# THE CAYLEY–HAMILTON THEOREM

This writeup begins with a standard proof of the Cayley–Hamilton Theorem, found in many books such as Hoffman and Kunze's linear algebra text, Jacobson's *Basic Algebra I*, and Serge Lang's *Algebra*. Then the writeup gives Paul Garrett's intrinsic reconstitution of the argument using multilinear algebra. Garrett's argument can be found in his algebra text (available in print and online at his website) and also in a separate writeup at his website.

## 1. Statement

**Theorem 1.1** (Cayley–Hamilton)**.** *Let $k$ be a field, let $V$ be a finite-dimensional vector space over $k$, and let $T$ be an endomorphism of $V$. Thus the characteristic polynomial of $T$ is*

$$f_T(x) = \det(x \cdot 1_V - T).$$

*Then*

$$f_T(T) = 0.$$

*That is, $T$ satisfies its own characteristic polynomial.*

## 2. A First Comment

A tempting-but-invalid approach to the Cayley–Hamilton theorem is to write

$$\text{“}f_T(T) = \det(T \cdot 1_V - T) = \det(T - T) = \det(0_V) = 0.\text{”}$$

However, the symbol-string is not justified: although $f_T(T)$ does connote substituting the endomorphism $T$ for the indeterminate $x$, the substitution does not imply taking the determinant of 0. One way to understand the substitution is first to expand the polynomial $\det(x \cdot 1_V - T)$ and only then replace $x$ by $T$. For example, if $n = 2$ and we choose a basis of $V$ then $T$ acquires a matrix $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, and

$$\det(x \cdot 1_V - T) = \det \begin{bmatrix} x - a & -b \\ -c & x - d \end{bmatrix}.$$

To take the determinant and then substitute the matrix of $T$ for $x$ is to compute

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 - (a + d) \begin{bmatrix} a & b \\ c & d \end{bmatrix} + (ad - bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

On the other hand, to substitute the matrix of $T$ for $x$ and then take the determinant initially seems to require ascribing some meaning to the quantity

$$\det \begin{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} - a & -b \\ -c & \begin{bmatrix} a & b \\ c & d \end{bmatrix} - d \end{bmatrix}.$$

But actually, a moment earlier when we took the determinant and then afterwards substituted the matrix of $T$ for $x$, we also substituted $I_2$ for 1. So here now, we similarly should scale the original matrix entries by $I_2$ before taking the determinant. We treat the determinant as the determinant of a two-by-two matrix of two-by-two matrices rather than the determinant of a four-by-four matrix. Doing so again gives $0_{2\times 2}$ in a nontrivial way,

$$\left|\begin{matrix} \begin{bmatrix} a & b \\ c & d\end{bmatrix} - aI_2 & -bI_2 \\ -cI_2 & \begin{bmatrix} a & b\\ c & d\end{bmatrix} - dI_2 \end{matrix}\right| = \left|\begin{matrix} \begin{bmatrix} 0 & b \\ c & d-a\end{bmatrix} & \begin{bmatrix} -b & 0 \\ 0 & -b\end{bmatrix} \\ \begin{bmatrix} -c & 0 \\ 0 & -c \end{bmatrix} & \begin{bmatrix} a-d & b \\ c & 0\end{bmatrix}\end{matrix}\right| = \begin{bmatrix} 0 & 0\\ 0 & 0\end{bmatrix}.$$

So, although the argument that the Cayley–Hamilton Theorem is a triviality because "$f_T(T) = \det(T - T) = \det(0_V) = 0$" is visibly wrong, the second method illustrated here feels so natural that something in the spirit of the incorrect argument really should work.

## 3. The Standard "Determinantal" Proof

Take an ordered basis $(v_1, \cdots, v_n)$ of $V$. Let $A$ denote the $n$-by-$n$ matrix of $T$ with respect to this basis, so that

$$T(v_i) = \sum_j a_{ij} v_j, \quad i = 1, \cdots, n.$$

Also, let $x$ be an indeterminate and give $V$ the structure of a $k[x]$-module by the rule

$$f(x)v = f(T)(v), \quad f(x) \in k[x],\ v \in V.$$

The action of a constant polynomial on the $k[x]$-module $V$ is simply scalar multiplication of the $k$-vector space $V$ by the constant. Thus the $k[x]$-action on $V$ gives the equations

$$\left\{\begin{array}{rlrrl} (x-a_{11})v_1 & -\,a_{12}\ v_2 - \cdots & & -\,a_{1n}\ v_n & = 0_V \\ -\,a_{21}\ v_1 & +\,(x-a_{22})v_2 - \cdots & & -\,a_{2n}\ v_n & = 0_V \\ \vdots & \quad\vdots & & \vdots & \ \ \vdots \\ -\,a_{n1}\ v_1 & -\,a_{n2}\ v_2 - \cdots & + & (x-a_{nn})v_n & = 0_V \end{array}\right\}.$$

Let $\mathrm{M}_n(k[x])$ denote the ring of $n$-by-$n$ matrices with entries in $k[x]$. The $k[x]$-module structure of $V$ extends to a $\mathrm{M}_n(k[x])$-module structure of the cartesian product $V^n$, under which the identities in the previous display gather in the form *matrix of polynomials times vector of vectors equals vector of vectors*,

$$(xI - A)\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} 0_V \\ 0_V \\ \vdots \\ 0_V \end{bmatrix}.$$

Let $(xI - A)^{\mathrm{adg}} \in \mathrm{M}_n(k[x])$ be the *adjugate matrix* or *cofactor matrix* or *classical adjoint matrix* of $xI - A$,

$$(xI - A)^{\mathrm{adg}}_{ij} = (-1)^{i+j}\det(xI - A)^{(ji)}, \quad i, j \in \{1, \cdots, n\}$$

(here $(xI - A)^{(ji)}$ is $xI - A$ with its $j$th row and $i$th column removed). Now carry out a calculation that begins with an $n$-vector whose entries are instances of the

$k[x]$-action on $V$, proceeds through steps that involve the $\mathrm{M}_n(k[x])$-action on $V^n$, intertwining the coordinates, and ends with an $n$-vector of $V$-values,

$$\begin{aligned}\begin{bmatrix}\det(xI-A)v_1\\ \det(xI-A)v_2\\ \vdots\\ \det(xI-A)v_n\end{bmatrix} &= \det(xI-A)I\begin{bmatrix}v_1\\ v_2\\ \vdots\\ v_n\end{bmatrix}\\ &= \Big((xI-A)^{\mathrm{adg}}(xI-A)\Big)\begin{bmatrix}v_1\\ v_2\\ \vdots\\ v_n\end{bmatrix}\\ &= (xI-A)^{\mathrm{adg}}\left((xI-A)\begin{bmatrix}v_1\\ v_2\\ \vdots\\ v_n\end{bmatrix}\right) = (xI-A)^{\mathrm{adg}}\begin{bmatrix}0_V\\ 0_V\\ \vdots\\ 0_V\end{bmatrix} = \begin{bmatrix}0_V\\ 0_V\\ \vdots\\ 0_V\end{bmatrix}.\end{aligned}$$

Because the first and last quantities are equal entrywise, the action of $\det(xI - A)$ on $V$ annihilates the basis $(v_1, \cdots, v_n)$, and so it annihilates all of $V$. Since $A$ is the matrix of $T$ with respect to the basis, this action is precisely the action of the endomorphism $f_T(T)$ of $V$ where $f_T(x)$ is the characteristic polynomial of $T$. (We have not bothered to give the standard linear algebra argument that $f_T(x)$ is independent of coordinates.) Thus $f_T(T)$ is the zero-endomorphism of $V$ as desired, and the proof is complete.

The argument is perhaps clear in the sense of being easy to follow step by step, but from the standpoint of algebraic structure it does not make clear at all (at least, not to the author of this writeup) what is happening.

- In the system of equations, $T$ seems to be acting in two different ways, and although the system looks like a matrix-by-vector multiplication, the entries of the vector are themselves vectors rather than scalars. For that matter, the usual definition of the matrix of $T$ with respect to the basis $(v_1, \cdots, v_n)$ is not our matrix $A$ but rather its transpose. Had we not snuck the transpose in silently at the beginning of the argument, it would have appeared later without explanation when we gathered together the system of equations using the $\mathrm{M}_n(k[x])$-action on $V^n$.
- The argument moves from $V$ up to a larger vector-of-vectors environment $V^n$ and then back down to $V$ in a way compatible with algebraic structure. Why does everything fit together and work?
- Coordinates are present in the form of the basis of $V$ and the cartesian product $V^n$. The adjugate matrix is very complicated in coordinates, and the proof in coordinates that $m^{\mathrm{adg}}m = \det(m)I$ is the sort of thing that, once mongered, nobody wants to think about again, ever.

Working in coordinates interferes with structural understanding. An incisive argument should set the appropriate intrinsic environment where a variant of the desirable symbol-string approach will work. As in the coordinate-based proof, $T-T$ should appear as a factor of $f_T(T)$, giving the result. An indication of how to proceed is supplied by a matrix that we saw earlier,

$$\begin{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} - aI_2 & -bI_2 \\ -cI_2 & \begin{bmatrix} a & b \\ c & d \end{bmatrix} - dI_2 \end{bmatrix}.$$

This is the tensor $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \otimes I_2 - I_2 \otimes \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ in coordinates. Thus, our method should be intrinsic multilinear algebra.

## 4. The Intrinsic Proof

While the commutative $k$-algebra $k[T]$ acts on the $k$-vector space $V$, the quantity

$$\text{“}x \cdot 1_V - T\text{,”}$$

with no value substituted for $x$, is sensible only as an element of the commutative $k[x]$-algebra obtained by extending the scalars of the original $k$-algebra $k[T]$ via the tensor product,

$$x \otimes 1_V - 1 \otimes T \quad \text{lies in} \quad k[x] \otimes_k k[T].$$

(Exercise: the tensor product of commutative $k$-algebras is again a commutative $k$-algebra in the only sensible way.) Give the tensor product a name to emphasize its ring structure,

$$R = k[x] \otimes_k k[T],$$

and similarly extend the scalars of the original vector space being acted on to create a free $k[x]$-module of rank $n = \dim_k(V)$,

$$M = k[x] \otimes_k V.$$

The extension of scalars naturally augments the action of $k[T]$ on $V$ to an action of $R$ on $M$,

$$R \times M \longrightarrow M,$$

specifically (giving only the action of monomials on monomials),

$$(f(x) \otimes h(T))(g(x) \otimes v) = f(x)g(x) \otimes h(T)v.$$

From general multilinear algebra, the action of $R$ on $M$ induces an adjugate action on $M$ as well. For each $r \in R$, the adjugate $r^{\text{adg}}$ need not lie in the subring $R$ of $\text{End}_{k[x]}M$, nor need it commute with $R$.

Introduce a name for the particular ring element of interest, to be used for the duration of the argument,

$$y = x \otimes 1_V - 1 \otimes T.$$

Since the characteristic polynomial of $T$ is $f_T(x) = \det(x \otimes 1_V - 1 \otimes T) = \det y$, general multilinear algebra says that

$$y^{\text{adg}}y \text{ acts as multiplication by } f_T(x) \text{ on } M.$$

And since $k[x]$ is an integral domain and $f_T(x) \neq 0$ in $k[x]$, general multilinear algebra also says that $y^{\text{adg}}$ commutes with $y$. Since $y$ is $x \otimes 1_V - 1 \otimes T$ and $x \otimes 1_V$ is central in $\text{End}_{k[x]}M$, it follows that $y^{\text{adg}}$ commutes with $1 \otimes T$, and hence $y$ has a property that does not hold for general elements of $R$,

$$y^{\text{adg}} \text{ commutes with all of } R.$$

Consider the ideal of $R$ generated by $y$,

$$I = yR.$$

The resulting quotient ring acts on the corresponding quotient module,

$$R/I \times M/IM \longrightarrow M/IM.$$

And because $y^{\mathrm{adg}}$ commutes with $R$, we have

$$y^{\mathrm{adg}}IM = Iy^{\mathrm{adg}}M \subset IM,$$

so that the action of $y^{\mathrm{adg}}$ on $M$ descends to an action on $M/IM$ as well. In the quotient ring $R/I$ we have for any polynomial $f(x) \in k[x]$,

$$f(x) \otimes 1_V + I = 1 \otimes f(T) + I.$$

That is:

*Working in $R/I$ is tantamount to substituting $T$ for $x$ in polynomials.*

Also, the calculation that for any $f(x) \otimes v \in M$,

$$\begin{aligned} f(x) \otimes v + IM &= (f(x) \otimes 1_V)(1 \otimes v) + IM \\ &= (1 \otimes f(T))(1 \otimes v) + IM = 1 \otimes f(T)v + IM \end{aligned}$$

shows that the quotient module is

$$M/IM \approx 1 \otimes V.$$

That is:

*Working in $M/IM$ is tantamount to working in the original $k$-vector space $V$.*

Now the Cayley–Hamilton argument is very quick. The fact that $y^{\mathrm{adg}}y$ acts as multiplication by $f_T(x)$ on $M$ is expressed as the relation

$$f_T(x) \otimes 1_V = y^{\mathrm{adg}}y \quad \text{(equality of endomorphisms of } M\text{)},$$

which descends to the condition

$$f_T(x) \otimes 1_V + I \text{ annihilates } M/IM,$$

or, by the nature of $R/I$ as just explained,

$$1 \otimes f_T(T) + I \text{ annihilates } M/IM.$$

Since $M/IM \approx 1 \otimes V$, this gives the result,

$$f_T(T) \text{ annihilates } V.$$

Strikingly, only at this last step of the proof of the Cayley–Hamilton theorem do we work with a $k$-endomorphism, as compared to $k[x]$-endomorphisms.

As mentioned earlier, the presence of $y = x \otimes 1_V - 1 \otimes T$ as a factor of $f_T(x) \otimes 1_V$ is the crux of the matter, making the reasoning of the previous paragraph the correct version of the naïve symbol-string argument from the beginning of the writeup.

Once one has taken in the intrinsic argument, one sees that it simply is the determinantal argument placed in a structurally coherent context where it is uncluttered by coordinates.