ON HILBERT SPACE OPERATORS AND OPERATOR ROOTS OF POLYNOMIALS

S. CATER

1. Introduction. The Cayley-Hamilton theorem states that any linear transformation (sometimes called "operator") on a finite dimensional vector space over a field is a root of its characteristic polynomial. On the other hand an operator on an infinite dimensional vector space need not be the root of any nonzero polynomial with scalar coefficients. It is our purpose to give necessary and sufficient conditions for a bounded operator on a complex Hilbert space to be the root of a nonzero polynomial with complex coefficients.

Significant in much linear algebra is the fact that an operator A on a finite dimensional vector space V over an algebraically closed field F must have an eigen value; more precisely, there is a scalar λ in F and a nonzero vector z in V such that $(A - \lambda)z = 0$. We make the following

DEFINITION. *An element λ in a field F is said to be an eigen value for the operator A on a (possibly infinite dimensional) vector space V over F if there exists at least one nonzero vector z in V such that $(A - \lambda)z = 0$. An operator A on V is said to be an eigen value producing (henceforth abbreviated "evp") operator if for each linear maniforld V' reducing A and $\neq V$, the operator A' induced by A on the quotient space V/V' has at least one eigen value.*

In particular if A is evp it has an eigen value, because (0) reduces A. One example of an evp operator is any operator on a finite dimensional vector space over an algebraically closed field. The central result of the present paper is that a bounded operator on a complex Hilbert space is evp if and only if it is the root of a nonzero polynomial with complex coefficients. Before we introduce Hilbert space operators we establish some algebraic machinery.

2. The structure of evp operators. In this section let A be an operator on a vector space V over the field F, and let V_λ be the linear manifold consisting of all vectors annihilated by some power of the operator $A - \lambda$, for each λ in F.

LEMMA 1. *Let λ and μ be distinct scalars and let z be a vector such that*

Received March 22, 1961, and in revised form April 25, 1961.

$$(A-\lambda)^{e+1}=0\neq(A-\lambda)^{e}z\ .$$

Then

$$(A-\lambda)^{e+1}(A-\mu)z=0\neq(A-\lambda)^{e}(A-\mu)z\ .$$

In particular if $\mu_1, \cdots, \mu_n$ is a finite set of scalars different from λ, then

$$(A-\lambda)^{e+1}w=0\neq(A-\lambda)^{e}w$$

where $w=(A-\mu_1)\cdots(A-\mu_n)z$.

Proof. Since A commutes with itself

$$(A-\lambda)^{e+1}(A-\mu)z=(A-\mu)(A-\lambda)^{e+1}z=0\ .$$

On the other hand
$$\begin{aligned}(A-\lambda)^{e}(A-\mu)z &= (A-\mu)(A-\lambda)^{e}z\\ &= (A-\lambda)^{e+1}z+(\lambda-\mu)(A-\lambda)^{e}z\\ &= (\lambda-\mu)(A-\lambda)^{e}z\neq 0\end{aligned}$$

because $\lambda\neq\mu$. The final statement follows by induction on n.

LEMMA 2. *The V_λ are linearly independent.*

Proof. Suppose V_λ contains a vector z which is also in the subspace spanned by $V_{\mu_1}, \cdots, V_{\mu_n}$ where $\mu_i\neq\lambda$. Say $z=z_1+\cdots+z_n$ where z_i is in V_{μ_i} and $(A-\mu_i)^{e_i}z_i=0$. Then

$$(A-\mu_1)^{e_1}\cdots(A-\mu_n)^{e_n}z=(A-\mu_1)^{e_1}\cdots(A-\mu_n)^{e_n}(z_1+\cdots+z_n)=0\ .$$

By Lemma 1, $z=0$.

LEMMA 3. *The V_λ together span V if and only if A is evp.*

Proof. Assume the V_λ span V. Let V' be a subspace of V reducing A and $\neq V$, and select z in $V-V'$. Then $z=z_1+\cdots+z_n$ where z_i is in V_{λ_i} and consequently there are scalars $\mu_1, \cdots, \mu_m$ such that $(A-\mu_1)\cdots(A-\mu_m)z=0$. Among the vectors

$$z, (A-\mu_m)z, (A-\mu_{m-1})(A-\mu_m)z, \cdots, (A-\mu_1)\cdots(A-\mu_m)z=0$$

the first is not in V' but the last is in V'. Let w be the last vector listed which is not in V'. Obviously there is a scalar μ such that $(A-\mu)w$ is in V'. It is clear that A', induced by A on V/V', has the eigen value μ. Hence A is evp.

Assume A is evp and let V' be the subspace spanned by the V_λ. We will show $V'=V$ by contradiction; suppose $V\neq V'$. Because A is evp there exists a scalar λ and a vector z in $V - V'$ such that $(A - \lambda)z$ is in V'. Say $\mu_1, \cdots, \mu_m$ are scalars such that $(A - \mu_1) \cdots (A - \mu_m)(A - \lambda)z = 0$. Clearly $[\pi_{\mu_i \neq \lambda}(A - \mu_i)]z$ is a vector which is annihilated by some power of $A - \lambda$; it is in V_λ. Now let $\lambda_1, \cdots, \lambda_n$ be those $\mu_i \neq \lambda$. We have

$$(A - \lambda_1)z = (A - \lambda)z + (\lambda - \lambda_1)z$$

and

$$(A - \lambda_2)(A - \lambda_1)z = (A - \lambda_2)(A - \lambda)z + (\lambda - \lambda_1)(A - \lambda)z + (\lambda - \lambda_2)(\lambda - \lambda_1)z \; .$$

Recalling that V' reduces A and repeating this argument we see finally that $[\pi_j(A - \lambda_j)]z = cz + w$ where w is in V' and c is the scalar $\pi_j(\lambda - \lambda_j)$. Since $c \neq 0$ and $cz + w$ is in V', z must be in V' contrary to the choice of z.

Thus A is evp if and only if V is the direct sum of all the V_λ. Of course in this general context infinitely many V_λ might be $\neq (0)$ and V_λ might not be annihilated by any power of $A - \lambda$. We leave to the reader the construction of examples demonstrating these possibilities.

LEMMA 4. *If A is evp then for each vector z in V there exists a nonzero polynomial p with coefficients in F (depending perhaps on z) such that $p(A)z = 0$.*

Proof. Assume A is evp and select z in V. By Lemma 3, z is in the subspace spanned by the V_λ and there exist scalars $\lambda_1, \cdots, \lambda_n$ such that $(A - \lambda_1)(A - \lambda_2) \cdots (A - \lambda_n)z = 0$. Then $(x - \lambda_1) \cdots (x - \lambda_n)$ is an appropriate polynomial.

LEMMA 5. *Let F be algebraically closed and suppose for each vector z in V there exists a nonzero polynomial p with coefficients in F such that $p(A)z = 0$. Then A is evp on V.*

Proof. Let V' be a subspace of V reducing A and $V' \neq V$. Select any z in $V - V'$. Let p be a polynomial such that $p(A)z = 0$. Since F is algebraically closed there exist scalars $c, \lambda_1, \cdots, \lambda_n$ such that

$$p(A) = c(A - \lambda_1) \cdots (A - \lambda_n)$$

and

$$(A - \lambda_1) \cdots (A - \lambda_n)z = 0 \; .$$

By the argument employed on the vectors

$$z, (A - \lambda_1)z, (A - \lambda_1)(A - \lambda_2)z, \cdots, (A - \lambda_1) \cdots (A - \lambda_n)z$$

in Lemma 3, it follows that A is evp.

3. Evp operators on Hilbert space. We now have all the algebraic machinery necessary to tackle the structure of bounded evp operators on complex Hilbert space. Until further notice assume that A is a bounded evp operator on the complex Hilbert space h.

LEMMA 6. *For each scalar λ there exists an integer N such that* $(A-\lambda)^N V_\lambda = (0)$.

Proof. Our proof is by contradiction; suppose there exists no integer N such that $(A-\lambda)^N V_\lambda = (0)$. Let Z_n be the null space of the operator $(A-\lambda)^n$. Then each Z_n is a closed subspace of h and together the Z_n span V_λ. The inclusions

$$Z_1 \subset Z_2 \subset Z_3 \subset \cdots$$

are all proper, for if $Z_n = Z_{n+1}$ for some n every vector annihilated by $(A-\lambda)^{n+1}$ is annihilated by $(A-\lambda)^n$ and likewise every vector annihilated by $(A-\lambda)^m$, $m > n$, is annihilated by $(A-\lambda)^n$ contrary to assumption. Setting $Z_0 = (0)$ select a unit vector z_n in $Z_n \cap Z_{n-1}^{\perp}$ for each index $n > 0$ and put $z = \sum_1^\infty 2^{-n^2} z_n$. By Lemma 4 there is a nonzero polynomial p with complex coefficients such that $p(A)z = 0$. For each index $n > 0$ we have $(A-\lambda)Z_n \subset Z_{n-1}$. For each index m

$$\begin{aligned} 0 &= (p(A)z, z_m) \\ &= ((p(A) - p(\lambda))z, z_m) + p(\lambda)(z, z_m) \\ &= 2^{-m^2}([p(A) - p(\lambda)][\sum_{n=m+1}^{\infty} 2^{m^2-n^2} z_n], z_m) + 2^{-m^2} p(\lambda) \end{aligned}$$

because $p(A) - p(\lambda)$ contains a factor of $A - \lambda$ and z_m is orthogonal to $[p(A) - p(\lambda)][\sum_{n=1}^{m} 2^{-n^2} z_n]$. Dividing out 2^{-m^2} we obtain

$$|p(\lambda)| \leqq 2^{-2m} \| p(A) - p(\lambda) \|$$

for and m, and clearly $p(\lambda) = 0$; indeed if q is a polynomial such that $q(\lambda) \neq 0$, then $(q(A)z, z_m) \neq 0$ for infinitely many indices m. Let q be the polynomial such that $p(A) = (A-\lambda)^e q(A)$ and $q(\lambda) \neq 0$. Now $0 = p(A)z = (A-\lambda)^e q(A)z$ and $q(A)z$ is in Z_e. But by the above argument $(q(A)z, z_m) \neq 0$ for infinitely many indices m and $z_m \perp Z_e$ for $m > e$, contradiction.

LEMMA 7. *There are at most finitely many complex scalars λ for which* $V_\lambda \neq (0)$.

Proof. Our proof is by contradiction; suppose there are infinitely many λ for which $V_\lambda \neq (0)$. Since A is bounded, the set of such λ is bounded. Let $\{\lambda_n\}$ be a sequence of distinct members of this set converging to a fixed scalar c such that $\lambda_n \neq c$ for all n. Let N be an integer such that $(A-c)^N V_c = (0)$. (It makes no difference if $V_c = (0)$ already.) Let Z_n be the null space of the operator $(A-c)^N(A-\lambda_1)\cdots(A-\lambda_n)$ for $n > 0$ and set $Z_0 = V_c$. All the inclusions

$$Z_0 \subset Z_1 \subset Z_2 \subset \cdots$$

are proper by Lemma 1, and $(A-\lambda_n)Z_n \subset Z_{n-1}$. Select in each $Z_n \cap Z_{n-1}^{\perp}$ a unit vector z_n and put $z = \sum_1^\infty 2^{-n^2} z_n$. There is a nonzero polynomial p such that $p(A)z = 0$. For each index m

$$\begin{aligned} 0 &= (p(A)z, z_m) \\ &= ([p(A) - p(\lambda_m)]z, z_m) + ([p(\lambda_m) - p(c)]z, z_m) + (p(c)z, z_m) \end{aligned}$$

and

$$\begin{aligned} 2^{-m^2}([p(A) - p(\lambda_m)][\sum_{n=m+1}^{\infty} 2^{m^2-n^2} z_n], z_m)& \\ + 2^{-m^2}[p(\lambda_m) - p(c)] + 2^{-m^2}p(c) = 0 \, .& \end{aligned}$$

Dividing out 2^{-m^2} we obtain

$$|p(c)| \leqq 2^{-2m} \| p(A) - p(\lambda_m) \| + | p(\lambda_m) - p(c) | \, .$$

As m tends to ∞, $\|p(A) - p(\lambda_m)\|$ tends to $\|p(A) - p(c)\|$ and $|p(\lambda_m) - p(c)|$ tends to 0. Clearly $p(c) = 0$; indeed if q is a polynomial such that $q(c) \neq 0$, then $(q(A)z, z_m) \neq 0$ for infinitely many indices m. Let q be the polynomial such that $p(A) = (A-c)^e q(A)$ and $q(c) \neq 0$. Then $0 = p(A)z = (A-c)^e q(A)z$ and $q(A)z$ is in $V_c = Z_0$. By the above argument $(q(A)z, z_m) \neq 0$ for infinitely many m but $z_m \perp Z_0$ for all $m > 0$, contradiction.

We are now able to prove our theorem. Let A be a bounded operator on the complex Hilbert space h and drop the assumption that A is evp.

THEOREM. *The following properties are equivalent for A.*

(1) *A is evp.*

(2) *There is a nonzero polynomial P with complex coefficients such that $P(A) = 0$ on h.*

Proof. Assume (1). By our development h is the direct sum of $V_{\lambda_1}, \cdots, V_{\lambda_n}$ and for each index $i = 1, \cdots, n$, there is an integer e_i such that $(A - \lambda_i)^{e_i} V_{\lambda_i} = (0)$. Clearly $\pi_1^n (A - \lambda_i)^{e_i} = 0$ on h and $\pi_1^n (x = \lambda_i)^{e_i}$ is an appropriate polynomial P.

Now assume (2). By Lemma 5, A is evp, and the proof is complete.

The next corollary is of particular interest because it states that a pointwise property is equivalent to a global property. The proof is an immediate consequence of Lemma 5 and our theorem.

COROLLARY 1. *The following properties are equivalent for A.*

(1) *For each vector z in h there exists a nonzero polynomial p (depending perhaps on z) such that $p(A)z = 0$.*

(2) *There exists a nonzero polynomial P such that $P(A) = 0$ on h.*

Any operator on a finite dimensional complex Hilbert space is of course evp. Such an operator A must also be bounded and the polynomials in A must constitute a uniformly closed subset of the ring of all operators; the latter fact is due to the theorem that any linear manifold in a finite dimensional normed linear space is closed in norm. Consequently we should not be surprised by

COROLLARY 2. *Let A be a bounded operator on a complex Hilbert space h. Then the following properties are equivalent for A.*

(1) *A is evp.*

(2) *the polynomials in A constitute a uniformly closed subset of $L(h)$, the ring of all bounded operators on h.*

Proof. First suppose A is evp. Let P be a nonzero polynomial of minimal degree such that $P(A) = 0$ and say n is the degree of P. Then any polynomial in A is equal to an appropriate polynomial in A of degree less than n. It follows that the polynomials in A form a finite dimensional linear manifold in the normed linear space $L(h)$ over the complex field, and this manifold must be closed in norm.

Now suppose the polynomials in A form a uniformly closed subset of $L(h)$. (The author is indebted to B. Yood for the remainder of the argument.) Select a positive scalar r greater than the spectral radius of A. Then $A - r$ is nonsingular and $(r - A)^{-1}$ is the limit of a convergent series in A and this series must converge to a polynomial $p(A)$. Thus $(r - A)p(A) = 1$ and A is a root of the nonzero polynomial $(r - x)p(x) - 1$.

4. Principal idempotents. Given operators A and B on the respective Hilbert spaces h and k we say A is similar to B if there exists an invertible linear tranformation S of h onto k such that S and S^{-1} are bounded and $SAS^{-1} = B$ on k. This section concerns the theorem on principal idempotents on finite dimensional vector spaces (see [7], pp. 175–7) which states that a matrix is similar to a diagonal matrix if and only if it is a linear combination of mutually orthogonal idempotent matrices (of course no condition of boundedness enters here). In particular if the field of scalars is the complex field this theorem remains valid when "normal matrix" replaces "diagonal matrix".

The following corollary generalizes this result by means of bounded normal evp operators. By the spectral mapping theorem all the spectral values of a bounded normal evp operator are roots of a certain nonzero polynomial; hence such an operator N has finite spectrum, and furthermore by the spectral theorem all of its spectral values are eigen values and N is a linear combination of finitely many mutually orthogonal projections the sum of which is the identity.

COROLLARY 3. *For an operator A mapping h into h the following are equivalent.*

(1) *A is similar to a bounded normal evp operator N on a Hilbert space.*

(2) *There exist a finite collection $E_1, \cdots, E_n$ of mutually orthogonal nonzero idempotent operators, each mapping h onto a closed subspace of h, and distinct scalars $\lambda_1, \cdots, \lambda_n$ such that $I = \sum_1^n E_i$ and $A = \sum_1^n \lambda_i E_i$.*

Proof. Assume (1). Then N is a linear combination of mutually orthogonal nonzero projections $P_1, \cdots, P_n$; say $N = \sum_1^n \lambda_i P_i$ and $I = \sum_1^n P_i$ where $\lambda_i \neq \lambda_j$ if $i \neq j$. There exists an invertible linear transformation S such that $SAS^{-1} = N$ and

$$A = S^{-1}NS = S^{-1}(\sum_1^n \lambda_i P_i)S = \sum_1^n \lambda_i S^{-1} P_i S ,$$

and obviously $E_i = S^{-1}P_iS$ suffices in (2). Of course the operators A, $E_1, \cdots, E_n$ are bounded because S, S^{-1}, N, $P_1, \cdots, P_n$ are bounded. Furthermore each linear manifold E_ih is closed in h because the range of a bounded idempotent operator on Hilbert space must be closed.

Now assume (2). Then each E_ih is a closed linear subspace of h and h is the direct linear sum of the E_ih. Let k be the direct orthogonal product of the Hilbert spaces E_ih. Let S_i be the linear transformation mapping h onto E_ih (regarded as a subspace of k) which is the identity on E_ih and which annihilates every vector in E_jh, $j \neq i$. Then $S = \sum_1^n S_i$ is an invertible linear transformation of h onto k. It suffices to show that S and S^{-1} are bounded; for if they are, the operators SE_iS^{-1} are mutually orthogonal projections on k and A is similar to the bounded normal evp operator $SAS^{-1} = \sum_1^n \lambda_i SE_iS^{-1}$ on k. That S^{-1} is bounded follows immediately form the inequality $\| S^{-1}z \|^2 \leqq n_2 \| z \|^2$, all z in k. By the closed graph theorem S is bounded also, and the proof is complete.

REFERENCES

1. G. Birkhoff and S. Maclane, *A Survey of Modern Algebra*, 1941.
2. J. Dixmier *Les Algebres D'operateurs dans L'espace Hiblertien*, 1957.
3. P. R. Halmos, *Finite Dimensional Vector Spaces*, 1958.
4. _______, *Introduction to Hilbert Space*, 1951.

5. M. Jacobson, *Lectures in Abstract Algebra, vol.* II, 1953.
6. N. H. McCoy, *Introduction to Modern Algebra*, 1960.
7. S. Perlis, *Theory of Matrices*, 1952.
8. M. H. Stone, *Linear Transformations on Hilbert Space*, 1932.

UNIVERSITY OF OREGON