2. G. H. Brown, Jr., On Halley's variation of Newton's method, this MONTHLY, 84 (1977) 726–728.
3. M. Davis and B. Dawson, On global convergence of Halley's iteration formula, Numer. Math., 24 (1975) 133–135.
4. H. Ehrmann, Konstruktion und Durchführung von Iterationsverfahren höherer Ordnung, Arch. Rational Mech. Anal., 4 (1959) 65–88.
5. E. Hansen and M. Patrick, A family of root finding methods, Numer. Math., 27 (1977) 257–269.
6. E. Halley, Methodus Nova, Accurata & Facilis Inveniendi Radices Aequationum Quarumcumque Generaliter, Sine Praevia Reductione, Philos. Trans. Roy. Soc. London, 18 (1694) 136–148.
7. P. Henrici, Elements of Numerical Analysis, Wiley, New York, 1964.
8. ———, Essentials of Numerical Analysis, Wiley, New York, 1982.
9. ———, Applied and Computational Complex Analysis, Wiley, New York, 1974.
10. A. M. Ostrowski, Solution of Equations and Systems of Equations, Academic Press, New York, 1973.
11. J. Stoer and R. Bulirsch, Introduction to Numerical Analysis, Springer-Verlag, New York, 1980.
12. J. F. Traub, Iterative Methods for Solution of Equations, Prentice-Hall, 1964.

A CONVERSE FOR THE CAYLEY-HAMILTON THEOREM

CARMEN CHICONE, N.J. KALTON*, AND IRA J. PAPICK
Department of Mathematics, University of Missouri-Columbia, Columbia, MO 65211

The Cayley-Hamilton theorem asserts that every $n \times n$ matrix A satisfies its characteristic polynomial, $\det(\lambda I - A)$. This note deals with the problem of characterizing those polynomials for which the Cayley-Hamilton theorem holds. Informally stated our result is that the polynomials which a square matrix satisfies are precisely the multiples (in a ring of polynomials) of the characteristic polynomial.

Some notation is necessary to make a precise statement of our theorem. If X denotes the $n \times n$ matrix of indeterminants (x_{ij}), then it is apparent that $\det(X)$ is a polynomial in n^2-variables. Moreover, if $F(x_{ij})$ is any polynomial in n^2-variables with coefficients in a commutative ring R with identity, then by $F(X)$ we shall mean $F(x_{ij})$. We now restate our problem: Characterize those polynomials $F(X)$ having the property that every $n \times n$ matrix A with entries in R satisfies the polynomial $F(\lambda I - A)$. We shall call such polynomials $F(X)$ Cayley-Hamilton polynomials, and we may now state a precise converse of the Cayley-Hamilton theorem.

THEOREM 1. *Let R be an infinite (commutative) integral domain, and let $F(X)$ be a polynomial in n^2-variables with coefficients in R. Then $F(X)$ is a Cayley-Hamilton polynomial if and only if $F(X) = \det(X)G(X)$, where $G(X)$ is a polynomial in n^2-variables with coefficients in R.*

Before proceeding we make two observations. The "if" direction of the theorem is clear. For if A is an $n \times n$ matrix with entries in R, and if $F(X) = \det(X)G(X)$, then $F(\lambda I - A) = \det(\lambda I - A)G(\lambda I - A)$. Hence, by the Cayley-Hamilton theorem, A satisfies the polynomial $F(\lambda I - A)$. Secondly, the theorem is false without some assumptions on R. For example if $R = Z_2$, the field with two elements, define $F(X) = (x_{12} + x_{21})x_{11}x_{22}$. Then given

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

with $a_{ij} \in Z_2$, we see that

$$F(\lambda I - A) = (a_{12} + a_{21})(\lambda - a_{11})(\lambda - a_{22}).$$

Thus $F(\lambda I - A) \equiv 0$ unless A is either upper or lower triangular and not diagonal, and in this case

$$F(\lambda I - A) = (\lambda - a_{11})(\lambda - a_{22}) = \det(\lambda I - A).$$

*Supported in part by NSF grant MCS-8001852.

Hence, $F(X)$ is a Cayley-Hamilton polynomial, but the determinant is not a factor of $F(X)$.

The above example is a homogeneous polynomial; a more transparent inhomogeneous example over Z_2 is $F(X) = x_{12}^2 - x_{12}$. Similar examples can be created for any finite field. Moreover, the latter example can be used for any infinite Boolean ring, so our assumptions on R are indeed necessary. (Recall that a ring R is called Boolean if $a^2 = a$ for each $a \in R$. It is straightforward to see that any Boolean ring R is commutative, and if R is infinite, then it is not an integral domain.)

We shall soon present an elementary proof of our theorem. However, it is worthwhile and interesting at this time to point out how the result follows in the special case when R is an algebraically closed field from the beautiful Hilbert Nullstellensatz [**2**, Proposition 7.4], which for our purposes may be stated as follows:

If J is an ideal in a polynomial ring over an algebraically closed field, and if $F(X)$ is a polynomial which vanishes at each point of the intersection of all zeros of the polynomials in J, then some power of $F(X)$, $(F(X))^n$, is an element of J.

First we need a lemma. (This lemma is undoubtedly well known. However, we were unable to find a reference for it, so we shall include a sketch of the proof here.)

LEMMA 1. *If R is an integral domain, then $\det(X)$ is an irreducible element of $R[X]$, where $R[X]$ denotes the polynomial ring in n^2-variables with coefficients in R.*

Proof. Suppose $\det(X) = F(X)G(X)$, where $F(X), G(X) \in R[X]$. Suppose F is not independent of x_{11}. Since $\det(X)$ is a linear function of the elements of the first column of $X, \{x_{11}, x_{21}, \ldots, x_{n1}\}$, we conclude that G is independent of these variables. A similar argument applied to each row in turn shows that G is a constant and hence a unit.

Alternate proof of the theorem for R an algebraically closed field. Let J be the set of Cayley-Hamilton polynomials in $R[X]$. It is easily seen that J is an ideal in $R[X]$. To complete the proof we will prove that J is the principal ideal generated by $\det(X)$. Since $\det(X) \in J$, it suffices to show that $J \subseteq P = (\det(X))$. Note that since $R[X]$ is a unique factorization domain [**2**, Theorem 6.14], and $\det(X)$ is irreducible, the ideal P is prime. Also notice that if A is an $n \times n$ singular matrix with entries in R, and $F(X) \in J$, then $F(A) = 0$. To see this, observe that zero is an eigenvalue of $-A$, and hence zero is a root of $F(\lambda I + A)$. Whence, $F(A) = 0$. Now by the Hilbert Nullstellensatz we have $(F(X))^n \in P$, for some $n \in N$. But, since P is a prime ideal, $F(X) \in P$ as required.

We are now ready to prove the theorem. Actually, we will deduce the theorem as a corollary from the following general lemma:

LEMMA 2. *Let R be an infinite integral domain, $S = R[X]$, and $F(X,\lambda) \in S[\lambda]$. Set $D(X,\lambda)$ $= \det(\lambda I - X)$. Then $F(A,A) = 0$ for all $A \in M_n(R)$ ($n \times n$-matrices over R) if and only if $D(X,\lambda)$ is a factor of $F(X,\lambda)$ in $S[\lambda]$.*

Proof. Since the "if" direction follows directly from the Cayley-Hamilton theorem, we need only prove the other direction. Write $F(X,\lambda) = p_m(X)\lambda^m + \cdots + p_0(X)$, where $p_i(X) \in S$. By assumption we have $p_m(A)A^m + \cdots + p_0(A)I = 0$ for each $A \in M_n(R)$. It follows that

$$F(X,X) = p_m(X)X^m + \cdots + p_0(X)I = 0.$$

To see this, observe that each entry of the matrix $F(X,X)$ is a polynomial in n^2-variables that vanishes identically, and hence is the zero polynomial [**1**, Exercise 1–4]. Now, since $D(X,\lambda)$ is a monic polynomial in the variable λ, we can apply the division algorithm in $S[\lambda]$ to write

$$F(X,\lambda) = Q(X,\lambda)D(X,\lambda) + R(X,\lambda),$$

where $Q(X,\lambda)$, $R(X,\lambda) \in S[\lambda]$ and degree $(R(X,\lambda)) = r < n$ [**2**, Theorem 6.2]. The proof will be complete when we show that $R(X,\lambda) = 0$. To reach this end first note that $R(X,X) = 0$. This follows from the equation

$$F(X,X) = Q(X,X)D(X,X) + R(X,X),$$

since by the Cayley-Hamilton theorem $D(X,X) = 0$ and, as shown above, $F(X,X) = 0$. Thus if

$$R(X,\lambda) = u_r(X)\lambda^r + \cdots + u_0(X),$$

we have $u_r(X)X^r + \cdots + u_0(X)I = 0$.

We now claim that the characteristic polynomial, $D(X,\lambda) \in S[\lambda]$ of X has n distinct roots (in some splitting field of the field of fractions of S). Once this is shown, the proof will be complete, for the minimal polynomial of the matrix X will be degree n and hence we can conclude that $u_r(X) = \cdots = u_0(X) = 0$, i.e., $R(X,\lambda) = 0$.

To establish our claim we note that the *discriminant* $\Delta(X)$ of $D(X,\lambda)$ is a homogeneous polynomial in the coefficients of $D(X,\lambda)$, and $\Delta(X) = 0$ if and only if $D(X,\lambda)$ has a repeated root [3, p. 288]. If $A \in M_n(R)$, then $\Delta(A)$ is the discriminant of $D(A,\lambda)$; hence by selecting A to be diagonal with n distinct eigenvalues (R is infinite!), we see that $\Delta(X)$ cannot vanish identically. This completes the proof.

Now to derive the theorem from Lemma 2. Let $G(X)$ be a Cayley-Hamilton polynomial, and define $F(X,\lambda) = G(\lambda I - X)$. Then $F(A,A) = 0$ for each $A \in M_n(R)$, and thus it follows from Lemma 2 that $D(X,\lambda)$ is a factor of $F(X,\lambda)$, i.e., $F(X,\lambda) = D(X,\lambda)Q(X,\lambda)$. Hence, putting $\lambda = 0$ and replacing X by $-X$ we have

$$G(X) = \det(X)Q(X,0),$$

where

$$Q(X,0) \in R[X].$$

References

1. W. Fulton, Algebraic Curves, W. A. Benjamin, New York, 1969.
2. T. Hungerford, Algebra, Holt, Rinehart and Winston, New York, 1974.
3. J. Uspensky, Theory of Equations, McGraw-Hill, New York, 1948.

AN ELEMENTARY CHARACTERIZATION OF WEAK CONVERGENCE OF MEASURES

JERRY JOHNSON
Department of Mathematics, Oklahoma State University, Stillwater, OK 74078

In 1977, Högnäs [2] solved a problem of P. Lévy which, roughly put, asks for conditions on a family $\{K_s\}$ of functions so that for each continuous function f on $[0,1]$ we have continuity of $s \to \int_0^1 f\,dK_s$. The solution follows immediately from his main theorem, whose proof is fairly long. This note presents a short, elementary proof of his theorem, using only a few basic results from a standard real variables text such as Royden [3].

Our notation is as follows: C denotes the continuous functions on $[0,1]$ with supremum norm $\|\cdot\|_\infty$. If f and K are functions on $[0,1]$, $\int f\,dK$ will denote the Riemann-Stieltjes integral when it exists (see [1] for the definition and basic facts). λ will denote Lebesgue measure on $[0,1]$.

The following five items used in the proof are stated in the *least* possible generality, reflecting the assumptions in the theorem.

(1) The Arzela-Ascoli Theorem: An equicontinuous bounded subset of C is relatively compact.
(2) The Stone-Weierstrass Theorem: The polynomials are dense in C.
(3) If μ is a signed measure on $[0,1]$, $f \in C$ and $K(x) = \mu[0,x]$ for $0 \leqslant x \leqslant 1$, then K is of bounded variation and $\int f\,d\mu = K(0)f(0) + \int f\,dK$.
(4) Integration by parts: If either $\int f\,dK$ or $\int K\,df$ exists, so does the other and $\int f\,dK + \int K\,df = Kf|_0^1$ (see [1, p. 195]).