ON THE MINIMAL POLYNOMIAL OF AN ALGEBRAIC OPERATOR

S. Karanasios and J. Maroulas
National Technical University of Athens
Department of Mathematics
Zografou Campus
GR 157 80 Zografou, Athens
Greece
skaran@math.ntua.gr
maroulas@math.ntua.gr

Abstract

In this paper we extent some results concerning the minimal polynomial of an algebraic operator from the finite dimentional case to infinite dimentions and characterize the degree of an algebraic operator. We also give some applications to linear control systems.

email: skaran@math.ntua.gr
maroulas@math.ntua.gr
AMS 1991 *Mathematics Subject Classification.* Primary 47B99, Secondary 47N10

1 Preliminaries and Notation

Let X be a finite-dimensional vector space. All operators on X are algebraic in the sense that to each operator A on X there exists a non-zero polynomial p such that $p(A) = 0$. Moreover, if the dimension of X is n, then the degree of p does not exceed n. (This is a consequence of the Cayley-Hamilton's theorem). In fact the minimal polynomial of A is the polynomial, among the annihilating polynomials of A, with the smallest degree. The degree of the minimal polynomial of A is called the **degree of A** and is denoted by $\deg(A)$. Also, by $\deg(p)$, we denote the degree of the polynomial p.

In infinite-dimensional spaces there are many operators which are not algebraic. But in this case the algebraic operators can be characterized in terms of their finite dimensional invariant subspaces (see [4]). Taking advantage of the fact that an algebraic operator A has finite dimensional invariant subspaces we characterize the degree of A, having in mind [2]. This is a generalisation from finite to infinite dimensional case. Finally we give applications defining finite dimensional controllable linear control systems by the operator A.

2 The minimal polynomial

Let X be a vector space and $\mathcal{B}(X)$ the set of all bounded linear operators on X.

Definition 1 *An operator $A \in \mathcal{B}(X)$ is called*
1. **algebraic** *if there exists a non trivial polynomial p such that $p(A) = 0$.*
2. **locally algebraic** *if for every $x \in X$ there exists a non trivial polynomial p such that $p(A)x = 0$.*
3. **algebraic of degree** n *if there exists a polynomial p of degree n such that $p(A) = 0$ and for any polynomial $q \neq 0$ of degree $\leq n-1$, $q(A) \neq 0$.*

Such a monic polynomial p is unique and is called **minimal polynomial** *of A. We denote it by p_A.*

The uniqueness of the minimal polynomial of A is easy to prove. Indeed, if p_1, p_2 are two monic polynomials of the same degree n such that $p_1(A) = p_2(A) = 0$ then $(p_1 - p_2)(A) = 0$ and $\deg(p_1 - p_2) \leq n - 1 < n$. Hence $p_1 = p_2$.

Definition 2 *Let x be a vector in X. A polynomial p is called an* **annihilating polynomial of x with respect to** $A \in \mathcal{B}(X)$ *if $p(A)x = 0$. If p is a monic polynomial of minimal degree then it is called* **minimal polynomial of x with respect to** A *and will be denoted by p_x.*

An operator $A \in \mathcal{B}(X)$ is called **boundedly locally algebraic** *if it is locally algebraic and the degree of the polynomials p_x is bounded independently of x.*

Remark 1 It is well known that boundedly locally algebraic operators are also algebraic. (see [1]).

Moreover $A \in \mathcal{B}(X)$ is algebraic if and only if the union of all finite dimentional invariant subspaces of A equals to X. (see [4]).

Also every operator $B \in \mathcal{B}(X)$ similar to an algebraic operator A of degree n is also algebraic of the same degree n. ($B = P^{-1}AP$ where P, $P^{-1} \in \mathcal{B}(X)$ and so $p(B) = P^{-1}p(A)P$ for any polynomial p)

Lemma 1 *Any annihilating polynomial of an algebraic operator $A \in \mathcal{B}(X)$ is divisible by its minimal polynomial p_A.*

Proof. Let p be any polynomial, such that $p(A) = 0$. Then there exists polynomials q and r such that $p = p_A \cdot q + r$, where $\deg(r) < \deg(p_A)$ and $p(A) = p_A(A)q(A) + r(A)$. Hence $r(A) = 0$ and therefore $r = 0$. □

Lemma 2 *The minimal polynomial p_x of any vector $x \in X$ with respect to A devides the minimal polynomial p_A of A.*

Proof. We have $p_A = q \cdot p_x + r$, where $\deg(r) < \deg(p_x)$. Then $p_A(A) = q(A) \cdot p_x(A) + r(A) \Rightarrow r(A) = -q(A) \cdot p_x(A)$, from which we get $r(A)x = -q(A)p_x(A)x = 0$ and hence $r = 0$. □

Proposition 1 *If X is a Banach space and $A \in \mathcal{B}(X)$ is an algebraic operator, then there exists a vector $x \in X$ such that its minimal polynomial p_x with respect to A coincides with the minimal polynomial p_A of A.*

Proof. It is well known that any ideal J in the ring $\mathcal{P}$ of polynomials of one variable is generated by a polynomial of minimal degree. Let $x \in X$ and the ideal

$$I_x = \{p \in \mathcal{P} : p(A)x = 0\}.$$

The ideal I_x is generated by a monic polynomial p_x of minimal degree which is the minimal polynomial of x with respect to A. If p_A is the minimal polynomial of A then $p_A \in I_x$ and, by Lemma 2, p_A is divisible by p_x. Therefore, considering all x in X, we get only a finite number of polynomials $p_{x_1}, p_{x_2}, ..., p_{x_k}$, since each p_{x_i}, $i = 1, ..., k$ devides p_A. Defining by

$$X_i = \{x \in X : p_{x_i}(A)x = 0\},\ i = 1, 2, ..., k$$

then $X = \cup X_i$ and consequently, in a similar way as in Theorem 4.8, [4], we have $X = X_j$ for some $j \in \{1, 2, ..., k\}$. Therefore, $p_{x_j}(A)X = 0$ and, by Lemma 1, it is implied that p_{x_j} is divisible by p_A. Hence, $p_A = p_{x_j}$. □

Corollary 1 *For an algebraic operator A on a Banach space X the degree of A is n if and only if for a certain vector $x \in X$ the vectors x, $Ax, ..., A^{n-1}x$ are linearly independent while the vectors x, $Ax, ..., A^n x$ are linearly dependent.*

The vector $x \in X$, in the above Corollary, is the vector for which we have $p_x = p_A$.

In the sequel X will be a Hilbert space. The next Proposition characterizes the first n coordinates, with respect to a suitable base, of any vector of X, where n is the degree of A.

Proposition 2 *Let A be an algebraic operator on a Hilbert space X with $deg(A) = n$. Then we can choose a basis $\{e_i\}$ of X such that for any vector $x = (x_1, x_2, ..., x_n, ...) \in X$ there exist $n \times 1$ matrices P, Q (where Q depends on x) and $n \times n$ matrix B, for which we have $x_k = Q^\top B^{k-1} P$, $k = 1, 2, ..., n$.*

Proof. If the degree of A is n then there exists a vector y such that $p_A = p_y$ and the vectors y, $Ay, ..., A^{n-1}y$ are linearly independent. Denoting the linear span of the vectors y, $Ay, ..., A^{n-1}y$ by

$$M = \text{span}\{y,\ Ay, ..., A^{n-1}y\},$$

we have that M is an invariant subspace of A with $\dim M = n$. Taking an orthonormal basis $\{m_i\}$ of M and extending it to be an othonormal basis $\{e_i\}$ of X, the restriction $A|_M$, of A to M, corresponds with respect to the basis $\{m_i\}$ to an $n \times n$ matrix B. Let $(y_1, y_2, ...)$ be the coordinates of y with respect to the base $\{e_i\}$ and $P = (y_1, y_2, ..., y_n)^\top$. Then, the matrix

C formed by the columns P, $BP, ..., B^{n-1}P$, is invertible. So, for any vector $x = (x_1, x_2, ..., x_n, 0, ...) \in M \subseteq X$, setting $Q^\top = (x_1, x_2, ..., x_n)C^{-1}$, we obtain $x = (Q^\top P, Q^\top BP, ..., Q^\top B^{n-1}P, 0, ...)$. □

Definition 3 *Let A be an algebraic operator, M be an invariant subspace of A of finite dimension and B the corresponding matrix of the operator $A|_M$ with respect to an orthonormal base of M. We say that the subspace M has the* **property** (d_n) *if* $\dim M = n$ *and for every* $x = (x_1, x_2, ...x_n, 0, ...) \in M$ *there exist* $n \times 1$ *matrices* P, Q *such that* $x_k = Q^\top B^{k-1}P,\ k = 1, 2, ..., n$.

Proposition 3 *Let A be an algebaraic operator on a Hilbert space X and M an invariant subspace of A with the property (d_n). Then* $\deg(A) \geq n$.

Proof. Let $\{e_1, e_2, ..., e_n\}$ be an orthonormal basis of M. To each vector e_i corresponds, by hypothesis, $n \times 1$ matrices P_i, Q_i, $i = 1, 2, ..., n$. Since $e_1, e_2, ..., e_n$ are linearly independent the matrix

$$C = \begin{pmatrix} Q_1^\top P_1 & \dots & Q_1^\top B^{n-1}P_1 \\ \vdots & \dots & \vdots \\ Q_n^\top P_n & \dots & Q_n^\top B^{n-1}P_n \end{pmatrix}$$

is invertible. Therefore the matrices $I, B, ..., B^{n-1}$ are linearly independent, otherwise the columns of C would be linearly dependent. Hence, $\deg(B) = \deg(A|_M) = n$ and therefore $\deg(A) \geq n$. □

Remark 2 It is implicit in the proof of Proposition 2 that the subspace $\text{span}\{y, Ay, ..., A^{n-1}y\}$ has the property (d_n), where $n = \deg(A)$. Therefore a consequence of Proposition 3 is the following:

Corollary 2 *We have*

$$\deg(A) = \max\{n : n = \dim M,\ M \text{ has the property } (d_n)\}.$$

Remark 3 An application to system theory can be obtained by the factorization of the matrix C, which is defined in the proof of Proposition 3, as follows.

We have

$$C = \begin{pmatrix} Q_1^{\mathsf{T}} P_1 & \dots & Q_1^{\mathsf{T}} B^{n-1} P_1 \\ \vdots & \dots & \vdots \\ Q_n^{\mathsf{T}} P_n & \dots & Q_n^{\mathsf{T}} B^{n-1} P_n \end{pmatrix} = \mathrm{diag}(Q_1, Q_2, \dots, Q_n)^{\mathsf{T}} \mathcal{C}(R, I_n \otimes B),$$

where $\mathcal{C}(R, I_n \otimes B) = [R, (I_n \otimes B)R, \dots, (I_n \otimes B)^{n-1}R]$, and $R = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix}$.

Therefore for the linear control system

$$\dot{z}(t) = (I_n \otimes B)z(t) + Ru(t),$$

by the equality $\mathrm{rank} C = \mathrm{rank}\mathcal{C}(R, I_n \otimes B) = n$, we get the following

Proposition 4 *Let A be an algebraic operator, M be an invariant subspace of A with the property d_n and B the corresponding matrix of the operator $A|_M$ with respect to an orthonormal basis of M. Then it is constructed the controllable system*

$$\dot{z}(t) = (I_n \otimes B)z(t) + Ru(t) \; ; \; R = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix},$$

where the matrices P_i, $i = 1, \dots, n$ correspond to the orthonormal basis and $u(t)$ is the control function.

Note that, when $\mathrm{deg} A = n$, the matrix B is non-derogatory and $P_1 = P_2 = \dots = P_n = P$, where P is the matrix defined in the proof of Proposition 2. Hence by $C = \mathrm{diag}(Q_1, Q_2, \dots, Q_n)^{\mathsf{T}}(I_n \otimes \mathcal{C}(P, B))$ we have the next simplified form of the corresponding controllable system

$$\dot{z}(t) = Bz(t) + Pu(t).$$

In this special case using Theorem 1, [3] it is implied that the minimal polynomial of B and that of P coincide.

References

[1] B. Aupetit, *A Primer on Spectral Theory*, Springer-Verlag, New York, Inc., 1991

[2] H. K. Farahat and W. Ledermann, *Matrices with prescribed characteristic polynomials*, Proc. Edinburgh Math. Soc., **11**, (1958), 143–146.

[3] J. Maroulas and S. Barnett, *A new Controllability Condition*, IEEE Transactions on automatic control, vol. **AC-24**, No. **4**, (1979), 635–637.

[4] H. Radjavi and P. Rosenthal, *Invariant subspaces*, Ergeb. Math. Grenzgeb. 77, Springer, Berlin, 1973.