## One strategy to find Invariant Subspaces

Suppose we have a linear transformation $L: V \rightarrow V$. and we pick any vector $\vec{v} \in V$.
Consider the sets:

$$
S_{\vec{v}, \ell}=\left\{\vec{v}, L(\vec{v}), \ldots, L^{\ell-1}(\vec{v})\right\}
$$

As we increase $\ell$, the set gets bigger, so either, $V$ is infinite dimensional, or eventually they become linearly dependent.
Suppose such an $\ell$ is the smallest values such that

$$
\vec{v}, L(\vec{v}), \ldots, L^{\ell-1}(\vec{v}), L^{\ell}(\vec{v})
$$

is linearly dependent and define:

$$
W_{\vec{v}}=\operatorname{Span}\left(S_{\vec{v}, \ell}\right)
$$

Note: $L^{\ell}(\vec{v}) \in W_{\vec{v}}$.

## Lemma

As defined above, $W_{\vec{v}}$ is $L$-invariant.
Proof: Let $\vec{w} \in W_{\vec{v}}$ be arbitrary.
Then $\vec{w} \in \operatorname{Span}\left(S_{\vec{v}, \ell}\right)$ and so we may write:

$$
\vec{w}=b_{0} \vec{v}+b_{1} L(\vec{v})+\cdots+b_{n-1} L^{\ell-1}(\vec{v})
$$

Then we have

$$
L(\vec{w})=b_{0} L(\vec{v})+b_{1} L^{2}(\vec{v})+\cdots+b_{n-1} L^{\ell}(\vec{v})
$$

but each vector on the right hand side is in $W_{\vec{v}}$, hence $L(\vec{w}) \in W_{\vec{v}}$.
Note: there may not be an invariant complementary subspace.

The matrix for $L$ on $W_{\vec{v}}$
Because, $L^{\ell}(\vec{v}) \in W_{\vec{v}}$ we know there exists $a_{1}, \ldots, a_{\ell} \in \mathbb{R}$ so that

$$
\overrightarrow{0}=L^{\ell}(\vec{v})+a_{\ell-1} L^{\ell-1}(\vec{v})+\cdots a_{1} L(\vec{v})+a_{0} \vec{v}
$$

## Lemma

With $W_{\vec{v}}$ as defined above, the matrix for $L$ acting on $W_{\vec{v}}$ with respect to the basis

$$
\vec{v}, L(\vec{v}), \ldots, L^{\ell-1}(\vec{v})
$$

is

$$
\left(\begin{array}{cccccc}
0 & & \cdots & & 0 & -a_{0} \\
1 & 0 & & & 0 & -a_{1} \\
0 & 1 & 0 & & 0 & -a_{2} \\
\vdots & & \ddots & & \vdots & \vdots \\
0 & & & 1 & 0 & -a_{\ell-2} \\
0 & & \cdots & & 1 & -a_{\ell-1}
\end{array}\right)
$$

## Proof:

For the first $\ell-1$-columns simply $L$-sends the $i$ th basis vector to the $(i+1)$ st. For the last column, this is precisely the relation we assumed.

## Characteristic Polynomial for Rational Canonical Form

## Theorem

Suppose $L$ is a linear transformation acting on a vector space $V$ with a matrix of the form:

$$
A=\left(\begin{array}{cccccc}
0 & & \cdots & & 0 & -a_{0} \\
1 & 0 & & & 0 & -a_{1} \\
0 & 1 & 0 & & 0 & -a_{2} \\
\vdots & & \ddots & & \vdots & \vdots \\
0 & & & 1 & 0 & -a_{\ell-2} \\
0 & & \cdots & & 1 & -a_{\ell-1}
\end{array}\right)
$$

the characteristic polynomial of $A$ is

$$
\operatorname{char}_{A}(x)=x^{\ell}+a_{\ell-1} x^{\ell-1}+\ldots+a_{1} x+a_{0}
$$

Proof Idea Proceed by induction computing the determinant of:

$$
\left(\begin{array}{cccccc}
x & & \cdots & & 0 & a_{0} \\
-1 & x & & & 0 & a_{1} \\
0 & -1 & x & & 0 & a_{2} \\
\vdots & & \ddots & & \vdots & \vdots \\
0 & & & -1 & x & a_{\ell-2} \\
0 & & \cdots & & -1 & x+a_{\ell-1}
\end{array}\right)
$$

by expansion along the top row or first column. We leave this as an exercise

## Minimal Polynomial for Rational Canonical Form

## Theorem

Suppose $L$ is a linear transformation acting on a vector space $V$ with a matrix of the form:

$$
A=\left(\begin{array}{cccccc}
0 & & \cdots & & 0 & -a_{0} \\
1 & 0 & & & 0 & -a_{1} \\
0 & 1 & 0 & & 0 & -a_{2} \\
\vdots & & \ddots & & \vdots & \vdots \\
0 & & & 1 & 0 & -a_{\ell-2} \\
0 & & \cdots & & 1 & -a_{\ell-1}
\end{array}\right)
$$

the minimal polynomial of $A$ is: $\quad \min _{A}(x)=x^{\ell}+a_{\ell-1} x^{\ell-1}+\ldots+a_{1} x+a_{0}$
Proof Idea Consider $P(x)=x^{\ell}+a_{\ell-1} x^{\ell-1}+\ldots+a_{1} x+a_{0}$. We note

$$
\overrightarrow{0}=L^{\ell}(\vec{v})+a_{\ell-1} L^{\ell-1}(\vec{v})+\cdots a_{1} L(\vec{v})+a_{0} \vec{v}
$$

and hence with we have $P(L)(\vec{v})=\overrightarrow{0}$ but then

$$
P(L)\left(L^{i}(\vec{v})\right)=P(L) \circ L^{i}(\vec{v})=L^{i} \circ P(L)(\vec{v})=L^{i}(\overrightarrow{0})=\overrightarrow{0}
$$

and as $\vec{v}, L(\vec{v}), \ldots, L^{\ell-1}(\vec{v})$ is a basis this proves $P(L)=0$.
We now claim the minimal polynomial has degree at least $\ell$, consider any polynomial of lower degree $Q(x)=b_{\ell-1} x^{\ell-1}+\ldots+b_{1} x+b_{0}$ for which $Q(L)=0$. Then $Q(L)(\vec{v})=0$ but

$$
Q(L)(\vec{v})=b_{\ell-1} L^{\ell-1}(\vec{v})+\cdots b_{1} L(\vec{v})+b_{0} \vec{v}
$$

but these vectors are LI and so the $b_{i}=0$, and hence $Q(x)=0$.

## Sufficient Condition for Complementary Subspaces

Everything about these spaces seems really easy to study.
There is one catch, we won't always have an invariant complementary subspace.
But there is hope, because for at least one choice of $\vec{v}$, there will be, this comes from the following theorem.

## Theorem

Suppose $L: V \rightarrow V$ is any linear transformation of a finite dimensional vector space, let $W_{\vec{v}}$ be any subspace as above then if the $\operatorname{dim}\left(W_{\vec{v}}\right)$ is maximal (there is no $\vec{v}^{\prime}$ with $\left.\operatorname{dim}\left(W_{\vec{v}}\right)<\operatorname{dim}\left(W_{\vec{v}^{\prime}}\right)\right)$ then there exists a complementary subspace so that

$$
V=W_{\vec{v}} \oplus W_{2}
$$

Proof will be done on the next bunch of slides.

## Corollary

There exists a collection $\vec{v}_{1}, \ldots, \vec{v}_{r}$ so that:

$$
V=W_{\vec{v}_{1}} \oplus \cdots \oplus W_{\vec{v}_{r}}
$$

Proof Idea: We need to use the general inductive idea about cutting vector spaces up. That is, just repeat the proceedure on $W_{2}$.
We note that given any finite dimensional vector space $V$ there must exist a $\vec{v}$ for which $\operatorname{dim}\left(W_{\vec{v}}\right)$ is maximal.

## Proof: Sufficient Condition for Complementary Subspaces I

Warning: Proof is ugly, you do not need to know it!!
First: Let $W^{\prime}$ be any complementary subspace, ie suppose $V=W_{\vec{v}} \oplus W^{\prime}$, and let

$$
\vec{e}_{1}, \ldots, \vec{e}_{r}
$$

be a basis for $W^{\prime}$.
Then we know that in the basis $\vec{v}, L(\vec{v}), \ldots, L^{\ell-1}(\vec{v}), \vec{e}_{1}, \ldots, \vec{e}_{r}$ that we have a matrix of the form:

$$
\left.\left(\begin{array}{cccccc}
0 & & \cdots & & 0 & -a_{0} \\
1 & 0 & & & 0 & -a_{1} \\
0 & 1 & 0 & & 0 & -a_{2} \\
\vdots & & \ddots & & \vdots & \vdots \\
0 & & & 1 & 0 & -a_{\ell-2} \\
0 & \cdots & & 1 & -a_{\ell-1} & {\left[\begin{array}{l} 
\\
\\
\\
\\
\\
\\
\end{array}\right.} \\
& & & & \\
& & \\
& & & & \\
C
\end{array}\right]\right)
$$

The matrix $C$ is $r$ by $r$ and $B$ is $\ell$ by $r$, the zero block is $r$ by $\ell$.

## Proof: Sufficient Condition for Complementary Subspaces II

 Given the original matrix $A$We will use the change of basis matrix (where $\vec{b}_{i}$ are the rows from $B$ ):


## Proof: Sufficient Condition for Complementary Subspaces III

 We want to calculate $M^{-1} A M$, but first we calculate $A M$.

The key is how block matricies multiply.

Proof: Sufficient Condition for Complementary Subspaces IV We now multiply by $M^{-1}$


Again, the key is how block matricies multiply.

## Proof: Sufficient Condition for Complementary Subspaces V

This change of basis keeps the same subspace $W_{\vec{v}}$ but replaces $W^{\prime}$ by $W^{\prime \prime}$ by using a new basis for the complementary space.
That is the new basis is:

$$
\vec{v}, L(\vec{v}), \ldots, L^{\ell-1}(\vec{v}), \vec{e}_{1}^{\prime}, \ldots, \vec{e}_{r}^{\prime} \quad \text { where } \quad \vec{e}_{j}^{\prime}=\vec{e}_{j}-\sum_{i=1}^{\ell-2}\left(\vec{b}_{i+1}, \vec{e}_{j}\right) L^{i-1}(\vec{v})
$$

where $\left(\vec{b}_{i+1}, \vec{e}_{j}\right)$ is just the $j$-th entrie of the row $\vec{b}_{i+1}$ from $B$.
In this basis we have the matrix for $L$ of the form:

$$
M^{-1} A M=\left(\begin{array}{cccccc}
0 & & \cdots & 0 & -a_{0} \\
1 & 0 & & 0 & -a_{1} \\
0 & 1 & 0 & & 0 & -a_{2} \\
\vdots & & \ddots & & \vdots & \vdots \\
0 & & 1 & 0 & -a_{\ell-2}
\end{array} \quad\left[\begin{array}{ccc}
\cdots & \vec{b}_{1}+\overrightarrow{b_{2}} C & \cdots \\
\cdots & \overrightarrow{b_{2}} C & \cdots \\
0 & \cdots & \\
\hline & 1 & -a_{\ell-1} \\
\cdots & \overrightarrow{b_{\ell}} C & \cdots \\
\cdots & \overrightarrow{0} & \cdots
\end{array}\right]\right)
$$

Now: we may repeat the process, but at each stage we obtain one additional row of zeros at the bottom.

## Proof: Sufficient Condition for Complementary Subspaces VI

By iterating the above process we may be sure that the matrix for $L$ is of the form:

$$
\left(\begin{array}{cccccc}
0 & & \cdots & & 0 & -a_{0} \\
1 & 0 & & & 0 & -a_{1} \\
0 & 1 & 0 & & 0 & -a_{2} \\
\vdots & & \ddots & & \vdots & \vdots \\
0 & & & 1 & 0 & \begin{array}{c}
-a_{\ell-2} \\
-a_{\ell-1}
\end{array} \\
0 & \cdots & & 1 & \\
{\left[\begin{array}{lllll}
\cdots & \vec{b} & \cdots \\
\cdots & \overrightarrow{0} & \cdots \\
& & & & \vdots \\
\cdots & \overrightarrow{\vec{~}} & \cdots \\
\cdots & \overrightarrow{0} & \cdots
\end{array}\right]} \\
& & 0 & &
\end{array}\right]
$$

So there exists a complementary space $W^{\prime \prime \prime}$ where the matrix is as above.
Next need to kill off the $\vec{b}$.
Note: It is possible to define a single matrix $M^{\prime}$ at the start which accomplishes the whole iterated process in a single step, the eventual result is that

$$
\vec{b}=\sum_{i=1}^{\ell} \vec{b}_{i} C^{i-1}
$$

The variant $M^{\prime}$ to use has

$$
\vec{b}_{j}^{\prime}=-\sum_{i=1}^{\ell-j} \vec{b}_{j+i} C^{i-1}
$$

We can likewise explicitly write out the new basis using these... what really matters is it exists!

That is with $M^{\prime}$ defined by



we have

$$
M^{\prime-1} A M^{\prime}=\left(\begin{array}{ccccc}
0 & & \cdots & 0 & -a_{0} \\
1 & 0 & & 0 & -a_{1} \\
0 & 1 & 0 & & 0 \\
-a_{2} \\
\vdots & & \ddots & & \vdots \\
\vdots \\
0 & & & 1 & 0 \\
0 & -a_{\ell-2} \\
0 & \cdots & & 1 & -a_{\ell-1}
\end{array}\left[\begin{array}{c}
\vec{b}_{1}+\vec{b}_{2} c+\vec{b}_{3} c^{2}+\cdots+\vec{b}_{\ell} c^{\ell-1} \\
\overrightarrow{0} \\
\overrightarrow{0} \\
\vdots \\
\overrightarrow{0} \\
\overrightarrow{0}
\end{array}\right]\right)
$$

This is a direct thing to check, it works like the previous case, we leave this as an exercise.

## Proof: Sufficient Condition for Complementary Subspaces VII

Lemma If $\left(b_{1}, \ldots, b_{n}\right)=\vec{b} \neq \overrightarrow{0}$ then we have a contradiction.
Define the vector $\vec{u}=\left(0, \ldots, 0, b_{1}, \ldots, b_{n}\right)$ that is, the first $\ell$ coordinates are 0 , follwed by the entries of $b$.

Looking at the shape of $A$, we easily compute

But these vectors are $\ell+1$ linearly independent vectors!
Why: The ( $\vec{b}, \vec{b}^{t}$ ) entry will be a pivot for the last $\ell$ columns, $\vec{b}^{t}$ gives a pivot somewhere in the first. and so $\operatorname{Dim}\left(W_{\vec{u}}\right)>\operatorname{Dim}\left(W_{\vec{v}}\right)$ a contradiction.

## Proof: Sufficient Condition for Complementary Subspaces VIII

 Now that we know $\vec{b}=0$, this tells us our matrix is actually of the form:$\left(\begin{array}{cccccc}0 & & \cdots & & 0 & -a_{0} \\ 1 & 0 & & & 0 & -a_{1} \\ 0 & 1 & 0 & & 0 & -a_{2} \\ \vdots & & \ddots & & \vdots & \vdots \\ 0 & & & 1 & 0 & -a_{\ell-2} \\ 0 & \cdots & & 1 & -a_{\ell-1}\end{array}\right]\left[\begin{array}{l}0 \\ {\left[\begin{array}{lllll} & & & & \end{array}\right]}\end{array}\right)$
and thus that we have a direct sum decompositon of

$$
V=W_{\vec{v}} \oplus W^{\prime}
$$

where both spaces are L-invariant!
There are other proofs of the above result. This proof relies on one clever/ugly matrix calculation.

## Now take a breath...

The point is that it is always possible to get a decomposition

$$
V=W_{\vec{v}_{1}} \oplus \cdots \oplus W_{\vec{v}_{r}}
$$

Which leads to a matrix structure

$$
A=\left(\begin{array}{llll}
A_{1} & & & \\
& A_{2} & & \\
& & \ddots & \\
& & & A_{r}
\end{array}\right)
$$

where each $A_{i}$ is of the form:

$$
A_{i}=\left(\begin{array}{cccccc}
0 & & \cdots & & 0 & -a_{0} \\
1 & 0 & & & 0 & -a_{1} \\
0 & 1 & 0 & & 0 & -a_{2} \\
\vdots & & \ddots & & \vdots & \vdots \\
0 & & & 1 & 0 & -a_{\ell-2} \\
0 & & \cdots & & 1 & -a_{\ell-1}
\end{array}\right)
$$

Note: The theorem gives us a strategy for finding the list of vectors, and a proceedure to carry out the recursion, however, for large matricies the process of finding the maximal blocks is a bit long to do by hand, as is computing the complementary invariant subspace to do the recursion.
None the less, for small matricies this process will end up being part of the algorithm to find the Jordan canonical form! We will revisit this when we come to it.

## Constructing Linear Transformations From Polynomials

So given any polynomial $P(x)=x^{\ell}+a_{\ell-1} x^{\ell-1}+\ldots+a_{1} x+a_{0}$ there is a matrix with minimal polynomial and characteristic polynomial $P(X)$.
If we denote the above matrix $A_{P}$ then we note the following:
If $P_{1}(x), \ldots, P_{r}(x)$ are a collection of polynomials then the matrix:

$$
A=\left(\begin{array}{llll}
A_{P_{1}} & & & \\
& A_{P_{2}} & & \\
& & \ddots & \\
& & & A_{P_{r}}
\end{array}\right)
$$

has characteristic polynomial and minimal polynomial

$$
\begin{aligned}
\operatorname{char}_{A}(x) & =P_{1}(x) \cdot P_{2}(x) \cdots P_{r}(x) \\
\min _{A}(x) & =\operatorname{LCM}\left(P_{1}(x), P_{2}(x), \ldots, P_{r}(x)\right)
\end{aligned}
$$

## Existance of a Rational Canonical Form

## Theorem

Every matrix is similar to one of the form

$$
A=\left(\begin{array}{llll}
A_{P_{1}} & & & \\
& A_{P_{2}} & & \\
& & \ddots & \\
& & & A_{P_{r}}
\end{array}\right)
$$

for some collection of polynomials $P_{1}(x), \ldots, P_{r}(x)$, such that

$$
\operatorname{char}_{A}(x)=P_{1}(x) \cdot P_{2}(x) \cdots P_{r}(x) \quad \min _{A}(x)=\operatorname{LCM}\left(P_{1}(x), P_{2}(x), \ldots, P_{r}(x)\right)
$$

Immediate from our previous work.
The above result is the main reason we did any of this, it will be useful when completing the construction of the Jordan Canonical Form.

Theorem (Cayley-Hamilton)
For any linear transformation $L: V \rightarrow V$ on a finite dimensional vector space we have

$$
\min _{L}(x) \mid \operatorname{char}_{L}(x)
$$

Immediate from the above as the LCM divides the product.

We will not prove this result, It is way outside the scope of this course. Theorem (Rational Canonical Form - One Variation)
Suppose $L: V \rightarrow V$ is a linear transformation of finite dimensional vector spaces, with characteristic polynomial $P(x)$ and minimal polynomial $Q(x)$, then there exist polynomials.

$$
Q_{1}(x)\left|Q_{2}(x)\right| Q_{3}(x)|\cdots| Q_{r}(x)=Q(x)
$$

such that $P(x)=Q_{1}(x) Q_{2}(x) Q_{3}(x) \cdots Q_{r}(x)$ and a basis for $V$ with respect to the matrix for $L$ is

$$
\left(\begin{array}{ccccc}
A_{1} & 0 & \cdots & & 0 \\
0 & A_{2} & \ddots & & \\
\vdots & \ddots & \ddots & & \vdots \\
0 & & 0 & A_{r-1} & 0 \\
0 & \cdots & 0 & A_{r}
\end{array}\right)
$$

Where $A_{i}$ is the matrix associated to $Q_{i}(x)$. All of $Q_{i}(x)$ and the change of basis matrix have entries in the field of coefficients of $L$ (namely if the original matrix for $L$ had entries in $\mathbb{Q}, \mathbb{Q}[\sqrt{2}], \mathbb{R}$ so can the new one, and so does the change of basis matrix.

The fact that we can do this without using complex numbers is in principal nice... but it makes proving the above much harder.

## Important Consequence We Need

Completing the details of the following are a bonus question on the assignment

## Theorem

If $L: V \rightarrow V$ has $\operatorname{char}_{L}(x)=x^{n}$, then there exists a basis for $V$ such that the matrix for $L$ has the form:

$$
\left(\begin{array}{ccccc}
A_{1} & 0 & \cdots & & 0 \\
0 & A_{2} & \ddots & & \\
\vdots & \ddots & \ddots & & \vdots \\
0 & & 0 & A_{r-1} & 0 \\
0 & & \cdots & 0 & A_{r}
\end{array}\right)
$$

Where each of that $A_{i}$ has the form:

$$
\left(\begin{array}{cccccc}
0 & 1 & 0 & & \cdots & 0 \\
0 & 0 & 1 & 0 & & \\
\vdots & & \ddots & \ddots & \ddots & \vdots \\
& & & 0 & 1 & 0 \\
0 & & \cdots & & 0 & 1 \\
0
\end{array}\right)
$$

## Corollary

If the characteristic polynomial of $L$ is instead $(x-\lambda)^{n}$ then the blocks $A_{i}$ above instead have $\lambda$ on the diagonal.

## Natural Questions About Rational Canonical Form

- How do I identify the rational canonical form of a matrix/linear transformation? We will not cover this, you do not need to know it.
- How do I find a basis that does this?

We will not cover this in general, you do not need to know it.
In practice finding maximal $W_{\vec{v}}$ is easy (because random vectors will work with high probability), My proof effectively gives an algorithm for finding the complementary space.

- So why did we do this?

The rational form is a really nice way to find a simple block decomposition, but it is awkward to use in general.
We will end up only using it in the special case where $\operatorname{char}_{A}(x)=x^{n}$, in which case things will be simpler. All the $a_{i}=0!$ !
So the next thing we need to talk about, is a decomposition that allows us to reduce to the case $\operatorname{char}_{A}(x)=x^{n}$.

