(VI.D) Generalized Eigenspaces

Let \( T : \mathbb{C}^n \to \mathbb{C}^n \) be a fixed linear transformation. For this section and the next, all vector spaces are assumed to be over \( \mathbb{C} \); in particular, we will often write \( V \) for \( \mathbb{C}^n \). In what follows, I will write “\( S \)” for an “arbitrary” transformation, which could be \( T \), or \( \sigma \mathbb{I} - T \), or its restriction to a subspace, etc.

We are looking for forms \( A = [T]_\ell \) can be put into (via \( S_B^{-1} A S_B \)) even if it is not diagonalizable. The structure underlying the rational canonical form was a decomposition of \( V = \mathbb{C}^n \) into \( T \)-cyclic subspaces in 1-1 correspondence with the nontrivial invariant factors \( \Delta_r(\lambda), \ldots, \Delta_n(\lambda) \) of \( A \). In the present section we describe the structure beneath the Jordan canonical form – which, unlike the rational form, actually reduces to \( D \) when \( A \) is diagonalizable (\( = S_B D S_B^{-1} \)). We can forget about most of the \( F[\lambda] \) stuff here; the theory is fortunately easier than that in the last two sections.

Recall that if \( A \) is diagonalizable with eigenvalues \( \{\sigma_1, \ldots, \sigma_m\} \),\(^1\) then \( V \) is the sum of the corresponding eigenspaces and in fact the geometric multiplicities add to \( n \):

\[
\sum \dim E_{\sigma_i}(A) = n.
\]

In the language of direct sums,

\[
V = E_{\sigma_1}(A) \oplus \cdots \oplus E_{\sigma_m}(A).
\]

What we claim is that there are “generalized” eigenspaces \( E^s_{\sigma_i} \) such that

\[
V = E^s_{\sigma_1}(A) \oplus \cdots \oplus E^s_{\sigma_m}(A)
\]

---

\(^1\)Here we mean the list of distinct eigenvalues, i.e. not repeated according to multiplicity.
even if $A$ is not diagonalizable.

The proof will require a few facts about stable image/kernel, and nilpotent transformations ($S : U \to U$ such that $S^k$ is the zero transformation for some $k$). Throughout it is important to remember that if $W \subseteq V$ is closed under the action of $T$ then the restriction of $T$ to $W$ makes sense as a linear transformation and is written $T|_W$ (and read “$T$ on $W$”).

**Stable Image and Kernel.** Given a transformation $S : V \to V$, the series of subspaces of $V$

$$\{0\} = \ker I \subseteq \ker S \subseteq \ker S^2 \subseteq \ldots$$

and

$$V = \im I \supseteq \im S \supseteq \im S^2 \supseteq \ldots$$

both level off at some point (since $V$ is finite dimensional). Let $K$ be sufficiently large that

$$\im S^K = \im S^{K+1} = \ldots$$

$$\ker S^K = \ker S^{K+1} = \ldots ;$$

these are called the **stable image** and **stable kernel** of $S$. An equivalent definition of these objects (subspaces of $V$) is:

$$\ker^s S = \left\{ \vec{w} \in V \mid S^k \vec{w} = 0 \text{ for some } k \right\}$$

$$\im^s S = \left\{ \vec{w} \in V \mid \text{for every } k, \exists \vec{v} \in V \text{ s.t. } \vec{w} = S^k \vec{v} \right\} .$$

**Remark 1.** The $\vec{v}$ such that $S^k \vec{v} = \vec{w}$ in the second definition are in general different for each $k$ (even for $k \geq K$).

We claim that

(i) $\im^s S \cap \ker^s S$, \hspace{1cm} (ii) $\im^s S + \ker^s S = V$.

To see (i), let $\vec{w} \in \im^s S \cap \ker^s S$; that is, $\vec{w} = S^k \vec{v}$ and $S^k \vec{w} = 0$, so that $0 = S^k(S^k \vec{v}) = S^{2k} \vec{v}$. But then $\vec{v} \in \ker S^{2K} = \ker^s S = \ker S^K$, so that $(\vec{w} =) S^k \vec{v} = 0$. 
To see (ii), apply rank-nullity to $S^K$ to get

$$\dim V = \dim(\text{im}S^K) + \dim(\ker S^K) = \dim(\text{im}^sS) + \dim(\ker^sS),$$

and the “modular law” $\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim W_1 + \dim W_2$ (cf. Exercise II.C.2) for subspaces $W_1, W_2 \subseteq V$ to get

$$(\text{im}^sS) + \dim(\ker^sS) = \dim(\text{im}^sS \cap \ker^sS) + \dim(\text{im}^sS + \ker^sS) \tag{i}$$

Combining this with (1), $\dim(\text{im}^sS + \ker^sS) = \dim V$ and (ii) follows. We rewrite (i) and (ii) as

$$V = \text{im}^sS \oplus \ker^sS.$$  

This is always true, for any $S : V \rightarrow V$. Moreover, since $S$ respects this decomposition (as you can check), one may speak of the restrictions $S|_{\ker^sS}$ and $S|_{\text{im}^sS}$. By definition some power $k$ of $S$ annihilates $\ker^sS$, and so $S|_{\ker^sS}$ is nilpotent. On the other hand,

$$\ker(S|_{\text{im}^sS}) = \ker S \cap \text{im}^sS \subseteq \ker^sS \cap \text{im}^sS = \{0\}$$

by (i) above, and thus $S|_{\text{im}^sS}$ is invertible. We have proved

**Proposition 2.** Given any $S : V \rightarrow V$, there is a direct-sum decomposition

$$V = U_0 \oplus W_0$$

respected by $S$, such that $S|_{W_0}$ is nilpotent and $S|_{U_0}$ is invertible.

Now let’s look more generally at the situation where $S$ respects a (possibly different) direct sum decomposition $V = U \oplus W$. We claim that

(a) $\ker S = (U \cap \ker S) + (W \cap \ker S)$, and
(b) $(U \cap \ker S) \cap (W \cap \ker S) = \{0\}$.

Now (b) is immediate since $U \cap W = \{0\}$. To see (a): take any $\vec{v} \in \ker S$ and write it $\vec{v} = \vec{u} + \vec{w}$ (possible because $V = U \oplus W$); clearly $0 = S\vec{v} = S\vec{u} + S\vec{w}$. Since $S$ respects $U \oplus W$, $S\vec{u} \in U$ and $S\vec{w} \in W$, but then $S\vec{u} = -S\vec{w}$ is a “problem” since $U \cap W = \{0\}$. So we must
have $Su = Sw = 0$. That means $u \in U \cap \ker S$, $w \in W \cap \ker S$, and since $\bar{v}$ is their sum we have proved (a).

Of course (a)+(b) $\implies \ker S = (U \cap \ker S) \oplus (W \cap \ker S)$, so applying this to $S^k$ we get

**Proposition 3.** Given $S : V \to V$ respecting some direct-sum decomposition

$$V = U \oplus W,$$

one has

$$\ker S^n = (U \cap \ker S) \oplus (W \cap \ker S).$$

**Nilpotent Transformations.** Every $S : V \to V$ has an eigenvalue (unless $V = \{0\}$), since the characteristic polynomial $f_S(\lambda)$ has a root in $\mathbb{C}$. (This is where we really need $V = \mathbb{C}^n$.) This eigenvalue has at least one nonzero eigenvector. What if zero is the only one?

**Proposition 4.** $S$ is nilpotent $\iff 0$ is its only eigenvalue.

**Proof.** ($\iff$) Suppose $0 = $ only eigenvalue of $S = $ only root of $f_S(\lambda)$. That is, $f_S(\lambda) = \lambda^n$. By Cayley-Hamilton, $S$ satisfies its own characteristic polynomial, so $S^n = 0$.

($\Rightarrow$) Suppose $S^k = 0$, and also suppose $\lambda$ is an eigenvalue of $S$. There is a nonzero $\bar{v}$ such that $S\bar{v} = \lambda\bar{v}$, and thus

$$0 = S^k\bar{v} = \lambda^k\bar{v} \implies \lambda^k = 0 \implies \lambda = 0.$$

□

**Stable Eigenspace.** Given $\lambda$ and eigenvalue of $S : V \to V$ ($\iff \lambda$ any root $\in \mathbb{C}$ of $f_S(\lambda)$), recall the definition

$$E_\lambda(S) := \ker(\lambda \mathbb{I} - S) = \{\bar{v} \in V \mid (\lambda \mathbb{I} - S)\bar{v} = 0\}$$

of the eigenspace of $\lambda$. Define the generalized or stable eigenspace

$$E^s_\lambda(S) := \ker(\lambda \mathbb{I} - S) = \{\bar{v} \in V \mid (\lambda \mathbb{I} - S)^k\bar{v} = 0 \text{ for some } k\}.$$

Clearly $E^s_\lambda(S) \supseteq E_\lambda(S)$. 
Now we return to our original $T : V \to V$ with distinct eigenvalues $\{\sigma_1, \ldots, \sigma_m\}$, and set

$$W_k = E_{\sigma_k}^s(T).$$

(These are not the $W_k$’s of §VI.C!) Clearly some power of $(\sigma_k \mathbb{I} - T)$ annihilates $W_k$, so that $(\sigma_k \mathbb{I} - T) |_{W_k}$ is nilpotent and has only eigenvalue 0. That is, if $\vec{v} \in W_k$ satisfies

$$(\sigma_k \mathbb{I} - T)\vec{v} = \lambda \vec{v},$$

then $\lambda = 0$. Therefore, if $\vec{v} \in W_k$ satisfies

$$T \vec{v} = \sigma \vec{v},$$

then

$$(\sigma_k \mathbb{I} - T)\vec{v} = (\sigma_k - \sigma) \vec{v}$$

and $\sigma_k - \sigma$ must be 0, i.e. $\sigma = \sigma_k$.

**Conclusion:** the only eigenvalue of $T |_{W_k}$ is $\sigma_k$.

Now consider for $i \neq j$ the intersection of two stable eigenspaces $W_i \cap W_j$.

The only eigenvalue of $T |_{W_i}$ is $\sigma_i$, while the only eigenvalue of $T |_{W_j}$ is $\sigma_j$. Since $\sigma_i \neq \sigma_j$, $T |_{W_i \cap W_j}$ can have no eigenvalue. This is absurd unless $W_i \cap W_j = \{0\}$.

**Proposition 5.** $E_{\sigma_i}^s(T) \cap E_{\sigma_j}^s(T) = \{0\}$ for all $i \neq j$.

We make one further observation concerning stable eigenspaces: how to find bases for them. You know how to find bases for kernels. Working in the standard basis of $\mathbb{C}^n$ (in terms of which $[T]_e = A$ by definition), find bases for

$$\ker(\sigma_i \mathbb{I} - A) \subseteq \ker\left\{(\sigma_i \mathbb{I} - A)^2\right\} \subseteq \ker\left\{(\sigma_i \mathbb{I} - A)^3\right\} \subseteq \ldots.$$
You stop when two successive bases have the same number of elements (once \( \ker S^k = \ker S^{k+1} \), all the remaining ones are the same as well: a nice exercise!).

**The Jordan Structure Theorem.** Here is what holds even when \( T \) is not semisimple (\( \Leftrightarrow A \) is not diagonalizable). We emphasize that the \( W_k \) have nothing to do with those in the preceding lecture.

**Theorem.** Let \( T : V \to V (V = \mathbb{C}^n) \) be a linear transformation, with distinct eigenvalues \( \{ \sigma_1, \ldots, \sigma_m \} \) and corresponding stable eigenspaces \( W_k = E^{s}_{\sigma_k}(T) = \ker^s(\sigma_k I - T) \). Then

\[
V = W_1 \oplus \cdots \oplus W_m
\]

where \( \dim W_k \) = algebraic multiplicity of \( \sigma_k \). Furthermore, \( T \) respects this decomposition.

**Proof.** We first prove the decomposition, by induction on \( m \). Set \( d_k = \dim W_k \), and \( A = [T]_S \).

Case \( m = 1 \) : \( \sigma_1 \) = the only eigenvalue of \( T \) on \( V \implies 0 = \) only eigenvalue of \( (\sigma_1 I - T) \) on \( V \implies (\sigma_1 I - T) \) nilpotent \( \implies (\sigma_1 I - T)^k = 0 \implies V = \ker^s(\sigma_1 I - T) = W_1 \).

Inductive Step : Assume the Theorem holds for transformations with \( m - 1 \) distinct eigenvalues, and let \( T \) be as above. Apply the discussion preceding Proposition 2 to \( (\sigma_m I - T) \) to get

\[
V = \ker^s(\sigma_m I - T) \oplus \im^s(\sigma_m I - T) =: W_m \oplus U_m,
\]

where \( \sigma_m I - T \) respects the decomposition. Moreover, since \( I \) also respects the direct sum (or, for that matter, and direct sum!), so do \( T \) and \( \sigma_k I - T \), \( k \neq m \). So we may speak of \( T |_{U_m} : U_m \to U_m \).

Since \( (\sigma_m I - T) \) is invertible on \( U_m \), \( \sigma_m \) cannot be an eigenvalue of \( T \) there. \(^2\)

Thus \( T |_{U_m} \) has eigenvalues \( \subseteq \{ \sigma_1, \ldots, \sigma_{m-1} \} \), and by induction

\[
U_m = \prime W_1 \oplus \cdots \oplus \prime W_{m-1},
\]

\(^2\)(\( \sigma_m I - T |_{U_m} \) invertible \( \implies \) \( \det (\sigma_m I - T |_{U_m}) \neq 0 \implies \sigma_m \) not a root of \( \det (\lambda I - T |_{U_m}) \).
where
\[ \prime W_k = \ker^s (\sigma_k I - T | U_m) = \ker^s (\sigma_k I - T) \cap U_m = W_k \cap U_k. \]

We must show that \( \prime W_k = W_k \).

Since \((k \neq m) \sigma_k I - T \) also respects the decomposition \( V = W_m \oplus U_m \), we have (Prop. 3)
\[ W_k = \ker^s (\sigma_k I - T) = \{ W_m \cap \ker^s (\sigma_k I - T) \} \oplus \{ U_m \cap \ker^s (\sigma_k I - T) \} = W_m \cap W_k \oplus U_m \cap W_k. \]

By Prop. 4, \( W_m \cap W_k = \{ 0 \} \) and so
\[ W_k = U_m \cap W_k = \prime W_k, \]
as desired.

\( T \) respects the direct sum: We need to show \( T(W_k) \subseteq W_k \). Take \( \bar{w} \in \ker^s (\sigma_k I - T) \), so that for \( \kappa \) sufficiently large \((\sigma_k I - T)\kappa \bar{w} = 0 \). But then \((\sigma_k I - T)^k T \bar{w} = T(\sigma_k I - T)^k \bar{w} = 0 \implies T \bar{w} \in \ker^s (\sigma_k I - T) \).

\( d_k = \) algebraic multiplicity of \( \sigma_k \) (as roots of \( p_A(\lambda) \)) : We take \( B_1, \ldots, B_m \) to be bases for \( W_1, \ldots, W_m \); the collection \( B = \{ B_1, \ldots, B_m \} \) yields a basis for \( V \) “subordinate to the direct sum”. Since \( T \) respects the direct sum, its matrix with respect to \( B \) splits into blocks down the diagonal (of dimensions \( d_1 \times d_1, \ldots, d_m \times d_m \)):
\[
[T]_B =: B = S_B^{-1} A S_B = \text{diag}\left\{ [T|_{W_1}]_{B_1}, \ldots, [T|_{W_m}]_{B_m} \right\} = \text{diag}\left\{ B_1, \ldots, B_m \right\}.
\]

Moreover, since \( A \sim B, \lambda I - A \sim \lambda I - B \) and \( f_A(\lambda) = f_B(\lambda) \). From
\[
\lambda I - B = \text{diag}\left\{ \lambda I_{d_1} - B_1, \ldots, \lambda I_{d_m} - B_m \right\}
\]
we have
\[
f_B(\lambda) = \det(\lambda I - B) = \prod_k \det(\lambda I_{d_k} - B_k) = f_{B_1}(\lambda) \cdots f_{B_m}(\lambda).
\]

Since the only eigenvalue of \( T|_{W_k} \) is \( \sigma_k \) (and \( B_k = [T|_{W_k}]_{B_k} \)) the only root of \( f_{B_k}(\lambda) \) is \( \sigma_k \). Since \( B_k \) is \( d_k \times d_k \), it follows that \( \deg(f_{B_k}) = d_k \).
and so $f_{B_k}(\lambda) = (\lambda - \sigma_k)^{d_k}$. But then $(f_A(\lambda) = )$

$$f_B(\lambda) = \prod (\lambda - \sigma_k)^{d_k}$$

and we are done. \hfill \square

**Exercises**

1. Find the stable eigenspaces of

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$ 

2. Suppose $A$ is an $8 \times 8$ matrix with $m_A(\lambda) = \lambda(\lambda - 1)^2(\lambda - 2)^3$

and $f_A = \lambda^2(\lambda - 1)^2(\lambda - 2)^4$. What are the dimensions of the eigenspaces and stable eigenspaces of $A$?

3. Show that for any transformation $S : C^n \rightarrow C^n$, $\ker(S^k) = \ker(S^{k+1})$ implies

(a) $\ker(S^k) = \ker(S^\ell)$ for all $\ell \geq k$, and

(b) $\im(S^k) = \im(S^\ell)$ for all $\ell \geq k$. [Hint for (b): use Rank + Nullity and (a).]