

## Cayley Hamilton theorem: LinearAlg

Jonathan L.F. King  
 University of Florida, Gainesville FL 32611-2082, USA  
 squash@ufl.edu  
 Webpage <http://squash.1gainesville.com/>  
 17 November, 2017 (at 12:05)

**Ques. Q1.** Suppose two  $\mathbf{F}$ -matrices are conjugate over the algebraic closure of  $\mathbf{F}$ . Are they conjugate over  $\mathbf{F}$ ?  $\square$

**Notation.** Use  $\varphi(x) := \text{Det}(\mathbf{M} - x\mathbf{I})$  for the *characteristic poly* of  $\mathbf{M}$ . I'll use symbol " $\simeq$ " with the following meaning: Suppose  $\varphi$  is the characteristic polynomial of an  $N \times N$  matrix  $\mathbf{M}$ , or of a trn  $\mathbf{T}: \mathbf{F}^N \rightarrow \mathbf{F}^N$ , and  $h$  is a polynomial. I'll write  $\varphi \simeq h$  to mean that

$$[-1]^N \cdot \varphi = h.$$

Use a similar convention for an alteration of the word "monic": The phrase

"Consider a degree- $K$  *m*-*monic* polynomial  $g \dots$ "

means that the high-order term of  $g(x)$  is  $[-1]^K x^K$ .

Let  $\mathbf{0}$  denote the zero-matrix or trn. Use  $\vec{0}$  for the zero vector.

**1: Cayley-Hamilton Theorem.** Over field  $\mathbf{F}$ , consider an  $N \times N$ -matrix  $\mathbf{M}$ . Then

$$\varphi(\mathbf{M}) = \mathbf{0}_{N \times N}.$$

So  $\mathbf{M}$  is a "root" of its own char-poly.  $\diamond$

**Proof when  $\mathbf{M}$  is upper-triangular.** In matrix  $\mathbf{M}$ , let  $\alpha_1, \alpha_2, \dots, \alpha_N \in \mathbf{F}$  be the diagonal entries; these are the eigenvalues of  $\mathbf{M}$ . Using the std basis, let  $\mathbf{E}_j := \text{Spn}(\{\mathbf{e}_1, \dots, \mathbf{e}_j\})$ ; so  $\mathbf{E}_0 = \{\mathbf{0}\}$ . Since  $\mathbf{M}$  is upper-triangular, the difference vector

$$\begin{aligned} 2: \quad \mathbf{d}_{j-1} &:= \mathbf{M}\mathbf{e}_j - \alpha_j \mathbf{e}_j \\ &\text{is in } \mathbf{E}_{j-1}, \end{aligned}$$

for each  $j \in [1..N]$ . We want to show that each such  $\mathbf{e}_j$  is annihilated by  $\varphi(\mathbf{M})$ .

For  $j \in [0..N]$ , factor the characteristic polynomial as  $\varphi \simeq L_j \cdot R_j$ , where the left&right are

$$\begin{aligned} L_j(x) &:= [x - \alpha_N] \cdot [x - \alpha_{N-1}] \cdot \dots \cdot [x - \alpha_{j+1}] ; \\ R_j(x) &:= [x - \alpha_j] \cdot [x - \alpha_{j-1}] \cdot \dots \cdot [x - \alpha_2] \cdot [x - \alpha_1]. \end{aligned}$$

[So  $L_0() \simeq \varphi()$  and  $R_0() = 1$ .] All powers of  $\mathbf{M}$  mutually commute, thus

$$\varphi(\mathbf{M}) \simeq L_j(\mathbf{M}) \cdot R_j(\mathbf{M}).$$

Hence ISTShow that

$$Q[j]: \quad R_j(\mathbf{M}) \text{ annihilates } \mathbf{E}_j.$$

Since all transformations annihilate  $\mathbf{E}_0$ , we need to prove  $Q[j-1] \Rightarrow Q[j]$ , for each  $j = 1, 2, \dots, N$ .

**Induction.** Fix a  $j \in [1..N]$  s.t  $Q[j-1]$ .

Firstly,  $R_j(\mathbf{M})$  annihilates  $\mathbf{e}_1, \dots, \mathbf{e}_{j-1}$ , since  $R_{j-1}(\mathbf{M})$  does, and  $R_j(\mathbf{M}) = [\mathbf{M} - \alpha_j \mathbf{I}] \cdot R_{j-1}(\mathbf{M})$ . Secondly, to kill off  $\mathbf{e}_j$  note that

$$\begin{aligned} R_j(\mathbf{M}) \cdot \mathbf{e}_j &= R_{j-1}(\mathbf{M}) \cdot [\mathbf{M} - \alpha_j \mathbf{I}] \cdot \mathbf{e}_j \\ &= R_{j-1}(\mathbf{M}) \cdot \mathbf{d}_{j-1}. \end{aligned}$$

This last product is  $\mathbf{0}$ , courtesy (2) and  $Q[j-1]$ .  $\blacklozenge$

**Proof of C-H using JCF.** We now handle a general  $\mathbf{M}$  by means of **JCF**, the Jordan Canonical Form thm. Let  $\mathbb{G}$  denote the algebraic closure of  $\mathbf{F}$ . Viewing  $\mathbf{M}$  as acting on  $\mathbb{G}^{\times N}$ , our  $\mathbf{M}$  is conjugate (i.e *similar*) to its Jordan Canonical Form. Since the JCF is upper-triangular, the previous proof finishes the argument in the general case.  $\blacklozenge$

**Elementary proof using a cyclic subspace.** The preceding argument used two non-trivial theorems: JCFThm, as well as the result that a field *has* an algebraic closure.

Here is an elementary proof of C-H thm, never leaving field  $\mathbf{F}$ . Consider a trn  $\mathbf{T}$  on a finite-dim'l  $\mathbf{F}$ -vectorspace and let  $\varphi$  be its characteristic poly. Fixing a vector  $\mathbf{v}_0 \neq \mathbf{0}$ , our goal is to show that

$$3: \quad \varphi(\mathbf{T})(\mathbf{v}_0) \text{ equals } \mathbf{0}.$$

**Exer:** *Why does this suffice?*

Iteratively define  $\mathbf{v}_{j+1} := T(\mathbf{v}_j)$  and stop at the first natnum  $N$  where  $T(\mathbf{v}_N)$  is in the vectorspace

$$\mathbf{W} := \text{Spn}(\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_N).$$

Define coeffs  $\alpha_j$  by

$$4: \quad T(\mathbf{v}_N) := \sum_{j+k=N} \alpha_j \mathbf{v}_k,$$

where such sums are taken over **natnums**  $j$  and  $k$ .

Notice that our  $\mathbf{W}$  is a  $T$ -invariant subspace. And the linearly-independent (*exercise!*) tuple

$$\mathcal{B} := (\mathbf{v}_0, \dots, \mathbf{v}_N)$$

is a basis for subspace  $\mathbf{W}$ .

**5: Companion Lemma.** *With  $T$ ,  $\mathbf{W}$  and  $\mathcal{B}$  as above, let  $\mathbf{M}$  be the  $[N+1] \times [N+1]$  matrix of  $T|_{\mathbf{W}}$  ( $T$  restricted to  $\mathbf{W}$ ) relative to ordered  $\mathbf{W}$ -basis  $\mathcal{B}$ . Then*

$$5a: \quad \mathbf{M} = \begin{bmatrix} 0 & & & & & & \alpha_N \\ 1 & 0 & & & & & \alpha_{N-1} \\ & 1 & 0 & & & & \alpha_{N-2} \\ & & \ddots & \ddots & & & \vdots \\ & & & 1 & 0 & & \alpha_3 \\ & & & & 1 & 0 & \alpha_2 \\ & & & & & 1 & 0 & \alpha_1 \\ & & & & & & 1 & \alpha_0 \end{bmatrix}.$$

And its characteristic polynomial is

$$5b: \quad \varphi_{\mathbf{M}}(x) \simeq x^{N+1} - \sum_{\substack{j+k=N, \\ \text{with } j,k \in \mathbb{N}}} \alpha_j x^k. \quad \diamond$$

**Remark.** A matrix of form (5a) is a **companion matrix**. It is “the companion matrix of polynomial (5b)”. Wikipedia has a nice write-up.  $\square$

**Proof of (5).** The  $[N+1] \times [N+1]$  matrix  $x\mathbf{I} - \mathbf{M}$  is

$$\begin{bmatrix} x & & & & & & -\alpha_N \\ -1 & x & & & & & -\alpha_{N-1} \\ & -1 & x & & & & -\alpha_{N-2} \\ & & \ddots & \ddots & & & \vdots \\ & & & -1 & x & & -\alpha_3 \\ & & & & -1 & x & -\alpha_2 \\ & & & & & -1 & x & -\alpha_1 \\ & & & & & & -1 & x - \alpha_0 \end{bmatrix}.$$

We compute its determinant by summing products over transversals. The main diagonal yields

$$\ddagger_0: \quad x^N \cdot [x - \alpha_0] \stackrel{\text{note}}{=} x^{N+1} - \alpha_0 x^N.$$

Now, in columns  $0, 1, \dots, N-1$  we either choose “ $x$ ” or “ $-1$ ”. In a column where we choose  $-1$ , the row of our choice *prevents* us from choosing  $x$  in the *next* column; so we must again choose  $-1$ . Thus: *Once we leave the main diagonal, we must stay on the first off-diagonal.*

So what is the contribution to  $\text{Det}(x\mathbf{I} - \mathbf{M})$  from a transversal with  $j \in [1..N]$  many  $-1$ ’s? It is

$$x^{N-j} \cdot [-1]^j \cdot [-\alpha_j] \cdot [\text{Sign of permutation}].$$

The sign of the perm is  $[-1]^j$ , so the  $j^{\text{th}}$ -transversal contribution to  $\varphi_{\mathbf{M}}(x)$  is

$$\ddagger_j: \quad -[\alpha_j \cdot x^{N-j}].$$

Adding ( $\ddagger_0$ ) to  $\sum_{j=1}^N (\ddagger_j)$  yields RhS(5b).  $\blacklozenge$

**Second Proof of C-H.** The given trn  $T$  and vector  $\mathbf{v}_0$  determine a  $T$ -invariant subspace  $\mathbf{W}$  and matrix  $\mathbf{M}$ , as above. An easy exercise –see the `Triangular Matrix Lemma` in the `jordan_decomp.latex` file– shows that  $\varphi_{\mathbf{M}}$  is a factor-poly of  $\varphi_T$ . So (3) will follow from showing that  $\varphi_{\mathbf{M}}(T)$  annihilates  $\mathbf{v}_0$ . And this follows from (5b) and (4).  $\blacklozenge$

**6: Corollary.** *Fix  $K \in \mathbb{Z}_+$  and an arbitrary degree- $K$  monic  $\mathbf{F}$ -poly  $g()$ . Then there exists a  $K \times K$  matrix over  $\mathbf{F}$  whose characteristic-poly equals  $g$ . Pf. Use matrix (5a) with  $K := N+1$ .  $\blacklozenge$*

**7: Application.** *Let  $\mathbf{F} := \mathbb{Z}_p$ , where  $p$  is prime. To produce a  $p \times p$   $\mathbf{F}$ -matrix  $\mathbf{M}$  with no  $\mathbf{F}$ -eigenvalues, pick a non-zero element  $\beta \in \mathbf{F}$ , and define*

$$g(x) := \beta + \prod_{\gamma \in \mathbf{F}} [x - \gamma].$$

*F*irinstance, consider  $p := 3$  and  $\beta := -1$ . Then

$$\begin{aligned} g(x) &= -1 + x[x - 1][x + 1] \\ &= x^3 - [x + 1] = x^3 - [\alpha_0 x^2 + \alpha_1 x + \alpha_2], \end{aligned}$$

using the notation of (5b), where  $\alpha_0 := 0$ ,  $\alpha_1 := 1$  and  $\alpha_2 := 1$ . Courtesy our (5a), then, matrix

$$M := \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \text{ has no eigenvalues in } \mathbb{Z}_3. \quad \diamond$$

## End Notes

First, we need a general lemma.

**8: Lemma.** Fix fields  $\mathbb{G} \supset \mathbf{F}$  and consider a collection  $\mathcal{C} \subset \mathbf{F}^{\times N}$  of vectors which is linearly dependent over  $\mathbb{G}$ . (Typically,  $\mathbb{G}$  is the algebraic closure of  $\mathbf{F}$ .) Then  $\mathcal{C}$  is already linearly dependent over  $\mathbf{F}$ . \(\diamond\)

*Proof.* View  $\mathbf{F}^{\times N}$ -vectors as column vectors, and use  $\vec{0}$  for the col-vec of all zeros. FTSOC, suppose we have a *non-trivial* dependence

$$\dagger: \quad \sum_{j=1}^7 \alpha_j \cdot \vec{c}_j = \vec{0},$$

for scalars  $\alpha_j \in \mathbb{G}$  and col-vecs in  $\vec{c}_j \in \mathcal{C}$ . Some  $\alpha_j \neq 0$ , so WLOG  $\alpha_1 \neq 0$ . By multiplying  $(\dagger)$  by  $1/\alpha_1$ , WLOG  $\boxed{\alpha_1 = 1}$ .

Now view  $\mathbb{G}$  as an  $\mathbf{F}$ -vectorspace. Collection  $\{1\}$  is L.I, so the axiom-of-choice says we can extend it to to an  $\mathbf{F}$ -basis  $\{1\} \sqcup \mathcal{E}$  for  $\mathbb{G}$ . [So  $\mathcal{E} \subset \mathbb{G}$ , and every  $\alpha \in \mathbb{G}$  can be uniquely written as an  $\mathbf{F}$ -linear-combination of “vectors” in  $\mathbb{G}$ .]

Define a linear map  $\text{Proj}:\mathbb{G} \rightarrow \mathbf{F}$  by  $1 \mapsto 1$  and, for each  $e \in \mathcal{E}$ , have  $\text{Proj}$  send  $e \mapsto 0$ . Whence  $\text{Proj}()$  is the identity on  $\mathbf{F}$ , and for  $\alpha, \beta \in \mathbb{G}$  and  $f \in \mathbf{F}$ :

$$\text{Proj}(\alpha + \beta) = \text{Proj}(\alpha) + \text{Proj}(\beta);$$

$$*: \quad \text{Proj}(\alpha \cdot f) = \text{Proj}(\alpha) \cdot f.$$

Applying map  $\text{Proj}^{\times N}:\mathbb{G}^{\times N} \rightarrow \mathbf{F}^{\times N}$  to  $(\dagger)$  yields

$$\ddagger: \quad \sum_{j=1}^7 \text{Proj}(\alpha_j) \cdot \vec{c}_j = \vec{0}$$

by  $(*)$ , since each entry in each  $\vec{c}_j$  is in  $\mathbf{F}$ .

And  $\text{Proj}(\alpha_1) = \text{Proj}(1) = 1$ , which is not zero. So  $(\ddagger)$  exposes a non-trivial  $\mathbf{F}$ -linear-dependence of  $\mathcal{C}$ . \(\diamond\)

### Minimal poly of M

See *jordan\_decomp. latex* for theorems used below. The *minimal polynomial* of an  $\mathbf{F}$ -matrix  $M$  is the smallest-degree monic  $\mathbf{F}$ -poly  $\Upsilon()$  such that  $\Upsilon(M) = \mathbf{0}$ . Applying (8) to collection  $\mathcal{C} := \{M^j\}_{j \in \mathbb{N}}$  shows, if we take the smallest-degree monic  $\mathbb{G}$ -poly, that we still get  $\Upsilon$ .

For a  $\lambda \in \mathbb{G}$ , consider the  $D \times D$  Jordan Block

$$J := \lambda\text{-JB}(D) := \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

It is the sum  $\lambda \mathbf{I} + \mathbf{N}$ , where  $\mathbf{N}$  is the nilpotent matrix  $0\text{-JB}(D)$ . For  $R \in \mathbb{N}$ , the Binomial thm applies, since  $\mathbf{N} \Leftrightarrow \mathbf{I}$ , to the  $R^{\text{th}}$ -power of  $J$  to assert

$$J^R = \sum_{j+k=R} \lambda^j \cdot \binom{R}{j,k} \cdot \mathbf{N}^k.$$

For  $R \in [0..D)$ , then,  $J^R$  has 1's on the  $R^{\text{th}}$  off-diagonal, and 0's on all higher diagonals. Thus  $\{\mathbf{I}, J, \dots, J^{D-1}\}$  is a lin-indep collection of matrices. And  $J^D = \lambda^D \mathbf{I}$ . So  $\text{Deg}(\Upsilon_J)$  equals  $D$ . Therefore,

$$\wp_J(x) \simeq \Upsilon_J(x) = [x - \lambda]^D.$$

**9: Fact.** The characteristic and minimum polynomials satisfy

$$\begin{aligned} \wp_M &= \wp_A \cdot \wp_B && ; \\ \Upsilon_M &= \text{Lcm}(\Upsilon_A, \Upsilon_B) && , \end{aligned}$$

when  $M := \text{Diag}(A, B)$  is a block-diagonal matrix. *Proof.* Immediate. \(\diamond\)

*Caveat.* Suppose M is *block upper-triangular*; it has square-blocks  $B_1, \dots, B_L$  along the diagonal, zeros south-west of these blocks, and possibly non-zero values north-east of these blocks. Certainly

$$\wp_M = \wp_{B_1} \cdot \wp_{B_2} \cdot \dots \cdot \wp_{B_L}.$$

However, the corresponding stmt for  $\Upsilon_M$  with Lcm is **false**.

As a CEX, the matrices  $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$  and  $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$  have the same  $1 \times 1$  diagonal-blocks, and the same char-poly, but different min-polys; they are  $x$  and  $x^2$ .  $\square$

An eigenvalue is a “*simple* eigenvalue” if its eigenspace is 1-dim’al.

**10: Coro.** A block-diagonal M has “equality”  $\wp_M \simeq \Upsilon_M$  IFF M has only simple eigenvalues.  $\diamond$

*Proof.* This follows from either (9) or (12).  $\diamond$

A *downtup*  $\vec{D}$  is a sequence of positive integers  $D_1 \geq D_2 \geq \dots \geq D_\varepsilon$ . It yields the JCF

$$\lambda\text{-JB}(\vec{D}) := \text{Diag}(\lambda\text{-JB}(D_1), \dots, \lambda\text{-JB}(D_\varepsilon))$$

of the general  $\lambda$ -nilpotent matrix. Use  $\text{Size}(\vec{D})$  for  $D_1 + \dots + D_\varepsilon$ .

For  $\mathbf{F}$ -matrix M, suppose that  $\lambda_1, \dots, \lambda_L$  are the distinct  $\mathbb{G}$ -eigenvalues. The eigenvalues yield a unique list  $\vec{D}^1, \vec{D}^2, \dots, \vec{D}^L$  of downtups (of varying lengths) so that

$$11: \quad \text{Diag}(\lambda_1\text{-JB}(\vec{D}^1), \dots, \lambda_L\text{-JB}(\vec{D}^L))$$

is the JCF of M.

**12: Theorem.** With  $\mathbf{F}$ -matrix M having JCF (11), its characteristic and minimum polys are

$$\wp_M(x) \simeq \prod_{\ell=1}^L [x - \lambda_\ell]^{\text{Size}(\vec{D}^\ell)} \quad \text{and}$$

$$\Upsilon_M(x) = \prod_{\ell=1}^L [x - \lambda_\ell]^{\text{Max}(\vec{D}^\ell)}.$$

Necessarily, these polynomials have all their coefficients in  $\mathbf{F}$ . *Proof.* Use (8) and (9).  $\diamond$

**13: Fact.** Suppose M is an  $N \times N$  matrix. Write its char and min polys as

$$\wp_M(x) \simeq [x - \lambda_1] \cdot \dots \cdot [x - \lambda_N]$$

$$\Upsilon_M(x) = [x - \beta_1] \cdot \dots \cdot [x - \beta_D].$$

[The  $\beta$ ’s form a sub-multiset of the  $\lambda$ ’s.] Then for each nz-scalar  $\sigma$ :

$$\wp_{\sigma M}(x) = [x - \sigma\lambda_1] \cdot \dots \cdot [x - \sigma\lambda_N] \cdot [-1]^N$$

$$= \sigma^N \cdot \wp_M\left(\frac{1}{\sigma}x\right),$$

and analogously for the min-poly.  $\diamond$

**Continuity.** Easily, the mapping  $M \mapsto \wp_M$  is cts. But neither  $\Upsilon_M$  nor JCF(M) varies continuously with M. For  $\beta \neq 0$ , define  $7 \times 7$  matrices

$$M_\beta := \begin{bmatrix} 0 & \beta & & & & & \\ & \ddots & \ddots & & & & \\ & & 0 & & & & \\ & & & \beta & & & \\ & & & & \ddots & & \\ & & & & & 0 & \\ & & & & & & 1 \end{bmatrix} \text{ and } J := \begin{bmatrix} 0 & 1 & & & & & \\ & \ddots & \ddots & & & & \\ & & 0 & & & & \\ & & & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \\ & & & & & & 1 \end{bmatrix}.$$

The JCF (Jordan Canonical Form) of M is  $J$ . <sup>$\heartsuit^1$</sup>  So the min-poly  $\Upsilon_M(x) = \Upsilon_J(x) = x^7$ . But as  $\beta \rightarrow 0$ , our  $M_\beta$  goes to  $\mathbf{0}_{7 \times 7}$ , whose min-poly is  $x$ . This example also shows that neither eigenspaces nor nilspaces vary ctsly.

**Invariant properties.** Suppose S is invertible. Since ST is conjugate (exercise) to TS, they have the same min-poly and char-poly. We now generalize char-poly to non-invertible:

**14: Lemma.** For  $S, T \in \text{MAT}_{N \times N}(\mathbf{F})$ : Products ST and TS have the same characteristic poly.  $\diamond$

<sup>$\heartsuit^1$</sup> INDIRECT: The only M eval is 0, yet  $\text{Rank}(M) = 7 - 1$ . So the nullspace, i.e 0-espce, is only 1-dimensional, hence JCF(M) has only one JB.

DIRECT: Consider ordered-basis  $\mathcal{V} := (\mathbf{v}_1, \dots, \mathbf{v}_7)$ , where we define  $\mathbf{v}_k := \beta^j \cdot \mathbf{e}_k$  with  $j + k = 7$ . The lefthand action of M, when expressed w.r.t  $\mathcal{V}$ , is J. Equivalently,  $J = C^{-1}MC$  where C is the diagonal matrix with entries  $\beta^6, \beta^5, \dots, \beta, 1$ .

*Proof.* We can proceed as follows if  $\mathbf{F}$  has a topology, with the field operations cts, so that  $\text{GL}(\mathbf{F}^{\times N})$  is dense in  $\text{LIN}(\mathbf{F}^{\times N})$ . For then, take invertible matrices  $\mathbf{S}_j$  which converge to  $\mathbf{S}$  and use that the char-poly varies continuously.  $\diamond$

Here is a standard “Algebraist’s argument” Let  $\tilde{\mathbf{F}}$  be the field generated by  $\mathbf{F}$  and  $N^2$  independent transcendentals. Let  $\tilde{\mathbf{S}}$  be a matrix obtained by putting a distinct transcendental in each position.

Since  $\tilde{\mathbf{S}}$  is  $\tilde{\mathbf{F}}$ -invertible,  $\tilde{\mathbf{S}}\mathbf{T}$  and  $\mathbf{T}\tilde{\mathbf{S}}$  have the same char-poly. Now apply the ring-hom  $\varphi: \tilde{\mathbf{F}} \rightarrow \mathbf{F}$  which sends each transcendental to its corresponding  $\mathbf{F}$ -element in  $\mathbf{S}$ . (I.e, plug in the  $\mathbf{S}$ -values for the corresponding transcendentals in  $\tilde{\mathbf{S}}$ .)  $\blacklozenge$

*Note.* We used that the above ring-hom  $\varphi: \tilde{\mathbf{F}} \rightarrow \mathbf{F}$  preserves determinants (since it preserves mult and addition) hence preserves charpolys.

However, this argument does not show that  $\mathbf{S}\mathbf{T}$  is conjugate to  $\mathbf{T}\mathbf{S}$ . Why? Well,  $\varphi$  can carry an invertible matrix to a *non*-invertible. Perhaps  $\varphi$  carries *every* matrix conjugating  $\tilde{\mathbf{S}}\mathbf{T}$  to  $\mathbf{T}\tilde{\mathbf{S}}$ , to a non-invertible puppy.

Here is an example: Let  $\mathbf{S} := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ ,  $\mathbf{T} := \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ . Then  $\mathbf{S}\mathbf{T}$  is the zero-matrix, but  $\mathbf{T}\mathbf{S}$  equals  $\mathbf{S}$ . So not only is  $\mathbf{S}\mathbf{T}$  not similar (not conjugate to)  $\mathbf{T}\mathbf{S}$ , they even have *different* minpolys, hence different JCFs. Since  $\mathbf{S}$  is the limit of  $\mathbf{S}_x := \begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}$  as  $x \searrow 0$ , we have another example showing that the minimum-poly and JCF do *not* vary ctsly.  $\square$

Filename: Problems/Algebra/LinearAlg/cayley\_hamilton.  
latex

As of: Tuesday 27May2003. Typeset: 17Nov2017 at 12:05.