Lemma 1

Suppose you have a diagonal operator of the form $D = \begin{pmatrix} \lambda & 0 & ... & 0 \\ 0 & \lambda & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & ... & \lambda \end{pmatrix}$ and some nilpotent matrix N with all entries above the diagonal. Then, $DN = ND$.

Proof:

$D = \lambda I$, so $DN = \lambda I N = N \lambda I = ND$.

Lemma 2

Suppose you have two $m \times m$ block diagonal operators:

$$A = \begin{pmatrix} A_1 & 0... & 0 & \\ 0 & A_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & ... & A_m \end{pmatrix} \quad B = \begin{pmatrix} B_1 & 0... & 0 & \\ 0 & B_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & ... & B_m \end{pmatrix}$$

Where each A_i, B_i is an $a_i \times a_i$ block. Then, we get:

$$AB = \begin{pmatrix} A_1B_1 & 0... & 0 & \\ 0 & A_2B_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & ... & A_mB_m \end{pmatrix}$$

That is, the diagonal blocks of the product is the matrix product of the diagonal blocks.

Proof:

For a general block diagonal matrix with m $a_i \times a_i$ diagonal blocks $T_1...T_m$, we can write:

$$T_{ij} = \begin{cases} \sum_{p=1}^{k-1} a_p < i \leq \sum_{p=1}^{k} a_p \text{ and } \sum_{p=1}^{k-1} a_p < j \leq \sum_{p=1}^{k} a_p : & (T_k)_{(i-\sum_{p=1}^{k-1} a_p, j-\sum_{p=1}^{k-1} a_p)} \\ \text{otherwise } : & 0 \end{cases}$$

For simplicity, if $\sum_{p=1}^{k-1} a_p < i \leq \sum_{p=1}^{k} a_p$, define $i_k = i - \sum_{p=1}^{k-1} a_p$ Also define, and so $0 < i_k \leq a_k$. Similarly, define $j_k = j - \sum_{p=1}^{k-1} a_p$, so $0 < j_k \leq a_k$. So, now we have:

$$AB_{i,j} = \sum_{q=0}^{n} A_{i,q} B_{q,j}$$

For the first factor to be nonzero, with $0 < i_k \leq a_k$, we need $0 < q_k \leq a_k$. Given this, for the second factor to be nonzero, we need $0 < j_k \leq a_k$. What this means is that everything term is 0 except the parts of the sum that satisfy the above conditions, so we get:

$$AB_{i,j} = \sum_{q_k=0}^{n-\sum_{p=1}^{k-1} a_p} A_{i_k,q_k} B_{q_k,j_k} = (A_k B_k)_{i_k,j_k}$$

As desired.

The Actual Goddamn Proof

Okay, luckily this part is pretty easy. Just put T in block diagonal form, so

$$T = \begin{pmatrix} T_1 & 0... & 0 & 0 \\ 0 & T_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & ... & T_m \end{pmatrix}$$

Let diag(T_i) equal the diagonal of one of the T_i blocks. Then, define

$$D = \begin{pmatrix} \text{diag}(T_1) & 0... & 0 & 0 \\ 0 & \text{diag}(T_2) & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & ... & \text{diag}(T_m) \end{pmatrix}$$

Which is obviously diagonal, and N as $T - D$. Then, N is nilpotent, since we just took the diagonal away from an upper triangular matrix, meaning it is now strictly upper triangular.

$$N = \begin{pmatrix} T_1 - \text{diag}(T_1) & 0... & 0 & 0 \\ 0 & T_2 - \text{diag}(T_2) & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & ... & T_m - \text{diag}(T_m) \end{pmatrix}$$

Now, each $T_i - \text{diag}(T_i$ is nilpotent. So, we know from Lemma 2 that DN can be found by multiplying each matrix on the diagonal of those operators together, and since the matrices on the diagonal of D are diagonal (with only one value on the diagonal) and the matrices on the diagonal of N are nilpotent, Lemma 1 shows each diag(T_i) commutes with each $T_i - \text{diag}(T_i)$. Since the diagonal blocks commute, and ND is just the product of their diagonal blocks, $ND = DN$.