# 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 July, 2016 (at *20:03*)

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

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

$$[-1]^N \cdot \wp \quad = \quad h\,.$$

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

> *"Consider a degree-$K$ m$\doteqdot$nic polynomial $g\ldots$"*

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 $\mathsf{M}$. Then*

$$\wp(\mathsf{M}) \quad = \quad \mathbf{0}_{N\times N}\,.$$

*So $\mathsf{M}$ is a "root" of its own char-poly.* ◊

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

$$\textbf{2:} \qquad \mathbf{d}_{j-1} := \mathsf{M}\mathbf{e}_j - \alpha_j \mathbf{e}_j \quad \text{is in } \mathbf{E}_{j-1},$$

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

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

$$L_j(x) := [x - \alpha_N] \cdot [x - \alpha_{N-1}] \cdot \ldots \cdot [x - \alpha_{j+1}]\,;$$
$$R_j(x) := [x - \alpha_j] \cdot [x - \alpha_{j-1}] \cdot \ldots \cdot [x - \alpha_2] \cdot [x - \alpha_1]\,.$$

[So $L_0() \doteqdot \wp()$ and $R_0() = 1$.] All powers of $\mathsf{M}$ mutually commute, thus

$$\wp(\mathsf{M}) \doteqdot L_j(\mathsf{M}) \cdot R_j(\mathsf{M})\,.$$

Hence ISTShow that

$$Q[j]\textbf{:} \qquad R_j(\mathsf{M}) \textit{ annihilates } \mathbf{E}_j\,.$$

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

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

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

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

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

*Proof of C-H using JCF.* We now handle a general $\mathsf{M}$ by means of ***JCF***, the Jordan Canonical Form thm. Let $\mathbb{G}$ denote the algebraic closure of $\mathbf{F}$. Viewing $\mathsf{M}$ as acting on $\mathbb{G}^{\times N}$, our $\mathsf{M}$ 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. ♦

**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 $\mathsf{T}$ on a finite-dim'al $\mathbf{F}$-vectorspace and let $\wp$ be its characteristic poly. Fixing a vector $\mathbf{v}_0 \neq \mathbf{0}$, our goal is to show that

$$\textbf{3:} \qquad \wp(\mathsf{T})(\mathbf{v}_0) \textit{ equals } \mathbf{0}.$$

Exer: *Why does this suffice?*

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

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

Define coeffs $\alpha_j$ by

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

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

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

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

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

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

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

*And its characteristic polynomial is*

5b: $$\wp_{\mathsf{M}}(x) \Leftrightarrow x^{N+1} - \sum_{\substack{j+k=N, \\ \text{with } j,k \in \mathbb{N}}} \alpha_j x^k. \qquad \diamondsuit$$

*Remark.* A matrix of form (5a) is a ***companion matrix***. It is "the companion matrix of polynomial (5b)". Wikipedia has a nice write-up. □

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

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

Now, in columns $0, 1, \ldots N-1$ we either choose "$x$" or "-1". In a column where we choose -1, the <u>row</u> 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 $\mathrm{Det}(x\mathbf{I} - \mathsf{M})$ 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 $\wp_{\mathsf{M}}(x)$ is

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

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

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

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

**7: Application.** *Let $\mathbf{F} := \mathbb{Z}_p$, where $p$ is prime. To produce a $p \times p$ $\mathbf{F}$-matrix $\mathsf{M}$ with <u>no</u> $\mathbf{F}$-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* $\beta := \text{-}1$. *Then*

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

*using the notation of* (5b), *where* $\alpha_0 := 0$, $\alpha_1 := 1$ *and* $\alpha_2 := 1$. *Courtesy our* (5a), *then, matrix* $\mathsf{M} := \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ *has no eigenvalues in* $\mathbb{Z}_3$. ♢

## 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* $\mathbf{F}$. ♢

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

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

for scalars $\alpha_j \in \mathbb{G}$ and col-vecs in $\vec{\mathbf{c}}_j \in \mathcal{C}$. Some $\alpha_j \neq 0$, so WLOG $\alpha_1 \neq 0$. By multiplying (†) 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 $\operatorname{Proj}: \mathbb{G} \to \mathbf{F}$ by $1 \mapsto 1$ and, for each $\mathbf{e} \in \mathcal{E}$, have Proj send $\mathbf{e} \mapsto 0$. Whence Proj() is the identity on $\mathbf{F}$, and for $\alpha, \beta \in \mathbb{G}$ and $f \in \mathbf{F}$:

$$\operatorname{Proj}(\alpha + \beta) = \operatorname{Proj}(\alpha) + \operatorname{Proj}(\beta);$$
$$*: \qquad \operatorname{Proj}(\alpha \cdot f) = \operatorname{Proj}(\alpha) \cdot f.$$

Applying map $\operatorname{Proj}^{\times N}: \mathbb{G}^{\times N} \to \mathbf{F}^{\times N}$ to (†) yields

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

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

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

### Minimal poly of M

*See* `jordan_decomp.latex` *for theorems used below.* The ***minimal polynomial*** of an $\mathbf{F}$-matrix $\mathsf{M}$ is the smallest-degree monic $\mathbf{F}$-poly $\Upsilon()$ such that $\Upsilon(\mathsf{M}) = \mathbf{0}$. Applying (8) to collection $\mathcal{C} := \{\mathsf{M}^j\}_{j \in \mathbb{N}}$ 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

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

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

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

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

$$\wp_{\mathsf{J}}(x) \Leftarrow\!\!\!\Rightarrow \Upsilon_{\mathsf{J}}(x) = [x - \lambda]^D.$$

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

$$\wp_{\mathsf{M}} = \wp_{\mathsf{A}} \cdot \wp_{\mathsf{B}} \qquad ;$$
$$\Upsilon_{\mathsf{M}} = \operatorname{Lcm}(\Upsilon_{\mathsf{A}}, \Upsilon_{\mathsf{B}}) \qquad ,$$

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

*Caveat.* Suppose M is *block upper-triangular*; it has square-blocks $\mathsf{B}_1, \ldots, \mathsf{B}_L$ along the diagonal, zeros south-west of these blocks, and possibly non-zero values north-east of these blocks. Certainly

$$\wp_{\mathsf{M}} = \wp_{\mathsf{B}_1} \cdot \wp_{\mathsf{B}_2} \cdot \ldots \cdot \wp_{\mathsf{B}_L} .$$

However, the corresponding stmt for $\Upsilon_{\mathsf{M}}$ with Lcm is **false**.

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

An eigenvalue is a "***simple*** eigenvalue" if its eigenspace is 1-dim'al.

**10: Coro.** *A block-diagonal* M *has "equality"* $\wp_{\mathsf{M}} \doteqdot \Upsilon_{\mathsf{M}}$ *IFF* M *has only simple eigenvalues.* ♢

*Proof.* This follows from either (9) or (12). ♦

A ***downtup*** $\vec{D}$ is a sequence of positive integers $D_1 \geqslant D_2 \geqslant \ldots \geqslant D_{\mathcal{E}}$. It yields the JCF

$$\lambda\text{-JB}(\vec{D}) := \mathit{Diag}\Big(\lambda\text{-JB}(D_1), \ldots, \lambda\text{-JB}(D_{\mathcal{E}})\Big)$$

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

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

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

is the JCF of M.

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

$$\wp_{\mathsf{M}}(x) \doteqdot \prod_{\ell=1}^{L} [x - \lambda_\ell]^{\text{Size}(\vec{D}^\ell)} \quad \text{and}$$
$$\Upsilon_{\mathsf{M}}(x) = \prod_{\ell=1}^{L} [x - \lambda_\ell]^{\text{Max}(\vec{D}^\ell)} .$$

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

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

$$\wp_{\mathsf{M}}(x) \doteqdot [x - \lambda_1] \cdot \ldots \cdot [x - \lambda_N]$$
$$\Upsilon_{\mathsf{M}}(x) = [x - \beta_1] \cdot \ldots \cdot [x - \beta_D] .$$

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

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

*and analogously for the min-poly.* ♢

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

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

The JCF (Jordan Canonical Form) of M is J.[♡1] So the min-poly $\Upsilon_{\mathsf{M}}(x) = \Upsilon_{\mathsf{J}}(x) = x^7$ . But as $\beta \to 0$, our $\mathsf{M}_\beta$ 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* $\mathsf{S}, \mathsf{T} \in \text{MAT}_{N\times N}(\mathbf{F})$*: Products* ST *and* TS *have the same characteristic poly.* ♢

[♡1] INDIRECT: The only M eval is 0, yet $\text{Rank}(\mathsf{M}) = 7 - 1$. So the nullspace, i.e 0-espace, is only 1-dimensional, hence JCF(M) has only one JB.

DIRECT: Consider ordered-basis $\mathcal{V} := (\mathbf{v}_1, \ldots, \mathbf{v}_7)$, 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, $\mathsf{J} = \mathsf{C}^{-1}\mathsf{MC}$ where C is the diagonal matrix with entries $\beta^6, \beta^5, \ldots, \beta, 1$.

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

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

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

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

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

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

Filename: `Problems/Algebra/LinearAlg/cayley_hamilton.latex`
As of: *Tuesday 27May2003.* Typeset: *17Jul2016 at 20:03.*