Given a P.I.D. D, let $M(r; a_1, \ldots, a_m) \coloneqq D^r \oplus \bigoplus_{i=1}^m \frac{D}{\langle a_i \rangle}$ where $r \in \mathbb{Z}_{\geq 0}$ and $a_i \in D - (D^\times \cup \{0\})$ satisfies that $a_1 \mid a_2 \mid \cdots \mid a_m$.

If M is an A-module, then the annihilator of M is:

$$\mathrm{Ann}_A(M) \coloneqq \{a \in A : am = 0 \text{ for all } m \in M\}.$$

Note that $\mathrm{Ann}_A(M)$ is always an ideal of A. After all:

$$\mathrm{Ann}_A(M) = \bigcap_{m \in M} \ker(a \mapsto am).$$

Similarly, if $m \in M$ then we define $\mathrm{Ann}_A(m) \coloneqq \{a \in A : am = 0\}$. This is also an ideal of A.

Lemma: $\mathrm{rank}(M(r; a_1, \ldots, a_m)) = r$.

Since there is an injection $D^r \hookrightarrow M(r; a_1, \ldots, a_m) =: M$, we know that:

$$r = \mathrm{rank}(D^r) \leq \mathrm{rank}(M).$$

Meanwhile, suppose $v_1, \ldots, v_{r+1} \in M$. Then note that $a_m \in \mathrm{Ann}_D(\bigoplus_{i=1}^m \frac{D}{\langle a_i \rangle}) - \{0\}$. Thus we can get that $a_m v_1, \ldots, a_m v_{r+1}$ are in $D^r \oplus \{0\}$. And since $\mathrm{rank}(D^r) = r$, there exists $c_1, \ldots, c_{r+1}$ not all zero such that $c_1 a_m v_1 + \ldots + c_{r+1} a_m v_r = 0$. Thus, the v_i are D-linearly dependent. And this proves that $\mathrm{rank}(M) < r + 1$. ■

Suppose D is an integral domain and M is a D-module. Then the torsion of M is defined to be:

$$\mathrm{Tor}(M) \coloneqq \{m \in M : \exists a \in D - \{0\} \ s.t. \ am = 0\}$$

In other words, $m \in \mathrm{Tor}(M)$ if $\mathrm{Ann}_D(m) \neq \{0\}$.

Lemma: $\mathrm{Tor}(M)$ is a submodule of M. (note this requires D to be a domain...)

Proof:

Suppose $m_1, m_2 \in \mathrm{Tor}(M)$ and $a_1, a_2 \in D$. We know there exists $c_1, c_2 \in D - \{0\}$ such that $c_1 m_1 = c_2 m_2 = 0$. Also as D is a domain, we know that $c_1 c_2 \neq 0$. That said:

$$c_1 c_2 (a_1 m_1 + a_2 m_2) = c_2 a_1 (c_1 m_1) + c_1 a_2 (c_2 m_2) = 0$$

Therefore $a_1 m_1 + a_2 m_2 \in \mathrm{Tor}(M)$. ■

Remark: If $\theta : M_1 \to M_2$ is a D-module isomorphism, then $\theta(\mathrm{Tor}(M_1)) = \mathrm{Tor}(M_2)$.

Why:

Suppose $\exists c \in D - \{0\}$ such that $cm = 0$. Then $c\theta(m) = \theta(cm) = \theta(0) = 0$. Hence, $\theta(m) \subseteq \mathrm{Tor}(M_2)$

This proves that $\theta(\mathrm{Tor}(M_1)) \subseteq \mathrm{Tor}(M_2)$. By symmetric reasoning, we can see that $\theta^{-1}(\mathrm{Tor}(M_2)) \subseteq \mathrm{Tor}(M_1)$. By applying θ to both sides, we get $\mathrm{Tor}(M_2) \subseteq \theta(\mathrm{Tor}(M_1))$. So, we can conclude that $\theta(\mathrm{Tor}(M_1)) = \mathrm{Tor}(M_2)$.

Consequently, if $\theta : M_1 \to M_2$ is a D-module isomorphism then $\mathrm{Tor}(M_1) \cong \mathrm{Tor}(M_2)$.

Lemma: $\mathrm{Tor}(M(r;a_1,\ldots,a_m)) = \frac{D}{\langle a_1\rangle}\oplus\cdots\oplus\frac{D}{\langle a_m\rangle}$ (where the latter is an internal direct sum).

Note that we also write $\frac{D}{\langle a_1\rangle}\oplus\cdots\oplus\frac{D}{\langle a_m\rangle} = M(0,a_1,\ldots,a_m)$.

Proof:
Note that $a_1\cdots a_m\cdot x = 0$ for any $x\in M(0,a_1,\ldots,a_m)$. Hence, the $\supseteq$ inclusion is clear.

On the other hand, suppose that $(v,x)\in\mathrm{Tor}(M)$ where $v\in D^r$ and $x\in M(0,a_1,\ldots,a_m)$. Then there exists $c\in D-\{0\}$ such that $(cv,cx)=0$. But $cv=0\Longrightarrow v=0$ as $c\neq 0$. So, $(v,x)\in M(0,a_1,\ldots,a_m)$. ■

If A is a unital commutative ring, we say $\mathfrak{a}_1,\ldots,\mathfrak{a}_m\lhd A$ are coprime if $\mathfrak{a}_i+\mathfrak{a}_j=A$ whenever $i\neq j$.

As a side note, suppose D is a P.I.D and let $a,b\in D$. Also let $d=\gcd(a,b)$ (see *page 519*). Then we claim that $\langle a\rangle+\langle b\rangle=\langle a,b\rangle$ is precisely equal to $\langle d\rangle$.

To see why, note that because $d\mid a$ and $d\mid b$, we know that $\langle a,b\rangle\subseteq\langle d\rangle$. On the other hand, as D is a P.I.D. we know that $\langle a,b\rangle=\langle c\rangle$ for some $c\in D$. But now as $c\mid a$ and $c\mid b$, we must have that $c\mid d$. Therefore, $\langle d\rangle\subseteq\langle c\rangle=\langle a,b\rangle$.

As a consequence, $\langle a\rangle+\langle b\rangle=A$ iff $\gcd(a,b)\in A^\times$. So this new definition of coprimeness purely generalizes the original definition.

Generalized Chinese Remainder Theorem: Suppose A is a unital commutative ring and $\mathfrak{a}_1,\ldots,\mathfrak{a}_m\lhd A$ are coprime. Then $\frac{A}{\mathfrak{a}_1\cap\cdots\cap\mathfrak{a}_m}\cong\frac{A}{\mathfrak{a}_1}\oplus\cdots\oplus\frac{A}{\mathfrak{a}_m}$ via the A-algebra isomorphism $x+(\mathfrak{a}_1\cap\cdots\cap\mathfrak{a}_m)\mapsto(x+\mathfrak{a}_1,\ldots,x+\mathfrak{a}_m)$

Proof:
Step 1: For all i, we have that $\mathfrak{a}_i+\bigcap_{j\neq i}\mathfrak{a}_j=A$.

Proof:
For all $j\neq i$, there exists $x_j\in\mathfrak{a}_j$ such that $x_j\equiv 1\pmod{\mathfrak{a}_i}$. Then, $y:=\prod_{j\neq i}x_j$ is in $\bigcap_{j\neq i}\mathfrak{a}_j$ and satisfies that $y\equiv 1\pmod{\mathfrak{a}_i}$. It follows that there exists $x_i\in\mathfrak{a}_i$ with $y+x=1$. And that shows that $\mathfrak{a}_i+\bigcap_{j\neq i}\mathfrak{a}_j=A$.

Step 2: $\theta:A\to\bigoplus_{i=1}^m\frac{A}{\mathfrak{a}_i}$ given by $\theta(x)=(x+\mathfrak{a}_1,\ldots,x+\mathfrak{a}_m)$ is a surjective A-algebra homomorphism (meaning it is an A-module homomorphism and a ring homomorphism).

As a side note, we can view $\bigoplus_{i=1}^m\frac{A}{\mathfrak{a}_i}$ as a ring by identifying it with the product ring $\prod_{i=1}^m\frac{A}{\mathfrak{a}_i}$. In general, an A-algebra M is an A-module such that M is also a ring and $a(m_1m_2)=(am_1)m_2=m_1(am_2)$ for all $a\in A$ and $m_1,m_2\in M$.

Proof:
The fact that θ is a homomorphism is obvious. What is less obvious is that θ is surjective. To prove this, it is necessary and sufficient to show that:

$$(0,\ldots,0,1+\mathfrak{a}_i,0,\ldots,0)\in\mathrm{im}(\theta)\text{ for all }i.$$

Fortunately by step 1, we know there exists $y_i\in\bigcap_{j\neq i}\mathfrak{a}_j$ such that $y_i\equiv 1\pmod{\mathfrak{a}_i}$. In turn $\theta(y_i)=(0,\ldots,0,1+\mathfrak{a}_i,0,\ldots,0)\in\mathrm{im}(\theta)$.

Step 3: $\ker(\theta) = \bigcap_{i=1}^{m} \mathfrak{a}_i$.
Hopefully this is obvious.

To finish, just use the first isomorphism theorem for rings as well as the first isomorphism theorem for A-modules. ■

Corollary: If M is a finitely generated D-module and D is a PID, then M is torsion-free (meaning $\mathrm{Tor}(M) = \{0\}$) if and only if M is free.

$(\Longleftarrow)$
If M is free then we know that $M \cong D^r$ for some positive integer r. But $\mathrm{Tor}(D^r) = \{0\}$ since D is a domain. Therefore, $\mathrm{Tor}(M) = \{0\}$ as well.

$(\Longrightarrow)$
Using the fundamental theorem of finitely generated modules over a PID, write $M \cong D^r \oplus M(0, a_1, \ldots, a_m)$. Then $\{0\} = \mathrm{Tor}(M) = M(0, a_1 \ldots, a_m)$ implies that $M \cong D^r$. So, M is free. ■

Other remarks:

- I used multiple times throughout the proof on *pages 588-591* that if D is an integral domain, then every ideal of D is torsion free.
- Note that if D is not a P.I.D then there exists an ideal $\mathfrak{a} \lhd D$ that is not principle. In turn, $\mathfrak{a}$ is not free.

  To see why, first note that $\mathrm{rank}(\mathfrak{a}) = 1$. After all, we clearly have that $\{a\}$ is D-linearly independent for any $a \in \mathfrak{a} - \{0\}$. Meanwhile suppose b is another element in $\mathfrak{a}$. Then, $b(a) + (-a)b = 0$ but $-a \neq 0$. So, $\{a, b\}$ is not D-linearly independent.

  As isomorphisms preserve the rank of modules and $\mathrm{rank}(D^n) = n$ for all integers n, we must have that if $\mathfrak{a}$ were free then we'd have that $\mathfrak{a} \cong D$ via a D-module homomorphism $\theta : D \to \mathfrak{a}$. But now we'd have a contradiction as $\mathfrak{a} = \langle \theta(1) \rangle$ is principal.

- By combining these remarks, we have shown that the $(\Longrightarrow)$ implication of the above corollary is false if D is merely a domain instead of a P.I.D.

Let D be a P.I.D. and let $\mathcal{P}_0 \subseteq D$ contain precisely one element from each equivalence class of companions containing irreducible elements (see *page 518* and recall that all P.I.Ds are U.F.Ds).

Note that because D is a U.F.D., $\gcd(a, b) = 1$, then $\langle ab \rangle = \langle a \rangle \cap \langle b \rangle$.
After all, the $\subseteq$ inclusion is trivial. Meanwhile, if $a \mid x$ and $b \mid x$, then because $\gcd(a, b) = 1$, the only way to not violate the unique factorization of x is if $ab \mid x$. This shows that $\langle a \rangle \cap \langle b \rangle \subseteq \langle ab \rangle$.

By induction we can conclude that $\langle a\rangle = \bigcap_{p\in\mathcal{P}_0}\langle p^{\nu_p(a)}\rangle$. And since D is a P.I.D., we know from two pages ago that all the ideals $\langle p^{\nu_p(a)}\rangle$ are coprime. Hence by the generalized Chinese remainder theorem, we have that:

$$\frac{D}{\langle a\rangle} = \frac{D}{\bigcap_{p\in\mathcal{P}_0}\langle p^{\nu_p(a)}\rangle} \cong \bigoplus_{p\in\mathcal{P}_0}\frac{D}{\langle p^{\nu_p(a)}\rangle}.$$

Now at last we are going to prove the uniqueness part of the theorem on *page 591*.

Theorem: If $M(r;a_1,\dots,a_m)\cong M(r',b_1,\dots,b_\ell)$, then $r=r'$, $m=\ell$, and $\langle a_i\rangle = \langle b_i\rangle$ for all i.

Proof:
To start out, we know that $r=\operatorname{rank}(M(r;a_1,\dots,a_m)) = \operatorname{rank}(M(r',b_1,\dots,b_\ell)) = r'$.
Also, we know that:

$$\begin{aligned} M(0;a_1,\dots,a_m) &= \operatorname{Tor}(M(r;a_1,\dots,a_m)) \\ &\cong \operatorname{Tor}(M(r',b_1,\dots,b_\ell)) = M(0;b_1,\dots,b_\ell). \end{aligned}$$

Hence, it now suffices to show that if $M(0;a_1,\dots,a_m)\cong M(0;b_1,\dots,b_\ell)$ then $m=\ell$ and $\langle a_i\rangle = \langle b_i\rangle$ for all i.

Part 1: Building a Counting Machine
Consider that:

$$M := M(0;c_1,\dots,c_m) = \bigoplus_{i=1}^m \frac{D}{\langle c_i\rangle} \cong \bigoplus_{i=1}^m\left(\bigoplus_{p\in\mathcal{P}_0}\frac{D}{\langle p^{\nu_p(c_i)}\rangle}\right) \cong \bigoplus_{p\in\mathcal{P}_0}\left(\bigoplus_{i=1}^m\frac{D}{\langle p^{\nu_p(c_i)}\rangle}\right)$$

We shall denote $M(c_1,\dots,c_m;p) := \bigoplus_{i=1}^m \frac{D}{\langle p^{\nu_p(c_i)}\rangle}$.

Importantly note that as $c_1 \mid c_2 \mid \cdots \mid c_m$, we know that $\nu_p(c_1)\le\cdots\le\nu_p(c_m)$ for all $p\in\mathcal{P}_0$. Since $M(c_1,\dots,c_m;p)$ has that additional structure, we'll try studying it.

- Note that if A is a commutative ring, then for any A-module N and $b\in A$ we have that $\ell_b(m) := bm$ is an A-module homomorphism from N to itself. It follows that $b\cdot M = \operatorname{im}(\ell_b)$ is an A-module.

- If $\{M_i\}_{i\in I}$ is a family of A-modules and $b\in A$, then $b\cdot\bigoplus_{i\in I}M_i = \bigoplus_{i\in I}(b\cdot M_i)$.

- Suppose $\theta: M_1\to M_2$ is an A-module isomorphism. Then $\theta(b\cdot M_1) = b\cdot M_2$.
  Hopefully this is easy to see.

- In particular, if D is a P.I.D. and $a\in D$, then we have that
  $$b\cdot\frac{D}{\langle a\rangle} = \frac{bD+\langle a\rangle}{\langle a\rangle} = \frac{\langle b,a\rangle}{\langle a\rangle} = \frac{\langle\gcd(a,b)\rangle}{\langle a\rangle}$$

  Consequently, if p_0, p are prime elements and r,k are nonnegative integers, then:

  $$p_0^r\cdot\frac{D}{\langle p^k\rangle} = \begin{cases} \frac{D}{\langle p^k\rangle} & \text{if } p_0\neq p \\ \frac{\langle p^{\min(r,k)}\rangle}{\langle p^k\rangle} & \text{if } p_0 = p \end{cases}$$

- $\frac{\langle \gcd(a,b)\rangle}{\langle a\rangle} \cong \frac{D}{\langle \frac{a}{\gcd(a,b)}\rangle}$ as D-modules.

  Proof:
  Let $d = \gcd(a,b)$ and define $\widehat{\theta} : D \to \frac{\langle d\rangle}{\langle a\rangle}$ by $\widehat{\theta}(x) := dx + \langle a\rangle$. Then $\widehat{\theta}$ is a surjective D-module homomorphism.

  Also, $x \in \ker(\widehat{\theta})$ iff $dx \in \langle a\rangle$, and that happens iff $dx = ay$ for some $y \in D$. But now as $d \mid a$ and D is an integral domain, we can say that $x = \frac{a}{d}y$ for some $y \in D$.

  Hence, $\ker(\widehat{\theta}) = \langle \frac{a}{d}\rangle$ and we finish by invoking the first isomorphism theorem. ■

Now for all integers $r \geq 0$ and $p_0 \in \mathcal{P}_0$, we have that:

$$\begin{aligned} p_0^r \cdot M &\cong \bigoplus_{p\in\mathcal{P}_0} p_0^r \cdot M(c_1,\ldots,c_m;p) \\ &= \bigoplus_{p\in\mathcal{P}_0} \left(\bigoplus_{i=1}^m p_0^r \cdot \frac{D}{\langle p^{\nu_p(c_i)}\rangle}\right) \\ &\cong \left(\bigoplus_{p\in\mathcal{P}_0-\{p_0\}} \left(\bigoplus_{i=1}^m \frac{D}{\langle p^{\nu_p(c_i)}\rangle}\right)\right) \oplus \bigoplus_{i=1}^m \frac{\langle {p_0}^{\min(r,\nu_{p_0}(c_i))}\rangle}{\langle {p_0}^{\nu_{p_0}(c_i)}\rangle} \end{aligned}$$

In order to cancel out the first ugly direct sum, we do an algebraic trick (which is referred to as taking a graded filtration). Consider the submodules:

$$M \supseteq p_0 \cdot M \supseteq p_0^2 \cdot M \supseteq \cdots$$

Now by recalling the lemma on *page 591*, we get that:

$$\frac{p_0^r\cdot M}{p_0^{r+1}\cdot M} \cong \frac{\bigoplus_{i=1}^m \frac{\langle {p_0}^{\min(r,\nu_{p_0}(c_i))}\rangle}{\langle {p_0}^{\nu_{p_0}(c_i)}\rangle}}{\bigoplus_{i=1}^m \frac{\langle {p_0}^{\min(r+1,\nu_{p_0}(c_i))}\rangle}{\langle {p_0}^{\nu_{p_0}(c_i)}\rangle}} \cong \bigoplus_{i=1}^m \frac{\left(\frac{\langle {p_0}^{\min(r,\nu_{p_0}(c_i))}\rangle}{\langle {p_0}^{\nu_{p_0}(c_i)}\rangle}\right)}{\left(\frac{\langle {p_0}^{\min(r+1,\nu_{p_0}(c_i))}\rangle}{\langle {p_0}^{\nu_{p_0}(c_i)}\rangle}\right)}.$$

Lemma: Suppose $M \supseteq N \supseteq L$ are D-modules. Then $(M/L)/(N/L) \cong (M/N)$

Proof:
We already know from the third isomorphism theorem for groups that $(x+L)+N/L \mapsto x+N$ is a group isomorphism. Then it's trivial to further check that this group isomorphism is also A-module homomorphism.

So, $\frac{p_0^r\cdot M}{p_0^{r+1}\cdot M} \cong \bigoplus_{i=1}^m \frac{\langle {p_0}^{\min(r,\nu_{p_0}(c_i))}\rangle}{\langle {p_0}^{\min(r+1,\nu_{p_0}(c_i))}\rangle}$.

But now consider that $\frac{\langle {p_0}^{\min(r,\nu_{p_0}(c_i))}\rangle}{\langle {p_0}^{\min(r+1,\nu_{p_0}(c_i))}\rangle} \cong \begin{cases} \{0\} & \text{if } \nu_{p_0}(c_i) \leq r \\ \frac{D}{\langle p_0\rangle} & \text{if } \nu_{p_0}(c_i) > r \end{cases}$

Therefore, we have that $\frac{p_0^r\cdot M}{p_0^{r+1}\cdot M} \cong \left(\frac{D}{\langle p_0\rangle}\right)^{\#\{i\,:\,\nu_{p_0}(c_i)>r\}}$ (where $\#S$ denotes the cardinality of the set S).

Lemma: If M is an A-module, $\mathfrak{b} \lhd A$, and $\theta : M \to (A/\mathfrak{b})^k$ is an A-module isomorphism, then we can view M as an $(A/\mathfrak{b})$-module by defining $(a+\mathfrak{b})m = am$. Furthermore, θ is then an $(A/\mathfrak{b})$-module isomorphism when M is equipped with this multiplication operation.

Proof:
Suppose $a_1 \equiv a_2 \pmod{\mathfrak{b}}$. Then $a_1 - a_2 \equiv 0 \pmod{\mathfrak{b}}$. In turn, for any $m \in M$ we must have that

$$\theta((a_1 - a_2)m) = (a_1 - a_2)\theta(m) = 0$$

And because θ is an isomorphism, that implies that $(a_1 - a_2)m = 0$. So, $a_1 m = a_2 m$ and we've proven that our scalar multiplication operation is well-defined. All the needed properties of this scalar multiplication are now easily seen as being inherited from the old scalar multiplication.

At last we get to the key observation. Note that as D is a P.I.D. and p_0 is irreducible, we know that $\frac{D}{\langle p_0 \rangle}$ is a field. Hence, our prior lemma says that $\frac{p_0^r \cdot M}{p_0^{r+1} \cdot M}$ is a $\frac{D}{\langle p_0 \rangle}$-vector space whose dimension is precisely the number of c_i for which $p_0^r \mid c_i$.

Part 2: Counting the irreducible factors of $a_1, \ldots, a_m$ and $b_1, \ldots, b_\ell$
Let $N_1 := M(0; a_1, \ldots, a_m)$ and $N_2 := M(0; b_1, \ldots, b_\ell)$, and suppose $N_1 \cong N_2$.

Remark: If M_1, M_2 are A-modules, $N \subseteq M_1$ is a submodule, and $\theta : M_1 \to M_2$ is an A-module isomorphism, then $\frac{M_1}{N} \cong \frac{M_2}{\theta(N)}$.

Proof:
Let $\pi : M_2 \to M_2/\theta(N)$ be the A-module homomorphism $m \mapsto m + \theta(N)$. Then, consider the map $\phi := \pi \circ \theta$. Clearly ϕ is surjective since both θ and π are. Also, $x \in \ker(\phi)$ iff $\theta(x) \in \theta(N)$. And since θ is injective, that happens iff $x \in N$. To finish off, we invoke the first isomorphism theorem.

As a result of the above remark plus an earlier remark that if $\theta : M_1 \to M_2$ is an A-module isomorphism then $\theta(b \cdot M_1) = b \cdot M_2$, we can now conclude for all for all $p \in \mathcal{P}_0$ and $r \in \mathbb{Z}_{\geq 0}$, we have that:

$$\left(\frac{D}{\langle p \rangle}\right)^{\#\{i \,:\, \nu_p(a_i) > r\}} \cong \frac{p^r \cdot N_1}{p^{r+1} \cdot N_1} \cong \frac{p^r \cdot N_2}{p^{r+1} \cdot N_2} \cong \left(\frac{D}{\langle p \rangle}\right)^{\#\{i \,:\, \nu_p(b_i) > r\}}$$

Since vector space isomorphisms preserve dimension, we can conclude that:
$\#\{i : \nu_p(a_i) > r\} = \#\{i : \nu_p(b_i) > r\}$ for all $p \in \mathcal{P}_0$ and $r \in \mathbb{Z}_{\geq 0}$

This let's us show that $m = \ell$ and that $\langle a_i \rangle = \langle b_i \rangle$ for all i.
To start off, let $M = \max\{\#\{i : \nu_p(a_i) > 0\} : p \in \mathcal{P}_0\}$.

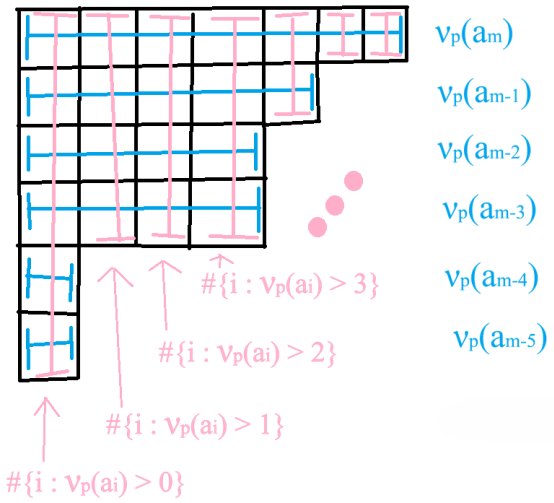
It's clear that $M \leq m$. Meanwhile, because $a_1 \notin A^\times$, there must exists some irreducible element p_0 such that $p_0 \mid a_1$. In turn, as $a_1 \mid a_2 \mid \cdots \mid a_m$, we have that $\#\{i : \nu_{p_0}(a_i) > 0\} = m$. Hence, we've proven that $M = m$.

By similar reasoning with the quantity $\max\{\#\{i : \nu_p(b_i) > 0\} : p \in \mathcal{P}_0\}$, we can thus conclude that:

$$m = \max\{\#\{i : \nu_p(a_i) > 0\} : p \in \mathcal{P}_0\} = \max\{\#\{i : \nu_p(b_i) > 0\} : p \in \mathcal{P}_0\} = \ell$$

Next, for any fixed $p \in \mathcal{P}_0$ consider the Young diagram such that the number of boxes in the $(r+1)$th column is equal to $\#\{i : \nu_p(a_i) > r\} = \#\{i : \nu_p(b_i) > r\}$.

Then the number of boxes in the kth row is precisely equal to $\nu_p(a_{-k+m+1})$ and $\nu_p(b_{-k+m+1})$. Below I've attached a sample diagram to show what I mean.



This proves that $\nu_p(a_i) = \nu_p(b_i)$ for all $p \in \mathcal{P}_0$ and $i \in \{1, \ldots, m\}$. The only way this is possible is if a_i, b_i are companions for each i. ■

Suppose M is an $F[x]$-module where F is a field. Then $\ell_x : M \to M$ defined by $\ell_x(m) = xm$ is an F-linear map. Also if $f(x) = c_n x^n + \ldots + c_1 x + c_0 \in F[x]$, then $f(x) \cdot m = \sum_{i=0}^{n} c_i \ell_x^i(m)$. Hence, the F-linear map ℓ_x uniquely determines the $F[x]$-module structure of M.

In a similar vein, if M_1, M_2 are $F[x]$-modules then $\theta : M_1 \to M_2$ is an $F[x]$-module homomorphism if and only if θ is an F-linear map, and $\theta(x \cdot m_1) = x \cdot \theta(m_1)$. To put another θ is an $F[x]$-module homomorphism iff if the following commutative diagram of F-module homomorphisms holds:

$$\begin{array}{ccc} M_1 & \xrightarrow{\theta} & M_2 \\ {\scriptstyle \ell_x}\downarrow & & \downarrow{\scriptstyle \ell_x} \\ M_1 & \xrightarrow{\theta} & M_2 \end{array}$$

$(\Longrightarrow)$
This is obvious.

$(\Longleftarrow)$
Suppose $f(x) = c_n x^n + \ldots + c_1 x + c_0$. Then:

$$\theta(f(x) \cdot m) = \theta(\sum_{i=0}^{n} c_i \ell_x^i(m)) = \sum_{i=0}^{n} c_i \theta(\ell_x^i(m)) = \sum_{i=0}^{n} c_i \ell_x^i(\theta(m)) = f(x) \cdot \theta(m).$$

We're going to use this to study linear algebra. Suppose F is a field and $a \in M_n(F)$. Then we are interested in the map $a : F^n \to F^n$ given by $|v\rangle \mapsto a\,|v\rangle$. So, we define the $F[x]$-

module V_a by taking the F-vector space F^n and defining $\ell_x(v) = a\,|v\rangle$. Hence:

$$(c_n x^n + \cdots + c_1 x + c_0) \cdot v = (c_n a^n + \cdots + c_1 a + c_0 I)|v\rangle$$

Lemma: $V_a \cong V_b$ as $F[x]$-modules if and only if a and b are similar matrices.

Proof:

$V_a \cong V_b$ as $F[x]$-modules iff there exists an $F[x]$ module isomorphism $\theta : V_a \to V_b$. But that happens iff there exists a bijective F-linear map $\theta : V_a \to V_b$ such that the below diagram commutes:

$$\begin{array}{ccccccc} F^n & = & V_a & \xrightarrow{\theta} & V_b & = & F^n \\ \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle \ell_x} & & \downarrow{\scriptstyle \ell_x} & & \downarrow{\scriptstyle b} \\ F^n & = & V_a & \xrightarrow{\theta} & V_b & = & F^n \end{array}$$

As a side note, if $a \in M_{n_1}(F), b \in M_{n_2}(F)$, and $n_1 \neq n_2$, then there doesn't exists an F-linear bijection $\theta : F^{n_1} \to F^{n_2}$. Hence, we lose nothing by assuming both V_a and V_b are equal as sets to F^n.

Equivalently, we can say there is some $g \in \mathrm{GL}_n(F)$ with $\theta(v) = g\,|v\rangle$ such that the below diagram commutes:

$$\begin{array}{ccc} F^n & \xrightarrow{g} & F^n \\ \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle b} \\ F^n & \xrightarrow{g} & F^n \end{array}$$

In other words, there exists an invertible matrix g such that $ga = bg$. And finally, we get that $b = gag^{-1}$ for some $g \in \mathrm{GL}_n(F)$. Hence, a and b are similar matrices. ■

Suppose $a \in M_{n_1}(F)$ and $b \in M_{n_2}(F)$. Then if we consider the direct sum $V_a \oplus V_b$ note that for any $(v, w) \in V_a \oplus V_b$ we have that:

$$x \cdot (v, w) = (a\,|v\rangle, b\,|v\rangle) = \begin{bmatrix} a & \mathbf{0} \\ \mathbf{0} & b \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix}$$

In particular, if we denote the block matrix $\begin{bmatrix} a & \mathbf{0} \\ \mathbf{0} & b \end{bmatrix}$ as $\mathrm{diag}(a, b)$, then we've shown that $V_a \oplus V_b \cong V_{\mathrm{diag}(a,b)}$.

Hopefully it is clear how this generalizes to larger finite direct sums, and how one would define a block matrix $\mathrm{diag}(a_1, \ldots, a_n)$.

Given a field F, we know that $F[x]$ is a P.I.D. Also if $a \in M_n(F)$ then we trivially have that V_a is finitely generated (because the standard basis for F^n will always generate all of V_a). Therefore, by the *fundamental theorem of finitely generated modules over a P.I.D*, we know that there exists a unique $r \in \mathbb{Z}_{\geq 0}$ as well as unique ideals $\langle f_1 \rangle \supseteq \langle f_2 \rangle \supseteq \cdots \supseteq \langle f_m \rangle$ in $F[x]$ such that:

$$V_a \cong (F[x])^r \oplus \tfrac{F[x]}{\langle f_1 \rangle} \oplus \cdots \oplus \tfrac{F[x]}{\langle f_m \rangle} \text{ as } F[x]\text{-modules.}$$

Equivalently, there exists unique monic polynomials $f_1 \mid f_2 \mid \cdots \mid f_m$ in $F[x]$ such that:
$$V_a \cong (F[x])^r \oplus \tfrac{F[x]}{\langle f_1 \rangle} \oplus \cdots \oplus \tfrac{F[x]}{\langle f_m \rangle} \text{ as } F[x]\text{-modules.}$$

But now we can immediately deduce that $r = 0$.

To see why compare the dimensions of both sides of the above equation as F-vector spaces. V_a would be an n-dimensional F-vector but $F[x]^r$ would be an infinite-dimensional F-vector space for any $r > 0$.

So, we can further refine our conclusion to the following:

Corollary: Suppose F is a field and $a \in M_n(F)$. There exists unique monic positive degree polynomials $f_1 \mid f_2 \mid \cdots \mid f_m$ such that $V_a \cong \tfrac{F[x]}{\langle f_1 \rangle} \oplus \cdots \oplus \tfrac{F[x]}{\langle f_m \rangle}$ as $F[x]$-modules.

We call the polynomials $f_1, \ldots, f_m$ above the invariant factors of the matrix a.

Corollary: If F is a field and $a, b \in M_n(F)$, then a is similar to b iff a and b have the same invariant factors.

Proof:
a is similar to b iff $V_a \cong V_b$ as $F[x]$-modules. But by the *fundamental theorem of finitely generated modules over a P.I.D*, the latter statement is equivalent to both a and b having the same invariant factors. ■

Question? Given any $f \in F[x]$, can we find a matrix $a \in M_n(F)$ such that $V_a \cong \tfrac{F[x]}{\langle f \rangle}$.

A reason to ask this question is that after finding matrices $a_1, \ldots, a_m$ such that $V_{a_i} \cong \tfrac{F[x]}{\langle f_i \rangle}$ for each $i \in \{1, \ldots, m\}$, we could then conclude that:
$$V_a \cong \tfrac{F[x]}{\langle f_1 \rangle} \oplus \cdots \oplus \tfrac{F[x]}{\langle f_m \rangle} \cong V_{a_1} \oplus \cdots \oplus V_{a_m} \cong V_{\text{diag}(a_1, \ldots, a_m)}$$

The answer to the above question is yes. To show this, first recall from math 100b and 100c that if $f(x) = c_n x^n + \cdots + c_1 x + c_0$ (where $c_n \neq 0$) then $\tfrac{F[x]}{\langle f \rangle}$ has the following F-basis (where $\overline{x^i}$ is just the equivalence class of $x^i \pmod{\langle f \rangle}$):
$$\mathcal{B} = (\overline{1}, \overline{x}, \ldots, \overline{x^{n-1}})$$

To see why, first note that by the long division theorem we can conclude that every equivalence class in $F[x]/\langle f \rangle$ contains a unique polynomial with degree less than n. It easily follows that $\mathcal{B}$ spans all of $F[x]/\langle f \rangle$. Also, as the unique polynomial of degree less than n equivalent to $0 \pmod{\langle f \rangle}$ is 0, we know that:
$$c_{n-1}\overline{x^{n-1}} + \cdots + c_1\overline{x^1} + c_0\overline{1} \equiv 0 \pmod{\langle f \rangle} \text{ iff all } c_i = 0.$$

Thus, we identify $\tfrac{F[x]}{\langle f \rangle} \to F^n$ as F-vector spaces via the mapping $\overline{g} \mapsto |\overline{g}\rangle_{\mathcal{B}}$ (where $|\overline{g}\rangle_{\mathcal{B}}$ is just the column vector for $\overline{g}$ with respect to the basis $\mathcal{B}$).

In other words, $a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 + \langle f \rangle \mapsto |(a_0, a_1, \ldots, a_{m-1})\rangle$.

We want to find a matrix $[\ell_x]_{\mathcal{B}} \in M_n(F)$ such that $[\ell_x]_{\mathcal{B}}|\overline{g}\rangle_{\mathcal{B}} = |\ell_x(\overline{g})\rangle_{\mathcal{B}}$. In other words, we want the following diagram to commute:

$$\begin{array}{ccc} \frac{F[x]}{\langle f\rangle} & \xrightarrow{|\cdot\rangle_{\mathcal{B}}} & F^n \\ {\scriptstyle \ell_x}\downarrow & & \downarrow{\scriptstyle [\ell_x]_{\mathcal{B}}} \\ \frac{F[x]}{\langle f\rangle} & \xrightarrow[|\cdot\rangle_{\mathcal{B}}]{} & F^n \end{array}$$

But this is easy. After all, given any linear map $T : V \to V$ and basis vectors $v_1, \ldots, v_n$, we can always write the matrix of T with respect to that basis as $\begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix}$. Hence we must have that:

$$\begin{aligned} [\ell_x]_{\mathcal{B}} &= \begin{bmatrix} |\ell_x(\overline{1})\rangle_{\mathcal{B}} & |\ell_x(\overline{x})\rangle_{\mathcal{B}} & \cdots & |\ell_x(\overline{x^{n-2}})\rangle_{\mathcal{B}} & |\ell_x(\overline{x^{n-1}})\rangle_{\mathcal{B}} \end{bmatrix} \\ &= \begin{bmatrix} |\overline{x}\rangle_{\mathcal{B}} & |\overline{x^2}\rangle_{\mathcal{B}} & \cdots & |\overline{x^{n-1}}\rangle_{\mathcal{B}} & |\overline{x^n}\rangle_{\mathcal{B}} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & \cdots & 0 & -c_0/c_n \\ 1 & 0 & \cdots & 0 & -c_1/c_n \\ 0 & 1 & \cdots & 0 & -c_2/c_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1}/c_n \end{bmatrix} \end{aligned}$$

This matrix is called the companion matrix of $f(x)$ and we denote it by c_f. (Note that if f is monic then $c_n = 1$...)

Thus we can conclude that $\frac{F[x]}{\langle f\rangle} \cong V_{c_f}$. Also, this leads to the following theorem.

Theorem (Rational Canonical Form): If $a \in M_n(F)$ and $f_1, \ldots, f_m$ are the invariant factors of a then a is similar to $\operatorname{diag}(c_{f_1}, \ldots, c_{f_m})$.

(This is the beginning of the final lecture I need to get through to fully catch up to the class. Unfortunately I didn't have time to do the third problem set.)

Firstly, we prove a uniqueness result related to the prior theorem. Note that given a matrix of the form:

$$a = \begin{bmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{bmatrix} \in M_n(F),$$

we can find a unique monic polynomial $f \in F[x]$ such that $a = c_f$.

Theorem (Uniqueness of the Rational Canonical Form): Suppose $f_1, \ldots, f_m \in F[x]$ are monic polynomials satisfing that $f_1 \mid f_2 \mid \cdots \mid f_m$ and the matrix a is similar to $\text{diag}(c_{f_1}, \ldots, c_{f_m})$. Then $f_1, \ldots, f_m$ are the invariant factors of the matrix a.

Proof:
$V_a \cong V_{\text{diag}(c_{f_1}, \ldots, c_{f_m})} \cong \bigoplus_{i=1}^m V_{c_{f_i}} \cong \bigoplus_{i=1}^m \frac{F[x]}{\langle f_i \rangle}$. Then as $f_1 \mid f_2 \mid \cdots \mid f_m$, we can conclude by the uniqueness part of the fundamental theorem of finitely generated modules over a P.I.D. that $f_1, \ldots, f_m$ are the invariant polynomials. ■

On *page 518* I wrote what it means for a unital commutative ring to be an F-algebra. However, for this next section it will be necessary to define what it means for a non-commutative unit ring to be an F-algebra. Also in general I want to prove some things I didn't think of proving back on page 518.

If F is a field and A is a noncommutative unital ring, we say that A is an F-algebra if there exists a ring homomorphism $f : F \to A$ such that $f(1_f) = f(1_A)$ and $f(c)a = af(c)$ for all $a \in A$ and $c \in F$.

- Note that if $A \neq \{0\}$ then f is necessarily injective. This is because $\ker(f)$ is an ideal of F. That said, the only ideals of F are F and $\{0\}$, and we know that $\ker(f) \neq F$. After all, $1_F \notin \ker(f)$ since $f(1_F) = 1_A \neq 0_A$. It follows that $\ker(f) = \{0\}$ and that proves that f is injective.

  The significance of the prior result is that if A is an F-algebra then we can view F as being a field embedded into A.

- If $c \in F$ and $a \in A$ satisfy that $ca = 0$, then we must have that either $c = 0$ or $a = 0$. After all, suppose $c \neq 0$ but $ca = 0$. Then $a = c^{-1}ca = c^{-1}0 = 0$.

- I mentioned on page 518 that A is an F-vector space when we consider defining $c \cdot a = f(c)a$ for all $c \in F$ and $a \in A$. This is easy to show. That said, we do need to explicity assume that $f(c)a = af(c)$ for all $a \in A$ and $c \in F$ in order for us to have that $c \cdot (a_1 a_2) = (c \cdot a_1)a_2 = a_1(c \cdot a_2)$.

- For an example of a non-commutative F-algebra that we'll be caring about, consider the set of matrices $M_n(F)$ and define $f : F \to M_n(F)$ by $c \mapsto cI$.

Lemma: Suppose A is an F-algebra (and $A \neq \{0\}$). If $\dim_F(A) < \infty$ then for all $a \in A$ there exists a unique nonconstant monic polynomial $m_{a;F}(x) \in F[x]$ such that for all $g \in F[x]$ we have that $g(a) = 0 \iff m_{a;F}(x) \mid g(x)$ in $F[x]$.

Note that $m_{a;F}$ is called the minimal polynomial of a over F.

Proof:
Let $e_a : F[x] \to A$ be the evaluation homomorphism given by $e_a(f(x)) := f(a)$. Then e_a is an F-algebra homomorphism (meaning it is a ring homomorphism from $F[x]$ into A and an F-linear map).

Note that e_a is a ring homomorphism specifically because $ca = ac$ for all $c \in F$ and $a \in A$. I won't prove that e_a is a homomorphism though.

We thus have that $\ker(e_a) \lhd F[x]$. Also, by the first isomorphism theorems for F-modules and rings, we have that $\frac{F[x]}{\ker(e_a)} \cong \text{im}(e_a) \subseteq A$ as an F-algebra. Since $\dim_F(A) < \infty$ and $\dim_F(F[x]) = \infty$, we must have that $\ker(e_a) \neq \{0\}$. Hence as $F[x]$ is a P.I.D. and F is a field, there must exist a unique monic polynomial $m_{a;F}(x) \in F[x]$ such that:

$$\ker(e_a) = \langle m_{a;F} \rangle.$$

Also note that $\ker(e_a) \neq F[x]$ as the only constant polynomial contained in $\ker(e_a)$ is the zero polynomial. Therefore, we can conclude that $m_{a;F} \neq 1$. And finally, $g \in \ker(e_a)$ iff $m_{a;F} \mid g$ in $F[x]$. ■

For an application of the prior lemma, note that if $a \in M_n(F)$ then $\text{Ann}(V_a) = \langle m_{a;F}(x) \rangle$.

Why?
$g(x) \cdot V_a = 0$ iff $\forall v \in F^n$, $g(a)|v\rangle = 0$. But the latter happens iff $g(a) = 0$, and that happens iff $m_{a;F}(x) \mid g(x)$.

Here's some more notes on annihilators.

Given a field F and $f \in F[x]$, we have that $\text{Ann}(\frac{F[x]}{\langle f \rangle}) = \langle f \rangle$ (where we are viewing $\frac{F[x]}{\langle f \rangle}$ as an $F[x]$-module).

Why?
If $g \in \langle f \rangle$ then $g(x) = f(x)h(x)$. In turn, for any $s(x) + \langle f \rangle$ in $F[x]/\langle f \rangle$ we have that $g(x)(s(x) + \langle f \rangle) = f(x)h(x)s(x) + \langle f \rangle \equiv 0 \pmod{f(x)}$. And this proves that $\langle f \rangle \subseteq \text{Ann}(\frac{F[x]}{\langle f \rangle})$. To show the other inclusion, suppose $g \in \text{Ann}(\frac{F[x]}{\langle f \rangle})$. Then $g(x)\,(1 + \langle f \rangle) = g(x) + \langle f \rangle = 0 + \langle f \rangle$. The only way this is possible is if $g \in \langle f \rangle$.

More generally, the above reasoning shows that if A is a unital ring and $\mathfrak{a} \lhd A$, then $\text{Ann}(\frac{A}{\mathfrak{a}}) = \mathfrak{a}$ (where we are viewing $A/\mathfrak{a}$ as an A-module).

Suppose $\{M_i\}_{i \in I}$ is a family of A-modules. Then $\text{Ann}(\bigoplus_{i \in I} M_i) = \bigcap_{i \in I} \text{Ann}(M_i)$.

Why?
If $a \in \text{Ann}(\bigoplus_{i' \in I} M_{i'})$ and we consider the projection $P_i : \bigoplus_{i' \in I} M_{i'} \to M_i$, then we know that $a \cdot m_i = P_i(a \cdot (m_{i'})_{i' \in I}) = P_i((0)_{i' \in I}) = 0$. In particular, this shows that $a \cdot m_i = 0$ for all $i \in I$ and $m_i \in M_i$. Hence $a \in \bigcap_{i \in I} \text{Ann}(M_i)$.

Conversely, if $a \in \bigcap_{i \in I} \text{Ann}(M_i)$ then we know that $a \cdot m_i = 0$ for any $m_i \in M_i$ and $i \in I$. In turn, $a \cdot (m_{i'})_{i' \in I} = (a \cdot m_{i'})_{i' \in I} = 0$.

As a corollary, if F is a field and $f_1, \ldots, f_m \in F[x]$, then:

$$\text{Ann}(\textstyle\bigoplus_{i=1}^m \frac{F[x]}{\langle f_i \rangle}) = \bigcap_{i=1}^m \langle f_i \rangle = \langle \text{lcm}(f_1, \ldots, f_m) \rangle.$$

To see what I specifically mean by a least common multiple, note that if D is a P.I.D. and $\langle d_1 \rangle, \ldots, \langle d_m \rangle$ are ideals in D, then there must exist some element $\ell \in D$ with $\langle \ell \rangle = \bigcap_{i=1}^m \langle d_i \rangle$. In turn, ℓ satisfies the property that $d_i \mid \ell$ for all i. Furthermore, if x satisfies that $d_i \mid x$ for all i, then $\ell \mid x$.

Proposition: Suppose F is a field. Then for all $a \in M_n(F)$ we have that $m_{a;F}(x) \in F[x]$ is the largest invariant factor of a.

Proof:

Suppose $f_1 \mid f_2 \mid \cdots \mid f_m$ are the invariant factors of a. Then, we know that:

$$\langle m_{a;F}(x)\rangle = \mathrm{Ann}(V_a) = \mathrm{Ann}(\tfrac{F[x]}{\langle f_1\rangle} \oplus \cdots \oplus \tfrac{F[x]}{\langle f_m\rangle}) = \textstyle\bigcap_{i=1}^m \langle f_i\rangle = \langle f_m\rangle$$

As a side note, if $M_1 \cong M_2$ as A-modules, then both $\mathrm{Ann}(M_1)$ and $\mathrm{Ann}(M_2)$ are equal subsets of A. In other words, it would be incorrect (or at least incredibly misleading) to write $\mathrm{Ann}(M_1) \cong \mathrm{Ann}(M_2)$ as opposed to $\mathrm{Ann}(M_1) = \mathrm{Ann}(M_2)$.

It follows that $m_{a;F}(x)$ and $f_m(x)$ are companion elements of $F[x]$. Yet as both are monic, the only way this is possible is if $m_{a;F}(x) = f_m(x)$. ■

The characteristic polynomial of a matrix $a \in M_n(A)$ (where A is commutative unital ring) is defined as the polynomial $\chi_a(t) := \det(tI - a)$ (where $tI - a \in M_n(A[t])$).

Lemma: If $a_1, \ldots, a_m$ are square matrices over some ring A, then:

$$\det(\mathrm{diag}(a_1, \ldots, a_m)) = \textstyle\prod_{i=1}^m \det(a_i).$$

Proof:

By noting that $\mathrm{diag}(a_1, \ldots, a_{m-1}, a_m) = \mathrm{diag}(\mathrm{diag}(a_1, \ldots, a_{m-1}), a_m)$, it suffices to prove that $\det(\mathrm{diag}(a, b)) = \det(a)\det(b)$ where $a \in M_{n_1}(A)$ and $b \in M_{n_2}(A)$ for some integers n_1 and n_2.

Next, we proceed by induction on n_1. For our base case, note that if $n_1 = 1$ (so that a is just a scalar in A), then we can take the Laplace expansion of the determinant formula to get that:

$$\det(\mathrm{diag}(a, b)) = (-1)^{1+1} a \det(b) + \textstyle\sum_{i=1}^{n_2} 0 = \det(a)\det(b).$$

As for the induction step, note again by the Laplace expansion formula that:

$$\begin{aligned}\det(\mathrm{diag}(a, b)) &= \textstyle\sum_{i=1}^{n_1} (-1)^{1+i} a_{1,i} \det(\mathrm{diag}(a(1,i), b)) \\ &= \textstyle\sum_{i=1}^{n_1} (-1)^{1+i} a_{1,i} \det(a(1,i))\det(b) \\ &= (\textstyle\sum_{i=1}^{n_1} (-1)^{1+i} a_{1,i} \det(a(1,i))) \cdot \det(b) = \det(a)\det(b).\ ■\end{aligned}$$

Lemma: If $b = gag^{-1}$ where $a, b, g \in M_n(A)$, then $\chi_a = \chi_b$.

Proof:

$$\begin{aligned}\chi_b = \det(tI - b) = \det(bI - gag^{-1}) &= \det(g(tI - a)g^{-1}) \\ &= \det(g)\det(tI - a)\det(g^{-1}) = \det(tI - a) = \chi_a.\end{aligned}$$

Lemma: Suppose f is a monic nonconstant polynomial in $F[t]$ (where F is a field). If c_f is the companion matrix then $\chi_{c_f} = f$.

Proof:

If $f(t) = t^n + c_{n-1}t^{n-1} + \ldots + c_1 t + c_0$, then:

$$tI - c_f = \begin{bmatrix} t & 0 & \cdots & 0 & c_0 \\ -1 & t & \cdots & 0 & c_1 \\ 0 & -1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t & c_{n-2} \\ 0 & 0 & \cdots & -1 & t + c_{n-1} \end{bmatrix}$$

By the laplace expansion formula, we get that:

$$\chi_{c_f}(t) = t \det \left(\begin{bmatrix} t & \cdots & 0 & c_1 \\ -1 & \cdots & 0 & c_2 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & t & c_{n-2} \\ 0 & \cdots & -1 & t + c_{n-1} \end{bmatrix} \right) + (-1)^{1+n} c_0 \det \left(\begin{bmatrix} -1 & t & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t \\ 0 & 0 & \cdots & -1 \end{bmatrix} \right)$$

$$= t\chi_{c_{\left(\frac{f-c_0}{t}\right)}} + (-1)^{1+n} c_0 (-1)^{n-1} = t\chi_{c_{\left(\frac{f-c_0}{t}\right)}} + c_0$$

And now it follows by doing induction on the degree of f that $\chi_{c_f} = f$.

Technically I still need to show a base case. Suppose $f(t) = t + c_0$. Then $tI - c_f = [t + c_0]$ and we trivially have that $\chi_{c_f}(t) = t + c_0 = f(t)$. ■

Corollary: If $f_1 \mid f_2 \mid \cdots \mid f_m$ are the invariant factors of a matrix $a \in M_n(F)$ then:
$m_{a;F}(x) = f_m(x)$ and $\chi_a(x) = \prod_{i=1}^m f_i(x)$.

Proof:
We already proved that $m_{a;F}(x) = f_m(x)$. To show the other equality, note that as a is similar to $\operatorname{diag}(c_{f_1}, \ldots, c_{f_m})$ we have that:

$$\chi_a = \chi_{\operatorname{diag}(c_{f_1}, \ldots, c_{f_m})} = \prod_{i=1}^m \chi_{c_{f_i}} = \prod_{i=1}^m f_i. \blacksquare$$

(Note that $\chi_{\operatorname{diag}(c_{f_1}, \ldots, c_{f_m})} = \prod_{i=1}^m \chi_{c_{f_i}}$ because of the first of our three prior lemmas plus the fact that $tI - \operatorname{diag}(c_{f_1}, \ldots, c_{f_m}) = \operatorname{diag}(tI - c_{f_1}, \ldots, tI - c_{f_m})$.)

Consequently, we get the following theorem:

Cayley Hamilton Theorem: Suppose F is a field and $a \in M_n(F)$. Then:

- $m_{a;F}(x) \mid \chi_a(x)$,
- If $p(x)$ is an irreducible factor of $\chi_a(x)$ then $p(x) \mid m_{a;F}(x)$.

2/2/2026

Math 220b Lecture Notes:

The Mittag-Leffler Problem on an open set G goes as follows:

Suppose $\{a_n\}_{n\in\mathbb{N}}$ is a sequence of distinct elments in G with no subsequential limits in G, and for each n suppose we have a function of the form:

$$g_n(z) = \frac{A_{n,m_n}}{(z-a_n)^{m_n}} + \frac{A_{n,m_n-1}}{(z-a_n)^{m_n-1}} + \cdots + \frac{A_{n,1}}{(z-a_n)^1},$$

where $A_{n,m_n} \neq 0$ and $m_n \in \mathbb{Z}_{>0}$ for all n. Then, does there exists a meromorphic function f on $\mathbb{C}$ with poles only at a_n and Laurent principal parts g_n on punctured disks about any a_n.

As we will show, the answer is yes.

Some remarks:

- If we only have finitely many a_n and g_n, then the problem is trivial. After all, we can then just set $f = \sum g_n$. This is why we now focus on the infinite case.

- In the infinite case, we can no longer guarentee that $\sum_{n=1}^{\infty} g_n$ converges. However, notice that if we subtract a function h_n which is holomorphic on G (i.e. has no poles), then that will not effect the principal part of the Laurent expansion about any particular point. Therefore, the general idea of the following proof is that we will try to find convergence enhancing corrections $h_n \in O(G)$ such that $\sum_{n=1}^{\infty}(g_n - h_n)$ converges.

  As a side note, the h_n will not be unique.

- One other observation is that if both f_1 and f_2 solve the same Mittag-Leffler problem, then we must have that $f_1 - f_2$ are entire functions. This is because the principal part of $f_1 - f_2$ about any singularity cancels and we are left with a power series.