The uniform normal form of a linear mapping

Richard Cushman

Dedicated to my friend and colleague Arjeh Cohen on his retirement

Let $V$ be a finite dimensional vector space over a field $k$ of characteristic 0. Let $A : V \to V$ be a linear mapping of $V$ into itself with characteristic polynomial $\chi_A$. The goal of this paper is to give a normal form for $A$, which yields a better description of its structure than the classical companion matrix. This normal form does not use a factorization of $\chi_A$ and requires only operations in the field $k$ to compute.

1 Semisimple linear mappings

We begin by giving a well known criterion to determine if the linear mapping $A$ is semisimple, that is, every $A$-invariant subspace of $V$ has an $A$-invariant complementary subspace.

Suppose that we can factor $\chi_A$, that is, find monic irreducible polynomials $\{\pi_i\}_{i=1}^m$, which are pairwise relatively prime, such that $\chi_A = \prod_{i=1}^m \pi_i^{n_i}$, where $n_i \in \mathbb{Z}_{\geq 1}$. Then

$$\chi'_A = \sum_{j=1}^m (n_j \pi_j^{n_j-1}) \prod_{i \neq j} \pi_i^{n_i} = (\prod_{\ell=1}^m \pi_\ell^{n_\ell-1}) \left( \sum_{j=1}^m (n_j \pi'_j) \prod_{i \neq j} \pi_i \right).$$

Therefore the greatest common divisor $\chi_A$ and its derivative $\chi'_A$ is the polynomial $d = \prod_{\ell=1}^m \pi_\ell^{n_\ell-1}$. The polynomial $d$ can be computed using the Euclidean algorithm. Thus the square free factorization of $\chi_A$ is the polynomial $p = \prod_{\ell=1}^m \pi_\ell = \chi_A/d$, which can be computed without knowing a factorization of $\chi_A$.

The goal of the next discussion is to prove

Claim 1.1 The linear mapping $A : V \to V$ is semisimple if $p(A) = 0$ on $V$.  

\footnotetext{Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada, T2N 1N4.}
Let $p = \prod_{j=1}^{m} \pi_j$ be the square free factorization of the characteristic polynomial $\chi_A$ of $A$. We now decompose $V$ into $A$-invariant subspaces. For each $1 \leq j \leq m$ let $V_j = \{ v \in V \mid \pi_j(A)v = 0 \}$. Then $V_j$ is an $A$-invariant subspace of $V$. For if $v \in V_j$, then $\pi_j(A)Av = A\pi_j(A)v = 0$, that is, $Av \in V_j$.

The following argument shows that $V = \bigoplus_{j=1}^{m} V_j$. Because for $1 \leq j \leq m$ the polynomials $\prod_{i \neq j} \pi_i$ are pairwise relatively prime, there are polynomials $f_j$, $1 \leq j \leq m$ such that $1 = \sum_{j=1}^{m} f_j \left( \prod_{i \neq j} \pi_i \right)$. Therefore every vector $v \in V$ can be written as

$$v = \sum_{j=1}^{m} f_j(A) \left( \prod_{i \neq j} \pi_i(A)v \right) = \sum_{j=1}^{m} f_j(A)v_j.$$ 

Since $\pi_j(A)\left( \prod_{i \neq j} \pi_i(A)v \right) = p(A)v = 0$, the vector $v_j \in V_j$. Therefore $V = \sum_{j=1}^{m} V_j$. If for $i \neq j$ we have $w \in V_i \cap V_j$, then for some polynomials $F_i$ and $G_j$ we have $1 = F_i \pi_i + G_j \pi_j$, because $\pi_i$ and $\pi_j$ are relatively prime. Consequently, $w = F_i(A)\pi_i(A)w + G_j(A)\pi_j(A)w = 0$. So $V = \sum_{j=1}^{m} \oplus V_j$.

We now prove

**Lemma 1.2** For each $1 \leq j \leq m$ there is a basis of the $A$-invariant subspace $V_j$ such that that matrix of $A$ is block diagonal.

**Proof.** Let $W$ be a minimal dimensional proper $A$-invariant subspace of $V_j$ and let $w$ be a nonzero vector in $W$. Then there is a minimal positive integer $r$ such that $A^r w \in \text{span}_k \{ w, Aw, \ldots, A^{r-1}w \} = U$. We assert: the vectors $\{ A^i w \}_{i=0}^{r-1}$ are linearly independent. Suppose that there are $a_i \in k$ for $1 \leq i \leq r-1$ such that $0 = a_0 w + a_1 Aw + \cdots + a_{r-1}A^{r-1}w$. Let $t \leq r-1$ be the largest index such that $a_t \neq 0$. So $A^t w = -\frac{a_{t-1}}{a_t} A^{t-1}w - \cdots - \frac{a_0}{a_t} w$, that is, $A^t w \in \text{span}_k \{ w, \ldots, A^{t-1}w \}$ and $t < r$. This contradicts the definition of the integer $r$. Thus the index $t$ does not exist. Hence $a_i = 0$ for every $0 \leq i \leq r-1$, that is, the vectors $\{ A^i w \}_{i=0}^{r-1}$ are linearly independent.

The subspace $U$ of $W$ is $A$-invariant, for

$$A(\sum_{j=0}^{r-1} b_j A^j w) = \sum_{j=0}^{r-2} b_j A^{j+1} w + b_{r-1} A^r w, \quad \text{where } b_j \in k$$

$$= \sum_{j=1}^{r-1} b_j A^j w + b_{r-1} \left( \sum_{\ell=0}^{r-1} a_{\ell} A^{\ell} w \right), \quad \text{since } A^r w \in U$$

$$= b_{r-1} a_0 w + \sum_{j=1}^{r-1} \left( b_{j-1} + b_{r-1} a_j \right) A^j w \in U.$$
Next we show that there is a monic polynomial \( \mu \) of degree \( r \) such that \( \mu(A) = 0 \) on \( U \). With respect the basis \( \{A^iw\}_{i=0}^{r-1} \) of \( U \) we can write \( A^r w = -a_0 w - \cdots - a_{r-1}A^{r-1}w \). So \( \mu(A)w = 0 \), where

\[
\mu(\lambda) = a_0 + a_1 \lambda + \cdots + a_{r-1} \lambda^{r-1} + \lambda^r.
\]

Since \( \mu(A)A^iw = A^i(\mu(A)w) = 0 \) for every \( 0 \leq i \leq r - 1 \), it follows that \( \mu(A) = 0 \) on \( U \).

By the minimality of the dimension of \( W \) the subspace \( U \) cannot be proper. But \( U \neq \{0\} \), since \( w \in U \). Therefore \( U = W \). Since \( U \subseteq V_j \), we obtain \( \pi_j(A)u = 0 \) for every \( u \in U \). Because \( \pi_j \) is irreducible, the preceding statement shows that \( \pi_j \) is the minimum polynomial of \( A \) on \( U \). Thus \( \pi_j \) divides \( \mu \). Suppose that \( \deg \pi_j = s < \deg \mu = r \). Then \( A^su' \in \text{span}_k \{w', \ldots, A^{s-1}w'\} = Y \) for some nonzero vector \( w' \) in \( U \). By minimality, \( Y = U \). But \( \dim Y = s < \dim U = r \), which is a contradiction. Thus \( \pi_j = \mu \).

Note that the matrix of \( A|U \) with respect to the basis \( \{A^iw\}_{i=0}^{r-1} \) is the \( r \times r \) companion matrix

\[
C_r = \begin{pmatrix}
0 & \cdots & \cdots & 0 & -a_0 \\
1 & 0 & \cdots & 0 & -a_1 \\
\vdots & 1 & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & 0 & -a_{r-2} \\
0 & \cdots & \cdots & 1 & -a_{r-1}
\end{pmatrix},
\]

where \( \pi_j = a_0 + a_1 \lambda + \cdots + a_{r-1} \lambda^{r-1} + \lambda^r \).

Suppose that \( U \neq V_j \). Then there is a nonzero vector \( w' \in V_j \setminus U \). Let \( r' \) be the smallest positive integer such that \( A^{r'} w' \in \text{span}_k \{w', Aw', \ldots, A^{r'-1}w'\} = U' \). Then by the argument in the preceding paragraph, \( U' \) is a minimal \( A \)-invariant subspace of \( V_j \) of dimension \( r' = r \), whose minimal polynomial is \( \pi_j \). Suppose that \( U' \cap U \neq \{0\} \). Then \( U' \cap U \) is a proper \( A \)-invariant subspace of \( U' \). By minimality \( U' \cap U = U' \), that is, \( U \subseteq U' \).

But \( r = \dim U = \dim U' = r' \). So \( U = U' \). Thus \( w' \in U' \) and \( w' \notin U \), which is a contradiction. Therefore \( U' \cap U = \{0\} \). If \( U \oplus U' \neq V_j \), we repeat the above argument. Using \( U \oplus U' \) instead of \( U \), after a finite number of repetitions we have \( V_j = \sum_{i=1}^{\ell} \oplus U_i \), where for every \( 0 \leq i \leq \ell \) the subspace \( U_i \) of \( V_j \) is \( A \)-invariant with basis \( \{A^ku_i\}_{k=0}^{r-1} \) and the minimal polynomial of \( A|U_i \) is \( \pi_j \). With respect to the basis \( \{A^ku_i\}_{(i,k)=(1,0)}^{(\ell,r-1)} \) of \( V_j \) the matrix of \( A \) is \( \text{diag}(C_r, \ldots, C_r) \), which is block diagonal.

\[\Box\]

For each \( 1 \leq j \leq m \) applying lemma 1.2 to \( V_j \) and using the fact that \( V = \sum_{j=1}^{m} \oplus V_j \) we obtain
Corollary 1.3 There is a basis of $V$ such that the matrix of $A$ is block diagonal.

Proof of claim 1.1 Suppose that $U$ is an $A$-invariant subspace $V$. Then by corollary 1.3, there is a basis $\varepsilon_U$ of $U$ such that the matrix of $A|U$ is block diagonal. By corollary 1.3 there is a basis $\varepsilon_V$ of $V$ which extends the basis $\varepsilon_U$ such that the matrix of $A$ on $V$ is block diagonal. Let $W$ be the subspace of $V$ with basis $\varepsilon_W = \varepsilon_V \setminus \varepsilon_U$. The matrix of $A|W$ is block diagonal. Therefore $W$ is $A$-invariant and $V = U \oplus W$ by construction. Consequently, $A$ is semisimple. \qed

2 The Jordan decomposition of $A$

Here we give an algorithm for finding the Jordan decomposition of the linear mapping $A$, that is, we find real semisimple and commuting nilpotent linear maps $S$ and $N$ whose sum is $A$. The algorithm we present uses only the characteristic polynomial $\chi_A$ of $A$ and does \textit{not} require that we know \textit{any} of its factors. Our argument follows that of [1].

Let $p$ be the square free factorization of $\chi_A$. Let $M$ be the smallest positive integer such that $\chi_A$ divides $p^M$. Then $M \leq \deg \chi_A$. Assume that $\deg \chi_A \geq 2$, for otherwise $S = A$. Write

$$S = A + \sum_{j=1}^{M-1} r_j(A)p(A)^j,$$ \hfill (3)

where $r_j$ is a polynomial whose degree is less than the degree of $p$. From the fact that $\chi_A$ divides $p^M$, it follows that $p(A)^M = 0$.

We want to determine $S$ in the form (3) so that

$$p(S) = 0.$$ \hfill (4)

From claim 1.1 it follows that $S$ is semisimple.

We have to find the polynomials $r_j$ in (3) so that equation (4) holds. We begin by using the Taylor expansion of $p$. If (3) holds, then

$$p(S) = p\left(A + \sum_{j=1}^{M-1} r_j(A)p(A)^j\right)$$

$$= p(A) + \sum_{i=1}^{M-1} p^{(i)}(A)\left(\sum_{j=1}^{M-1} r_j(A)p(A)^j\right)^i,$$
where $p^{(i)}$ is $\frac{1}{n}$ times the $i$th derivative of $p$

$$= p(A) + \sum_{i=1}^{M-1} \sum_{k=1}^{M-1} c_{k,i} p(A)^k p^{(i)}(A). \tag{5}$$

Here $c_{k,i}$ is the coefficient of $z^k$ in $(r_1 z + \cdots + r_{M-1}z^{M-1})^i$. Note that $c_{k,i} = 0$ if $k > i$. A calculation shows that when $k \leq i$ we have

$$c_{k,i} = \sum_{\alpha_1 + \cdots + \alpha_k = i \atop \alpha_1 + 2\alpha_2 + \cdots + (k-1)\alpha_k = k} \frac{i!}{\alpha_1! \cdots \alpha_{k-1}!} r_1^{\alpha_1} \cdots r_{k-1}^{\alpha_{k-1}}. \tag{6}$$

Interchanging the order of summation in (5) we get

$$p(S) = p(A) + \sum_{i=1}^{M-1} \left( r_i(A)p^{(1)}(A) + e_i(A) \right)p(A)^i,$$

where $e_1 = 0$ and for $i \geq 2$ we have $e_i = \sum_{j=1}^i c_{i,j} p^{(j)}$. Note that $e_i$ depends on $r_1, \ldots, r_{i-1}$, because of (6).

Suppose that we can find polynomials $r_i$ and $b_i$ such that

$$r_i p^{(1)} + e_i = b_i p - b_{i-1}, \tag{7}$$

for every $1 \leq i \leq M - 1$. Here $b_0 = 1$. Then

$$\sum_{i=1}^{M-1} \left( r_i(A)p^{(1)}(A)+e_i(A) \right)p(A)^i = \sum_{i=1}^{M-1} \left( b_i(A)p(A)-b_{i-1}(A) \right)p(A)^i = -p(A),$$

since $p^M(A) = 0$ and $b_0 = 1$, which implies $p(S) = 0$, see (5).

We now construct polynomials $r_i$ and $b_i$ so that (7) holds. We do this by induction. Since the polynomials $p$ and $p^{(1)}$ have no common nonconstant factors, their greatest common divisor is the constant polynomial 1. Therefore by the Euclidean algorithm there are polynomials $g$ and $h$ with the degree of $h$ being less than the degree of $p$ such that

$$gp - hp^{(1)} = 1. \tag{8}$$

Let $r_1 = h$, and $b_1 = g$. Using the fact that $b_0 = 1$ and $e_1 = 0$, we see that equation (8) is the same as equation (7) when $i = 1$. Let $d_1 = 0$ and $q_0 = q_1 = 0$. Now suppose that $n \geq 2$. By induction suppose that the polynomials
\( r_1, \ldots, r_{n-1}, e_1, \ldots, e_{n-1}, q_1, \ldots, q_{n-1} \) and \( b_1, \ldots, b_{n-1} \) are known and that \( r_i \) and \( b_i \) satisfy (7) for every \( 1 \leq i \leq n-1 \). Using the fact that the polynomials \( r_1, \ldots, r_{n-1} \) are known, from formula (6) we can calculate the polynomial \( e_n = \sum_{j=2}^{n} c_{i,n} p^{(j)} \). For \( n \geq 2 \) define the polynomial \( d_n \) by

\[
d_n = q_{n-1} + h \sum_{i=1}^{n} g^{n-i} e_i.
\]

Note that the polynomials \( q_{n-1}, g = b_1, h = r_1, \) and \( e_i \) for \( 1 \leq i \leq n-1 \) are already known by the induction hypothesis. Thus the right hand side of (9) is known and hence so is \( d_n \). Now define the polynomials \( q_n \) and \( r_n \) by dividing \( d_n \) by \( p \) with remainder, namely

\[
d_n = q_n p + r_n.
\]

Clearly, \( q_n \) and \( r_n \) are now known. Next for \( n \geq 2 \) define the polynomial \( b_n \) by

\[
b_n = -p^{(1)} q_n + g \sum_{i=1}^{n} g^{n-i} e_i.
\]

Since the polynomials \( p^{(1)}, q_n, g = b_1, \) and \( e_i \) for \( 1 \leq i \leq n \) are known, the polynomial \( b_n \) is known. We now show that equation (7) holds.

**Proof.** We have already checked that (7) holds when \( n = 1 \). By induction we assumed that it holds for every \( 1 \leq i \leq n-1 \). Using the definition of \( b_n \) (11) and the induction hypothesis we compute

\[
b_n p - b_{n-1} = \left[ -p^{(1)} pq_n + pg \sum_{i=1}^{n} g^{n-i} e_i \right] - \left[ -p^{(1)} q_{n-1} + g \sum_{i=1}^{n-1} g^{n-1-i} e_i \right]
\]

\[
= -p^{(1)} (q_n p - q_{n-1}) + pg \sum_{i=1}^{n} g^{n-i} e_i - \sum_{i=1}^{n-1} g^{n-i} e_i
\]

\[
= -p^{(1)} (-r_n + d_n - q_{n-1}) + (hp^{(1)} + 1) \sum_{i=1}^{n} g^{n-i} e_i - \sum_{i=1}^{n-1} g^{n-i} e_i,
\]

using (8) and (10)

\[
= p^{(1)} r_n - hp^{(1)} \sum_{i=1}^{n} g^{n-i} e_i + hp^{(1)} \sum_{i=1}^{n} g^{n-i} e_i + \sum_{i=1}^{n} g^{n-i} e_i - \sum_{i=1}^{n-1} g^{n-i} e_i,
\]

using (9)

\[
= p^{(1)} r_n + e_n.
\]

\[\square\]
This completes the construction of the polynomial \( r_n \) in (3). Repeating this construction until \( n = M - 1 \) we have determined the semisimple part \( S \) of \( A \). The commuting nilpotent part of \( A \) is \( N = A - S \).

\[\square\]

3 Uniform normal form

In this section we give a description of the uniform normal form of a linear map \( A \) of \( V \) into itself. We assume that the Jordan decomposition of \( A \) into its commuting semisimple and nilpotent summands \( S \) and \( N \), respectively, is known.

3.1 Nilpotent normal form

In this subsection we find the Jordan normal form for a nilpotent linear transformation \( N \).

Recall that a linear transformation \( N : V \to V \) is said to be nilpotent of index \( n \) if there is an integer \( n \geq 1 \) such that \( N^{n-1} \neq 0 \) but \( N^n = 0 \). Note that the index of nilpotency \( n \) need not be equal to \( \dim V \). Suppose that for some positive integer \( \ell \geq 1 \) there is a nonzero vector \( v \), which lies in \( \ker N^{\ell} \setminus \ker N^{\ell-1} \). The set of vectors \( \{v, Nv, \ldots, N^{\ell-1}v\} \) is a Jordan chain of length \( \ell \) with generating vector \( v \). The space \( V^\ell \) spanned by the vectors in a given Jordan chain of length \( \ell \) is a \( N \)-cyclic subspace of \( V \). Because \( N^\ell v = 0 \), the subspace \( V^\ell \) is \( N \)-invariant. Since \( \ker N|V^\ell = \ker N|V^\ell = \text{span}_k\{N^{\ell-1}v\} \), the mapping \( N|V^\ell \) has exactly one eigenvector corresponding to the eigenvalue 0.

**Claim 3.1.1** The vectors in a Jordan chain are linearly independent.

**Proof.** Suppose not. Then \( 0 = \sum_{i=0}^{\ell-1} \alpha_i N^i v \), where not every \( \alpha_i \in k \) is zero. Let \( i_0 \) be the smallest index for which \( \alpha_{i_0} \neq 0 \). Then

\[ 0 = \alpha_{i_0} N^{i_0}v + \cdots + \alpha_{\ell-1} N^{\ell-1}v. \quad (12) \]

Applying \( N^{\ell-1-i_0} \) to both sides of (12) gives \( 0 = \alpha_{i_0} N^{\ell-1}v \). By hypothesis \( v \not\in \ker N^{\ell-1} \), that is, \( N^{\ell-1}v \neq 0 \). Hence \( \alpha_{i_0} = 0 \). This contradicts the definition of the index \( i_0 \). Therefore \( \alpha_i = 0 \) for every \( 0 \leq i \leq \ell - 1 \). Thus the vectors \( \{N^i v\}_{i=0}^{\ell-1} \), which span the Jordan chain \( V^\ell \), are linearly independent.

\[\square\]

With respect to the standard basis \( \{N^{\ell-1}v, N^{\ell-2}v, \ldots, Nv, v\} \) of \( V^\ell \) the
matrix of \( N|V^\ell \) is the \( \ell \times \ell \) matrix
\[
\begin{pmatrix}
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & \ldots & \ldots & 1 \\
\end{pmatrix},
\]
which is a Jordan block of size \( \ell \).

We want to show that \( V \) can be decomposed into a direct sum of \( N \)-cyclic subspaces. In fact, we show that there is a basis of \( V \), whose elements are given by a dark dot \( \bullet \) or an open dot \( \circ \) in the diagram below such that the arrows give the action of \( N \) on the basis vectors. Such a diagram is called the Young diagram of \( N \).

![Young diagram of N](https://via.placeholder.com/150)

Figure 3.1.1. The Young diagram of \( N \).

Note that the columns of the Young diagram of \( N \) are Jordan chains with generating vector given by an open dot. The black dots form a basis for the image of \( N \), whereas the open dots form a basis for a complementary subspace in \( V \). The dots on or above the \( \ell^{th} \) row form a basis for \( \ker N^\ell \) and the black dots in the first row form a basis for \( \ker N \cap \im N \). Let \( r_\ell \) be the number of dots in the \( \ell^{th} \) row. Then \( r_\ell = \dim \ker N^\ell - \dim \ker N^{\ell-1} \). Thus the Young diagram of \( N \) is unique.

Claim 3.1.2 There is a basis of \( V \) that realizes the Young diagram of \( N \).

Proof. Our proof follows that of Hartl [2]. We use induction of the dimension of \( V \). Since \( \dim \ker N > 0 \), it follows that \( \dim \im N < \dim V \). Thus by
the induction hypothesis, we may suppose that \( \text{im} N \) has a basis which is the union of \( p \) Jordan chains \( \{w_i, Nw_i, \ldots, N^{m_i}w_i\} \) each of length \( m_i \). The vectors \( \{N^{m_i}w_i\}_{i=1}^p \) lie in \( \text{im} N \cap \ker N \) and in fact form a basis of this subspace. Since \( \ker N \) may be larger than \( \text{im} N \cap \ker N \), choose vectors \( \{y_1, \ldots, y_q\} \) where \( q \) is a nonnegative integer such that \( \{N^{m_i}w_1, \ldots, N^{m_p}w_p, y_1, \ldots, y_q\} \) form a basis of \( \ker N \).

Since \( w_i \in \text{im} N \) there is a vector \( v_i \) in \( V \) such that \( w_i = Nv_i \). We assert that the \( p \) Jordan chains

\[
\{v_i, Nv_i, \ldots, N^{m_i+1}v_i\} = \{v_i, w_i, Nw_i, \ldots, N^{m_i}w_i\}
\]

each of length \( m_i + 2 \) together with the \( q \) vectors \( \{y_j\} \), which are Jordan chains of length 1, form a basis of \( V \). To see that they span \( V \), let \( v \in V \). Then \( Nv \in \text{im} N \). Using the basis of \( \text{im} N \) given by the induction hypothesis, we may write

\[
Nv = \sum_{i=1}^p \sum_{\ell=0}^{m_i} \alpha_{i\ell} N^\ell w_i = N\left( \sum_{i=1}^p \sum_{\ell=0}^{m_i} \alpha_{i\ell} N^\ell v_i \right).
\]

Consequently,

\[
v - \sum_{i=1}^p \sum_{\ell=0}^{m_i} \alpha_{i\ell} N^\ell v_i = \sum_{i=1}^p \beta_i N^{m_i+1} v_i + \sum_{\ell=1}^q \gamma_{i\ell} y_{i\ell},
\]

since the vectors

\[
\{N^{m_1}w_1, \ldots, N^{m_p}w_p, y_1, \ldots, y_q\} = \{N^{m_1+1}v_1, \ldots, N^{m_p+1}v_p, y_1, \ldots, y_q\}
\]

form a basis of \( \ker N \). Linear independence is a consequence of the following counting argument. The number of vectors in the Jordan chains is

\[
\sum_{i=1}^p (m_i + 2) + q = \sum_{i=1}^p (m_i + 1) + (p + q) = \dim \text{im} N + \dim \ker N = \dim V.
\]

\[\square\]

We note that finding the generating vectors of the Young diagram of \( N \) or equivalently the Jordan normal form of \( N \), involves solving linear equations with coefficients in the field \( k \) and thus only operations in the field \( k \).
3.2 Some facts about $S$

We now study the semisimple part $A$.

Lemma 3.2.1 $V = \ker S \oplus \im S$. Moreover the characteristic polynomial $\chi_S(\lambda)$ of $S$ can be written as a product of $\lambda^n$, where $n = \dim \ker S$ and $\chi_{S|\im S}$, the characteristic polynomial of $S|\im S$. Note that $\chi_{S|\im S}(0) \neq 0$

Proof. $\ker S$ is an $S$-invariant subspace of $V$. Since $Sv = 0$ for every $v \in \ker S$, the characteristic polynomial of $S|\ker S$ is $\lambda^n$.

Because $S$ is semisimple, there is an $S$-invariant subspace $Y$ of $V$ such that $V = \ker S \oplus Y$. The linear mapping $S|Y : Y \to Y$ is invertible, for if $Sy = 0$ for some $y \in Y$, then $S(y + u) = 0$ for every $u \in \ker S$. Therefore $y + u \in \ker S$, which implies that $y \in \ker S \cap Y = \{0\}$; that is, $y = 0$. So $S|Y$ is invertible. Suppose that $y \in Y$, then $y = S((S|Y)^{-1}y) \in \im S$. Thus $Y \subseteq \im S$. But $\dim \im S = \dim V - \dim \ker S = \dim Y$. So $Y = \im S$.

Since $\ker S \cap \im S = \{0\}$, we see that $\lambda$ does not divide the polynomial $\chi_{S|\im S}(\lambda)$. Consequently, $\chi_{S|\im S}(0) \neq 0$. Since $V = \ker S \oplus \im S$, where $\ker S$ and $\im S$ are $S$-invariant subspaces of $V$, we obtain

$$\chi_S(\lambda) = \chi_{\ker S}(\lambda) \cdot \chi_{S|\im S}(\lambda) = \lambda^n \chi_{S|\im S}(\lambda).$$

Lemma 3.2.2 The subspaces $\ker S$ and $\im S$ are $N$-invariant and hence $A$-invariant.

Proof. Suppose that $x \in \im S$. Then there is a vector $v \in V$ such that $x = Sv$. So $Nx = N(Sv) = S(Nv) \in \im S$. In other words, $\im S$ is an $N$-invariant subspace of $V$. Because $\im S$ is also $S$-invariant and $A = S + N$, it follows that $\im S$ is an $A$-invariant subspace of $V$. Suppose that $x \in \ker S$, that is, $Sx = 0$. Then $S(Nx) = N(Sx) = 0$. So $Nx \in \ker S$. Therefore $\ker S$ is an $N$-invariant and hence $A$-invariant subspace of $V$.

3.3 Description of uniform normal form

We now describe the uniform normal form of the linear mapping $A : V \to V$, using both its semisimple and nilpotent parts.

Since $A|\ker S = N|\ker S$, we can apply the discussion of §3.1 to obtain a basis of $\ker S$ which realizes the Young diagram of $N|\ker S$, which say has $r$ columns. For $1 \leq \ell \leq r$ let $F_{q\ell}$ be the space spanned by the generating vectors of Jordan chains of $N|\ker S$ in $\ker S$ of length $m_\ell$.

By lemma 3.2.1 $A|\im S$ is a linear mapping of $\im S$ into itself with invertible semisimple part $S|\im S$ and commuting nilpotent part $N|\im S$. Using
the discussion of §3.1 for every $r + 1 \leq \ell \leq p$ let $F_{\ell}$ be the set of generating vectors of the Jordan chains of $N|\text{im} S$ in $\text{im} S$ of length $m_\ell$, which occur in the $p - (r + 1)$ columns of the Young diagram of $N|\text{im} S$.

Now we prove

**Claim 3.3.1** For each $1 \leq \ell \leq p$ the space $F_{\ell}$ is $S$-invariant.

**Proof.** Let $v^\ell \in F_{\ell}$. Then $\{v^\ell, Nv^\ell, \ldots, N^{m_{\ell}-1}v^\ell\}$ is a Jordan chain in the Young diagram of $N$ of length $m_\ell$ with generating vector $v^\ell$. For each $1 \leq \ell \leq r$ we have $F_{\ell} \subseteq \ker S$. So trivially $F_{\ell}$ is $S$-invariant, because $S = 0$ on $F_{\ell}$. Now suppose that $r + 1 \leq \ell \leq p$. Then $F_{\ell} \subseteq \text{im} S$ and $S|\text{im} S$ is invertible. Furthermore, suppose that for some $\alpha_j \in k$ with $0 \leq j \leq m_\ell - 1$ we have $0 = \sum_{j=0}^{m_{\ell}-1} \alpha_j N^j(Sv^\ell)$. Then $0 = S(\sum_{j=0}^{m_{\ell}-1} \alpha_j N^j v^\ell)$, because $S|\text{im} S$ and $N|\text{im} S$ commute. Since $S|\text{im} S$ is invertible, the preceding equality implies $0 = \sum_{j=0}^{m_{\ell}-1} \alpha_j N^j v^\ell$. Consequently, by lemma 3.1.1 we obtain $\alpha_j = 0$ for every $0 \leq j \leq m_\ell - 1$. In other words, $\{Sv^\ell, N(Sv^\ell), \ldots, N^{m_{\ell}-1}(Sv^\ell)\}$ is a Jordan chain of $N|\text{im} S$ in $\text{im} S$ of length $m_\ell$ with generating vector $Sv^\ell$. So $Sv^\ell \in F_{\ell}$. Thus $F_{\ell}$ is an $S$-invariant subspace of $\text{im} S$ and hence is an $S$-invariant subspace of $V$, since $V = \text{im} S \oplus \ker S$.

An $A$-invariant subspace $U$ of $V$ is uniform of height $m - 1$ if $N^{m-1}U \neq \{0\}$ and $N^mU = \{0\}$ and $\ker N^{m-1}U = NU$. For each $1 \leq \ell \leq r$ let $U_{\ell}$ be the space spanned by the vectors in the Jordan chains of length $m_\ell$ in the Young diagram of $N|\ker S$ and for $r + 1 \leq \ell \leq p$ let $U_{\ell}$ be the space spanned by the vectors in the Jordan chains of length $m_\ell$ in the Young diagram of $N|\text{im} S$.

**Claim 3.3.2** For each $1 \leq \ell \leq p$ the subspace $U_{\ell}$ is uniform of height $m_{\ell} - 1$.

**Proof.** By definition $U_{\ell} = F_{\ell} \oplus NF_{\ell} \oplus \cdots \oplus N^{m_{\ell}-1}F_{\ell}$. Since $N^{m_{\ell}} F_{\ell} = \{0\}$ but $N^{m_{\ell}-1} F_{\ell} \neq \{0\}$, the subspace $U_{\ell}$ is $A$-invariant and of the height $m_{\ell} - 1$. To show that $U_{\ell}$ is uniform we need only show that $\ker N^{m_{\ell}-1} \cap U_{\ell} \subseteq NU_{\ell}$ since the inclusion of $NU_{\ell}$ in $\ker N^{m_{\ell}-1}$ follows from the fact that $N^{m_{\ell}} F_{\ell} = 0$. Suppose that $u \in \ker N^{m_{\ell}-1} \cap U_{\ell}$, then for every $0 \leq i \leq m_{\ell} - 1$ there are unique vectors $f_i \in F_{\ell}$ such that $u = f_0 + Nf_1 + \cdots + N^{m_{\ell}-1}f_{m_{\ell}-1}$. Since $u \in \ker N^{m_{\ell}-1} \cap U_{\ell}$ we get $0 = N^{m_{\ell}-1}u = N^{m_{\ell}-1}f_0$. If $f_0 \neq 0$, then the preceding equality contradicts the fact that $f_0$ is a generating vector of a Jordan chain of $N$ of length $m_\ell$. Therefore $f_0 = 0$, which means that $u = N(f_1 + \cdots + N^{m_{\ell}-2}f_{m_{\ell}-1}) \in NU_{\ell}$. This shows that $\ker N^{m_{\ell}-1} \cap U_{\ell} \subseteq NU_{\ell}$. Hence $\ker N^{m_{\ell}-1} \cap U_{\ell} = NU_{\ell}$, that is, the subspace $U_{\ell}$ is uniform of height $m_{\ell} - 1$. 

11
Now we give an explicit description of the uniform normal form of the linear mapping $A$. For each $1 \leq \ell \leq p$ let $\chi_{S|F_{q\ell}}$ be the characteristic polynomial of $S$ on $F_{q\ell}$. From the fact that every summand in $U_{q\ell} = F_{q\ell} \oplus NF_{q\ell} \oplus \cdots \oplus N^{m_{\ell}-1}F_{q\ell}$ is $S$-invariant, it follows that the characteristic polynomial $\chi_{S|U_{q\ell}}$ of $S$ on $U_{q\ell}$ is $\chi_{S|F_{q\ell}}^{m_{\ell}}$. Since $V = \sum_{\ell=1}^{p} \oplus U_{q\ell}$, we obtain

$$\chi_S = \prod_{\ell=1}^{p} \chi_{S|F_{q\ell}}^{m_{\ell}}.$$ 

Choose a basis $\{u_{j,\ell}^{\ell}\}_{j=1}^{q_{\ell}}$ of $F_{q\ell}$ so that the matrix of $S|F_{q\ell}$ is the $q_{\ell} \times q_{\ell}$ companion matrix $C_{q_{\ell}}$ associated to the characteristic polynomial $\chi_{S|F_{q\ell}}$. When $1 \leq \ell \leq r$ the companion matrix $C_{q_{\ell}}$ is 0 since $S|F_{q_{\ell}} = 0$. With respect to the basis $\{u_{j,\ell}^{\ell}, Nu_{j,\ell}^{\ell}, \ldots, N^{m_{\ell}-1}u_{j,\ell}^{\ell}\}_{j=1}^{q_{\ell}}$ of $U_{q\ell}$ the matrix of $A|U_{q\ell}$ is the $m_{q_{\ell}}q_{\ell} \times m_{q_{\ell}}q_{\ell}$ matrix

$$D_{m_{q_{\ell}}} = \begin{pmatrix} C_{q_{\ell}} & 0 & 0 & \cdots & 0 \\ I & C_{q_{\ell}} & 0 & \cdots & 0 \\ 0 & I & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I & C_{q_{\ell}} \\ 0 & \cdots & 0 & 0 & I \end{pmatrix}.$$ 

Since $V = \sum_{\ell=1}^{p} \oplus U_{q\ell}$, the matrix of $A$ is $\text{diag}(D_{m_{q_{1}}}, \ldots, D_{m_{q_{p}}})$ with respect to the basis $\{u_{j,\ell}^{\ell}, Nu_{j,\ell}^{\ell}, \ldots, N^{m_{\ell}-1}u_{j,\ell}^{\ell}\}_{j=1}^{q_{\ell}}$. We call preceding matrix the uniform normal form for the linear map $A$ of $V$ into itself. We note that this normal form can be computed using only operations in the field $k$ of characteristic 0.

Using the uniform normal form of $A$ we obtain a factorization of its characteristic polynomial $\chi_A$ over the field $k$.

**Corollary 3.3.3** $\chi_A(\lambda) = \prod_{\ell=1}^{p} \chi_{S|F_{q\ell}}^{m_{\ell}}(\lambda) = \lambda^n \prod_{\ell=r+1}^{p} \chi_{S|F_{q\ell}}^{m_{\ell}}(\lambda)$, where $n = \sum_{\ell=1}^{r} m_{\ell} = \dim \ker S$.

**References**

[1] N. Burgoyne and R. Cushman, *The decomposition of a linear mapping*, J. Lin. Alg. & Appl., 8 (1974) 515–519.

[2] J. Hartl, *Ein einfacher Induktionsbeweis für die Existenz der Jordan-Normalform*, Arch. Math. (Basel) 50 (1988) 323–327.