

THE CAYLEY HAMILTON AND FROBENIUS THEOREMS VIA THE LAPLACE TRANSFORM

WILLIAM A. ADKINS AND MARK G. DAVIDSON

ABSTRACT. The Cayley Hamilton theorem on the characteristic polynomial of a matrix A and Frobenius' theorem on minimal polynomial of A are deduced from the familiar Laplace transform formula $\mathcal{L}(e^{At}) = (sI - A)^{-1}$. This formula is extended to a formal power series ring over an algebraically closed field of characteristic 0, so that the argument applies in the more general setting of matrices over a field of characteristic 0.

A traditional use of the structure theory of a linear operator has been the calculation of the matrix exponential e^{At} and its use in solving the system of differential equations $\mathbf{y}' = A\mathbf{y}$, when A is a real or complex $n \times n$ matrix [2]. Zieber [6] and Schmidt [5] reversed this process and applied knowledge of the basic form of e^{At} to a derivation of the main results on the structure of A as a linear operator. Their approach started with an application of the Cayley Hamilton theorem to deduce the form of each entry of e^{At} as a solution of a constant coefficient linear differential equation. Since the Cayley Hamilton theorem can be viewed as part of the structure theory of a linear operator, it seems natural to ask if, by means of a different starting point for the analysis of e^{At} , one can also deduce this result from information about e^{At} . It turns out that the Laplace transform formula [4],

$$(1) \quad \mathcal{L}(e^{At}) = (sI - A)^{-1}$$

provides such a starting point. This formula applies to real and complex matrices A , but if k is any algebraically closed field of characteristic 0, one can define a Laplace transform on a k -linear subspace of the ring of formal power series $k[[t]]$ in a natural manner so that equation (1) remains valid. We will present our proof of the Cayley Hamilton theorem in this context. Moreover, this approach extends to a simple proof of the Frobenius characterization of the minimal polynomial of

2000 *Mathematics Subject Classification.* 15A24, 15A54.

Key words and phrases. Cayley Hamilton theorem, Frobenius theorem, Laplace transform, Matrix exponential, characteristic polynomial, minimal polynomial, formal power series.

A which is traditionally deduced from the invariant factors of A ([1], Page 315, [3], Page 201).

If k is a field, which we will assume to be algebraically closed of characteristic 0, then $k(s)$ denotes the rational functions in the indeterminate s , $k^0(s)$ the proper rational functions, that is, the rational functions where the degree of the numerator is less than the degree of the denominator, and $k[[t]]$ the formal power series over k in the indeterminate t . The space $\mathcal{E}[[t]]$ is the k -linear subspace of $k[[t]]$ generated by the formal power series $\varphi_{n,a}(t) = t^n e^{at}$ for $n = 0, 1, \dots$, and $a \in k$, where $e^{at} = \sum_{m=0}^{\infty} \frac{1}{m!} a^m t^m$. The family of series $\varphi_{n,a}(t)$ are easily seen to be linearly independent, and hence are a basis for $\mathcal{E}[[t]]$, which can be thought of (somewhat loosely) as a formal analog of the functions of exponential growth on which the Laplace transform is commonly defined.

Since a proper rational function can be expanded in partial fractions, a k -basis of $k^0(s)$ is $\mathcal{B} = \{1/(s-a)^n : a \in k, n = 1, 2, \dots\}$. We observe that the bijective correspondence $\varphi_{n,a}(t) \longleftrightarrow 1/(s-a)^{n+1}$, where $a \in k$ and $n = 0, 1, \dots$, establishes a linear isomorphism between $\mathcal{E}[[t]]$ and $k^0(s)$. This linear isomorphism will be modified slightly in the following definition to conform to the traditional formulas for the Laplace transform.

Definition 1. Define the *formal Laplace transform* $\mathcal{L} : \mathcal{E}[[t]] \rightarrow k^0(s)$ to be the k -linear map determined by the formula

$$\mathcal{L}(t^n e^{at}) = \left(\frac{n!}{(s-a)^{n+1}} \right).$$

The *formal inverse Laplace transform* \mathcal{L}^{-1} is given on the basis \mathcal{B} by

$$\mathcal{L}^{-1} \left(\frac{1}{(s-a)^{n+1}} \right) = \frac{t^n}{n!} e^{at}.$$

If $A(t) \in M_n(\mathcal{E}[[t]])$ is an $n \times n$ matrix of formal power series, then $\mathcal{L}(A(t)) \in M_n(k^0(s))$ is the $n \times n$ matrix of proper rational functions obtained by applying \mathcal{L} to each entry of the matrix $A(t)$. There is a similar interpretation of $\mathcal{L}^{-1}(B(s))$ for $B(s) \in M_n(k^0(s))$.

Note that if $A \in M_n(k)$ then $e^{At} = \sum_{m=0}^{\infty} \frac{1}{m!} A^m t^m$ is the unique matrix $\Phi(t) \in M_n(k[[t]])$ which solves the initial value problem

$$(2) \quad \Phi'(t) = A\Phi(t), \quad \Phi(0) = I_n.$$

Our goal is to show that the individual entries of e^{At} actually lie in the subspace $\mathcal{E}[[t]]$ so that e^{At} can be expressed as

$$(3) \quad e^{At} = \sum_{i=1}^m \sum_{j=0}^{n_i-1} M_{ij} t^j e^{a_i t} = \sum_{i=1}^m \sum_{j=0}^{n_i-1} M_{ij} \varphi_{j, a_i}(t)$$

where $M_{ij} \in M_n(k)$. This formula was derived by Ziebur and Schmidt using the Cayley Hamilton theorem. Our observation is that this formula also follows from equation (1). If $k = \mathbb{C}$, equation (1) is a well known property of the Laplace transform, but for an arbitrary k , it is necessary to derive equation (1) from the definition of the formal Laplace transform. The properties of the formal Laplace transform \mathcal{L} that we need are delineated in the following result.

Proposition 2.

(L1) $\mathcal{L}(y'(t)) = s\mathcal{L}(y(t)) - y(0)$ for all $y(t) \in \mathcal{E}[[t]]$.

(L2) If $r(s) = q(s)/p(s) \in k^0(s)$ is a rational function with $p(s)$ monic of degree n , then $\mathcal{L}^{-1}(r(s))(0)$ is the coefficient of s^{n-1} in $q(s)$.

(L3) If $A \in M_n(k)$ then

$$\mathcal{L}^{-1}((sI - A)^{-1}) = e^{At}.$$

Proof. (L1) It is sufficient to check this for $y(t) = \frac{t^n}{n!} e^{at}$. If $n \geq 1$, then $y'(t) = \frac{t^{n-1}}{(n-1)!} e^{at} + a \frac{t^n}{n!} e^{at}$ and

$$\mathcal{L}(y'(t)) = \frac{1}{(s-a)^n} + \frac{a}{(s-a)^{n+1}} = \frac{s}{(s-a)^{n+1}} = s\mathcal{L}(y(t)) - y(0).$$

If $n = 0$ then $y' = ae^{at}$ so

$$\mathcal{L}(y'(t)) = \frac{a}{s-a} = \frac{s}{s-a} - 1 = s\mathcal{L}(y(t)) - y(0).$$

(L2) From definition (1) the constant term of the series $\mathcal{L}^{-1}\left(\frac{1}{(s-a)^{n+1}}\right)$ is 0 unless $n = 0$, in which case the constant term is 1. The result follows from the fact that if $r(s)$ is expanded in partial fractions

$$r(s) = \sum_{i=1}^m \sum_{j=1}^{r_j} \frac{c_{ij}}{(s-a_i)^j},$$

the coefficient of s^{n-1} in $q(s)$ is $c_{11} + c_{21} + \cdots + c_{m1}$.

(L3) Let $p(s) = \det(sI - A)$ be the characteristic polynomial. If $B(s) = (sI - A)^{-1}$ then the adjoint formula applied to $sI - A$ shows that each entry of $B(s)$ is a rational function and the entries have the form $c_{\mu\nu}(s)/p(s)$ where the degree of $c_{\mu\nu}(s)$ is less than $n - 1$ if $\mu \neq \nu$, while $c_{\mu\mu}(s)$ is a monic polynomial of degree $n - 1$. Then (L2) shows that $\Phi(t) = \mathcal{L}^{-1}(B(s))$ has $\Phi(0) = I_n$. Since $(sI - A)B(s) = I$ it follows that $sB(s) - I = AB(s)$. But this equation, combined with the differentiation rule (L1), states that $\Phi(t)$ is a solution of the differential equation (2). Hence

$$\mathcal{L}^{-1}((sI - A)^{-1}) = \Phi(t) = e^{At}.$$

□

Equation (3) is now an immediate consequence of Proposition (2) (L3), and the fact that $(sI - A)^{-1} \in M_n(k^0(s))$.

Theorem 3 (Cayley Hamilton). *If $p(s)$ is the characteristic polynomial of A then $p(A) = 0$.*

Proof. The characteristic polynomial of A , $p(s) = \prod_{i=1}^m (s - a_i)^{n_i}$, determines the series $\varphi_{j,a_i}(t) = t^j e^{a_i t}$ which appear in (3) since $p(s)$ is the denominator of every entry of $(sI - A)^{-1}$. If Equation (3) is differentiated l times and then evaluated at 0 we get

$$(4) \quad A^l = \sum_{i=1}^m \sum_{j=0}^{n_i-1} M_{ij} \varphi_{j,a_i}^{(l)}(0).$$

Expanding $\varphi_{j,a_i}(t) = t^j e^{a_i t}$ in a series, observing that the coefficient of t^l is 0 if $l < j$, $\frac{1}{(l-j)!} a_i^{l-j}$ if $l \geq j$, and noting that $\varphi_{l,0}(t) = t^l$ gives

$$(5) \quad \varphi_{j,a_i}^{(l)}(0) = \begin{cases} 0 & \text{if } l < j \\ \frac{l!}{(l-j)!} a_i^{l-j} & \text{if } l \geq j \end{cases} = \varphi_{l,0}^{(j)}(a_i).$$

If $q(t) = c_0 + c_1 t + \dots + c_N t^N$ is an arbitrary polynomial, then Equations (4) and (5) give

$$(6) \quad \begin{aligned} q(A) &= \sum_{l=0}^N c_l A^l = \sum_{i=1}^m \sum_{j=0}^{n_i-1} \sum_{l=0}^N c_l M_{ij} \varphi_{j,a_i}^{(l)}(0) \\ &= \sum_{i=1}^m \sum_{j=0}^{n_i-1} M_{ij} \sum_{l=0}^N c_l \varphi_{l,0}^{(j)}(a_i) = \sum_{i=1}^m \sum_{j=0}^{n_i-1} M_{ij} q^{(j)}(a_i). \end{aligned}$$

Since the characteristic polynomial $p(s) = \prod_{i=1}^m (s - a_i)^{n_i}$, has $p^{(j)}(a_i) = 0$ for $0 \leq j \leq n_i - 1$, Equation (6) gives

$$p(A) = \sum_{i=1}^m \sum_{j=0}^{n_i-1} M_{ij} p^{(j)}(a_i) = 0.$$

□

Thus the Cayley Hamilton theorem is a consequence of the Laplace transform formula $e^{At} = \mathcal{L}^{-1}((sI - A)^{-1})$, and in conjunction with the previously observed work of Zeibur and Schmidt, we have the observation that the structure theory of a linear endomorphism follows from the basic existence theorem for systems of ordinary differential equations.

As noted above, the characteristic polynomial of A , $p(s) = \prod_{i=1}^m (s - a_i)^{n_i}$, determines the series $\varphi_{j,a_i}(t) = t^j e^{a_i t}$ which appear in (3) since $p(s)$ is a denominator of every entry of $(sI - A)^{-1}$, and hence it determines the constant matrices M_{ij} which appear in equation (6). If we refine the common denominator of the rational functions in $(sI - A)^{-1}$, then we get a refined version of equations (3) and (6). This observation allows for a proof of Frobenius's characterization of the minimal polynomial of A along the same lines as the proof of Cayley Hamilton given above.

Theorem 4 (Frobenius). *If $A \in M_n(k)$, $p(s) = \det(sI - A)$, $\Delta(s)$ is the greatest common divisor (in $k[s]$) of the $(n - 1) \times (n - 1)$ minors of $sI - A$, and $f(s) = p(s)/\Delta(s)$, then $f(s)$ is the minimal polynomial of A .*

Proof. As in the proof of (L3) note that each entry of $(sI - A)^{-1}$ is a rational function $c_{\mu\nu}(s)/p(s)$ where $c_{\mu\nu}(s)$ is \pm an $(n - 1) \times (n - 1)$ minor of $sI - A$. Then

$$(*) \quad \frac{c_{\mu\nu}(s)}{p(s)} = \frac{c_{\mu\nu}(s)/\Delta(s)}{p(s)/\Delta(s)} = \frac{b_{\mu\nu}(s)}{f(s)},$$

where $\{b_{\mu\nu}(s) : 1 \leq \mu, \nu \leq n\}$ is a relatively prime set of polynomials. Since $f(s)$ divides $p(s) = \prod_{i=1}^m (s - a_i)^{n_i}$ we can write $f(s) = \prod_{i \in \Gamma} (s - a_i)^{r_i}$ where $\Gamma = \{i : 1 \leq i \leq m \text{ and } 1 \leq r_i \leq n_i\}$. Using the representation (*) for the entries of $(sI - A)^{-1}$, equation (3) becomes

$$(7) \quad e^{At} = \sum_{i \in \Gamma} \sum_{j=0}^{r_i-1} M_{ij} t^j e^{a_i t} = \sum_{i \in \Gamma} \sum_{j=0}^{r_i-1} M_{ij} \varphi_{j,a_i}(t)$$

where $M_{ij} \in M_n(k)$, and then using (6) exactly as in the proof of Cayley Hamilton, we conclude that

$$f(A) = \sum_{i \in \Gamma} \sum_{j=0}^{r_i-1} M_{ij} f^{(j)}(a_i) = 0.$$

Thus, the minimal polynomial $m_A(s)$ of A divides $f(s)$, and to show that $f(s) = m_A(s)$ it is sufficient to show that $f_\gamma(A) \neq 0$ for each $f_\gamma(s) = f(s)/(s - a_\gamma)$ with $\gamma \in \Gamma$. To see that $f_\gamma(A) \neq 0$ we argue as follows. Since $\{b_{\mu\nu}(s) : 1 \leq \mu, \nu \leq n\}$ is relatively prime, for each $\gamma \in \Gamma$ there is an index μ, ν for which $b_{\mu\nu}(a_\gamma) \neq 0$. If we write $f(s) = (s - a_\gamma)^{r_\gamma} g_\gamma(s)$ where $g_\gamma(a_\gamma) \neq 0$, then the partial fraction decomposition for $b_{\mu\nu}(s)/f(s)$ can be written in the form

$$(8) \quad \frac{b_{\mu\nu}(s)}{f(s)} = \sum_{j=0}^{r_\gamma-1} c_j (s - a_\gamma)^{j-r_\gamma} + \frac{h_\gamma(s)}{g_\gamma(s)}.$$

Multiplying by $f(s)$ and evaluating at a_γ gives $c_0 g_\gamma(a_\gamma) = b_{\mu\nu}(a_\gamma) \neq 0$ so the coefficient $c_0 \neq 0$. Hence the term $t^{r_\gamma-1} e^{a_\gamma t}$ appears in $\mathcal{L}^{-1}(b_{\mu\nu}(s)/f(s))$ with a nonzero coefficient so that the matrix $M_{\gamma, r_\gamma-1} \neq 0$ in the expansion (7). Computing $f_\gamma(A)$ from equation (6) using expansion (7) gives

$$f_\gamma(A) = \sum_{i \in \Gamma} \sum_{j=0}^{r_i-1} M_{ij} f_\gamma^{(j)}(a_i).$$

However, $f_\gamma^{(j)}(a_i) = 0$ except when $i = \gamma$ and $j = r_\gamma - 1$. Thus

$$f_\gamma(A) = M_{\gamma, r_\gamma-1} (r_\gamma - 1)! g_\gamma(a_\gamma) \neq 0,$$

and hence $m_A(s) = f(s)$. \square

Remark 5. (1) In the above proof the set $\Gamma = \{1 \leq i \leq m \text{ and } 1 \leq r_i \leq n_i\}$ indexes those roots a_i of $p(s)$ which are also roots of $f(s) = p(s)/\Delta(s)$. Of course, once we know that $f(s)$ is the minimal polynomial of A , then we know that $\Gamma = \{1 \leq i \leq m\}$, i.e., every root of the characteristic polynomial is also a root of the minimal polynomial. The proof above did not need that $\Gamma = \{1 \leq i \leq m\}$, but it is in fact easy to verify this directly in the spirit of the above analysis. To see this, note that since $\Delta(s)$ divides each element of the matrix $\text{Adj}(sI - A)$ it follows that $\det \text{Adj}(sI - A) = q(s)\Delta^n(s)$. Then taking determinants of the matrix equation $(sI - A)\text{Adj}(sI - A) = p(s)I$ gives $q(s)\Delta^n(s) = p(s)^{n-1}$, so that

$$(**) \quad q(s)\Delta(s) = f(s)^{n-1}$$

since $f(s) = p(s)/\Delta(s)$. If $p(a) = 0$ then $f(a) = 0$ or $\Delta(a) = 0$. In the first case we are done; in the second case (***) shows that $f(a) = 0$ also.

- (2) The argument we have given for the Cayley Hamilton and Frobenius theorems uses that k is (1) algebraically closed so that an explicit basis could be described for $k^0(s)$ which facilitates the description of \mathcal{L}^{-1} and (2) of characteristic 0 so that the series defining e^{At} makes sense and the formal differential equation $\mathbf{y}' = \mathbf{A}\mathbf{y}$, $\mathbf{y}(0) = \mathbf{y}_0$ has a unique solution. By embedding k in an algebraic closure \bar{k} , it is not necessary to assume that k is algebraically closed.

REFERENCES

- [1] William A. Adkins and Steven H. Weintraub, *Algebra: An Approach via Module Theory*, Springer-Verlag New York, 1992.
- [2] Morris W. Hirsch and Stephen Smale, *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, 1974.
- [3] Nathan Jacobson, *Basic Algebra I, Second Edition*, W. H. Freeman, 1985.
- [4] N. J. Pullman, *Matrix Theory and its Applications*, Marcel Dekker, 1976.
- [5] E. J. P. Georg Schmidt, An alternative approach to canonical forms of matrices, *American Mathematical Monthly*, **93**, (1986), 176-184.
- [6] Allen D. Ziebur, On determining the structure of A by analyzing e^{At} , *SIAM Review*, **12**, (1970), 98-102.

DEPARTMENT OF MATHEMATICS, LOUISIANA STATE UNIVERSITY, BATON ROUGE,
LA 70803

E-mail address: adkins@math.lsu.edu, davidson@math.lsu.edu