More on Jordan Form

I will show uniqueness of the Jordan form of a linear operator in a more direct way than is done in the book. We suppose that $T: V \to V$ is a linear operator on a finite dimensional vector space V and the characteristic polynomial of T is $p(x) = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}$ and the minimal polynomial of T is $q(x) = (x - c_1)^{e_1} \cdots (x - c_k)^{e_k}$. We know already that $e_i \le d_i$ for all i and we may also suppose that $e_i > 0$ since p and q have the same roots by Thm 3, page 193.

Then we proved that there is a basis $\mathcal{B}$ of V so that if $A = [T]_\mathcal{B}$ then $A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_k \end{bmatrix}$ where each A_i is a $d_i \times d_i$ matrix and $A_i = \begin{bmatrix} J_{\ell_{i,1} c_i} & 0 & \cdots & 0 \\ 0 & J_{\ell_{i,2} c_i} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & J_{\ell_{i,n_i} c_i} \end{bmatrix}$ where (using different notation from the book) $J_{\ell c}$ is an $\ell \times \ell$ Jordan block with c on the diagonal entries, 1 just below each diagonal entry and 0 everywhere else. Of course we have $\ell_{i,1} + \ell_{i,2} + \cdots + \ell_{i,n_i} = d_i$. We also may as well reorder the basis so that the Jordan blocks have nonincreasing size, so $\ell_{i,1} \ge \ell_{i,2} \ge \cdots \ge \ell_{i,n_i}$.

The uniqueness we will show is that the integers $\ell_{i,1}, \ell_{i,2}, \cdots, \ell_{i,n_i}$ are completely determined by T which means that T has only one possible Jordan form, up to shuffling the order of the Jordan blocks.

Right now I will just state the results, and will fill in the details later. Let m_{ij} be the number of m so that $\ell_{i,m} = j$, in other words, the number of $j \times j$ Jordan blocks in A_i. It suffices to show that the numbers m_{ij} are uniquely determined by T.

0) $\ell_{i,1} = e_i$ so the exponent of $(x - c_i)$ in the minimal polynomial gives the size of the largest Jordan block in A_i.
1) n_i is the dimension of the space of characteristic vectors for the characteristic value c_i. In other words $n_i = \dim NS(T - c_i I)$.
2) $\dim NS(T - c_i I)^2 = 2n_i - m_{i1}$, so m_{i1} is determined by T.
3) $\dim NS(T - c_i I)^3 = 3n_i - m_{i2} - 2m_{i1}$, so m_{i2} is determined by T.
4) In general, $\dim NS(T - c_i I)^{j+1} = (j+1)n_i - m_{i,j} - 2m_{i,j-1} - \cdots - jm_{i,1}$, so m_{ij} is determined by T for each j. Consequently the sizes of the Jordan blocks are uniquely determined by T.

Before proving these facts, we'll prove a couple simple lemmas:

Lemma 1. *Suppose $W_1, \ldots, W_k$ are independent subspaces of V and $U_i \subset W_i$ are subspaces. Then $U_1, U_2, \ldots, U_k$ are independent subspaces of V.*

Proof: This is really trivial. Suppose $\alpha_1 + \alpha_2 + \cdots + \alpha_k = 0$ and $\alpha_i \in U_i$. Then $\alpha_i \in W_i$ for all i, so each $\alpha_i = 0$ by independence of the W_i. ∎

Lemma 2. *Suppose $W_1, \ldots, W_k$ are independent subspaces of V such that $W_1 \oplus W_2 \oplus \cdots \oplus W_k = V$. Suppose also that each W_i is invariant under a linear operator $S: V \to V$. Let $S_i: W_i \to W_i$ be the restriction of S to W_i. Then $NS(S) = NS(S_1) \oplus NS(S_2) \oplus \cdots \oplus NS(S_k)$. In particular, the nullity of S is the sum of the nullities of the S_i.*

Proof: Suppose $\beta \in NS(S)$, so $S\beta = 0$. We may write $\beta = \beta_1 + \beta_2 + \cdots + \beta_k$ where $\beta_i \in W_i$. We have $0 = S\beta = S\beta_1 + S\beta_2 + \cdots + S\beta_k = S_1\beta_1 + S_2\beta_2 + \cdots + S_k\beta_k$. Since $S_i\beta_i \in W_i$ we know by independence of the W_i that each $S_i\beta_i = 0$. Thus $\beta_i \in NS(S_i)$ and we see that $\beta \in NS(S_1) + NS(S_2) + \cdots + NS(S_k)$. Conversely, if $\alpha \in NS(S_1) + NS(S_2) + \cdots + NS(S_k)$ then $\alpha = \alpha_1 + \cdots + \alpha_k$ with $\alpha_i \in NS(S_i)$, but then $S\alpha = S\alpha_1 + \cdots + S\alpha_k = 0 + \cdots + 0 = 0$ so $\alpha \in NS(S)$. So we have shown $NS(S) = NS(S_1) + NS(S_2) + \cdots + NS(S_k)$. But by Lemma 1, $NS(S_1) + NS(S_2) + \cdots + NS(S_k) = NS(S_1) \oplus NS(S_2) \oplus \cdots \oplus NS(S_k)$ ∎

To show 1)-4) above, we apply Lemma 2 to $S = (T - c_i I)^\ell$ using the decomposition into cyclic subspaces given by the Jordan form. Each of these cyclic subspaces has the form $Z(\alpha, N_j)$ where $\alpha \in V$, $N_j = (T - c_j I)$, $N_j^m(\alpha) = 0$, and $N_j^{m-1} \ne 0$. In this case the cyclic subspace has dimension m and will correspond to one of the $m \times m$ Jordan blocks in A_j. First of all, if $j \ne i$ then the null space of the restriction of S to $Z(\alpha, N_j)$ is 0. (The usual proof of this involves ideals, but there are many other proofs. I leave it as an exercise.) So we may suppose $j = i$. Then $S = N_i^\ell$. So if $\ell \geq m$ we know the null space of the restriction of S to $Z(\alpha, N_i)$ is all of $Z(\alpha, N_i)$ and so has dimension m. If $\ell < m$ then the null space is the span of $N_i^{m-\ell}\alpha, N_i^{m-\ell+1}\alpha, \ldots, N_i^{m-1}\alpha$ and thus has dimension ℓ.

So apply all this when $\ell = 1$ we see that the null space of the restriction of $T - c_iI$ to $Z(\alpha, N_i)$ always has dimension 1, so by Lemma 2 the dimension of $NS(T - c_iI)$ is the number of such cyclic subspaces, which is the same as the number of Jordan blocks in A_i. So 1) is shown.

Letting $\ell = 2$ we see that the null space of the restriction of $(T - c_iI)^2$ to $Z(\alpha, N_i)$ has dimension 2 if $m \geq 2$ and dimension 1 if $m = 1$. So by Lemma 2 the dimension of $NS(T - c_iI)^2$ is two times the number of Jordan blocks in A_i which are 2×2 or larger, plus the number of 1×1 Jordan blocks in A_i. This is the same as twice the number of Jordan blocks in A_i minus the number of 1×1 Jordan blocks in A_i. So 2) is shown.

For general ℓ the dimension of $(T - c_iI)^\ell$ is ℓ times the number of Jordan blocks in A_i which are $\ell \times \ell$ or larger plus n times the number of $n \times n$ Jordan blocks in A_i for each $n < \ell$, so 4) (and 3)) are shown.

Here is an example. Suppose $T: V \to V$ has characteristic values 2 and 3 and no others. Suppose $(T - 2I)$ has nullity 3, $(T - 2I)^2$ has nullity 5, $(T - 2I)^j$ has nullity 6 for all $j > 2$, $(T - 3I)$ has nullity 2, and $(T - 3I)^j$ has nullity 3 for all $j > 1$. What is the Jordan form of T? From 1) we know there are $n_1 = 3$ Jordan blocks for characteristic value 2 and $n_2 = 2$ Jordan blocks for characteristic value 3. From 2) we know that $5 = 2 \cdot 3 - m_{11}$ so $m_{11} = 1$. Also $3 = 2 \cdot 2 - m_{21}$ so $m_{21} = 1$. So each characteristic value has a single 1×1 block. From 2) we know $6 = 3 \cdot 3 - m_{12} - 2 \cdot 1$ so $m_{12} = 1$. Also $3 = 3 \cdot 2 - m_{22} - 2 \cdot 1$ so $m_{22} = 1$. So each characteristic value has a single 2×2 block. From 3) we know $6 = 4 \cdot 3 - m_{13} - 2 \cdot 1 - 3 \cdot 1$ so $m_{13} = 1$. So in the end, we see the Jordan blocks are one of each of $J_{12}, J_{22}, J_{32}, J_{13}, J_{23}$.

Note that $\dim NS(T - c_iI)^{j+1} - \dim NS(T - c_iI)^j = n_i - m_{ij} - m_{i,j-1} - \cdots - m_{i1} =$ the number of Jordan blocks in A_i which are bigger than $j \times j$. So this difference in nullities cannot increase as j increases. Moreover, if this difference is ever 0 then we have reached the end, there are no Jordan blocks in A_i bigger than $j \times j$.