subset \prod X_i$ is characterized by sequences $\{(x_i)_{i \in I} \mid \text{under the map } X_i \to X_j \text{ for } i \leq j,\ x_i \mapsto x_j\}$. For example, if $I = \mathbb{N}$, then define $i \to j$ iff $i \geq j$, giving a system $\cdots \to X_3 \to X_2 \to X_1$. The inverse limit is just the set of sequences such that one element maps to the next under these maps.

Localization of Rings. For the rest of this lecture, A will be a commutative ring with unit.

Example 7.1. If A is a domain, the field of total fractions $K = \mathrm{Frac}(A) = \{a/b \mid a, b \in A, b \neq 0\}$ modulo the relation of lowest terms is a familiar object, e.g. $\mathbb{Z} \mapsto \mathbb{Q}$, $k[x] \mapsto k(x)$. This is a useful construction.

Example 7.2. For a more analytic construction, let A be the ring of holomorphic functions on some domain $D \subset \mathbb{C}$ (this ring is given by pointwise multiplication, and is thus commutative). One often talks about the germ of a function $f \in A$ at some $x \in D$. This is an equivalence class of functions that agree on some neighborhood U of x.[30] Since these functions are analytic, this is equivalent to taking the Taylor expansion at x.

There are actually several rings going on here: there's $A = \mathrm{Hol}(D)$, and the ring of germs at x, denoted A_x or sometimes $\mathcal{O}_x$ or $\mathcal{O}_x^{\mathrm{hol}}$. There's a map $A \to A_x$ giving the equivalence class of a function.

Then, there is also the notion of a Taylor expansion: if $t = z - x$, then one considers the ring of formal power series $\mathbb{C}[[t]] = \{\sum_{n \geq 0} c_n (z - x)^n\}$. Since these are formal power series, one ignores any questions of convergence.

The words I just said can be collected into this diagram:

$$\begin{array}{ccc} A & \longrightarrow & \mathcal{O}_x^{\mathrm{hol}} \\ & \searrow^{\varphi} & \downarrow \\ & & \mathbb{C}[[t]] \end{array}$$

where φ takes the Taylor expansion of a function, which does end up being a ring homomorphism. The induced homomorphism $\mathcal{O}_x^{\mathrm{hol}} \to \mathbb{C}[[t]]$ is also a ring homomorphism, since this is just the ring of power series with a positive radius of convergence. For D sufficiently well-behaved, this is a localization, and A_x and $\mathbb{C}[[t]]$ are both local rings, examples of completion.

Example 7.3. Let $A = k[x_1, \ldots, x_n]$, i.e. the ring of algebraic functions on affine space, and let $\mathbf{0} = (0, \ldots, 0)$. Then, $A_{\mathbf{0}} = \{f(x_1, \ldots, x_n)/g(x_1, \ldots, x_n) \mid f, g \in A, g(x_1, \ldots, x_n) \neq 0\}$. If k is such that the notion of a neighborhood makes sense, then the functions in A are all defined within a neighborhood of zero. Once again, the notion of power series arises, since rational functions lead to them: $f/(1 - \blacksquare) = f(1 + \blacksquare + \blacksquare^2 + \cdots)$, giving a commutative diagram

$$\begin{array}{ccc} A = k[x_1, \ldots, x_n] & \hookrightarrow & A_{\mathbf{0}} \\ & \searrow & \downarrow \\ & & k[[x_1, \ldots, x_n]] \end{array}$$

Thus, $A_{\mathbf{0}}$ is the localization at the maximal ideal $\mathfrak{m} = (x_1, \ldots, x_n)$, and the map $A_{\mathbf{0}} \to k[[x_1, \ldots, x_n]]$ is called completion.

Definition. If A is a ring, then $S \subset A$ is multiplicative if $1 \in S$ and whenever $x, y \in S$, then $xy \in S$.

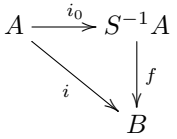
[30] Formally, the germ of a function f is the set of functions $g \in \mathrm{Hol(U)}$ such that $g|_V = f|_V$ for some open $V \subset U$. Then, the ring of germs is A modulo this equivalence relation.

We want to obtain a ring $S^{-1}A$ which is the smallest A-algebra in which (the images of) elements of S are invertible. It will be a modification of A by formally adding inverses of S to A. That is, define the category $\mathcal{C}_S = \{(B, i) \mid i : A \to B \text{ is a ring homomorphism}, i(S) \subset B^\times\}$. In words, this is the category of A-algebras in which S contains entirely invertible elements. Using categories, one can get a good definition of "smallest."

The definition of the ring $S^{-1}A$ will be embedded in the theorem asserting its existence:

Theorem 7.1. *$\mathcal{C}_S$ has an initial object, which is called the localization of A at S and is denoted $S^{-1}A$.*

More concretely, notice that there is a canonical homomorphism $A \xrightarrow{i_0} S^{-1}A$, since $S^{-1}A$ is an A-algebra, and if B is any other ring that satisfies this, then there is a unique f such that



commutes (which is just what it means to be initial).

Proof of Theorem 7.1. Define an equivalence relation on $S \times A$ (written as formal symbols $a/s = (s, a)$) in which $a/s \sim a'/s'$ if there exists some $t \in S$ such that $(as' - a's)t = 0$.[31]

Define $S^{-1}/A = (S \times A)/\sim$, with addition given by $a/s + a'/s' = (a's + as')/ss'$ and multiplication given by $(a/s)(a'/s') = (aa')/(ss')$. Then, the map $i_0 : A \to S^{-1}A$ sending $a \mapsto a/1$ (which is why we require $1 \in S$, though alternatively one could send $a \mapsto as/s$). These give $S^{-1}A$ the structure of an A-algebra.

This ring homomorphism extends i in the sense that $\psi \circ i_0 = i$: $\psi(i_0(a)) = \psi(a/1) = i(a)i(1)^{-1} = i(a)$ for any $a \in A$. Moreover, it is the unique homomorphism of A-algebras to do so: if $f : S^{-1}A \to B$ is an A-algebra homomorphism such that $f \circ i_0 = i$, then $f(a/s)f(s) = f(a)$, but $f(s) = i(s)$ since f factors through i. But since $i(s)$ is invertible, this forces $f(a/s) = i(a)i^{-1}(s)$, which means that $f = \psi$.

Thus, $(S^{-1}A, i_0)$ is initial in $\mathcal{C}_S$. ⊠

Example 7.4.

(1) If A is a domain, let $S = A \setminus 0$, so that $S^{-1}A = \mathrm{Frac}(A)$, as discussed in Example 7.1.
(2) In fact, if A is any commutative ring and S is the set of elements of A that aren't zero divisors, then S is multiplicative. Sometimes the notation $S^{-1}A = \mathrm{Frac}(A)$ is used, but this object is not in general a field, so it is instead called a ring of fractions. For example, $\mathbb{Z} \times \mathbb{Z} \rightsquigarrow \mathbb{Q} \times \mathbb{Q}$. This object is not a field, since it has zero divisors.
(3) For a more minimal example, let $S = A^\times$. From the universal property, $\mathcal{C}_S$ is just the category of A-algebras, since $i(S)$ is invertible in any A-algebra, and the initial object is A itself.
(4) For a less minimal example, let $S = \{f^n \mid n \geq 0\}$ for some $f \in A$. This localization is often denoted A_f; its elements look like a/f^n, and (as will appear in the homework) $A_f \cong A[x]/(fx - 1)$ as A-algebras.

Example 7.5. Here is an important example: let $\mathfrak{p}$ be a prime ideal, and $S = A \setminus \mathfrak{p}$. (This is a generalization of the field of fractions discussed above, which set $\mathfrak{p} = 0$.) This works because $A \setminus I$ is multiplicative iff I is prime, which follows from the definition. The resulting ring is often denoted $S^{-1}A = A_{\mathfrak{p}}$, the localization of A at $\mathfrak{p}$.

There's an "intuitive" picture of this when $A = \mathbb{C}[k]$: there are two kinds of maximal ideals. $\mathfrak{m}_z = (x - z)$ for any $z \in \mathbb{C}$, and (0). Thus, $\mathrm{Spec}\,\mathbb{C}[x]$ looks like the complex plane plus the generic point (0): though this is a small ideal, it should be thought of as a very "big" point, since $\mathrm{Spec}(A_{\mathfrak{m}_z}) = \{(0), \mathfrak{m}_z\}$, and all of the other maximal ideals vanish. This is a topological space concentrated at $\mathfrak{m}_z$, and illustrates what can be local about localization.[32]

Here are a couple words of warning about localization:

(1) Different sets S may yield the same localization $S^{-1}A$. For a simple example, throwing units into S doesn't change anything. More generally, let $\tilde{S} = \{a \in A \mid a \mid s \text{ for some } s \in S\}$ (that is, the set of elements a such that $s = ab$ for an $s \in S$ and $b \in A$). Then, $S^{-1}A \cong \tilde{S}^{-1}A$, and the isomorphism is canonical, because the categories $\mathcal{C}_S$ and $\mathcal{C}_{S'}$ are identical: in any A-algebra, if $i(s)$ is invertible, then its divisors ought to be as well. Thus, since the categories are the same, their initial objects are as well.
(2) Localization doesn't always enlarge rings. Sometimes, it's more like taking a quotient, so that it's not injective. For example, take $A = \mathbb{Z}/6$ and $S = \{1, 3, 5\}$, so that $S^{-1}A \cong \mathbb{Z}/2$. This is because anything odd must go to 1, and $2/1 \sim 6/3 \sim 0/1 = 0$. Oops! In fact, $\mathbb{Z}/6 \to S^{-1}A$ is the same map as the quotient, reducing mod 2.

[31] Aside: the intuitive equivalence relation would be $a/s \sim a'/s'$ iff $as' = a's$, but this isn't an equivalence relation in some cases where A isn't an integral domain, because then it wouldn't be transitive. For example, let $A = \mathbb{Z}/6$ and $S = \{1, 3, 5\}$, which is certainly multiplicative. Then, $2/1 \sim 6/3$, but $6/3 = 0/3 \sim 0/1$. Yet $2/1 \not\sim 0/1$, which is a problem.

[32] I didn't really get this example, nor the picture with a big fat point at the origin. Maybe I'm missing the point.

Example 7.6. If $A = \mathbb{Z}$ and $S = \mathbb{Z} \setminus 0$, then $S^{-1}A = \mathbb{Q}$. If S is the set of nonzero even numbers, then $S^{-1}A = \mathbb{Q}$ still, because all of $\mathbb{Z}$ are divisors of elements of S. Similarly, one could take $S = \{1, 100, 101, \dots\}$, so S is less important than the resulting set.

Localization of Modules. If A is a ring and S a multiplicative subset of it, there is a category of $S^{-1}A$-modules. Every $S^{-1}A$-module can be realized as an A-module by $F : N \to N$, the forgetful functor (also called restriction of scalars).

Theorem 7.2. *This functor F admits a left adjoint $M \mapsto S^{-1}M$.*

More carefully, if M is an arbitrary $S^{-1}A$-module, then any element of $\mathrm{Hom}_A(M, N)$ (or, equivalently, $\mathrm{Hom}_A(M, F(N))$, which is the same thing) can be automatically extended to an $S^{-1}A$-linear element of $\mathrm{Hom}(S^{-1}M, N)$. The quickest way to do this is to take $S^{-1}M = S^{-1}A \otimes_A M$, but that will come later.

Proof. Basically, you end up constructing the tensor product: let $S^{-1}M = (S \times M)/(m/s \sim m'/s'$ if $(sm' - s'm)t = 0$ for some $t \in S)$. ⊠

This construction can be applied to ideals $I \subset A$, giving localized ideals $S^{-1}I \to S^{-1}A$ (this map will be injective, but we can't show that just yet). One also has to check that $S^{-1}I$ is an ideal of $S^{-1}A$, but this basically follows from the definition. This construction also has a lot of nice properties: $S^{-1}(I + J) = S^{-1}I + S^{-1}J$, $S^{-1}(IJ) = (S^{-1}I)(S^{-1}J)$, and $S^{-1}(I \cap J) = S^{-1}I \cap S^{-1}J$. Finally, $(S^{-1}A)/(S^{-1}I) \cong S^{-1}(A/I)$, because the quotient is an A-algebra and thus an A-module.

Sometimes, a proper ideal of A will localize to all of A, as when $I \cap S \neq \emptyset$. Then, after localization, I contains an invertible element, so $S^{-1}I = S^{-1}A$.

8. Behavior of Ideals Under Localization: 10/16/13

Given a ring homomorphism $A \xrightarrow{f} B$, one obtains a map $\{\text{ideals in } A\} \to \{\text{ideals in } B\}$ given by $I \mapsto J = f(I)$. Conversely, $\{\text{ideals in } B\} \to \{\text{ideals in } A\}$ given by $J \subset B \mapsto f^{-1}(J)$ (which is an ideal, which is not hard to check). These two operations are called expansion and contraction, respectively.

Lemma 8.1. *Every ideal in $S^{-1}A$ is obtained as an expansion of an ideal in A.*

Proof. Let $J \subset S^{-1}A$ be an ideal. One wants to find an ideal that maps to J, so that its contraction I is $\{a \in A \mid a/1 \in J\}$. Then, J is the expansion of I under the natural map $A \to S^{-1}A$.

That $J \supset f(I)$ is obvious, since $I = f^{-1}(J)$. For the other direction, if $x \in J$, then $x = a/s$ for an $a \in A$ and an $s \in S$, so $x = (a/1)(1/s)$. Then, $a/1 = f(a) \in J$, since $a/1 = xs \in J$, and thus $a \in I$ by the construction of J, so x is generated by $f(I)$, so $x \in (f(I))$. ⊠

Intuitively, this says that there are *fewer* ideals in the target $S^{-1}A$ than in the source; they look like a subset.

Lemma 8.2. *$\{$prime ideals in $S^{-1}A\} \to \{$prime ideals in $A\}$ under contraction is an injective map, and the image consists of prime ideals $\mathfrak{p} \subset A$ such that $\mathfrak{p} \cap S = \emptyset$.*

That this map exists was a homework exercise (i.e. that the preimage of a prime ideal is prime), and that it's injective is a consequence of Lemma 8.1 and the fact that contraction followed by expansion is an isomorphism.

Proof of Lemma 8.2.

(1) If $\mathfrak{q} \subset S^{-1}A$ is such that $f^{-1}(\mathfrak{q}) \cap S \neq \emptyset$, then there exists an $s \in f^{-1}(\mathfrak{q}) \cap S$, meaning that $s/1 \in \mathfrak{q}$, but $s/1$ is invertible in $S^{-1}A$, which is a contradiction.

(2) If $\mathfrak{p} \subset A$ is a prime ideal such that $\mathfrak{p} \cap S = \emptyset$, take $\mathfrak{q}$ to be the expansion of $\mathfrak{p}$: $\mathfrak{q} = ((f(\mathfrak{p})) = \{a/s \mid a \in \mathfrak{p}, s \in S\}$. This is a prime ideal, because $\mathfrak{p}$ doesn't intersect S, so $\mathfrak{q} \subsetneq S^{-1}A$.[33]

It turns out that $\mathfrak{q}$ is also a prime ideal, because if $a/s, b/t \in \mathfrak{q}$, then $(ab)/(st) = c/r$ where $c \in \mathfrak{p}$, since any element of $\mathfrak{q}$ is of that form, and $(abr - cst)s' = 0$ for some $s' \in S$. Reducing mod $\mathfrak{p}$, this means that $(abr - cst)s' \equiv 0 \bmod \mathfrak{p}$, but since $s' \notin \mathfrak{p}$ and $c \in \mathfrak{p}$, then $abr \in \mathfrak{p}$. Thus, since $r \in S$, then $ab \in \mathfrak{p}$, which means one of a or b is in $\mathfrak{p}$, since $\mathfrak{p}$ is prime. Thus, one of a/s and $b/t \in \mathfrak{q}$, so $\mathfrak{q}$ is also prime. To finish the proof, it remains to check that contraction on $\mathfrak{q}$ gives $\mathfrak{p}$ back again, but this is not hard. ⊠

[33] Since proper ideals of A can be sent to A by expansion, it is meaningful to check this.

This is a very nice description of where the prime ideals go: the ones that intersect S are thrown away. Applying it, one can understand the ideals of $A_\mathfrak{p}$ (i.e. $(A \setminus \mathfrak{p})^{-1}A$) by looking at the following diagram:

$$\begin{array}{ccc} \{\text{prime ideals in } A_\mathfrak{p}\} & \hookrightarrow & \{\text{prime ideals of } A\} \\ & \searrow^{\sim} & \downarrow \\ & & \{\text{prime ideals } \mathfrak{p}' \subset A,\ \mathfrak{p}' \subset \mathfrak{p}\} \end{array}$$

That is, prime ideals in $A_\mathfrak{p}$ are in bijection with prime ideals of A *contained in* $\mathfrak{p}$.

Corollary 8.3. *$\mathfrak{p}$ is the unique maximal ideal in $A_\mathfrak{p}$.*

This is because $\mathfrak{p}A_\mathfrak{p}$ is given by expansion.

Definition. A local ring is a commutative ring with a unique maximal ideal.

This is an important class of rings.

Example 8.1. Local rings often arise from function theory, such as $\mathcal{O}^{\text{hol}}$ from Example 7.2. This is a local ring, and its unique maximal ideal is $\mathfrak{m}$, the set of holomorphic functions that vanish at x.

There's a useful criterion for local rings:

Proposition 8.4. *Let A be a ring and $I \subset A$ be an ideal. Then, A is a local ring with maximal ideal I iff $A \setminus I = A^\times$.*

Proof. In the forward direction, if $a \notin I$ is such that (a) is a proper ideal, then $(a) \subset \mathfrak{m}$ for some maximal ideal $\mathfrak{m}$ of A by Zorn's lemma, but this doesn't happen, because then $\mathfrak{m} = I$. Thus, $(a) = A$, so a is a unit.

Conversely, if $A \setminus I = A^\times$, then if J is any proper ideal of A, then $J \subset I$, because J cannot contain any elements of $A \setminus I$, which are all units. Thus, since I is proper (since $1 \notin I$), then it the sole maximal ideal of A, and thus A is a local ring. ⊠

Morally, everything not in the denominator is a unit in a local ring.

Definition. A property $(*)$ of a ring A (resp. A-module M) is called a local property if A (resp. M) has $(*)$ iff $A_\mathfrak{p}$ (resp. $M_\mathfrak{p}$) has the property for all prime ideals $\mathfrak{p}$ of A.

Lemma 8.5. *$x \cdot M = 0$ is a local property.*

Remark. More than the lemma statement is true, since the proof deals with the strongest case. In fact, if $x \cdot A_\mathfrak{m} = 0$ for all maximal ideals $\mathfrak{m}$ of A, then $x \cdot M = 0$.

Definition. If $\mathcal{C}$ and $\mathcal{D}$ are categories, then a functor $F : \mathcal{C} \to \mathcal{D}$ is exact if it sends exact sequences to exact sequences; that is, if $X \xrightarrow{f} Y \xrightarrow{g} Z$ is exact at Y, then $F(X) \xrightarrow{F(f)} F(Y) \xrightarrow{F(g)} F(Z)$ is exact in $\mathcal{D}$.

Theorem 8.6. *The localization functor $M \mapsto S^{-1}M$ from the category of A-modules to the category of $S^{-1}A$-modules is exact.*

Proof. Let $0 \to M' \xrightarrow{\alpha} M \xrightarrow{\beta} M'' \to 0$ be a short exact sequence of A-modules and $S \subset A$ be multiplicative. Consider the maps $S^{-1}M' \xrightarrow{\tilde{\alpha}} S^{-1}M \xrightarrow{\tilde{\beta}} S^{-1}M''$ induced by α and β (i.e. $\tilde{\alpha} : m'/s \mapsto \alpha(m')/s$ and $\tilde{\beta} : m/s \mapsto \beta(m)/s$). Thus:

(1) $\tilde{\alpha}$ is injective, because if $m'/s \mapsto 0 \in S^{-1}M$ for some $m'/s \in S^{-1}M'$, then there exists a $t \in S$ such that $t(1 \cdot \alpha(m') - s \cdot 0)$, so that $t\alpha(m') = 0$. Thus, $\alpha(tm') = 0$, so $tm' = 0$, since α is injective. Thus, $t(1 \cdot m' - 0 \cdot s) = 0$, so since $t \in S$ this implies that $m'/s = 0$ in $S^{-1}M'$.

(2) $\tilde{\beta}$ is surjective, because if $m''/s \in S^{-1}M''$, then because β is surjective there is a preimage $m \in \beta^{-1}(m'')$, so that $\tilde{\beta} : m/s \mapsto m''/s$.

(3) $\text{Im}(\tilde{\alpha}) = \ker(\tilde{\beta})$:
- Suppose $m/s \in \text{Im}(\tilde{\alpha})$, so that $\tilde{\alpha}(m'/s') = m/s$. Then, since $\tilde{\alpha}(m'/s') = \alpha(m')/s'$, then $s' = s$. Then, $\tilde{\beta}(m/s) = \tilde{\beta}(\tilde{\alpha}(m'/s)) = \tilde{\beta}(\alpha(m')/s) = \beta(\alpha(m))/s = 0/s = 0$, so $m/s \in \ker(\tilde{\beta})$.
- Suppose $m/s \in \ker(\tilde{\beta})$, so that $ms \mapsto 0$. Thus, there exists a $t \in S$ such that in M'', $t\beta(m) = 0$, so $\beta(tm) = 0$, so there exists an $m' \in M'$ such that $\alpha(m') = tm$, because the original sequence is exact. Thus, $\tilde{\alpha}(m'/st) = \alpha(m')/st = tm/st = m/s$, so $m/s \in \text{Im}(\tilde{\alpha})$. ⊠

That localization is exact is very important: it preserves injectivity and surjectivity, for example.

Proposition 8.7. *Localization of modules commutes with the direct limit: $S^{-1}(\varinjlim_i M_i) \cong \varinjlim_i (S^{-1}M_i)$.*

This can be proven by definition, but it holds true in much greater generality: there exists a forgetful functor from the category of $S^{-1}A$-modules to the category of A-modules that remembers the A-action on the set but not its ring structure. Then, localization is the left adjoint to this functor; that is, if $M \in \underline{A\text{-Mod}}$ and $N \in \underline{S^{-1}A\text{-Mod}}$, then $\mathrm{Hom}_{S^{-1}A}(S^{-1}M, N) = \mathrm{Hom}_A(M, N)$, where the forgetful functor is applied to N on the right, so one says they're adjoint. This means that Proposition 8.7 can be reformulated as follows.

Proposition 8.8. *Let $\mathcal{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathcal{D}$ be a pair of functors (so that F is the left adjoint to G) and such that direct limits exist in both $\mathcal{C}$ and $\mathcal{D}$. Then, F commutes with direct limits.*

Proof. Consider a system of objects $M_i \in \mathcal{D}$ (concretely, think of these M_i as A-modules). Then, $F(\varinjlim M_i)$ is the desired goal, but instead of building up an isomorphism to $\varinjlim F(M_i)$, one could instead show that for any test object $N \in \mathcal{C}$, $\mathrm{Hom}_{\mathcal{C}}(F(\varinjlim M_i), N) \cong \mathrm{Hom}_{\mathcal{C}}(\varinjlim(F(M_i)), N)$. By the Yoneda lemma, is is sufficient to show that the two are the same, so that F commutes with direct limits.

This can be shown by setting up a group homomorphism between their respective Hom spaces compatible with the maps $N_1 \to N_2$, so

$$\mathrm{Hom}_{\mathcal{C}}(F(\varinjlim M_i), N) = \mathrm{Hom}_{\mathcal{D}}(\varinjlim M_i, G(N))$$

because F and G are adjoint. Then, because $\mathrm{Hom}(\quad, M)$ is contravariant, then

$$\begin{aligned} &= \varprojlim \mathrm{Hom}_{\mathcal{D}}(M_i, G(N)) \\ &= \varprojlim \mathrm{Hom}_{\mathcal{C}}(F(M_i), N), \end{aligned}$$

because F and G are adjoint, so they certainly are term by term. Thus, the necessary equality is shown. ⊠

This is a proof by abstract nonsense; we needed no information about G, other than that it existed. However, the proof did depend on the general categorical facts that $\mathrm{Hom}_{\mathcal{C}}(\varinjlim X_i, Y) = \varprojlim \mathrm{Hom}_{\mathcal{C}}(X_i, Y)$ and $\mathrm{Hom}_{\mathcal{C}}(Y, \varprojlim X_i) = \varprojlim \mathrm{Hom}_{\mathcal{C}}(Y, X_i)$.

Note that in some categories, direct or inverse limits don't exist. Both exist in $\underline{A\text{-Mod}}$, but for commutative rings with identity, the inverse limit is a subring of the product ring (so it exists), but the direct limit might not, since the infinite direct sum of rings lacks 1.[34]

9. Nakayama's Lemma: 10/18/13

Theorem 9.1 (Nakayama's lemma)**.** *Let A be a local ring and $\mathfrak{m}$ its unique maximal ideal, and let M be a finitely generated A-module such that $M = \mathfrak{m}M = \{\sum a_i x_i \mid a_i \in \mathfrak{m}, x_i \in M\}$. Then, $M = 0$.*

The condition of finite generation is necessary: let A be a domain that is not a field, that is also a local ring (the requirement that A isn't a field guarantees that the maximal ideal is nonzero), such as $\mathbb{Z}_{(p)} = \{a/b \mid p \nmid b\}$, with unique maximal ideal $(p)\mathbb{Z}_{(p)}$. Let $M = \mathrm{Frac}(A)$ (in the case $\mathbb{Z}_{(p)}$, $M = \mathbb{Q}$); then, $M = (p)M$, since every element of $\mathbb{Q}$ can be divided by an element of (p), but $M \neq 0$.

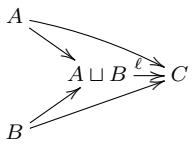
To prove Theorem 9.1, one fact will be needed from linear algebra.

Lemma 9.2. *If $B = (b_{ij}) \in \mathrm{Mat}_n(A)$, where A is a commutative ring, then B is invertible iff $\det(B) \in A^\times$.*

Proof. If $BC = 1$ for some $C \in \mathrm{Mat}_n(A)$, then $\det(B)\det(C) = 1$, so $\det(B) \in A^\times$.

Conversely, if $\det(B)$ is invertible, then construct the inverse by letting C be the adjoint matrix B^* scaled: $C = B^*(\det B)^{-1}$, so that $C = B^{-1}$. ⊠

[34] There's some categorical lawyerese going on here, in that the coproduct of commutative rings is actually tensor product rather than direct sum, because the categorical direct sum $A \sqcup B$ of A and B is such that for every object C, there exists a unique morphism ℓ such that



commutes. If one thinks about the diagram enough, there's no canonical way to do this for $A \times B$, and instead one wants $A \otimes_{\mathbb{Z}} B$ (or the infinite product and tensor product). Nonetheless, there are still categories in which direct or inverse limits don't exist, such as the category of finite sets.

Proof of Theorem 9.1. Since M is finitely generated, there exists a surjection $A^n \overset{\pi}{\twoheadrightarrow} M$. Let $N = \ker(\pi)$, so that one obtains the short exact sequence $0 \to N \to A^n \overset{\pi}{\to} M \to 0$. Let $e_1, \dots, e_n$ denote the standard basis of A^n and $x_i = \pi(e_i)$, so that $\{x_1, \dots, x_n\}$ generates M. Thus, since $M = \mathfrak{m}M$, then each x_i can be written a $x_i = \sum a_j y_j$ for $a_j \in \mathfrak{m}$ and $y_j \in M$, and in turn each y_j can be written in terms of the generators. Once this is done, each coefficient x_j appears in the sum multiple times, so collect the coefficients and write $x_i = \sum b_{ij} x_j$ for some $b_{ij} \in \mathfrak{m}$. Let $B \in \mathrm{Mat}_n(A)$ have $(i,j)^{\text{th}}$ entry equal to b_{ij}, and view B as an A-linear map $A^n \to A^n$.

Then, $\mathrm{Im}(\mathrm{id} - B) \subset N$, because

$$(\mathrm{id} - B) \cdot e_i = e_i - \sum_{j=1}^{n} b_{ij} e_j \overset{\pi}{\longmapsto} x_i - \sum_{j=1}^{n} b_{ij} x_j = 0.$$

Since the entries of B all lie in $\mathfrak{m}$, then $\det(\mathrm{id} - B)$ is of the form $1 + m$ for some $m \in \mathfrak{m}$, which can be shown by induction on n:

- In the base case, $n = 1$ and $B = [b_{11}]$. Then, $\det(\mathrm{id} - B) = \det(1 - b_{11}) = 1 + (-b_{11})$, so $m = b_{11}$.
- In the general case, calculate the determinant of $\mathrm{id} - B$ by taking minors along the first row, and let M_i be the minor obtained by removing the i^{th} column and first row from B. Then,

$$\begin{aligned}\det(B) &= \sum_{j=1}^{n} (\mathrm{id} - B)_{1j} \det(M_j) = (1 - b_{11}) \det(M_1) + \sum_{j=2}^{n} b_{1j} \det(M_j) \\ &= (1 - b_{11})(1 - m_1) + \sum_{j=2}^{n} b_{1j} m_j,\end{aligned}$$

  for some $m_1, \dots, m_n \in \mathfrak{m}$, by the inductive assumption.

$$= 1 + \underbrace{b_{11} m_1 - m_1 - b_{11} + \sum_{j=2}^{n} b_{1j} m_j}_{\in \mathfrak{m}}.$$

Thus, $\det(\mathrm{id} - B) \in 1 + \mathfrak{m}$. Since $\det(\mathrm{id} - B) \notin \mathfrak{m}$ and A is local, then $\det(\mathrm{id} - B) \in A^\times$, which by a theorem of linear algebra implies that $\mathrm{id} - B$ is invertible. In particular, $\mathrm{Im}(\mathrm{id} - B) = A^n$, so $N = A^n$ and thus $\pi = 0$, since $N = \ker(\pi)$. However, since π is surjective, this forces $M = 0$. ⊠

Remark. The crucial aspect of this argument is that $1 + \mathfrak{m} \subset A^\times$. Thus, if A is a commutative ring and $I \subset A$ is an ideal such that any element of the form $1 + I$ is invertible, then the statement of the theorem still holds: if M is a finitely generated A-module such that $M = IM$, then $M = 0$.

For example, if A is any commutative ring and I is the nilpotent radical (the ideal of all nilpotent elements), then $1 + I \subset A^\times$, because $(1 + x)^{-1} = 1 + x + x^2 + \cdots$, and this terminates eventually.

Corollary 9.3. *Let M be an A-module for a local ring A with maximal ideal $\mathfrak{m}$. Let $x_1, \dots, x_n \in M$. Suppose their images in $M/\mathfrak{m}M$ form a k-basis. Then $\{x_1, \dots, x_n\}$ generate M as an A-module.*

Proof. Again denote $\pi : A^n \to M$ sending $e_i \to x_i$. Unlike the previous part, however, the goal here is to prove that π is surjective; if this can be shown, then it implies that the x_i generate M, which is what needs to be demonstrated.

Let $L = \mathrm{coker}(\pi)$, so that there is an exact sequence $A^n \overset{\pi}{\to} M \to L \to 0$. The sequence $A^n/\mathfrak{m}A^n \overset{\overline{\pi}}{\to} M/\mathfrak{m}M \to L/\mathfrak{m}L \to 0$ is still exact, which implies that $L/\mathfrak{m}L = \mathrm{coker}(\overline{\pi})$, where $\overline{\pi}$ denotes the map induced by π on the quotient. Let $\overline{e}_i$ and $\overline{x}_i$ be the images of e_i and x_i, respectively, under their respective quotients, so that $\overline{\pi} : \overline{e}_i \mapsto \overline{x}_i$. However, the $\overline{x}_i$ generate $M/\mathfrak{m}M$ and they are in the image of $\overline{\pi}$, so $\mathrm{Im}(\overline{\pi}) = M/\mathfrak{m}M$, and $\overline{\pi}$ is surjective. Thus, $\mathrm{coker}(\overline{\pi}) = L/\mathfrak{m}L = 0$, which means that $L = \mathfrak{m}L$.

Since M is finitely generated and there is a surjection $g : M \twoheadrightarrow L$ given by the exact sequence above, then L is generated by the images of the generators of M under g; thus, L is also finitely generated. Thus, Theorem 9.1 applies, and implies that $L = 0$. Thus, $\mathrm{coker}(\pi) = 0$, so π is surjective. ⊠

There's a nuance here: just because the quotient is finitely generated doesn't imply that the original is too. For example, if $A = \mathbb{Z}_{(p)}$ and $M = \mathbb{Q}$, then $M/\mathfrak{m}M = 0$, which is certainly finitely generated.

Proposition 9.4. *Let A be a commutative ring and $f : A^n \to A^n$ be A-linear. If f is surjective, then f is an isomorphism.*

Notice that this doesn't work if f is merely injective, such as $\mathbb{Z} \overset{\cdot 2}{\to} \mathbb{Z}$.

Proof of Proposition 9.4. Let $K = \ker(f)$, and let $M = N = A^n$ so that they can be distinguished. Thus, the proposition statement induces the short exact sequence $0 \to K \to M \xrightarrow{f} N \to 0$. Once nice property of free A-modules is that there exists a section $N \to M$ given by $s : e_i \to \tilde{e}_i \in f^{-1}(e_i)$, so that $f \circ s = \mathrm{id}$. Thus, the module M splits: $M \cong s(N) \oplus K$,

The goal is to show that $K = 0$, which is equivalent to $K_{\mathfrak{p}} = 0$ for every prime ideal $\mathfrak{p}$ of A. Thus, pick such a $\mathfrak{p}$ and localize: $M_{\mathfrak{p}} \cong s(N_{\mathfrak{p}}) \oplus K_{\mathfrak{p}}$ (though it's important to check that the section behaves well with respect to localization, so that $s(N)_{\mathfrak{p}} = s(N_{\mathfrak{p}})$). This isomorphism is as $A_{\mathfrak{p}}$-modules. Now, mod out by $\mathfrak{p}$, giving the induced section $\overline{s} : N_{\mathfrak{p}}/\mathfrak{p}N_{\mathfrak{p}} \to M_{\mathfrak{p}}/\mathfrak{p}M_{\mathfrak{p}}$. This is a section of the similarly defined $\overline{f}$, which is still a surjection. Thus, $M_{\mathfrak{p}}/\mathfrak{p}M_{\mathfrak{p}} \cong s(N_{\mathfrak{p}}/\mathfrak{p}N_{\mathfrak{p}}) \oplus K_{\mathfrak{p}}/\mathfrak{p}K_{\mathfrak{p}}$, and s is injective, so $K_{\mathfrak{p}}/\mathfrak{p}K_{\mathfrak{p}} = 0$. Thus, $K_{\mathfrak{p}} = 0$. All that remains to check is finite generation, but this is easy, since a finite direct summand of finitely generated modules is finitely generated. ⊠

Corollary 9.5. *Since the only requirement was that the sequence split, A^n can be replaced with any finitely generated projective A-module.*

This sort of proof is very common in commutative algebra: a statement about general rings can be reduced to a problem on local rings. Then, reducing mod $\mathfrak{p}$ makes it a statement about vector spaces, which can be handled by linear algebra.

Completion. Related to localization is the notion of completion, such as $k[x] \rightsquigarrow k[[x]] = \varprojlim_n k[x]/(x^n)$. One can view this as a sequence of surjections $\cdots \to k[x]/(x^3) \to k[x]/(x^2) \to k[x]/(x) = k$. The inverse limit maps surjectively onto everything, so it is "on top" of the whole chain. This is called the completion of $k[x]$ at the ideal (x), and in the limit, is a sort of replacement for the original ring.

Definition. More generally, if A is a commutative ring and $I \subset A$ is an ideal, then the completion $\widehat{A}_{/I} = \varprojlim_n A/I^n$. This can be denoted $\widehat{A}$ if I is known from context.

Often, one takes A to be local, and completes with respect to its maximal ideal $\mathfrak{m}$. Then, take a look at the last bit of the tower:

$$\begin{array}{ccc} \mathfrak{m}/\mathfrak{m}^2 & \hookrightarrow & A/\mathfrak{m}^2 \\ & & \downarrow \\ & & A/\mathfrak{m} = k \end{array}$$

This is a short exact sequence.

Definition. $\mathfrak{m}/\mathfrak{m}^2$, which is a k-vector space, is called the cotangent space.

Example 9.1. A motivating example for the name is as follows: let A be the ring of C^∞ germs at $\mathbf{0} \in \mathbb{R}^n$. Then, $\mathfrak{m}$ is the set of germs that vanish at $\mathbf{0}$, $\mathfrak{m}^2$ the set of germs that vanish at the origin to second order (in the Taylor expansion), etc. Thus, $\mathfrak{m}/\mathfrak{m}^2$ is the set of linear terms in the Taylor expansion. This is a vector space upon which one can apply differential operators, such as those in the tangent space $\mathrm{Tang}_{\mathbf{0}} \mathbb{R}^n$, generated by $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}$. Then, this is a natural dual to $\mathfrak{m}/\mathfrak{m}^2$, because $\mathrm{Tang}_{\mathbf{0}} \mathbb{R}^n \times \mathfrak{m}/\mathfrak{m}^2 \to \mathbb{C}$ given by $\left\langle \frac{\partial}{\partial x_i}, f \right\rangle \mapsto \frac{\partial f}{\partial x_i}(\mathbf{0})$ is something called a perfect pairing.

Now, if $\widehat{A}$ is the completion of A with respect to $\mathfrak{m}$, then $\widehat{A} \cong \mathbb{R}[[x_1, \ldots, x_n]]$, though this isomorphism is noncanonical. When one takes a singular curve or surface rather than a manifold, such as $A = k[x, y]/(x^2 - y^2 - y^3)$, completing at (x, y) gives $\widehat{A}[[a, b]]/(ab)$, which is a lot simpler to deal with. For example, if $k = \mathbb{C}$, this just looks like the coordinate axes, which was the localization of the more complicated curve $x^2 - y^2 - y^3 = 0$.

10. The p-adic Integers: 10/23/13

Recall the notion of completion: if I is an ideal of A, take $A \supset I \supset I^2 \supset \cdots$, and let $\{I^n \mid n \in \mathbb{N}\}$ be a directed set with the canonical maps (i.e. inclusion $I_i \hookrightarrow I_j$ if $i \geq j$). Then, the completion is $\hat{A}_{/I} = \varprojlim_n A/I^n$. More generally, one can take any set $\{I_i\}_{i \in \mathcal{I}}$ of ideals of A, and view it as a directed set by $i \to j$ whenever $I_i \subset I_j$, so that $A/I_i \twoheadrightarrow A/I_j$. Then, the completion is $\hat{A}_{/I} = \varprojlim_{i \in \mathcal{I}} A/I_i$.

It turns out there exists a unique topology on A such that $\{I_i\}$ forms a neighborhood basis[35] of $0 \in A$ by declaring that a set is a neighborhood of 0 if it contains an ideal I_i, and extending this to the rest of A by declaring it invariant under translation, i.e. $\{a + I_n\}$ forms a neighborhood basis for $a \in A$, and this uniquely determines the topology (along with addition being continuous).

[35]This means that for any open neighborhood U of the origin, there is some $I_n \subset U$ and $0 \in I_n$.

Then, the completion of A as described above is actually the completion with respect to this topology, i.e. the set of Cauchy sequences under an equivalence relation of convergence. A Cauchy sequence in A is a sequence $(x_i)_{i\in I}$ such that for every $i \in I$, there exists an $N(i) \in \mathbb{N}$ such that $x_n \equiv x_m \pmod{I_i}$ for all $m, n \geq N(i)$. Two Cauchy sequences are considered equivalent if their difference goes to zero, i.e. $x_n \equiv y_n \pmod{I_i}$ for $n > N(i)$. This is very similar to the notion of completeness in a metric space, but the sequence of ideals replaces the metric.

For a specific example we will consider the p-adic integers. Take $A = \mathbb{Z}$ and pick some prime p, and consider the sequence $(p) \supset (p^2) \supset (p^3) \supset \cdots$. Then, $\hat{A}_{/(p)} = \varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$ is called the ring of p-adic integers, denoted $\mathbb{Z}_p$. By definition, an element in this limit is a sequence (x_n), where $x_n \in \mathbb{Z}/p^n\mathbb{Z}$ and such that $x_{n+1} \equiv x_n \ (p^n)$. Then, $x_1 = a_0 \in \{0, 1, \dots, p-1\}$ and $x_2 = a_0 + pa_1$, where $a_1 \in \mathbb{Z}/p\mathbb{Z}$ again. Since such choices exhaust everything mod $\mathbb{Z}/p^2\mathbb{Z}$, then keep going: $x_3 = a_0 + pa_1 = p^2a_2$ for $a_0, a_1, a_2 \in \mathbb{Z}/p\mathbb{Z}$, and so on. Thus, to give a p-adic integer is the same as giving a sequence $(a_0, a_1, a_2, \dots)$ in $\mathbb{Z}/p\mathbb{Z}$, and there are no restrictions on them. In some sense, this is akin to a set of formal power series, and $\mathbb{Z}/p \leftrightarrow \prod_{n=0}^{\infty} \mathbb{Z}/p\mathbb{Z}$, though this is only as sets, and the ring structure on $\mathbb{Z}_p$ is much more interesting.

Using topology, it's possible to make sense of this infinite sequence.

Definition. Let $x \in \mathbb{Z}_p \setminus 0$. Define the p-adic valuation $v_p(x)$ to be the largest n such that $x \equiv 0 \ (p^n)$.

This is equal to the number of leading zeros in the expansion $x = (a_1, a_2, \dots)$ given above; $v_p(x)$ is the first index such that $a_{v_p(x)}$ is nonzero.

Since every ring maps into its completion, then $\mathbb{Z} \subset \mathbb{Z}_p$, so if $x \in \mathbb{Z} \setminus 0$, then the p-adic norm of x is the largest $n \in \mathbb{N}$ such that $p^n \mid x$. Thus, if $p \nmid x$, then $v(x) = 0$, and if $pq = x$ for $(q, p) = 1$, then $v_P(x) = 1$, and so forth.

This defines a map $v_p : \mathbb{Z}_p \setminus 0 \to \mathbb{Z}_{\geq 0}$ sending an element to its p-adic valuation. Here are some properties of this map:

- $v_p(xy) = v_p(x) + v_p(y)$, because if $x = p^n a_n + p^{n+1}a_{n+1} + \cdots$ and $y = p^m b_m + \cdots$, then $xy = p^{m+n}(a_n b_m) + \cdots$, and since $a_n, b_m \neq 0$, then their product is still nonzero.
- $v_p(1) = 0$, since $p \nmid 1$.
- $v_p(x+y) \geq \min(v_p(x), v_p(y))$; if $n \neq m$, this is strict equality, but if $n = m$, then it's possible that $a_n + b_n = 0$, and then maybe $a_{n+1} + b_{n+1} = 0$, and so on, so it's not possible to obtain an exact bound for the entire set.

This valuation can be used to define a metric on $\mathbb{Z}_p$, which is really just a different way of saying the same thing. Let

$$d(x, y) = \begin{cases} 0, & x = y \\ p^{-v_p(x-y)}, & x \neq y. \end{cases}$$

Concretely, $d(x, 0) = p^{-n}$, where n is the largest integer such that $p^n \mid x$. Thus, x is closer to zero if it is more divisible by p. Note that any other positive integer can be used as the exponent index in the formula for $d(x, y)$, and the topology is the same.

With this metric, the infinite series $a_0 + pa_1 + p^2a_2 + \cdots$ makes sense for all $a_i \in \mathbb{Z}$, and is a well-defined element of $\mathbb{Z}_p$, since it has a well-defined residue class modulo every p^n.[36]

Consequently, one can define some functions which don't exist on the integers. Assuming $p \neq 2$,[37] define the exponential map $\exp : p\mathbb{Z}_p \to 1 + p\mathbb{Z}_p$ (i.e. it's necessary for the valuation to be at least 1) given by sending $x \mapsto 1 + x + x^2/2 + x^3/6 + \cdots$. To make sense of such a sequence, it's necessary to be able to divide by these numbers. Fortunately, we have the following lemma to the rescue.

Lemma 10.1. *If $n \in \mathbb{Z}$ is prime to p, then n is invertible in $\mathbb{Z}_p$.*

Proof. Explicitly construct an inverse, similarly to finding the inverse of a formal power series. Then, $n = a_0 + pa_1 + p^2a_2 + \cdots$ with $a_0 \neq 0$, so take $b_0 + b_1p + \cdots$ such that their product is 1, or $a_0b_0 \equiv 1 \ (p)$. Thus, the first step is to pick any $b_0 \in \mathbb{Z}$ such that $a_0b_0 \equiv 1 \ (p)$ (which can be done because $a_0 \not\equiv 0 \ (p)$), and then a $b_1 \in \mathbb{Z}$ such that $b_1a_0 + a_1b_0 \equiv 0 \ (p)$, and so on. The error term is pushed to be divisible by higher and higher powers of p, so it must be zero. ⊠

Thus, $\exp(x)$ makes sense up to $x^p/p!$ $p!$ is pretty clearly noninvertible in $\mathbb{Z}_p$. But it's a nice fact that when $v_p(x) \geq 1$, then $v_p(x^p) \geq v_p(p!)$, because there is exactly one p in $p!$, and at least p in x^p. Thus, $x^p/p! \in p^{p-1}\mathbb{Z}_p$ (i.e. it has valuation at least $p-1$), so in general, the n^{th} term has p-adic valuation at least n, except past $x^p/p!$, where it's at least $n-1$, and so on.

Thus, it's necessary to compare $v_p(x^n) - v_p(n!)$ for all n, which is the p-adic valuation of $x^n/n!$. It turns out that $v_p(x^n) - v_p(n!) \geq n - n(1/p + 1/p^2 + \cdots)$, because of all $k \leq n$, roughly n/p of them are divisible by p,a nd then

[36] This residue class can be found by truncating the sequence early. This is clearly compatible with the criterion for the inverse limit, so it lies within the inverse limit.

[37] There is a version that works for $\mathbb{Z}_2$, but it requires restricting the domain a bit more.

n/p^2 are divisible by p^2, and so on. Since this simplifies to $(1-1/(p-1))n$, then this just grows linearly with n. Thus, the sequence converges in $\mathbb{Z}_p$, and $\exp(x) \equiv 1 \bmod p$.

There is also a logarithm $\log : 1 + p\mathbb{Z}_p \to p\mathbb{Z}_p$ given by a Taylor series:

$$y \mapsto \log(1+(y-1)) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(y-1)^n}{n}.$$

The same argument as for the exponential holds: $(y-1) \in p\mathbb{Z}_p$, so $(y-1)^n \in p^n\mathbb{Z}_p$, and the denominator is in $p^k\mathbb{Z}_p$ roughly n/p^k of the time Thus, the log converges.

It happens that the exp and log are inverses of each other, and $\exp : (p\mathbb{Z}_p, +) \to (1+p\mathbb{Z}_p, \cdot)$ is a group homomorphism (because $(1+\cdots)(1+\cdots) = 1+\cdots$ still). It's necessary to check that $\exp(x+y) = \exp(x)\exp(y)$, but this isn't too bad now that we know the power series works out nicely. Since the logarithm is also a group homomorphism, then $(p\mathbb{Z}_p, +) \cong (1+p\mathbb{Z}_p, \cdot)$. This isomorphism is very useful: for example, if $x \in 1+p\mathbb{Z}_p$ (i.e. $x \equiv 1 \bmod p$), then as long as $(n,p) = 1$, then $\sqrt[n]{x} = \exp(\log(\sqrt[n]{x})) = \exp((1/n)\log x)$, since n is known to be invertible by Lemma 10.1. This fact also works in any $\mathbb{Z}_p$-module. Other n^{th} roots may exist, but this is the unique one that is congruent to 1 mod p.

This serves as a useful example of how by passing to a completion, a ring becomes more flexible than it used to be, even admitting analysis.

One can localize $\mathbb{Z}_p$ at $\{p^n \mid n \geq 0\}$, so that one obtains $\{p^n\}^{-1}\mathbb{Z}_p = \{x/p^n \mid x \in \mathbb{Z}_p, n \geq 0\}$. This object happens to be a field, called the field of p-adic numbers, and is denoted $\mathbb{Q}_p$. Elements of $\mathbb{Q}_p$ are Laurent series in p: $\{p^n a_n + p^{n+1}a_{n+1} + \cdots \mid a \in \mathbb{Z}/p, n \in \mathbb{Z}\}$; that is, every sequences has to start from somewhere, but that somewhere can be any integer. The valuation can be extended from $\mathbb{Z}_p$ to $\mathbb{Q}_p$, giving $v_p : \mathbb{Q}_p \setminus 0 \to \mathbb{Z}$; the valuation can now be negative. The valuation of some element is precisely the index n of the starting point of its Laurent series.

It would be convenient to generalize this situation.

Definition. A discrete valuation field (DVF) is a field K along with a map $v : K^* \twoheadrightarrow \mathbb{Z}$,[38] such that

(1) v is a group homomorphism $(K^*, \cdot) \to (\mathbb{Z}, +)$.
(2) $v(x,y) \geq \min(v(x), v(y))$.

Note that condition 1 implies that $v(xy) = v(x) + v(y)$, so $\mathbb{Q}_p$ is a discrete valuation field.

Definition. A discrete valuation ring (DVR) is a domain A such that:

(1) its fraction field K is a discrete valuation field,[39]
(2) $A = v^{-1}(\mathbb{Z}_{\geq 0}) \cup \{0\}$.

In essence, one starts with a DVF and takes all elements with nonnegative valuation as a subring. That this is a ring has to be checked: that it's closed under multiplication follows from part 1 of the definition of a DVF, and that it's closed under addition follows from part 2.

Example 10.1. Along with $\mathbb{Z}_p \subset \mathbb{Q}_p$, one also has $k[[x]] \subset k((x))$ (power series and Laurent series, respectively). These are both complete in the sense presented earlier in this lecture; one example that isn't complete is $\mathbb{Z}_{(p)} = \{a/b \mid b \nmid p\} \subset \mathbb{Q}$. In this last example, $\mathbb{Q}$ is a DVF with $v_p : \mathbb{Q}^\times \to \mathbb{Z}$ (i.e. using the p-adic valuation for any prime p), so $\mathbb{Z}_{(p)}$ is the ring of rational numbers with nonnegative valuation (so that the denominator is prime to p).

Notice that $\mathbb{Z}_{(p)} \subset \mathbb{Z}_p$ as rings, because all $b \nmid p$ are invertible in $\mathbb{Z}_p$, so by the universal property of localization, $\mathbb{Z}_{(p)} \hookrightarrow \mathbb{Z}_p$. This also means that $\mathbb{Z}_p = \varprojlim_n \mathbb{Z}_{(p)}/p^n\mathbb{Z}_{(p)}$, which is the same idea as the construction with $\mathbb{Z}$, but with a local ring.

Similarly, let $K = k(x)$, the field of rational functions over some field k, and let $v : K^* \to \mathbb{Z}$ send an $f \in K^*$ to the order of vanishing of f at $x = 0$.[40] This means that if $v(f) < 0$, then f has a pole at 0.

Then, if $A = \{0\} \cup v^{-1}(\mathbb{Z}_{\geq 0})$, then $A = \{g(x)/h(x) \mid h(0) \neq 0\}$, a construction similar to $\mathbb{Z}_p$. This is the same thing as $A = k[x]_{(x)}$, localizing at the prime ideal (x). Localizing a ring at a prime ideal is a good way to obtain DVRs, but doesn't work always or even often.

Proposition 10.2. *A is a discrete valuation ring iff A is a local ring and a principal ideal domain.*

Partial proof. If A is a DVR, then there is a map $v : K^* \to \mathbb{Z}$, where K is the field of fractions of A. Take an $a \in K$ such that $v(a) = 1$; then one can show that $v^{-1}(\mathbb{Z}_{\geq 1}) \cup \{0\}$ is a maximal ideal of A; call it $\mathfrak{m}$. Then, $\mathfrak{m} = (a)$. Thus, $A \setminus \mathfrak{m} = A^\times$ (since it has everything with valuation zero), so $\mathfrak{m}$ is maximal, and all ideals of A are of the form (a^n) or (0). Thus, A is a PID.

In the other direction, try to produce a valuation $v : A \setminus 0 \to \mathbb{Z}$ by sending x to the largest n such that $x \in \mathfrak{m}^n$. ⊠

[38] One could create an entire map $K \to \mathbb{Z} \cup \{\infty\}$ by formally adding the symbol ∞ and defining it to be greater than every other integer, and life still works.

[39] This isn't just a condition; the valuation on K is an important part of the data.

[40] This means that if $f(x) = g(x)/h(x) = x^n g_1(x)/x^m h_1(x)$, then $v(f) = n - m$.

Part 5. Tensor Algebra

11. Discrete Valuation Rings and Tensor Products: 10/25/13

Definition. If A is a DVR, we saw in the last class that it is a local ring that is a PID. Its unique maximal ideal is $\mathfrak{m} = (x)$, generated by any x such that $v(x) = 1$. Thus, $x \in \mathfrak{m} \setminus \mathfrak{m}^2$. Such an x is called a uniformizer.

For example, in $\mathbb{Z}_p$, a uniformizer is an element divisible by p exactly once: $pa_1 + p^2a_2 + \cdots$ such that $a_1 \not\equiv 0 \ (p)$. In $k[[x]]$, a uniformizer is $c_1x + c_2x^2 + \cdots$ such that $c_1 \neq 0$.

All ideals of a DVR A are of the form $\mathfrak{m}^n = (x^n)$ or (0).

Using this idea of a DVR, it is possible to obtain a short proof of the structure theorem for finitely generated modules over a PID. To be precise, the torsion part of Theorem 3.2 will be shown: that if A is a PID and M is a finitely generated A-module, then $M \cong A/(a_1) \oplus \cdots \oplus A/(a_n)$ such that $a_1 \mid \cdots \mid a_n$.

Proof. Just as in the usual proof, reduce to the case where M is killed by p^N, using the Chinese Remainder theorem. Thus, $A/p^N \circlearrowright M$, and the goal is to write M as a direct sum of A/p^{e_i}.

But this is the same as localizing M at the prime p (i.e. at (p), which is a prime ideal). Again using the Chinese Remainder theorem, let $M[p^\infty] = \{\alpha \in M \mid p^N x = 0 \text{ for some } n \in \mathbb{N}\}$, which is a submodule of M.

Exercise 11.1. Show that $M[p^\infty] = M_p$; this depends on the fact that M is torsion.

Knowing this, it's possible to view $M[p^\infty]$ as an A_p-module.

Claim. A_p is a DVR.

Proof. This is immediate from Proposition 10.2, because A_p is local, and it's a PID because A is.

Alternatively, one can explicitly define a valuation on A_p. Let $K = \mathrm{Frac}(A_p)$ and define $v_p : K^* \to \mathbb{Z}$ by generalizing the p-adic definition on $\mathbb{Q}$. Specifically, $a/b = p^m q_1/p^n q_2$; then, $v_p(a/b) = m - n$. It remains to check the axioms for a DVF for K, and that $A_p = v_p^{-1}(\mathbb{Z}_{\geq 0}) \cup \{0\}$. ⊠

Now, M_p is a finitely generated torsion A_p-module, ad we want to write it as $\bigoplus A_p/p^n A_p$, since $A_p/p^n A_p \cong A/p^n$, so this implies the theorem. To simplify the notation, write A for A_p and M for M_p. Let $\mathfrak{p}$ denote the maximal ideal of A and π be a uniformizer.

Since M is finitely generated, then write $A^m \xrightarrow{f} A^n \to M \to 0$ (though it's necessary to argue that the kernel of the projection is also finitely generated). Then, f is given by an $n \times m$ matrix (a_{ij}), and the goal is to transform this into a diagonal matrix with diagonal entries $\pi_1^{e_1}, \ldots, \pi_\ell^{e_\ell}, 0, \ldots, 0$ using elementary row and column operations. This is implied by one of the variations of the structure theorem, but since A is a DVR, it can be done directly:

(1) Find an a_{ij} in the matrix with the smallest valuation, i.e. least divisible by powers of π, and place it in position $(1, 1)$. Now, a_{11} has the smallest valuation.
(2) Do some elementary row operations, since $v(a_{1i}) > v(a_{11})$, then $a_{11} \mid a_{1i}$ for each i, so it's possible to clear out the rest of the column out.
(3) Do the same thing on the first row, so that a_{11} is the only nonzero entry in the first row and column.
(4) Now, induct on the smaller matrix consisting of the second through last rows and columns.

The crucial property is that the element with the smallest valuation divides everything else, which doesn't hold true for PIDs in general. ⊠

This argument works whenever the localization of a ring is a DVR, along with the finiteness condition used in the reduction steps. This motivates a generalization.

Definition. A Dedekind domain A is a domain such that:

(1) $A_\mathfrak{p}$ is a DVR for any nonzero prime ideal $\mathfrak{p}$ of A.
(2) A is Noetherian.

The second constraint is the finiteness condition, albeit somewhat implicitly.

Example 11.1. Let X be a compact Riemann surface, e.g. the projective space $\mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$, or the torus generated by quotienting $\mathbb{C}$ by some lattice. This means that it is a surface with a complex structure. Fix some points $x_1, \ldots, x_n \in X$.

Let A be the set of holomorphic functions on $X \setminus \{x_1, \ldots, x_n\}$ which are meromorphic at $x_1, \ldots, x_n$ (i.e. there are possibly poles at the x_i, but everywhere else it's well-behaved). Then, the ring structure on A turns it into a Dedekind domain.

For example, if $X = \mathbb{P}^1(\mathbb{C})$ and $x = \infty$, then A is the subring of $\mathrm{Hol}(X)$ that are meromorphic at ∞ (i.e. excluding functions like $\exp(z)$, which has an essential singularity at infinity). Thus, $A = \mathbb{C}[z]$. This specific case ends up as a PID; all PIDs are Dedekind domains, but the converse isn't true.

Theorem 11.1 (Structure Theorem for Dedekind Domains)**.** *Suppose A is a Dedekind domain and M is a finitely generated A-module. Then,*

$$M \cong A^{r-1} \oplus I \oplus \underbrace{\bigoplus_{i=1}^{s} A/\mathfrak{p}_i^{e_i}}_{\mathrm{Tor}_A(M)}$$

for some $r > 0$, $s \geq 0$, an ideal I of A, and prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ of A, subject to the following uniqueness conditions:

- *r and s are unique,*
- *I is unique up to multiplication by a principal ideal,*
- *and the $\{(\mathfrak{p}_1, e_1), \ldots, (\mathfrak{p}_s, e_s)\}$ are unique up to reordering.*

Tensor Products. Fix a commutative ring A, so that the tensor product of two A-modules M and N can be discussed.

In linear algebra, the tensor product $V \otimes_k W$ of two vector spaces V and W with bases $\{e_i\}$ and $\{f_i\}$, respectively, is a vector space with basis $\{e_i \otimes f_j\}$. However, to make a definition it's important to avoid the notion of basis, especially once the result is generalized to non-free A-modules, for which the notion of basis doesn't make any sense.

For any A-module L, define $\mathrm{Bil}_A(M, N; L) = \{\text{bilinear } M \times N \xrightarrow{f} L\}$ (i.e. $f(a_1 m_1 + a_2 m_2, n) = a_1 f(m_1, n) + a_2 f(m_2, n)$ and $f(m, a_1 n_1 + a_2 n_2) = a_1 f(m, n_1) + a_2 f(m, n_2)$); notice that neither of these are linear functions on $M \times N$.

Definition. Recall that by the Yoneda lemma, there is an embedding $\mathcal{C}^{\mathrm{op}} \hookrightarrow \mathrm{Fun}(\mathcal{C}, \underline{\mathrm{Set}})$ given by $X \mapsto h_X$, where $h_X : Y \mapsto \mathrm{Mor}_{\mathcal{C}}(X, Y)$, and this functor is fully faithful. Then, a functor in $\mathrm{Im}(h)$ is called representable in $\mathcal{C}$, and the object that maps to it is said to represent it.[41]

As with localization, the definition shall be encompassed in the theorem.

Theorem 11.2. *The functor $\underline{A\text{-Mod}} \to \underline{\mathrm{Set}}$ sending $L \mapsto \mathrm{Bil}_A(M, N; L)$, where M and N are fixed and L varies in $\underline{A\text{-Mod}}$, is representable. The object representing this is denoted $M \otimes_A N$.*

This theorem can be reformulated to state that there exists an A-module X such that $h_X \cong \mathrm{Bil}_A(M, N; -)$;[42] in other words, for each A-module L, one wants a bijection between $\mathrm{Hom}_A(X, L) \cong \mathrm{Bil}_A(M, N; L)$, such that this bijection respects maps between A-modules.

Proof of Theorem 11.2. The proof gives a concrete construction of X, though it's a bit of a tautology. If $f \in \mathrm{Bil}_A(M, N; L)$, then $f : M \times N \to L$ is a set map; in general it doesn't preserve the abelian group structure on these sets. Then, extend f to the free $\mathbb{Z}$-module with basis $M \times N$ (as a set). Then, some parts of this huge module $\mathbb{Z}[M \times N]$ have to be killed: let $X = \mathbb{Z}[M \times N]/R$, where R is the subgroup generated by the following elements:

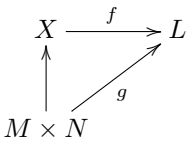
(1) $(m_1 + m_2, n) - (m_1, n) - (m_2, n)$ for all $m_1, m_2 \in M$ and $n \in N$,
(2) $(m, n_1 + n_2) - (m, n_1) - (m, n_2)$ for all $m \in M$ and $n_1, n_2 \in N$, and
(3) $(am, n) - (m, an)$ for all $m \in M$, $n \in N$, and $a \in A$.

Though this definition isn't all that easy to work with, there's not actually all that much to check. For simplicity, denote the image of (m, n) in X as $m \otimes n$. Then, the quotient relations in R imply relations among these tensors $m \otimes n$.

Define an A-action on X by $a(m \otimes n) = am \otimes n$, which is also equal to $m \otimes an$ by the third relation. This isn't a complete description of the A-module structure for a general element in X, which is of the form $\sum_{i=1}^{n} m_i \otimes n_i$, because each element is a finite weighted sum of the basis elements, but the weights can be absorbed into the m_i or the n_i by the third relation in R. Then, $a \cdot \sum m_i \otimes n_i = \sum (a m_i) \otimes n_i$. This ends up being well-defined, because if $\sum m_i \otimes n_i = \sum m_i' \otimes n_i'$, then their difference is in the quotient and it ends up working out.

Exercise 11.2. Show that this action gives X an A-module structure.[43]

Next, if L is some other A-module, then $\mathrm{Hom}_A(X, L) \subset \mathrm{Hom}_{\mathbb{Z}}(X, L)$: suppose $f : X \to L$ is A-linear. Then, by definition, there is a set map $g : M \times N \to L$ such that the following diagram commutes:



[41] Note that h is in general far from being surjective.

[42] $\mathrm{Bil}_A(M, N, -)$ is pronounced "A-bilinear of M, N, blah."

[43] An alternate path is to define an A-module structure on $\mathbb{Z}[M \times N]$ and show that the quotient is stable under this action.

Then, the quotient relations force g to be bilinear.

Next, it's necessary to show that the map $\Phi : \mathrm{Hom}_A(X, L) \to \mathrm{Bil}(M, N; L)$ is a bijection. Since $f \in \mathrm{Hom}(X, L)$ is determined by its value on the generators $m \otimes n$, then Φ must be injective, for if two functions f_1 and f_2 yield the same $g : M \times N \to L$, then f_1 and f_2 must agree on these generators and thus be identical.

For surjectivity, take the quotient, which isn't quite trivial. Let $\Psi : \mathrm{Bil}_A(M, N; L) \to \mathrm{Hom}_A(X, L)$ be given by taking a bilinear $g : M \times N \to L$, extending it linearly to $\mathbb{Z}[M \times N] \to L$, and then check that it factors through the quotient R (i.e. R is killed by this extension of g). Then, one can directly check that Φ and Ψ are inverses.

Finally, it remains to check that this is functorial, i.e. that it's compatible with maps $L_1 \to L_2$. This is not a hard exercise. ⊠

12. More Tensor Products: 10/30/13

> *"If you really want to impress your friends and confound your enemies, you can invoke* tensor products *in [the context of Fourier series]... People run in terror from the ⊗ symbol. Cool." – Brad Osgood*

Recall that the tensor product $M \otimes_A N$ is the unique object in the category of A-modules such that $\mathrm{Hom}_A(M \otimes_A N, -) = \mathrm{Bil}_A(M \times N, -)$. From this definition, one can deduce some functorial properties of $\otimes$: for example, if $f : M_1 \to M_2$ and $g : N_1 \to N_2$, then they induce a map $f \otimes g : M_1 \otimes_A N_1 \to M_2 \otimes_A N_2$, because this is the same as a bilinear map $M_1 \times N_1 \to M_2 \otimes_A N_2$, given by $(x, y) \mapsto f(x) \otimes g(y)$ (though it remains for the spectators to check that this map is indeed A-bilinear). Then, from the definition, and in particular *not* the construction, one obtains the map $f \otimes g : x \otimes y \mapsto f(x) \otimes g(y)$.

Proposition 12.1. *There is a canonical isomorphism $M \otimes_A N \to N \otimes_A M$.*

Proof. Let $\alpha : M \times N \to N \otimes_A M$ be given by $(x, y) \mapsto y \otimes x$, which induces a $u : M \otimes_A N \to N \otimes_A M$ (it remains to check that α is bilinear, but this isn't hard). Similarly, let $\beta : N \times M \to M \otimes_A N$ send $(y, x) \mapsto x \otimes y$, inducing a $v : N \otimes_A M \to M \otimes_A N$.

Then, $v \circ u : M \otimes_A N \to N \otimes_A M \to M \otimes_A N$ sends $x \otimes y \mapsto y \otimes x \mapsto x \otimes y$, since it comes from $(x, y) \mapsto y \otimes x \mapsto x \otimes y$, so $v \circ u = \mathrm{id}$, and similarly for $u \circ v$. ⊠

This proof used something of the knowledge of the construction of $M \otimes_A N$, i.e. that it is generated by simple tensors. But there's a more categorical argument: $M \otimes_A N$ and $N \otimes_A M$ represent the same functor $\mathrm{Bil}_A(M \times N, -) = \mathrm{Bil}_A(N \times M, -)$, since the notion of a bilinear map doesn't depend on the order of the indices, and $M \times N \cong N \times M$ canonically. Thus, by Yoneda's lemma, the objects representing them, $M \otimes_A N$ and $N \otimes_A M$, must be canonically isomorphic.

Here, one uses Yoneda's lemma to say that $\mathcal{C}^{\mathrm{op}} \to \mathrm{Fun}(\mathcal{C}, \underline{\mathrm{Set}})$ sending $X \mapsto h_X : Y \to \mathrm{Mor}_{\mathcal{C}}(X, Y)$ is fully faithful, so the sets of morphisms are the same. Thus, if $X_1 \mapsto h_1$ and $X_2 \mapsto h_1$, where $h_1 \cong h_2$, then this isomorphism will correspond to some isomorphism $X_1 \to X_2$. In other words, there are maps in both directions between h_1 and h_2, and these pull back to maps between X_1 and X_2. These are u and v in the above proof; the image of $u \circ v$ under h is the identity, and of course $\mathrm{id} \xrightarrow{\sim} \mathrm{id}$, which is among the criteria for a fully faithful functor. Moreover, since there's a bijection, *only* the identity is sent to the identity.

Proposition 12.2. *Suppose M is a free A-module with basis S and N is a free A-module with basis T. Then, $M \otimes_A N$ is a free A-module with basis indexed by $S \times T$.*

Proof. Since M is free, then $\mathrm{Hom}_A(M, L) = \mathrm{Maps}(S, L)$; that is, every A-linear map out of M is uniquely determined by what it does on the basis elements, and any set map $S \to L$ induces a unique A-linear homomorphism. This follows from the universal property of free modules. A similar property applies to bilinear maps: $\mathrm{Bil}_A(M \times N, L) = \mathrm{Maps}(S \times T, L) = \mathrm{Hom}_A(F_{S \times T}, L)$, where $F_{S \times T}$ is the free A-module with a basis indexed by $S \times T$.

Thus, for any target L, these have the same Hom space, so the two modules must be the same, and thus $M \otimes_A N \cong F_{S \times T}$ canonically. ⊠

In particular, if S and T are finite sets, then $M \cong A^m$ and $N \cong A^n$, so $M \otimes_A N \cong A^{mn}$.

Proposition 12.3. *If $M = A/I$, then $(A/I) \otimes_A N \cong N/IN$, where*

$$IN = \left\{ \sum_{i=1}^{m} a_i n_i \mid a_i \in I, n_i \in N \right\},$$

i.e. the submodule generated by elements of the form an, with $a \in I$ and $n \in N$.

Proof. This can be checked by the definition of the tensor product: a bilinear map $A/I \times N \xrightarrow{f} L$ requires specifying $(1.x) \mapsto f(1,x) \in L$. Additionally, for any $a \in I$, $f(a,x) = f(0,x) = 0$, but on the other hand, $f(a,x) = af(1,x)$. Thus, $f(1,x)$ is killed by I.

Conversely, suppose that $g : N \to L$ is such that $g(x)$ is killed by I for all x. Then, g uniquely extends to a bilinear map $A/I \times N \to L$ by sending $(a,x) \mapsto a \cdot g(x)$.[44] This formula is forced because it must be bilinear: $(1,x) \mapsto g(x)$ and the rest must follow.

The conclusion is that $\mathrm{Bil}_A(A/I \times N, L) = \mathrm{Hom}_A(N, L[I])$, where $L[I]$ is the submodule of L that is killed by I. Thus, this is also equal to $\mathrm{Hom}_A(N/NI, L)$, because any $g : N \to L[I]$ is such that $\mathrm{Im}(g)$ is killed by I, so it goes to zero in the quotient, and thus $g(x \bmod NI)$ is always well-defined. Conversely, any $g : N/IN \to L$ can be viewed as a map $\overline{g} : N \to L$, but then it's easy to check that $\mathrm{Im}(g)$ is killed by I.

Thus, $A/I \otimes_A N \cong N/IN$ canonically, because they represent the same functor $\mathrm{Hom}_A(N, L[I])$. ⊠

In textbooks, one often sees a more explicit proof; creating such a map is not difficult, but checking the isomorphisms is a bit uglier, relying on the concrete construction.

Exactness. One often uses short exact sequences to compute the tensor product of more complicated modules in terms of simpler ones. For example, in $M' \xrightarrow{f} M \xrightarrow{g} M'' \to 0$, M might be more complicated than M' and M''.

Theorem 12.4. *If $M' \xrightarrow{f} M \xrightarrow{g} M'' \to 0$ is an exact sequence, then*

$$M' \otimes_A N \xrightarrow{f \otimes \mathrm{id}_N} M \otimes_A N \xrightarrow{g \otimes \mathrm{id}_N} M'' \otimes_A N \longrightarrow 0$$

is still exact.

A functor that obeys this is called right exact; thus, $- \otimes_A N$ is a right exact functor from A-modules to A-modules. There exist examples for which it doesn't preserve injectivity.

Lemma 12.5. *If $0 \to \mathrm{Hom}_A(X'', Y) \to \mathrm{Hom}_A(X, Y) \to \mathrm{Hom}_A(X', Y)$ is exact, then so is $X' \to X \to X'' \to 0$.*

This proof will also be on the homework, and states that the Hom functor transforms a right exact sequence into a left exact sequence.

Proof of Theorem 12.4. It suffices to show that $\mathrm{Hom}_A(M^*, N)$ is left exact, but since these objects are just sets of bilinear maps, then the goal is to check that

$$0 \longrightarrow \mathrm{Bil}_A(M'' \times N, Y) \longrightarrow \mathrm{Bil}_A(M \times N, Y) \longrightarrow \mathrm{Bil}_A(M' \times N, Y)$$

is left exact. This, however, is relatively easy. For example, to test exactness in the middle, if $\varphi : M \times N \to Y$ goes to 0 in $\mathrm{Bil}_A(M' \times N, Y)$, then $M' \times N \to M \times N \to Y = 0$, so $\varphi|_{\mathrm{Im}(f) \times N} = 0$ as well. Thus, φ induces a bilinear map $M/\mathrm{Im}(f) \times N \to Y$. However, $M/\mathrm{Im}(f) = M''$, so $\varphi \in \mathrm{Im}(f)$, and so on. ⊠

Partial proof of Lemma 12.5. A similar line of reasoning takes care of the lemma. For the exactness in the middle, suppose $\varphi : X \to Y$ is sent to zero in $X' \to Y$. Then, $X' \xrightarrow{\alpha} X \xrightarrow{\varphi} Y = 0$ given by composition, so $\varphi|_{\mathrm{Im}(\alpha)} = 0$, and thus φ induces a map $\varphi : X/\mathrm{Im}(\alpha) \to Y$, and thus a map $X' \to Y$.

In general, the tensor product is not left exact: consider $0 \to \mathbb{Z} \xrightarrow{\cdot 2} \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$, and let N be any abelian group. Then, the induced sequence $\mathbb{Z} \otimes_{\mathbb{Z}} N \to \mathbb{Z} \otimes_Z N \to \mathbb{Z}/2\mathbb{Z} \otimes_Z N$ simplifies to

$$N \xrightarrow{\cdot 2} N \xrightarrow{\bmod 2} N/2N \longrightarrow 0,$$

since $A \otimes_A N = N$ for any A-module N (using Proposition 12.3, with $I = (0)$). But this first map isn't always injective, since there's a problem whenever N has 2-torsion (e.g. just take $N = \mathbb{Z}/2$).

Later, it will be possible to measure the failure of this exactness using homological algebra.

Definition. An A-module N is called flat if $- \otimes_A N$ is an exact functor, i.e. for any short exact sequence $0 \to M' \to M \to M'' \to 0$, the resulting sequence $0 \to M' \otimes N \to M \otimes N \to M'' \otimes N \to 0$ is also short exact.

Really, the only condition is that $M' \otimes N \hookrightarrow M \otimes N$, since the rest holds for all modules N. Thus, N is flat iff for any injective map $M' \hookrightarrow M$, the resulting map $M' \otimes N \to M \otimes N$ is also injective.

Though it won't be proven here, it's easy to check that if A is a domain and M is a flat A-module, then M is torsion-free. This argument generalizes the previous example with $\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}$. However, there's no generalization: the notion of a torsion-free A-module makes no sense if A isn't a domain.

Examples of flat modules include:

[44] Technically, it should be $(a,x) \mapsto \tilde{a} \cdot g(x)$, where $\tilde{a}$ is a preimage of $a \in A/I$, but this is not a huge deal because it's independent of a choice of $\tilde{a}$.

- Free modules of finite rank, because $M \otimes_A A^n \cong M^n$, so injections are obviously preserved. This is still true when the free module has infinite rank, though this requires more of an argument.
- Another important class of flat A-modules is the localizations $S^{-1}A$ for any multiplicative set $S \subset A$.

To check the last item in particular, one needs the following lemma, though it's important in its own right.

Lemma 12.6. *Let $S \subset A$ be a multiplicative set and M be an A-module. Then, $S^{-1}M \cong S^{-1}A \otimes_A M$, where the isomorphism is as $S^{-1}A$-modules (and therefore also as A-modules).*

This can be checked by appealing to the functorial definition, because $\mathrm{Hom}_{S^{-1}A}(S^{-1}M, L) = \mathrm{Hom}_A(M, L)$, which we discussed before. Thus, there is a pair of adjoint functors between the category of A-modules and $S^{-1}A$-modules, which when boiled down to the definition of localization gives an equality.

Assuming the lemma, we can show that the localization of A at S is flat, because $- \otimes_A S^{-1}A = S^{-1}(-)$, and the latter is known to be exact. Thus, $S^{-1}A$ is flat as an A-module.

For example, one can localize from $\mathbb{Z}$ to $\mathbb{Q}$, so $\mathbb{Q}$ is a flat $\mathbb{Z}$-module, though it's not free. In some cases, all flat modules are free, as in local rings, but not all cases.

13. Tensor, Symmetric, and Skew-Symmetric Algebras: 11/1/13

"Where is everyone? The time change is on Sunday, right?"

Last time, it was stated without proof that $S^{-1}M \cong S^{-1}A \otimes_A M$. This can be proven in a more general context: suppose $A \xrightarrow{f} B$, so that B can be viewed as an A-algebra; in particular, B is an A-module. Then, if M is another A-module, then $B \otimes_A M$ is *a priori* an A-module, but also has a B-module structure, with action of a $b \in B$ given by $\mathrm{act}(b) : (b_1, m) \mapsto bb_1 \otimes m$. By the definition of tensor product, this is A-linear, but the action must give the structure of a B-module (basically, ensuring that it's compatible with multiplication within B), so it's necessary to check this.

The functor $B \otimes_A - : A\text{-Mod} \to B\text{-Mod}$ is called extension of scalars (from A to B; since A and B act on the modules, then they are called scalars, and often, B is larger than A). There's also the forgetful functor $B\text{-Mod} \to A\text{-Mod}$, which is called restriction of scalars.

Proposition 13.1. *$B \otimes_A -$ is left adjoint to restriction of scalars.*

Proof sketch. The goal is to show that $\mathrm{Hom}_B(B \otimes_A M, N) \cong \mathrm{Hom}_A(M, N)$. The correspondence is described as:

$$\varphi : B \otimes_A M \to N \longmapsto \psi : M \to N \text{ given by } m \mapsto \varphi(1 \otimes m).$$

On the right, $\psi : M \to N \mapsto \widetilde{\psi} : B \times M \to N$ is A-bilinear, so it extends to an A-linear $B \otimes_A M \to N$. It remains to check these are inverses of each other, etc. ⊠

Proposition 13.2. *$S^{-1}M \cong S^{-1}A \otimes_A M$ as A-modules.*

Proof. For any $S^{-1}A$-module N, $\mathrm{Hom}_{S^{-1}A}(S^{-1}M, N) = \mathrm{Hom}_A(M, N)$. Recall that $S^{-1}(-)$ is left adjoint to the forgetful functor $S^{-1}A\text{-Mod} \to A\text{-Mod}$, but $S^{-1}A \otimes_A -$ is also left adjoint to the same forgetful functor. Thus, $\mathrm{Hom}_{S^{-1}A}(S^{-1}A \otimes_A M, N) = \mathrm{Hom}_A(M, N)$. Thus, $S^{-1}M$ and $S^{-1}A \otimes_A M$ represent the same functor $S^{-1}A\,\mathrm{Mod} \to \mathrm{Set}$ given by $N \mapsto \mathrm{Hom}_A(M, N)$. Since they give the same Hom space, they must be canonically isomorphic as A-modules. ⊠

This is a useful test: to show that two objects are isomorphic, one can show that they represent the same functor in some category.

Tensor Product of Algebras. Suppose A is a commutative ring, and B and C are commutative A-algebras.

Proposition 13.3. *$B \otimes_A C$, which is a priori an A-module, also has a natural ring structure, giving a ring homomorphism $A \to B \otimes_A C$ that makes it an A-algebra.*

In effect, this just says that the tensor product of two rings is still a ring. The proof of this proposition will be deferred while multi-tensors are mentioned. This is a straightforward generalization of the tensor product outlined in the exercises; let $M_1, \dots, M_r, N$ be A-modules and $L_A(M_1 \times \cdots \times M_r; N)$ be the set of A-multilinear functions $M_1 \times \cdots \times M_r \to N$.

Claim. $L_A(M_1 \times \cdots \times M_r, -)$ is representable, and the object that represents it is denoted $M_1 \otimes_A \cdots \otimes_A M_r$.

In other words, for any A-module N, $\mathrm{Hom}_A(M_1 \otimes_A \cdots \otimes_A M_r) = L_A(M_1 \times \cdots \times M_r, N)$.

From this definition, we can deduce the associativity of the tensor product: $M_1 \otimes_A (M_2 \otimes_A M_3) \cong (M_1 \otimes_A M_2) \otimes_A M_3 = M_1 \otimes_A M_2 \otimes_A M_3$. This is true because all 3 represent the same functor, which sends an A-module L to the set of trilinear maps $M_1 \times M_2 \times M_3 \to L$. This generalizes in the reasonable way to greater numbers of factors, and means that parentheses in the construction of the multi-tensor product are unimportant. Since each of the three

constructions is slightly different, it is not completely tautological to show that the functors are the same; but it is not particularly difficult.

Proof of Proposition 13.3. Define a map $B \times C \times B \times C \to B \otimes_A C$ by $(b_1, c_1, b_2, c_2) \mapsto b_1 b_2 \otimes c_1 c_2$. It's clear that this is A-linear in each argument, so this is an A-quadrilinear function. Thus, this induces an A-linear map $B \otimes_A C \otimes_A B \otimes_A C \xrightarrow{\mu} B \otimes_A C$. By the associativity of the tensor product, this is equal to the map $(B \otimes_A C) \otimes_A (B \otimes_A C) \xrightarrow{\mu} B \otimes_A C$. Thus, again using the universal property, there is a unique bilinear map ψ such that the following diagram commutes:

$$\begin{array}{ccc} (B \otimes_A C) \times (B \otimes_A C) & \xrightarrow{\psi} & B \otimes_A C \\ \downarrow{\scriptstyle (b,c) \mapsto b \otimes c} & & \downarrow{\scriptstyle \wr} \\ (B \otimes_A C) \otimes_A (B \otimes_A C) & \longrightarrow & B \otimes_A C \end{array}$$

This induces the required ring structure, though it's necessary to check that this actually is in fact associative and commutative, and that $1 \otimes 1$ is the unit. Finally, one will have to show that the ring homomorphism making $B \otimes_A C$ into an A-algebra is $a \mapsto a \otimes 1 = 1 \otimes a$. (Since this is a tensor product over A, scalars can be pushed around like this.) ⊠

Though it's possible to write down a ring homomorphism and check it, this fuller argument cleanly ensures that it works on sums of simple tensors.

Example 13.1. If A is an abelian group, then $A \otimes_{\mathbb{Z}} \mathbb{Z}[x_1, \dots, x_n] \cong A[x_1, \dots, x_n]$ (in the exercises, this was checked as groups, but it also holds true for rings). It's also true that $A[x] \otimes_A A[y] \cong A[x, y]$.

Be careful, though: $A \otimes_{\mathbb{Z}} \mathbb{Z}[[x]] \to A[[x]]$ is in general not surjective. In essence, this is because finite linear combinations don't map to infinite power series very well. This is because $\sum a_i \otimes f_i$ maps to something whose coefficients are all linear combinations of the $a_1, \dots, a_n$. Thus, to show that surjectivity fails, pick a power series whose set of coefficients isn't finitely generated, e.g. if $A = \mathbb{Q}$, $\sum_{n \geq 1} x^n / n \notin \mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}[[x]]$.

Tensor Algebras.

Definition. Given an A-module M, the tensor algebra of M is

$$T(M) = A \oplus M \oplus (M \otimes_A M) \oplus (M \otimes_A M \otimes_A M) \oplus \cdots = A \oplus \bigoplus_{n=1}^{\infty} M^{\otimes n}.$$

Here, the notation $M^{\otimes n}$ refers to $M \otimes_A \cdots \otimes_A M$, where there are n terms in the tensor product.

Claim. $T(M)$ is naturally an associative A-algebra.

Proof sketch. Define multiplication as follows: for an $x_1 \otimes \cdots \otimes x_m \in M^{\otimes m}$ and a $y_1 \otimes \cdots \otimes y_n \in M^{\otimes n}$, their product is

$$(x_1 \otimes \cdots \otimes x_m)(y_1 \otimes \cdots \otimes y_n) = x_1 \otimes \cdots \otimes x_m \otimes y_1 \otimes \cdots \otimes y_n \in M^{\otimes (m+n)},$$

giving an A-bilinear map $M^{\otimes m} \otimes M^{\otimes n} \to M^{\otimes (m+n)}$, though it's necessary to check that this multiplication behaves well on sums of tensors, and in particular that it is well-defined.

Since every element in $T(M)$ is a finite linear combination of these, one obtains a map $T(M) \times T(M) \to T(M)$ with the required properties. ⊠

This algebra is associative (as are all algebras in this class), but it is *not* commutative: in particular, looking just at $M \times M \to M^{\otimes 2}$, $(x, y) \mapsto x \otimes y \neq y \otimes x$ in general.

Proposition 13.4. *Let R be an associative A-algebra. Then,* $\mathrm{Hom}_{A\text{-Alg}}(T(M), R) \cong \mathrm{Hom}_A(M, R)$*, and the isomorphism is canonical. That is, the forgetful functor A-Alg $\to A$-Mod has T as its left adjoint.*

Proof sketch. Again, build maps in both directions.

Suppose $\varphi : T(M) \xrightarrow{A\text{-Alg}} R$; then, take $\psi = \varphi|_M : M \to R$, which is clearly A-linear.

Conversely, if $\psi : M \to R$ is A-linear, first define $\varphi_n : M \times \cdots \times M \to R$ by $(x_1, \dots, x_n) \mapsto \psi(x_1)\psi(x_2) \cdots \psi(x_n)$, which is multilinear, so it extends to $\widetilde{\psi}_n : M^{\otimes n} \to R$. Then, let $\varphi = \bigoplus \widetilde{\psi}_n$ (i.e. $\varphi|_{M^{\otimes n}} = \tilde{\psi}$). This is *a priori* A-linear, but it remains to be checked that it's also a ring homomorphism, and then that these two maps are inverses of each other. ⊠

Symmetric Algebras. There's a similar universal property that gives a commutative A-algebra called the symmetric algebra. Let $\mathcal{C}$ denote the category of commutative A-algebras.

Proposition 13.5. *The forgetful functor $\mathcal{C} \to A$-Mod admits a left adjoint, denoted S, i.e. for every A-module M, there exists a commutative A-algebra $S(A)$ such that $\mathrm{Hom}_{\mathcal{C}}(S(M), B) = \mathrm{Hom}_A(M, B)$ (where the forgetfulness of B is assumed in the right side).*

Proof sketch. The construction is given by quotienting $T(M)$ by the universal relations that force the quotient to be commutative. This is a general procedure: if R is a ring, then $[R, R] = \mathrm{Span}\{xy - yx \mid x, y \in R\}$, so that $[R, R]$ is a two-sided ideal (which remains to be checked). If you're lucky, then $1 \notin [R, R]$, so one can quotient to obtain $R^{\mathrm{ab}} = R/[R, R]$, which is a commutative ring (if $1 \in [R, R]$, then the quotient is zero, which is sad). This is the smallest set of relations that force commutativity, so apply this procedure to $T(M)$: take the ideal generated by $x \otimes y - y \otimes x$ for all $x, y \in M$, and quotient by it.[45] ⊠

From the construction, it's easy to show that $\mathrm{Hom}_{\mathcal{C}}(S(M), B) = \mathrm{Hom}_A(M, B)$, because Proposition 13.4 applies to B as a (not necessarily commutative) A-algebra: $\mathrm{Hom}_{A\text{-Alg}}(T(M), B) = \mathrm{Hom}_A(M, B)$. So it only remains to check that $\mathrm{Hom}_{A\text{-Alg}}(T(M), B) = \mathrm{Hom}_{\mathcal{C}}(S(M), B)$, i.e., that such a map factors through the quotient, which is not that bad.

The most important examples of a symmetric algebra is the polynomial ring.

Claim. Suppose M is a free A-module of rank n. Then, $S(M) = A[x_1, \dots, x_n]$.

Proof sketch. Pick a basis of M, as $M = Ae_1 \oplus \cdots \oplus Ae_n$. Given some A-linear map $M \to A[x_1, \dots, x_n]$, that sends $e_i \mapsto x_i$ (which extends uniquely to an A-linear map because M is free), this is by Proposition 13.5 equivalent to $\varphi : S(M) \to A[x_1, \dots, x_n]$.

In the other direction, consider the map $A[x_1, \dots, x_n] \to T(M)$ given by

$$\prod_{i=1}^{n} x_i^{\ell_i} \longmapsto \bigotimes_{i=1}^{n} \bigotimes^{\ell_i} x_i \in M^{\otimes(\ell_1 + \cdots + \ell_n)}.$$

Taking the quotient, this yields a map $\psi : A[x_1, \dots, x_n] \to S(M)$, and it can be shown this is an inverse to φ. ⊠

This can be generalized; if M is a free A-module with basis S, then $S(M)$ is the set of polynomials over A in $|S|$ variables, even if M isn't finite-dimensional. Similarly, $T(A^n) = A\langle x_1, \dots, x_n \rangle$, and a similar result holds for the infinite-dimensional case.

Skew-symmetric Algebras. Note: the rest of this lecture was stated to be about exterior algebras and then later corrected to what follows. Skew-symmetric algebras aren't as important as exterior algebras (the proper definition will be provided in the next lecture). That said, what is said below is mostly still true.

Given an A-module M, one wants a universal, graded, anti-commutative A-algebra $\Lambda_{\mathrm{skew}}(M)$. This will end up being nearly commutative, and is also given by a quotient of the tensor algebra:

$$\Lambda_{\mathrm{skew}}(M) = T(M)/(x \otimes y + y \otimes x \mid x, y \in M)$$

(i.e. taking linear combinations of these generators), which is *not* commutative. The image of an $x \otimes y \in T(M)$ in $\Lambda_{\mathrm{skew}}(M)$ is denoted $x \wedge y$, and the ideal forces the relation $x \wedge y + y \wedge x = 0$ (there was in general no relation in $T(M)$), or, equivalently, $x \wedge y = -y \wedge x$. This is what is meant by anti-symmetric or anti-commutative.

More generally, to switch the order of pure wedges, the sign doesn't always change: given an $x_1 \wedge \cdots \wedge x_k, y_1 \wedge \cdots \wedge y_\ell \in \Lambda_{\mathrm{skew}}(M)$, $x_1 \wedge \cdots \wedge x_k$ is the image of $x_1 \otimes \cdots \otimes x_k \in T(M)$ (and similarly for the other term). Multiplication in $\Lambda_{\mathrm{skew}}(M)$ is also denoted $\wedge$, but this isn't actually ambiguous, because $(x_1 \wedge \cdots \wedge x_k) \wedge (y_1 \wedge \cdots \wedge y_\ell) = x_1 \wedge \cdots \wedge x_k \wedge y_1 \wedge \cdots \wedge y_\ell$ (since this is an associative algebra), then this is OK. However, to switch the order the sign might have to change:

$$(x_1 \wedge \cdots \wedge x_k) \wedge (y_1 \wedge \cdots \wedge y_\ell) = (-1)^{k\ell}(y_1 \wedge \cdots \wedge y_\ell) \wedge (x_1 \wedge \cdots \wedge x_k),$$

because $k\ell$ pairs need to be exchanged and each flips the sign. Thus, even numbers of wedges commute with everything.

The construction of $\Lambda_{\mathrm{skew}}(M)$ is universal in the category of graded, skew-symmetric A-algebras, and satisfies the same adjoint property, which is further developed in the exercises.

The notion of exterior algebra has applications to the real world: every topological space X has an associated ring $H^*(X)$ called the cohomology ring, which is a graded anti-commutative algebra. Thus, this category is useful. In fact, sometimes the cohomology ring is an exterior power, e.g. the n-dimensional torus T^n has $H^*(T^n) = \Lambda(\mathbb{Z}^{\oplus n})$.

[45] These lie in $M^{\otimes 2}$, but by left- or right-multiplying by other things, one obtains all of the other necessary elements.

14. Exterior Algebras: 11/6/13

"Old McDonald had a form, $e_i \wedge e_i = 0$."

The correct definition of an exterior algebra is as follows:

Definition. Let I be the ideal of $T(M)$ generated by all elements of the form $x \otimes x$ for $x \in M$. Then, the exterior algebra of an A-module M is $\Lambda(M) = T(M)/I$.

The resulting A-algebra is the largest quotient of the tensor algebra that has elements of the form $x \otimes x \to 0$.

The skew-symmetric algebra is slightly different: the ideal given was $I_{\text{skew}} = (x \otimes y - y \otimes x \mid x, y \in M)$. Then, $I_{\text{skew}} \subset I$, because $x \otimes y + y \otimes x = (x+y) \otimes (x+y) - x \otimes x - y \otimes y$. Recall that this is akin to the discussion on bilinear forms, where in some cases, alternating and skew-symmetric forms aren't the same. But at least we have that $\Lambda_{\text{skew}}(M) \twoheadrightarrow \Lambda(M)$.

Example 14.1. For a pathological example, let A be a field of characteristic 2. Then, since addition and subtraction are the same thing, I_{skew} is equal to the ideal generated by $x \otimes y - y \otimes x$ (since they're the same generating elements). Thus, $\Lambda_{\text{skew}}(M) = \text{Sym}_A(M)$, so it looks like a polynomial algebra, and is in particular infinite-dimensional when $M \neq 0$. However, when M is free of rank n, then $\Lambda(M)$ has rank 2^n.

Focusing on exterior algebras, since $T(M)$ and I are both graded, then the notion of degree still exists in the quotient $\Lambda(M)$, so one has $\Lambda^i(M) = \{x_1 \wedge \cdots \wedge x_i \mid x_1, \ldots, x_i \in M\}$, implying that $\Lambda(M) = \bigoplus_{i \geq 0} \Lambda^i(M)$.

Lemma 14.1. *Suppose that M is a free A-module of rank n; then,*

$$\{x_{i_1} \wedge \cdots \wedge x_{i_k} \mid 1 \leq i_1 < \cdots < i_k \leq n\} \tag{3}$$

generate $\Lambda^k(M)$ as an A-module.

Proof. Notice that by the relations in I, $x_i \wedge x_i = 0$ and $x_i \wedge x_j = -x_j \wedge x_i$. $M^{\otimes k}$ has the basis $\{x_{i_1} \otimes \cdots \otimes x_{i_k} \mid 1 \leq i_1, \ldots, i_k \leq n\}$ (the ordering is arbitrary, so there are n^k elements). After the quotient $M^{\otimes k} \to \Lambda^k(M)$,

$$x_{i_1} \otimes \cdots \otimes x_{i_k} \longmapsto \begin{cases} 0, & \text{if there is a repetition in the } (i_1, \ldots, i_n) \\ \pm x_{j_1} \wedge \cdots \wedge x_{j_k}, & \text{where the } j_\ell \text{ are the } i_\ell \text{ in strictly ascending order.} \end{cases}$$

Thus, their image is (3), so this set does generate $\Lambda^k(M)$. ⊠

Lemma 14.2. *If M is a free A-module of rank n with basis $\{x_1, \ldots, x_n\}$, then $\det(M) = \Lambda^n(M)$ is free of rank 1, with basis $\{x_1 \wedge \cdots \wedge x_n\}$.*

Proof. By Lemma 14.1, $x_1 \wedge \cdots \wedge x_n$ generates $\det(M)$, so $a \overset{\iota}{\mapsto} a(x_1 \wedge \cdots \wedge x_n)$ is surjective. Now, define $\varphi : \Lambda^n(M) \to A$ by taking $\widetilde{\varphi} : M \times \cdots \times M \to A$ by $(v_1 \cdots, v_n) \mapsto \det(V)$, where V is the $n \times n$ matrix with v_i as its i^{th} column, for all i. Thus, this map induces a $\psi : M^{\otimes n} \to A$, so it's necessary to check that ψ factors through the quotient (the n^{th}-degree part I_n of I), i.e. $\psi(I_n) = 0$.

More concretely, $I_n = \{\sum \cdots \otimes x_i \otimes x_i \cdots\}$, or linear combinations of things where two neighbors are the same. But these all map to zero because the determinant of a matrix with two identical columns must be zero. Thus, ψ factors through I, so φ is well-defined. It remains to check that ψ and ι are inverses, but this is not difficult. ⊠

Note that invoking ι is unnecessary, since an injective map between free A-modules of the same dimension is automatically surjective. Also, the proof uses important properties of the determinant, but, oddly enough, not the fact that it's independent of basis.

Proposition 14.3. *Let M be a free A-module of rank n. Then, $\Lambda^k(M)$ is a free A-module of rank $\binom{n}{k}$.*

Proof. Pick a basis $x_1, \ldots, x_n$ of M. By Lemma 14.1, $\dim \Lambda^k(M) \leq \binom{n}{k}$, but it remains to show that these generators are linearly independent. To do this, observe that there is a pairing $\Lambda^i(M) \times \Lambda^{n-i}(M) \to \Lambda^n(M) : (\omega, \omega') \mapsto \omega\omega'$. Under this pairing, the basis elements map as follows:

$$(x_{i_1} \wedge \cdots \wedge x_{i_k}, x_{j_1} \wedge \cdots \wedge x_{j_{n-k}}) \longmapsto \begin{cases} \pm x_1 \wedge \cdots \wedge x_n, & \text{if } \{i_1, \ldots, i_k, j_1, \ldots, j_{n-k}\} \text{ has no repetitions} \\ 0, & \text{otherwise.} \end{cases}$$

Now, suppose

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} c_{i_1 \cdots i_k} x_{i_1} \wedge \cdots \wedge x_{i_k} = 0,$$

for some $c_{i_1 \cdots i_k} \in A$. Then, $\omega \wedge (x_{j_1} \wedge \cdots \wedge x_{j_{n-k}}) = \pm c_{i_1 \cdots i_k}(x_1 \wedge \cdots \wedge x_n)$ if the $x_{i_1}, \ldots, x_{i_k}, x_{j_1}, \ldots, x_{j_{n-k}}$ exhaust all of the basis elements of M, and is zero otherwise (i.e. there's some sort of repetition). But since $x_1 \wedge \cdots x_n$ is a basis for $\Lambda^n(M)$, then $\omega \wedge (x_{j_1} \wedge \cdots \wedge x_{j_{n-k}}) = 0$ forces $c_{i_1 \cdots i_k} = 0$, so the geneating set is linearly independent. ⊠

A useful corollary of this is that $\Lambda^k(M) = 0$ if $k > n$, so $\operatorname{rank}_A \Lambda(M) = 2^n$.

Definition. A pairing $(\cdot,\cdot) : M \times N$ is an A-bilinear map, which implies that $f_m = (m, -) : N \to A$ is automatically A-linear for each A (and similarly in the other argument); thus, one obtains a map $M \to N^\vee = \operatorname{Hom}_A(N, A)$ given by $m \mapsto f_m$. The pairing is called perfect[46] if this is an isomorphism.

This definition can use any rank-1 free A-module in place of A, which is useful because it allows things to be more independent of basis, especially because there is no canonical identification $\Lambda^n(M) \xrightarrow{\sim} A$.

One criterion for a perfect pairing is that if M and N are free A-modules of finite rank n, then an A-bilinear map $(\cdot,\cdot) : M \times N \to A$ is perfect iff there exist a basis $\{x_1, \dots, x_n\}$ of M and a basis $\{y_1, \dots, y_n\}$ of N such that $(x_i, y_j) = \delta_{ij}$.[47] This pair of bases is called a dual basis, and in the case of free A-modules, a perfect pairing is equivalent to the existence of such a dual basis.

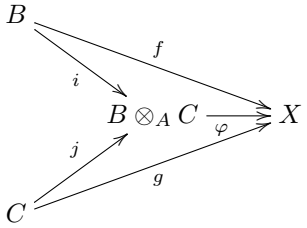
Proposition 14.4. *The pairing $\Lambda^k(M) \times \Lambda^{n-k}(M) \xrightarrow{\wedge} \Lambda^n(M)$ is perfect.*

To prove this, one would show that $\{x_{i_1} \wedge \cdots \wedge x_{i_k}\}$ and $\{x_{j_1} \wedge \cdots \wedge x_{j_{n-k}}\}$ form a dual basis under this pairing.

Theorem 14.5. *If B and C are commutative A-algebras, then $B \otimes_A C$ is the direct sum of B and C in the category of commutative A-algebras.*

Recall that thus means that:

(1) there are (not necessarily injective) A-algebra maps (i.e. A-linear ring homomorphisms) $i : B \to B \otimes_A C$ and $j : C \to B \otimes_A C$, and
(2) this is the universal solution to the diagram: whenever $f : B \to X$ and $g : C \to X$, then there exists a unique A-algebra homomorphism φ such that the following diagram commutes:



Once this is known, the proof isn't so bad, and the theorem is rather useful, as in the following example.

Proposition 14.6. *For any A-modules M and N, $\operatorname{Sym}_A(M \oplus N) \cong \operatorname{Sym}_A(M) \otimes_A \operatorname{Sym}_A(N)$, and the isomorphism is canonical.*

Proof sketch. This proof uses the typical strategy of showing both objects represent the same functor from the category $\mathcal{C}$ of commutative A-algebras to the category of sets, using the Yoneda lemma.

Let B be some other commutative A-algebra. Then, $\operatorname{Hom}_{\mathcal{C}}(\operatorname{Sym}(M \oplus N), B) = \operatorname{Hom}_A(M \oplus N, B)$, since Sym_A is left adjoint to the forgetful functor from $\mathcal{C}$ to the category of A-modules. But by the definition of the direct sum of modules, $\operatorname{Hom}_A(M \oplus N, B) = \operatorname{Hom}_A(M, B) \times \operatorname{Hom}_A(N, B)$. On the right, $\operatorname{Hom}_{\mathcal{C}}(\operatorname{Sym}(M) \otimes_A \operatorname{Sym}(N), B) = \operatorname{Hom}_{\mathcal{C}}(\operatorname{Sym}(M), B) \times \operatorname{Hom}_{\mathcal{C}}(\operatorname{Sym}_A(N), B) = \operatorname{Hom}_A(M, B) \times \operatorname{Hom}_A(N, B)$ by adjointness. Thus, the two are isomorphic, and the isomorphism is canonical because no basis was chosen. ⊠

Note that if M and N are free, this says things we already knew about polynomial rings.

One can do essentially the same thing with exterior algebras.

Proposition 14.7. *If M and N are A-modules, then $\Lambda(M \oplus N) \cong \Lambda(M) \otimes_A \Lambda(N)$ as A-algebras.*

The fact that one can put an A-algebra structure on $\Lambda(M) \otimes_A \Lambda(N)$ is true, but it's not immediate, since these two rings are noncommutative, since $(\omega_i \otimes \theta_j)(\omega_k \otimes \theta_\ell) = (-1)^{jk}(\omega_i \wedge \omega_k) \otimes (\theta_j \wedge \theta_\ell)$. This is called the Kozul sign convention, which allows this structure on A-algebras that are commutative only up to sign.

Proof sketch of Proposition 14.7. Effectively the same proof will be given as for Proposition 14.6, though this time $\mathcal{C}$ will denote the category of $\mathbb{Z}_{\geq 0}$-graded alternating algebras.[48] These are $B = \bigoplus_{n=0}^{\infty} B_n$, with $x^2 = 0$ for any $x \in B_1$ and $x_i x_j = (-1)^{ij} x_j x_i$ for any $x_i \in B_i$ and $x_j \in B_j$.

It turns out that $B \otimes_A B'$ is the direct sum in $\mathcal{C}$, because Λ is left adjoint to the forgetful functor $\mathcal{C} \to A$-Mod. This requires a proof, but it's basically the same words as the previous proposition.

[46] The notion of a "perfect pairing" is the cutest thing I've heard of in mathematics since the Gale-Shapley marriage algorithm.

[47] δ_{ij} means 0 if $i \neq j$ and 1 if $i = j$.

[48] This category doesn't seem to have a good catchy name, even though it appears in nature. In this proof, it's just cooked up to handle the specific proposition.

The direct product in each of these categories is just the direct product as rings.

The infinite direct sum exists in the category of A-algebras, and you can think about what this would look like. One can make an infinite tensor product, though it's much easier to work with the restricted version in which all but finitely many factors in a tensor are 1. In other words,

$$\bigotimes_{n=1}^{\infty} B_n = \varinjlim_n B_1 \otimes_A \cdots \otimes_A B_n,$$

where the directed map sends $x_1 \otimes \cdots \otimes x_n \mapsto x_1 \otimes \cdots \otimes x_n \otimes 1$. The infinite tensor product as defined above is the direct sum in the category of commutative A-algebras.

None of this stuff is too foreign; for example, if $B = A[x_i]$, the limit is just $A[x_1, x_2, \dots]$, the ring of polynomials in countably many variables.

Part 6. Homological Algebra

15. Complexes and Projective and Injective Modules: 11/8/13

The purpose of homological algebra is to discuss the non-exactness of functors. Many functors have already come up in this class, the most basic of which are $\mathrm{Hom}_A(X, -)$, $\mathrm{Hom}_A(-, X)$, and $X \otimes_A -$.

$\mathrm{Hom}_A(X, -)$ is left exact: that is, if $0 \to M' \to M \to M'' \to 0$ is a short exact sequence, then $0 \to \mathrm{Hom}_A(X, M') \to \mathrm{Hom}_A(X, M) \to \mathrm{Hom}_A(X, M'')$ is exact, but the last map is not necessarily surjective. Similarly, $\mathrm{Hom}_A(-, X)$ is a left exact contravariant functor; given the same short exact sequence above, we know only that $0 \to \mathrm{Hom}_A(M'', X) \to \mathrm{Hom}_A(M, X) \to \mathrm{Hom}_A(M', X)$ is exact. Tensoring is right exact: it preserves surjectivity but not always injectivity.

Homological algebra extends short inexact sequences into longer exact ones, which provide information about how far a map is from being injective or surjective.

Homological algebra happens in abelian categories, which don't need to be discussed in huge detail. Basically, it's necessary for finite direct sums to exist, as well as kernels and cokernels of morphisms. Here are some examples:

- If R is any associative algebra, then the category of left R-modules is abelian.
- If R is Noetherian, then the category of finitely generated R-modules is abelian; the Noetherian condition is necessary to force subobjects of finitely generated objects to be finitely generated.
- If G is a group, then one has the category of (finitely-generated) complex representations of G, i.e. $\mathrm{Rep}(G)$, the set of $\mathbb{C}$-vector spaces with a G-action, and with morphisms G-equivalent $\mathbb{C}$-linear maps.

We also need the notion of complexes in an abelian category.

Definition. Let $\mathcal{C}$ be an abelian category. Then, a complex $X^\bullet$ of $\mathcal{C}$ is a chain of maps

$$\cdots \longrightarrow X^{-2} \xrightarrow{f^{-2}} X^{-1} \xrightarrow{f^{-1}} X^0 \xrightarrow{f^0} X^1 \xrightarrow{f^1} X^2 \xrightarrow{f^2} \cdots$$

such that $X^i \in \mathcal{C}$, $f^i \in \mathrm{Hom}_{\mathcal{C}}(X^i, X^{i+1})$, and (most importantly) $f^i \circ f^{i-1} = 0$ for each i.

These can be infinite or finite in either direction (the latter corresponding to an infinite chain of zero objects and zero maps).

Definition. The i^{th} cohomology of the above complex $X^\bullet$ is $H^i(X^\bullet) = \ker(f^i)/\operatorname{Im}(f^{i-1})$.

$\ker(f^i)$ and $\operatorname{Im}(f^{i-1})$ are both $\mathcal{C}$-subobjects of X^i, and since $f^i \circ f^{i-1} = 0$, then $\operatorname{Im}(f^{i-1}) \subset \ker(f^i)$ and therefore the quotient makes sense (and the notion of abelian categories means it's possible to take quotients).

In summary, the cohomology is a sequence of (typically smaller) objects associated to some complex.

Definition. A map between two complexes $(X^\bullet, f^\bullet)$ and $(Y^\bullet, g^\bullet)$ is a set of maps $\phi^\bullet$ such that the following diagram commutes for each i:

$$\begin{array}{ccccccc}
\cdots & \longrightarrow & X^{i-1} & \xrightarrow{f^{i-1}} & X^i & \longrightarrow & \cdots \\
 & & \downarrow{\scriptstyle \phi^{i-1}} & & \downarrow{\scriptstyle \phi^i} & & \\
\cdots & \longrightarrow & Y^{i-1} & \xrightarrow[g^{i-1}]{} & Y^i & \longrightarrow & \cdots
\end{array}$$

This is also denoted $(X^\bullet, f^\bullet) \xrightarrow{\phi^\bullet} (Y^\bullet, g^\bullet)$.

Definition. Let $\phi^\bullet, \psi^\bullet : (X^\bullet, f^\bullet) \to (Y^\bullet, g^\bullet)$. Then, a homotopy between $\phi^\bullet$ and $\psi^\bullet$ is a sequence of maps h^i that shifts the grading:

$$\begin{array}{ccccccccccc}
\cdots & \longrightarrow & X^{i-1} & \longrightarrow & X^i & \xrightarrow{f^i} & X^{i+1} & \longrightarrow & \cdots \\
 & & \downarrow{\scriptstyle h^{i-1}} & & \downarrow{\scriptstyle h^i} & & \downarrow{\scriptstyle h^{i+1}} & & \downarrow{\scriptstyle h^{i+2}} & & \\
\cdots & \longrightarrow & Y^{i-1} & \xrightarrow[g^{i-1}]{} & Y^i & \longrightarrow & Y^{i+1} & \longrightarrow & \cdots
\end{array}$$

such that $h^{i+1} \circ f^i + g^{i-1} \circ h^i = \phi^i - \psi^i$.

This means that the above diagram doesn't commute, so be careful. Additionally, the sign convention could be $h^{i+1} \circ f^i - g^{i-1} \circ h^i$, and everything still works (albeit with different component maps h^i).

Lemma 15.1.

(1) $\phi^\bullet : (X^\bullet, f^\bullet) \to (Y^\bullet, g^\bullet)$ induces a map $H^i(\phi^\bullet) : H^i(X^\bullet) \to H^i(Y^\bullet)$.
(2) If there is a homotopy $h^\bullet$ between $\phi^\bullet$ and $\psi^\bullet$ (where both are $X^\bullet \to Y^\bullet$), then $H^i(\phi^\bullet) = H^i(\psi^\bullet)$ for all i, with the equality taken in $\mathrm{Mor}_{\mathcal{C}}(H^i(X^\bullet), H^i(Y^\bullet))$.

The intuition for the first part of the lemma is that $\ker(f^i) \to \ker(g^i)$ and $\mathrm{Im}(f^i) \to \mathrm{Im}(g^i)$, so the map factors through the quotient.

Definition. A complex $(X^\bullet, f^\bullet)$ is called acyclic if all cohomology objects vanish: $H^i(X^\bullet) = 0$ for all i. In this case, the chain of maps $\cdots \to X^{i-1} \xrightarrow{f^{i-1}} X^i \to \cdots$ is called a long exact sequence.

More generally, $(X^\bullet, f^\bullet)$ is exact at the i^{th} place if $H^i(X^\bullet) = 0$, or equivalently that $\mathrm{Im}(f^{i-1}) = \ker(f^i)$.

Short exact sequences thus become a special case of these more general exact sequences.

Lemma 15.2. *If there exist maps r_i such that*

$$\begin{array}{ccccccccccc}
\cdots & \longrightarrow & X^{i-1} & \longrightarrow & X^i & \longrightarrow & X^{i+1} & \longrightarrow & \cdots \\
 & & \downarrow & \swarrow{\scriptstyle r_i} & \downarrow & \swarrow{\scriptstyle r_{i+1}} & \downarrow & & \\
\cdots & \longrightarrow & X^{i-1} & \longrightarrow & X^i & \longrightarrow & X^{i+1} & \longrightarrow & \cdots
\end{array}$$

and $f^{i-1} \circ r_i + r_{i+1} \circ f^i = \mathrm{id}_X$ for all i, then $(X^\bullet, f^\bullet)$ is acyclic.

Proof. Using Lemma 15.1, (r_i) is a homotopy between $\mathrm{id}_{(X^\bullet, f^\bullet)}$ and 0. Thus, they induce the same maps on the cohomology: $H^i(\mathrm{id}) = H^i(0)$ as functions $H^i(X^\bullet) \to H^i(X^\bullet)$. However, $H^i(\mathrm{id}) = \mathrm{id}$ and $H^i(0) = 0$, so $H^i(X^\bullet) = 0$ (i.e. the zero object in $\mathcal{C}$). ⊠

Exercise 15.1. Since no proof of Lemma 15.1 was given, formulate a proof of Lemma 15.2 that doesn't depend on it.

The basic idea of calculations in homological algebra is to replace an arbitrary object $X \in \mathcal{C}$ with a complex of nicer objects in $\mathcal{C}$. This is called a resolution. Generally, one is given a "nice" class of objects $\mathcal{P} \subset \mathcal{C}$, such as projective, injective, or free modules, depending on context.

Definition. Given the above, a right resolution of X is a complex

$$0 \longrightarrow Y^0 \xrightarrow{f^0} Y^1 \xrightarrow{f^1} Y^2 \xrightarrow{f^2} \cdots$$

(i.e. $Y^i = 0$ if $i < 0$), such that $Y^i \in \mathcal{P}$, $X = H^0(Y^\bullet)$ (so that $X = \ker(f^0)$), and $H^i(Y^\bullet) = 0$ for all $i > 0$. In other words, the cohomology vanishes everywhere except at zero, where it is equal to X.

Similarly, a left resolution of X is a complex

$$\cdots \longrightarrow Y^{-2} \longrightarrow Y^{-1} \longrightarrow Y^0 \longrightarrow 0$$

such that $Y^i \in \mathcal{P}$, $H^0(Y^\bullet) = X$, and $H^i(Y^\bullet) = 0$ for all $i < 0$.

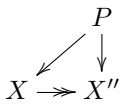
Projective Modules. Of course, these are fine definitions, but to put them to use one must have a good candidate for $\mathcal{P}$.

Definition. $P \in \mathrm{Obj}(\mathcal{C})$ is called projective if the functor $\mathrm{Mor}_{\mathcal{C}}(P, -) : \mathcal{C} \to \mathrm{Ab}$ is exact.[49]

[49]Here, Ab is the category of abelian groups. That $\mathrm{Mor}_{\mathcal{C}}(P, Q)$ is abelian is a consequence of the axioms of an abelian category. Also, in this lecture, Mor is used for general categories and Hom for modules, where $\mathrm{Hom}_A(M, N)$ has additional structure.

Since $\mathrm{Mor}_{\mathcal{C}}(P, -)$ is always left exact, the condition is equivalent to it sending a surjective map to a surjective map: $X \twoheadrightarrow X''$ must imply that $\mathrm{Hom}_{\mathcal{C}}(P, X) \twoheadrightarrow \mathrm{Hom}_{\mathcal{C}}(P, X'')$.

Another way of saying this is that whenever $X \twoheadrightarrow X''$, every map $P \to X''$ lifts to a (not necessarily unique) map $P \to X$ such that the following diagram commutes:



Example 15.1. Free A-modules are projective A-modules. Let F_S be free on a basis S (not necessarily finite). Then, an A-linear map $F_S \to X''$ is equivalent to a set map $S \to X''$, which can of course be lifted. Pick some arbitrary preimage $f : S \to X$, which induces a map $\widetilde{f} : F_S \to X$.

Lemma 15.3. *In $\mathcal{C} = A$-Mod, P is a projective A-module iff P is a direct summand in a free A-module.*

Proof. For any projective A-module P, there exists a free A-module F such that $F \overset{\pi}{\twoheadrightarrow} P$. Then, the map $\mathrm{id} : P \to P$ lifts to a section $s : P \to F$ of π, i.e. $\pi \circ s = \mathrm{id}_P$.

Using the inclusion maps given from the direct sum, there is an induced map $\ker(\pi) \oplus s(P) \overset{\sim}{\to} P$, which can be shown to be an isomorphism. The kernel is zero because the two factors don't intersect, which is guaranteed because $\pi \circ s = \mathrm{id}_p$. For surjectivity, write $x \in F$ as $s \circ \pi(x) + x - s(\pi(x))$, but the latter two terms fall in $\ker(\pi)$.

Conversely, if $F = P \oplus Q$, then suppose $P \oplus Q \to X''$ lifts to $P \oplus Q \to X$. Every map $P \oplus Q \to X$ is of the form (p, q) where $p : P \to X$ and $q : Q \to X$ are $\mathcal{C}$-morphisms, so the map $P \oplus Q \to X''$ is of the form (p, q), and it lifts to some maps $(\widetilde{p}, \widetilde{q})$. Thus, p lifts to $\widetilde{p}$, and so on.

Thus, for any $p : P \to X''$, one can extend it to $(p, 0) : P \oplus Q \to X''$, which extends to some $(\widetilde{p}, \widetilde{q}) : P \oplus Q \to X$. Thus, $\widetilde{p}$ is a lift $P \to X$, so P is projective. ⊠

Observe that this is harder to generalize to more general abelian categories, since the technicalities force the more specific use of A-modules.

Example 15.2. This all still works in the case of noncommutative algebras. Let $R = \mathrm{Mat}_n(k)$, so that $V = k^n$ is a left R-module.

Claim. V is a projective R-module, and in fact every finitely generated R-module is projective.

Proof. For a proof of the first part, $R \curvearrowright R$ is a free module of rank 1 (by matrix multiplication from the left), but R breaks into n pieces, one for each column, and each column is a direct summand: $R = \bigoplus V_i$, where V_i is the set of matrices whose only nonzero entries are in column i (and thus is R-stable), and each $V_i \cong V$. Thus, V is projective, and $R = V^{\oplus n}$.

The second part follows from the more interesting fact that every finitely generated R-module is a direct sum of copies of V. ⊠

Generally, in the commutative world, projective modules are larger than the ring itself, but as above, this isn't the case in the noncommutative world.

Exercise 15.2. If A is a local ring, then finitely generated projective A-modules are free.

Injective Modules. Dual to the notion of projective objects is that of injective objects.

Definition. $I \in \mathrm{Obj}(\mathcal{C})$ is called injective if $\mathrm{Hom}_{\mathcal{C}}(-, I) : \mathcal{C}^{\mathrm{op}} \to \mathrm{Ab}$ is exact.

A simple characterization of injective objects is that if $X' \hookrightarrow X$, then a map $f : X' \to I$ lifts to a map $X \to I$. Unfortunately, there's no explicit description akin to Lemma 15.3, so it's harder to construct injective A-modules. At this point, it's not even clear if they exist. Fortunately, though, there are lots of injective modules:

Theorem 15.4. *For any A-module X, there exists an inclusion $X \hookrightarrow I$, where I is an injective A-module.*

The proof is nontrivial and outlined in the exercises. A crucial step happens in which $\mathcal{C}$ is the category of abelian groups.

Definition. An abelian group I is called divisible if for every $n \in \mathbb{N}$, the map $I \overset{\times n}{\to} I$ is surjective, i.e. for every $i \in I$ and $n \in \mathbb{N}$, there exists an $i' \in I$ such that $ni' = i$.

This preimage is not necessarily unique; though $\mathbb{Q}$ is the most obvious example of a divisible group, $\mathbb{Q}/\mathbb{Z}$ is one in which the preimage of 0 is $1/n$.

Proposition 15.5. *Let I be an abelian group. Then, I is injective iff I is divisible.*

16. The Derived Functor $\mathrm{Ext}^i(M,N)$: 11/13/13

"When discussing homological algebra, it's important to be exact in one's speech."

Throughout this lecture, let A be a commutative ring and M and N be A-modules.

The functors $\mathrm{Ext}^i(M,N)$ (well, $\mathrm{Ext}^i_A(M,N)$ if you're being really meticulous) are derived functors of the Hom functor, which means that $\mathrm{Ext}^0(M,N) = \mathrm{Hom}(M,N)$ $\mathrm{Ext}^1(M,N)$ measures the failure of right exactness of Hom, etc. What this means precisely will be discussed in a bit.

Fix N, so that $\mathrm{Hom}(-,N)$ is a contravariant left exact functor $(\underline{A\text{-Mod}})^{\mathrm{op}} \to \mathrm{Ab}$.[50] The process of deriving a functor works for any contravariant left exact functor, though the specific details will be discussed primarily for the Hom functor.

Given some A-module M, pick a projective resolution for M, i.e. a complex

$$\cdots \longrightarrow P_2 \xrightarrow{f_2} P_1 \xrightarrow{f_1} P_0 \xrightarrow{f_0} M \longrightarrow 0$$

that is a long exact sequence, and such that each P_i is a projective A-module.[51]

Remark. This can always be done, because it's possible to do it with a free resolution, because there always exists a free module F_0 such that $F_0 \overset{f_0}{\twoheadrightarrow} M$ (given by taking the free module on the generators of M). Then, repeat the same thing on $\ker(f_0)$ to get a free module F_1 such that $F_1 \twoheadrightarrow \ker(f_0)$, and repeat; the maps f_i are given by composing through the kernel as in the following diagram:

$$\begin{array}{ccccccccc} \cdots \longrightarrow & F_2 & \xrightarrow{f_2} & F_1 & \xrightarrow{f_1} & F_0 & \xrightarrow{f_0} & M \\ & & \searrow & \uparrow & \searrow & \uparrow & & \\ & & & \ker(f_1) & & \ker(f_0) & & \end{array}$$

In some sense, first f_0 is found, then f_1 is calculated using its kernel, then f_2 from $\ker(f_1)$, and so on, working towards the left.

It is common to leave M out of the projective resolution entirely, since it can be quickly recovered as $\mathrm{Im}(f_0)$ and the result tends to be cleaner, as in the following definition. Note that this means the sequence is no longer exact at the last term, since $\mathrm{coker}(f_1) = M$.

Definition. Consider a projective resolution $\cdots \to P_2 \xrightarrow{f_2} P_1 \xrightarrow{f_1} P_0 \to 0$, let $\mathrm{Hom}(P_\bullet, M)$ denote the induced complex

$$0 \longrightarrow \mathrm{Hom}(P_0,N) \longrightarrow \mathrm{Hom}(P_1,N) \longrightarrow \cdots$$

The change in direction occurs because $\mathrm{Hom}(-,N)$ is contravariant. Then, define $\mathrm{Ext}^i(M,N) = H^i(\mathrm{Hom}(P_\bullet,N))$.

A priori, this definition depends on the choice of projective resolution for M. However, it will end up being independent of this choice in a way that will be clarified.

Lemma 16.1. *$\mathrm{Ext}^i(M,N)$ is independent of the choice of projective resolution of M in a canonical way.*

Proof. Suppose $P_\bullet$ and $Q_\bullet$ are both projective resolutions of M. Then $P_0, Q_0 \twoheadrightarrow M$, so since P_0 is projective, the map $P_0 \twoheadrightarrow M_0$ lifts to a map $\phi_0 : P_0 \to Q_0$ so that the following diagram commutes:

$$\begin{array}{ccccccccc} \cdots \longrightarrow & P_2 & \xrightarrow{f_2} & P_1 & \xrightarrow{f_1} & P_0 & \xrightarrow{f_0} & M_0 \\ & & & & & \downarrow{\scriptstyle \phi_0} & & \| \\ \cdots \longrightarrow & Q_2 & \xrightarrow{g_2} & Q_1 & \xrightarrow{g_1} & Q_0 & \xrightarrow{g_0} & M_0 \end{array}$$

The next connecting morphism is a little trickier. Since these sequences are exact, then $\phi_0 : \ker(f_0) \to \ker(g_0)$ works just as well, so one can lift $\phi_0 \circ f_1$ to a map $\phi_1 : P_1 \to Q_1$ as in the following diagram:

$$\begin{array}{ccc} P_1 & \xrightarrow{f_1} & \ker(f_0) \\ \downarrow{\scriptstyle \phi_1} & \searrow^{\phi_0 \circ f_1} & \downarrow{\scriptstyle \phi_0} \\ Q_1 & \xrightarrow{g_1} & \ker(g_0) \end{array}$$

[50] Since the Hom space has a canonical A-module structure, this can be given as a functor into the category of A-modules again.

[51] Notice that unlike previous complexes, this one uses subscripts rather than superscripts, and increases to the left, rather than to the right. This is a common convention: that X_i is in degree $-i$.

Thus, the filling exists. This can be inductively extended to all i, so the filling (ϕ_i) exists:

$$\begin{array}{ccccccccc}
\cdots & \longrightarrow & P_2 & \xrightarrow{f_2} & P_1 & \xrightarrow{f_1} & P_0 & \xrightarrow{f_0} & M_0 \\
 & & \downarrow{\scriptstyle \phi_2} & & \downarrow{\scriptstyle \phi_1} & & \downarrow{\scriptstyle \phi_0} & & \| \\
\cdots & \longrightarrow & Q_2 & \xrightarrow{g_2} & Q_1 & \xrightarrow{g_1} & Q_0 & \xrightarrow{g_0} & M_0
\end{array}$$

Taking $\mathrm{Hom}(-, N)$, this becomes

$$\begin{array}{ccccccc}
0 & \longrightarrow & \mathrm{Hom}(P_0, N) & \longrightarrow & \mathrm{Hom}(P_1, N) & \longrightarrow & \cdots \\
 & & \uparrow{\scriptstyle \phi_0^*} & & \uparrow{\scriptstyle \phi_1^*} & & \\
0 & \longrightarrow & \mathrm{Hom}(Q_0, N) & \longrightarrow & \mathrm{Hom}(Q_1, M) & \longrightarrow & \cdots
\end{array}$$

in which the (ϕ_i^*) are induced by $\mathrm{Hom}(-, N)$; since this functor is contravariant, they must point the other way. This sequence of maps induces $H^i(\phi^*) : H^i(\mathrm{Hom}(Q_\bullet, N)) \to H^i(\mathrm{Hom}(P_\bullet, N))$.

Then, the same argument with P_i and Q_i switched gives maps $\psi_i : Q_i \to P_i$ such that the relevant diagram commutes, and then an induced $H^i(\psi^*) : H^i(\mathrm{Hom}(P_\bullet, N)) \to H^i(\mathrm{Hom}(Q_\bullet, N))$.

Are these maps inverses of each other? It seems quite improbable, but is in fact the case. This can be shown by taking $\psi \circ \phi$ and calculating the induced map $H^i((\psi \circ \phi)^*) : H^i(\mathrm{Hom}(P_\bullet, N)) \to H^i(\mathrm{Hom}(P_\bullet, N))$. It will be shown to be the identity, and then the same proof works in the other direction.

Lemma 16.2. *Suppose $P_\bullet$ is a projective resolution of M and $Q_\bullet$ is a resolution of M (i.e. not necessarily projective). If $\alpha_i, \beta_i : P_i \to Q_i$, then (α_i) and (β_i) are homotopic to each other.*

$$\begin{array}{ccccccccc}
\cdots & \longrightarrow & P_2 & \longrightarrow & P_1 & \longrightarrow & P_0 & \longrightarrow & M \\
 & & {\scriptstyle \beta_2}\downarrow\downarrow{\scriptstyle \alpha_2} & & {\scriptstyle \beta_1}\downarrow\downarrow{\scriptstyle \alpha_1} & & {\scriptstyle \beta_0}\downarrow\downarrow{\scriptstyle \alpha_0} & & \| \\
\cdots & \longrightarrow & Q_2 & \longrightarrow & Q_1 & \longrightarrow & Q_o & \longrightarrow & M
\end{array}$$

This lemma can be applied to $\psi \circ \phi, \mathrm{id} : P_\bullet \to P_\bullet$, so $\psi \circ \phi \sim \mathrm{id}_{P_\bullet}$. The same thing can be played after applying $\mathrm{Hom}_A(-, N)$: one has two maps $(\psi \circ \phi)^*, \mathrm{id} : \mathrm{Hom}(P_\bullet, N) \to \mathrm{Hom}(P_\bullet, N)$, which are therefore homotopic. And, of course, by Lemma 15.1, homotopic maps induce the same cohomology. The same argument works for the composition in the other direction.

But wait! There's more! It's still necessary to show that this isomorphism $H^i(\phi^*)$ is independent of the choice of ϕ; this is what the lemma statement means by "canonical." So consider some other chain of maps ϕ'. Thanks to Lemma 16.2, $\phi, \phi' : P_\bullet \to Q_\bullet$, so $\phi \sim \phi'$, and therefore their induced maps ϕ^* and ϕ'^* are homotopic, so they have the same induced cohomology map. Thus, this map (the isomorphism) is indeed canonical. ⊠

Thus, when computing $\mathrm{Ext}^i(M, N)$, any resolution can be chosen, even the free ones.

Since $P_1 \to P_0 \to M \to 0$ is exact, then so is

$$0 \longrightarrow \mathrm{Hom}(M, N) \longrightarrow \mathrm{Hom}(P_0, N) \xrightarrow{f_0^*} \mathrm{Hom}(P_1, N),$$

so $\ker(f_0^*) = \mathrm{Hom}(M, N)$, and therefore $\mathrm{Ext}^0(M, N) = \mathrm{Hom}(M, N)$.

Example 16.1. Suppose $A = \mathbb{Z}$, and that one wants to compute $\mathrm{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/n, \mathbb{Z}/m)$.

$\mathbb{Z}/n$ has a free resolution in two steps: $0 \to \mathbb{Z} \xrightarrow{\cdot n} \mathbb{Z} \to 0$ (i.e. the cokernel of $\cdot n$ is $\mathbb{Z}/n$). Then, apply $\mathrm{Hom}_{\mathbb{Z}}(-, \mathbb{Z}/m)$, so there is an exact sequence

$$0 \longrightarrow \mathrm{Hom}(\mathbb{Z}, \mathbb{Z}/m) \xrightarrow{f} \mathrm{Hom}(\mathbb{Z}, \mathbb{Z}/m) \longrightarrow 0$$

where f is induced by multiplication by n in the first argument, so it's still the operation of $\cdot n$ on abelian groups. Thus, one obtains $0 \to \mathbb{Z}/m \xrightarrow{\cdot n} \mathbb{Z}/m \to 0$.

Then, $\mathrm{Ext}^0(\mathbb{Z}/m, \mathbb{Z}/n) = H^0 = \ker(n : \mathbb{Z}/m \to \mathbb{Z}/n)$. If $m = \prod p^{\beta_p}$ and $n = \prod p^{\alpha_p}$ are the prime decompositions of m and n, then this is

$$\prod_{p \text{ prime}} \ker\left(\mathbb{Z}/p^{\beta_p} \xrightarrow{p^{\alpha_p}} \mathbb{Z}/p^{\beta_p}\right) = \prod_p \begin{cases} \mathbb{Z}/p^{\beta_p}, & \text{if } \alpha_p \geq \beta_p, \\ p^{\beta_p - \alpha_p}\mathbb{Z}/p^{\beta_p}, & \text{otherwise.} \end{cases}$$

Thus, if $\ell = \gcd(m, n)$ and $m = \ell m'$, then $\mathrm{Hom}(\mathbb{Z}/m, \mathbb{Z}/n) = \mathrm{Ext}^0(\mathbb{Z}/m, \mathbb{Z}/n) = m'\mathbb{Z}/m\mathbb{Z}$. As abelian groups, this is isomorphic to $\mathbb{Z}/\ell\mathbb{Z}$, but this is less canonical.

Then, $\operatorname{Ext}^1(\mathbb{Z}/n, \mathbb{Z}/m) = \operatorname{coker}(\mathbb{Z}/m \xrightarrow{n} \mathbb{Z}/m) = \mathbb{Z}/(m\mathbb{Z} + n\mathbb{Z}) \cong \mathbb{Z}/\ell\mathbb{Z}$, which is this time completely canonical.

The general idea of such a calculation is to pick a free resolution and then do some computation.

Proposition 16.3. *Suppose $0 \to M' \xrightarrow{i} M \xrightarrow{\pi} M'' \to 0$ us a short exact sequence. Then, there is a long exact sequence*

$$\begin{array}{ccccccc} 0 & \longrightarrow & \operatorname{Hom}(M'', N) & \longrightarrow & \operatorname{Hom}(M, N) & \longrightarrow & \operatorname{Hom}(M', N) \\ & & \downarrow & & & & \\ & & \operatorname{Ext}^1(M'', N) & \longrightarrow & \operatorname{Ext}^1(M, N) & \longrightarrow & \operatorname{Ext}^1(M', N) \\ & & \downarrow & & & & \\ & & \operatorname{Ext}^2(M'', N) & \longrightarrow & \cdots & & \end{array}$$

The existence of the connecting map $\operatorname{Ext}^i(M', N) \to \operatorname{Ext}^{i+1}(M'', N)$ is the interesting bit, as well as how it fits into the exactness.

This proposition is useful in that if $\operatorname{Ext}^i(M', N)$ and $\operatorname{Ext}^i(M'', N)$ are known, then the idea of the size of $\operatorname{Ext}^i(M, N)$ is known, and often more information than that (e.g. because some of the terms vanish).

Proof of Proposition 16.3. Choose a projective resolution $(P'_\bullet, f'_\bullet)$ for M' and a projective resolution $(P''_\bullet, f''_\bullet)$ for M''. Then, the goal will be to generate a short exact sequence for M that allows things to commute nicely. The obvious choice is $P_i = P'_i \oplus P''_i$, but some care must be taken in choosing the maps.

To construct a map $P_0 \to M$, build a map on each of its two components: $f_{0,1} : P'_0 \to M$ is given by $i \circ f'_0$, and since P''_0 is projective, then the map $f''_0 : P''_0 \to M''$ lifts through $M \twoheadrightarrow M''$ to a map $f_{0,2} : P''_0 \to M$. Then, let $f_0 = (f_{0,1}, f_{0,2})$. Then, since each $f_{0,i}$ commutes in its respective component, this choice of f_0 makes the following diagram commute:

$$\begin{array}{ccc} 0 & & 0 \\ \downarrow & & \downarrow \\ P'_0 & \xrightarrow{f'_0} & M' \\ \downarrow & & \downarrow{\scriptstyle i} \\ P'_0 \oplus P''_0 & \xrightarrow{f_0} & M \\ \downarrow & & \downarrow \\ P''_0 & \xrightarrow{f''_0} & M'' \\ \downarrow & & \downarrow \\ 0 & & 0 \end{array}$$

In the general inductive step, a similar procedure works: one can show that the red sequence of kernels in the following diagram are exact:

$$\begin{array}{ccccccc} 0 & & 0 & & 0 & & \\ \downarrow & & \downarrow & & \downarrow & & \\ P'_n & \twoheadrightarrow & \ker(f'_{n-1}) & \longrightarrow & P'_{n-1} & \xrightarrow{f'_{n-1}} & \cdots \\ \downarrow & & \downarrow & & \downarrow & & \\ P'_n \oplus P''_n & \longrightarrow & \ker(f_n) & \longrightarrow & P'_{n-1} \oplus P''_{n-1} & \xrightarrow{f_{n-1}} & \cdots \\ \downarrow & & \downarrow & & \downarrow & & \\ P''_n & \twoheadrightarrow & \ker(f''_{n-1}) & \longrightarrow & P''_{n-1} & \xrightarrow{f''_{n-1}} & \cdots \\ \downarrow & & \downarrow & & \downarrow & & \\ 0 & & 0 & & 0 & & \end{array}$$

Then, the next maps in the resolutions are the compositions of the blue arrows, and the process continues.

The upshot is that one can choose projective resolutions of M', M, and M'' such that they are termwise split short exact: $0 \to P'_\bullet \to P_\bullet \to P''_\bullet \to 0$, where the sequence is in the category of complexes. Applying $\operatorname{Hom}(-, N)$, one gets maps $\operatorname{Hom}(P''_\bullet, N) \to \operatorname{Hom}(P_\bullet, N) \to \operatorname{Hom}(P'_\bullet, N)$, which is still termwise split short exact because all of the P_i are projective.

17. More Derived Functors: 11/15/13

The proof of Proposition 16.3 was left unfinished, so it will be completed here.

Lemma 17.1 (Snake)**.** *Suppose the following diagram of A-modules is commutative and has exact rows:*

$$\begin{array}{ccccccc} & & X' & \longrightarrow & X & \longrightarrow & X'' \longrightarrow 0 \\ & & \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle f''} \\ 0 & \longrightarrow & Y' & \longrightarrow & Y & \longrightarrow & Y'' \end{array}$$

Then, there is an exact sequence

$$\ker(f') \to \ker(f) \to \ker(f'') \xrightarrow{\delta} \operatorname{coker}(f') \to \operatorname{coker}(f) \to \operatorname{coker}(f''),$$

and δ is canonical.

The bit of the proof involving δ is the only nontrivial part. Also note that the ends of the sequence aren't zero; the first and last maps might not be injective and surjective, respectively. For a full proof, consult [1].

Continuation of the proof of Proposition 16.3. Today, somewhat different notation for the long exact sequence was used:

$$\begin{array}{ccccccccc} & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & X^{-1} & \longrightarrow & Y^{-1} & \longrightarrow & Z^{-1} & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f_{-1}} & & \downarrow{\scriptstyle g_{-1}} & & \downarrow{\scriptstyle h_{-1}} & & \\ 0 & \longrightarrow & X^0 & \xrightarrow{\alpha_0} & Y_0 & \xrightarrow{\beta_0} & Z^0 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f_0} & & \downarrow{\scriptstyle g_0} & & \downarrow{\scriptstyle h_0} & & \\ 0 & \longrightarrow & X^1 & \xrightarrow{\alpha_1} & Y_1 & \xrightarrow{\beta_1} & Z^1 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle g_1} & & \downarrow{\scriptstyle h_1} & & \end{array}$$

In this diagram, $X^0/\operatorname{Im}(f_{-1}) \to Y^0/\operatorname{Im}(g_{-1}) \to Z^0/\operatorname{Im}(h_{-1}) \to 0$ because all the squares are commutative and $Y^0 \twoheadrightarrow Z^0$. It's also exact at $Y^0/\operatorname{Im}(g_{-1})$, but not necessarily on the left. Similarly, $0 \to \ker(f_1) \to \ker(g_1) \to \ker(h_1)$ is left exact.

Thus, we have the following diagram:

$$\begin{array}{ccccccccc} & & X^0/\operatorname{Im}(f_{-1}) & \longrightarrow & Y_0/\operatorname{Im}(g_{-1}) & \longrightarrow & Z^0/\operatorname{Im}(h_{-1}) & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \overline{f}_0} & & \downarrow{\scriptstyle \overline{g}_0} & & \downarrow{\scriptstyle \overline{h}_0} & & \\ 0 & \longrightarrow & \ker(f_1) & \longrightarrow & \ker(g_1) & \longrightarrow & \ker(h_1) & & \end{array}$$

where the maps f_i, g_i, and h_i factor through these quotients (as $\overline{f}_i$, and so on) because this is in a complex. Then, using the Snake lemma, the following sequence is exact at the middle four points:

$$\ker(\overline{f}_0) \to \ker(\overline{g}_0) \to \ker(\overline{h}_0) \xrightarrow{\delta} \operatorname{coker}(\overline{f}_0) \to \operatorname{coker}(\overline{g}_0) \to \operatorname{coker}(\overline{h}_0).$$

Then, $\ker(\overline{f}_0) = \ker(f_0)/\operatorname{Im}(f_1) = H^0(X)$, and similarly for $\overline{g}_0$ and $\overline{h}_0$. Thus, the above sequence is actually

$$H^0(X) \to H^0(Y) \to H^0(Z) \xrightarrow{\delta} H^1(X) \to H^1(Y) \to H^1(Z).$$

This argument can be applied to the next row to continue the sequence, and so on, yielding connecting homomorphisms $\delta^i : H^i(Z) \to H^{i+1}(X)$.

Concretely, suppose $z^0 \in \ker(h_0)$. Since β_0 is surjective, then z^0 has a preimage y^0, so $y^1 = g_0(y^0)$ is killed by β_1, since the square commutes. Then, y^1 has a preimage x^1 under α_1. This is the required thing: $x^1 = \delta(z^0)$, though it's necessary to check that $x^1 \in \ker(f_1)$. Since $y^1 \in \operatorname{Im}(g^0)$, then $g_1(y^1) = 0$, and therefore $\alpha_2 \circ f_1(x^1) = 0$, and thus $f_1(x^1) = 0$ because α_2 is injective, and so $x^1 \in \ker(f_1)$, so one obtains something in $H^1(X)$. It yet remains to check that this is well-defined modulo the image, but this isn't so hard. ⊠

Last time, it was shown that any short exact sequence $0 \to M' \to M \to M'' \to 0$ can be resolved by a short exact sequence $0 \to P'_\bullet \to P_\bullet \to P''_\bullet \to 0$ of projective resolutions (of M', M, and M'' respectively). Then, one obtains a sequence

$$0 \longrightarrow \operatorname{Hom}(P''_\bullet, N) \longrightarrow \operatorname{Hom}(P_\bullet, N) \longrightarrow \operatorname{Hom}(P'_\bullet, N) \longrightarrow 0,$$

where $P_\bullet = P'_\bullet \oplus P''_\bullet$, though as shown the maps require a little thought. Then, apply the previous construction to get a long exact sequence

$$\cdots \to \operatorname{Ext}^i(M'', N) \to \operatorname{Ext}^i(M, N) \to \operatorname{Ext}^i(M', N) \xrightarrow{\delta} \operatorname{Ext}^{i+1}(M'', N) \to \cdots \tag{4}$$

This long exact sequence is very useful for proving things, such as the following:

- Suppose $0 \to M' \to M \to M'' \to 0$ is short exact and $\operatorname{Ext}^i(M', N) = 0$ and $\operatorname{Ext}^i(M'', N) = 0$. Then, (4) becomes $\cdots \to 0 \to \operatorname{Ext}^i(M, N) \to 0 \to \cdots$, which means that $\operatorname{Ext}^i(M, N) = 0$.
- Similarly, suppose that $\operatorname{Ext}^i(M', N) = 0$ and $\operatorname{Ext}^{i+1}(M, N) = 0$. Then, (4) becomes $\cdots \to 0 \xrightarrow{\delta} \operatorname{Ext}^{i+1}(M'', N) \to 0 \to \cdots$, so $\operatorname{Ext}^{i+1}(M'', N) = 0$.

As an application of this idea, if I want to compute $\operatorname{Ext}^i(M'', N)$, I could take a free resolution $0 \to M' \to A^{\oplus n} \to M'' \to 0$. Then, using Proposition 16.3, the following sequence is exact:

$$\operatorname{Ext}^i(M'', N) \to \operatorname{Ext}^i(A^{\oplus n}, N) \to \operatorname{Ext}^i(M', N) \to \operatorname{Ext}^{i+1}(M'', N) \to \operatorname{Ext}^{i+1}(M'', N).$$

However, by definition, if M is projective, then it is its own projective resolution, so $\operatorname{Ext}^i(M, N) = 0$ whenever $i > 0$, because the complex only has one nonzero term, and $\operatorname{Ext}^0(M, N) = \operatorname{Hom}(M, N)$. In particular, $\operatorname{Ext}^i(A^{\oplus n}, N) = 0$, so the sequence looks like

$$\operatorname{Ext}^i(M'', N) \to 0 \to \operatorname{Ext}^i(M', N) \to \operatorname{Ext}^{i+1}(M'', N) \to 0 \to \cdots$$

so $\operatorname{Ext}^i(M', N) \cong \operatorname{Ext}^{i+1}(M'', N)$ whenever $i > 0$. This can simplify some computations.

It's hard to say anything more when $i = 0$. $\operatorname{Ext}^1(M'', N) = \operatorname{Hom}(M', N)/\operatorname{Im}(\operatorname{Hom}(A^{\oplus n}, N))$, which is the definition, but this is the space of maps $M' \to N$ modulo maps that can be lifted to $A^{\oplus n} \to 0$ (since $M' \subset A^{\oplus n}$). Since N isn't required to be injective, this is interesting. Thus, $\operatorname{Ext}^1(M, N)$ measures the degree to which extending maps into N from M' to $A^{\oplus n}$ fails. This is one useful interpretation of $\operatorname{Ext}^1(M, N)$; there are others.

For example, another interpretation of $\operatorname{Ext}^1(M, N)$ is that it classifies extensions of M by N (hence the name) up to isomorphism, i.e. short exact sequences $0 \to N \to X \to M \to 0$, where this is isomorphic to another extension $0 \to N \to X' \to M \to 0$ if an f exists such that

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N & \longrightarrow & X & \longrightarrow & M & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle f} & & \| & & \\ 0 & \longrightarrow & N & \longrightarrow & X' & \longrightarrow & M & \longrightarrow & 0 \end{array}$$

commutes. However, while $\operatorname{Ext}^1(M, N)$ is an abelian group and even an A-module, the set of isomorphisms is abstractly just a set, and algebraic operations have to be defined on it, which are developed in the exercises. For example, multiplication by an $a \in A^\times$ sends $0 \to N \xrightarrow{\alpha} X \xrightarrow{\beta} M \to 0$ to a short exact sequence $0 \to N \xrightarrow{\alpha} X \xrightarrow{a\beta} M \to 0$, which is in general a different extension class. This is a very useful interpretation of $\operatorname{Ext}^1(M, N)$, and the higher Ext functors also admit such an interpretation. For example, $\operatorname{Ext}^n(M, N)$ has a similar meaning on sequences of the form $0 \to N \to X_1 \to \cdots \to X_n \to M \to 0$.

The covariant case. So far, $\operatorname{Ext}^i(M, N)$ has been defined by fixing N, yielding a contravariant functor $\operatorname{Ext}^i(-, N)$. What if M were fixed and N varied? $\operatorname{Hom}(M, -)$ is a covariant left exact functor from the category of A-modules to itself, and the goal is to use derived functors to understand how inexact it is. Temporarily denote the functor derived from $\operatorname{Hom}(M, -)$ as $\operatorname{\exists xt}^i(M, -)$, which is also a functor from A-modules to A-modules (this is nonstandard notation, and will be dropped in a bit). Formally, pick an injective resolution $0 \to I^0 \to I^1 \to I^2 \to \cdots$ of N and define $\operatorname{\exists xt}^i(M, N) = H^i(\operatorname{Hom}(M, I^\bullet))$, i.e. applying $\operatorname{Hom}(M, -)$ to each I. It's easy to see that $\operatorname{\exists xt}^0(M, N) = \operatorname{Hom}(M, N)$, and that $\operatorname{\exists xt}^i(M, N) = 0$ if $i < 0$.

There is an analogous statement to Lemma 16.1 for $\operatorname{\exists xt}^i(M, N)$, in that it is canonically independent of choice of injective resolution of N. The proof is essentially the same: show that any two choices induce homotopies of complexes. There's also an analogous statement to Proposition 16.3: that given some short exact sequence $0 \to N' \to N \to N'' \to 0$, there's a compatible short exact sequence of injective resolutions $0 \to I'^\bullet \to I^\bullet \to I''^\bullet \to 0$ such that after applying $\operatorname{Hom}(M, -)$ and the snake lemma, one has a long exact sequence

$$\cdots \to \operatorname{\exists xt}^i(M, N') \to \operatorname{\exists xt}^i(M, N) \to \operatorname{\exists xt}^i(M, N'') \to \operatorname{\exists xt}^{i+1}(M, N') \to \cdots$$

(Seriously, the proof is the same, just reversing the arrows.)

Theorem 17.2. *There is a natural isomorphism* $\operatorname{Ext}^i(M, N) \cong \operatorname{\exists xt}^i(M, N)$.

As in most cases where the word "natural" is invoked, it's entirely unclear what it actually means, but it will be clarified in a few paragraphs.

It's not hard to figure out that if M is fixed, then $N \to N'$ gives a natural transformation $\mathrm{Ext}^i(M,N) \to \mathrm{Ext}^i(M,N')$, so $\mathrm{Ext}^i(M,-)$ is a functor. One can also view $\mathrm{Ext}^i(-,-)$ as a bifunctor $\underline{A\text{-Mod}}^{\mathrm{op}} \times \underline{A\text{-Mod}} \to \underline{A\text{-Mod}}$. Likewise, $\exists\mathrm{xt}^i(-,-)$ is a bifunctor in the same manner. Then, the naturality condition in the theorem states that these bifunctors are isomorphic, which is why the $\exists$xt notation is never used. The proof of this theorem involves observing that a lot of diagrams that look like they ought to commute do in fact commute.

Note that it is in general easier to find a projective resolution than an injective resolution, so using $\mathrm{Ext}^i(-,M)$ is generally a lot easier for the purposes of computation. In some sense, injective modules are bigger objects: not so much in abelian groups, where $\mathbb{Q}/\mathbb{Z}$ is a good example of an injective module, but in general, e.g. $A = \mathbb{Z}[x]$, it's not very easy to write down injective modules. Abstractly, projective and injective objects look very symmetric, but this isn't the case in reality, because the categories themselves are biased. This is because the forgetful functor $\underline{A\text{-Mod}} \to \underline{\mathrm{Set}}$ has a left adjoint (that takes a set S and returns the free A-module having a basis indexed by S), but cannot have a right adjoint, so the notion of free modules for projective modules has no corresponding notion for injective modules. If for some specific A a right adjoint exists, then it will be equally as easy to construct projective and injective modules.

It's not necessary to go all the way to an injective resolution to compute $\exists\mathrm{xt}^i(M,N)$ when M is fixed. Instead of an injective resolution $N \to I^\bullet$, one can take a complex $I^\bullet$ such that $\mathrm{Ext}^{>0}(M, I^j) = 0$ for all j. Then, this $I^\bullet$ is as good as an injective resolution. This is a weaker condition; it requires only checking for a single M, while for injectivity it would be necessary for it to hold for all M. The proof of this equivalence is not entirely trivial, though.

Proposition 17.3. *Let I be an A-module. Then, the following are equivalent:*

(1) I is injective.
(2) $\mathrm{Ext}^i(M,I) = 0$ for any A-module M and $i > 0$.
(3) $\mathrm{Ext}^1(M,I) = 0$ for any A-module M.
(4) $\mathrm{Ext}^1(A/J, I) = 0$ for all ideals J of A.

Proof. (1) $\implies$ (2) is done by taking I as its own injective resolution. The projective case can be seen earlier in today's notes. Then, (2) $\implies$ (3) is trivial, and similarly (3) $\implies$ (4) is clear. Thus, consider only the remaining things to show:

- (3) $\implies$ (1). Suppose $j : M' \hookrightarrow M$ and $M' \to I$; then, to show that I is injective, one would want an extension of these to $M \to I$. But the obstruction to this is just $\mathrm{Ext}^1(M,I)$: let $M'' = \mathrm{coker}(j)$, so that $0 \to M' \to M \to M'' \to 0$ is a short exact sequence. Then, there is an associated exact sequence $\mathrm{Hom}(M,I) \to \mathrm{Hom}(M',I) \to \mathrm{Ext}^1(M'',I) = 0$, so the first arrow is surjective and therefore the lift exists.
- (4) $\implies$ (1). The same argument as above shows that every map $J \to I$ extends to the whole ring $A \to I$, which as demonstrated in the exercises implies I is injective. Thus, one could add even more equivalences to this list. ⊠

There is an analogous version of Proposition 17.3 for projective modules, outlined in the exercises. Some of the conditions look slightly different, though.

There's one more functor to derive, $\otimes$. The functor $M \otimes_A -$ sends A-modules to A-modules, and is in general not exact. The resulting functor is called $\mathrm{Tor}_i^A(M,-)$.[52] This is given by $\mathrm{Tor}_i^A(M,N) = H^{-i}(M \otimes_A P_\bullet)$, where $P_\bullet$ is a projective (or free) resolution of N, so $M \otimes P_\bullet$ is taken termwise. Since $M \otimes -$ is covariant, this is only nonzero when $i \le 0$.

Definition. An A-module is called M flat if $M \otimes_A -$ is an exact functor.

Since $M \otimes_A -$ is already known to be right exact, then only left-exactness or (equivalently) preserving injectivity needs checking.

18. Flatness and the Tor Functors: 11/18/13

Recall that $\mathrm{Tor}_i^A(M,N)$ was defined by taking a projective resolution $P_\bullet \to M$ and then defining $\mathrm{Tor}_i^A(M,N) = H^{-i}(P_\bullet \otimes_A M)$. In some sense, this is done by fixing N and letting M vary. We also defined A-flat modules (how musical!) as those modules M where $M \otimes_A -$ or equivalently $- \otimes_A M$ are exact.

Alternatively, one could take a projective resolution $Q_\bullet \to N$, which leads to an *a priori* different functor $H^{-i}(M \otimes_A Q_\bullet)$. Thankfully, though, there is a canonical isomorphism $\mathrm{Tor}_i^A(M,N) \cong H^{-i}(M \otimes_A Q_\bullet)$, which satisfies all sorts of functoriality and compatibility conditions.

[52] The i is in the lower index here because this corresponds to some negative index of the cohomology of some complex.

Example 18.1.

- Free A-modules are flat, because tensoring with a free module just produces copies of the other product.
- Projective A-modules are flat, since they're direct summands of free modules; a fuller proof is deferred to the exercises.
- Any localization $S^{-1}A$ of A is flat, which follows from a combination of two statements:
  - $S^{-1}A \otimes_A - = S^{-1}(-)$, and:
  - Localization is an exact functor, which was proven in the exercises.
- If M_i is flat for every $i \in I$, where I is an inductive system (i.e. directed set), then $\varinjlim_i M_i$ is also flat. In particular, the (infinite) direct sum of flat modules is also flat. This can be proven by checking the definition: if $N' \hookrightarrow N$, just check that $(\varinjlim M_i) \otimes N' \to (\varinjlim_i M_i) \otimes N$ is injective. This follows from the following two facts:
  - Direct limits commute with the tensor product; specifically, $(\varinjlim_i M_i) \otimes N = \varinjlim_i (M_i \otimes N)$. This was shown in the exercises.
  - The limit of exact sequences is still exact (also shown in the exercises); in particular, injective maps are preserved.

  This statement is particularly strong: it implies localization preserves flatness, because every localization can be written as a direct limit of copies of A, e.g. $\mathbb{Q} = \varinjlim_n M_n$, where $M_n = (1/n)\mathbb{Z}$ (isomorphic to $\mathbb{Z}$ as groups), with natural maps $M_n \hookrightarrow M_{n'}$ whenever $n \mid n'$. Thus, $\mathbb{Q}$ is a flat $\mathbb{Z}$-module.

Thus, flatness is more general than projectivity or freeness, because $\mathbb{Q}$ is neither projective nor free: if $F = \mathbb{Z}[\mathbb{Q}]$, then any lift s of $\mathrm{id} : \mathbb{Q} \to \mathbb{Q}$ must send $s(1) = \sum c_x e_x$ for some $c_x \in \mathbb{Z}$ and $e_x \in \mathbb{Q}$ that are the basis elements $e_x \mapsto x$ via $F \twoheadrightarrow \mathbb{Q}$.. Let N be larger than any of the c_x (which is easy because there are a finite number of them); then, $s(1/N) = \sum c_x/N \cdot e_x$. Thus, this isn't an integer, which is a problem. In fact, $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}, \mathbb{Z}[\mathbb{Q}]) = \{0\}$. However, if $M \to M''$ and M'' is finitely generated, then $\mathbb{Q} \to M''$ must be the zero map, because $\mathbb{Q} \to \mathbb{Z}$ and $\mathbb{Q} \to \mathbb{Z}/n$ are forced to be the zero map, and this trivially lifts. Notice that just checking on finitely generated modules fails, which is interesting. However, for flatness, such a check does work.

Theorem 18.1. *Let M be an A-module. Then, the following are equivalent:*

(1) M is A-flat.
(2) For any $N' \hookrightarrow N$, $N' \otimes M \to N \otimes N$ is still injective.
(3) For any ideal $I \hookrightarrow A$, $I \otimes_A M \to A \otimes_A M = M$ is injective.
(4) For any ideal $I \hookrightarrow A$, $\mathrm{Tor}_1(M, A/I) = 0$.
(5) For any A-module N, $\mathrm{Tor}_1(M, N) = 0$.
(6) For any A-module N, $\mathrm{Tor}_i(M, N) = 0$ for all $i > 0$.

Proof. The proof will follow the following diagram of implications, in approximately clockwise order:

$$\begin{array}{ccccc} (1) & \Longleftrightarrow & (2) & \Longrightarrow & (3) \\ \Downarrow & \searrow & & & \Downarrow \\ (6) & \Longrightarrow & (5) & \Longleftarrow & (4) \end{array}$$

- First, the easiest: (1) $\implies$ (2) by definition.
- Similarly, (2) $\implies$ (3) because it's just a special case.
- (3) $\implies$ (4): consider the short exact sequence $0 \to I \xrightarrow{f} A \to A/I \to 0$. Then, apply Tor to obtain the long exact sequence

  $$\cdots \to \mathrm{Tor}_1(A, M) \to \mathrm{Tor}_1(A/I, M) \to I \otimes M \xrightarrow{f} A \otimes M \to A/I \otimes M \to 0,$$

  but $\mathrm{Tor}_1(A, M) = 0$, because A is a free (and therefore flat) A-module. Thus, $\mathrm{Tor}_1(A/I, M) = \ker(f)$, but f is injective, so $\mathrm{Tor}_1(A/I, M) = 0$.
- (4) $\implies$ (5): this argument involves "approximating" N by A/I using two reductions. This is a common trick in commutative algebra. First, write $N = \varinjlim N_i$, where all of the N_i are finitely generated (e.g. all finitely generated submodules of N, directed by inclusion). Then, just as tensor products commute with direct limits, so does the Tor functor (which will be proven in the exercises). Thus, $\varinjlim_\alpha \mathrm{Tor}_i(M, N_\alpha) = \mathrm{Tor}_i(M, \varinjlim_\alpha N_\alpha)$.[53] Then, it's enough to show that $\mathrm{Tor}_1(M, N_\alpha) = 0$ for all finitely generated submodules N_α, so assume N is finitely generated and induct on the number of generators n needed to generate N.

  If $n = 0$, then $N = 0$, so there's nothing to prove.

[53] Notice that this is completely untrue for Ext, because $\mathbb{Q}$ isn't projective.

If $n \geq 1$, then pick a set of generators $x_1, \dots, x_n$ of N. Let $N' \subset N$ be the submodule generated by the first $n-1$ elements. Then, N/N' is generated by one element, so $N/N' \cong A/I$ for some ideal $I \subset A$. Thus, the short exact sequence $0 \to N' \to N \to N/N' \to 0$ is actually $0 \to N' \to N \to A/I \to 0$, which induces the long exact sequence

$$\cdots \longrightarrow \operatorname{Tor}_1(M, N') \longrightarrow \operatorname{Tor}_1(M, N) \longrightarrow \operatorname{Tor}_1(M, A/I) \longrightarrow \cdots$$

But $\operatorname{Tor}_1(M, N') = 0$ by the inductive hypothesis and $\operatorname{Tor}_1(M, A/I) = 0$ by (4), so the middle term must also be zero.

- (5) $\implies$ (1): Suppose $\operatorname{Tor}_1(M, -) = 0$. Then, given the short exact sequence $0 \to N' \to N \to N'' \to 0$, one has the long exact sequence

$$\operatorname{Tor}_1(M, N'') = 0 \longrightarrow N' \otimes M \longrightarrow N \otimes M \longrightarrow N'' \otimes M \longrightarrow 0$$

and therefore $N' \otimes M \hookrightarrow N \otimes M$.
- (6) $\implies$ (5) because the latter is just a special case.
- (1) $\implies$ (6): Suppose M is flat. By the rest of the proof, this means that $\operatorname{Tor}_1(M, N) = 0$. For $\operatorname{Tor}_2(M, N)$ (which will illustrate the general argument), the goal will be to relate it to Tor_1 of some other module.

  Let F be a free resolution of M: $0 \to M' \to F \to M \to 0$, where M' is the kernel of the surjection $F \twoheadrightarrow M$.

  Claim. M is A-flat.

  Proof. Let $N' \hookrightarrow N$. Then, because M and F are both flat, then one has the diagram

$$\begin{array}{ccccccc} N' \otimes M' & \xrightarrow{\star} & N' \otimes F & \longrightarrow & N' \otimes M & \longrightarrow & 0 \\ \downarrow{\scriptstyle ?} & & \downarrow & & \downarrow & & \\ N \otimes M & \longrightarrow & N \otimes F & \longrightarrow & N \otimes M & \longrightarrow & 0 \end{array}$$

  If $\star$ were injective, then ? would be too (since it's part of a composition of injective things). Then, take the exact sequence $\operatorname{Tor}_1(N', M) \to N' \otimes M' \xrightarrow{\star} N' \otimes F$, but $\operatorname{Tor}_1(N', M) = 0$ because M is flat, so $\star$ is injective.[54] ⊠

  Thus, apply $\operatorname{Tor}(-, N)$ to yield the sequence

$$\operatorname{Tor}_2(F, N) \to \operatorname{Tor}_2(M, N) \to \operatorname{Tor}_1(M', N) \to \operatorname{Tor}_1(F, N),$$

  but since F is free (a projective module works just fine here, too), then $\operatorname{Tor}_2(F, N) = 0$, and since M' is flat, then $\operatorname{Tor}_1(M, N) = 0$. Thus, $\operatorname{Tor}_2(M, N) = 0$ as well, and then induct. ⊠

Computing Tor. Let A be a domain and $a \in A \setminus 0$. Then, what is $\operatorname{Tor}_1(A/(a), M)$?

Let $f : A \xrightarrow{\cdot a} A$, which is injective because A is a domain. Then, $0 \to A \xrightarrow{f} A \to A/(a) \to 0$ is exact, so one has the exact sequence

$$0 \to \operatorname{Tor}_1(A/(a), M) \to M \xrightarrow{a} M \to M/aM \to 0,$$

so $\operatorname{Tor}_1(A/(a), M) = M[a] = \ker(m \mapsto am)$ (i.e. the submodule of M killed by a). In particular, if M is torsion-free, then $\operatorname{Tor}_1(A/(a), M) = 0$.

If A is a PID and M is torsion-free, then this further implies that M is flat. If one further assumes that M is finitely generated, then it's free, which makes it even easier, but this result is more general.

More interestingly, let $A = k[x, y]$ so that $k = A/(x, y)$. Then, calculate $\operatorname{Tor}_1(k, M)$.

One has $A \twoheadrightarrow k$ with kernel (x, y), so only two polynomials are necessary to express things in the kernel. Thus, take $A \oplus A \to A$ that sends $(1, 0) \mapsto x$ and $(0, 1) \mapsto y$, which is also not injective. Its kernel consists of pairs $(f(x, y), g(x, y))$ such that $xf(x, y) + yg(x, y) = 0$. Since A is a UFD, then $y \mid f$ and $x \mid g$, so this is actually of the form $(y, -x) \cdot h$ for any $h \in k[x, y]$. Thus, this is really a free A-module, isomorphic to A, so there's a three-term free resolution of k. After tensoring with M, this becomes $M \xrightarrow{s} M \oplus M \xrightarrow{t} M$, where $s : \gamma \mapsto (y\gamma, -x\gamma)$ and $t : (\alpha, \beta) \mapsto x\alpha + y\beta$.

This sequence might not be exact, which is why we're computing Tor in the first place. $\operatorname{Tor}_2(k, M) = \ker(s) = M[x] \cap M[y]$ (i.e. those things killed by x and y). $\operatorname{Tor}_1(k, M)$ is a little more complicated: $\operatorname{Tor}_1(k, M) = \{(\alpha, \beta) \mid x\alpha + y\beta = 0\}/\{y\gamma, -x\gamma \mid \gamma \in M\}$. If $M = k$, then s and t are just the zero maps, in which case $\operatorname{Tor}_1(k, k) = k \oplus k$ and $\operatorname{Tor}_2(k, k) = k$.

These calculations are true more generally for polynomials in n variables, but they require a little more thought. For example, if $A = k[x_1, \dots, x_n]$, then $\operatorname{Tor}_i^A(k, k) \cong \bigwedge^i(k^n)$. More intrinsically, if there exists a vector space V of

[54] Here, we're using something which is implicitly non-obvious: that $\operatorname{Tor}_1(M, -) = 0$ iff $\operatorname{Tor}_1(-, M) = 0$. Be observant as to which side varies.

dimension n such that $A = \mathrm{Sym}(V)$, then $\mathrm{Tor}_i^A(k,k) \cong \Lambda^i(V)$, completely canonically (unlike the previous example, which requires choosing generators). In particular, $\mathrm{Tor}_i^A(k,k) = 0$ when $i > \dim(V)$.

Part 7. Representation Theory of Finite Groups

19. Group Representations and Maschke's Theorem: 11/20/13

"What motivates the study of group representations? I'm not Schur." – Ravi Fernando

Before talking about representations, it's important to have some examples of finite groups in mind.

- Finite abelian groups are finitely generated $\mathbb{Z}$-modules, so they are of the form $\mathbb{Z}/n_1 \times \cdots \mathbb{Z}/n_k$. The simplest such groups are $\mathbb{Z}/p$ when p is prime.
- The dihedral groups D_{2n} are slightly non-abelian. These groups are given by the symmetries of the n-gon, so there are n rotations by $2\pi k/n$ for $k \in \mathbb{Z}/n$, but also n reflections, each that fixes a different line through the origin. This is a group of order $2n$; it has a normal subgroup of order n, the group of rotations (isomorphic to $\mathbb{Z}/n$), and the quotient is $\mathbb{Z}/2$.

  One writes $1 \to \mathbb{Z}/n \to D_{2n} \to \mathbb{Z}/2 \to 1$ to state that $\mathbb{Z}/n$ is normal in D_{2n}; notice that 1 is used instead of 0 in this short exact sequence. The shortness of this sequence is what was meant by "slightly non-abelian;" solvable groups such as this one are built out of abelian groups.
- S_n is a very non-abelian group, the group of permutations on n letters, as is A_n, the group of even permutations on n letters.
- There are also groups that are interesting in other subjects of mathematics, such as

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,b,c,d \in \mathbb{F}_p, ad - bc = 1 \right\},$$

  usually denoted $\mathrm{SL}_2(\mathbb{F}_p)$. This is a finite, non-abelian group under matrix multiplication. More generally, one could take $n \times n$ matrices over $\mathbb{F}_p$, denoted $\mathrm{SL}_n(\mathbb{F}_p)$, or even over any finite ring R, yielding $\mathrm{SL}_n(R)$.

Fix a field k and let G be a finite group.

Definition. A representation of G is a pair (ρ, V) where V is a finite-dimensional k-vector space and $\rho : G \to \mathrm{GL}(V)$ (where $\mathrm{GL}(V)$ is the group of k-linear automorphisms of V) is a group homomorphism.

Infinite-dimensional representations exist, but matter more so in the context of infinite groups, and thus aren't a concern right now.

Another useful way of understanding the definition is that ρ is just an action of G on V.

Recall the notion of a group ring $k[G]$, a k-vector space spanned by symbols $[g]$ for $g \in G$. The ring structure is given by

$$\left(\sum a_g \cdot [g]\right)\left(\sum b_j \cdot [h]\right) = \sum a_g b_h \cdot [gh] = \sum_{g \in G} [g] \left(\sum_{h_1 h_2 = g} a_{h_1} b_{h_2} \right).$$

Here, multiplication within brackets is the group multiplication. Recall further that if G is noncommutative, then so is $k[G]$.

Lemma 19.1. *A representation of G is the same as a finite-dimensional $k[G]$-module.*

Proof.

- Given a representation (ρ, V), place a $k[G]$-module structure on V in which $\sum a_g \cdot [g]$ acts by $\sum a_g \rho(g)$, which is still in $\mathrm{End}(V)$ because it's linear.
- Conversely, given a finite-dimensional $k[G]$-module M, it is a vector space, and define $\rho(g) = [g] \cdot M$. ⊠

Though this is a bit tautological, both points of view are helpful: a homomorphism into a matrix group or a module over a noncommutative algebra.

Example 19.1. Suppose $G = \mathbb{Z}/n$. Then, a representation of G is the same as giving a vector space V and an operator $\sigma : V \to V$ such that $\sigma^n = 1$, since the action of G is determined by its generator.

Nonetheless, it's still not entirely trivial to compute its structure. If $k = \mathbb{C}$, then σ can be diagonalized as $\sigma \sim \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, where each λ_i is an n^{th} root of unity. Thus, V is the direct sum of one-dimensional representations, because the eigenvectors given by the λ_i are G-stable. Each one-dimensional representation acts by an n^{th} root of unity, and thus they are classified.

Over $\mathbb{R}$, it's slightly more complicated, since σ might not be diagonalizable. However, it can be decomposed into 2×2 blocks and diagonal entries; the latter are again the eigenvalues, and each 2×2 block is of the form $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, where $\theta \in 2\pi\ell/n$ for an $\ell \in \mathbb{Z}/n$. The single entries are roots of unity in $\mathbb{R}$, so they must be ± 1. Thus, in this case, V is a direct sum of one-dimensional and two-dimensional representations.

Definition. A representation (ρ, V) of G is irreducible if $V \neq 0$ and for any G-stable subspace $V' \subsetneq V$, $V' = 0$.[55]

This implies that one-dimensional representations are always irreducible, but that two-dimensional ones might be reducible. It also illustrates that over $\mathbb{R}$, every representation of $\mathbb{Z}/n$ is a direct sum of irreducible representations of degrees 1 and 2, which can be classified: the one-dimensional representations are just ± 1, or just 1 when n is odd, and the two-dimensional representations are those given by the θ as defined above. However, different choices of θ sometimes give rise to the same representation. Thus, the actual set of angles is $\{0, 2\pi/n, \dots, 2\pi(n-1)/n\}/(\theta \sim -\theta)$, where angles are always mod 2π. Thus, the number of irreducible representations depends on the parity of n: it is $[n/2] + 1$. This is because

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \sim \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{pmatrix},$$

so these two matrices give the same representation. Equivalently, these matrices have the same complex eigenvalues.

Notice that an easy representation theory over $\mathbb{C}$ becomes more interesting over $\mathbb{R}$, because the latter isn't algebraically closed. Looking at the representations of $\mathbb{Z}/n$ over $\mathbb{Q}$, the quotient relationship is different: one has $\{1, e^{2\pi i/n}, \dots, e^{2\pi(n-1)/n}\}/\operatorname{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ (i.e. we kill the action of this group). More algebraically, this just becomes $\{0, 1, \dots, n-1\}/(\mathbb{Z}/n\mathbb{Z})^*$, because cyclotomic fields are nice. The orbits under this quotient are in bijection with divisors of n.

The case of $\mathbb{R}$ is actually very similar, because $\operatorname{Gal}(\mathbb{C}/\mathbb{R})$ sends $\theta \mapsto -\theta$, so it still works, but there are more representations.

Now, consider $G = \mathbb{Z}/p\mathbb{Z}$ and $k = \mathbb{F}_p$. There is a representation $a \mapsto \left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right) = \rho_a$, where $V = \mathbb{F}_p \cdot e_1 \oplus \mathbb{F}_p \cdot e_2$. Since e_1 and e_2 are fixed by these matrices, then $V_1 = \mathbb{F}_p \cdot e_1 \subset V$, but V_1 isn't a direct summand of V as irreducible representations, which is different from how things work in characteristic zero. In other words, there is no decomposition $V = V_1 \oplus V_2$ into G-stable subspaces: if it existed, then V_1 would be generated by an element of the form $xe_1 + ye_2$ with $y \neq 0$, so multiplying by ρ_a gives $(x + ay, y) \in V$, but then their difference is in V_2, so $V_1 \cap V_2 \neq \emptyset$, which is bad. Thus, in characteristic p, things becomes more complicated.

Definition. In the context of a representation ρ of a group G into a k-vector space V, a G-map f is a k-linear map such that the following diagram commutes for all $g \in G$.

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle \rho(g)}\downarrow & & \downarrow{\scriptstyle \rho(g)} \\ V & \xrightarrow{f} & V \end{array}$$

Definition. The set of G-maps $V \to V$ is called $\operatorname{End}_G(V)$, the G-endomorphisms of V. $\operatorname{End}_G(V)$ admits a ring structure: addition is pointwise, and multiplication is by composition.

Theorem 19.2 (Maschke)**.** *If $\#G$ is prime to* $\operatorname{char}(k)$,[56] *then any representation of G is a direct sum of irreducible representations.*

Proof. The idea of the proof is to split out one irreducible summand at a time.

Suppose $V_1 \subset V$ is G-stable; then, we want a projector $\pi : V \to V_1$. The inclusion $V_1 \hookrightarrow V$ already exists, and the goal is to find a one-sided inverse. This would imply $V = V_1 \oplus \ker(\pi)$, but since we want this to be G-stable, it's necessary for π to commute with the G-maps. Once this is done, it can be repeated inductively on $\ker(\pi)$, which has strictly smaller dimension than V, so the process terminates eventually.

Let π_1 be any projector onto V_1; it's not necessarily G-stable. It will be made to be so by averaging over G, an important notion in the representation theory of finite groups. Right now, it's not necessarily the case that $\pi_1 = g^{-1} \circ \pi_1 \circ g$ for all $g \in G$, so take

$$\pi = \frac{1}{\#G} \sum_{g \in G} g^{-1} \circ \pi_1 \circ g. \tag{5}$$

Notice that this step requires the characteristic assumption.

[55] A subspace V' of V is called G-stable if the action of G sends V' to itself.

[56] This will always be understood to include the case $\operatorname{char}(k) = 0$.

π is thus still a k-linear map $V \to V_1$, and it's a G-map: for any $h \in G$,

$$h^{-1} \circ \pi \circ h = \frac{1}{\#G} \sum_{g \in G} h^{-1} g^{-1} \pi g h = \frac{1}{\#G} \sum_{g \in G} (gh)^{-1} \pi (gh) = \pi,$$

because gh ranges over all of G exactly once, because h is fixed.

To check that π is a projection, take $V_1 \hookrightarrow V \xrightarrow{\pi} V_1$; is their composition the identity? Of course: take a $v_1 \in V_1$, so that $g(v_1) = v_1$ for all $g \in G$, and $\pi|_{V_1} = \mathrm{id}$, so

$$\frac{1}{\#G} \sum_{g \in G} g^{-1} \pi g(v_1) = \frac{1}{\#G} \sum_{g \in G} v_1 = \frac{\#G}{\#G} v_1 = v_1. \qquad \boxtimes$$

Maschke's theorem holds whenever one can average as in (5), i.e. whenever it makes sense to sum over G and then divide out. In particular, this argument applies to compact topological groups that aren't necessarily finite. For example, one has the Lie group $\mathrm{SU}(2)$, for which the sum becomes an integral $\int_G g^{-1} \pi g \, \mathrm{d}g$, where $\int_G \mathrm{d}g = 1$ (so that the average is taken, not the sum). Since $\mathbb{Z}_p$ is also topological, but is the direct limit of finite groups, something between a sum and an integral can handle it, so one also obtains similar theorems in this case.

Lemma 19.3 (Schur)**.** *If (ρ, V) is an irreducible representation of a finite group G, then $\mathrm{End}_G(V)$ is a division k-algebra (i.e. all nonzero elements are invertible).*

Proof. If $T : V \to V$ is a nonzero G-map, then $\ker(T)$ s a G-stable subspace of V. Since V is irreducible, then $\ker(T) = 0$ or $\ker(T) = V$, but $T \neq 0$, so $\ker(T) \subsetneqq V$. Thus, $\ker(T) = 0$, so T is injective, and therefore an isomorphism, so it is also invertible.

It's also necessary to check that T^{-1} is a G-map, but this is not hard. $\boxtimes$

This simple lemma is surprisingly useful: take $k = \mathbb{R}$ and G finite, and let V be an irreducible $\mathbb{R}$-vector space. Then, $\mathrm{End}_G(V)$ is a finite-dimensional division algebra containing $\mathbb{R}$, so there are three possibilities: $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$. These are the only division algebras over $\mathbb{R}$, and all of them appear in representations:

- Take the trivial representation $\rho(g) = 1$ for all $g \in G$, so that one has $\mathbb{R}$.
- If $G = \mathbb{Z}/3 = \langle x \rangle$ and $\rho(x) = \sigma = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ for $\theta = 2\pi/3$, then the matrices that commute with this action are $\mathrm{End}_G(V)$, which is a similar angle matrix for any angle, composed with a scaling, This action is multiplication by the complex number $re^{i\theta}$, so thus $\mathrm{End}_G(V) = \mathbb{C}$.
- Consider the group $G = \{\pm 1, \pm i, \pm j, \pm k\}$ in the Hamilton quaternions, i.e. $i^2 = j^2 = k^2 = 1$, $ij = k$, $jk = i$, and $ki = j$. This is a group of order 8 contained in $\mathbb{H}$, so there's an inclusion $G \hookrightarrow \mathbb{H}$, and then $\mathbb{H}$ acts on itself by left multiplication: $\mathbb{H} \to \mathrm{GL}(\mathbb{H})$. Composing these, G has a four-dimensional real representation ρ. It's a bit of work to show this is irreducible, but then $\mathrm{End}_G(\mathbb{H}) \supseteq \mathbb{H}$, because right multiplication gives an inclusion $\mathbb{H} \to \mathrm{End}_G(\mathbb{H})$, because left and right multiplication commute, and specifically, this action and that of G. Since $\mathbb{H}$ is a division algebra, this must be injective, and then one can show it's an isomorphism.

Later, we will introduce character theory in part to detect which case happens.

Theorem 19.4. *Suppose $\#G$ is prime to $\mathrm{char}(k)$, and let $V_1, \ldots, V_n$ be all of the irreducible representations of G over k (i.e. up to isomorphism; choose one representative from each isomorphism class). First, there are in fact only finitely many of them. Second, let $D_i = \mathrm{End}_{G_i}(V_i)$, which is a division algebra by Lemma 19.3; then, each V_i is a vector space over D_i, and in fact $\mathrm{End}_{D_i}(V_i) \cong \mathrm{Mat}_{d_i}(D_i)$, where $d_i = \dim_{D_i}(V_i)$.*

Then, there is a ring isomorphism

$$k[G] \xrightarrow{\sim} \prod_{i=1}^{r} \mathrm{End}_{D_i}(V_i).$$

The notion of dimension of a module over a division algebra turns out to work; modules over division algebras behave in much the same way as vector spaces.

20. Group Characters: 12/4/13

"Oh yeah? Well what if I don't love [representation theory]?"
"Then it will build character." – Calvin and Hobbes (more or less)

Let G be a finite group and V be a representation of G over an algebraically closed field k (i.e. $k = \overline{k}$, which is the most commonly used notation), and let the representation be given by $\rho : G \to \mathrm{Aut}_k(V)$. Assume also that $\mathrm{char}(k) \nmid \#G$, so that Maschke's theorem applies: every finite-dimensional representation of G is a direct sum of irreducible representations.

For concreteness, one can think of $k = \mathbb{C}$.

Definition. The character of V is $\chi_V(g) = \mathrm{Tr}(g \mid V)$ (i.e. writing g as a matrix over V; the specific matrix doesn't matter because the trace will still be the same).

These functions have several nice properties. First, characters are invariant under conjugation: $\chi_V(hgh^{-1}) = \chi_V(g)$ for any $g, h \in G$, because $\rho(hgh^{-1}) = \rho(h)\rho(g)\rho(h)^{-1}$, so $\rho(hgh^{-1})$ and $\rho(g)$ are similar matrices, and thus have the same trace. Note, however, that it is generally *not* true that $\chi(gh) = \chi(g)\chi(h)$, unless V is one-dimensional.

Definition. A function $f : G \to k$ is called a class function if $f(hgh^{-1}) = f(g)$ for all $g, h \in G$, i.e. it is invariant under conjugation.

Thus, characters are class functions. The term "class function" is used because these functions can be thought of as functions on the conjugacy classes of G; later, it will be shown that the characters of irreducible representations of G over k form a basis for the space of class functions.

Given a representation V, G also acts on the dual space $V^* = \mathrm{Hom}_k(V, k)$ by $\varphi \overset{g}{\mapsto} g \circ \varphi$ given by $g \circ \varphi(v) = \varphi(g^{-1} \cdot v)$ for $v \in V$. The inverse is necessary so that $g_1 \cdot (g_2 \cdot \varphi) = g_1 g_2 \cdot \varphi$ rather than $g_2 g_1 \cdot \varphi$; they need to be in the right order. In this context, V^* is called the dual representation to V.

On characters, $\chi_{V^*}(g) = \chi_V(g^{-1})$. This is because if one chooses a basis for V, the action of g has some matrix A in this basis. Then, in the dual basis for V^*, g-action gives another matrix B. One can check that $B = (A^{-1})^{\mathrm{T}}$ (the transpose-inverse), so when one takes the trace, the transpose dies.

Given two representations V and W, their direct sum is also a representation, and the characters add: $\chi_{V \oplus W} = \chi_V + \chi_W$. Similarly, their tensor product is a representation by $g \cdot (v \otimes w) = (g \cdot v) \otimes (g \cdot w)$, which extends to the whole tensor product linearly. Then, $\chi_{V \otimes W} = \chi_V \cdot \chi_W$; that is, pointwise multiplication of the original characters. This is because if one picks bases for V and W, sending the action of g to A and B respectively, then its action on $V \otimes W$ is $A \otimes B$, i.e. the matrix whose $(i, j)^{\text{th}}$ block entry is $b_{ij}A$. Then, taking the trace shows that the result is the product.

Induction. The above techniques provide useful ways to construct new representations, but generally $V \oplus W$ and $V \otimes W$ aren't irreducible. Something called induction is more useful.

Let $H < G$ and W be a representation of H, given by $\sigma : H \to \mathrm{Aut}_k(W)$. Then, define the vector space $\mathrm{Ind}_H^G(W) = \{f : G \to W \mid f(gh) = \sigma(h^{-1})f(g) \text{ for all } g \in G, h \in H\}$. Addition and scalar multiplication in $\mathrm{Ind}_H^G(W)$ are unsurprising, but it is also a representation of G, with action given by $f \mapsto g \cdot f$, where $(g \cdot f)(x) = f(g^{-1}x)$; in other words, $g \cdot f : G \to W$.

Example 20.1. If W is the trivial representation, then $\mathrm{Ind}_H^G(k) = \{f : G/H \to k\}$, since the condition forces invariance under right translation by H. But G still acts by left translation. This representation is reducible, since it contains the constant functions, which form a one-dimensional k-vector space, and this containment is strict whenever $H \neq G$.

Sometimes, though, the induced representation is irreducible. If $G = D_{2n}$ and $H = \mathbb{Z}/n$, then H is a normal subgroup of G, realized by all of its rotations. Then, $G/H \cong \mathbb{Z}/2\mathbb{Z}$. The irreducible representations of H are classified by the n^{th} roots of unity in k, because by Schur's lemma, all of the irreducible representations of an abelian group are one-dimensional when k is algebraically closed. Let $k(\zeta)$ be the space given when $1 \in G$ acts as $\zeta \in k^*$, where ζ is some n^{th} root of unity.

Then, $\mathrm{Ind}_H^G(k(\zeta))$ is always two-dimensional, because if $f : G \to k(\zeta)$ is such that $f(gh) = \rho_\zeta(h)^{-1}f(g)$, then f is determined by its value on a set of representatives on G/H; there's a unique way to extend these to a full function.

More generally, $\dim \mathrm{Ind}_H^G(W) = [G : H] \dim W$.

Suppose n is odd and $\zeta \neq 1$ (since that case gives a reducible representation, as seen above). Then, $\mathrm{Ind}_H^G(k(\zeta))$ is irreducible: suppose that it weren't, and then, because it's two-dimensional, there will be a one-dimensional subrepresentation $V \subset \mathrm{Ind}_H^G(k(\zeta))$ that is stable under G. But since k^* is an abelian group (under multiplication), then any one-dimensional representation must factor through the abelianization $G/[G, G] = G/H$ as $G \to G/H \to k^*$, so V is either trivial or the sign representation, sending a rotation to 1 and a reflection to -1. If V is trivial, then as seen above the constant functions are a subspace of $\mathrm{Ind}_H^G(k(\zeta))$, but the constant functions don't satisfy $f(gh) = \rho_\zeta(h)^{-1}f(g)$, and the sign representation has a similar problem.

When n is even and $\zeta \neq \pm 1$ (since this time, -1 is an n^{th} root of unity), $\mathrm{Ind}_H^G(k(\zeta))$ is irreducible. When $\zeta = 1$, $\mathrm{Ind}_H^G(k) = \{f : G/H \to k\} = \text{trivial} \oplus \text{sign}$, and $\mathrm{Ind}_H^G(k(-1)) = V \oplus V'$ for some nontrivial one-dimensional representations V and V'. Specifically, they are nontrivial when restricted to $k(-1)$ as H-representations.

However, not all of these representations are distinct: $\mathrm{Ind}_H^G(k(\zeta)) = \mathrm{Ind}_H^G(k(\zeta^{-1}))$, and this is still true when n is odd. Geometrically, all of the roots of unity form an n-gon inscribed in the unit circle; then, ζ and ζ' give the same values if they have the same real part, i.e. there's a vertical line between them. This means when n is odd, there is one reducible representation corresponding to $\zeta = 1$, and $(n-1)/2$ irreducible ones, and when n is even, two

reducible representations (because both 1 and -1 are special) and $(n-2)/2$ irreducible ones. In both cases, this gives the complete list of irreducible representations.

Using the Group Ring. One way to understand representations is to use the ring structure of the group ring.

Suppose V is an irreducible representation of a group G over a field k, writing once again $\rho : G \to \mathrm{Aut}_k(V)$. Then, there is a k-linear map $k[G] \xrightarrow{\pi_V} \mathrm{End}(V)$ given by $[g] \mapsto \rho(g)$. This is actually a ring homomorphism (of noncommutative rings), because ρ is a representation, so $[g] \cdot [h] = \rho(gh) = \rho(g)\rho(h)$. But it also interacts nicely with the action of G. The left translation action of G on $k[G]$, called $L_g : k[G] \to k[G]$[57] sends $[x] \mapsto [gx]$, which extends to a k-linear group automorphism (*not* a ring automorphism) of $k[G]$. Similarly, one has a right translation action $R_g : k[G] \to k[G]$ given by $[x] \mapsto [xg^{-1}]$ (again so that things come out in the right[58] order: $R_{g_1} \circ R_{g_2} = R_{g_1 g_2}$). These two actions obviously commute with each other, so there is a $G \times G$ action on $k[G]$ as (g_1, g_2) acts by $L_{g_1} \circ R_{g_2}$ (and the order can be switched). This is the extra structure we want — and it helps us put some extra structure on $\mathrm{End}(V)$, too. $\mathrm{End}(V) = V \otimes V^*$, and G acts on both components, so $G \times G$ acts on $\mathrm{End}_G(V)$ by $(g_1, g_2)(v \otimes v^*) = (g_1 v) \otimes (g_2 v^*)$. Note that one could also write $\mathrm{End}(V) = V^* \otimes V$, but the notation given above is more conventional.

Lemma 20.1. *$\pi_V : k[G] \to \mathrm{End}(V)$ is a map of $(G \times G)$-representations (i.e. it's equivariant under the $G \times G$ action); that is, L_g becomes the action on the first component V, and R_g becomes the action on the second component V^*.*

Proof. Exercise.

Now, let $\{V_i\}$ be the complete list of irreducible representations of G up to isomorphism, given by choosing one from each isomorphism class, and let $\pi : k[G] \to \prod_i \mathrm{End}(V_i)$ be given as the product of these π_{V_i} (so that it is a $(G \times G)$-map and a ring homomorphism).

Claim. Then, π is an isomorphism.

Proof. The goal is to produce an inverse $\mathrm{End}(V_i) \to k[G]$. Since $\mathrm{End}(V_i) = V_i \otimes V_i^*$, then produce a function in $k[G]$ from $v \otimes v^*$ by sending $v \otimes v^* \mapsto [x \in G \mapsto v^*(x^{-1} \cdot v)]$.

To convince you that this is correct, consider the following, where $V = V_i$ is used to simplify the notation:

$$\begin{array}{ccc} v \otimes v^* & \longmapsto & [x \xmapsto{\varphi} v^*(x^{-1}v)] \\ \downarrow & & \downarrow \\ gv \otimes v^* & \longmapsto & [x \xmapsto{(L_g\varphi)(x)} v^*(x^{-1} \cdot gv)] \end{array}$$

Then, since $(L_g\varphi)(x) = \varphi(g^{-1}x)$, then the last map sends x to $v^*((g^{-1}x)^{-1} \cdot v)$, so the diagram commutes. The right G-action also makes its diagram commute in a similar "completing the square"-type proof, which is left as an exercise.

The left translation function was defined on the generators of $k[G]$, but can be extended in the obvious way. The map $m_i : \mathrm{End}(V_i) \to k[G]$ is called the matrix coefficient: $v^*(x^{-1}v)$ is just writing x^{-1} in the standard basis and then taking a specific component. Moreover, it is nonzero, which can be shown by choosing good values for v and v^*. So now there's a good candidate for the inverse map

$$\mathrm{End}(V_i) \xrightarrow{m_i} k[G] \xrightarrow{\pi} \prod_i \mathrm{End}(V_i)$$

The first question is: why is this product even finite? Otherwise, there's no chance of an isomorphism whatsoever. However, this question can be answered.

Claim. There exist only finitely many irreducible representations up to isomorphism.

Proof. Suppose V is irreducible; then, the map $V \otimes V^* \to k[G]$ is nonzero. Then, G acts on the first factor by left translation, but since V is irreducible, then $V \otimes V \cong V^{\oplus d}$ as a $k[G]$-module (that is, the second action is temporarily ignored). Thus, V appears as a subrepresentation of $k[G]$ when the latter is viewed as a G-representation by left translation. But $k[G]$ is a finite-dimensional vector space, so there can be only finitely many such irreducible representations up to isomorphism. ⊠

[57] The function L_g is called "left translation by g" or sometimes just "left g."

[58] No pun intended.

Now, because the product is known to be finite, the above composition of maps becomes this chain of $(G \times G)$-maps:

$$\prod_i \operatorname{End}(V_i) \xrightarrow{m=(m_i)} k[G] \xrightarrow{\pi=(\pi_i)} \prod_i \operatorname{End}(V_i).$$

Since V_i and V_i^* are irreducible, then $\operatorname{End}(V_i)$ is an irreducible representation of $G \times G$. But different irreducibles don't talk to each other in this direct sum, and they were chosen so as not to be isomorphic, so the composition map $\operatorname{End}(V_i) \to k[G] \to \operatorname{End}(V_j)$ must be zero when $i \neq j$. Then, it remains to compute the map when $i = j$; it ends up being scalar multiplication, where the scalar comes out to $(\dim V_i)^{-1} \# G$. This in some sense encodes information about the representation, and since it's nonzero, then $\pi \circ m_i$ is an isomorphism of vector spaces. This proves π is surjective.

Suppose that π weren't injective; then, $f \in \ker(\pi)$ acts as zero on each irreducible representation V_i. But $k[G] = \bigoplus V_i^{\oplus d_i}$ for some d_i, so f must also act on $k[G]$ by zero, just by left-multiplication (maybe with an inverse in there somewhere). But $k[G]$ has an identity, so this can't possibly work. Thus, $f = 0$, so π is injective, too. ⊠

21. Character Tables: 12/6/13

> *"To first-order approximation you can ignore [this inverse], but when you're teaching it to someone else you have to be careful."*

We ended the last lecture with $\bigoplus \operatorname{End}(V_i) \to k[G] \to \bigoplus \operatorname{End}(V_i)$, where k is an algebraically closed field such that $\operatorname{char}(k) \nmid \# G$, and that both maps are $G \times G$ maps and bijections. Having both of these G-actions made $\operatorname{End}(V_i)$ irreducible, and the composition of the two actions was just multiplication by $\# G / \dim V$.

The second map was $[g] \mapsto (\rho_i(g)_i$, and the first was called "matrix coefficients:" if one tensors a covector and a vector, $v_i^* \otimes v_i \mapsto [g \mapsto v_i^*(\rho_i(g)^{-1} \cdot v_i)]$.

Here are some consequences of this isomorphism:

(1) $\sum (\dim V_i)^2 = \dim k[G] = \# G$.
(2) $\dim V_i \mid \# G$, so the multiplication given by composing the two actions is actually by an integer. These maps preserve an integral structure, which draws in algebraic number theory (in essence, since it must be an algebraic integer and must be rational, then it must be an integer).
(3) The number of irreducible representations of G is equal to its number of conjugacy classes.

Each V_i is associated with a character χ_{V_i}, which is a class function on G, i.e. it is constant on all conjugacy classes. Thus, one can draw a character table that collects all of the character values.

Proof of (3). Consider $G \hookrightarrow G \times G$ diagonally, sending $g \mapsto (g, g)$, and restrict to this diagonal action, acting on $k[G]$ by conjugating the basis elements. In $\operatorname{End}(V_i)$, it's also conjugation, and the inverse trickles down somewhere from the definitions of these actions.

G-invariants on $k[G]$ are just class functions, as a vector space, and on the endomorphism ring are $\operatorname{End}_G(V_i)$, so the isomorphism of class functions to $\bigoplus_i \operatorname{End}_G(V_i)$. But since k is algebraically closed, then by Schur's lemma, $\operatorname{End}_G(V_i) = k$.

Thus, the character table is actually a change-of-basis matrix: one basis is given by the conjugacy classes of G (i.e. 1 on a given class and 0 on all others), and another is given from $\bigoplus \operatorname{End}_G(V_i)$, which ends up becoming the characters of the irreducible representations. ⊠

The above proof contained some other useful information, too.

Example 21.1. Let $G = S_4$. Then, the conjugacy classes are $\{1\}$, (1 2), (1 2)(3 4), (1 2 3), and (1 2 3 4), because the length of a cycle is invariant under conjugation. Thus, without much further knowledge, the character table looks like Table 1. In this table, "triv." denotes the trivial representation, and the sign representation sends even

	1	(1 2)	(1 2)(3 4)	(1 2 3)	(1 2 3 4)
triv.	1	1	1	1	1
sgn.	1	−1	1	1	−1
std.	3	1	−1	0	−1
ρ_4					
ρ_5					

Table 1. The partial character table of S_4, constructed by just writing down some irreducible representations. Some trickery is necessary to generate the last two rows.

permutations to 1 and odd ones to -1. Then, $S_4 \circlearrowright k^4$ by permuting the basis elements, but $\operatorname{Span}\{(1,1,1,1)\}$ is

S_4-stable; its 3-dimensional complement, the "standard representation," is irreducible and appears in the table. Its character can be computed from the character of the four-dimensional representation: since these are permutations and therefore act as permutation matrices, $\mathrm{Tr}(g)$ is equal to the number of fixed points when g acts on $\{1,2,3,4\}$. Thus, to obtain the reduced character, subtract 1 from the values.

The remaining representations can be obtained by tensoring those already enumerated, and thus multiplying the characters. This leads to irreducible representations, though many of them are things we've already listed. However, tensoring the standard and sign representations makes something new.

The last representation is d-dimensional, where $24 = 1^2 + 1^2 + 3^2 + 3^2 + d^2$, because the sum of the squares of the dimensions of $\#S_4$, so it must be a two-dimensional representation. To describe it better, though, requires the orthogonality relations, which impose a (not quite literal) orthogonal structure on this matrix.

Proposition 21.1 (Orthogonality Relations)**.** *Let $\chi_1, \dots, \chi_r$ be the characters of the irreducible representations of G over k, up to isomorphism (i.e. one representative from each isomorphism class). Then,*

(1)
$$\frac{1}{\#G}\sum_{g\in G}\chi_{V_i}(g)\chi_{V_j}(g) = \begin{cases} 1, & V_i \cong V_j^*, \\ 0, & \text{otherwise.} \end{cases}$$

(2)
$$\sum_{i=1}^{r}\chi_{V_i}(g_1)\chi_{V_i}(g_2) = \begin{cases} \#C_G(g_1), & g_1 \sim g_2^{-1} \\ 0, & g_1 \not\sim g_2^{-1}. \end{cases}$$

Here, $\sim$ denotes conjugacy to and $C_G(g_1) = \{h \in G \mid hg_1 = g_1h\}$ is the centralizer of g_1, the set of elements it commutes with.

An equivalent reformulation of part (1) is that

$$\frac{1}{\#G}\sum_{g\in G}\chi_{V_i}(g)\chi_{V_j}(g^{-1}) = \begin{cases} 1, & i = j \\ 0, & \text{otherwise,} \end{cases} \tag{6}$$

because $\chi_{V^*}(g) = \chi_V(g^{-1})$ as shown the other day. In English, the first relation means that the inner product of two rows (over the elements of the group, not over the conjugacy classes) is nonzero iff they are dual. The second relation provides information about the columns, saying that they are orthogonal unless the conjugacy classes are inverse to each other.

In S_4 in particular, every conjugacy class is closed under the taking of inverses, which will make life a little simpler. In particular, it allows the last row to be found. Another trick allows $\chi_5((1\ 2))$ and $\chi_5((1\ 2\ 3\ 4))$ to be found: since $\mathrm{sign} \otimes \rho_5 = \rho_5$ (since there isn't room for any more irreducible representations), then multiplying $\chi_5((1\ 2))$ by -1 cannot change it, and similarly with $(1\ 2\ 3\ 4)$, so they must go to zero. Thus, the complete character table is in Table 2.

	1	(1 2)	(1 2)(3 4)	(1 2 3)	(1 2 3 4)
triv.	1	1	1	1	1
sgn.	1	−1	1	1	−1
std.	3	1	−1	0	−1
std ⊗ sgn	3	−1	−1	0	1
ρ_5	2	0	2	−1	0

TABLE 2. The complete character table of S_4, using orthogonality relations to generate the final rows.

Proof of Proposition 21.1. For part 1, specifically the reformulation (6), which is equivalent, consider the G-representation $\mathrm{Hom}(V_j, V_i)$, and look at its G-action by conjugation. $g \circ \varphi = \rho_j(g) \circ \varphi \circ \rho_j(g^{-1})$, so writing $\mathrm{Hom}(V_j, V_i) = V_j^* \otimes V_i$, then $\chi_{\mathrm{Hom}(V_j,V_i)} = \chi_{V_i}(g)\chi_{V_i^*}(g) = \chi_{V_i}(g)\chi_{V_j}(g^{-1})$. Thus,

$$\frac{1}{\#G}\sum_{g\in G}\chi_{\mathrm{Hom}(V_j,V_i)}(g) = \mathrm{Tr}\left(\overbrace{\frac{1}{\#G}\sum_{g\in G} g\Big|_{\mathrm{Hom}(V_j,V_i)}}^{e}\right).$$

But e is a projector: it's easy to check that $e^2 = e$ by averaging out over the standard basis of the group algebra, and the effect of e is to project $V \twoheadrightarrow V^G$ (i.e. the G-invariant subset). This is easy to show: if $v \in V$ is already G-invariant, then $e \cdot v = v$ (since by definition it can't do anything), and for v in general, averaging over G produces something G-invariant.

Then, the trace of a projector is just the dimension of the subspace it projects onto, so $\mathrm{Tr}(e) = \dim \mathrm{Hom}(V_i, V_j)^G = \dim \mathrm{Hom}_G(V_i, V_j)$, which is equal to 1 or 0 depending on whether they're the same.

The second relation is a formal consequence of the first, and that $\chi_{V^*}(g) = \chi_V(g^{-1})$, since any matrix whose columns are orthonormal also has orthonormal rows. But a direct proof exists. Since $\mathrm{End}(V_i)$ acts on V_i and V_i^* in two components, then

$$\begin{aligned}\sum_{i=1}^{r} \chi_{V_i}(g_1)\chi_{V_i^*}(g_2) &= \sum_{i=1}^{r} \chi_{\mathrm{End}(V_i)}(g_1, g_2^{-1}) \\ &= \mathrm{Tr}\left((g_1, g_2^{-1}) \mid \bigoplus \mathrm{End}(V_i)\right) \\ &= \mathrm{Tr}((g_1, g_2^{-1} \mid k[G]).\end{aligned}$$

Here, $\mathrm{Tr}(a \mid V)$ again indicates the trace of a over V. The action on $k[G]$ is by left and right translation, sending $[x] \mapsto g_1 x g_2$. Thus, it's a permutation action, so its trace is the number of fixed points under this action, i.e. $\{x \in G \mid x^{-1} g_1 x = g_2^{-1}\}$, which is empty if g_1 and g_2^{-1} aren't conjugate, and is equal to $C_G(g_1)$ if they are: if $x_0^{-1} g_1 x_0 = g_2^{-1}$, then any other x in that set is of the form $y x_0$ such that y commutes with g_1, so there is a bijection ith $C_G(g_1)$, albeit noncanonically. Another way of thinking about this is that $C_G(g_1)$ acts on the set transitively, so they must have the same cardinality. ⊠

Working specifically in $\mathbb{C}$, these two orthogonality relations can be restated as follows: constrct the normalized character table by $\chi_V(g) \mapsto \chi_V(g)/\sqrt{\#C_G(g)}$; then, this table is a unitary matrix, i.e. $\sum a_i \bar{b}_i = 0$ if $a_i \neq b_i$, and $\sum |a_i|^2 = 1$.

This depends on the fact that for ever complex representation of G, the averaging trick shows that it contains a G-invariant positive-definite Hermitian form. In other words, $V \cong \overline{V^*}$ (i.e. first taking the dual, then the conjugate; the isomorphism is as vector spaces. In some sense, this is changing the $\mathbb{C}$-action, as conjugation is a $\mathbb{C}$-linear isomorphism). Thus, $\chi_V(g) = \overline{\chi_V(g^{-1})}$.

A similar argument shows that over $\mathbb{R}$, each real representation admits a positive definite invariant quadratic form, which means that all representations are self-dual.

Since the characters are traces of some matrix, $\chi_V(g) = \mathrm{Tr}(g \mid V)$ such that $g^N = 1$ (since in a finite group, all elements must have finite order), then $\chi_V(g)$ is a sum of n^{th} roots of unity, as the eigenvalues are all roots of unity. Thus, the character actually lies in $\mathbb{Q}^{\mathrm{ab}} = \bigcup_{\zeta_n \text{ root of unity}} \mathbb{Q}(\zeta_n)$, an infinite extension of $\mathbb{Q}$ (the "abelian" name is kind of poorly chosen). But it's even possible to do better: it's necessarily generated by n^{th} roots, so it's in $\mathbb{Q}(\zeta_n)$. But $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}/n)^\times$, so this Galois group permutes the rows of the character table! This is because a Galois automorphism permutes these matrices, thus sending representations to representations. But $(\mathbb{Z}/n)^\times$ also permutes the columns: if k is invertible mod n, i.e. $k \in (\mathbb{Z}/n)^\times$, then $g \mapsto g^k$ sends conjugacy classes to conjugacy classes.

These two actions have different structures, but impressively (and very usefully), they are compatible. Let $\sigma_k \in \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ be given by $k \in (\mathbb{Z}/n)^\times$; then, $\chi_V(g) = \chi_{\sigma_k V}(g^k)$ and $\chi_{\sigma_k V}(g) = \chi_V(g^k)$; that is, these diagonal points are equal in the character table. This is pretty nifty, and in the case $k = -1$ (which is always invertible) reduces to $\chi_V(g) = \chi_{V^*}(g^{-1})$.

References

[1] Clayburgh, Jill. *It's My Turn*. Columbia Pictures Corporation, 1980.

[2] Milnor, J.W., Dale Husemöller. *Symmetric Bilinear Forms*, Springer-Verlag, Ergebnisse der Mathematik und ihrer Grenzgebiete, 1973.