Integr Equat Oper Th
Vol. 17 (1993)

0378-620X/93/010107-16$1.50+0.20/0

COMPLETE CONTROLLABILITY AND SPECTRUM ASSIGNMENT IN INFINITE DIMENSIONAL SPACES

A.S. Markus and V.R. Olshevsky

INTRODUCTION

The concept of complete controllability of a system plays an important role in the theory of systems with a control, [1]–[3]. A large number of conditions is known which are equivalent to complete controllability (see, for example, [2], §34). Some of these were generalized to the infinite dimensional case in [4]–[7].

In the present paper two further criteria for complete controllability in infinite dimensional spaces are shown: Property 5), defined in Theorem 3 (for the finite dimensional case the dual statement – a criterium for observability – was obtained by Hautus, [8]), and Property 6), defined in Theorem 10 (in the finite dimensional case this property coincides with conditon 3^o from [2], §34).

We have also obtained new proofs of the theorems on the equivalence in Hilbert spaces of complete controllability and the assignability of the spectrum from [5], [7], and of some propositions from [6]. Examples are presented which show that in Banach spaces the assignability of the spectrum does not always follow from complete controllability. Properties which are dual to complete controllabiliy are investigated in this paper as well. They can be applied to the study of observability in infinite dimensional spaces. Finally, the connections with the theory of holomorphic operator functions are indicated.

1. SOME CRITERIA FOR COMPLETE CONTROLLABILITY.

1.1 By $\mathcal{L}(U,X)$ we shall denote the set of all bounded linear operators acting from one complex Banach space U into another one X. Instead of $\mathcal{L}(X,X)$ we shall write $\mathcal{L}(X)$. By $\sigma(A)$ we shall denote the spectrum of the operator $A\in\mathcal{L}(X)$. If $B\in\mathcal{L}(U,X)$, then Im B is the set of all values of B and Ker B is its kernel.

The identity operator is denoted by I, but if it is necessary to indicate the space X on which it acts, then by I_X. By $X_1\oplus...\oplus X_n$ we shall denote the direct sum of the Banach spaces $X_1,\dots,X_n$, where the norm is defined by the identity

$$\|(x_1,\dots,x_n)\| = (\sum_1^n \|x_k\|^2)^{1/2}.$$

In the case where $X_1 = X_2 = \dots = X_n$ we shall denote this sum by X^n. Operators which act from $X_1 \oplus \dots \oplus X_n$ into $U_1 \oplus \dots \oplus U_m$ we write down in the form of operator matrices of order $m \times n$. For $m=1$ we obtain operator rows, and for $n=1$ operator columns. The direct sum of the subspaces Y and Z of the space X will be denoted by $Y \dot{+} Z$.

We consider the equation:

$$\frac{dx}{dt} = Ax + Bu, \tag{1.1}$$

where $A \in \mathcal{L}(X)$, $B \in \mathcal{L}(U,X)$, and $x(t)$ and $u(t)$ are vector functions with values in X and U, respectively ($u(t)$ is called the *control*). The equation (1.1) is called *completely controllable* (the terms entirely controllable, exactly controllable zero-controllable are also in use), if for arbitrary $x_o \in X$ there exist a number $T > 0$ and a piecewise continuous control $u(t)$ $(0 \le t \le T)$ such that the solution $x(t)$ of (1.1) with the initial condition $x(0) = x_o$ and the control $u(t)$ has the property $x(T) = 0$.

Various criteria for complete controllability for finite dimensional X and U are contained in [1]–[3]. In the paper [4] the criterium of Kalman, [1], §19, has been transferred to the infinite dimensional case, i.e., it was shown that equation (1.1) is completely controllable if and only if the condition

$$\text{1)} \qquad \operatorname{Im} K^m(B,A) = X$$

is met, where $K^m(B,A) = (B \;\; AB \;\; A^2B \dots A^{m-1}B)$, i.e., the operator $K^m(B,A) (\in \mathcal{L}(U^m, X))$ is given by the identity

$$K^m(B,A)\begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} = \sum_{k=1}^m A^{k-1}Bu_k$$

In [4] it was shown that if condition 1) is met, then one can restrict oneself to piecewise constant controls $u(t)$ which have not more than $m-1$ points of discontinuity.

1.2 Basically, we shall be interested in the generalizations to the infinite dimensional case of the criteria of Popov–Langenhope–Wonham, [3], §§3.2 and 5.2 (the assignment of the spectrum). For the proof of the following statement the authors in [6] used results from the spectral theory of holomorphic operator functions; we adduce, however, a more simple proof.

THEOREM 1. *If there exists an operator* $F \in \mathcal{L}(X,U)$ *such that*

(1.2) $$\sigma(A) \cap \sigma(A-BF)=\emptyset,$$

then the operator $K^m(B,A)$ *is the right invertible for some natural number* m.

PROOF. Because of (1.2) one finds a bounded open set $\Omega(\subseteq\mathbb{C})$ such that $\sigma(A)\subset\Omega$, $\sigma(A-BF)\cap\bar{\Omega}=\emptyset$ and such that its boundary Γ consists of a finite number of simple, closed, rectifiable curves. Having integrated the identity $(A-BF-\lambda I)^{-1}-(A-\lambda I)^{-1}= =(A-\lambda I)^{-1}BF(A-BF-\lambda I)^{-1}$ over the curve Γ we obtain

$$2\pi i I=\int_\Gamma (A-\lambda I)^{-1}BF(A-BF-\lambda I)^{-1}d\lambda.$$

We approximate the integral by a finite sum such that this sum remains an invertible operator

$$G=\sum_{k=1}^{m}(A-\xi_k I)^{-1}BF(A-BF-\xi_k I)^{-1}\Delta\lambda_k$$

We multiply this identity from the left by $\Pi(A-\xi_k I)$ and from the right by $G^{-1}\Pi(A-\xi I)^{-1}$. At the right-hand side we multiply out the parentheses which stand in each summand to the left of the product BF, and we combine similar terms. We obtain the identity $I=\sum_{j=0}^{m-1}A^jBQ_j$, where $Q_j\in\mathcal{L}(X,U)$. The theorem is proved.■

Besides Property 1) we shall consider the following properties of operator pairs B,A:

2) The operator $K^m(B,A)$ is right invertible;

3) For an arbitrary (nonempty) compact set $K\subset\mathbb{C}$ there exists an operator $F\in\mathcal{L}(X,U)$ such that $\sigma(A-BF)=K$ (for $\dim X<\infty$ it is assumed that K consists of at most $\dim X$ points);

4) For $\lambda\in\mathbb{C}$ arbitrary one finds an operator $F\in\mathcal{L}(X,U)$ such that $(A-\lambda_0 I-BF)^m=0$.

Property 3) is usually called the *assignability of the spectrum*. It is evident that 1) follows from 2).

From Theorem 1 follows immediately

THEOREM 2. *Each of the properties* 3), 4) *implies Property* 2).

REMARK 1. If Property 4) holds (for just one $\lambda_0\notin\sigma(A)$) then Property 2) holds as well with the same m.

Indeed, according to Theorem 1 there exists a number l such that

(1.3) $$\sum_{j=0}^{l-1}A^jBQ_j=I.$$

From condition 4) it follows that

$$A^m=\sum_{k=0}^{m-1}A^kBS_k \qquad (S_k\in\mathcal{L}(X,U))$$

If $l>m$, then, after replacing A^{l-1} in (1.3) by $A^{l-m-1}\sum_{k=0}^{m-1} A^k BS_k$, we reduce the highest power of A in identity (1.3) by one. Hence the required conclusion follows.

1.3 We introduce a condition which is equivalent to 1). For finite dimensional X and U the dual result (a criterium for observability) was obtained by Hautus, [8]. The operator which is considered below acts from $X \oplus U$ into X.

THEOREM 3. *The condition* 1) *is equivalent to the property*

5) $$\operatorname{Im}(A-\lambda I, B)=X \text{ for arbitrary } \lambda\in\mathbb{C}.$$

PROOF. It is evident that the identity $\operatorname{Im}(A,B)=X$ follows from 1). On the other hand, Property 1) is preserved if A is replaced by $A-\lambda I$, and hence 5) follows from it.■

We assume that condition 5) holds. We fix some $R>\|A\|$. By a theorem of Leiterer, [9], there exist for arbitrary $f\in X$ vector functions $x(\lambda)$ and $v(\lambda)$, holomorphic in the disc $|\lambda|<R$ and with values in X and U, respectively, such that $(A-\lambda I)x(\lambda)+Bv(\lambda)=f$ $(|\lambda|<R)$. We rewrite this identity in the following form

$$x(\lambda)+(A-\lambda I)^{-1}Bv(\lambda)=(A-\lambda I)^{-1}f \qquad (|\lambda|<R, \lambda\notin\sigma(A)). \tag{1.4}$$

Let $r\in(\|A\|,R)$ and let the expansion of $v(\lambda)$ in its Maclaurin series be of the form

$$v(\lambda)=\sum_{k=0}^{\infty} v_k\,\lambda^k \quad (v_k\in U).$$

As

$$-\frac{1}{2\pi i}\int_{|\lambda|=r} \lambda^k\,(A-\lambda I)^{-1}d\lambda = A^k\ ,$$

we obtain, after integration of (1.4) over the circle $|\lambda|=r$, that

$$\sum_{k=0}^{\infty} A^k B v_k = f. \tag{1.5}$$

Let $l^\infty(U)$ be the space of all bounded sequences in U with the norm $\|(u_k)_1^\infty\| = \sup\|u_k\|$. It is easy to see that the operators

$$Su=\sum_{k=0}^{\infty}\frac{A^kB}{r^k}u_k, \qquad S_n u=\sum_{k=0}^{n}\frac{A^kB}{r^k}u_k \qquad (u=(u_k)_1^\infty)$$

belong to $\mathcal{L}(l^\infty(U),X)$, whereas

$$\|(S-S_n)u\|\le\sum_{k=m+1}^{\infty}\left(\frac{\|A\|}{r}\right)^k\|B\|\,\|u_k\|\le\|u\|\,\|B\|\sum_{k=m+1}^{\infty}\left(\frac{\|A\|}{r}\right)^k.$$

Consequently,

$$\|S-S_n\| \leq \|B\| \sum_{k=n+1}^{\infty} \left(\frac{\|A\|}{r} \right) \to 0 \ (n \to \infty).$$

The vectors v_k from identity (1.5) allow the estimate

$$\|v_k\| \leq M r^{-k} \qquad (M = \max_{|\lambda| \leq r} \|v(\lambda)\|),$$

and hence $Su = f$, where $u = (r^k v_k)_1^{\infty} \in l^{\infty}(U)$.

Since the vector f was arbitrary one has $\operatorname{Im} S = X$. Hence it follows from (1.6) that $\operatorname{Im} S_n = X$ as well for n sufficiently large. This now means that condition 1) is satisfied. The theorem is proved.■

1.4 In the papers [5]–[7] it was established that the Properties 1)–5) are equivalent for Hilbert spaces X and U (an independent proof of this will be presented in Section 2). In Banach spaces the mentioned statement is, generally speaking, incorrect. We present some examples where 2) does not follow from 1) and hence none of the Properties 2), 3), 4) follows from complete controllability.

THEOREM 4. *In each of the following cases Property* 2) *does not follow from Property* 1):

a) $X = U$ and it coincides with one of the following spaces:

l^p, L^p $(1 < p \leq \infty,\ p \neq 2)$, c_o, C, L^1;

b) $U = l^1$, X is an arbitrary separable Banach space which is not isomorphic to l^1;

c) $X = L^2$, $U = L^p (2 < p \leq \infty)$.

For a proof of the Theorem it suffices to find an operator $B \in \mathcal{L}(U,X)$ such that $\operatorname{Im} B = X$ but B has no right inverse. Associating with such an operator B the operator $A = 0$, we obtain the required example. The proof of the existence of an operator with such properties can be found for the case a) in [10] (which contains the detailed references), for the case b) in [11], pp.54, 108 and for the case c) in [12], p.70 $(p < \infty)$ and [11], pp.57, 111 $(p = \infty)$.

REMARK 2. If U is a Hilbert space then $\operatorname{Ker} K^m(B,A)$ is always complemented and hence 2) follows from 1). In that case X is automatically isomorphic to a Hilbert space if there exists just one pair of operators A,B for which condition 1) is satisfied. On the other hand, case c) of Theorem 4) shows that when X is a Hilbert space but U isn't, then the implication 1)⇒2) is, generally speaking, incorrect.

REMARK 3. We give nontrivial examples of pairs X and U for which the Properties 1) and 2) are equivalent:
α) $X = l^1$, U is an arbitrary Banach space; β) $X = L^2$, $U = L^p$, $1 < p < 2$.
In case α) the equivalence of 1) and 2) follows from [13], p.72, and in case β) it follows from [14]. The case β) is interesting if it is compared with the case c) of Theorem 4.

2. DUAL PROBLEMS. PROOF OF THE EQUIVALENCE OF THE CONSIDERED PROPERTIES FOR HILBERT SPACES.

2.1 Let X and U be Banach spaces, $A \in \mathcal{L}(X)$, $C \in \mathcal{L}(X,U)$. By $K_m(C,A)$ we denote the operator column

$$K_m(C,A) = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{m-1} \end{pmatrix},$$

i.e, the operator from $\mathcal{L}(X,U^m)$, given by the identity $K_m\ (C,A)f = (CA^{k-1}f)_1^m \quad (f \in X)$.
We consider the following properties of the operator pair C,A:
1*) The operator $K_m(C,A)$ is an isomorphic embedding;
2*) The operator $K_m(C,A)$ is left invertible;
3*) For an arbitrary nonempty compact set $K \subset \mathbb{C}$ there exists an operator $F \in \mathcal{L}(U,X)$ such that $\sigma(A - FC) = K$ (for $\dim X = \infty$ it is assumed that K consists of at most $\dim X$ points);
4*) For $\lambda_0 \in \mathbb{C}$ arbitrary one can find an operator $F \in \mathcal{L}(U,X)$ such that $(A - \lambda_0 I - FC)^m = 0$;
5*) For arbitrary $\lambda \in \mathbb{C}$ the operator $\begin{pmatrix} A - \lambda I \\ C \end{pmatrix}$ is an isomorphic embedding *of* X in $X \oplus U$.
We observe that the properties 1*)–5*) are dual to 1)–5), i.e., if the operators $A \in \mathcal{L}(X)$, $B \in \mathcal{L}(U,X)$ possess some property k), then the operators $A^* \in \mathcal{L}(X^*)$, $B^* \in \mathcal{L}(X^*,U^*)$ have the property k*), and if $A \in \mathcal{L}(X)$, $C \in \mathcal{L}(X,U)$ have the property k*), then the operators $A^* \in \mathcal{L}(X^*)$, $C^* \in \mathcal{L}(U^*,X^*)$ possess the property k) $(k = 1,\ldots,5)$.
It is evident that 1*) follows from 2*). In analogy to the Theorems 2–4 and Remark 3 the following statements are proved (for reflexive spaces X and U they are corollaries form the results mentioned in Section 1).

THEOREM 5. *Each of the properties* 3*), 4*) *implies* 2*).

THEOREM 6. *The properties* 1*) *and* 5*) *are equivalent.*

THEOREM 7. *In each of the cases listed below* 2*) *does not follow from* 1*)
a) $X = U$ *and it coincides with one of the spaces* l^p, L^p $(1 \le p \le \infty,\ p \ne 2)$, C;
b) $U = l^\infty$, X *is an arbitrary separable Banach space*;
c) $X = L^2$, $U = L^p$ $(1 < p < 2)$.

REMARK 4. In the following cases the properties 1*) and 2*) are equivalent:

α) $X = c_0$, U is an arbitrary separable Banach space; β) $X = L^2$, $U = L^p$ $(2 < p < \infty)$.

2.2 By $\rho(f,M)$ we shall denote the distance between a vector f and a set M and by $M_\varepsilon(C)$ the set of all $f \in X$ for which the inequality $\|Cf\| \le \varepsilon \|f\|$ is satisfied. Let $r(C,A)$ be the largest number l such that for arbitrary $\varepsilon > 0$ there exists a nonzero vector $f \in M_\varepsilon(C)$ satisfying the inequalities $\rho(A^k f, M_\varepsilon(C)) \le \varepsilon \|f\|$ $(k = 1,\dots,l-1)$.
If among these numbers l there is no largest, then we put $r(C,A) = \infty$, and in the case where $M_\varepsilon(C) = \{0\}$ for some $\varepsilon > 0$ we shall consider $r(C,A) = 0$. We shall investigate the following property of an operator pair C,A:

6*) $$r(C,A) < \infty.$$

The definition of the quantity $r(C,A)$ is rather tedious, hence we shall clarify its meaning in some important special cases. We denote by $r_0(C,A)$ the largest number l such that for arbitrary $\varepsilon > 0$ there exists a nonzero vector $f \in \operatorname{Ker} C$ which satisfies the inequalities $\rho(A^k f, \operatorname{Ker} C) \le \varepsilon \|f\|$ $(k = 1,\dots,l-1)$. By $r_1(C,A)$ we shall denote the smallest number l such that for arbitrary $f \in \operatorname{Ker} C$ $(f \ne 0)$ at least one of the vectors $\{A^k f\}_1^l$ is not in Ker C. If Ker $C = \{0\}$ then we put $r_0(C,A) = r_1(C,A) = 0$.

LEMMA 1. *If* $\overline{\operatorname{Im} C} = \operatorname{Im} C$ *then* $r_0(C,A) = r(C,A)$. *If* dim $X < \infty$ *then* $r_1(C,A) = r(C,A)$.

PROOF. If Im C is closed, then there exists, as is well-known, ([15], Lemma VI.6.1), a number κ such that for arbitrary $u \in \operatorname{Im} C$ one finds $x \in X$ for which $Cx = u, \|x\| \le \kappa \|u\|$. Hence for arbitrary $f \in M_\varepsilon(C)$ one finds $g \in \operatorname{Ker} C$ such that $\|f - g\| \le \kappa \|Cf\| \le \kappa\varepsilon \|f\|$. From this it is not difficult to deduce the first statement of the lemma. In order to obtain the other statement it is necessary to use the fact that for dim Ker $C < \infty$ it follows from the relations $f_n \in \operatorname{Ker} C$, $\|f_n\| = 1$, $\rho(A^k f_n, \operatorname{Ker} C) \to 0$ $(k = 1,\dots,l)$ that there exists a vector f (the limit of some subsequence of $\{f_n\}$) for which $\|f\| = 1$, $A^k f \in \operatorname{Ker} C$ $(k = 0,1,\dots,l)$.■

Thus, in the finite dimensional case condition 6*) means that an arbitrary nonzero vector $f \in \operatorname{Ker} C$ is mapped out of the subspace Ker C by applying the operator A to it at most $r = r(C,A)$ times. The equivalence of this condition to the conditions 1*)–5*) in the afore-mentioned case is well-known (see, for example, [2], p.351, condition 3). We shall show that the conditions 1*)–6*) are equivalent for arbitrary Hilbert spaces S and U (as has been observed already, this was proved for the conditions 1*)–5*) with different methods in [5]–[7]).

2.3 We shall establish some properties of the quantity $r(C,A)$.

LEMMA 2. *If condition* 1*) *is met, then* $r(C,A) \le m-1$.

PROOF. Condition 1* means that for some $\delta > 0$

$$\sum_{k=0}^{m-1} \|CA^k f\| \ge \delta \|f\| \, (f \in X). \tag{2.1}$$

If $f \in X$ and $g \in M_\varepsilon(C)$, then

$$\|CA^k f\| \le \|Cg\| + \|C(A^k f - g)\| \le \varepsilon\|g\| + \|C\|\,\|A^k f - g\|,$$

and hence

$$\|CA^k f\| \le 2\varepsilon\|A^k f\| + \|C\|\rho(A^k f, M_\varepsilon(C)). \tag{2.2}$$

From (2.1) and (2.2) it follows that for $f \in M_\varepsilon(C)$

$$\|C\|\sum_{k=1}^{m-1}\rho(A^k f, M_\varepsilon(C)) \ge \delta\|f\| - 2\varepsilon\sum_{k=0}^{m-1}\|A^k f\| \ge (\delta - 2\varepsilon\sum_{k=0}^{m-1}\|A^k\|)\|f\|. \tag{2.3}$$

We choose $\varepsilon(>0)$ sufficiently small, in order that $2\varepsilon\sum_{k=0}^{m-1}\|A^k\| + \varepsilon\|C\|(m-1) < \delta$.
Then we obtain from (2.3)

$$\max_{1\le k<m} \rho(A^k f, M_\varepsilon(C)) > \varepsilon\|f\| \quad (f \in M_\varepsilon(C), f \ne 0).$$

Consequently, $r(C,A) \le m-1$. The Lemma is proved.■

LEMMA 3. *There exists a number $\gamma > 0$ (depending on C and A) such that $r(C_1,A) \le r(C,A)$ if $\|C - C_1\| < \gamma$.*

PROOF. If $r(C,A) = r < \infty$, then for some $\varepsilon > 0$

$$\max_{1\le k\le r} \rho(A^k f, M_\varepsilon(C)) > \varepsilon\|f\| \qquad (f \in M_\varepsilon(C), f \ne 0),$$

We choose $\gamma \in (0,\varepsilon)$. As $M_{\varepsilon-\gamma}(C_1) \subset M_\varepsilon(C)$ for $\|C_1 - C\| < \gamma$, one has

$$\max_{1\le k\le r} \quad \rho(A^k f, M_{\varepsilon-\gamma}(C_1)) > (\varepsilon-\gamma)\|f\| \qquad (f \in M_{\varepsilon-\gamma}\ (C_1), f \ne 0),$$

and hence $r(C_1,A) \le r$. The Lemma is proved.■

LEMMA 4. *For arbitrary $\lambda \in \mathbb{C}$ one has $r(C, A-\lambda I) = r(C,A)$.*

This statement follows from the inequality

$$\rho((A-\lambda I)^k f,\ M_\varepsilon(C)) \le \sum_{j=0}^{k}\binom{k}{j}|\lambda|^{k-j}\rho(A^j f, M_\varepsilon(C)).$$

From Lemma 4 follows

LEMMA 5. *For the deduction of 4*) from 6*) it suffices to consider the case $\lambda = 0$, and for the deduction of 3*) from 6*) the case $0 \in K$ suffices.*

2.4 Here we prove some simple auxiliary propositions.

LEMMA 6. *Let X and U be Hilbert spaces and let $C\in\mathcal{L}(X,U)$. For arbitrary $\varepsilon>0$ one can find decompositions of the spaces X and U into orthogonal sums: $X=X_\varepsilon\oplus X^\varepsilon$, $U=U_\varepsilon\oplus U^\varepsilon$ such that $CX_\varepsilon\subset U_\varepsilon$, $CX^\varepsilon\subset U^\varepsilon$ and $\|Cf\|\le\varepsilon\|f\|(f\in X_\varepsilon)$, $\|Cf\|\ge\varepsilon\|f\|$ $(f\in X^\varepsilon)$.*

PROOF. We consider the polar decomposition $C=VD$ of the operator C, where $D\ge 0$ $(D\in\mathcal{L}(X))$ and $V(\in\mathcal{L}(X,U))$ is a partially isometric operator which is isometric on Im D. It is easy to convince oneself that for X_ε one can choose the spectral subspace of the operator D corresponding to the interval $[0,\varepsilon]$.■

LEMMA 7. *Under the assumptions of Lemma 6 one finds for arbitrary $\varepsilon>0$ an operator $R_\varepsilon\in\mathcal{L}(U)$ such that $\|C-R_\varepsilon C\|<\varepsilon$ and Im $(R_\varepsilon C)$ is closed.*

For this it suffices to choose for R_ε the orthogonal projection on the subspace U^ε (see Lemma 6).

LEMMA 8. *Let the conditions of Lemma 6 be satisfied and let $\|Cf_n\|\to 0$ $(f_n\in X)$. Then $\rho(f_n,X_\varepsilon)\to 0$ for arbitrary $\varepsilon>0$.*

PROOF. Let $f_n=g_n+h_n$ $(g_n\in X_\varepsilon, h_n\in X^\varepsilon)$. As $(Cg_n,Ch_n)=0$ one has $\|Cf_n\|\ge\|Ch_n\|\ge\varepsilon\|h_n\|$. Hence $h_n\to 0$. The Lemma is proved.■

2.5 In this section will be proved

THEOREM 8. *If X and U are Hilbert spaces then the property 6*) implies the properties 3*) and 4*). Moreover, one can assume that the number m in 4*) is not larger than $r(C,A)+1$.*

PROOF. 1. At first we shall consider separately the simple case dim $X<\infty$ (in order that the basic ideas of the further proof become more transparent).

We put $r=r(C,A)=r_1(C,A)$ (Lemma 1) and we denote by M_k the subspace of all $f\in X$ for which $A^jf\in\operatorname{Ker} C$ $(j=0,1,\dots,k)$. Evidently, $\operatorname{Ker} C=M_0\supseteq M_1\supseteq\dots\supseteq M_r=\{0\}$. From the set $\{1,\dots,r\}$ we select those numbers $k_1<k_2<\dots<k_s$ for which $M_{k_i}\ne M_{k_i-1}$. Evidently, $k_s=r$. We set $G_i=M_{k_{i-1}}\ominus M_{k_i}$ and let $\{l_{ij}\}_{j=1}^{p_i}$ be a basis in G_i. It is not difficult to verify that the vectors A^kl_{ij} $(j=1,\dots,p_i;\ k=0,\dots,k_i-1;\ i=1,\dots,s)$ form a basis in Ker C. By M we shall denote the subspace spanned by the vectors $A^{k_i}l_{ij}$ $(j=1,\dots,p_i;$ $i=1,\dots,s)$ and by N the orthogonal complement in X of Ker $C\oplus M$.

As the subspace $M\oplus N$ is a direct complement of Ker C one can, by an appropriate choice of the operator F, define the operator FC, and hence also $A-FC$, in an arbitrary way on this subspace. In particular, if $(A-FC)f=0$ for all $f\in M\oplus N$, then, as is easily seen, the identity $(A-FC)^{r+1}=0$ will hold. This proves 4*) (see Lemma 5).

For the proof of 3*) we shall select F such that the operator $A-FC$ maps N into itself, having a prescribed spectrum (some subset of K) and

$$(A-FC)A^{k_i}l_{ij} = -\sum_{t=0}^{k_i-1} a_{ij}^t A^t l_{ij} \qquad (j=1,\dots,p_i;\ i=1,\dots,s),$$

where the a_{ij}^t are certain numbers. Each of the subspaces $\mathcal{R}_{ij}$ $(j=1,\dots,p_i; i=1,\dots,s)$ spanned by the vectors $\{A^t l_{ij}\}_{t=0}^{k_i}$ will be invariant for the operator $A-FC$, whereas the characteristic polynomial of the restriction $(A-FC)|_{\mathcal{R}_{ij}}$ is

$$p_{ij}(\lambda) = \lambda^{k_i} + \sum_{t=0}^{k_i-1} a_{ij}^t \lambda^t.$$

Thus it is possible to make the spectrum of $(A-FC)|_{\mathcal{R}_{ij}}$ equal to an arbitrary set, consisting of at most k_i+1 points. Consequently, the spectrum of $A-FC$ can be an arbitrary set K, containing not more than $\dim X$ points (one might assume that K consists of exactly $\dim X$ points, among which some may be identical; under this assumption the number of repetitions of some number λ_0 will concide with the multiplicity of the eigenvalue λ_0 of the operator $A-FC$).

2. We turn to the examination of the case $\dim X=\infty$. We shall show, first of all, that is suffices to carry out the proof for an operator C which has a closed set of values (this step is also used in [5], [7]). By Lemma 7 an operator $R_\varepsilon \in \mathcal{L}(U)$ is found, for arbitrary $\varepsilon>0$, such that $\|C-R_\varepsilon C\| \le \varepsilon$ and $\overline{\operatorname{Im}\ (R_\varepsilon C)} = \operatorname{Im}\ (R_\varepsilon C)$. According to Lemma 3 $r(R_\varepsilon C,A) \le r(C,A)$. It is easy to see that from the validity of the conditions 3*) or 4*) for the operators $R_\varepsilon C, A$ follows their validity for C,A as well.

Using the established results we shall assume in the sequel that $\overline{\operatorname{Im}\ C} = \operatorname{Im}\ C$. We shall denote the value of $r(C,A)=r_0(C,A)$ (Lemma 1) by r. According to the definition of $r_0(C,A)$ there exists $\varepsilon_0>0$ such that

$$\max_{1\le k\le r} \|(I-P)A^k f\| \ge \varepsilon_0 \|f\| \, (f \in \operatorname{Ker}\ C), \tag{2.4}$$

where $P(\in \mathcal{L}(X))$ is the orthogonal projection onto $\operatorname{Ker}\ C$. We fix some $\varepsilon \in (0,\varepsilon_0)$. After application of Lemma 6 to the subspaces $\operatorname{Ker}\ C$, $X\ominus\operatorname{Ker}\ C$ and the operator $(I-P)A$ we obtain the existence of a subspace $H_1 \subset \operatorname{Ker}\ C$, with the properties $\|(I-P)Af\| \le \varepsilon \|f\|$ $(f\in H_1)$, $\|(I-P)Af\| \ge \varepsilon\|f\|$ $(f \in \operatorname{Ker} C \ominus H_1)$ and $(I-P)AH_1$ is orthogonal to $(I-P)A(\operatorname{Ker} C\ominus H_1)$. Next we apply Lemma 6 to the subspaces $H_1, X\ominus\operatorname{Ker}\ C$ and the operator $(I-P)A^2$ and so on. As result we obtain a set of subspaces $H_1 \supseteq H_2 \supseteq \dots \supseteq H_r$ such that

$$\|(I-P)Af\| \le \varepsilon\|f\| \ (f\in H_k), \ \ \|(I-P)A^k f\| \ge \varepsilon \|f\| \ (f \in H_{k-1}\ominus H_k)$$

and $(I-P)A^k H_k$ is orthogonal to $(I-P)A^k(H_{k-1}\ominus H_k)$. From inequality (2.4) it follows that $H_r=\{0\}$. We shall denote H_{r-1} by H.

3. We observe that (2.4) implies the inequality

$$\max_{0\le k<r}\|A^k\|\,\|Af\|\ge\varepsilon_0\|f\| \quad (f\in\operatorname{Ker} C),$$

and hence A maps Ker C isomorphically onto $A(\operatorname{Ker} C)$.

Now we show that the subspaces $A(\operatorname{Ker} C)$ and H intersect only at the zero vector and that their direct sum is closed. Assume that this is not the case. Then one can find sequences $f_n\in\operatorname{Ker} C$ and $h_n\in H$ such that $\|Af_n-h_n\|\to 0$, $\|h_n\|\ge\delta>0$. Consequently

(2.5) $$\|A^{k+1}f_n-A^kh_n\|\to 0 \qquad (n\to\infty; k=0,\ 1,\dots).$$

As $h_n\in H$ one has

(2.6) $$(I-P)h_n=0, \qquad \|(I-P)A^kh_n\|\le\varepsilon\|h_n\| \quad (k=1,\dots,r-1).$$

From (2.5) and (2.6) follows the inequality

$$\|(I-P)A^{k+1}f_n\|\le 2\varepsilon\|h_n\|\le 3\varepsilon\|Af_n\|\le 3\varepsilon\|A\|\,\|f_n\|,$$

where $n\ge n_0$, $k=0,\ 1,\dots,\ r-1$. We require that the number ε also meets the condition $3\varepsilon\|A\|<\varepsilon_0$. Then we obtain a contradiction with (2.4). That $A(\operatorname{Ker} C)\dot{+}H$ is closed has been shown. Hence follows the existence of a direct complement G to the subspace H such that $A(\operatorname{Ker} C)\subseteq G$.

4. We write down the matrix of the operator A with respect to the decomposition $X=H\dot{+}G$:

$$A=\begin{pmatrix}0 & A_1\\ A_3 & A_2\end{pmatrix}$$

(the block in the left upper corner is equal to 0, as $AH\subseteq A(\operatorname{Ker} G)\subseteq G$). Relative to this decomposition one has $C=(0\quad C_2)$, where $0\in\mathcal{L}(H,U)$, $C_2\in\mathcal{L}(G,U)$. Evidently, $\operatorname{Ker} C_2=\operatorname{Ker} C\cap G$. We observe also that $A_1(\operatorname{Ker} C_2)=\{0\}$, since $A(\operatorname{Ker} C_2)\subseteq G$.

A fundamental moment in the proof is the ascertainment of the inequality

(2.7) $$r_0(C_2,A_2)\le r-1.$$

To this end we put

$$E=\begin{pmatrix}0 & A_1\\ A_3 & 0\end{pmatrix}, \qquad F=\begin{pmatrix}0 & 0\\ 0 & A_2\end{pmatrix}$$

and by $P_2\in\mathcal{L}(G)$ we shall denote the orthogonal projection onto Ker C_2. Having assumed that $r_0(C_2,A_2)\ge r$ one can choose a sequence $h_n\in\operatorname{Ker} C_2$ such that $\|h_n\|=1$ and

(2.8) $$\|(I_G-P_2)A_2^kh_n\|\to 0 \quad (n\to\infty;\ k=1,\dots,r-1).$$

From the identities $A^k=(E+F)^k$ and

$$E(\text{Ker } C_2)=\{0\} \tag{2.9}$$

it follows that

$$A^k h_n=\sum_{j=1}^{k=1} G_{kj}EF^j h_n+F^k h_n, \tag{2.10}$$

where $G_{kj}\in\mathcal{L}(X)$. With regard to (2.9) we have

$$EF^j h_n=E(I_G-P_2)F^j h_n=E\begin{pmatrix} 0 \\ (I_G-P_2)\ A_2^j\ h_n\end{pmatrix}. \tag{2.11}$$

From the relations (2.10), (2.11) and (2.8) it follows that

$$\|(I-P)A^k h_n\|\to 0 \qquad (n\to\infty; k=1,\dots,r), \tag{2.12}$$

whence for $k=1$ follows, with help of Lemma 8, that $\|h_n-h_n^{(1)}\|\to 0$, where $h_n^{(1)}\in H_1$. From this and from (2.12) for $k=2$ we have $\|(I-P)A^2h_n^{(1)}\|\to 0$, and, again by Lemma 8, that $\|h_n^{(1)}-h_n^{(2)}\|\to 0$ $(h_n^{(2)}\in H_2)$. Repeating the given argument we obtain that $\|h_n-h_n^{(r-1)}\|\to 0$ $(n\to\infty)$, where $h_n^{(r-1)}\in H_{r-1}=H$. However, this contradicts the fact that Ker $C_2\dot{+}H$ is closed, and hence inequality (2.7) is proved.

5. We pass to the direct proof of Theorem 8, in which we proceed by induction to $r=r(C,A)$. For $r=0$ the statement of the Theorem is proved without difficulty. In this case Ker $C=\{0\}$ and as $\overline{\text{Im } C}=\text{Im } C$ the operator C has a left inverse $C^{(-1)}$. Putting $F=AC^{(-1)}$ we obtain $A-FC=0$ and putting $F=(A-D)C^{(-1)}$, where $\sigma(D)=K$, we have $A-FC=D$. Now we assume that the statements of the Theorem have been verified for $r-1$. We shall establish their validity for r.

From the inequality (2.7) follows, because of the induction assumption, the existence of an operator $F_2\in\mathcal{L}(U,G)$ such that $(A_2-F_2C_2)^r=0$. We put

$$F=\begin{pmatrix} F_1 \\ F_2\end{pmatrix} \qquad (F_1\in\mathcal{L}(U,H))$$

and we shall choose F_1 such that $(A-FC)^{r+1}=0$. Evidently,

$$A-FC=\begin{pmatrix} 0 & A_1-F_1C_2 \\ D & A_2-F_2C_2\end{pmatrix},$$

where $D\in\mathcal{L}(H,G)$. Let $R\in\mathcal{L}(U)$ be the orthogonal projection onto the subspace Im $C_2=\text{Im } C$ and $C_2^{(-1)}\in\mathcal{L}(\text{Im } C,\text{Im}(I_G-P_2))$ the inverse of the operator C_2, acting from $\text{Im}(I_G-P_2)$ into Im C. We put $F_1=A_1C_2^{(-1)}R$.

Evidently, $A_1-F_1C_2=0$ and hence

$$(2.13) \qquad A-FC=\begin{pmatrix} 0 & 0 \\ D & A_2-F_2C_2 \end{pmatrix},$$

and since $(A_2-F_2C_2)^r=0$ one has

$$(2.14) \qquad (A-FC)^r=\begin{pmatrix} 0 & 0 \\ D_1 & 0 \end{pmatrix}, \qquad (D_1\in\mathcal{L}(H,G)).$$

From (2.13) and (2.14) it follows that $(A-FC)^{r+1}=0$. This proves Property 4*) for $\lambda_0=0$, and hence for arbitrary λ_0 (Lemma 5). For the proof of 3*) we act in an analogous way, but we select F_2 such that $\sigma(A_2-F_2C_2)=K$. This can be done by virtue of the induction assumption, since $\dim G=\infty$ (the latter identity follows from the relations $\dim X=\infty$ and $AH\subseteq G$). But then it follows form (2.13) that $\sigma(A-FC)=K\cup\{0\}$. Hence, in case $0\in K$, Property 3*) is established, and again it remains to refer to Lemma 5. The theorem is proved.■

2.6 From the Theorems 5, 6 and 8 follows

THEOREM 9. *If X and U are Hilbert spaces, then the properties* 1*)–6*) *are equivalent.*

By duality we obtain the validity of

THEOREM 10. *If X and U are Hilbert spaces then the properties* 1)–5) *and also the property*

6) $\qquad r(B^*,A^*)<\infty$

are equivalent.

REMARK 5. From the proofs it is clear that the smallest constant in the Properties 1*), 2*) and 4*) is the same and coincides with $r(C,A)+1$. The analogous statement is valid for the Properties 1), 2) and 4) and the number $r(B^*,A^*)+1$.

REMARK 6. The equivalence of the conditions 1)–5) (respectively 1*)–5*)) for Hilbert spaces X and U was established by different methods in the papers [5]–[7]. We note that these statements are valid under the condition that only U is a Hilbert space. Indeed, then it follows from the existence of just one pair of operators A,B (respectively, C,A) for which Property 1) (respectively 1*)) holds that X is isomorphic to a Hilbert space (see Remark 2).

REMARK 7. If only X is a Hilbert space then the Properties 1)–5) (and 1*)–5*)) are, generally speaking, not equivalent. We present examples where these equivalences hold nevertheless: If $X=l^2$, $U=l^2\oplus l^p$ $(2<p<\infty)$ then the Properties 1)–5) are equivalent and if $X=l^2$, $U=l^2\oplus l^p$ $(1<p<2)$ then the Properties 1*)–5*) are equivalent.

2.7 We shall explain in brief the relation of the results of Section 2 with the theory of holomorphic operator functions (for details see [6],[16]). Let $W(\lambda)$ be an operator function with values in $\mathcal{L}(U)$, which is holomorphic on some region $\Omega(\subseteq\mathbb{C})$, where the set $\Sigma(W)$ of those $\lambda\in\Omega$ for which the operator $W(\lambda)$ is not invertible (the spectrum of $W(\lambda)$) is compact. An operator $A\in\mathcal{L}(X)$ is called a *spectral linearization* of the operator function $W(\lambda)$ on the region Ω, if $\sigma(A)\subseteq\Omega$ and

$$\begin{pmatrix} W(\lambda) & 0 \\ 0 & I_X \end{pmatrix} = E(\lambda)\begin{pmatrix} I_U & 0 \\ 0 & \lambda I_X - A \end{pmatrix} F(\lambda) \qquad (\lambda\in\Omega),$$

where $E(\lambda)$, $F(\lambda)$ are holomorphic operator functions with values in $\mathcal{L}(U\oplus X)$ which are invertible for all $\lambda\in\Omega$.

We put

$$C=\frac{1}{2\pi i}\int_{\Gamma} \mathcal{P}F^{-1}(\lambda)\ \mathcal{T}(\tau I_X - A)^{-1}\ d\lambda \qquad (\in\mathcal{L}(X,U)),$$

where $\Gamma(\subset\Omega)$ is a closed rectifiable curve, enclosing $\Sigma(W)$, $\mathcal{P}$ is the canonical projection from $U\oplus X$ onto U, $\mathcal{T}$ is the canonical embedding of X into $U\oplus X$. The operator pair (C,A) is called a (right) *spectral pair* for the operator function $W(\lambda)$ on the region Ω. Spectral pairs play an important role in the study of various properties of holomorphic operator functions. We note two important results on spectral pairs: (1) If (C,A) is a spectral pair, then the operator $K_m(C,A)$ is left invertible; (2) If the operator $K_m(C,A)$ is left invertible, the operator A is invertible and $(A^{-1}-FCA^{-1})^m=0$ for some $F\in\mathcal{L}(U,X)$, then there exists an operator polynomial $P(\lambda)$ of degree $\leq m$ for which C,A is a spectral pair on $\mathbb{C}$ ([6]). It is evident that from the left invertibility of $K_m(C,A)$ follows that $K_m(CA^{-1},A^{-1})$ is left invertible. If U is a Hilbert space then one has by virtue of the results of this section that 2*) $\Rightarrow$ 4*), and hence the identity $(A^{-1}-FCA^{-1})^m=0$ for some $F\in\mathcal{L}(U,X)$ follows from 2*), which allows to stengthen some results of the paper [6]. In particular, in the Theorems 5.1, 6.2 and 6.3 of [6] one can discard the condition $\bar{\Omega}\neq\mathbb{C}$ and in the Theorems 7.1 and 7.2 the separability condition on H and G, whereas in the Theorems 7.1 and 7.3 one can replace $3m-2$ by m (this observation follows from the results of the paper [7] as well).

REFERENCES

1. N.N. Krasovskii, The Theory of Control by Motion, Moskow, 1960 (Russian).

2. V.M. Popov, Hyperstability of Automatic Systems, Moskow, 1970 (Russian).

3. W.M. Wonham, Linear Multivariable Control: a geometric approach, Springer–Verlag, Berlin, 1979.

4. V.I. Korobov and R. Rabakh, Exact controllability in Banach space, Dif. Equations 15 (1979), 1531–1537.

5. G. Eckstein, Exact controllability and spectrum assignment, in "Topics in Modern Operator Theory", Birkhäuser, Basel, 1981, 81–94.

6. M.A. Kaashoek, C.V.M. van der Mee and L. Rodman, Analytic operator functions with compact spectrum III. Hilbert space case: inverse problem and applications, J. Operator Theory 10 (1983), 219–250.

7. K. Takahashi, Exact controllability and spectrum assignment, J. Math. Anal. and Appl. 104 (1984), 537–545.

8. M.L.J. Hautus, Controllability and observability conditions of linear autonomous systems, Indag. Math. 31 (1969), 443–448.

9. J. Leiterer, Banach coherent analytic Fréchet sheaves, Math. Nachr. 85 (1978), 91–109.

10. K. Astala and H.–O. Tylli, On semi–Fredholm operators and the Calkin algebra, J. London Math. Soc. 34 (1986), 541–551.

11. J. Lindenstrauss and L. Tzafriri, Classical Banach Spaces I, Springer–Verlag, Berlin, 1977.

12. G. Bennett, Lectures on matrix transformations of ℓ^p spaces , in "Notes in Banach spaces", Austin, London, 1980, 39–80.

13. J. Diestel, Sequences and Series in Banach Spaces, Springer–Verlag, Berlin, 1984.

14. A. Pelczynski, Projections in certain Banach spaces, Studia Math. 19 (1960), 209–228.

15. N. Dunford and J.T. Schwarz, Linear operators I, Wiley–Interscience, New York, 1958.

16. M.A. Kaashoek, C.V.M. van der Mee and L. Rodman, Analytic operator functions with compact spectrum II. Spectral pairs and factorization, Integral Equations Operator Theory 5 (1982), 791–827.

Current addresses of the authors:

A. Markus
Department of Mathematics
and Computer Sciences
Ben Gurion University
of the Negev
Beer Sheva 84105
Israel

V. Olshevsky
School of Mathematical Sciences
Raymond and Beverly Sackler
Faculty of Exact Sciences
Tel Aviv University
Ramat Aviv 69978
Israel

Translator: Philip Thijsse (Rotterdam)

This paper was originally published in Diferentjial'nye uravnenia i matematicheskaya fizika, Mat. Issled., 106(1989), 97–113, Kishinev, 1989 (Russian).
The authors thank the translator of the paper and the editor of this journal.

MSC: 47A10, 93B05, 93B07, 93B55