

## INVARIANT SUBSPACES OF POLYNOMIALLY COMPACT OPERATORS ON BANACH SPACE

ALLEN R. BERNSTEIN

**This paper contains a proof of the following:**

**MAIN THEOREM.** Let  $T$  be a bounded linear operator on an infinite-dimensional Banach space  $B$  over the complex numbers. Suppose there exists a polynomial  $p(\lambda) \neq 0$  with complex coefficients such that  $p(T)$  is compact (completely continuous). Then  $T$  leaves invariant at least one closed linear subspace of  $B$  other than  $\{0\}$  or  $B$ .

Since in most of the common nonreflexive Banach spaces (e.g.  $l_1$ ,  $C[0, 1]$ ,  $AP$ , etc.) weakly compact operators have compact squares (cf. [7], pp. 511 and 580), one can conclude in particular from the above theorem that these operators have proper invariant subspaces.

The proof of the main theorem is carried out within the framework of A. Robinson's Theory of Nonstandard Analysis and follows lines similar to the proof presented in [3] for the special case where  $B$  is a Hilbert space, which settled a question raised by P. Halmos and K. Smith [9]. That proof made strong use of the fact that in a separable Hilbert space it is possible to choose a countable orthonormal basis. In a general separable Banach space, of course, one does not even know whether there is a basis, much less an "orthonormal" basis. However this difficulty may be overcome by the introduction of *metric projections* to take the place of projections in Hilbert space, as was done by Aronszajn and Smith in their proof of the existence of invariant subspaces for compact operators [1], and by the introduction of a *semi-basis* to take the place of the orthonormal basis in Hilbert space.

The proof in [3] was carried out within the framework of a nonstandard model of the real numbers. This was possible because the elements of a separable Hilbert space may be represented as sequences of complex numbers which in turn may be defined as ordered pairs of real numbers. However, in the case of a general Banach space no such convenient representation is apparent so that a more general version of nonstandard analysis is necessary. The remainder of this section is an introduction to such a version drawn essentially from the work of A. Robinson ([11]-[14]).

The class  $T$  of *types* is defined inductively as follows. (i) 0 is a type; (ii) if  $\tau_1, \dots, \tau_n$  are types,  $n \geq 1$ , then  $(\tau_1, \dots, \tau_n)$  is a type; (iii)  $T$  is the smallest class satisfying (i) and (ii).

A *higher order structure* is defined to be a generalized sequence

$\{A_\tau\}_{\tau \in \mathcal{T}}$  of sets (i.e., a mapping which assigns to each type  $\tau$  a set  $A_\tau$ ) such that  $A_0$  is nonempty and if  $\tau \neq 0$ ,  $\tau = (\tau_1, \dots, \tau_n)$ , then  $A_\tau$  is a set of subsets of  $A_{\tau_1} \times A_{\tau_2} \times \dots \times A_{\tau_n}$ . If  $n = 1$  then the preceding Cartesian product is understood to be just the set  $A_{\tau_1}$  and  $A_\tau$  is then a set of subsets of  $A_{\tau_1}$ . If  $a \in A_0$  then  $a$  is called an *individual* of  $M$  and if  $Q \in A_\tau$ ,  $\tau \neq 0$ , then  $Q$  is called a *relation* of  $M$ . If for each  $\tau \neq 0$ ,  $\tau = (\tau_1, \dots, \tau_n)$ ,  $A_\tau$  is the set of *all* subsets of  $A_{\tau_1} \times A_{\tau_2} \times \dots \times A_{\tau_n}$ , then  $M$  is called the *complete structure* of  $A_0$ .

Now let  $B$  be the set of elements of a Banach space over the set of complex numbers  $C$ . Let  $A_0 = B \cup C$  and let  $M = \{A_\tau\}_{\tau \in \mathcal{T}}$  be the complete structure of  $A_0$ . Thus among the relations of  $M$  are  $B$  and  $C$  together with all their subsets, relations, relations of relations, etc. In particular the set of real numbers  $R$  and the set of positive integers  $Z$ , being subsets of  $C$ , will be relations of  $M$ . Furthermore we may single out the relations under which  $B$  is a Banach space over the complex numbers  $C$ . Thus for example, we can find in  $M$  the relation  $S(x, y, z)$  which holds just in case  $x, y, z \in C$  and  $x - y = z$  and the relation  $A(x, y, z)$  which holds just in case  $x, y, z \in B$  and  $x + y = z$ . In this manner all the algebraic operations over  $B$  and  $C$  are relations of  $M$ . Similarly we can find in  $M$  the relation  $N(x, y)$  which holds just in case  $x \in B$ ,  $y \in R$  and the norm of  $x$ ,  $\|x\|$ , is equal to  $y$ . A function of  $n$  variables in this system is a relation of  $n + 1$  variables in which the first  $n$  coordinates determine the  $(n + 1)$ st coordinate uniquely. Thus, for example, a sequence of elements of  $B$ ,  $\{x_n\}$ , which is a function from  $Z$  into  $B$ , is a two-placed relation  $S(n, y)$  which holds just in case  $y = x_n$ .

Next we choose a suitable higher-order predicate language  $L$  which includes distinct constant symbols for each individual and relation of  $M$ .  $L$  contains variables ranging over each type and quantification is permitted over each variable. A suitable notion of when a sentence of  $L$  is true in a higher order structure  $M$  (under a suitable correspondence between the constant symbols of  $L$  and the individuals and relations of  $M$ ) may then be defined (see [12] or [2]). It then follows from a completeness theorem of Henkin ([10]) that there exists a non-standard higher order structure  $*M = \{B_\tau\}_{\tau \in \mathcal{T}}$ , where  $B_0$  properly includes  $A_0$ , such that under a suitable correspondence between the constant symbols of  $L$  and the elements of the  $B_\tau$ 's, any sentence of  $L$  which is true in  $M$  is true also in  $*M$ . Furthermore any constant symbol of  $L$  which denotes some individual  $a \in A_0$  in  $M$  will again denote  $a$  in  $*M$ , now regarded as an element of  $B_0$ . If a sentence of  $L$  is known to be true in  $M$ , we say its truth in  $*M$  is established by *transfer* from  $M$  to  $*M$ . In  $*M$  the variables of  $L$  range only over the elements of the  $B_\tau$ 's so that, for example, a sentence of  $L$

which asserts that a certain property is true for all sets of individuals will be true in  $*M$  if it is true for each element of  $B_{(0)}$ . Any element of one of the  $B_\tau$ 's,  $\tau \neq 0$ , is called an *interior* (or *admissible*) relation.

Now let  $Q$  be a relation of  $M$ ,  $Q \in A_\tau$ .  $Q$  is denoted by some constant symbol  $\beta$  of  $L$  and  $\beta$  must denote in  $*M$  some element of  $B_\tau$  which we write as  $*Q$ .  $*Q$  will have the same properties as  $Q$  to the extent to which these can be expressed as sentences of  $L$ . In particular  $B, C, R$ , and  $Z$  will extend in  $*M$  to the sets  $*B, *C, *R, *Z$  which properly include  $B, C, R$ , and  $Z$  respectively. Thus the nonstandard Banach space  $*B$  will contain in addition to all standard points of  $B$ , other nonstandard points which are not elements of  $B$ . Similarly  $*Z$  will contain elements which do not belong to  $Z$  and these are called *infinite integers*.

Any function  $f$  which is a relation of  $M$  corresponds to a relation  $*f$  of  $*M$  which again must be a function since this fact can be expressed as a sentence of  $L$ . If, in addition,  $f$  is a function from individuals to individuals, then  $*f$  will be an extension of  $f$ . Thus for example a sequence  $\{x_n\}$  of elements of  $B$  (or  $C$ ) which is a function  $f$  from  $Z$  into  $B$  (or  $C$ ), extends to a "sequence"  $*\{x_n\}$  which is just the function  $*f$  defined now for all  $n \in *Z$  and taking values in  $*B$  (or  $*C$ ). If  $m \in *Z$ , then  $x_m$  will designate the element of  $*B$  (or  $*C$ ) corresponding to  $m$  in  $*\{x_n\}$ , i.e.,  $x_m = *f(m)$ . Since  $*f(m) = f(m)$  for all  $m \in Z$ , this designation is unambiguous.

If  $\Sigma$  is the set of sequences in  $M$ , then we refer to any element of  $\Sigma$  as an  $S$ -sequence. Its extension to  $*M$ ,  $*\Sigma$ , will contain all standard sequences, that is all extensions  $*\{x_n\}$  of  $S$ -sequences  $\{x_n\}$  in  $\Sigma$ . In addition,  $*\Sigma$  will contain other elements which must be functions from  $*Z$  into  $*A_0$  since this fact can be expressed as a sentence of  $L$ . Any element  $\{t_n\}$  of  $*\Sigma$  we call a quasi-standard sequence, or briefly a  $Q$ -sequence.

For the algebraic operations on  $B$  and  $C$  we use the same symbol for their extensions to  $*M$ . Thus for example the extension of the relation of addition,  $+$ , for complex numbers is in  $*M$  a relation over elements of  $*C$  and will again be denoted by  $+$ . Likewise the norm  $\|\cdot\|$  which in  $M$  is a function from  $B$  into  $R$  extends to a function from  $*B$  into  $*R$  which is again denoted by  $\|\cdot\|$ .

Let  $c$  be an element of  $*C$ . If  $|c| < r$  for all  $r \in R$  then  $c$  is called *infinitesimal*. If  $|c| < s$  for some  $s \in R$  then  $c$  is called *finite*. In this case there is a unique complex number  ${}^{\circ}c \in C$  such that  $c - {}^{\circ}c$  is *infinitesimal*. We write  $c_1 \sim c_2$  if  $c_1 - c_2$  is *infinitesimal* and it is easily verified that this is an equivalence relation over  $*C$ . The proof of the next theorem may be found in [11] or [13].

**THEOREM 1.1.** *The  $S$ -sequence  $\{s_n\}$  of standard real numbers*

converges to the standard real number  $s$ ,  $\lim_{n \rightarrow \infty} s_n = s$ , if and only if  $|s - s_\omega|$  is infinitesimal for all infinite positive integers  $\omega$ .

Let  $x$  be an element of  ${}^*B$ . If  $x \in B$  we say  $x$  is *standard*. If  $\|x\| < r$  for some  $r \in R$ ,  $x$  is called *norm-finite*. If  $\|x\| < r$  for all  $r \in R$ , then  $x$  is called *infinitesimal*. We write  $x \sim y$  for two points  $x, y \in {}^*B$  such that  $x - y$  is infinitesimal and this is an equivalence relation over  ${}^*B$ . It is easy to verify that if  $x_1 \sim y_1$  and  $x_2 \sim y_2$  then  $x_1 + x_2 \sim y_1 + y_2$  and that  $cx_1 \sim cy_1$  providing  $c$  is a finite element of  ${}^*C$ . If there is a standard  $y \in {}^*B$  (i.e.  $y \in B$ ) such that  $x \sim y$  then  $x$  is called *near-standard*. In this case there is a unique such  $y$  and we write  $y = {}^\circ x$ .  ${}^*B$  will in general contain points which are norm-finite but not near-standard.

**THEOREM 1.2.** *Let  $x \in {}^*B$  be not near-standard. Then there is a standard  $\varepsilon > 0$  such that  $\|x - b\| > \varepsilon$  for all standard  $b$  in  ${}^*B$ .*

*Proof.* Suppose that the conclusion of the theorem is false. For each  $k \in Z$ , we may therefore pick an element  $b_k$  in  $B$  such that  $\|x - b_k\| < 1/2k$ . Then if  $m, n > k$ ,

$$\|b_m - b_n\| \leq \|b_m - x\| + \|x - b_n\| < \frac{1}{2m} + \frac{1}{2n} < \frac{1}{2k} + \frac{1}{2k} = \frac{1}{k}.$$

Thus  $\{b_n\}$  is a Cauchy sequence and converges to some point  $b$  in  $B$ . Now choose any standard  $\varepsilon > 0$ , and let  $n \in Z$  be greater than  $1/\varepsilon$  and such that  $\|b - b_n\| < \varepsilon/2$ . Then

$$\|x - b\| \leq \|x - b_n\| + \|b_n - b\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This shows that  $\|x - b\|$  is less than any standard  $\varepsilon > 0$  and must therefore be infinitesimal. Since  $b$  is standard,  $x$  is near-standard and this contradiction proves the theorem.

**THEOREM 1.3.** *Let  $S$  be a compact set of points of  $B$ . Then every point of  ${}^*S$  is near-standard.*

*Proof.* Suppose there is a point  $x \in {}^*S$  which is not near-standard. Using the previous theorem, let  $\varepsilon$  be a standard positive number such that  $\|x - b\| > \varepsilon$  for all standard points  $b$ . Since  $S$  is compact it contains standard points  $b_1, b_2, \dots, b_n, n \in Z$ , such that it is true in  $M$  that "for every point  $\xi$  in  $S$  at least one of the numbers  $\|\xi - b_1\|, \|\xi - b_2\|, \dots, \|\xi - b_n\|$  is smaller than  $\varepsilon$ ". (We are excluding the trivial case where  $S$  is empty.) The statement in quotes may be expressed as a sentence of  $L$  and therefore is true also in  ${}^*M$ . Specifying

$\xi$  to  $x$ , we conclude that  $\|x - b_i\| < \varepsilon$  for some  $i$ ,  $1 \leq i \leq n$ . This contradicts our assumption that  $\|x - b\| > \varepsilon$  for all standard points  $b$  and completes the proof.

Let  $V$  be a relation of  $M$  of type  $(\tau)$  so that  $V$  is a set consisting of elements of type  $\tau$ . If the elements of  $V$  are described by some common mathematical name then this name may be used again to refer to the elements of  $*V$ . For example, if  $V$  is the set of bounded operators on  $B$ , the set of projections on  $B$ , or the set of linear subspaces of  $B$ , then the elements of  $*V$  will be referred to as "bounded operators on  $*B$ ", "projections on  $*B$ ", or "linear subspaces of  $*B$ ". We may also use the expressions "Q-operator", "Q-projection", etc.

Let  $W$  be the relation of  $M$  which is the set of all the bounded operators on  $B$ . There is a real-valued norm function,  $\|\cdot\|$ , defined on elements  $T \in W$  by

$$\|T\| = \sup_{\substack{x \in B \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}.$$

As we pass to  $*M$ , this function extends to a function from  $*W$  (the set of bounded operators on  $*B$ ) to  $*R$  which will again be denoted by  $\|\cdot\|$  and which has the property that  $\|Tx\| \leq \|T\|\|x\|$  for all  $T \in *W$  and all  $x \in *B$ . For  $T \in W$ , we denote its extension to an operator on  $*B$  by  $*T$  as usual.

**THEOREM 1.4.** *Let  $T$  be a bounded operator on  $B$ . Then  $*T$  transforms every norm-finite point of  $*B$  into a norm-finite point.*

*Proof.* If  $x$  is a norm-finite point of  $*B$ , then

$$\|*Tx\| \leq \|*T\|\|x\| = \|T\|\|x\|$$

since  $\|*T\| = \|T\|$ . Thus  $\|*Tx\|$  is less than the finite element of  $*R$ ,  $\|T\|\|x\|$ , and hence  $*Tx$  is norm-finite.

A bounded linear operator  $T$  on  $B$  is called *compact* (or completely continuous) if it transforms every bounded set of points of  $B$  into a conditionally compact set, i.e., a set whose closure is compact. Finally we have the following fundamental characterization of compact operators given by Robinson.

**THEOREM 1.5.** *Let  $T$  be a compact operator on  $B$ . Then  $*Tx$  is near-standard for all norm-finite  $x$  in  $*B$ .*

*Proof.* If  $x$  is norm-finite then  $\|x\| < r$  for some standard  $r \in R$ . The set  $D = \{\xi \mid \|\xi\| < r\}$  in  $B$  is bounded and is therefore mapped by  $T$  on a set whose closure,  $A$ , is compact. If the corresponding sets

in  $*B$  are  $*D$  and  $*A$  respectively then  $*D$  contains  $x$  (since  $x$  satisfies the defining condition of  $D$ ) and so  $*A$  contains  $*Tx$ . But  $*A$  contains only near-standard points, by 1.3, so  $*Tx$  is near-standard, proving 1.5.

**2. Semi-bases in Banach space.** Let  $X$  be a set of elements of the infinite-dimensional complex Banach space  $B$ . The smallest closed linear subspace  $S$  of  $B$  which contains all the elements of  $X$  is called the subspace of  $B$  *generated* by  $X$ . Alternately we say  $X$  *spans*  $S$  and write  $\text{sp } X = S$ . If  $X$  is a finite set, then  $\text{sp } X$  is just the set of linear combinations of elements of  $X$ . Similarly, for a sequence  $\{x_i\}$  of elements of  $B$ , we designate the smallest closed linear subspace containing all the  $x_i$  by  $\text{sp } \{x_i\}$ .

The following results through (2.7) may be found in [1].

Let  $x$  be an arbitrary nonzero element of  $B$ . The closed subspace of  $B$ ,  $\text{sp } \{x, Tx, T^2x, \dots\} \neq \{0\}$ , is clearly an invariant subspace for  $T$ . Therefore we shall limit ourselves to the case where

$$(2.1) \quad B = \text{sp } \{x, Tx, T^2x, \dots\}.$$

In particular, (2.1) implies immediately that  $B$  is separable. Furthermore, since we are assuming  $B$  is infinite-dimensional it must be the case that all the elements  $T^n x$ ,  $n = 0, 1, 2, \dots$ , are unequal to zero and are linearly independent.

A Banach space norm,  $\|\cdot\|$ , is called *strictly convex* if for all  $x, y$  in  $B$ ,  $x \neq y$  and  $\|x\| = \|y\| \neq 0$  imply that  $\|x + y\|$  is strictly less than  $\|x\| + \|y\|$ . By a theorem of Clarkson [4], one may define in any separable Banach space an equivalent strictly convex norm. Since the property of being a proper invariant subspace does not change upon passage to an equivalent norm, we shall suppose hereafter that the norm in  $B$  is strictly convex.

Consider any arbitrary finite-dimensional subspace  $S \subset B$ . For each  $x \in B$  we may consider the minimal distance  $\rho(x, S)$  from  $x$  to  $S$ . Since  $S$  is finite-dimensional, the shortest distance is attained and by the strict convexity of the norm it can be immediately shown that there exists a *unique* point  $Px \in S$  which realizes this minimal distance, i.e.,

$$\|x - Px\| = \rho(x, S) = \min_{y \in S} \|x - y\|.$$

$Px$  represents a bounded projection operation on  $B$  which in general is nonlinear.  $P$  is called the *metric projection* on  $S$ . We list here a few properties of  $P$  which are immediate consequences of its definition.

$$(2.2) \quad P^2 = P$$

$$(2.3) \quad P(\lambda x) = \lambda(Px) \quad \text{for every } \lambda \in C, x \in B.$$

$$(2.4) \quad P(x + y) = x + Py \quad \text{for every } x \in S, y \in B.$$

$$(2.5) \quad \|x - Px\| \leq \|x\|, \quad \|Px\| \leq 2\|x\| \quad \text{for every } x \in B.$$

$$(2.6) \quad \|x - Px\| \leq \|x - y\| + \|y - Py\| \quad \text{for every } x, y \in B.$$

(2.7) If  $S' \subset S$  and  $P'$  is the metric projection on  $S'$ , then

$$\|x - Px\| \leq \|x - P'x\| \quad \text{for every } x \in B.$$

DEFINITION. A sequence  $\{e_i\}$ ,  $i = 1, 2, 3, \dots$  of elements of  $B$  is a *semi-basis* of  $B$  if

$$(2.8) \quad \|e_i\| = 1; \quad i = 1, 2, 3, \dots$$

$$(2.9) \quad \text{sp}\{e_i\} = B.$$

(2.10) For each  $n > 1$ ,  $P_{n-1}e_n = 0$  where  $P_{n-1}$  is the metric projection of  $B$  onto  $\text{sp}\{e_1, e_2, \dots, e_{n-1}\}$ .

For a given semi-basis  $\{e_i\}$  of  $B$ , we denote by  $B_n$  the subspace  $\text{sp}\{e_1, e_2, \dots, e_n\}$  and by  $P_n$  the metric projection of  $B$  onto  $B_n$ . Then

(2.11) For any integer  $n \in Z$  and any complex number  $\alpha \in C$ , if

$$z = x + \alpha e_{n+1} \quad \text{where } x \in B_n, \text{ then } z - P_n z = \alpha e_{n+1}.$$

*Proof.* By (2.4) and (2.3),  $P_n z = x + \alpha P_n e_{n+1}$ . But  $P_n e_{n+1} = 0$  by (2.10), so that  $P_n z = x$ . Therefore  $z - P_n z = x + \alpha e_{n+1} - x = \alpha e_{n+1}$ , which establishes (2.11).

As in the preceding section we may now suppose that the elements of  $B$  are individuals within a complete structure  $M$  which includes also the field of complex numbers  $C$ , and within it, the field of real numbers  $R$  and the set of positive integers  $Z$ , where  $C \supset R \supset Z$ . As we pass to an enlargement  $*M$  of  $M$ ,  $B$  is extended to a set  $*B$ , and  $C, R$ , and  $Z$  are extended to  $*C, *R$ , and  $*Z$  respectively, where  $*C \supset *R \supset *Z$ .

Now let  $\{e_i\}$  be a semi-basis of  $B$ . As we pass to  $*M$ ,  $\{e_i\}$  is extended to a sequence  $*\{e_i\}$  of elements of  $*B$ , where  $i$  now ranges over  $*Z$ . If  $\omega$  is an infinite positive integer, then  $B_\omega$  will designate the subspace of  $*B$  spanned by  $\{e_1, e_2, \dots, e_\omega\}$ .  $B_\omega$  consists of all elements of  $*B$  which can be expressed in the form  $\sum_{i=1}^\omega \alpha_i e_i$  with  $\alpha_i \in *C$ ,  $i = 1, 2, \dots, \omega$ . However this summation sign cannot now be regarded as representing the iteration of the addition operation in the ordinary manner, but must be regarded as defined by transfer from  $M$ . For this purpose, let  $S_m(n, x)$  be a relation of  $M$  which is a function from  $\{i \in Z \mid 0 < i \leq m\}$  into  $C$ , and let  $V$  be the set of all such relations as  $m$  varies over  $Z$ . That is,  $V$  is the set of all finite sequences of elements of  $C$ . Now let  $R(y, z)$  be the relation in  $M$  which holds just

in case  $y \in V$ , say  $y = S_m(n, x)$ , and  $z = \sum_{i=1}^m \beta_i e_i$ , where  $\beta_i$  is the unique element of  $C$  such that  $S_m(i, \beta_i)$  holds. Then the statement that  $z = \sum_{i=1}^{\omega} \alpha_i e_i$  is equivalent to the statement that there is an element  $y \in {}^*V$  such that  ${}^*R(y, z)$  holds, and  $y = S_{\omega}(n, x)$  is a function from  $W = \{i \in {}^*Z \mid 0 < i \leq \omega\}$  into  ${}^*C$  such that for each  $i \in W$ ,  $S_{\omega}(i, \alpha_i)$  holds.

The subspace  $B_{\omega}$  of  ${}^*B$  is "finite-dimensional" in the sense of non-standard analysis. That is,  $B_{\omega}$  satisfies any sentence of  $L$  which is true for all finite-dimensional subspaces of  $B$ . In particular we may define a metric projection  $P_{\omega}$  from the whole space  ${}^*B$  onto  $B_{\omega}$  and which satisfies (2.3)–(2.7) and (2.11), where of course we must replace  $C, B$ , and  $Z$  by  ${}^*C, {}^*B$ , and  ${}^*Z$ . We then have the following two theorems which are analogous to similar theorems concerning non-standard Hilbert space.

**THEOREM 2.12.** *If  $x$  is a standard point of  ${}^*B$  (i.e.,  $x \in B$ ), then  $\|x - P_{\omega}x\|$  is infinitesimal for all infinite integers  $\omega$ .*

*Proof.* Let  $x \in B$ . Then it follows from (2.9) that for any standard  $\varepsilon > 0$ , there exists an  $n \in Z$  and  $y \in B_n$  such that  $\|x - y\| < \varepsilon$ . Since  $P_n$  is the nearest point operator onto  $B_n$ ,  $\|x - y\| \geq \|x - P_n x\|$ , and therefore  $\|x - P_n x\| < \varepsilon$ . Now if  $m > n$ , then  $B_n \subset B_m$  and we may use (2.7) to obtain  $\|x - P_m x\| < \varepsilon$ . Thus the  $S$ -sequence  $\{\|x - P_n x\|\}$  converges to 0, so by Theorem 1.1,  $\|x - P_{\omega}x\|$  is infinitesimal for all infinite integers  $\omega$ .

**THEOREM 2.13.** *If  $x$  is a near-standard element of  ${}^*B$ , then  $x$  is norm-finite and  $\|x - P_{\omega}x\|$  is infinitesimal for all infinite integers  $\omega$ .*

*Proof.* Let  $x \in {}^*B$  and suppose  $\|x - {}^{\circ}x\|$  is infinitesimal for some  ${}^{\circ}x \in B$ . Then by the triangle inequality,

$$\|x\| \leq \|x - {}^{\circ}x\| + \|{}^{\circ}x\| \leq 1 + \|{}^{\circ}x\|$$

which shows that  $x$  is norm-finite. Furthermore, it follows from (2.6) that for any infinite integer  $\omega$ ,

$$\|x - P_{\omega}x\| \leq \|x - {}^{\circ}x\| + \|{}^{\circ}x - P_{\omega}{}^{\circ}x\|.$$

Now  $\|x - {}^{\circ}x\|$  is infinitesimal by assumption and  $\|{}^{\circ}x - P_{\omega}{}^{\circ}x\|$  is infinitesimal by the previous theorem, thus  $\|x - P_{\omega}x\|$  is infinitesimal, proving (2.13).

It is interesting to note that (2.13) does not follow directly from (2.12) as it would if  $P_{\omega}$  were a linear operator. Instead, we had to use also the property (2.6) which  $P_{\omega}$  possesses since for  $\xi \in B$ ,  $P_{\omega}\xi$  is the point of  $B_{\omega}$  which is nearest to  $\xi$ .

3. **Infinite matrices.** We start out by working in the standard model  $M$ . Let  $[a_{jk}]$  be an infinite matrix where  $j$  and  $k$  range over  $Z$  and the entries  $a_{jk}$  are in the field of complex numbers  $C$ .  $[a_{jk}]$  is called *column-finite* if for each  $k$  there is an  $n$  such that  $a_{jk} = 0$  for all  $j > n$ .

Let  $[a_{jk}]$  and  $[b_{jk}]$  be column-finite matrices with entries in  $C$ . We may form the matrix sum  $[d_{jk}] = [a_{jk}] + [b_{jk}]$  in the usual way where the entries  $d_{jk}$  are given by the formula,

$$(3.1) \quad d_{jk} = a_{jk} + b_{jk} .$$

It is clear that  $[d_{jk}]$  must then also be column-finite. Similarly, we may form the matrix product  $[c_{jk}] = [a_{jk}] \cdot [b_{jk}]$  by means of the formula,

$$(3.2) \quad c_{jk} = \sum_{i=1}^{\infty} a_{ji} b_{ik} .$$

Since for a given  $k$ ,  $b_{ik} = 0$  except possibly for finitely many  $i$ , the sum in 3.2 reduces to a finite sum and therefore the product matrix is well-defined. It can easily be verified that if  $[a_{jk}]$  and  $[b_{jk}]$  are column-finite then  $[c_{jk}]$  is likewise column-finite.

In what follows the infinite summation sign,  $\sum_{i=1}^{\infty} \xi_i$ , where  $\xi_i \in B$ , will be used only when all but a finite number of the  $\xi_i$  are equal to zero, in which case it denotes the sum of the nonzero terms.

Now let us choose an arbitrary but definite semi-basis  $\{e_i\}$  of  $B$ . (We shall see in §5 that there do exist semi-bases of  $B$ .) We shall keep this choice fixed through the remainder of the section.

Let  $T$  be a bounded linear operator on  $B$ . Then we say the matrix  $[a_{jk}]$  represents  $T$  relative to  $\{e_i\}$  if

$$(3.3) \quad [a_{jk}] \text{ is column-finite and}$$

$$(3.4) \quad \text{for each } k, Te_k = \sum_{j=1}^{\infty} a_{jk} e_j \text{ (where by (3.3) this reduces to a finite sum).}$$

The next theorem follows directly from the above definition and the Definitions (3.1) and (3.2).

**THEOREM 3.5.** *Let  $T$  and  $U$  be bounded linear operators on  $B$  and suppose that  $[a_{jk}]$  represents  $T$  and  $[b_{jk}]$  represents  $U$ , both relative to  $\{e_i\}$ . Then (i)  $[a_{jk}] + [b_{jk}]$  represents  $T + U$  relative to  $\{e_i\}$ , and (ii)  $[a_{jk}] \cdot [b_{jk}]$  represents  $T \cdot U$  relative to  $\{e_i\}$ .*

If  $[a_{jk}]$  is a column-finite matrix, we denote the matrix  $[a_{jk}]^n$  (the product of  $[a_{jk}]$  taken  $n$  times) by  $[a_{jk}^{(n)}]$ . It follows from repeated use of (3.5) that if  $[a_{jk}]$  represents  $T$  relative to  $\{e_i\}$ , then  $[a_{jk}^{(n)}]$  represents  $T^n$  relative to  $\{e_i\}$ .

A matrix  $[a_{jk}]$  is called *almost superdiagonal* if  $a_{jk} = 0$  for  $j > k + 1$ . An almost superdiagonal matrix is thus column-finite.

**THEOREM 3.6.** *Let  $[a_{jk}]$  be an almost superdiagonal matrix. Then for any positive integers  $s, m, n$*

$$(3.7) \quad a_{n+m,n}^{(s)} = 0 \quad \text{if } s < m.$$

$$(3.8) \quad a_{n+m,n}^{(m)} = \prod_{i=0}^{m-1} a_{n+i+1,n+i}.$$

*Proof of 3.7. (by induction on  $s$ .)* If  $s = 1$ , then for all  $n$ ,  $a_{n+m,n}^{(1)} = a_{n+m,n} = 0$  if  $1 < m$  since  $[a_{jk}]$  is almost superdiagonal.

Now let  $s > 1$  and assume that for any  $n$ ,  $a_{n+m,n}^{(s-1)} = 0$  whenever  $s - 1 < m$ . Fixing  $n$ , the definition of matrix multiplication gives for any  $m$

$$(3.9) \quad a_{n+m,n}^{(s)} = \sum_{i=1}^{\infty} a_{n+m,i}^{(s-1)} a_{i,n}.$$

Suppose  $s < m$ . Then  $s - 1 < m - 1$  and by the induction hypothesis,  $a_{n+m,i}^{(s-1)} = 0$  for  $i \leq n + 1$ . But for  $i > n + 1$ ,  $a_{i,n} = 0$  since  $[a_{jk}]$  is almost superdiagonal; hence all the terms under the summation sign in (3.9) must be equal to 0, so  $a_{n+m,n}^{(s)} = 0$ , establishing (3.7).

*Proof of 3.8. (by induction on  $m$ .)* For the case  $m = 1$ ,

$$a_{n+1,n}^{(1)} = a_{n+1,n} = \prod_{i=0}^0 a_{n+i+1,n+i}.$$

Now let  $m > 1$  and assume that for all  $n$ ,

$$a_{n+m-1,n}^{(m-1)} = \prod_{i=0}^{m-2} a_{n+i+1,n+i}.$$

Fixing  $n$ , we have by the definition of matrix multiplication,

$$a_{n+m,n}^{(m)} = \sum_{i=1}^{\infty} a_{n+m,i}^{(m-1)} a_{i,n}.$$

For  $i \leq n$ ,  $a_{n+m,i}^{(m-1)} = 0$  by (3.7), and for  $i > n + 1$ ,  $a_{i,n} = 0$  since  $[a_{jk}]$  is almost superdiagonal. Therefore only the  $(n + 1)^{\text{st}}$  term in the summation is nonzero, and using this together with the induction hypothesis,

$$\begin{aligned} a_{n+m,n}^{(m)} &= a_{n+m,n+1}^{(m-1)} a_{n+1,n} = a_{n+1+m-1,n+1}^{(m-1)} a_{n+1,n} \\ &= \left( \prod_{i=0}^{m-2} a_{n+1+i+1,n+1+i} \right) a_{n+1,n} \\ &= \left( \prod_{i=1}^{m-1} a_{n+i+1,n+i} \right) a_{n+1,n} = \prod_{i=0}^{m-1} a_{n+i+1,n+i}. \end{aligned}$$

This establishes (3.8), proving (3.6).

We recall here that the spectrum of a bounded linear operator  $T$ ,  $\sigma(T)$ , is defined to be the set of complex numbers  $\lambda$  with the property that there does not exist a bounded linear operator on  $B$  which is the inverse of  $(\lambda I - T)$ , where  $I$  is the identity operator on  $B$ . It is a standard fact (cf. [7], p. 567) that  $\sigma(T)$  is a nonempty compact set and  $\sup |\sigma(T)| = \lim_{n \rightarrow \infty} \|T^n\|^{1/n}$ . A bounded linear operator  $T$  is called *quasi-nilpotent* if  $\sigma(T) = \{0\}$  and this will be the case if and only if  $\lim_{n \rightarrow \infty} \|T^n\|^{1/n} = 0$ . We shall see in § 5 that the only difficult case in the proof of the main theorem arises when  $T$  is quasi-nilpotent.

An infinite matrix  $[a_{jk}]$  may be regarded as a function  $f$  from  $Z \times Z$  into  $C$  and as such is represented by a relation of  $M$ . As we pass to  $*M$ ,  $f$  becomes extended to  $*f$  which is a function from  $*Z \times *Z$  into  $*C$ . We may then regard  $*f$  as an infinite matrix which we denote by  $*[a_{jk}]$ , where now  $j$  and  $k$  range over  $*Z$  and the entries  $a_{jk}$  are elements of  $*C$ . For any  $j, k \in *Z$  we designate the  $(j, k)^{\text{th}}$  coordinate of  $*[a_{jk}]$ ,  $*f(j, k)$ , again by  $a_{jk}$ . Since  $*f(j, k) = f(j, k)$  for  $j, k \in Z$ , this designation is unambiguous.

If  $\{x_i\}$  is a  $Q$ -sequence of elements of  $*B$  with the property that there is an  $n \in *Z$  such that  $x_i = 0$  for all  $i > n$ , then we use the infinite summation sign,  $\sum_{i=1}^{\infty} x_i$ , to designate  $\sum_{i=1}^n x_i$ . The latter is well defined in the manner described in § 2.

Now if  $T$  is a bounded linear operator on  $B$  which is represented by  $[a_{jk}]$  relative to  $\{e_i\}$ , then  $*T$  is represented by  $*[a_{jk}]$  relative to  $*\{e_i\}$ . That is, for any  $k \in *Z$ , (i) there is an  $n \in *Z$  such that  $a_{jk} = 0$  for all  $j > n$  and (ii)  $*Te_k = \sum_{j=1}^{\infty} a_{jk}e_j$ . For any finite integer  $n$ , we may consider the  $Q$ -operator  $(*T)^n$ .  $(*T)^n$  is represented by the matrix  $*[a_{jk}^{(n)}]$  whose coefficients  $a_{jk}^{(n)}$  satisfy the relations (3.7) and (3.8) for infinite as well as finite integers by transfer from  $M$ . Similarly if  $\nu$  is an infinite integer then we may consider the  $Q$ -operator  $(*T)^\nu$ . Here, however,  $(*T)^\nu$  cannot strictly be regarded as the iteration of the product operation but must be regarded as defined by transfer from  $M$  in a manner similar to the way the sum of  $\nu$  terms was defined in § 2. Then  $\|(*T)^\nu\|^{1/\nu}$ , the positive  $\nu^{\text{th}}$  root of  $\|(*T)^\nu\|$  may be similarly defined. It then follows from (1.1) that  $T$  is quasi-nilpotent if and only if  $\|(*T)^\nu\|^{1/\nu}$  is infinitesimal for all infinite  $\nu$ .  $(*T)^\nu$  may be represented by a matrix  $[a_{jk}^{(\nu)}]$ , where  $j$  and  $k$  range over  $*Z$  and the coefficients  $a_{jk}^{(\nu)}$  again satisfy (3.7) and (3.8) for all  $j, k \in *Z$  by transfer from  $M$ .

**THEOREM 3.10.** *Let  $T$  be a quasi-nilpotent bounded linear operator on  $B$  which is represented by an almost superdiagonal matrix  $[a_{jk}]$  relative to  $\{e_i\}$ . Then there exists an infinite integer  $\omega$  such*

that  $a_{\omega+1,\omega}$  is infinitesimal.

*Proof.* Let  $\nu$  be an arbitrary but definite infinite positive integer. Then since  $T$  is quasi-nilpotent,  $\|(*T)^\nu\|^{1/\nu}$  is infinitesimal. Next, let  $\mu$  be another infinite positive integer. The  $S$ -sequence  $\{e_i\}$  in  $B$  extends to the standard sequence  $*\{e_i\}$  in  $*B$  whose  $\mu^{\text{th}}$  element is  $e_\mu$ . Transferring (2.8) to  $*M$ ,  $\|e_i\| = 1$  for all  $i$  in  $*Z$ , in particular  $\|e_\mu\| = 1$ . Then

$$(3.11) \quad \|(*T)^\nu e_\mu\| \leq \|(*T)^\nu\| \|e_\mu\| = \|(*T)^\nu\|.$$

Now since  $[a_{jk}^{(\nu)}]$  represents  $(*T)^\nu$  relative to  $*\{e_i\}$ , it follows from 3.7 that

$$(*T)^\nu e_\mu = \sum_{j=1}^{\infty} a_{j\mu}^{(\nu)} e_j = \sum_{j=1}^{\mu+\nu} a_{j\mu}^{(\nu)} e_j = \sum_{j=1}^{\mu+\nu-1} a_{j\mu}^{(\nu)} e_j + a_{\mu+\nu,\mu}^{(\nu)} e_{\mu+\nu}.$$

Then denoting the metric projection of  $*B$  onto  $\text{sp}\{e_1, \dots, e_{\mu+\nu-1}\}$  by  $P_{\mu+\nu-1}$  as usual, it follows from (2.11) that

$$(*T)^\nu e_\mu - P_{\mu+\nu-1}(*T)^\nu e_\mu = a_{\mu+\nu,\mu}^{(\nu)} e_{\mu+\nu}.$$

Hence using this together with (2.5),

$$(3.12) \quad |a_{\mu+\nu,\mu}^{(\nu)}| = \|a_{\mu+\nu,\mu}^{(\nu)} e_{\mu+\nu}\| \leq \|(*T)^\nu e_\mu\| \leq \|(*T)^\nu\|.$$

We now use (3.8) to obtain

$$(3.13) \quad |a_{\mu+\nu,\mu}^{(\nu)}| = \left| \prod_{i=0}^{\nu-1} a_{\mu+i+1,\mu+i} \right| = \prod_{i=0}^{\nu-1} |a_{\mu+i+1,\mu+i}|.$$

Now suppose that for no  $i$ ,  $0 \leq i < \nu$ , is it the case that  $a_{\mu+i+1,\mu+i}$  is infinitesimal. Then

$$\text{minimum}_{0 \leq i < \nu} |a_{\mu+i+1,\mu+i}| > \delta > 0$$

for some standard  $\delta$ . Then by (3.13),

$$\delta^\nu < |a_{\mu+\nu,\mu}^{(\nu)}|.$$

But then using (3.12),

$$\delta = (\delta^\nu)^{1/\nu} < |a_{\mu+\nu,\mu}^{(\nu)}|^{1/\nu} \leq \|(*T)^\nu\|^{1/\nu}.$$

But  $\|(*T)^\nu\|^{1/\nu}$  is infinitesimal, so it cannot be greater than the standard positive number  $\delta$ . This contradiction shows that for some  $k$ ,  $0 \leq k < \nu$ ,  $a_{\mu+k+1,\mu+k}$  is infinitesimal. Letting  $\omega = \mu + k$ , we obtain the theorem.

4. Subspaces in nonstandard Banach space. Let  $\{e_i\}$  be a semi-

basis for  $B$  and let  $\omega$  be an infinite integer. Defining  $B_\omega$  and  $P_\omega$  as before, let  $E$  be an (internal) linear subspace of  $B_\omega$ . Define  ${}^\circ E$ , a subset of  $B$ , by:  $x \in {}^\circ E$  if and only if  $x \in B$  and  $\|x - x'\|$  is infinitesimal for some  $x' \in E$ .  $E$  is "finite-dimensional" in the sense of non-standard analysis because it is a subspace of the "finite-dimensional" space  $B_\omega$ , so there is a metric projection  $P_E$  of  $*B$  onto  $E$ . Then since  $P_E$  is the nearest point operator onto  $E$ ,  $\|x - x'\| \geq \|x - P_E x\|$  for all  $x$  in  $B$  and  $x' \in E$ , and it follows that for  $x$  in  $B$ ,  $x \in {}^\circ E$  if and only if  $\|x - P_E x\|$  is infinitesimal.

The next three theorems can be proved in the same way as in the Hilbert space case ([3]).

**THEOREM 4.1.** *Given  $E$  as above,  ${}^\circ E$  is a closed linear subspace of  $B$ .*

*Proof.* Let  $x, y \in {}^\circ E$  and  $\lambda \in C$ . There exist elements  $x', y' \in E$  such that  $\|x - x'\|$  and  $\|y - y'\|$  are infinitesimal. Then  $x' + y' \in E$  and

$$\|(x + y) - (x' + y')\| \leq \|x - x'\| + \|y - y'\|$$

so that the left hand side of this inequality is also infinitesimal. Hence  $x + y$  belongs to  ${}^\circ E$ . Also,  $\lambda x' \in E$  and

$$\|\lambda x - \lambda x'\| = |\lambda| \|x - x'\|$$

is infinitesimal, so  $\lambda x \in {}^\circ E$ . This shows that  ${}^\circ E$  is linear in the algebraic sense.

Now let  $x_n \rightarrow x$  where the  $x_n$  are defined for standard  $n \in Z$  and belong to  ${}^\circ E$ . In order to prove that  ${}^\circ E$  is closed we have to show that  $x$  belongs to  ${}^\circ E$ . By 2.6, for each  $n \in Z$ ,

$$\|P_E x - x\| \leq \|x - x_n\| + \|x_n - P_E x_n\|.$$

For a given standard  $\varepsilon > 0$ , choose  $n \in Z$  such that  $\|x - x_n\| < \varepsilon/2$ . Since for each  $n \in Z$ ,  $x_n \in {}^\circ E$  and consequently  $\|x_n - P_E x_n\|$  is infinitesimal, it follows that  $\|P_E x - x\| \leq \varepsilon$ . Since  $\varepsilon$  was an arbitrary standard positive number,  $\|P_E x - x\|$  is infinitesimal, hence  $x \in {}^\circ E$  and the proof is complete.

Let  $T$  be a bounded linear operator on  $B$  and let  $\omega$  be an infinite integer. We define the operator  $T_\omega$  on  $B_\omega$  to be the restriction of  $P_\omega * T P_\omega$  to  $B_\omega$ . Then  $\|T_\omega\| \leq \|P_\omega\|^2 \|*T\| \leq 4 \|*T\| = 4 \|T\|$  so that  $T_\omega$  has finite norm.

**THEOREM 4.2.** *Let  $E$  be an internal linear subspace of  $B_\omega$  which is invariant for  $T_\omega$ , i.e.,  $T_\omega E \subseteq E$ . Then  ${}^\circ E$  is invariant for  $T$ ,  $T {}^\circ E \subseteq {}^\circ E$ .*

*Proof.* Choose any  $x \in {}^\circ E$ . Then  $P_E x \sim x$  and since  $*T$  has finite norm,  $*TP_E x \sim *Tx$ . Since  $*Tx$  is standard,  $*TP_E x$  is near-standard and we may apply (2.13) to conclude that  $P_\omega *TP_E x \sim *TP_E x$ . Putting the last two relations together,  $T_\omega P_E x = P_\omega *TP_E x \sim *Tx = Tx$  where  $T_\omega P_E x$  is in  $E$  since  $E$  is invariant for  $T_\omega$ .  $Tx$  is thus infinitely close to an element of  $E$ , so  $Tx \in {}^\circ E$ . This shows that  ${}^\circ E$  is invariant for  $T$  which proves the theorem.

The number of dimensions of  $B_\omega$  as defined within the language  $L$  is  $\omega$ ,  $d(B_\omega) = \omega$ . In this sense  $B_\omega$  is “finite-dimensional”. Similarly, with every (internal) linear subspace  $E$  of  $B_\omega$  there is associated an integer  $d(E) \in {}^*Z$  (or  $d(E) = 0$ ) which may be finite or infinite, and which has the properties of a dimension to the extent to which these can be expressed as sentences of  $L$ .

**THEOREM 4.3.** *Let  $E, E_1$  be linear subspaces of  $B_\omega$  such that  $E \subset E_1$  and  $d(E_1) = d(E) + 1$ . Then  ${}^\circ E \subseteq {}^\circ E_1$  and any two points of  ${}^\circ E_1$  are linearly dependent modulo  ${}^\circ E$ .*

*Proof.* Since  $E \subset E_1$ , it is trivial that  ${}^\circ E \subseteq {}^\circ E_1$ . Now let  $x, y \in {}^\circ E_1$ . There exist  $x', y'$  in  $E_1$  such that  $x \sim x'$  and  $y \sim y'$ . Since the dimension of  $E_1$  exceeds that of  $E$  only by 1 there is a representation

$$(3.4) \quad x' = \lambda y' + z \text{ or vice versa, where } \lambda \in {}^*C \text{ and } z \in E.$$

If  $\lambda$  is finite, then it possesses a standard part  ${}^\circ \lambda$  and  $\lambda y'$  is infinitely close to the standard point  ${}^\circ \lambda y$ . Therefore  $z = x' - \lambda y'$ , being the difference of two near-standard points must itself possess a standard part  ${}^\circ z$ ,  ${}^\circ z \sim z$ . Now  $x' \sim x$ ,  $\lambda y' \sim {}^\circ \lambda y$ , and  $z \sim {}^\circ z$ , so substituting this in (3.4) we obtain  $x \sim {}^\circ \lambda y + {}^\circ z$  where  ${}^\circ z \in {}^\circ E$  since  $z \in E$ . But two standard numbers cannot be infinitely close unless they are equal, so  $x = {}^\circ \lambda y + {}^\circ z$  and therefore  $x$  and  $y$  are linearly dependent modulo  ${}^\circ E$ .

If  $\lambda$  is infinite then we may rewrite (3.4) as,  $y' = (1/\lambda)x' + (-1/\lambda)z$ . Now  $1/\lambda$  is infinitesimal, *a-fortiori* finite, and  $(-1/\lambda)z \in E$  so we have precisely the same case as was considered in the previous paragraph with  $y$  and  $x$  interchanged. Therefore we again conclude that  $y$  and  $x$  are linearly dependent modulo  ${}^\circ E$ , which completes the proof of the theorem.

**5. Proof of the main theorem.** We shall first show that the proof of the main theorem reduces to the case where  $T$  is quasi-nilpotent, that is, we shall show that the following theorem implies the main theorem.

**THEOREM 5.1.** *Let  $T$  be a quasi-nilpotent bounded linear operator on  $B$ . Suppose there exists a polynomial  $p(\lambda) \neq 0$  with complex*

coefficients such that  $p(T)$  is compact. Then  $T$  leaves invariant at least one closed linear subspace of  $B$  other than  $\{0\}$  or  $B$ .

*Proof that (5.1) implies the main theorem.* Suppose that (5.1) is true. Let  $T$  be a bounded linear operator satisfying the hypothesis of the main theorem, that is, suppose  $p(\lambda) \neq 0$  is a complex polynomial such that  $p(T)$  is compact. We wish to show that  $T$  has a proper (i.e.  $\neq\{0\}$ ,  $\neq B$ ) invariant (closed, linear) subspace.

We may restrict our attention to the case where  $\sigma(T)$  is connected, for suppose  $\sigma(T)$  is disconnected. Then  $\sigma = \sigma_1 \cup \sigma_2$  where  $\sigma_1$  and  $\sigma_2$  are disjoint nonempty sets, both open and closed in the relative topology of  $\sigma(T)$ . In other words  $\sigma_1$  and  $\sigma_2$  are nonempty disjoint spectral sets in the sense of [6] and we may find proper subspaces of  $B$  which actually reduce  $T$ ; in particular then,  $T$  has a proper invariant subspace. Therefore we shall suppose that  $\sigma(T)$  is connected.

Since  $\sigma(T)$  is connected,  $p(\sigma(T))$  must again be connected since  $p(\lambda)$  is continuous on  $\sigma(T)$ . By the spectral mapping theorem (cf. [15], p. 263),  $p(\sigma(T)) = \sigma(p(T))$ , thus  $\sigma(p(T))$  is connected. Now  $p(T)$  is compact, so  $\sigma(p(T))$  is at most denumerable and has no point of accumulation except possibly the point 0. Also since  $B$  is infinite-dimensional,  $0 \in \sigma(p(T))$  (cf. [5], p. 319). We wish to show that 0 is the only point in  $\sigma(p(T))$ . Suppose that  $\sigma(p(T))$  contains a point  $\lambda_0 \neq 0$ . Then since  $\lambda_0$  is not a point of accumulation of  $\sigma(p(T))$ , there is an  $\epsilon > 0$  such that the sphere  $\{\lambda \mid |\lambda - \lambda_0| \leq \epsilon\}$  contains no other points of  $\sigma(p(T))$ . Then if we let

$$(5.2) \quad \sigma_1 = \sigma(p(T)) \cap \{\lambda \mid |\lambda - \lambda_0| < \epsilon\}$$

$$(5.3) \quad \sigma_2 = \sigma(p(T)) \cap \{\lambda \mid |\lambda - \lambda_0| < \epsilon\},$$

we obtain  $\sigma(p(T)) = \sigma_1 \cup \sigma_2$  where  $\sigma_1$  and  $\sigma_2$  are nonempty disjoint sets, both open and closed in the relative topology of  $\sigma(p(T))$ . This means that  $\sigma(p(T))$  is disconnected and this contradiction shows that 0 must be the only point of  $\sigma(p(T))$ , i.e.,

$$(5.4) \quad \sigma(p(T)) = p(\sigma(T)) = \{0\}.$$

Since  $p(\lambda)$  can have only finitely many zeros, it follows from (5.4) that  $\sigma(T)$  must contain only finitely many points. Then since  $\sigma(T)$  is connected, this means that  $\sigma(T)$  contains only one point, say  $\lambda_0$ .

Now let  $V = T - \lambda_0 I$ . Since  $\sigma(T) = \{\lambda_0\}$ ,  $\sigma(V) = \{0\}$ , so  $V$  is a quasi-nilpotent bounded linear operator. Letting  $q(\lambda) = p(\lambda + \lambda_0)$ , we have that  $q(V) = p(V + \lambda_0 I) = p(T)$  which is compact. Furthermore since  $p(\lambda) \neq 0$ , we must have  $q(\lambda) \neq 0$ . Then since we are assuming (5.1) to be true, there must be a proper closed linear subspace  $E$  of  $B$  such that  $VE \subseteq E$ . But then  $TE = (V + \lambda_0 I)E \subseteq E$ , so that  $E$  is

a proper invariant subspace also for  $T$ . This completes the proof of the fact that (5.1) implies the main theorem.

We shall now prove (5.1). Our method like that of [1] and [3] is based on the elementary fact that in a *finite-dimensional* vector space of dimension  $m$ , say, any linear operator possesses a chain of invariant subspaces

$$(5.5) \quad \{0\} = E_0 \subset E_1 \subset E_2 \subset \cdots \subset E_m \text{ where } d(E_j) = j, \quad 0 \leq j \leq m.$$

To begin with we work in the standard space  $B$ .

Let  $T$  be a bounded linear operator on  $B$  and let  $x$  be an element of  $B$  of norm 1. We define a sequence  $\{e_i\}$  of elements of  $B$  as follows.

$$(5.6) \quad e_1 = x$$

$$(5.7) \quad e_n = \frac{T^{n-1}x - P_{n-1}T^{n-1}x}{\|T^{n-1}x - P_{n-1}T^{n-1}x\|} \quad \text{for } n > 1$$

where  $P_{n-1}$  is the metric projection of  $B$  onto  $\text{sp}\{x, Tx, \dots, T^{n-2}x\}$ . As we remarked in the beginning of this chapter, the set  $\{x, Tx, T^2x, \dots\}$  is linearly independent, so

$$T^{n-1}x \notin \text{sp}\{x, Tx, T^2x, \dots, T^{n-2}x\};$$

therefore since  $P_{n-1}T^{n-1}x \in \text{sp}\{x, Tx, T^2x, \dots, T^{n-2}x\}$ ,  $T^{n-1}x \neq P_{n-1}T^{n-1}x$  and the above is a valid definition.

We shall now show by induction that for any  $n \in \mathbb{Z}$ ,

$$(5.8) \quad \text{sp}\{e_1, e_2, \dots, e_n\} = \text{sp}\{x, Tx, \dots, T^{n-1}x\}.$$

For the case  $n = 1$ ,  $e_1 = x = T^0x$  by (5.6).

Now let  $n > 1$  and assume that

$$\text{sp}\{e_1, e_2, \dots, e_{n-1}\} = \text{sp}\{x, Tx, \dots, T^{n-2}x\}.$$

$P_{n-1}T^{n-1}x \in \text{sp}\{x, Tx, \dots, T^{n-2}x\}$ , therefore by (5.7),  $e_n$  is a linear combination of elements of  $\{x, Tx, \dots, T^{n-1}x\}$ ; hence using this together with our induction hypothesis,  $\text{sp}\{e_1, e_2, \dots, e_n\} \subseteq \text{sp}\{x, Tx, \dots, T^{n-1}x\}$ . To obtain the reverse inclusion, observe that by (5.7) again,  $T^{n-1}x$  is a linear combination of  $e_n$  and  $P_{n-1}T^{n-1}x$ . But

$$P_{n-1}T^{n-1}x \in \text{sp}\{x, Tx, \dots, T^{n-2}x\} = \text{sp}\{e_1, \dots, e_{n-1}\}$$

by the induction hypothesis, so

$$\text{sp}\{x, Tx, \dots, T^{n-1}x\} \subseteq \text{sp}\{e_1, e_2, \dots, e_n\}.$$

This establishes (5.8).

Next we shall show that  $\{e_i\}$  is a semi-basis for  $B$  by verifying (2.8)–(2.10). To verify (2.8) we need only remark that

- (i)  $\|e_i\| = 1$  for each  $i \in Z$ , by construction.
- (ii) Since (5.8) is valid for each  $n \in Z$ ,  $\text{sp}\{e_1, e_2, e_3, \dots\} = \text{sp}\{x, Tx, T^2x, \dots\}$ . But by (2.1),  $\text{sp}\{x, Tx, T^2x, \dots\} = B$ , therefore  $\text{sp}\{e_i\} = B$ . This verifies (2.9).
- (iii) Observe first that since

$$\text{sp}\{e_1, e_2, \dots, e_{n-1}\} = \text{sp}\{x, Tx, \dots, T^{n-2}x\},$$

$P_{n-1}$  is the metric projection onto  $B_{n-1} = \text{sp}\{e_1, e_2, \dots, e_{n-1}\}$  so that our notation is consistent with that of (2.10). Now applying  $P_{n-1}$  to both sides of (5.7),

$$P_{n-1}e_n = \frac{P_{n-1}T^{n-1}x - P_{n-1}T^{n-1}x}{\|T^{n-1}x - P_{n-1}T^{n-1}x\|} = 0,$$

where we have used the properties (2.4) and (2.3) of  $P_{n-1}$ . This verifies (2.10) and together with (i) and (ii) shows that  $\{e_i\}$  is a semi-basis for  $B$ .

For any integer  $k \in Z$ ,  $e_k \in \text{sp}\{e_1, e_2, \dots, e_k\} = \text{sp}\{x, Tx, \dots, T^{k-1}x\}$ , so that

$$Te_k \in \text{sp}\{Tx, T^2x, \dots, T^kx\} \subseteq \text{sp}\{e_1, e_2, \dots, e_{k+1}\}.$$

Thus  $Te_k = \sum_{j=1}^{k+1} \beta_{jk}e_j$  for some  $\beta_{jk} \in {}^*C$ ,  $0 < j \leq k + 1$ . After choosing such  $\beta_{jk}$  for each  $k \in Z$ , we define the matrix  $[a_{jk}]$  by (i)  $a_{jk} = \beta_{jk}$  if  $j \leq k + 1$ , and (ii)  $a_{jk} = 0$  if  $j > k + 1$ . Then  $[a_{jk}]$  is an almost super-diagonal matrix which represents  $T$  relative to the semi-basis  $\{e_i\}$ .

Now suppose that  $T$  is quasi-nilpotent and that

$$p(\lambda) = c_0 + c_1\lambda + c_2\lambda^2 + \dots + c_n\lambda^n, \quad c_n \neq 0,$$

is a complex polynomial such that  $p(T)$  is compact. Using (3.10), let  $\omega$  be an infinite integer such that  $\alpha_{\omega+1, \omega}$  is infinitesimal. For brevity we denote by  $P$  the metric projection  $P_\omega$  of  ${}^*B$  onto  $B_\omega$  and by  $T'$  the  $Q$ -operator  $P^*TP$ . We next show by induction on finite integers  $n$  that

$$(5.9) \quad ({}^*T)^n \xi \sim (T')^n \xi$$

for norm-finite  $\xi \in B_\omega$ ,  $n \in Z$ . It will then follow by applying (5.9) to the monomials of

$$p({}^*T) = c_0 {}^*I + c_1 {}^*T + c_2 ({}^*T)^2 + \dots + c_n ({}^*T)^n$$

and

$$p(T') = c_0 {}^*I + c_1 T' + c_2 (T')^2 + \dots + c_n (T')^n$$

that

$$(5.10) \quad p(*T)\xi \sim p(T')\xi$$

for norm-finite  $\xi \in B_\omega$ .

To establish (5.9) for the case  $n = 1$ , let  $\xi$  be a norm-finite element of  $B_\omega$ ; then  $\xi = P\xi = \sum_{k=1}^\omega \alpha_k e_k$  for some  $\alpha_k \in *C$ ,  $k = 1, 2, \dots, \omega$ . By (2.11)  $\xi - P_{\omega-1}\xi = \alpha_\omega e_\omega$ , and using (2.5)

$$|\alpha_\omega| = \|\alpha_\omega e_\omega\| = \|\xi - P_{\omega-1}\xi\| \leq \|\xi\|$$

so that  $\alpha_\omega$  is finite. Now

$$*T\xi = \sum_{k=1}^\omega \alpha_k *Te_k = \sum_{k=1}^\omega \alpha_k \sum_{j=1}^\infty a_{jk} e_j = \sum_{k=1}^\omega \alpha_k \sum_{j=1}^{\omega+1} a_{jk} e_j$$

since  $a_{jk} = 0$  for  $j, k$  satisfying  $k < \omega + 1 < j$ , where we are using the fact that  $[a_{jk}]$  is almost superdiagonal. Switching summation signs we obtain

$$*T\xi = \sum_{j=1}^{\omega+1} \left( \sum_{k=1}^\omega \alpha_k a_{jk} \right) e_j.$$

But  $a_{\omega+1,k} = 0$  for  $k < \omega$  so that  $\sum_{k=1}^\omega \alpha_k a_{\omega+1,k} = \alpha_\omega a_{\omega+1,\omega}$ , thus

$$*T\xi = \sum_{j=1}^\omega \left( \sum_{k=1}^\omega \alpha_k a_{jk} \right) e_j + \alpha_\omega a_{\omega+1,\omega} e_{\omega+1}.$$

Hence by (2.11),

$$*T\xi - P*T\xi = \alpha_\omega a_{\omega+1,\omega} e_{\omega+1}.$$

But  $|\alpha_\omega|$  and  $\|e_{\omega+1}\|$  are finite and  $a_{\omega+1,\omega}$  is infinitesimal; thus  $\|*T\xi - P*T\xi\|$  is infinitesimal, or since  $\xi$  was an arbitrary norm-finite element of  $B_\omega$  and  $T' = P*TP$ ,

$$(5.11) \quad *T\xi \sim T'\xi \text{ for norm-finite } \xi \in B_\omega, \text{ proving (5.9) for the case } n = 1.$$

Now let  $n \geq 2$  and assume that  $(*T)^{n-1}\xi \sim (T')^{n-1}\xi$  for any norm-finite  $\xi \in B_\omega$ . Since  $*T$  has finite norm,

$$(5.12) \quad (*T)^n \xi = *T(*T)^{n-1}\xi \sim *T(T')^{n-1}\xi.$$

$(T')^{n-1}\xi$  is a norm-finite element of  $B_\omega$ , so we may apply (5.11) with  $(T')^{n-1}\xi$  in place of  $\xi$  to obtain

$$*T(T')^{n-1}\xi \sim T'(T')^{n-1}\xi = (T')^n \xi.$$

This together with (5.12) shows that  $(*T)^n \xi \sim (T')^n \xi$ , which establishes (5.9) and also then (5.10).

We denote by  $T_\omega$  the restriction of  $T'$  to  $B_\omega$ .  $T_\omega$  is a linear operator on  $B_\omega$ , for let  $y \in B_\omega$ . Then

$$y \in \text{sp} \{e_1, e_2, \dots, e_\omega\} = \text{sp} \{x, {}^*Tx, ({}^*T)^2x, \dots, ({}^*T)^{\omega-1}x\}$$

so  $y = \sum_{i=0}^{\omega-1} a_i ({}^*T)^i x$  for some  $a_i \in {}^*C$ ,  $0 \leq i \leq \omega - 1$ . Then using (2.4) and (2.3), and the fact that  $Py = y$ ,

$$T_\omega y = P {}^*Ty = P \sum_{i=0}^{\omega-1} a_i ({}^*T)^{i+1} x = \sum_{i=0}^{\omega-2} a_i ({}^*T)^{i+1} x + \alpha_{\omega-1} P ({}^*T)^\omega x.$$

This shows that  $T_\omega$  is linear. (An equivalent remark is made also in [1].) Since  $B_\omega$  is “finite-dimensional” in the sense of nonstandard analysis, there exists a chain of subspaces as in (5.5) with  $m = \omega$ , such that  $T_\omega E_j \subseteq E_j$ ,  $j = 0, 1, \dots, \omega$ . The  $E_j$  are internal linear subspaces of  ${}^*B$  and are “finite-dimensional” in the sense of nonstandard analysis since they are subspaces of  $B_\omega$ . Let  $Q_j$  be the metric projection from  ${}^*B$  onto  $E_j$ ,  $j = 0, 1, \dots, \omega$ , so that  $Q_\omega = P$ .

Choose any  $\xi \in H$ ,  $\|\xi\| = 1$ .  $p(T)\xi = p({}^*T)\xi$  must be different from zero, otherwise  $\xi, T\xi, \dots, T^n\xi$  would be linearly dependent contrary to what we have assumed (in (2.1)). By (2.12),  $\xi \sim P\xi$  and since  $p({}^*T)$  and  $p(T')$  both have finite norm,  $p({}^*T)\xi \sim p({}^*T)P\xi$  and  $p(T')\xi \sim p(T')P\xi$ . But since we have chosen  $\omega$  to satisfy the conclusion of Theorem (3.10), by (5.10)  $p({}^*T)P\xi \sim p(T')P\xi$ ; therefore putting the last three relations together,  $p({}^*T)\xi \sim p(T')\xi$ . Now  $p({}^*T)\xi$  is a standard point unequal to 0, so  $p(T')\xi$  is not infinitesimal and consequently  $\|p(T')\xi\| > r$  for some standard positive  $r$ .

Consider the expressions

$$(5.13) \quad r_j = \|p(T')\xi - p(T')Q_j\xi\|, \quad j = 0, 1, 2, \dots, \omega$$

and note that  $r_j \leq \|p(T')\| \|\xi - Q_j\xi\|$ . We have  $r_0 = \|p(T')\xi\|$  so  $r_0 > r$ . Also  $\|\xi - Q_\omega\xi\| = \|\xi - P\xi\|$  is infinitesimal, hence  $r_\omega < r/2$ . It follows that there exists a smallest positive integer  $\mu$  which may be finite or infinite, such that  $r_\mu < r/2$  but  $r_{\mu-1} \geq r/2$ .

With every  $E_j$ , we associate the closed linear subspace  ${}^\circ E_j$  of  $B$  which was defined in § 4. Now  ${}^\circ E_{\mu-1}$  cannot coincide with  $B$ , in particular it cannot include  $\xi$ . For if it did, then  $\|\xi - Q_{\mu-1}\xi\|$  would be infinitesimal, so  $r_{\mu-1}$ , which is bounded by  $\|p(T')\| \|\xi - Q_{\mu-1}\xi\|$  would be infinitesimal contrary to the choice of  $\mu$ .

On the other hand  ${}^\circ E_\mu$  cannot reduce to  $\{0\}$ . Consider the point  $\eta = p(T')Q_\mu\xi$ .  $\eta \in E_\mu$  since  $Q_\mu\xi \in E_\mu$  and  $E_\mu$  is invariant under  $p(T_\omega)$  and, equivalently, under  $p(T')$ . By (5.10), since  $Q_\mu\xi \in B_\omega$ ,

$$\eta = p(T')Q_\mu\xi \sim p({}^*T)Q_\mu\xi = (p(T))Q_\mu\xi,$$

where the right-hand side is near-standard, by 1.5 since  $Q_\mu\xi$  is norm-finite and  $p(T)$  is compact. It follows that  $\eta$  possesses a standard part,  ${}^\circ\eta$ , and that  ${}^\circ\eta$  belongs to  $E_\mu$ . If  ${}^\circ\eta = 0$ , then  $\eta$  would be infinitesimal and it would then follow from (5.13) that

$$r_\mu \geq \|p(T')\xi\| - \|p(T')Q_\mu\xi\| > r - \zeta$$

where  $\zeta = \|\eta\|$  is infinitesimal. This would imply that  $r_\mu > r/2$ , contrary to the choice of  $\mu$ . We conclude that  ${}^\circ E_\mu$  contains a point different from 0, namely  ${}^\circ\eta$ .

Both  ${}^\circ E_{\mu-1}$  and  ${}^\circ E_\mu$  are invariant for  $T$ , by 4.2. If neither were a proper invariant subspace of  $B$  for  $T$  we would have  ${}^\circ E_{\mu-1} = \{0\}$ ,  ${}^\circ E_\mu = B$ . But this contradicts (4.3), proving (5.1), and this in turn establishes the main theorem.

#### REFERENCES

1. N. Aronszajn and K. T. Smith, *Invariant subspaces of completely continuous operators*, Ann. of Math. **60** (1954), 345-350.
2. A. R. Bernstein, *Invariant subspaces for linear operators*, Doctoral Dissertation, University of California, Los Angeles, 1965.
3. A. R. Bernstein and A. Robinson, *Solution of an invariant subspace problem of P.R. Halmos and K.T. Smith*, Pacific J. Math. **16** (1966), 421-431.
4. J. A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc. **40** (1936), 396-414.
5. J. Dieudonne, *Foundations of Modern Analysis*, Academic Press, New York and London, 1960.
6. N. Dunford, *Spectral theory I, convergence to projections*, Trans. Amer. Math. Soc. **54** (1943), 185-217.
7. N. Dunford and J. T. Schwarz, *Linear Operators, Part I*, Interscience, New York, 1958.
8. P. R. Halmos, *Finite-Dimensional Vector Spaces*, Van Nostrand, Princeton, 1958.
9. ———, *A glimpse into Hilbert space*, Lectures on Modern Mathematics, Vol. 1, ed. T. L. Saaty, New York and London, 1963, 1-22.
10. L. Henkin, *Completeness in the theory of types*, Journal of Symbolic Logic **15** (1950), 81-91.
11. A. Robinson, *Non-standard analysis*, Proceedings of the Royal Academy of Sciences, Amsterdam (A) **64** (1961), 432-440.
12. ———, *Complex-function theory over non-Archimedean fields* (Technical scientific note No. 307, U.S. Air Force Contract No. 61(052)-187), Jerusalem, 1962.
13. A. Robinson, *Introduction to Model Theory and to the Metamathematics of Algebra*, North-Holland, Amsterdam, 1963.
14. ———, *On generalized limits and linear functionals*, Pacific J. Math. **14** (1964), 269-283.
15. A. E. Taylor, *Introduction to Functional Analysis*, Wiley, New York, 1958.

Received February 25, 1966. This paper is taken in the main from the author's doctoral dissertation written at the University of California, Los Angeles under the direction of Professor Abraham Robinson. Financial support was received from the National Science Foundation, contract nos. GP-1812 and GP-4257.

UNIVERSITY OF WISCONSIN