RATIONAL CANONICAL AND JORDAN FORMS

PATRICK BROSNAN

1. Introduction

These are some notes about polynomials and rational canonical form for Math 405. They mostly cover the material in Chapers 4, 6 and 7 of *Linear Algebra* by Hoffman and Kunze. But the proof of the existnce of rational canonical form given here in Theorems 4.8 and 4.10 uses an argument involving duality which seems to make the proof shorter.

1.1. **Notation.** I write $\mathbb{N} = \{0, 1, \ldots\}$, $\mathbb{P} = \{1, 2, \ldots\}$. I write $\chi(T)$ for the characteristic polynomial of a linear operator T on a finite dimensional vector space.

2. Consequences of the Euclidean algorithm for polynomial rings

In this section, F is a field and $F[x]$ is the ring of polynomials. We know that $F[x]$ is a principal ideal domain. In other words, every ideal in $F[x]$ is of the form $pF[x]$ where p is a polynomial.

Definition 2.1. Suppose $p_1, \ldots, p_k$ are polynomials in $F[x]$ which are not all 0. Set $I = \langle p_1, \ldots, p_k \rangle$. Let d denote the monic generator of I. We call d the *greatest common divisor of the* p_i and write $d = \gcd(p_1, \ldots, p_k)$. If $d = 1$, we say that the polynomials $p_1, \ldots, p_k$ is *relatively prime*. Note that, by definition, there exists polynomials $q_1, \ldots, q_k \in F[x]$ such that

$$d = p_1 q_1 + \cdots + p_k q_k.$$

So, if $p_1, \ldots, p_k$ are relatively prime, we can find $q_1, \ldots, q_k$ such that $\sum_{i=1}^k p_i q_i = 1$.

Lemma 2.2. *Suppose $p, q \in F[x]$ are not both 0. Set $d = \gcd(p, q)$. If $e|p$ and $e|q$, then $e|d$.*

Proof. Write $d = ap + bq$ with $a, b \in F[x]$. Then it is obvious. □

Corollary 2.3. *Suppose $p, q \in F[x]$ are not both 0. Set $d = \gcd(p, q)$. Then $\gcd(p/d, q/d) = 1$.*

Proof. Suppose $e|(p/d)$ and $e|(q/d)$ for some monic polynomial e. Then $ed|p$ and $ed|q$. So $ed|d$. So $e = 1$. □

Theorem 2.4. *Suppose p, q are relatively prime polynomials in $F[x]$, and $f \in F[x]$. Then $p|qf \Leftrightarrow p|f$.*

Proof. ($\Leftarrow$) is obvious. ($\Rightarrow$): Write $1 = ap + bq$. Then $p|qf \Rightarrow p|(ap + bq)f = f$. □

Corollary 2.5. *Suppose p, q are relatively prime polynomials and f is a polynomial which is divisible by both p and q. Then $pq|f$.*

Proof. Suppose $f = pa$ for some $a \in F[x]$. Then $q|f \Rightarrow q|a$. So $a = qb$ for some $b \in F[x]$. So $f = pqb$. □

Definition 2.6. A non-constant polynomial p is *reducible* if $p = ab$ for a, b non-constant polynomials. Otherwise not-constant polynomial p which is not reducible is called *irreducible*.

Theorem 2.7. *Suppose p is irreducible and $a, b \in F[x]$. Then $p|ab \Rightarrow p|a$ or $p|b$.*

Proof. We can assume that p is monic. Then $\gcd(p, a)$ is either 1 or p. If $\gcd(p, a) = p$ then $p|a$. Otherwise $\gcd(p, a) = 1$. So $p|b$ by Theorem 2.4. □

Theorem 2.8. *Suppose f is a non-constant, monic, polynomial. Then there exists irreducible polynomials $p_1, \ldots, p_k$ such that $f = p_1 \cdots p_k$. These are unique up to reordering.*

The expression $f = p_1 \cdots p_k$ is called the *factorization of f into irreducibles*.

Proof. Suppose there exists a non-constant, monic polynomial which cannot be written as a product of irreducibles. Then let f be a non-constant, monic polynomial of smallest possible degree which cannot be written as such a product. The polynomial f cannot itself be irreducible (obviously). So we must have $f = ab$ with $\deg a, \deg b < \deg f$. But then a and b can be written as products or irreducibles. So so can f.

Supppose there exists a monic, non-constant polynomial with two different factorizations into irreducibles. We can pick one of smallest possible degree. Write $f = p_1 \cdots p_n = q_1 \cdots q_m$ for the two factorizations. If $p_i = q_j$ for some i, j then we can factor out p_i and see that f/p_i has two different factorizations. So we can asssume that the sets $\{p_1, \ldots, p_n\}$ and $\{q_1, \ldots, q_m\}$ are disjoint. But $p_1|f$. So p_1 must divide one of the q_i. So $p_1 = q_i$ for some i. Contradiction.

□

Definition 2.9. We say $a \in F$ is a root of a polynomial $f \in F[x]$ if $f(a) = 0$. We say F is algebraically closed if every non-constant polynomial has a root in F. For example, $\mathbb{C}$ is algebraically closed.

Proposition 2.10. *Suppose $f \in F[x]$ and $a \in F$. Then $f(a) = 0 \Leftrightarrow (x - a)|f$.*

Proof. Write $f = (x - a)g + r$ with r constant. Then $f(a) = r$. □

Theorem 2.11. *Suppose F is algebraically closed. Then a monic polynomial $p \in F[x]$ is irreducible iff $p = x - a$ for some $a \in F$.*

Proof. Obviously $(x - a)$ is irreducible. Suppose p is monic, irreducible. Let a be a root of p. Then $p = (x - a)q$ for some monic $q \in F[x]$. This is a contradiction unless $q = 1$. □

3. Annihilators and cyclic subspaces

In this section V is a finite dimensional vector space over a field F and $T \in L(V)$ is a linear operator on V.

Definition 3.1. Write $\mathrm{Ann}(T) = \{f \in F[x] : f(T) = 0\}$.

We know that $\mathrm{Ann}(T)$ is a non-zero ideal in $F[x]$. So that leads us to the following definition.

Definition 3.2. Write $\min(T)$ for the monic generator of $\mathrm{Ann}(T)$.

Definition 3.3. A subspace $W \subset V$ is called *T-stable* if $TW \subset W$.

Example 3.4. Suppose $\alpha \in V$. Then set

$$F[T]\alpha := \langle \alpha, T\alpha, T^2\alpha, \ldots \rangle.$$

This is the *cyclic subspace generated by* α. It is pretty obvious that $F[T]\alpha$ is a T-stable subspace of V. A subspace W of V is said to be *cyclic* if $W = F[T]\alpha$ for some $\alpha \in W$.

Lemma 3.5. *Suppose A and B are T-stable subspaces of V. Then so are $A + B$ and $A \cap B$.*

Proof. Easy exercise. □

Definition 3.6. Suppose W is a T-stable subspace. We write $T_{|W}$ (or occasionally $T|W$) for the restriction of T to W. So $T_{|W}$ is the linear operator on W sending any $w \in W$ to Tw. Write $\operatorname{Ann}(T, W)$ for $\operatorname{Ann}(T_{|W})$ and $\min(T, W)$ for $\min(T_{|W})$. If $\alpha \in V$, write $\operatorname{Ann}(T, \alpha) := \operatorname{Ann}(T, F[T]\alpha)$ and $\min(T, \alpha) := \min(T, F[T]\alpha)$. If T is fixed, as it often will be for us, we drop it from the notation and just write $\operatorname{Ann}(W), \min(W), \operatorname{Ann}(\alpha)$ and $\min(\alpha)$.

Proposition 3.7. *We have* $\operatorname{Ann}(T, \alpha) = \{f \in F[x] : f(T)\alpha = 0\}$.

Proof. Almost obvious. Just need to realize that $f(T)\alpha = 0 \Rightarrow f(T)q(T)\alpha = 0$ for any $q \in F[x]$. □

Proposition 3.8. *Suppose $W \subset V$ is T-stable. Then* $\operatorname{Ann}(W) \supset \operatorname{Ann}(V)$. *Consequently* $\min(W) | \min(V)$.

Proof. Obvious. If $f \in \operatorname{Ann}(T, V)$ then $f(T)\alpha = 0$ for all $\alpha \in V$. So obviously $f(T)\alpha = 0$ for all $\alpha \in W$. □

Proposition 3.9. *Suppose $V = F[T]\alpha$ is cyclic. Let $f = \min(\alpha) = x^n + a_{n-1}x^{n-1} + \cdots a_0$. For each $i \in \mathbb{N}$, set $\alpha_i := T^i\alpha$. Then*

(1) *$B = (\alpha_0, \ldots, \alpha_{n-1})$ is a basis for V;*
(2) *$\alpha_n = -\sum_{i=0}^{n-1} a_i\alpha_i$.*
(3) *In the basis B we have*

$$T = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

Proof. Suppose $\beta \in V$. We can write $\beta = p(T)\alpha$ for some $p \in F[x]$. Write $p = qf + r$ with $q, r \in F[t]$ and $\deg r < \deg f$. Then $p(T) = r(T)$ since $f(T) = 0$. So we can write $\beta = r(T)\alpha$ with $\deg r < n$. This shows that B spans. To show that B is a basis, suppose $r(T)\alpha = 0$ for some r of degree less than n. Then $f|r$. So $r = 0$. This proves (1).

(2) is obvious because, by definition, $f(T)\alpha = 0$. (3) follows directly from (2). □

Definition 3.10. The matrix in Proposition 3.9 (3) is called the *companion matrix* of $F[T]\alpha$. Note that, if V is cyclic, the companion matrix depends only on the minimal polynomial f of T.

Theorem 3.11. *Suppose $V = F[T]\alpha$ is cyclic. Then* $\min(T)$ *is equal to the characteristic polynomial $\chi(T)$ of T.*

Proof. We just need to show that the characteristic polynomial of the companion matrix is equal to the polynomial $f = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ (as in Proposition 3.9 (3)). For this, write $f = xg + a_0$ and induct on n. It is obvious when $n = 1$; so assume $n > 1$. The characteristic polynomial of T is

$$h := \begin{vmatrix} x & 0 & 0 & \cdots & 0 & a_0 \\ -1 & x & 0 & \cdots & 0 & a_1 \\ 0 & -1 & x & \cdots & 0 & a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & x + a_{n-1} \end{vmatrix} \tag{3.11.1}$$

Let S denote the matrix

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_1 \\ 1 & 0 & 0 & \cdots & 0 & -a_2 \\ 0 & 1 & 0 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

By induction, the characteristic polynomial of S is g. Then, expanding out the determinate in (3.11.1) in minors on the first column, we see that $h = xg + (-1)^{n-1}(-1)^{n-1}a_0 = xg + a_0 = f$. □

4. Canonical Form

Here again V is a finite dimensional vector space over F and $T \in L(V)$.

Lemma 4.1. *Suppose $E_1, E_2 \in L(V)$ satisfy the following*

(1) $E_1 + E_2 = \mathrm{Id}$;
(2) $E_1E_2 = E_2E_1 = 0$.

Then E_1 and E_2 are the projectors onto complementary subspaces.

Proof. We just need to show that $E_i^2 = E_i$ for $i = 1, 2$. But we have $E_1 = E_1(E_1 + E_2) = E_1^2 + E_1E_2 = E_1^2$. □

Theorem 4.2. *Suppose $f := \min(T) = p_1p_2$ with p_1, p_2 relatively prime. Set $W_i = \ker p_i$. Then*

$$V = W_1 \oplus W_2.$$

Moreover both subspaces in the above decomposition are T-stable, and $\min(T, W_i) = p_i$.

Proof. It is obvious that $\ker p_i(T)$ is T-stable $i = 1, 2$.

Pick $a_1, a_2 \in F[x]$ such that $a_1p_1 + a_2p_2 = 1$. Then set $h_i = a_ip_i$, $E_i = h_i(T)$. We have $E_1E_2 = a_1(T)a_2(T)f(T) = 0 = E_2E_1$, and $E_1 + E_2 = \mathrm{Id}$. So E_1 and E_2 are projections onto complementary subspaces. It follows that $V = \ker E_1 \oplus \ker E_2$.

Now, $E_i = a_i(T)p_i(T)$. So $\ker p_i(T) \subset \ker E_i$. On the other hand, suppose $\alpha \in \ker E_1$. Then $p_1(T)\alpha = p_1(T)(E_1 + E_2)\alpha = p_1(T)E_2(T)\alpha = p_1(T)a_2(T)p_2(T)\alpha = a_2(T)f(T)\alpha = 0$. So $\ker p_1(T) = \ker E_1$. And similarly $\ker p_2(T) = \ker E_2$.

It is clear that $p_1(T) = 0$ on W_1. Suppose $g(T) = 0$ on W_1. Then $g(T)p_2(T) = 0$. So $f|gp_2$. So $p_1|g$. □

Lemma 4.3. *Suppose W_1 and W_2 are T-stable subspaces of V with $\min(W_i) = g_i$. Set $W = W_1 + W_2$ and suppose that $\gcd(g_1, g_2) = 1$. Then*

(a) $W_1 \cap W_2 = \{0\}$. *So $W = W_1 \oplus W_2$.*

(b) $\min(W) = g_1g_2$.

Proof. (a): Suppose $w \in W_1 \cap W_2$. Then $\min(w)|g_1$ and $\min(w)|g_2$. So $\min(w)|\gcd(g_1, g_2) = 1$. So $\min(w) = 1$. So $w = 0$.

(b): Suppose $f \in F[x]$. Then $f(T) = 0$ iff $f(T)_{|W_1} = f(T)_{|W_2} = 0$. This happens iff $g_1|f$ and $g_2|f$. Since g_1 and g_2 are relatively prime, this happens iff $g_1g_2|f$. So $\min(V) = g_1g_2$. □

Lemma 4.4. *Suppose $\alpha_1, \alpha_2 \in V$ and set $g_i = \min(\alpha_i)$. If $\gcd(g_1, g_2) = 1$, then $\min(\alpha_1 + \alpha_2) = g_1g_2$.*

Proof. Set $W_i = F[T]\alpha_i$. Then $W_1 \cap W_2 = \{0\}$ since $\gcd(g_1, g_2) = 1$. So, if we set $W = W_1 + W_2$, the sum is direct and $W = W_1 \oplus W_2$. Then, set $\alpha = \alpha_1 + \alpha_2$. For $f \in F[x]$, $f(T)\alpha = 0 \Leftrightarrow f(T)\alpha_1 = f(T)\alpha_2 = 0 \Leftrightarrow g_1|f$ and $g_2|f$. Since g_1, g_2 are relatively prime, this happens iff $g_1g_2|f$. □

Theorem 4.5. *Suppose $\min(V) = f$. Then there is an $\alpha \in V$ with $\min(\alpha) = f$.*

Proof. If $T = 0$ so that $f = 1$, then the result is obvious. So assume that T is non-zero. Write $f = \prod_{i=1}^{r} p_i^{d_i}$ with the p_i distinct irreducibles and $d_i \in \mathbb{P}$, and induct on r.

If $r = 1$, then $f = p_1^{d_1}$. So, for every $\beta \in V$, we have $\min(\beta) = p_1^{e(\beta)}$ for some integer $e(\beta)$ satisfying $0 \le e \le d_1$. Let e be the maximum value of $e(\beta)$ obtained as β ranges over all elements of V. Then $p_1(T)^e = 0$ on V. So $e = d_1$. Therefore $d_1 = e(\alpha)$ for some $\alpha \in V$ and the result follows.

Now assume $r > 1$. Write $f = gp_r^{d_r}$ with $\gcd(g, p_r) = 1$, and write $V = W \oplus W_r$ with $W = \ker g(T)$ and $W_r = \ker p_r^{d_r}(T)$. By induction, can find $\beta \in W$ such that $\min(\beta) = \min(W) = g$. And by the proof for $r = 1$, we can find $\gamma \in W_r$ such that $\min(\gamma) = p_r^{d_r}$. Set $\alpha = \beta + \gamma$. Then $\min(\alpha) = gp_r^r = f$. □

Lemma 4.6. *We have $\mathrm{Ann}(T^*, V^*) = \mathrm{Ann}(T, V)$.*

Proof. For $X, Y \in L(V)$, we have $(X + Y)^* = X^* + Y^*$ and $(XY)^* = Y^*X^*$. Moreover, it is easy to see that $X = 0 \Leftrightarrow X^* = 0$. So $p(T)^* = p(T^*)$ for $p \in F[x]$. Therefore, $p(T) = 0 \Leftrightarrow p(T)^* = 0 \Leftrightarrow p(T^*) = 0$. □

Lemma 4.7. *Suppose $W \subset V^*$ is a T^*-stable subspace. Then $W^\perp \subset V$ is T-stable.*

Proof. Suppose $v \in W^\perp$ and $\lambda \in W$. Then $\langle \lambda, Tv \rangle = \langle T^*\lambda, v \rangle = 0$. So $Tv \in W^\perp$. □

Theorem 4.8. *Suppose $\alpha \in V$ is an element with $f = \min(\alpha) = \min(T, V)$. Then there exists a T-stable subspace K such that*

$$V = (F[T]\alpha) \oplus K.$$

Proof. Set $H = F[T]\alpha$. Then $\min(T^*, H^*) = \min(T, H) = f$. So there exists a $\bar{\lambda} \in H^*$ such that $\min(\bar{\lambda}) = f$. Find a $\lambda \in V^*$ such that $\lambda_{|H} = \bar{\lambda}$. (Exercise 3.5.12 in Hoffman-Kunze shows that we can do this.) Then $p(T)\lambda = 0 \Rightarrow p(T)\bar{\lambda} = 0$. So $f|\min(\lambda)$. But $\min(T^*, V^*) = \min(T, V) = f$. So we have $f = \min(\lambda)$. Now set $K = (F[T]\lambda)^\perp$. By Lemma 4.7 K is T-stable.

If $p(T)\alpha \in K$ then $\langle q(T^*)\lambda, p(T)\alpha \rangle = \langle p(T^*)\lambda, q(T)\alpha \rangle = 0$ for all $q \in F[T]$. So $p(T^*)\bar{\lambda} = 0$. So $f|p$. So $p(T)\alpha = 0$. This shows that $H \cap K = \{0\}$. On the other hand, $\dim F[T]\lambda = \deg f = \dim H$. So $\dim K = \dim V - \dim H$. So, by the Hausdorff dimension formula it follows that $V = H \oplus K$. □

Theorem 4.9 (Cayley-Hamilton)**.** *Suppose $T \in L(V)$ with $\dim V < \infty$. Then the minimal polynomial of T divides the characteristic polynomial.*

Proof. Pick $\alpha \in V$ with $f = \min(\alpha) = \min(T, V)$. Then write $V = H \oplus K$ with $H := F[T]\alpha$. We have $\chi(T) = \chi(T_{|H})\chi(T_{|K}) = f\chi(T_{|K})$. □

Theorem 4.10 (Rational canonical form)**.** *We can find elements $\alpha_1, \dots, \alpha_r$ of V such that*

(1) $V = F[T]\alpha_1 \oplus \cdots \oplus F[T]\alpha_r$;
(2) $\min(\alpha_i) | \min(\alpha_{i-1})$ *for* $i = 2, \dots, r$.

Moreover, the polynomials $\min(\alpha_i)$ *are unique. They are called the* **elementary divisors** *of T.*

Proof. Apply Theorem 4.8 inductively to prove the existence.

For uniqueness, suppose $V, T \in L(V)$ is a pair where the polynomials associated to the decomposition above are not unique and assume that V has minimal dimension for this property. So we have

$$\begin{aligned} V &= F[T]\alpha_1 \oplus \cdots \oplus F[T]\alpha_r \\ &= F[T]\beta_1 \oplus \cdots \oplus F[T]\beta_s \end{aligned}$$

with $\min(\alpha_i) | \min(\alpha_{i-1})$ and $\min(\beta_i) | \min(\beta_{i-1}$ for $i \geq 2$. Set $f_i = \min(\alpha_i)$, $g_i = \min(\beta_i)$. Then $f_1 = \mathrm{Ann}(T, V) = g_1$.

Suppose $p_i = q_i$ for $i = 1, \dots, j-1$, but $p_j \neq q_j$. By switching the p's and q's we can assume that $\deg p_j \leq \deg q_j$. Then

$$\begin{aligned} p_j(T)V &= F[T]p_j(T)\alpha_1 \oplus \cdots \oplus F[T]p_j(T)\alpha_{j-1} \\ &= \left(F[T]p_j(T)\beta_1 \oplus \cdots \oplus F[T]p_j(T)\beta_{j-1}\right) \oplus F[T]p_j(T)\beta_j \oplus \cdots \oplus p_j(T)F[T]\beta_s. \end{aligned}$$

We have $\min(p_j(T)\alpha_i) = \min(p_j(T)\beta_i) = p_i/p_j$ for $i < j$. So, for $i < j$, $\dim F[T]p_j(T)\alpha_i = \dim F[T]p_j(T)\beta_i$. Therefore

$$F[T]p_j(T)\beta_j \oplus \cdots \oplus p_j(T)F[T]\beta_s = 0.$$

So, since $p_j(T)\beta_j = 0$, $q_j | p_j$. But this implies that $q_j = p_j$ since $\deg q_j \geq \deg p_j$. Contradiction. □

Proposition 4.11. *The product of the elementary divisors is the characteristic polynomial of T.*

Proof. Obvious. Since $\chi(T|F[T]\alpha_i) = \min(T|F[T]\alpha_i)$. □

Definition 4.12. We say T is *nilpotent* if $T^k = 0$ for some $k \in \mathbb{N}$. Then $\min(T) = x^n$ where n is the smallest positive integer such that $T^n = 0$. We say that n is the *nilpotence index of T*.

Proposition 4.13. *Suppose T is nilpotent and V is cyclic. Then the companion matrix of T is*

$$T = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

Proof. Obvious. □

It follows from the proposition that any nilpotent T is similar to a matrix having blocks of the form in Proposition 4.13 with size less than or equal to the nilpotence index.

5. Jordan form

Here V is a finite dimensional vector space over a field F which we assume to be algebraically closed. We fix $T \in L(V)$, and let $f = \min(T)$. We assume that $T \neq 0$ so that $f \neq 1$. Since F is algebebraically closed we can factor

$$f = \prod_{i=1}^{r} (x - a_i)^{d_i}$$

where $a_i \in F$ and $d_i, r \in \mathbb{P}$.

Theorem 5.1. *Set $W_i = \ker(T - a_i)^{d_i}$ for $i = 1, \dots r$. Then*

$$V = W_1 \oplus \cdots \oplus W_r.$$

Proof. Induct on r. If $r = 1$ it is obvious. Otherwise write $f = g(x - a_r)^{d_r}$. We get

$$V = W \oplus W_r$$

where $W = \ker g(T)$, and $\min(T|W) = g$. Now apply induction. □

Corollary 5.2. *We can find a basis for V such that T is a block matrix with $\ell \times \ell$ blocks T_{ij}, $i = 1, \dots, r$, $j = 1, \dots, m(i)$ for some $m(i) \in \mathbb{P}$, of the following form.*

$$T_{ij} = \begin{pmatrix} a_i & 0 & 0 & \cdots & 0 & 0 \\ 1 & a_i & 0 & \cdots & 0 & 0 \\ 0 & 1 & a_i & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_i \end{pmatrix}.$$

Moreover, if the blocks T_{ij} are of size λ_{ij}, then $\lambda_{i1} + \cdots + \lambda_{im(i)} = d_i$.

Proof. Set $T_i = T_{|W_i}$. Then set $N_i = T_i - a_i\mathrm{Id}$. Since $(T - a_i)^{d_i} = 0$, N_i is nilpotent. So we can write N_i as a sum of blocks as in Proposition 4.13. Then $T_i = a_i\mathrm{Id} + N_i$, and the result follows. We say a matrix in the above form is in *Jordan canonical form*. □

Example 5.3. Suppose we have

$$T = \begin{pmatrix} 3 & 1 & 0 \\ -1 & 1 & 0 \\ 3 & 2 & 2 \end{pmatrix}.$$

Let's compute the Jordan form of T. The first step is to compute the characteristic polynomial. This is $(x - 2)^3$. One way to do this is to just multiply it out. There is also a clever way to do it using row and column operations.

Now, set

$$N = T - 2 = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 3 & 2 & 0 \end{pmatrix}.$$

We have

$$N^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, N^3 = 0.$$

So the minimal polynomial of N is x^3 and therefore the minimal polynomial of T is $(x-2)^3$. It follows that T has Jordan form

$$J := \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

Note that we did not have to compute a basis B for which $[T]_B$ has this form to do this computation.

But suppose we want such a basis. It's not hard. We just need to find a vector α such that $F[N]\alpha = V$. In other words, we need a vector not in the kernel of N^2. So e_1 will do. Set $B = (e_1, Ne_1, N^2e_1) = (e_1, e_1 - e_2 + e_3, e_3)$. Then $[T]_B = J$.

Now, let's give the clever way compute χT using row and column operations. Set

$$S = \begin{pmatrix} x-3 & -1 & 0 \\ 1 & x-1 & 0 \\ 3 & -2 & x-2 \end{pmatrix}.$$

The idea is to do row and column operations (including permuting rows and permuting columns) in such a way as to reduce S to a diagonal matrix. Keeping track of the sign of the determinant. Suppose R_1, R_2 and R_3 are the rows and p is a polynomial. Then the matrix with rows $R_1, R_2 + pR_1, R_3$ has the same determinant as S. This allows us to reduce the degrees in the the matrix using long division.

So here's the computation:

$$\begin{aligned}
|S| &= -\begin{vmatrix} 1 & x-1 & 0 \\ x-3 & -1 & 0 \\ 3 & -2 & x-2 \end{vmatrix} = -\begin{vmatrix} 1 & x-1 & 0 \\ 0 & -1-(x-1)(x-3) & 0 \\ 0 & -2-3(x-1) & x-2 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & 0 \\ 0 & x^2-4x+4 & 0 \\ 0 & -3x+1 & x-2 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 0 \\ 0 & -3x+1 & x-2 \\ 0 & (x-2)^2 & 0 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & 0 \\ 0 & x-2 & -3x+1 \\ 0 & 0 & x^2-4x+4 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & x-2 & -5 \\ 0 & 0 & (x-2)^2 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & 0 \\ 0 & 5 & x-2 \\ 0 & -(x-2)^2 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 5 & x-2 \\ 0 & 0 & (x-2)^3/5 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (x-2)^3 \end{vmatrix}.
\end{aligned}$$

So $|S| = (x-2)^3$. Note that we only really had to do the first three parts of the computation to get down to a lower triangular matrix. However, I wanted to do this computation to the end so that I could state the following fact.

Theorem 5.4. *Suppose $T \in L(V)$ and let $S = x\mathrm{Id} - T$. Using row and column operations as above, we can reduce S to a diagonal matrix of the form*

$$\begin{pmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & p_n \end{pmatrix} \tag{5.4.1}$$

where the p_i are monic polynomils and $p_1|p_2|\cdots p_n$. Then the p_i are the elementary divisors of T (in reverse order with 1's omitted).

Proof. See §7.4 of Hoffman and Kunze. □

Example 5.5. What is the Jordan canonical form of the matrix

$$N = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$$

with $a, b \in \mathbb{C}$? The easiest way to figure this out is to note that $Ne_3 = be_1 + ce_2$, $N^2 e_3 = ace_1$ and $N^3 e_3 = 0$. So the minimal polynomial divides x^3. If a, c are both non-zero, then $V = \mathbb{C}[N]e_3$. So, using the ordered basis e_3, Ne_3, N^2e_3, we see that

$$N \sim \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{5.5.1}$$

If $a = b = c = 0$, then clearly N is the 0-matrix. Otherwise, N has Jordan form

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{5.5.2}$$

and elementary divisors x^2 and x. This is easy to see abstractly because the list of elementary divisors of N is not x^3 and not x, x, x. So the only thing it can be is x^2, x. And this shows that the Jordan form must be as above.

However, to see it explicitly it helps to work in cases. Let's do the case where $a = 0$ but $c \neq 0$ explicitly using the method of proof in Theorem 4.8. This is certainly not the easiest way to do the computation. But it may help in understanding how the proof of Theorem 4.8 works. Set $H = \mathbb{C}[N]e_3 = \langle, e_3, be_1 + ce_2\rangle$. Let e_1^*, e_2^*, e_3^* denote the dual basis the the standard basis e_1, e_2, e_3. Set $\lambda = e_2^*$ and let $\bar{\lambda}$ denote the restriction of λ to H. Then $(N^*\bar{\lambda}, e_3) = (\bar{\lambda}, Ne_3) = (\bar{\lambda}, be_1 + ce_2) = (e_2^*, be_1 + ce_2) = c$. And, since $N^2 = 0$, $(N^*)^2 = 0$ as well. So $\mathbb{C}[N^*]\bar{\lambda} = H^*$. So, following the argument in the proof of Theorem 4.8, set $K := (\mathbb{C}[N^*]\lambda)^\perp$. We have $Ne_2^* = ce_3^*$. So $K = \langle e_2^*, e_3^*\rangle^\perp = \langle e_1\rangle$. We get $V = H \oplus K = \langle e_3, be_1 + ce_2\rangle \oplus \langle e_1\rangle$. for the Jordan blocks. In other words, with respect to the ordered basis $B = (e_3, be_1 + ce_2, e_1)$, N has the matrix in (5.5.2).

Example 5.6. Set

$$T = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}.$$

Let's find the elementary divisors of T both by hand and using the method in Theorem 5.4. First we do the row and column operations needed. We write $A \approx B$ if A and B differ by an elementary row or column operation. (Not to be confused with $\sim$ for similarity of matrices. But hopefully this will be clear from the context.) We have

$$\begin{aligned} S &= \begin{pmatrix} x-4 & 0 & -1 \\ -2 & -3 & -2 \\ -1 & 0 & -4 \end{pmatrix} \approx \begin{pmatrix} -1 & 0 & -4 \\ x-4 & 0 & -1 \\ -2 & x-3 & -2 \end{pmatrix} \\ &\approx \begin{pmatrix} -1 & 0 & 0 \\ x-4 & 0 & (x-4)^2-1 \\ -2 & x-3 & -2x+6 \end{pmatrix} \approx \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & x^2-8x+15 \\ 0 & x-3 & -2x+6 \end{pmatrix} \\ &\approx \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-3 & -2x+6 \\ 0 & 0 & (x-3)(x-5) \end{pmatrix} \approx \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-3 & 0 \\ 0 & 0 & (x-3)(x-5) \end{pmatrix}. \end{aligned}$$

This tell us that the elementary divisors are $(x-3)$ and $(x-3)(x-5)$. It follows that T has diagonal Jordan form. That is

$$T \sim \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

But, since we haven't prove Theorem 5.4 yet. Maybe we don't trust it. So let's find the Jordan form another way. The above computation tells us at least that $\chi(T) = (x-3)^2(x-5)$. So we have

$$T - 5 = \begin{pmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{pmatrix}.$$

This has row-echelon form

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

So $(1, 2, 1)$ generates the eigenspace of T with eigenvalue 5.

Now

$$T - 3 = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{pmatrix} = (1/2)(T-3)^2.$$

This row reduces to

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This shows us that that vectors $(-1, 0, 1)$ and $(0, 1, 0)$ generate the eigenspace with eigenvalue 3. So again we get that T is diagonalizable into the form above.

Department of Mathematics, 1301 Mathematics Building, University of Maryland, College Park, MD 20742

E-mail address: pbrosnan@umd.edu