MATH 412 NOTEBOOK

JUSTIN KATZ

1. Some problems on rings and ideals

Let C be compact, and define

$$R = \{f : C \to \mathbb{C} | f \text{ is continuous}\},$$

and equip R with pointwise addition and multiplication to make it a commutative ring with 1.

Claim 1. *For any proper ideal I of R, there is some point $x \in C$ such that $f(x) = 0$ for all $f \in R$.*

Proof. Suppose I is such that for all $x \in C$ there is some function $f_x \in R$ such that $f_x(x) \neq 0$. Further, because each f_x is continuous, there is some nonempty neighborhood U_{f_x} of x such that f_x is nonzero on all of U_{f_x}. The collection $\{U_{f_x}\}_{x \in C}$ forms an open cover of C, so by compactness we can reduce to a finite subcover $\{U_{f_{x_1}}, ..., U_{f_{x_n}}\}$. Rename the functions $f_1, ..., f_n$ and the corresponding opens $U_1, ..., U_n$. By strong closure of I in R, and the fact that complex conjugation is continuous on $\mathbb{C}$, the nonnegative functions $f_i \cdot \bar{f}_i$ are also in R. Their sum, $\sum_{i=1}^{n} f_i \cdot \bar{f}_i$ is a non vanishing continuous function on C and thus has a continuous inverse. By strong closure, $1 \in I$ so $I = R$. □

The above claim demonstrates that the only possible candidates for maximal ideals in R are those of the form

$$I_x = \{f \in R : f(x) = 0\}.$$

Indeed, every ideal in R has *at least one* common zero, so every ideal is contained in some I_x. The cause for worry is that there may be further relaxations on the collection of continuous functions that vanish at x that might lead to a bigger proper ideal than I_x. The next claim shows that this cannot happen.

Claim 2. *For $x \in C$ the ideal I_x is maximal*

Proof. First, recall that in any commutative ring with 1, A with ideal J. we have the equivalence

$$A \text{ is maximal } \iff A/J \text{ is a field.}$$

(To convince oneself, recall that an ideal of a quotient A/J is of the forml K/J where K is a superideal of J, and conversely a superideal of J will descend to an ideal of A/J, but fields have no nontrivial proper ideals.)

Introduce the evaluate-at-x homomorphism: $\phi_x : R \to \mathbb{C}$ defined by $\phi_x(f) = f(x)$. One should convince oneself that ϕ_x is indeed a homomorphism by appealing to the pointwise definition of the ring operations. Observe that ϕ_x surjects R onto $\mathbb{C}$ by considering the image under ϕ_x of the constant functions on C. Thus, by one of the isomorphism theorems, we have $R/I_x = R/\ker(\phi_x) \cong \phi_x(R) = \mathbb{C}$. Maximality of I_x follows from $\mathbb{C}$ being a field. □

2. Some problems on modules over a PID

Let A be a PID, $n \geq 2$ be an integer, and $v = (a_1, \ldots, a_n) \in A^{\oplus n}$ be such that $\gcd(a_1, .., a_n) = 1$.

Claim 3. *If a nonzero $v' = (a'_1, \ldots, a'_n) \in A^{\oplus n}$ is such that $av' \in Av$ for a nonzero a, then $v' \in Av$ itself.*

Proof. If $av' = bv$ then in each component a divides ba_i, so[1] $a/\gcd(a, b)$ divides each component a_i, meaning $a/gcd(a, b) \cdot v' = b/gcd(a, b)v$ but $\gcd(v) = 1$ so $a/\gcd(a, b) = 1$, which means $v' = b/\gcd(a, b) \cdot v$. Thus $v' \in Av$. □

With this claim in hand, we can show that quotienting a free module by a submodule generated by a primitive vector is torsion free.

Claim 4. *The quotient $A^{\oplus n}/Av$ is torsion free.*

Proof. Let π be the canonical projection of $A^{\oplus n}$ onto $A^{\oplus n}/Av$. Let $q \in A^{\oplus n}/Av$ and suppose for some a, $aq = 0_{A^{\oplus n}/Av}$. Write $q = \pi(s)$ for some $s \in A^{\oplus n}$, so that $aq = a\pi(s) = \pi(as) = 0_{A^{\oplus n}/Av}$, because π respects algebra. This is to say that $as \in \ker(\pi)$, which by definition of the canonical projection onto a quotient is exactly Av. By the last claim, this means $s \in \ker(\pi)$ so $q = \pi(s) = 0$. Thus, $A^{\oplus n}/Av$ is torsion free. □

[1]Notation may be misleading, I am not assuming that $gcd(a, b)$ is invertible, I am thinking division with (zero) remainder

By the structure theorem for finitely generated modules over a PID, there is a basis $(f_1, .., f_n)$ of $A^{\oplus n}/Av$ and a nonzero ideal $\mathfrak{a}_1$ such that

$$A^{\oplus n} = Af_1 \oplus \cdots \oplus Af_1$$
$$Av = \mathfrak{a}_1 f_1$$
$$A^{\oplus n}/Av = (A/\mathfrak{a}_1 f_1) \oplus Af_2 \oplus \cdots \oplus Af_n$$

and $\mathfrak{a}_1$ annihilates $A/\mathfrak{a}_1$.

Claim 5. $\mathfrak{a}_1 = A$

Proof. By the last claim, $A^{\oplus n}/Av$ is torsion free. By the structure theorem, the nonzero ideal $\mathfrak{a}_1$ annihilates the $A/\mathfrak{a}_1$ component of vectors in $A^{\oplus n}/Av$. Thus the $A/\mathfrak{a}_1$ component must be trivial, proving the claim. □

Because $S = \mathfrak{a}_1 f_1 = Af_1$ we can take $f_1 = v$, and $\{v, f_2, \cdots, f_n\}$ is a basis for $A^{\oplus n}$. Let $\{e_1, \cdots, e_n\}$ be the standard basis for $A^{\oplus n}$ and define the A-linear map $T : A^{\oplus n} \to A^{\oplus n}$ by $T(e_1) = v$, $T(e_2) = f_2$, and so on, then extending by linearity. As a bijection between bases, T is an A-module isomorphism, so it has an A-invertible determinant. The matrix of T in the standard basis is

$$[T] = \begin{bmatrix} v & f_2 & \cdots & f_n \end{bmatrix}.$$

When $n \geq 2$, $v \neq f_n$, and we can scale the last column (which is *not* the first column) by $1/(\det T)$. By the multilinearity of the determinant, the resulting matrix will have determinant 1.

3. More problems on modules over a PID

The special linear group over the integers $SL_2(\mathbb{Z})$ acts on the upper half plane $\mathcal{H}$ by fractional linear transformations:

$$\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \times \tau \mapsto \frac{a\tau + b}{c\tau + d}.$$

This exercise determines the points in $\mathcal{H}$ which are fixed by nontrivial elements of $SL_2(\mathbb{Z})$.

Claim 6. *If* $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in SL_2(\mathbb{Z})$ *fixes an point* $\tau \in \mathcal{H}$ *then* $|a + d| < 2$.

Proof. If $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \tau = \tau$ then $a\tau + b = c\tau^2 + d\tau$. Immediately c is nonzero, because τ is not rational. Furthermore, the roots to the above quadratic polynomial in τ are

$$\tau = \frac{-(d-a) \pm \sqrt{(d-a)^2 + 4cb}}{2c}.$$

In order for τ to be in the upper half plane, $\sqrt{(d-a)^2 + 4cb}$ must be purely imaginary. That is, $(d-a)^2 + 4cb < 0$. Expanding the square, adding a convenient $-2ad + 2ad = 0$, and using the fact that $ad - bc = 1$ gives the computation

$$(d-a)^2 + 4cb = d^2 - 2ad - a^2 + 4cb = d^2 + 2ad - a^2 + 4cb - 4ad = (d+a)^2 - 4 < 0$$

and the claim is immediate. □

Notice that that the above quadratic equation shows that any element of $SL_2(\mathbb{Z})$ can fix at most two points, and so long as they are both not real, the two points will conjugates of one another, so exactly one of the fixed points will be in the upper half plane. Thus, each element of $SL_2(\mathbb{Z})$ can fix at most one point in $\mathfrak{H}$. One should note that the claim says that for any $\gamma \in SL_2(\mathbb{Z})$ that fixes a point, $\mathrm{Tr}(\gamma) \in \{-1, 0, 1\}$. This observation leads to a quick proof of

Claim 7. *If γ has order 2 then $\gamma = -I$*

Proof. To say that γ has order 2 is to say that γ satisfies the polynomial $x^2 - 1$. Thus, the minimal polynomial of γ divides $x^2 - 1$. The candidates for the minimal polynomial for γ are

- $x^2 - 1$
- $x - 1$
- $x + 1$

We can immediately eliminate the second listed polynomial, because the only matrix that satisfies $x - 1$ is the identity, and we have assumed that the order of γ is strictly more than 1. On the other hand, by the Cayley-Hamilton theorem, γ satisfies its characteristic polynomial $x^2 - \mathrm{Tr}(\gamma)x + 1$, so the minimal polynomial of γ must divide $x^2 - \mathrm{Tr}(\gamma)x + 1$. Further, the minimal polynomial of γ will divide the difference $(x^2 - 1) - (x^2 - \mathrm{Tr}(\gamma)x + 1) = \mathrm{Tr}(\gamma)x - 2$ which is linear. Thus, the only possible minimal polynomial of γ is $x + 1$ so $\gamma = -I$. □

In the given classification of point fixing elements of $SL_2(\mathbb{Z})$, a claim is made that for any vectors $u, v \in \mathbb{Z}^2$ if $u\mathbb{Z} + v\mathbb{Z} = \mathbb{Z}^2$ then $\det \begin{bmatrix} u & v \end{bmatrix} \in \{1, -1\}$. To verify this, observe that by assumption any $x \in \mathbb{Z}^2$ can be written as an integer linear combination of u and v. In particular $(1, 0)$ and $(0, 1)$

can be written as linear combinations of u and v, so define the $\mathbb{Z}$ module map $T : \mathbb{Z}^2 \to \mathbb{Z}^2$ by $u \mapsto (1,0)$ and $v \mapsto (0,1)$. The observation in the sentence before last rephrases as T is a $\mathbb{Z}$ module isomorphism, meaning T is invertible as a $\mathbb{Z}$ module map. Thus $\det T \in \mathbb{Z}^\times = \pm 1$. To conclude, observe that the matrix representation of T with respect to the standard basis of $\mathbb{Z}^2$ is $\det [u \quad v]$.

The writeup proves that any order 6 element of $SL_2(\mathbb{Z})$ is conjugate to $\left[\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right]^{\pm 1}$ by considering $\mathbb{Z}^2$ as a $\mathbb{Z}[\zeta_6]$ module, and applying the modules over a PID theorem to find that $\mathbb{Z}^2$ and $\mathbb{Z}[\zeta_6]$ are isomorphic as $\mathbb{Z}[\zeta_6]$ modules. Exploiting the isomorphism, and using the relation on the generators of $\mathbb{Z}[\zeta_6]$, the writeup proves the claim. To prove similar results for order 4 and order 6 elements of $SL_2(\mathbb{Z})$ the argument is practically identical, modulo using ζ_4 and ζ_3 instead, and using facts about its minimal polynomial.

Claim 8. *If γ has order 4 then γ is conjugate to $\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]^{\pm 1}$ in $SL_2(\mathbb{Z})$. If γ has order 3 then γ is conjugate to $\left[\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right]^{\pm 1}$.*

Proof. Let γ have order 4. Let ζ_4 be a primitive 4^{th} root of unity, so that $\zeta_4^2 + 1 = 0$ (the other 4^{th} roots are ± 1). As in the writeup, we can view $\mathbb{Z}^2$ as a $\mathbb{Z}[\zeta_4]$ module by letting $(a + b\zeta_4) \cdot v = (a \operatorname{id} + b\gamma_4) \cdot v$ for all $v \in \mathbb{Z}^2$. With the norm $a + b\zeta_4 \mapsto a^2 + b^2$, one can verify $\mathbb{Z}[\zeta_4]$ is Euclidean, thus principle. By the structure theorem for modules over PID's, $\mathbb{Z}^2$ is a sum of quotients of $\mathbb{Z}[\zeta_4]$, but all nonzero ideals of $\mathbb{Z}[\zeta_4]$ have rank 2 as abelian groups, forcing torsion on each nonfree summand, but $\mathbb{Z}^2$ is torsion free, so there can be only one summand, and it must be torsion free. Thus $\mathbb{Z}[\zeta_4]$ and $\mathbb{Z}^2$ are isomorphic as $\mathbb{Z}[\zeta_4]$ modules. Let $\phi : \mathbb{Z}[\zeta_4] \to \mathbb{Z}^2$ be one such isomorphism. Set $u = \phi(1)$ and $v = \phi(\zeta_4)$. Because $\mathbb{Z}^2 = \mathbb{Z}u + \mathbb{Z}v$, the map generated by $u \mapsto (1,0)$ and $v \mapsto (0,1)$ is a $\mathbb{Z}$ module isomorphism, thus invertible, thus has unit determinant. In coordinates $\det [u \quad v] = \pm 1$. Further, $\gamma \cdot u = \gamma \cdot \phi(1) = \phi(\zeta_4 \cdot) = \phi(\zeta_4) = v$, whereas $\gamma \cdot v = \gamma \cdot \phi(\zeta_4) = \phi(\zeta_4^2) = \phi(-1) = -\phi(1) = -u$. Thus $\gamma [u \quad v] = [v \quad -u] = [u \quad v] \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ and $\gamma [v \quad u] = [-u \quad v] = [v \quad u] \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] = [v \quad u] \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]^{-1}$. One of $[u \quad v]$ or $[v \quad u]$ is in $SL_2(\mathbb{Z})$, and correspondingly, γ is conjugate to one of $\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]^{\pm 1}$.

The proof is almost verbatim in the case that γ has order 3, but in that case the minimal polynomial for γ is $x^2 + x + 1$. Going through the same song and dance, defining an isomorphism $\phi : \mathbb{Z}[\zeta_3] \to \mathbb{Z}^2$, with $u = \phi(1)$ and $v = \phi(\zeta_3)$, observe $\gamma u = \phi(\zeta_3) = v$ and $\gamma v = \phi(\zeta_3^2) = \phi(-\zeta_3 - 1) = -v - u$. Thus, we have $\gamma [u \quad v] = [-v \quad -u - v]$. The latter is the matrix $[u \quad v]$, after adding the first column to the second, switching columns, and then negating. This is done by the matrix $\left[\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right]$, and the conjugacy result follows. $\square$

To complete the classification of elliptic points, one first notes that ζ_3 is fixed by $\pm \left[\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right]^{\pm 1}$ because $\bar{\zeta_3} = \zeta_3^{-1}$, while i is fixed by $\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]^{\pm 1}$ because $-1/i = i$. Furthermore for any $g \in SL_2(\mathbb{Z})$ the translate $g\zeta_3$ is fixed by the conjugate $g \left[\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right] g^{-1}$. Similarly, gi is fixed by $g \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right] g^{-1}$. Consequently, the only elliptic points for all of $SL_2(\mathbb{Z})$ are the translates $SL_2(\mathbb{Z})i$ and $SL_2(\mathbb{Z})\zeta_3$. Because i and ζ_3 cannot be taken to one another by the modular group, the first modular curve $Y(1)$ has exactly two elliptic points, $SL_2(\mathbb{Z})i$ and $SL_2(\mathbb{Z})\zeta_3$. The isotropy subgroup $SL_2(\mathbb{Z})_i$ can be computed directly, suppose $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] i = i$ then $ai + b = -c + id$ so that $a = d$ and $b = -c$, the determinant condition demands $ad - bc = a^2 + b^2 = 1$. If $a \neq 0$ then $a = \pm 1$ and $d = \pm 1$ and the resulting matrix is the identity or the negative identity. Otherwise, $a = 0$ and the solutions are $b = -1, c = 1$ or $b = 1, c = -1$. Similarly, if $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \zeta_3 = \zeta_3$ then $a\zeta_3 + b = c\zeta_3^2 + d\zeta_3 = (d - c)\zeta_3 + c$ so $a = d - c$ and $c = b$ along with the determinant condition $ad - bc = 1$ meaning $d^2 - cd - c^2 = 1$. If $c = 0 = b$ then $d = \pm 1$ and $a = \pm 1$ and the resulting matrix is $\pm \mathrm{id}$. If $c \neq 0$ then $c = \pm 1 = -b$ and $d = \pm 1$, forcing $a = 0$. Both possibilities are covered by the subgroup $\langle \left[\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right] \rangle$, so $SL_2(\mathbb{Z})_{\zeta_3} = \langle \left[\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right] \rangle$.

To see that the isotropy subgroup of a point in $\mathfrak{H}$ is finite cyclic, first note that $\mathfrak{H} \approx SL_2(\mathbb{R})/SL_2(\mathbb{R})_i \approx SL_2(\mathbb{R})/SO_2(\mathbb{R})$ as $SL_2(\mathbb{R})$ spaces under the isomorphism $gSO_2(\mathbb{R}) \mapsto g(i)$. Then compute that for any $\tau = x + iy \in \mathfrak{H}$, the composite $\left[\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right] \begin{bmatrix} \sqrt{y} & 0 \\ 0 & 1/\sqrt{y} \end{bmatrix}$ takes i to τ. Define $s_\tau = \left(\left[\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right] \begin{bmatrix} \sqrt{y} & 0 \\ 0 & 1/\sqrt{y} \end{bmatrix} \right)^{-1}$ so that $s_\tau \tau = i$. We can now write the inverse of the isomorphism $gSO_2(\mathbb{R}) \mapsto g(i)$ as $\tau \mapsto s_\tau SO_2(\mathbb{R})$. Observe that for all $g \in SL_2(\mathbb{R})_\tau \approx SL_2(\mathbb{R})_{s_\tau SO_2(\mathbb{R})}$, we have $gs_\tau SO_2(\mathbb{R}) = s_\tau SO_2(\mathbb{R})$, meaning $s_\tau^{-1} g s_\tau SO_2(\mathbb{R}) = SO_2(\mathbb{R})$, so that $s_\tau^{-1} g s_\tau \in SO_2(\mathbb{R})$. Thus $s_\tau^{-1} SL_2(\mathbb{R})_\tau s_\tau$ is a subgroup of $SO_2(\mathbb{R})$. Further, every element of $SL_2(\mathbb{R})_\tau$ has finite order, and the only possible order for its elements are $2, 3, 4$, and 6, so that $SL_2(\mathbb{R})_\tau$ is discrete in $SO_2(\mathbb{R})$. A discrete subspace of a compact space is finite, so $SL_2(\mathbb{R})_\tau$ is finite. Finally, recall that finite subgroups of $SO_2(\mathbb{R})$ are cyclic, because $SO_2(\mathbb{R})$ are rotations of the circle.

4. Hom group formation as right adjoint to forgetful functor

In class, we did not deeply investigate the fully general notion of functor adjunction, [2] but I found that seeing the general definition helped clarify the situation. Recall from linear algebra that the adjoint of a linear map T (on a finite dimensional vector space V), with respect to the

[2] I understand the intention in not bogging down the class with distracting abstraction when there are rich phenomena to be investigated

hermitian pairing $\langle \cdot, \cdot \rangle$ is the unique linear map T^* such that for all $v, w \in V$ the *adjunction relation* $\langle Tv, w \rangle = \langle v, T^* w \rangle$ holds. Because V is finite dimensional, V is (very non)canonically isomorphic to its dual V^* by the assignment $v \mapsto \langle \cdot, v \rangle$ (such an isomorphism requires a choice of basis and inner product). The adjunction relationship between T and T^* says that the pairing between a T transformed vector and an untransformed dual vector is the same as the pairing of that vector untransformed and that dual vector T^* transformed.

Similarly, an adjunction relation between functors asserts that the collection of mappings from a transformed object in one category to an untransformed object in another category is naturally isomorphic to the collection of mappings from the untransformed first object to the adjointly transformed second object.

To be precise, let $\mathcal{C}$ and $\mathcal{D}$ be categories, and $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ and $\mathcal{G} : \mathcal{D} \to \mathcal{C}$ be functors. We say that $\mathcal{F}$ is adjoint to $\mathcal{G}$ if both of the following conditions hold:

(1) For every object X of $\mathcal{C}$ and every object Y of $\mathcal{D}$, we have an isomorphism $i_{X,Y} : \hom_{\mathcal{D}}(\mathcal{F}X, Y) \to \hom_{\mathcal{C}}(X, \mathcal{G}Y)$
(2) The above collection of isomorphisms is natural in X and Y.

By a collection of isomorphisms being natural in X, we mean that for any two objects X, X' of $\mathcal{C}$, any object Y of $\mathcal{D}$, any map $f : X' \to X$, and any $g \in \hom_{\mathcal{D}}(\mathcal{F}X, Y)$ the following equality of functions holds

$$i_{X',Y}(\mathcal{G}(g) \circ f) = (i_{X,Y}(g)) \circ f$$

Similarly, when we say a collection of isomorphism is natural in Y, we mean that for any two objects Y, Y' of $\mathcal{D}$, any object X of $\mathcal{C}$, any map $f : Y \to Y'$, and any $g \in \hom_{\mathcal{D}}(\mathcal{F}X, Y)$ we have

$$i_{X,Y'}(f \circ g) = \mathcal{F}(f) \circ (i_{X,Y}(g)).$$

Literally speaking, what I have just described is $\mathcal{G}$ being the *left* adjoint to $\mathcal{F}$. The *right* adjoint to a functor $\mathcal{F}$ is a functor (G) such that there are a collection of isomorphisms $i_{X,Y} : \hom_{\mathcal{D}}(Y, \mathcal{F}X) \to \hom_{\mathcal{C}}(\mathcal{G}Y, X)$ is natural in X and Y.

In class, we demonstrated that tensor product formation is the left adjoint to the forgetful functor, in the sense that for rings A, B with identities, and a map $A \to B$ such that $1_A \mapsto 1_B$, the

functor

$$\begin{aligned} B \otimes_A - \quad &: \{A \text{ modules}\} \to \{B \text{ modules}\} & M &\mapsto B \otimes_A M \\ \mathrm{id} \otimes_A - \quad &: \{A \text{ module maps}\} \to \{B \text{ module maps}\} & f &\mapsto \mathrm{id} \otimes_A f \end{aligned}$$

is left adjoint to the restriction functor of A modules (and maps) to A modules (and maps).

On the other hand (haha)

Claim 9. *A-module Hom group formation*

$$\begin{aligned} \hom_A(B, -) \quad &: \{A \textit{ modules}\} \to \{B \textit{ modules}\} & N &\mapsto \hom_A(B, N) \\ g \circ - \quad &: \{A \textit{ module maps}\} \to \{B \textit{ module maps}\} & f &\mapsto g \circ f \end{aligned}$$

is right adjoint to (B to A module) restriction.

The density of the symbol patterns can make one's head spin. The idea is that we have a functor $\mathcal{F}$ that takes an A-module N, and returns the B module $F(N) = \hom_A(B, N)$, and takes an A module map $f : N \to N'$ and returns a B module map $\mathcal{F}(f) : \hom_A(B, N) \to \hom_A(B, N')$ which takes $(B \xrightarrow{A} N) \xmapsto{B} (B \xrightarrow{A} N')$. That is, if f takes A module vectors to A module vectors, then $\mathcal{F}(f)$ takes A module maps out of B to A module maps out of B. We see the A module maps out of B as B module *vectors*. See the template that I turned in for proof.

5. Parroting of intrinsic proof of Cayley-Hamilton theorem

Before we do anything, note that for A, a commutative ring with 1, and M, N two A algebras, we can make the tensor product $M \otimes_A N$ into a commutative algebra by defining multiplication componentwise on monomials: for all $m, m' \in M$ and $n, n' \in N$ say

$$(m \otimes_A n) \cdot (m' \otimes_A n') = (mm') \otimes_A (nn').$$

There is one thing that we *must* check in order for $M \otimes_A N$ to be a commutative A algebra, and another thing that we must check in order for what we've done to be even remotely sensible: check that this algebra satisfies the universal mapping property for tensor products of algebras. To check the former, one first needs to check that the properties of a commutative algebra are compatible with the structure we've put on $M \otimes_A N$, first on monomials. Then one must show that distributivity and commutativity allow us to lift the algebra structure from the monomials to

the whole module, sensibly. To check the universal mapping property, one needs to go through a similar sort of ritual as what we did for module tensor products. Both are fairly banal.

To prove the Cayley-Hamilton theorem, we must first develop a suitable environment. Let k be a field (of characteristic 2), and let V be a finite dimensional k vector space. Let T be an endomorphism of V, so that the k algebra $k[T]$ acts on V in the only reasonable way. Now, enlarge the ring of scalars of the k-algebra $k[T]$ to include polynomials, $k[x] \otimes_k k[T]$. As noted above, this tensor product is as commutative algebras, which we might as well think of as a ring. Denote it R. Further, we can let the ring $k[x] \otimes_k k[T]$ act on the (monomials in the) vector space (with enlarged scalars) $k[x] \otimes_k V$ by $(f(x) \otimes g(T)) \cdot (q(x) \otimes v) = (f(x)q(x) \otimes g(T)v)$ (and extending by linearity from the monomials). Call the new k vector space M but note that we are actually viewing M as an R-module. Be aware that in the first component, the action is by multiplication of polynomials, wheras in the second component the action is by application of linear transformation. For any $r \in R$ there is an adjugate $r^{adj} \in \mathrm{End}_{k[x]}(M)$, but there is no reason to believe that $r^{adj} \in R$ or that r^{adj} commutes with all of R.

Introduce the main actor $y = x \otimes \mathrm{id} - 1 \otimes T \in R$. We care about y because the characteristic polynomial of T is given by $f_T(x) = \det(x \otimes \mathrm{id} - 1 \otimes T) = \det y$. That is, $y^{adj} * y$ acts on M as multiplication by $f_T(x)$. The adjugate of y has properties not shared by the adjugate of general elements of R, it commutes with all of R. To see this, note that k is a domain so that $y^{adj}y = yy^{adj}$. This means $y^{adj}(x \otimes \mathrm{id} - 1 \otimes T) = y^{adj}(x \otimes \mathrm{id}) - y^{adj}(1 \otimes T)$, and because multiplication in $k[x]$ is commutative and id commutes with everything $y^{adj}(x \otimes \mathrm{id}) - y^{adj}(1 \otimes T) = (x \otimes \mathrm{id})y^{adj} - y^{adj}(1 \otimes T)$. Thus the commutativity of the whole expression forces $y^{adj}(1 \otimes T) = (1 \otimes T)y^{adj}$. Notice that a generic scalar in R can be written

$$\begin{aligned} f(x) \otimes g(T) &= (\sum_{i=0}^{p} a_i x^i) \otimes (\sum_{j=0}^{q} b_j T^j) \\ &= \sum_{i=0}^{p} a_i \sum_{j=0}^{q} b_j (x^i \otimes T^j) \\ &= \sum_{i=0}^{p} a_i \sum_{j=0}^{q} b_j (x^i \otimes \mathrm{id})(1 \otimes T^j) \\ &= \sum_{i=0}^{p} a_i \sum_{j=0}^{q} b_j (x \otimes \mathrm{id})^i (1 \otimes T)^j \end{aligned}$$

so the commutativity of y^{adj} with $1 \otimes T$ means y^{adj} commutes with all of R. Consider the ideal $I = yR$ and the action of the resulting quotient ring R/I on the quotient module M/IM, written elementwise as $(f(x) \otimes g(T) + I) \cdot (q(x) \otimes v + IM) = (f(x)g(x) \otimes g(T)v) + IM$. Again, because for a general $r \in R$, the adjugate r^{adj} need not be in R, so there is no guarantee that the action of r^{adj} sensibly acts on the quotient M/IM. Happily, the commutativitiy of y^{adj} with R allows descension to the quotient[3]

$$y^{adj}IM = y^{adj}yRM = yy^{adj}RM = yRy^{adj}M \subset yRM = IM.$$

In the quotient R/I, a polynomial in x in the first component along with the identity in the second component is the same as 1 in the first component along with that polynomial of T in the second. To see this, compute

$$x \otimes \mathrm{id} - 1 \otimes T + I = 0 + I$$

so

$$x \otimes \mathrm{id} + I = 1 \otimes T + I,$$

and in particular

$$(x \otimes \mathrm{id})^j + I = (1 \otimes T)^j + I,$$

which, by virture of componentwise multiplication in R,

$$x^j \otimes \mathrm{id} + I = 1 \otimes T^j + I.$$

Summing over j with suitible coefficients and using bilinearity of the tensor product shows that for any polynomial $f(x) \in k[x]$,

$$f(x) \otimes \mathrm{id} + I = 1 \otimes f(T) + I.$$

To reiterate, we have just shown that while working in R/I we are free to substitute T for x in polynomials. Similarly, in the quotient M/IM a polynomial in the first component and a vector in the second is the same as 1 in the first component along with the same polynomial evaluated at T

[3]This is something we talked about on one of the first days of class. In order for a map on the whole object to be sensible on a quotient of that object (meaning: is constant on equivalence classes) it suffices for the map to take the kernel *into* itself. The map need hit the whole kernel, which is why we're fine with the containment $y^{adj}IM \subset IM$, as opposed to equality.

applied to that vector in the second. Indeed, compute for any $f(x) \otimes v \in M$, by definition of the action of R on M

$$f(x) \otimes v + IM = (f(x) \otimes \mathrm{id})(1 \otimes v) + IM,$$

and by the last computation

$$(f(x) \otimes \mathrm{id})(1 \otimes v) + IM = (1 \otimes f(T))(1 \otimes v) + IM,$$

which, again by the definition of the action, is to say

$$f(x) \otimes v + IM = 1 \otimes f(T)v + IM.$$

Further, recall $IM = yRM = (x \otimes \mathrm{id} - 1 \otimes T)(k[x] \otimes V) = xk[x] \otimes V - k[x] \otimes TV$. Thus, the vector $1 \otimes Tv$ is zero modulo IM. Thus, only the monomial coming from the constant term in the polynomial f will be nonzero in M/IM. That is to say that every $f(x) \otimes v = 1 \otimes$constant term of $f \cdot$, making M/IM isomorphic to V. To reiterate, working in the quotient M/IM is the same as working in the original vector space V. Furthermore, to manipulate a vector in M/IM to a form that is easily recognizably sitting in a copy of V, we take the polynomial in the first component, evaluate it at T, and apply that linear map to the vector in the second component and the resulting vector is in V (note that this is not the best reduced form for that vector, but it is in V nonetheless).

We are now set up to prove the Cayley-Hamilton theorem. Recall that k is any field, V is any finite dimensional vector space over k, T is any endomorphism of V, and $f_T(x) = \det(x \otimes \mathrm{id} - 1 \otimes T)$ is the characteristic polynomial of f.

Theorem 1. *T satisfies its characteristic polynomial. That is, $f_T(T) = \det(T \otimes \mathrm{id} - 1 \otimes T) = 0$.*

Proof. By construction, $y^{adj}y$ annihilates $M/IM (= M/yRM)$. In the discussion above we found that $y^{adj}y$ acts on M as multiplication by $f_T(x) \otimes \mathrm{id}$, which is to say $f_t(x) \otimes \mathrm{id} + I$ annihilates M/IM. Furthermore, in R/I a polynomial in the first component and id in the second is the same as 1 in the first component and that polynomial evaluated at T in the second, that is ${}_t(x) \otimes \mathrm{id} + I = 1 \otimes f_T(T) + I$ so $1 \otimes f_T(T) + I$ annihilates M/IM. Finally, M/IM is isomorphic to V, so that the operator $f_T(T)$ annihilates V, proving the theorem. □