The Invariant Subspace Problem

by

Jonathan A. Noel

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF SCIENCE (HONS.)

in the

Department of Mathematics & Statistics

THOMPSON RIVERS UNIVERSITY

We accept this thesis as conforming to the required standards:

Robb Fry
Dept. of Mathematics & Statistics
Thesis Supervisor

Examiner 2

Examiner 3

Dated April 23, 2011, Kamloops, British Columbia, Canada

THOMPSON RIVERS UNIVERSITY

DEPARTMENT OF MATHEMATICS & STATISTICS

Signature of Author

Abstract

The notion of an invariant subspace is fundamental to the subject of operator theory. Given an operator T on a Banach space X, a closed subspace M of X is said to be a non-trivial invariant subspace for T if $T(M) \subseteq M$ and $M \neq \{0\}, X$. A famous unsolved problem, called the "invariant subspace problem," asks whether every bounded linear operator on a Hilbert space (more generally, a Banach space) admits a non-trivial invariant subspace.

In this thesis, we discuss the greatest achievements in solving this problem for special classes of linear operators. We include several positive results for linear operators related to compact operators and normal operators, and negative results for certain linear operators on Banach spaces. Our goal is to build up the theory from the basics, and to prove the main results in a way that is accessible to a student who is relatively new to the world of functional analysis.

Acknowledgements

First and foremost, I would like to thank my supervisor Robb Fry for agreeing to work with a self-described graph theorist. I appreciate the time and effort taken to read and respond to several drafts of this thesis, guide me through difficult material, and help me fill many gaps in my knowledge of basic functional analysis. Despite facing some difficult circumstances, Robb has been an excellent mentor and teacher. I was first inspired to study mathematics by learning analysis from Robb; it is only fitting that I am finishing my degree in the same way.

I consider myself very fortunate to have studied at Thompson Rivers University, and will always be grateful to the professors in the Department of Mathematics and Statistics who have helped me along the way. Special thanks go out to Rick Brewster for introducing me to the incredible world of mathematical research.

I also thank my family and friends, for without them I would not be the person, or the mathematician, that I am today. The completion of this thesis signals a new era in my life, where I will have to leave the ones that I love in order to pursue my goals. It has become clearer than ever that the people closest to you are the most easily taken for granted. I am forever grateful to my parents and siblings for their constant love and support.

Contents

Chapter 1

Introduction

There are many fundamental problems in mathematics which remain unsolved. Often they can be stated in relatively simple terms, requiring little background knowledge, and yet somehow their solutions continue to elude humankind. The archetypal example of this is the famous 'twin prime conjecture' in number theory.[1]

One of these tantalizing open problems is the so-called 'invariant subspace problem' in functional analysis. Given a linear operator T on a vector space X, a subset S of X is said to be T-*invariant* if $T(S) \subseteq S$. A vector subspace M of X which is T-invariant is called an *invariant subspace* for T. All linear operators have invariant subspaces; For example $\{0\}$ and X are obviously invariant under T and so these are referred to as *trivial* invariant subspaces. The problem, in a general form, is stated as follows:

The Invariant Subspace Problem. If T is a bounded linear operator on a Banach space X, does it follow that T has a non-trivial closed invariant subspace?

Note, for a non-zero vector x the linear span of $\{x, Tx, T^2x, \dots\}$ is T-invariant, usually not equal to X, but may not be closed. Thus, the major difficulty of the invariant

[1]The twin prime conjecture states that there are infinitely many prime numbers p such that $p+2$ is also prime.

subspace problem comes from requiring that the T-invariant subspace be *simultaneously* non-trivial and closed.

In Chapter 2 we show that the problem is solved easily in the case that X is either finite-dimensional or non-separable. As it turns out, it is also solved in the negative for certain Banach spaces. The reduction to infinite-dimensional separable Hilbert spaces, however, remains one of the most famous and elusive open problems in functional analysis.

The Invariant Subspace Problem (as it stands today). If T is a bounded linear operator on an infinite-dimensional separable Hilbert space H, does it follow that T has a non-trivial closed invariant subspace?

For the remainder of the thesis, let us simply say *invariant subspace* when referring to a *closed invariant subspace*. In Chapter 3 we present some of the most important achievements in proving that certain operators have non-trivial invariant subspaces. This includes theorems of von Neumann and Lomonosov on compact operators, and Brown's results for operators related to normal operators. In the final chapter, we exhibit an example from C. J. Read of a bounded operator on the classical Banach space ℓ_1 having only the trivial invariant subspaces. In addition, we provide a brief discussion of the history and controversy surrounding the first known counterexamples on Banach spaces. For a list of standard definitions and notation, please refer to the appendices.

Chapter 2

Preliminaries

In this chapter, we give detailed solutions to the invariant subspace problem for Banach spaces which are either finite-dimensional (too small) or non-separable (too large). Although these reductions are quite straightforward, the solutions raise some important themes which shall be returned to throughout the thesis.

2.1 Eigenvectors and Finite-Dimensional Spaces

As with most problems in functional analysis, the invariant subspace problem only remains unsolved for infinite-dimensional spaces. In this section, we provide a solution to the finite-dimensional case by making clever use of the Fundamental Theorem of Algebra. We use freely the fact that finite-dimensional subspaces of normed spaces are closed. The proof of this is not difficult, and follows easily from the fact that finite-dimensional normed spaces are complete.

Note. In this section, X denotes a real or complex normed space of dimension $n \geq 0$ and T is an arbitrary linear operator on X.

The solution is built up from an elementary fact regarding n-dimensional spaces.

Remark 2.1.1. Given $x \in X$, the vectors of the set

$$S_n(x,T) = \{x, Tx, T^2x, \ldots, T^nx\}$$

are linearly dependent.

This is simply because the above set has cardinality $n+1$, and therefore cannot be linearly independent. Thus, we can ensure the existence of scalars $\alpha_0, \alpha_1, \ldots, \alpha_n$ so that:

$$\alpha_0 x + \alpha_1 Tx + \alpha_2 T^2 x + \cdots + \alpha_n T^n x = 0.$$

Let us define a polynomial $p(t) = \sum_{i=0}^{n} \alpha_i t^i$. Applying Corollaries D.8 and D.9, we can rewrite p as $p(t) = r_m(t)r_{m-1}(t)\ldots r_1(t)$ for $m \leq n$ where each polynomial r_i, $1 \leq i \leq m$, has degree 1 or 2. We can assume degree 1 if scalars are taken to be complex. We use this to prove the following result.

Proposition 2.1.2. *Suppose that $n \geq 1$. Then every operator T on X has an invariant subspace M of dimension 1 or 2. If X is complex, then M can be chosen to have dimension 1.*

Proof. Let x be an arbitrary vector of $X \backslash \{0\}$ and, as above, define a polynomial $p(t) = r_m(t)r_{m-1}(t)\ldots r_1(t)$ so that $p(T)x = 0$. Let us choose j to be the minimum index so that $r_j(T)r_{j-1}(T)\ldots r_1(T)x = 0$ and define $u = r_{j-1}(T)\ldots r_1(T)x = 0$ (if $j = 1$, simply let $u = x$). By minimality of j, we have that $u \neq 0$ and $r_j(T)u = 0$.

Recall that r_j has degree 1 or 2. First suppose that $r_j(t) = \alpha t + \beta$ for $\alpha \neq 0$. Then, we have $r_j(T)u = (\alpha T + \beta I)u = 0$. In other words,

$$Tu = -\alpha^{-1}\beta u.$$

We define $M = \langle \{u\} \rangle$. The space M is easily seen to be T-invariant, and M is 1-

dimensional since $u \neq 0$. Note that this is always possible when X is complex, since we may assume $deg(r_j) = 1$.

On the other hand, we may have $r_j(t) = \alpha t^2 + \beta t + \lambda$ where $\alpha \neq 0$. In this case we obtain:

$$T^2 u = -\alpha^{-1}\beta T u - \alpha^{-1}\lambda u$$

We simply choose $M = \langle\{u, Tu\}\rangle$, which is seen to be T-invariant and has dimension either 1 or 2. □

The ideas in the above proof are closely related to the well-known notions of eigenvalues and eigenvectors. A scalar λ is called an *eigenvalue* for T if there exists a non-zero vector x such that $Tx = \lambda x$, or equivalently, $(T - \lambda I)x = 0$; In this case, x is called an *eigenvector for T corresponding to λ*, or simply an *associated eigenvector* when T and λ are clear from context.

In proving Proposition 2.1.2, we used the fact that eigenvalues are equivalent to 1-dimensional invariant subspaces. Indeed, if λ is an eigenvalue for T with associated eigenvector x, then $\langle\{x\}\rangle$ is a 1-dimensional T-invariant subspace. On the other hand, if M is 1-dimensional and T-invariant, then any non-zero vector x of M is an eigenvector for T since the vectors of $\{x, Tx\} \subseteq M$ are linearly dependent. Let us now characterize operators on finite-dimensional spaces with non-trivial invariant subspaces.

Theorem 2.1.3. *Let X be an n-dimensional Banach space and $T : X \to X$ a linear operator. Then T has a non-trivial invariant subspace if, and only if, either $n \geq 3$ or $n = 2$ and T has an eigenvalue.*

Proof. First, if $n = 0$ or 1, then the only subspaces of X are $\{0\}$ and X and so T cannot have a non-trivial invariant subspace. In the case that $n = 2$, the only non-trivial subspaces of X are 1-dimensional. As noted in the paragraph before the theorem, the existence of a 1-dimensional invariant subspace for T is equivalent to T having an eigenvalue.

Finally if $n \geq 3$, then T has an invariant subspace M of dimension 1 or 2 by Proposition 2.1.2. Since the dimension of M differs from that of X and $\{0\}$, we must have that M is non-trivial. The result follows. $\square$

Any linear operator on a 2-dimensional complex space satisfies the hypotheses of Theorem 2.1.3 by Proposition 2.1.2. Consider the following example showing that linear operators on 2-dimensional real spaces may not have eigenvalues.

Example 2.1.4. Define an operator T_θ which rotates each vector in $\mathbb{R}^2$ by θ radians counter clockwise about the origin. The explicit definition of this operator on a vector $x = (x_1, x_2) \in \mathbb{R}^2$ is given as follows:

$$T_\theta x = (x_1 \cos(\theta) - x_2 \sin(\theta), x_1 \sin(\theta) + x_2 \cos(\theta))$$

Provided that $\mathbb{R}^2$ is equipped with the Euclidean norm, the operator T_θ is norm preserving; That is, $\|T_\theta x\| = \|x\|$ for every vector $x \in \mathbb{R}^2$ (by the identity $\cos^2(\theta) + \sin^2(\theta) = 1$). It follows that the only possible eigenvalues for T_θ are $\lambda = 1$ or -1. In either case, the equation $T_\theta x = \lambda x$ for a non-zero vector x implies that $\sin(\theta) = 0$. Therefore, T_θ has an eigenvalue if, and only if, θ is an integer multiple of π.

2.2 Cyclic Subspaces and Separability

The solution for finite-dimensional spaces breaks down in infinitely many dimensions for an obvious reason: There may be (and often is) no $n \geq 0$ such that the vectors of $S_n(x, T)$ are linearly dependent. In fact, the vectors of $S(x, T) = \{x, Tx, T^2x, \dots\}$ may even be linearly independent, as the next example shows.

Example 2.2.1. For $1 \leq p \leq \infty$ consider the *unilateral shift* operator U_+ on ℓ_p defined by $U_+(x_0, x_1, \dots) = (0, x_0, x_1, \dots)$. If x is non-zero then it is easy to see that the vectors of $S(x,T)$ are linearly independent. This is because distinct vectors in $S(x,T)$ begin with a different number of zeros.

However, U_+ does have a non-trivial invariant subspace (in fact, many). For example, the set of sequences with zero in the first coordinate (the range of U_+) is closed in ℓ_p and U_+-invariant.

Note. Throughout this section, let X be a Banach space over a field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. We let $T : X \to X$ denote an arbitrary bounded linear operator on X.

As it turns out, a natural object to study in the general case is the notion of a cyclic subspace. Given non-zero $x \in X$, we define the *cyclic subspace generated by* x, denoted $W(x,T)$, to be the smallest closed subspace of X containing $S(x,T)$. As an explicit formula, we can express it as

$$W(x,T) = [S(x,T)]. \tag{2.1}$$

For a non-zero vector $x \in X$, we say that x is a *cyclic vector* for T or that x is T*-cyclic* if $W(x,T) = X$.

Proposition 2.2.2. *The operator T has only the trivial invariant subspaces if, and only if, every non-zero vector of X is a cyclic vector for T.*

Proof. If M is a non-trivial invariant subspace for T, then for every non-zero vector $x \in M$ we have $W(x,T) \subseteq M$ and so x is not a cyclic vector for T. On the other hand, if $x \in X$ is a non-zero non-cyclic vector for T, then $W(x,T) \neq X$. Also, since $x \neq 0$ and $x \in W(x,T)$, we have $W(x,T) \neq \{0\}$. It follows that $W(x,T)$ is a non-trivial invariant subspace for T. The result follows. $\square$

It is sometimes useful to define cyclic subspaces in terms of polynomial combinations of T. Since $\{p(T)x : p(t) \in \mathbb{F}[t]\} = \langle\{T^n x : n \geq 0\}\rangle$ we have that

$$W(x,T) = \overline{\{p(T)x : p(t) \in \mathbb{F}[t]\}}. \tag{2.2}$$

Using this idea, it is immediate that cyclic subspaces are separable. Indeed, define a set Q by $Q = \mathbb{Q}$ if $\mathbb{F} = \mathbb{R}$ and $Q = \{a + bi : a, b \in \mathbb{Q}\}$ if $\mathbb{F} = \mathbb{C}$. The set D of polynomials over Q is easily seen to be countable. The fact that $\{p(T)x : p \in D\}$ is dense in $W(x,T)$ follows from density of $\mathbb{Q}$ in $\mathbb{R}$. We omit this argument here.

Theorem 2.2.3. *If X is non-separable, then every bounded linear operator $T : X \to X$ has a non-trivial invariant subspace.*

Proof. Let T be a bounded linear operator on a non-separable Banach space X, and choose $x \in X \backslash \{0\}$. Since $W(x,T)$ is separable and X is not, we have that $W(x,T) \neq X$. So, x is not T-cyclic. The result follows by Proposition 2.2.2. □

2.2.1 Some Additional Facts About Cyclic Vectors

The final results of this section highlight important properties of cyclic vectors that are used later in our discussion of Lomonosov's Theorem (Section 3.3) and Read's counterexample on ℓ_1 (Sections 4.2 - 4.9).

Proposition 2.2.4. *A vector $x \in X$ is T-cyclic if, and only if, for every non-empty open set U of X there is some $p(t) \in \mathbb{F}[t]$ such that $p(T)x \in U$.*

Proof. First suppose that x is T-cyclic and let U be an open set of X containing a point y. Then, since $y \in X = [\{p(T)x : p(t) \in \mathbb{F}[t]\}]$ there must be some polynomial p such that $p(T)x \in U$. On the other hand, if x is not T-cyclic, then $U = X - W(x,T)$ is a non-empty open set such that $p(T)x \notin U$ for any polynomial p. The result follows. □

Proposition 2.2.5. *Suppose that x_0 is T-cyclic. If x is a vector of X such that $x_0 \in W(x,T)$, then x is T-cyclic.*

Proof. Recall, $W(x_0,T) = X$ is the smallest invariant subspace of T containing x_0. Therefore, since $W(x,T)$ is T-invariant and contains x_0 we must have $X \subseteq W(x,T)$, the reverse inclusion being trivial. Therefore x is T-cyclic. □

Proposition 2.2.6. *An operator T on X has only the trivial invariant subspaces if, and only if, every unit vector is T-cyclic.*

Proof. The argument is simple. We have that T has only the trivial invariant subspaces if, and only if, every non-zero vector of X is T-cyclic by Proposition 2.2.2. However, a non-zero vector x is T-cyclic if, and only if, the unit vector $\|x\|^{-1}x$ is T-cyclic by Proposition 2.2.5. The result follows. □

The main result of Chapter 4 applies the following corollary.

Corollary 2.2.7. *Let T be an operator on X with cyclic vector x_0. The operator T has only the trivial invariant subspaces if, and only if, for every unit vector x and $\varepsilon > 0$ there is a polynomial q so that*

$$\|q(T)x - x_0\| < \varepsilon.$$

Chapter 3

Positive Results

We begin to investigate the more 'interesting' case of the invariant subspace problem: infinite-dimensional separable Banach spaces. Several of the crucial techniques apply only to complex spaces, so they are our central focus.

Note. In this chapter, X denotes an arbitrary Banach space and H a Hilbert space. Both spaces are assumed to be complex, infinite-dimensional and separable.

This chapter samples some of the important breakthroughs in showing that certain classes of bounded operators do indeed possess non-trivial invariant subspaces. Before presenting these results, we must cover important background on spectral theory and become familiar with the class of compact operators.

3.1 Spectral Theory and Compact Operators

Earlier in Section 2.1, we demonstrated the usefulness of eigenvalues in characterizing operators on finite-dimensional spaces with non-trivial invariant subspaces. Spectral theory extends the concept of eigenvalues to infinite-dimensional spaces in a natural way.

Definition 3.1.1. Given a bounded linear operator T on X we define the *spectrum* of T, denoted by $\sigma(T)$, to be the set of scalars $\alpha \in \mathbb{C}$ such that $T - \alpha I$ is not invertible (bijective). The *point spectrum* of T, denoted by $\sigma_p(T)$, is the set of all scalars $\lambda \in \mathbb{C}$ such that $T - \lambda I$ is not injective.

On an n-dimensional space, it is well known that a linear operator T is injective if, and only if, it is surjective and therefore we have $\sigma(T) = \sigma_p(T)$ in this case. This is not necessarily true for operators on infinite-dimensional spaces, although $\sigma_p(T) \subseteq \sigma(T)$ clearly holds.

Lemma 3.1.2. *If $T : X \to X$ is a bounded linear operator, then $\sigma_p(T)$ is precisely the set of eigenvalues for T.*

Proof. Suppose that $\lambda \in \sigma_p(T)$. Since $T - \lambda I$ is not injective, there are distinct vectors x and y in X such that $(T - \lambda I)(x) = (T - \lambda I)(y)$. This implies that $T(x - y) = \lambda(x - y)$. Since $x \neq y$ we have $x - y \neq 0$ and so $x - y$ is an eigenvector for T corresponding to λ.

Now, suppose that λ is an eigenvalue for T and let $x \neq 0$ be an associated eigenvector. Then we have $Tx = \lambda x$, which implies $(T - \lambda I)x = 0$. Since $T - \lambda I$ also maps the zero vector to 0, we have that $T - \lambda I$ is not injective. Therefore, $\lambda \in \sigma_p(T)$, as desired. □

We introduce the well-known concept of an eigenspace. While it may seem that this definition belongs more in Section 2.1, it was actually not necessary for our treatment of finite-dimensional spaces. Eigenspaces are needed briefly, however, for our proof of of Theorem 3.3.2 at the end of the chapter.

Definition 3.1.3. If $T : X \to X$ is a linear operator with eigenvalue λ, the *eigenspace* corresponding to λ is defined to be the set $W_\lambda = \{x \in X : Tx = \lambda x\}$.

It is clear that any eigenspace W_λ for T is T-invariant since T simply acts as a multiple of the identity on W_λ. It is easily shown that an eigenspace is also a closed subspace, provided that T is bounded.

Lemma 3.1.4. *If λ is an eigenvalue for a bounded linear operator $T : X \to X$, then W_λ is a non-trivial closed invariant subspace for X.*

As one can imagine, operators which are closely related to one another tend to share invariant subspaces. Consider, for example, the following definition.

Definition 3.1.5. Let T be a linear operator on X.

(a) We say that a linear operator A on X *commutes* with T if $AT = TA$.

(b) A subspace M of X is said to be a *hyperinvariant* subspace for T if $A(M) \subseteq M$ for every operator A which commutes with T.

Proposition 3.1.6. *Let T be an operator on X and suppose that λ is an eigenvalue for T with eigenspace W_λ. Then W_λ is a hyperinvariant subspace for T.*

Proof. Let A be an operator which commutes with T and fix any $x \in W_\lambda$. We have $TAx = ATx = A\lambda x = \lambda Ax$. Thus, $TAx = \lambda Ax$ and so $Ax \in W_\lambda$ as desired. □

Understanding the spectral properties of bounded operators is very important to many areas of operator theory. One of the crucial ideas is that the spectrum is bounded, and therefore contained within some closed disc in $\mathbb{C}$.

Definition 3.1.7. Suppose that $T : X \to X$ is a bounded operator such that $\sigma(T)$ is non-empty. The *spectral radius* of T, denoted by $r_\sigma(T)$, is defined by $r_\sigma(T) = \sup\{|\alpha| : \alpha \in \sigma(T)\}$.[1]

[1] It can also be shown that $\sigma(T)$ is compact, and therefore the supremum in the definition of r_σ can be replaced by a maximum, see [35, Section 3.3].

The following theorem gives us a very useful formula for calculating the spectral radius of a bounded operator. This result is quite well known, but the proof is quite involved and would distract from the main focus of this thesis; See for example [35, Section 3.3].

Theorem 3.1.8 (Gelfand Spectral Radius Formula)**.** *If $T : X \to X$ is bounded, then $\|T^n\|^{1/n} \to r_\sigma(T)$.*

Of special importance is the class of operators with spectral radius equal to zero, which are involved in a few of the results discussed later in the chapter.

Definition 3.1.9. A bounded linear operator $T : X \to X$ such that $\|T^n\|^{1/n} \to 0$ is said to be *quasinilpotent*.[2]

Example 3.1.10. We provide two examples of quasinilpotent operators on the Banach space ℓ_p for $1 \leq p \leq \infty$. First, consider the linear operator $T : \ell_p \to \ell_p$ defined by $T(x_0, x_1, \dots) = (0, x_0, 0, x_2, 0, \dots)$. Clearly, $T^2 = 0$ and therefore T is quasinilpotent (in fact, T is nilpotent).

For our next example, let $\alpha = (\alpha_n)_{i=1}^{\infty}$ be a sequence of positive real numbers such that $\alpha_n < n^{-n}$ for all $n \geq 1$. We let $T_\alpha : \ell_p \to \ell_p$ be the 'weighted backwards shift' operator $T_\alpha(x_0, x_1, \dots) = (\alpha_1 x_1, \alpha_2 x_2, \dots)$. It is easy to argue that for any $n \geq 2$ and vector x satisfying $\|x\| \leq 1$, we have $\|T_\alpha^n x\| \leq \alpha_n \alpha_{n-1} \dots \alpha_1 < \alpha_n$. Thus, $\|T_\alpha^n\|^{1/n} \leq n^{-1}$ and so T_α is quasinilpotent.

The next lemma highlights a simple property of quasinilpotent operators, which is used to prove a special case of Lomonosov's Theorem in Section 3.3.

Lemma 3.1.11. *If T is a quasinilpotent operator on X, then for every scalar c we have $\|(cT)^n\| \to 0$.*

---

[2] Quasinilpotent operators generalize nilpotent operators. An operator $T : X \to X$ is *nilpotent* if $T^n = 0$ for some $n \geq 1$.

Proof. Given $n \geq 1$ let us define $a_n = \|(cT)^n\|$, which is seen to equal $|c|^n\|T^n\|$. The sequence a_n must approach zero as $n \to \infty$ (in fact, rather quickly) since the sequence $a_n^{1/n} = |c|\|T^n\|^{1/n}$ approaches zero. □

3.1.1 Compact Operators

We give a brief but informative introduction to compact operators. With the exception of Section 3.4, compact operators are involved in all of the invariant subspace theorems in this chapter. The belief is that once the reader obtains a sufficient 'feel' for compact operators, these main results should seem somewhat natural and intuitive.

Definition 3.1.12. Let Y be a Banach space. A linear operator $K : X \to Y$ is said to be *compact* if $K(S)$ is relatively compact for every bounded subset S of X.

The definition of compact operators as it appears above is due to Riesz [46]. The notion of a *completely continuous operator*, first studied by Hilbert [31], is equivalent to a compact operator on a separable Hilbert space. The theory of compact operators is very rich, especially on Hilbert spaces. This section merely samples some of the most well-known and useful results.

Proposition 3.1.13. *Every compact linear operator $K : X \to Y$ is bounded.*

Proof. Recall, $\|K\| = \sup\{\|Kx\| : \|x\| \leq 1\}$. Since $S = \{x : \|x\| \leq 1\}$ is bounded and K is compact, we have that $\overline{K(S)}$ is compact. Of course, compact sets must be bounded. Indeed, if $F \subseteq X$ is unbounded, then the collection $\{B_n(0) : n \geq 1\}$ would be an open cover for F having no finite subcover and therefore F cannot be compact. Thus, $K(S) \subseteq \overline{K(S)}$ is also bounded. The result follows. □

As it turns out, simple examples of compact operators are not hard to come by. In fact, all bounded finite-rank operators are compact. This comes from the basic result from analysis: *A subset of a finite-dimensional normed space is relatively compact if, and only if, it is bounded.* Hence, the following result.

Proposition 3.1.14. *Every bounded finite-rank operator from a Banach space into a Banach space is compact.*

Definition 3.1.15. Let $\mathcal{K}(X, Y)$ denote the set of all compact operators mapping X to Y. We simply write $\mathcal{K}(X)$ in the case that $Y = X$.

The set $\mathcal{K}(X, Y)$ is a closed subspace of $\mathcal{B}(X, Y)$ as illustrated by the next proposition. Actually $\mathcal{K}(X)$ is a closed ideal in $\mathcal{B}(X)$. Recall, a subspace M of $\mathcal{B}(X)$ is called an *ideal* if $TA, AT \in M$ for every $T \in \mathcal{B}(X)$ and $A \in M$.[3]

Proposition 3.1.16. *We have the following:*

(a) If $K_1, K_2 \in \mathcal{K}(X, Y)$ are compact and α is a scalar, then $K_1 + K_2$ and αK_1 are compact.

(b) If $(K_n)_{n=0}^{\infty} \subseteq \mathcal{K}(X, Y)$ converges to a bounded operator K, then K is compact.

(c) If $T : X \to Y$ and $A : Y \to Z$ are bounded and one of T or A is compact, then TA is compact.

Using Propositions 3.1.14 and 3.1.16 (b), it follows that the norm limit of a monotone sequence of bounded finite-rank operators is compact. We use this fact to obtain our first example of a compact operator having infinite-dimensional range.

Example 3.1.17. We consider an operator on ℓ_p for $1 \leq p \leq \infty$. Let $(\alpha_i)_{i=0}^{\infty}$ be a sequence of positive real numbers such that $\alpha_i \to 0$ and let D be the 'diagonal operator' on ℓ_p defined by $D(x_0, x_1, \dots) = (\alpha_0 x_0, \alpha_1 x_1, \dots)$. Consider the sequence $(D_i)_{i=0}^{\infty}$ of

[3] Generally, subset I of a ring R is an *ideal* if I is closed under addition and for each $i \in I$ and $r \in R$ we have $ir, ri \in I$.

bounded finite-rank operators such that $D_i(x_0, x_1, \dots) = (\alpha_0 x_0, \alpha_1 x_1, \dots \alpha_i x_i, 0, 0, \dots)$. Each D_i is compact by Proposition 3.1.14. Given a vector x such that $\|x\| \leq 1$ we have $\|Dx - D_i x\| \leq \alpha_{i+1}$. Therefore, $D_i \to D$ and so D is compact. By a similar argument, the operator T_α from Example 3.1.10 is compact as well as being quasinilpotent.

As one can imagine, it is very useful to represent compact operators as the norm limit of a sequence of bounded finite-rank operators. The space X is said to have the *approximation property* if for every Banach space Y the set of bounded finite-rank operators mapping Y to X is dense in $\mathcal{K}(Y, X)$. Many important spaces have the approximation property including Hilbert spaces and ℓ_p spaces, $1 \leq p < \infty$.

A famous and longstanding open problem in functional analysis, called the 'approximation problem,' asked whether every Banach space has the approximation property. A counterexample was finally given in a 1973 paper by Per Enflo [18] on a separable Banach space.[4] Enflo is also famous for discovering the first counterexample for the invariant subspace problem on a Banach space [20]. See Chapter 4 for more information.

From the ideas presented so far, one should recognize that, in some sense, compact operators bridge the gap between bounded finite-rank operators and general bounded operators. Compact operators even have spectral properties that are very similar to operators on finite-dimensional spaces, as illustrated by the next theorem.

Theorem 3.1.18 (Riesz [46])**.** *If K is a compact operator on X, then we have $\sigma(K) = \sigma_p(K) \cup \{0\}$.*

A consequence of Theorem 3.1.18 is that non-zero scalar operators are not compact. This result can also be proved directly using the well-known fact that the unit

[4] Enflo's counterexample also solved a problem of Stanisław Mazur, posed in 1936. Mazur's problem was written in the so-called "Scottish Book" of open problems kept by Polish mathematicians who frequented the Scottish Café in Lwów. He offered the reward of a live goose to anyone who could come up with a solution. More than thirty years later, Mazur was true to his word. After lecturing on his solution in Warsaw, Enflo was awarded with a live goose.

sphere S_X is compact if, and only if, X is finite-dimensional.

Corollary 3.1.19. *If X is infinite-dimensional, then the only compact scalar operator on X is the zero operator.*

Proof. For some scalar $\lambda \neq 0$, define a scalar operator T by $Tx = \lambda x$ for all $x \in X$. We have that T is bijective and so 0 is not an element of $\sigma(T)$. Therefore T is not compact by Theorem 3.1.18. Clearly the zero operator is compact. The result follows. □

We obtain another immediate corollary by Lemma 3.1.11 and Theorem 3.1.18.

Corollary 3.1.20. *If $K : X \to X$ is compact with no eigenvalues, then K is quasinilpotent. Moreover, for an arbitrary scalar c we have that $\|(cK)^n\| \to 0$.*

3.2 von Neumann's Result and Some Extensions

Given a bounded finite-rank operator $F : H \to H$, the range of F is an immediate non-trivial invariant subspace for F. Since Hilbert spaces have the approximation property, it can be seen that each compact operator K on H admits a sequence of 'approximately K-invariant' finite-dimensional subspaces. While this argument does not prove that compact operators satisfy the invariant subspace problem, it does seem somewhat reasonable that this may be true.

Sometime during the 1930s John von Neumann proved that compact operators have non-trivial invariant subspaces, but did not decide to publish it. The proof was rediscovered and finally published by N. Aronszajn and K. T. Smith [7] in 1954.

Theorem 3.2.1 (von Neumann, proved in [7])**.** *Every compact operator on H has a non-trivial invariant subspace.*

While von Neumann's original proof uses orthogonal projections, and therefore applies only to Hilbert spaces, Aronszajn and Smith also included an alternative proof that extends to general Banach spaces.

von Neumann's Theorem resisted generalization for more than a decade after the Aronszajn and Smith paper, and not for lack of interest. Finally, in 1966 Bernstein and Robinson [13] extended the result to the slightly larger class of polynomially compact operators. A linear operator T on a Banach space is said to be *polynomially compact* if there is a non-zero polynomial $p \in \mathbb{C}[t]$ such that $p(T)$ is compact.

Theorem 3.2.2 (Bernstein and Robinson [13])**.** *Every polynomially compact operator on H has a non-trivial invariant subspace.*

Clearly all compact operators are polynomially compact by considering the polynomial $p(t) = t$, however, the converse is not true. Consider the following example.

Example 3.2.3. Let K be a compact operator on X and let α be any scalar such that $\alpha \notin \sigma(K)$. The operator $T = K - \alpha I$ is bijective, and therefore not compact by Theorem 3.1.18 (ie. $0 \notin \sigma(T - \alpha I)$). However, T is polynomially compact with polynomial $p(t) = t + \alpha$. Also, any nilpotent operator is polynomially compact (and may not be compact) by considering the polynomial $p(t) = t^n$ for sufficiently large n.

An interesting aspect of Bernstein and Robinson's proof is that it used the relatively new techniques of non-standard analysis, which builds up the foundations of analysis based on a rigorous definition of 'infinitesimal' numbers. Shortly after, the proof was translated into standard analysis by Halmos [26].

The next major generalization was achieved by Arveson and Feldman [8] in 1968. First, consider the following definition.

Definition 3.2.4. For a bounded linear operator T on X, the *uniformly closed algebra generated by T*, denoted by $\mathfrak{A}(T)$, is defined to be the subspace $[\{I, T, T^2, \dots\}]$ of $\mathcal{B}(X)$.

Alternatively, $\mathfrak{A}(T)$ is the smallest closed subspace of $\mathcal{B}(X)$ containing T and I which is closed under function composition.

If T is a bounded operator, then $\mathfrak{A}(T)$ can be thought of as the closure of the set of polynomial combinations of T, or the set of all operators which can be norm approximated by polynomial combinations of T.

Theorem 3.2.5 (Arveson and Feldman [8]). *If $T : H \to H$ is a bounded quasinilpotent operator such that $\mathfrak{A}(T)$ contains a non-zero compact operator, then T has a non-trivial invariant subspace.*

Some further generalizations were also discovered. For example, Arveson and Feldman's result was extended to Banach spaces [23], and to the following: *If T is quasinilpotent and the closure of the set of rational functions of T contains a non-zero compact operator, then T has a non-trivial invariant subspace* [36,38]. Also, the Arveson and Feldman's proof highlighted a new notion of *quasitriangular* operators, which would be extracted and studied by Halmos [27].

3.3 Lomonosov's Theorem

While the techniques of von Neumann and subsequent generalizations yielded many interesting and surprising theorems during the 1950s and 60s, their effectiveness was reaching its limit by the 70s. Just as this was occurring, a young mathematician named Victor Lomonosov introduced a new and powerful technique [34]. Recall the definition of a hyperinvariant subspace, Definition 3.1.5 (b).

Theorem 3.3.1 (Lomonosov [34]). *If A is a non-scalar operator on X which commutes with a non-zero compact operator K, then A has a non-trivial hyperinvariant subspace.*

Lomonosov's Theorem was a significant breakthrough for several reasons. For one, it applies to Banach spaces and not just Hilbert spaces. Even restricted to Hilbert spaces, however, Lomonosov's Theorem is still more general than anything that was previously known; see Proposition 3.3.3. Moreover, his technique allowed for a short and elegant proof.

Theorem 3.3.1 describes a 'commuting chain' of operators $K - A - T$ such that

- K is non-zero and compact,
- A is nonscalar, which implies that
- T has a non-trivial invariant subspace.

We provide a simple proof discovered by Hilden [37] of a weak version of the theorem. It contains many of the same ideas as Lomonosov's original proof, but avoids the technical Schauder Fixed Point Theorem. The tradeoff is that Hilden's proof only applies to commuting chains $K - T$ of length two. In Section 4.10 we give a delightful argument from [50] showing that Lomonosov's Theorem cannot be extended to commuting chains of length four.

Theorem 3.3.2 (Lomonosov [34]). *If $T : X \to X$ commutes with a non-zero compact operator K, then T has a non-trivial invariant subspace.*

Proof (Hilden, proved in [37]). The proof is by contradiction. Suppose to the contrary that T does not have a non-trivial invariant subspace. First, we may assume that $\|K\| = 1$ as the operator $\|K\|^{-1}K$ is compact and also commutes with T.

We argue that K cannot have any eigenvalues. Indeed, if K had an eigenvalue λ, then we would have $W_\lambda \neq X$ since K is non-scalar (Corollary 3.1.19). By Lemma 3.1.4 and Proposition 3.1.6 the subspace W_λ would be a non-trivial invariant subspace for T, a contradiction. Therefore, we assume that K has no eigenvalues. By Corollary 3.1.20 we deduce the following:

$$\|(cK)^n\| \to 0 \text{ for every scalar } c \in \mathbb{C} \tag{3.1}$$

Next, let us choose some $x_0 \in X$ such that $\|Kx_0\| > \|K\| = 1$. This implies $\|x_0\| > 1$ as well by the definition of the operator norm. Let U denote the set $B_1(x_0)$. By definition of x_0, neither the closure of U nor $K(U)$ contains the zero vector (this is the source of our contradiction). For each polynomial $p \in \mathbb{C}[t]$, let $\theta(p) = p(T)^{-1}(U)$. Since $p(T)$ is continuous, we have that $\theta(p)$ is open. Since we are assuming that T has only the trivial invariant subspaces, every non-zero vector of X must be a cyclic vector for T (Proposition 2.2.2). It follows that for every non-zero $x \in X$ there is some polynomial $p \in \mathbb{C}[t]$ such that $p(T)x \in U$ by Proposition 2.2.4. Thus, the collection $\{\theta(p)\}_p$ is an open cover $\overline{K(U)}$ (in fact, it covers all of $X \backslash \{0\}$). The set U is bounded and so $\overline{K(U)}$ is compact since K is a compact operator. It follows that there is a finite subcollection of $\{\theta(p)\}_p$ which covers $\overline{K(U)}$. Thus, we may let F be a finite set of polynomials such that $\{\theta(p)\}_{p\in F}$ covers $\overline{K(U)}$. Let $c = \max\{\|p(T)\| : p \in F\}$.

The rest of the proof has been appropriately called "Hilden's ping-pong technique." Since $Kx_0 \in K(U) \subseteq \overline{K(U)}$ we have that $Kx_0 \in \theta(p_1)$ for some $p_1 \in F$. Therefore, $p_1(T)Kx_0 \in U$. It follows that $Kp_1(T)Kx_0 \in K(U)$ and so we may choose $p_2 \in F$ such that $p_2(T)Kp_1(T)Kx_0 \in U$. Continuing this process for any positive integer n gives $p_n(T)Kp_{n-1}(T)K \dots p_1(T)Kx_0 \in U$ where $p_i \in F$ for each i. By (3.1), given arbitrary $\varepsilon > 0$ we may choose n large enough so that $\|(cK)^n x_0\| < \varepsilon$. We obtain the following inequality:

$$\|p_n(T)Kp_{n-1}(T)K \dots p_1(T)Kx_0\| = \|p_n(T)p_{n-1}(T) \dots p_1(T)K^n x_0\| \text{ (as } TK = KT)$$

$$\leq \|p_n(T)\|\|p_{n-1}(T)\| \dots \|p_1(T)\|\|K^n x_0\| \leq \|(cK)^n x_0\| < \varepsilon$$

Thus, we have a sequence of points in U converging to 0, contradicting the fact that $0 \notin \overline{U}$. Therefore, T must have a non-trivial invariant subspace. The result follows. □

All of the main results from Section 3.2 follow from Theorem 3.3.2, as we shall now demonstrate.

Proposition 3.3.3. *Let T and A be non-zero bounded linear operators on X. If the uniformly closed algebra generated by T contains A, then T commutes with A.*

Proof. Suppose that $A \in \mathfrak{A}(T)$. Given $\varepsilon > 0$, let p be a polynomial such that $\|p(T) - A\| < \frac{\varepsilon}{2\|T\|}$. We obtain the following inequality:

$$\|AT - TA\| \leq \|AT - p(T)T\| + \|Tp(T) - TA\| \text{ (since } p(T)T = Tp(T))$$

$$\leq \|A - p(T)\|\|T\| + \|T\|\|p(T) - A\| < \varepsilon$$

Therefore, we must have $\|AT - TA\| = 0$, which implies $AT = TA$. The result follows. □

Corollary 3.3.4. *If $T : H \to H$ is a bounded operator such that the uniformly closed algebra generated by T contains a non-zero compact operator, then T has a non-trivial invariant subspace.*

Many operator theorists were curious whether Lomonosov's Theorem could solve the invariant subspace problem, at least for complex separable Hilbert spaces. That is, it was not clear whether a bounded operator on a Hilbert space could fail to satisfy the hypotheses of Theorem 3.3.1. The first natural candidate was the unilateral shift U_+ on ℓ_2; however, an example of a non-scalar operator commuting with both U_+ and a compact operator was discovered by Cowen [16]. Finally in 1980 Hadwin et al. [24] discovered a class of operators, called 'quasianalytic shifts,' which do not satisfy Lomonosov's hypotheses, ending the seven year search.

We mention that Theorem 3.3.1 has been extended to real spaces by Hooker [32]. He proves that a bounded linear operator on a real or complex space which commutes with a compact operator and does not satisfy an irreducible polynomial equation has a non-trivial hyperinvariant subspace. On a complex space, this only rules out scalar operators by the Fundamental Theorem of Algebra, but the case for real spaces is more complicated. Fortunately, he also proves that a bounded linear operator on a real space commuting with a compact operator cannot satisfy any irreducible polynomial equation anyways.

Finally, we would like to point out that Lomonosov's Theorem can often provide us with not only an invariant subspace, but a sequence of nested invariant subspaces. For instance, let $K : X \to X$ be compact with invariant subspace M_1. It is easily shown that the restriction K_1 of K to M_1 is also compact. So, provided that the dimension of M_1 is at least 2, we have that K_1 has a non-trivial invariant subspace $M_2 \subsetneq M_1$, which is seen to be an invariant subspace for K as well. A similar property of normal operators is discussed in the next section.

3.4 Operators Related to Normal Operators

We begin with a standard definition from linear algebra.

Definition 3.4.1. Let T be a bounded linear operator on H. A bounded operator T^* is said to be the *adjoint* of T if for all $x, y \in H$ we have $\langle Tx, y \rangle = \langle x, T^*y \rangle$.

The fact that every bounded operator has a unique adjoint is well-known. The concept of an adjoint is extremely important to the study of linear operators on Hilbert spaces. In some sense, adjoints extend the idea of complex conjugation. To illustrate this, note that the adjoint of a scalar operator λI is $\overline{\lambda}I$. Indeed, for $x, y \in H$ we have

$$\langle \lambda x, y\rangle = \lambda\langle x, y\rangle = \lambda\overline{\langle y, x\rangle} = \overline{\overline{\lambda}\langle y, x\rangle} = \overline{\langle \overline{\lambda}y, x\rangle} = \langle x, \overline{\lambda}y\rangle.$$

The notion of normal operators (and matrices) is very well-studied. We provide this definition here.

Definition 3.4.2. A bounded linear operator N on H is said to be *normal* if N commutes with N^*.

Normal operators generalize self-adjoint and unitary operators. Recall, an operator T is self-adjoint if $T^* = T$ and unitary if $T^* = T^{-1}$.[5] Note, for example, that for a real scalar λ the operator λI is self-adjoint (since λ is self-conjugate).

The fact that normal operators have non-trivial invariant subspaces has been known for some time. This follows from a few results including Fuglede's Theorem and the Spectral Theorem for normal operators. We only mention Fuglede's Theorem briefly, for a deeper examination see [40, Section 1.5].

Theorem 3.4.3 (Fuglede's Theorem [22])**.** *Let T and N be bounded operators where N is normal. If T commutes with N, then T commutes with N^*.*

The study of operators related to normal operators has been an interesting and fruitful area of research. A natural generalization is the following, due to Halmos [25].

Definition 3.4.4. Let T be an operator on H. We say that T is *subnormal* if T is the restriction of a normal operator to an invariant subspace. That is, if there is a Hilbert space $H' \supseteq H$ and a normal operator N on H' that is equal to T on H.

Every normal operator is obviously subnormal by letting $H' = H$, but the converse is not true. In 1950 Halmos [25] asked whether every subnormal operator has

[5] Equivalently, an operator T is unitary if it has dense range and preserves the inner product. That is, $\langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in H$.

a non-trivial invariant subspace. This problem was finally solved by Scott Brown [14] in 1978. His solution made clever use of a deep result of Sarason [48] and introduced powerful new techniques.

Theorem 3.4.5 (Brown [14])**.** *Every subnormal operator on H has a non-trivial invariant subspace.*

Brown's theorem is equivalent to the following surprising result: *If M is an infinite-dimensional invariant subspace for a normal operator N, then M contains a proper subspace other than $\{0\}$ which is also N-invariant.* His work was applied and generalized by many researchers, including Brown himself who extended his result in [15] to hyponormal operators with 'thick' spectra.

Definition 3.4.6. An operator T on H is said to be *hyponormal* if $\|T^*x\| \leq \|Tx\|$ for all $x \in H$.

As it turns out, subnormal operators are hyponormal but the converse is not true. Brown's full result on hyponormal operators cannot be stated without developing a lot of necessary background. An interesting special case is given as follows.

Theorem 3.4.7 (Brown [15])**.** *If T is a hyponormal operator on H such that $\sigma(T)$ has non-empty interior, then T has a non-trivial invariant subspace.*

This concludes our treatment of operators on Banach and Hilbert spaces known to have non-trivial invariant subspaces. We admit that this merely scratches the surface of invariant subspace theorems, especially on Hilbert spaces. For a more in-depth study of invariant subspaces for operators on Hilbert spaces, see [40].

Chapter 4

Counterexamples on Banach Spaces

The theorems of Lomonosov and Brown together constitute the strongest evidence for a positive answer to the invariant subspace problem. The results of this chapter, however, contrast sharply. We discuss the known counterexamples to the problem on Banach spaces, and outline the proof of the simplest such example on ℓ_1.

4.1 History and Controversy

A few years after Lomonosov's Theorem was proved came another monumental breakthrough in the study of the invariant subspace problem. In 1975 Per Enflo discovered the first example of an operator on a Banach space having only the trivial invariant subspaces. He gave an outline of the proof in 1976 [19]. However, his full solution was not submitted until 1981 and did not appear in print until 1987 [20]. The delay was mainly due to the sheer complexity of Enflo's construction. It was so formidable that few people had the time and energy to work through the fine details. This, combined with the fact that early versions contained several minor errors, made Enflo's paper a nightmare for referees.

> Davie [17]: "Enflo's successful completion of the task is a remarkable achievement; however the latter part of his paper is so impenetrable that it is destined to be admired rather than read."

> Radjavi and Rosenthal [39]: "We've heard lots of rumors of the form: so-and-so spent two months working very hard on the manuscript, found minor correctable mistakes, got about 1/3 of the way through, then quit in exhaustion."

This is all in spite of the fact that, on the first level, Enflo's idea is both natural and interesting. He notices that a linear operator T with a cyclic vector x acts as a shift on the set $\{x, Tx, T^2x, \dots\}$. He likens this shift behaviour to the action of multiplying a polynomial $p \in \mathbb{C}[t]$ by the independent variable t. His goal then, is to construct a suitable norm $\|\cdot\|$ on the set of polynomials that achieves the following: (1) the operator T on the completion X which maps $T : p(t) \mapsto tp(t)$ is bounded, and (2) every non-zero $x \in X$ is T-cyclic. In other words, his idea is to construct the Banach space to suit his operator, not the other way around.

As Enflo's paper crawled through the publication process, C. J. Read developed a counterexample of his own and submitted it for publication [41]. The paper was of similar length and complexity to Enflo's, however it was published much earlier in 1984. He did not cite Enflo's work, except to say the following:

> Read [41]: "This is the only counterexample which the author knows to be valid. P. Enflo has produced two preprints purporting to contain examples of operators without an invariant subspace, obtained by very different methods to our own. However, these preprints have been in existence since 1976 and 1981, and neither has yet been published."

The fact that Read would publish his example first, without giving proper credit to Enflo, was seen by many mathematicians to be in bad form. This was made worse by the fact that several facets of Read's construction were actually based on Enflo's ideas. Like Enflo, Read begins with a shift operator on the set of finitely non-zero sequences (isomorphic to $\mathbb{C}[t]$) and constructs the space as he goes.

> Beauzamy [11]: "To give such precision is uncommon, and would be of no interest, if C. Read had not, several times, unelegantly and unsuccessfully, tried to claim priority towards the solution of the problem. This behaviour might be condemned with stronger words, but we remember we are presently writing for posterity."

There is no question among mathematicians that Enflo's ideas were behind the *first* counterexample on a Banach space, and for that he should receive full credit. Since then, however, many others have emerged. Beauzamy [10] provides a significant simplification of Enflo's original example (although, it is still very complicated). Apart from this, most advances have been developed by Read. The first such achievement came in 1985 when he managed to give a counterexample on the classical Banach space ℓ_1 [42]. A simplification was given by A. M. Davie, which appears in [11, Chapter XIV]. Another construction by Read of a linear opearator on ℓ_1 in [43] is the simplest known counterexample to date. We outline this construction in Section 4.2.

Read has also discovered counterexamples having additional interesting properties. For one, in 1997 he showed that his example from [43] could be modified to have the additional property of being quasinilpotent [45]. Even more interesting is an example given by Read in 1988 of an operator T on ℓ_1 such that every non-zero vector is hypercyclic [44] (as opposed to just cyclic). We say that a vector x of X is *hypercyclic* for T if the set $\{x, Tx, T^2x, \dots\}$, called the *orbit* of x, is dense in X. This example proves the incredible result: *There is an operator T on ℓ_1 such that no non-trivial closed set $S \subseteq \ell_1$ is T-invariant.*

Lastly, we make note of some recent work showing that the invariant subspace problem actually has a positive solution on certain infinite-dimensional separable Banach spaces. Argyros and Haydon [6] have shown that there is such a space X so that every operator $T : X \to X$ can be written as $T = K + \lambda I$ where K is compact and λ is a scalar.[1] By Lomonosov's Theorem, such operators are guaranteed to have non-trivial invariant subspaces, solving the problem positively in this case.

4.2 A Counterexample on ℓ_1

We guide the reader through the construction of the simplest known counterexample to the invariant subspace problem on the Banach space ℓ_1, due to C. J. Read [43]. See Appendix E for some of the relevent notation and definitions. The example is constructed as a simple shift operator on a basis $(e_i)_{i=0}^{\infty}$ for F, different from the canonical basis $(f_i)_{i=0}^{\infty}$. The properties of T are therefore directly determined by the construction of $(e_i)_i$, which we shall provide shortly.

The construction and proof revolve around a somewhat mysterious increasing sequence $\mathbf{d} = (d_i)_{i=1}^{\infty}$ of natural numbers. Often we require that $\mathbf{d}$ increases sufficiently rapidly so that certain deductions can be made. The phrase '$\mathbf{d}$ increases sufficiently rapidly' is taken to mean that for some $m > 1$ the number d_m is bounded below by a real-valued function of $d_1, d_2, \ldots, d_{m-1}$. We can ensure that the sequence $\mathbf{d}$ is well-defined, provided that only finitely many such bounds are required.

Note. We define $(a_i)_{i=1}^{\infty}$ and $(b_i)_{i=1}^{\infty}$ to be the sequences of odd and even terms of $\mathbf{d}$: $a_i = d_{2i-1}$ and $b_i = d_{2i}$ for $i \geq 1$. Let $a_0 = 1$, $v_0 = 0$ and define $v_n = n(a_n + b_n)$ for

[1] This solves the so-called 'scalar-plus-compact' problem.

$n \geq 1$.

As previously mentioned, the linear operator T is defined so that for each $i \geq 0$ we have $Te_i = e_{i+1}$. Since $(e_i)_{i=0}^{\infty} = (T^i e_0)_{i=0}^{\infty}$ spans F, which is dense in ℓ_1, we have that e_0 is T-cyclic regardless of how $(e_i)_i$ is defined. Our main goal of the proof is outlined as follows.

Objective. We show that for each unit vector x and $\varepsilon > 0$ there is some polynomial q so that

$$\|q(T)x - e_0\| < \varepsilon.$$

The fact that T has no non-trivial invariant subspaces follows by Corollary 2.2.7. Since $\mathbf{d}$ is unbounded it suffices to show that for any $k \geq 1$ there is a polynomial q so that $\|q(T)x - e_0\| < \frac{3}{a_{k-1}}$. In order to achieve this, we focus on an associated vector y based on x such that

- y is an element of some special compact set $K_{k,n}$ for $n > k$;
- We can obtain sufficient control over $\|q(T)(x - y)\|$.

We then decompose x into $x = (x - y) + y$. The polynomial q is actually defined via a compactness argument *after* strategically choosing $y \in K_{k,n}$. Our conclusion to the proof has the following form:

$$\|q(T)x - e_0\| \leq \|q(T)(x - y)\| + \|q(T)y - e_0\|$$

$$\leq \frac{1}{a_{k-1}} + \frac{2}{a_{k-1}} = \frac{3}{a_{k-1}} < \varepsilon.$$

On the way to this conclusion, we need several bounds regarding polynomial combinations of T on y. Let us move on with the construction of $(e_i)_i$.

4.3 Constructing $(e_i)_i$

The first application of $\mathbf{d}$ is to induce a partitioning $\mathscr{P}$ of the natural numbers into finite intervals. The purpose of $\mathscr{P}$ is to facilitate the construction of $(e_i)_i$. For each $i \geq 1$, the vector e_i is then defined based on which set of $\mathscr{P}$ contains i.

Definition 4.3.1. We define $\mathscr{P}$ as the collection of all intervals in $\mathbb{N}$ of the following form, for $1 \leq r \leq n$.

$$(v_{n-1}, a_n) \text{ if } r = 1; \tag{4.1.1}$$

$$(v_{n-r+1} + (r-1)a_n, ra_n) \text{ if } r > 1; \tag{4.1.2}$$

$$[ra_n, v_{n-r} + ra_n]; \tag{4.2}$$

$$(na_n + (r-1)b_n, r(a_n + b_n)); \tag{4.3}$$

$$[r(a_n + b_n), na_n + rb_n]. \tag{4.4}$$

Note that we may assume that $a_n > v_{n-1} + 1$ and $b_n > (n-1)a_n + 1$ for all $n \geq 1$ to ensure that intervals described in (4.1.1), (4.1.2), and (4.3) are non-empty.

We are required to show that $\mathscr{P}$ is indeed a partitioning of the natural numbers. We do not prove this rigorously, as it should be easy enough to see after going over a few examples.

Example 4.3.2. Let us consider the sets of $\mathscr{P}$ for fixed $n = 2$. The interval $(v_1, v_2]$ can be written as a union of the following (disjoint, non-empty) sets:

$$(v_1, a_2) \cup [a_2, v_1 + a_2] \cup (v_1 + a_2, 2a_2) \cup [2a_2, 2a_2] \cup$$

$$(2a_2, a_2 + b_2) \cup [a_2 + b_2, 2a_2 + b_2] \cup (2a_2 + b_2, v_2) \cup [v_2, v_2]$$

Each set of $\mathscr{P}$ for $n = 2$ appears in the above union.[2] A more general argument shows that the sets of $\mathscr{P}$ for fixed n form a partitioning of $(v_{n-1}, v_n]$. Since $v_0 = 0$ we have that $\mathscr{P}$ partitions $\mathbb{N}$. We suggest that the reader tries to duplicate this example for $n = 3$ to get a better 'feel' for how the partition works in general.

Therefore, we may assume that every integer $i \geq 1$ is contained in precisely one set of $\mathscr{P}$. We are now in position to define the sequence $(e_i)_{i=0}^{\infty}$. We warn the reader that the full definition is complex, and is difficult to comprehend out of context. Fortunately, it is not yet necessary to grasp the fine details of the definition, only a few minor facts which are clarified shortly.

Definition 4.3.3. Define

$$f_0 = e_0. \tag{4.5}$$

Given $I \in \mathscr{P}$ and $i \in I$, we define e_i in the following way:

$$\text{If } I \text{ is as in (4.1.1) or (4.1.2), let } f_i = 2^{(h-i)/\sqrt{a_n}} e_i \text{ where } h = (r - 1/2)\, a_n; \tag{4.6}$$

$$\text{If } I \text{ is as in (4.2), let } f_i = a_{n-r}(e_i - e_{i-ra_n}); \tag{4.7}$$

$$\text{If } I \text{ is as in (4.3), let } f_i = 2^{(h-i)/\sqrt{b_n}} e_i \text{ where } h = (r - 1/2)\, b_n; \tag{4.8}$$

$$\text{If } I \text{ is as in (4.4), let } f_i = e_i - b_n e_{i-b_n}. \tag{4.9}$$

Remark 4.3.4. Let $1 \leq r \leq n$. For some of the arguments given later, it is important to notice the following:

(a) $f_{ra_n} = a_{n-r}(e_{ra_n} - e_0)$ by (4.7).

(b) If $i \in (na_n + b_n, 2(a_n + b_n))$, then $e_i = 2^{(i-\frac{3}{2}b_n)/\sqrt{b_n}} f_i$ by (4.8).

---

[2] Taking the sets in the order that they appear in the union, the sets are described in (4.1.1), (4.2), (4.1.2), (4.2), (4.3), (4.4), (4.3), (4.4). The sequence of r values is $1, 1, 2, 2, 1, 1, 2, 2$.

(c) If $i \in [a_n + b_n, na_n + b_n]$, then $e_i = f_i + b_n e_{i-b_n}$ by (4.9).

4.4 First Properties of $(e_i)_i$

For each integer $m \geq 0$ let E_m be the span of $(e_i)_{i=0}^m$. In Definition 4.3.3, each e_i is defined by writing f_i in terms of $e_0, e_1, \ldots, e_i$ where e_i makes a non-zero contribution. Using this fact, we obtain a simple but important lemma.

Lemma 4.4.1. *For $m \geq 0$ we have $E_m = F_m$. Moreover, $\cup_m E_m = \cup_m F_m = F$.*

Proof. Each vector f_i for $0 \leq i \leq m$ can be written as $f_i = \sum_{j=0}^{i} \lambda_{m,j} e_j$ where $\lambda_{i,i} \neq 0$. This is easily seen to be invertible. That is, e_i can be written as $\sum_{j=0}^{i} \beta_i e_i$ where $\beta_i \neq 0$. Therefore, we obtain $E_m = F_m$. The fact that $\cup_m E_m = \cup_m F_m = F$ follows immediately. □

Recall that the construction of each vector e_i for $i \geq 1$ depends heavily on the sequence $\mathbf{d}$. However, for fixed i the vector e_i only depends on *certain* terms of $\mathbf{d}$. Consider the following.

Note 4.4.2. Let $i \geq 1$ and $n \geq 1$ be given.

- If $i \leq na_n$, then the definition of e_i depends only on the choice of (at most) $a_1, b_1, \ldots, b_{n-1}, a_n$;
- If $i \leq v_n$, then the definition of e_i depends only on the choice of (at most) $a_1, b_1, \ldots, a_n, b_n$.

For each integer $m \geq 0$ let $J_m : F_m \to F_m$ be the linear isomorphism such that

$$J_m(e_i) = f_i$$

for $0 \leq i \leq m$.

In the case that $m = na_n$ for some $n \geq 1$, the vectors $e_0, e_1, e_2, \ldots, e_m$ depend only on the choice of $a_1, b_1, \ldots, b_{n-1}, a_n$ by Note 4.4.2. Thus, J_m (and, in particular, its norm) is affected only by the choice of these terms of $\mathbf{d}$ and so we may define a real-valued function M_n such that:

$$M_n(a_1, b_1, \ldots, b_{n-1}, a_n) \geq \max\{\|J_{na_n}\|, \|J_{na_n}^{-1}\|\}. \tag{4.10}$$

Similarly, by Note 4.4.2 there is a function N_n mapping into the real numbers so that:

$$N_n(a_1, b_1, \ldots, a_n, b_n) \geq \max\{\|J_{v_n}\|, \|J_{v_n}^{-1}\|\}. \tag{4.11}$$

The purpose of functions M_n and N_n is to highlight the fact that, for example, b_n can be chosen large with respect to $(a_i)_{i=0}^{n}$ and $(b_i)_{i=0}^{n-1}$ without affecting the value of $\|J_{na_n}\|$.

4.5 The Linear Operator

We define the linear operator T which is the central focus of the counterexample. For now, we consider T as a linear operator on F. Later we see that T can be extended uniquely to a bounded operator on ℓ_1.

Definition 4.5.1. Given the sequence $(e_i)_{i=0}^{\infty}$, we define $T : F \to F$ to be the unique linear operator such that

$$T(e_i) = e_{i+1}$$

for each $i \geq 0$.

In some sense, T is based on the unilateral shift U_+. This makes it somewhat surprising that T would not have any non-trivial invariant subspaces, since there are many U_+-invariant subspaces which are quite easy to describe. For example, the range of U_+ suffices, see Example 2.2.1. The construction of T has a fundamental difference in the fact that $T^m(F) = \langle\{e_i : i \geq m\}\rangle$ is dense in ℓ_1. In particular, $\overline{T(F)}$ contains e_0, as we shall now demonstrate.

Lemma 4.5.2. *Given $n > k \geq 1$, we have*

$$\|e_{(n-k+1)a_n} - e_0\| = \frac{1}{a_{k-1}}.$$

Proof. Letting $r = n - k + 1$ we recognize that $f_{ra_n} = a_{n-r}(e_{ra_n} - e_0)$ by Remark 4.3.4 (a). So, we have $\|e_{ra_n} - e_0\| = \frac{1}{a_{n-r}} = \frac{1}{a_{k-1}}$. □

4.5.1 A Truncated Version Of T

At certain points it is useful to study the action of T on certain finite subspaces. For this, we define a truncated version of T on F_m.

Definition 4.5.3. Let $T_m : F_m \to F_m$ be the linear operator so that

$$T_m(e_i) = \begin{cases} e_{i+1} & \text{for } 0 \leq i \leq m-1, \\ 0 & \text{for } i = m. \end{cases}$$

One useful property of T_m is that it allows us to 'isolate' certain vectors e_i. To see what we mean, consider the following lemma.

Lemma 4.5.4. *Let $y = \sum_{i=\alpha}^{m} \lambda_i e_i$ where $\lambda_\alpha \neq 0$. Then there is some polynomial r so that $r(T_m)y = e_\alpha$ and $deg(r) \leq m - \alpha$.*

Proof. Given $j \geq 1$ we have $(T_m)^j y = \sum_{i=\alpha}^{m} \lambda_i (T_m)^j e_i = \sum_{i=\alpha+j}^{m} \lambda_{i-j} e_i$ by definition of

T_m. So, for $j \geq 1$ we have

$$y - \lambda_{\alpha+j}\lambda_\alpha^{-1}(T_m)^j y = \sum_{i=\alpha}^{\alpha+j-1} \lambda_i e_i + \sum_{i=\alpha+j}^{m} (\lambda_i - \lambda_{\alpha+j}\lambda_\alpha^{-1}\lambda_{i-j})e_i$$

$$= \sum_{i=\alpha}^{\alpha+j-1} \lambda_i e_i + \sum_{i=\alpha+j+1}^{m} (\lambda_i - \lambda_{\alpha+j}\lambda_\alpha^{-1}\lambda_{i-j})e_i.$$

Essentially, we have taken a vector $y = \sum_{i=\alpha}^{m} \lambda_i e_i$ and constructed a new vector $y' = \sum_{i=\alpha}^{m} \beta_i e_i$ where $\beta_{\alpha+j} = 0$. Now it should be clear that there is a polynomial r satisfying the above properties. □

4.5.2 A Norm On The Set Of Polynomials

For ease of demonstration, it is useful to define a norm function $\|\cdot\|_p$ on the set $\mathbb{F}[t]$ of polynomials. We define $\|\cdot\|_p$ in a way that is analogous to the standard 1-norm on F, where polynomials $1, t, t^2, \ldots$ are taken to be the standard unit vectors in $\mathbb{F}[t]$.

Definition 4.5.5. Given a polynomial $p = \sum_{i=0}^{m} \alpha_i t^i$ over $\mathbb{F}$, define $\|p\|_p = \sum_{i=0}^{m} |\alpha_i|$.

By applying a slight variation of Proposition E.1, it is clear that $\|T\|_e = 1$. Using this fact, we obtain the following lemma.

Lemma 4.5.6. *If $x \in F$ and p is a polynomial, then $\|p(T)x\|_e \leq \|p\|_p \|x\|_e$.*

Proof. Let $p(t) = \sum_{i=0}^{m} \alpha_i t^i$. We obtain

$$\|p(T)x\|_e \leq \|p(T)\|_e \|x\|_e$$

$$\leq \sum_{i=0}^{m} |\alpha_i| \|T^i\|_e \|x\|_e \leq \|p\|_p \|x\|_e,$$

as desired. □

4.6 A Compactness Argument

We define an operator to help us construct a sequence of compact sets which is crucial to the proof.

Definition 4.6.1. For $n > k \geq 1$ let $\tau_{k,n} : F_{na_n} \to F_{na_n}$ be the unique linear operator so that:

$$\tau_{k,n}(e_i) = \begin{cases} e_i & \text{for } 0 \leq i < (n-k)a_n, \\ 0 & \text{for } (n-k)a_n \leq i \leq na_n. \end{cases}$$

The linear operator $\tau_{k,n}$ is a standard projection $F_{na_n} \to F_{(n-k)a_n-1}$, where $n > k \geq 1$. Our sequence of compact sets $K_{k,n}$ is defined as follows.

Definition 4.6.2. For $n > k \geq 1$ let us define a set $K_{k,n} \subseteq F_{na_n}$ in the following way:

$$K_{k,n} = \left\{ y \in F_{na_n} : \|y\| \leq a_n \text{ and } \|\tau_{k,n} y\| \geq \frac{1}{a_n} \right\}.$$

Lemma 4.6.3. *For $n > k \geq 1$ the set $K_{k,n}$ is non-empty and compact.*

Proof. We have that $e_0 \in K_{k,n}$ since $\|e_0\| = \|f_0\| = 1 \leq a_n$ and $\|\tau_{k,n} e_0\| = \|e_0\| = 1 \geq \frac{1}{a_n}$ for all $n \geq 2$. Therefore $K_{k,n}$ is non-empty.

Now, to prove compactness it suffices to show that $K_{k,n}$ is closed and bounded since F_{na_n} is finite-dimensional. The linear operator $\tau_{k,n}$ is continuous since F_{na_n} is finite-dimensional. It follows that the $\tau_{k,n}^{-1}(S)$ is closed for any closed subset S of F_{na_n}. The intersection of two closed sets is closed, and so

$$K_{k,n} = \{y \in F_{na_n} : \|y\| \leq a_n\} \cap \tau_{k,n}^{-1}\left(\left\{ y \in F_{na_n} : \|y\| \geq \frac{1}{a_n} \right\} \right)$$

is closed. Also, $y \in K_{k,n}$ implies $\|y\| \leq a_n$ by definition, and so $K_{k,n}$ is bounded. □

Let us give our first of many results regarding polynomial combinations of T and T_{na_n} on vectors of $K_{k,n}$.

Lemma 4.6.4. *Let $n > k \geq 1$ and choose $y \in K_{k,n}$. Then there is a polynomial p so that $t^{a_n} \mid p$, $deg(p) \leq na_n$, and $p(T_{na_n})y = e_{(n-k+1)a_n}$.*

Proof. Let us write $y = \sum_{i=\alpha}^{na_n} \lambda_i e_i$ where $\lambda_\alpha \neq 0$. By Lemma 4.5.4 there is a polynomial r so that $r(T_{na_n})y = e_\alpha$ and $deg(r) < na_n - \alpha$. Since $y \in K_{k,n}$ we have that $\tau_{k,n}y \neq 0$ and so $\alpha < (n-k)a_n$. Therefore $j = (n-k+1)a_n - \alpha > a_n$ and the polynomial $p(t) = t^j r(t)$ satisfies the desired properties. □

Here is how the compactness argument works. By Lemma 4.6.4 for every $y \in K_{k,n}$ there is a polynomial p so that

$$t^{a_n} \mid p, \tag{4.12}$$

$$deg(p) \leq na_n. \tag{4.13}$$

$$\|p(T_m)y - e_{(n-k+1)a_n}\| < \frac{1}{a_n}, \tag{4.14}$$

By Lemma 4.6.4 and compactness of $K_{k,n}$ there is a finite set $P_{k,n}$ of polynomials satisfying (4.12) and (4.13) so that for every $y \in K_{k,n}$ there is some $p \in P_{k,n}$ satisfying (4.14). Now, notice that the choice polynomials which constitute $P_{k,n}$ are dependent only on the definition of vectors $(e_i)_{i=0}^{na_n}$. It follows by Note 4.4.2, and the fact that $P_{k,n}$ is finite, that for each $n \geq 2$ there is a real-valued function L_n so that

$$\|p\|_p \leq L_n(a_1, b_1, \ldots, b_{n-1}, a_n) \text{ for each } k < n \text{ and } p \in P_{k,n}.$$

Let us summarize the results of this section.

Proposition 4.6.5. *Let $n > 1$ be given. There is a real-valued function L_n with the following property. For each k so that $1 \leq k < n$ and $y \in K_{k,n}$ there is a polynomial p satisfying*

$$t^{a_n} \mid p,$$

$$deg(p) \leq na_n,$$

$$\|p\|_p \leq L_n(a_1, b_1, \ldots, b_{n-1}, a_n),$$

$$\|p(T_{na_n})y - e_0\| \leq \frac{1}{a_n} + \frac{1}{a_{k-1}}.$$

Proof. The polynomial p is chosen from the set $P_{k,n}$ so that $\|p(T_{na_n})y - e_{(n-k+1)a_n}\| < \frac{1}{a_{k-1}}$. The fact that p satisfies the first three properties is simply by definition. For the fourth property we invoke the triangle inequality. We have

$$\|p(T_{na_n})y - e_0\| \leq \|p(T_{na_n})y - e_{(n-k+1)a_n}\| + \|e_{(n-k+1)a_n} - e_0\|$$

$$\leq \frac{1}{a_n} + \frac{1}{a_{k-1}}.$$

by Lemma 4.5.2. The result follows. □

4.7 Tweaking Our Polynomials

In this section, we would like to construct, for each $y \in K_{k,n}$, a polynomial q satisfying $\|q(T)y - e_0\| < \frac{2}{a_{k-1}}$ in such a way that q has many other desirable properties. As it turns out, the polynomial that we want is $q(t) = \frac{t^{b_n}}{b_n}p(t)$ where p is chosen as in Proposition 4.6.5. It is clear that q satisfies the following:

$$t^{a_n+b_n} \mid q; \tag{4.15}$$

$$deg(q) \leq na_n + b_n; \tag{4.16}$$

$$\|q\|_p = \frac{\|p\|_p}{b_n} \leq \frac{L_n(a_1, b_1, \ldots, b_{n-1}, a_n)}{b_n}. \tag{4.17}$$

We obtain the following bound via the triangle inequality:

$$\|q(T)y - e_0\| \leq \left\| q(T)y - \frac{T^{b_n}}{b_n} p(T_{na_n})y \right\| +$$

$$\left\| \frac{T^{b_n}}{b_n} p(T_{na_n})y - p(T_{na_n})y \right\| + \|p(T_{na_n})y - e_0\| \tag{4.18}$$

Notice that we have already proven $\|p(T_{na_n})y - e_0\| \leq \frac{1}{a_n} + \frac{1}{a_{k-1}}$ in Proposition 4.6.5. Therefore, we may obtain control over $\|q(T)y - e_0\|$ by observing bounds on $\left\| q(T)y - \frac{T^{b_n}}{b_n} p(T_{na_n})y \right\|$ and $\left\| \frac{T^{b_n}}{b_n} p(T_{na_n})y - p(T_{na_n})y \right\|$. This is done in Lemmas 4.7.3 and 4.7.2 respectively.

Let us assume that $\mathbf{d}$ increases so rapidly that for each n the values a_n and b_n satisfy

$$b_n \geq L_n(a_1, b_1, \ldots, b_{n-1}, a_n) M_n(a_1, b_1, \ldots, b_{n-1}, a_n) a_n^2, \tag{4.19}$$

$$b_n \geq L_n(a_1, b_1, \ldots, b_{n-1}, a_n) 2^{(n-1)a_n + 1} a_{n-1}, \tag{4.20}$$

$$a_n \geq a_{n-1} N_{n-1}(a_1, b_1, \ldots, a_{n-1}, b_{n-1}). \tag{4.21}$$

$$b_n \geq 4na_n, \tag{4.22}$$

$$a_n \geq 3a_{n-1}. \tag{4.23}$$

These bounds are applied in the coming results.

Lemma 4.7.1. *Given $n > k \geq 1$ choose $y \in K_{k,n}$ and let p be a polynomial as in Proposition 4.6.5. We have that*

$$\|p(T)y\|_e \leq L_n(a_1, b_1, \ldots, b_{n-1}, a_n) M_n(a_1, b_1, \ldots, b_{n-1}, a_n) a_n.$$

Proof. By Lemma 4.5.6 and by Lemma E.3 we get

$$\|p(T)y\|_e \leq \|p\|_p \|y\|_e \leq \|p\|_p \|J_{na_n}\| \|y\|.$$

This is bounded above by $L_n(a_1, b_1, \ldots, b_{n-1}, a_n) M_n(a_1, b_1, \ldots, b_{n-1}, a_n) a_n$ by (4.10), definition of p, and the fact that $\|y\| \leq a_n$. □

Lemma 4.7.2. *Given $n > k \geq 1$ choose $y \in K_{k,n}$, and let p be a polynomial as in Proposition 4.6.5. We have*

$$\left\| \frac{T^{b_n}}{b_n} p(T_{na_n}) y - p(T_{na_n}) y \right\| \leq \frac{1}{a_n}.$$

Proof. Let us write $p(T_{na_n})y = \sum_{i=a_n}^{na_n} \lambda_i e_i$. Recall that $e_i = f_i + b_n e_{i-b_n}$ whenever $i \in [a_n + b_n, na_n + b_n]$ by Remark 4.3.4 (c). Therefore

$$\left\| \frac{T^{b_n}}{b_n} p(T_{na_n}) y - p(T_{na_n}) y \right\| = \left\| \frac{T^{b_n}}{b_n} \left(\sum_{i=a_n}^{na_n} \lambda_i e_i \right) - \sum_{i=a_n}^{na_n} \lambda_i e_i \right\|$$

$$= \left\| \sum_{i=a_n+b_n}^{na_n+b_n} \frac{\lambda_{i-b_n}}{b_n} e_i - \sum_{i=a_n}^{na_n} \lambda_i e_i \right\| = \left\| \sum_{i=a_n+b_n}^{na_n+b_n} \frac{\lambda_{i-b_n}}{b_n} (f_i + b_n e_{i-b_n}) - \sum_{i=a_n}^{na_n} \lambda_i e_i \right\|$$

$$= \left\| \sum_{i=a_n+b_n}^{na_n+b_n} \frac{\lambda_{i-b_n}}{b_n} f_i \right\| = \frac{1}{b_n} \sum_{i=a_n}^{na_n} |\lambda_i|.$$

This last expression is equal to $\frac{1}{b_n}\|p(T_{na_n})y\|_e$. Applying Lemma 4.7.1, we get

$$\frac{1}{b_m} \|p(T_{na_n}) y\|_e \leq \frac{1}{b_n} L_n(a_1, b_1, \ldots, b_{n-1}, a_n) M_n(a_1, b_1, \ldots, b_{n-1}, a_n) \|y\|.$$

Recall that $y \in K_{k,n}$ and so $\|y\| \leq a_n$. Therefore, by (4.19) this final expression is bounded above by $\frac{1}{a_n}$, as desired. The result follows. □

Lemma 4.7.3. *Given $n > k \geq 1$ choose $y \in K_{k,n}$. Let p be a polynomial as in Proposition 4.6.5, and $q(t) = \frac{t^{b_n}}{b_n} p(t)$. We have*

$$\left\| q(T)y - \frac{T^{b_n}}{b_n} p(T_{na_n})y \right\| \leq \frac{1}{a_n}.$$

Proof. Since $t_n^a \mid p$, $deg(p) \leq na_n$ and $y \in F_{na_n}$ we can write $p(T)y = \sum_{i=a_n}^{2na_n} \lambda_i e_i$ for scalars λ_i. Note that $p(T_{na_n})y = \sum_{i=a_n}^{na_n} \lambda_i e_i$ since T_{na_n} is truncated. So, $p(T)y - p(T_{na_n})y = \sum_{i=na_n+1}^{2na_n} \lambda_i e_i$. We get

$$q(T)y - \frac{T^{b_n}}{b_n} p(T_{na_n})y = \frac{T^{b_n}}{b_n} \left(p(T)y - p(T_{na_n})y \right) = \frac{T^{b_n}}{b_n} \left(\sum_{i=na_n+1}^{2na_n} \lambda_i e_i \right)$$

Therefore

$$\left\| q(T)y - \frac{T^{b_n}}{b_n} p(T_{na_n})y \right\| = \left\| \sum_{i=na_n+b_n+1}^{2na_n+b_n} \frac{\lambda_{i-b_n}}{b_n} e_i \right\| \leq \sum_{i=na_n+b_n+1}^{2na_n+b_n} \left\| \frac{\lambda_{i-b_n}}{b_n} e_i \right\|$$

$$\leq \frac{\|p(T)y\|_e}{b_n} \max\{\|e_i\| : na_n + b_n < i \leq 2na_n + b_n\},$$

where

$$\|p(T)y\|_e \leq L_n(a_1, b_1, \ldots, b_{n-1}, a_n) M_n(a_1, b_1, \ldots, b_{n-1}, a_n) a_n$$

by Lemma 4.7.1. So, given (4.19), we obtain

$$\left\| q(T)y - \frac{T^{b_n}}{b_n} p(T_{na_n})y \right\| \leq \frac{\max\{\|e_i\| : na_n + b_n < i \leq 2na_n + b_n\}}{a_n}.$$

Now it suffices to show that $\|e_i\| \leq 1$ whenever $i \in (na_n + b_n, 2na_n + b_n]$. By (4.22) we have $b_n \geq 4na_n > 2(n-1)a_n$. This implies that $2na_n + b_n < 2(a_n + b_n)$, and so $(na_n + b_n, 2na_n + b_n] \subseteq (na_n + b_n, 2(a_n + b_n))$. Thus, we have $e_i = 2^{(i - \frac{3}{2}b_n)/\sqrt{b_n}} f_i$ by

Remark 4.3.4 (b) and so $\|e_i\| = 2^{(i-\frac{3}{2}b_n)/\sqrt{b_n}}$. Dealing with the exponent, we see that $b_n + 2na_n \geq i$ and (4.22) imply

$$i - \frac{3}{2}b_n \leq b_n + 2na_n - \frac{3}{2}b_n = 2na_n - \frac{1}{2}b_n \leq 0.$$

So, we have $\|e_i\| \leq 1$. The result follows. □

Let us summarize the results of this section and the last.

Proposition 4.7.4. *Let $n > k \geq 1$ and choose some $y \in K_{k,n}$. There is a polynomial q satisfying the following:*

$$t^{a_n+b_n} \mid q;$$

$$deg(q) \leq na_n + b_n;$$

$$\|q\|_p \leq \frac{1}{2^{(n-1)a_n+1}a_{k-1}},$$

$$\|q(T)y - e_0\| \leq \frac{2}{a_{k-1}}.$$

Proof. The polynomial q is chosen as in Lemma 4.7.3. The fact that q satisfies the first two properties is simply by definition. For the third, we have

$$\|q\|_p = \frac{L_n(a_1, b_1, \ldots, b_{n-1}, a_n)}{b_n} \leq \frac{1}{2^{(n-1)a_n+1}a_{k-1}},$$

by (4.17) and (4.20).

For the fourth property of q, we combine the bounds in (4.18), Proposition 4.6.5, and Lemmas 4.7.2 and 4.7.3 to obtain:

$$\|q(T)y - e_0\| \leq \frac{1}{a_n} + \frac{1}{a_n} + \left(\frac{1}{a_n} + \frac{1}{a_{k-1}}\right)$$

We have that this final expression is bounded above by $\frac{2}{a_{k-1}}$ by (4.23). The result follows. □

4.8 Final Arguments

The work up to this point has allowed us to obtain a certain amount of control over vectors in $K_{k,n}$. However, there are still a few crucial questions which remain unanswered.

- Is T bounded on F?
- How should we decompose an arbitrary unit vector x into $(x-y)+y$ for $y \in K_{k,n}$?

We take the time now to settle these questions, before moving on to prove the main result. The special decomposition of unit vectors makes use of a special finite-rank linear operator, Q_n^0, which we define now.

Definition 4.8.1. For $n \geq 1$, let $Q_n^0 : F \to F_{na_n}$ be the linear operator such that

$$Q_n^0(f_i) = \begin{cases} f_i & \text{for } 0 \leq i \leq na_n, \\ -a_{m-r}e_{i-ra_m} & \text{for } i \in [ra_m, v_{m-r} + ra_m], \text{ where } 1 \leq m-n < r \leq m, \\ 0 & \text{otherwise.} \end{cases}$$

Let us obtain an explicit upper bound on $\|Q_n^0\|$.

Lemma 4.8.2. *We have* $\|Q_n^0\| \leq a_n$.

Proof. Recall that $\|Q_n^0\| = \sup\{\|Q_n^0 f_i\| i \geq 0\}$ by Proposition E.1. We have that $Q_n^0 f_i = f_i$ or 0 unless $j \in [ra_m, ra_m + v_{m-r}]$ for $0 < m-n < r \leq m$. In this case we have

$$\|Q_n^0 f_i\| = a_{m-r}\,\|e_{i-ra_n}\| \leq a_{n-1}\left\|J_{v_{n-1}}^{-1} f_{i-ra_n}\right\|$$

$$\leq a_{n-1}N_{n-1}(a_1, b_1, \ldots, a_{n-1}, b_{n-1}) \leq a_n$$

by (4.11) and (4.21). The result follows. □

The next two lemmas are crucial to the proof. However, we have chosen to omit their proofs as they are overly technical and not particularly enlightening. For proofs, see [43].

Lemma 4.8.3. *For any unit vector x and $k \geq 1$ there is some $n > k$ so that $y = Q_n^0 x \in K_{k,n}$.*

Lemma 4.8.4. *Provided that* **d** *increases sufficiently rapidly, then for every $n \geq 1$ we have:*

$$\|T\| < 2$$

and

$$\|T^{a_n+b_n}(I - Q_n^0)\| < 2.$$

By Lemmas 4.8.2 and 4.8.4 we have that T and Q_n^0 can be extended uniquely to bounded operators on ℓ_1 (by Proposition C.16). For the rest of the proof, the symbols T and Q_n^0 refer to their extensions to ℓ_1. The next result gives us control over the value of $\|q(T)(I - Q_n^0)\|$, which is needed to prove the main result.

Proposition 4.8.5. *Let y be a vector in $K_{k,n}$ and choose a polynomial q as in Proposition 4.7.4. We have $\|q(T)(I - Q_n^0)\| < \frac{1}{a_{k-1}}$.*

Proof. Let r be the polynomial so that $q(t) = t^{a_n+b_n} r(t)$. Note that $deg(r) \leq (n-1)a_n$. Also, q and r have the same coefficients and so $\|q\|_p = \|r\|_p$. We have that

$$\|q(T)(I - Q_n^0)\| = \|r(T)T^{a_n+b_n}(I - Q_n^0)\| \leq \|r(T)\|\|T^{a_n+b_n}(I - Q_n^0)\|$$

$$< 2\|r(T)\| \leq 2\|r\|_p\|T\|^{deg(r)} \leq \|q\|_p 2^{(n-1)a_n+1} \leq \frac{1}{a_{k-1}},$$

by Lemma 4.8.4 and Proposition 4.7.4. □

4.9 Every Unit Vector Is T-Cyclic

We are now in position to prove the main result.

Theorem 4.9.1. *The operator $T : \ell_1 \to \ell_1$ has only the trivial invariant subspaces.*

Proof. It suffices to show that every unit vector is T-cyclic. So, we let x be any unit vector and let $\varepsilon > 0$ be arbitrary. Choose k large enough so that $\frac{3}{a_{k-1}} < \varepsilon$. Our goal is to prove that there exists a polynomial q so that $\|q(T)x - e_0\| \leq \frac{3}{a_{k-1}}$.

By Lemma 4.8.3 we may let $n > k$ so that $y = Q_n^0 \in K_{k,n}$. Let us choose a polynomial q as in Proposition 4.7.4. We have

$$\|q(T)x - e_0\| \leq \|q(T)(x-y)\| + \|q(T)y - e_0\|$$

$$\leq \|q(T)(I - Q_n^0)\|\|x\| + \|q(T)y - e_0\|$$
$$< \frac{1}{a_{k-1}} + \frac{2}{a_{k-1}} = \frac{3}{a_{k-1}} < \varepsilon$$

by Proposition 4.8.5, definition of q, and the fact that $\|x\| = 1$. The result follows. □

Surprisingly, Read [43] also shows that under additional conditions on $\mathbf{d}$ we can ensure that either: (1) T^k has no non-trivial invariant subspaces for any integer $k \geq 1$, or alternatively (2) T^k has non-trivial invariant subspaces for all $k \geq 1$.

4.10 Sharpness of Lomonosov's Theorem

Recall that Lomonosov's Theorem applies to a 'commuting chain' of operators $K - A - T$ where A is non-scalar and K is compact. Intuitively, one may wonder whether the result can be generalized to a longer chain, perhaps $K - A_1 - A_2 - T$ where both A_1 and A_2 are non-scalar. We provide a clever observation of Troitsky [50] showing that this is not possible. Throughout this section, let T be Read's operator, which is defined earlier in the chapter. We suppose, without loss of generality, that every term of the sequence $\mathbf{d}$ is even.

Let us define $A_1 = T^2$. Clearly A_1 is non-scalar and commutes with T. Also, let A_2 to be the unique operator on ℓ_1 satisfying the following:

$$A_2 e_i = \begin{cases} e_i & i \text{ is even;} \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 4.10.1. *We have* $A_2 f_i = \begin{cases} f_i & i \textit{ is even;} \\ 0 & \textit{otherwise.} \end{cases}$

Proof. Consider the cases indicated by Definition 4.3.3. In any of them, we have that f_i can be written as a linear combination of vectors e_i, e_{i-ra_n} and e_{i-b_n} for some values of r and n. Notice that the indices $i, i - ra_n$ and $i - b_n$ have the same parity since a_n and b_n are even. The result follows. □

Corollary 4.10.2. *The operator A_2 commutes with A_1.*

Proof. Consider operators $A_1 A_2 = T^2 A_2$ and $A_2 A_1 = A_2 T^2$. We have, by definition:

$$T^2 A_2 e_i = \begin{cases} T^2 e_i = e_{i+2} & i \text{ is even;} \\ 0 & \text{otherwise.} \end{cases}$$

and

$$A_2 T^2 e_i = \begin{cases} A_2 e_{i+2} = e_{i+2} & i \text{ is even;} \\ A_2 e_{i+2} = 0 & \text{otherwise.} \end{cases}$$

Therefore, we have $A_1 A_2 = A_2 A_1$. The result follows. □

Now, simply define an operator K on ℓ_1 such that

$$Kf_0 = \begin{cases} f_i & i = 0; \\ 0 & \text{otherwise.} \end{cases}$$

We have that K is a bounded finite-rank operator, and therefore compact by Proposition 3.1.14. It is easily seen that K commutes with A_2, completing the example.

It is interesting to note that T^2 is seen to have a non-trivial invariant subspace by Lomonosov's Theorem, provided that each term of $\mathbf{d}$ is even. In fact, as Troitsky [50] points out, if $m \geq 2$ divides each term of $\mathbf{d}$, then a similar argument can be used with T^m in place of T^2.

Chapter 5

Conclusion

There are a few questions which continue to linger, of which the following is most obvious: (1) Do bounded operators on infinite-dimensional separable Hilbert spaces have non-trivial invariant subspaces? Given the fact that nobody knows the answer to the first question, one might ask a second: (2) Is there at least a consensus among mathematicians regarding what the answer *should* be?

Among the many open problems in mathematics, the invariant subspace problem is a somewhat exceptional in the fact that the answer to (2) is certainly "no."[1] When it comes to this problem, most mathematicians remain noncommittal, and it is quite rare for anyone to put forth a guess.[2] While the powerful results of Lomonosov and Brown pull in one direction, the impressive counterexamples of Enflo and Read pull in the other.

Enflo has noted in [20] that there are some major challenges in constructing counterexamples on reflexive Banach spaces (which include Hilbert spaces), and feels that this offers some weak evidence for a positive answer to the problem. To this day,

[1]There are very few mathematicians who deny the truth of, for example, the twin prime conjecture.
[2]It is called the invariant subspace *problem*, after all, and not *conjecture*.

these challenges have not been overcome as there are still no examples on reflexive Banach spaces.

On the other hand, some subsequent work in this area may suggest a negative answer. First, the problem on Banach spaces has been shown to be not only false, but *very* false; We have already mentioned Read's example in [44] of a bounded linear operator so that non-zero vectors are hypercyclic. Also, in [11] Beauzamy cites an example of a bounded operator T on a Hilbert space that he has constructed with a hypercyclic vector x_0 so that $p(T)x_0$ is hypercyclic whenever p is a polynomial with complex coefficients (we have not been able to find this example in print). While it is still possible that certain vectors may not be T-cyclic, this is the closest anyone has come to a counterexample on a Hilbert space.

Another valid and pertinent question is the following: (3) Why should we care about invariant subspaces in the first place? Part of the allure of the problem is simply that it is so simple to state, and yet disproportionally difficult to solve. However, there must be some more practical reason for people to direct so much time and energy towards it. There are very few (if any) instances where researchers assume the truth of the invariant subspace problem and derive important or surprising consequences. So, what is the point?

History has shown that the study of the invariant subspaces of a certain class of operators is often one of the first steps towards a rich and useful structural theory for those operators, as is elegantly explained in [12]. One might consider, for example, the work of Ringrose [47] on compact operators, which was inspired by the results of Aronszajn and Smith. Also, as was briefly mentioned in Section 3.2, a very deep and rewarding study of quasitriangular operators began as a result of Arveson and Feldman's invariant subspace theorem. This culminated in an impressive series of papers by Apostol and others which established a surprising link between quasitriangular and the more well-known semi-Fredholm operators, see [1–5]. The work of Brown on operators related to normal operators also contained a wealth of information on the structure of subnormal and hyponormal operators and initiated an interesting line of research. As noted in [33], the understanding of invariant subspaces can also be applied to develop the theory of functional calculus.

As a final question, we pose the following: (4) Will the invariant subspace problem ever be solved? It is, of course, impossible for us to judge whether the solution might be beyond the potential of human understanding. However, we tend to side with humanity on this particular issue. The future promises to bring many bright minds to fill the shoes of great operator theorists such as Enflo, Read, Lomonosov and Brown; Many of the secrets behind bounded operators will surely be unlocked. In the meantime, however, perhaps one of the most remarkable things about the invariant subspace problem is its ability to keep operator theorists humble.

Appendix A

Vector Spaces

Vector spaces are the main objects of interest in functional analysis. In order to understand the content of this thesis, is important to have a good handle of the basic properties of vector spaces, norms and (to a lesser extent) inner products. Most of the definitions here can be found in any linear algebra textbook, such as [9].

Definition A.1. Given a field $\mathbb{F}$, a set V is called a *vector space over* $\mathbb{F}$ if there are operations $+ : V \times V \to V$ and $\cdot : \mathbb{F} \times V \to V$ such that for $u, v, w \in V$ and $a, b \in \mathbb{F}$ the following hold:

1. $u + v = v + u$ (commutativity of vector addition);
2. $(u + v) + w = u + (v + w)$ (associativity of vector addition);
3. There exists $0 \in V$ such that for any $x \in V$ we have $0 + x = x$ (vector additive identity);
4. For any $x \in V$ there exists $-x \in V$ such that $x + (-x) = 0$ (vector additive inverse);
5. $a(bu) = (ab)u$ (associativity of scalar multiplication);

6. $(a + b)u = au + bu$ (distributivity of scalar sums);

7. $a(u + v) = au + av$ (distributivity of vector sums);

8. For $1 \in \mathbb{F}$, $1u = u$ (scalar multiplication by identity).

If V is a vector space over $\mathbb{F}$, the elements of V and $\mathbb{F}$ are referred to as *vectors* and *scalars* respectively. The operation $+$ is called *vector addition* and $\cdot$ is called *scalar multiplication*. We say that V is a *real* or *complex* vector space if $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ respectively.

Definition A.2. A subset W of a vector space V is called a *subspace* of V if W is closed under the operations of addition and scalar multiplication on V.

Remark A.3. Every subspace W of a vector space V is a vector space under the operations of vector addition and scalar multiplication inherited from V.

Definition A.4. Let S be a subset of a vector space V over $\mathbb{F}$. A vector $v \in V$ of the form $v = \sum_{i=1}^{n} a_n v_n$ where $n \geq 1$, $a_i \in \mathbb{F}$, and $v_i \in S$ for $1 \leq i \leq n$ is called a *linear combination* of vectors in S.

Definition A.5. If V is a vector space and $S \subseteq V$ is non-empty, we let the *linear span* (or simply the *span*) of S be the set $\langle S \rangle$ of all linear combinations of vectors in S. In the case that $S = \emptyset$, we define $\langle S \rangle = \{0\}$. Alternatively, $\langle S \rangle$ is the intersection of all subspaces of V containing S.

Remark A.6. If S is a subset of a vector space V, then $\langle S \rangle$ is a subspace of V. Moreover, S is a subspace of V if, and only if, $\langle S \rangle = S$.

Definition A.7. The vectors of a subset S of a vector space V are said to be *linearly dependent* if there is some positive integer n, vectors $v_1, v_2, \ldots, v_n \in S$, and scalars $a_1, a_2, \ldots, a_n$, not all zero, such that $\sum_{i=1}^{n} a_n v_n = 0$. Otherwise, the vectors of S are said to be *linearly independent*.

Definition A.8. If V is a vector space, a subset β of V is called a *basis* for V if the vectors of β are linearly independent and $\langle\beta\rangle = V$.

The next Theorem ensures the existence of a vector space basis. The proof relies on the widely accepted (but somewhat controversial) 'Axiom of Choice' from the ZFC axioms of set theory.

Theorem A.9. *Given any linearly independent subset S of a vector space V there is a basis β for V such that $S \subseteq \beta$. In particular, every vector space has a basis.*

Theorem A.10. *If V is a vector space, then any two bases of V have the same cardinality.*

Definition A.11. The *dimension* of a vector space V, denoted by $dim(V)$, is defined to be the cardinality of a basis β for V. If $dim(V) = n$ where n is an integer, we say that V is *finite-dimensional* or *n-dimensional*, otherwise V is said to be *infinite-dimensional*.

Remark A.12. If the vectors of $S \subseteq V$ are linearly independent, then $|S| \leq dim(V)$.

Remark A.13. If W is a subspace of V, then $dim(W) \leq dim(V)$. In particular, every subspace of a finite-dimensional vector space is finite-dimensional.

Appendix B

Norms and Inner Products

In the most simple vector spaces, such as $\mathbb{R}^n$, it is intuitive that each vector is determined by a length and a direction in relation to other vectors; however, the concepts of length and direction are not found in the basic definition of a vector space (Definition A.1). In fact, for certain vector spaces we are not able to define these concepts in a natural way. For this reason, we distinguish the vector spaces where length (norm) and direction relative to other vectors (inner product) can be defined to suit our intuition.

Definition B.1. Given a vector space V over a field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, a *norm* on V is a function $\|\cdot\| : V \to \mathbb{R}^+$ such that for $u, v \in V$ and $a \in \mathbb{F}$ the following hold:

1. $\|u\| = 0$ if, and only if, u is the zero vector (positive definiteness);

2. $\|au\| = |a| \, \|u\|$ (positive scalability);

3. $\|u + v\| \leq \|u\| + \|v\|$ (triangle inequality).

A vector space V with a norm is called a *normed space*.

Definition B.2. If V is a vector space over a field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, an *inner product* on V is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ such that for $x, y, z \in V$ and $a, b \in \mathbb{F}$ the following hold:

1. $\langle x, x \rangle \geq 0$ where equality holds if, and only if, $x = 0$ (positive definiteness);

2. $\overline{\langle x, y \rangle} = \langle y, x \rangle$ where the bar denotes complex conjugation (conjugate symmetry);

3. We have $\langle ax + by, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle$ (linearity of the first argument).

A vector space V with an inner product is called an *inner product space*.

Note. The notation for an inner product should not be confused with that of a linear span (Definition A.5). An inner product has two arguments, both of which are vectors, while the linear span has one argument, a set.

Definition B.3. For a subset S of an inner product space V, the *orthogonal complement* is defined to be the set $S^{\perp} = \{x \in V : \langle x, y \rangle = 0 \text{ for all } y \in S\}$ (read "S perp").

Remark B.4. If W is a subspace of V, then $W^{\perp}$ is a subspace of V.

Remark B.5. If V is an inner product space, then the function $\|\cdot\| : V \to \mathbb{R}^{+}$ defined by $\|x\| = \sqrt{\langle x, x \rangle}$ is a norm on V. This is referred to as the *norm induced by* $\langle \cdot, \cdot \rangle$, or simply the *induced norm*. Unless stated otherwise, the norm on an inner product space is the induced norm.

B.1 Some Norm Topology

Topological notions such as open, closed and compact sets are crucial to many areas of mathematics, including functional analysis. One of the most important topologies for a set is the *metric topology* induced by a distance function called a *metric*. Normed spaces inherit a natural distance function from their norm: the distance between vectors x and y is defined by $\|x - y\|$. Thus, a very natural topology for a normed space is the *norm topology*, also called the *strong topology*, induced by this distance function.

Definition B.6. Given a vector $x_0 \in X$ and a positive real number δ, we define the *open ball of radius* δ *around* x_0 to be the set $B_\delta(x_0) = \{x : \|x - x_0\| < \delta\}$.

Definition B.7. Given a normed space V we say that a subset S of V is *open* if for every $x_0 \in S$ there exists $\delta > 0$ such that $B_\delta(x_0) \subseteq S$. A subset F of V is said to be *closed* if $V \backslash F$ is open.

Remark B.8. If $\{U_\alpha\}_\alpha$ and $\{U_i\}_{i=1}^n$ are collections of open sets, then $\cup_\alpha U_\alpha$ and $\cap_i U_i$ are open.

Remark B.9. Every open ball is an open set of X.

Definition B.10. Given a normed space V, we define the *open unit ball* and *unit sphere* of V to be the sets $B_V = B_1(0)$ and $S_V = \{x \in V : \|x\| = 1\}$ respectively.

Continuous functions are vital to many areas of mathematics. Through the study of continuous functions we are given a way of comparing, identifying, and understanding the topological properties of different mathematical objects. The following definition is stated for normed spaces in particular, but applies to general topological spaces.

Definition B.11. Let V and W be normed vector spaces. A function $f : V \to W$ is *continuous* if $f^{-1}(U)$ is an open set of V whenever U is an open set of W.

Definition B.12. Given a subset S of a normed space V a point $x_0 \in V$ is said to be a *limit point* of S if for every open set U containing x_0 there exists some $x \in S \backslash \{x_0\}$ such that $x \in U$. We let S' denote the set of limit points of S. We call the set $\overline{S} = S \cup S'$ the *closure* of S. Alternatively, $\overline{S}$ is the intersection of all closed sets containing S.

Remark B.13. A subset $S \subseteq V$ is closed if, and only if, $S = \overline{S}$.

Definition B.14. We say that a subset S of a normed space V is *dense* in V if $\overline{S} = V$.

Definition B.15. Given a subset S of a normed vector space V, the *closed linear span* of S is the set $[S]$ defined by $[S] = \overline{\langle S \rangle}$. Alternatively, $[S]$ is the intersection of all closed subspaces of V containing S.

Remark B.16. For any subset S of a normed vector space V, the closed linear span of S is a subspace of V.

Definition B.17. For a subset S of a normed space V a collection $\mathscr{C}$ is said to be a *cover* for S if $S \subseteq \cup_{U \in \mathscr{C}} U$. In the case that each element of $\mathscr{C}$ is open in V, we say that $\mathscr{C}$ is an *open cover* for S.

Definition B.18. We say a subset S of a normed space V is *compact* if every open cover of S has a finite subcover; That is, a finite subcollection which is also a cover for S.

Definition B.19. A subset S of a normed space V is said to be *relatively compact* if $\overline{S}$ is compact.

Definition B.20. A subset S of a normed space V is said to be *bounded* if there is a real number M such that $\|x\| \leq M$ for every $x \in S$. A sequence $(x_n)_{n=0}^{\infty} \subseteq V$ is *bounded* if the set $S = \{x_n : n \geq 0\}$ is bounded.

Definition B.21. Let $(x_n)_{n=0}^{\infty}$ be a sequence in a normed space V. Given $x \in V$, we say that $(x_n)_{n=0}^{\infty}$ *converges (strongly)* to x, written $x_n \to x$, if for every open set U containing x there exists some $N \in \mathbb{N}$ such that if $n > N$, then $x_n \in U$. If $(x_n)_{n=0}^{\infty}$ converges to some $x \in V$, then we simply say that $(x_n)_{n=0}^{\infty}$ *converges.*

Remark B.22. For a subset $S \subseteq V$, the closure of S is precisely the set of vectors $x \in V$ such that there exists a sequence $(x_n)_{n=0}^{\infty} \subseteq S$ converging to x.

B.2 Banach Spaces and Hilbert Spaces

The real numbers have the important property of being complete. In real analysis, the property of completeness is usually introduced in terms of suprema and the 'least upper bound property' of $\mathbb{R}$ as an ordered field. Completeness can also be looked at in terms of the convergence of Cauchy sequences. In this way, the concept of completeness generalizes to arbitrary metric spaces and, in particular, normed spaces.

Definition B.23. Given a normed space V we say that a sequence $(x_n)_{n=0}^{\infty} \subseteq V$ is a *Cauchy sequence* if for every $\varepsilon > 0$ there exists an integer N such that if $n, m > N$, then $\|x_n - x_m\| < \varepsilon$. A normed space V is said to be *complete* if every Cauchy sequence in V converges.

Definition B.24. Given a normed space V, the *completion* of V is the smallest complete normed space X containing V.

Definition B.25. A normed space X which is complete is called a *Banach space*.

Remark B.26. Every closed subspace of a Banach space is a Banach space.

Example B.27. The following are examples of Banach spaces:

1. The n-dimensional real or complex space $\mathbb{R}^n$ or $\mathbb{C}^n$ with term wise vector addition and scalar multiplication and the *Euclidean norm* $\|x\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$ for $x \in X$;

2. The space ℓ_p for $1 \leq p < \infty$ is the set of all sequences $x = (x_n)_{n=0}^{\infty} \in \mathbb{R}$ or $\mathbb{C}$ such that $\sum_n |x_n|^p$ converges. Vector addition and scalar multiplication are defined term wise (in the usual way) and the norm on ℓ_p, sometimes called the *p-norm*, is defined by $\|x\| = \sqrt[p]{\sum_n |x_n|^p}$;

3. The set of all sequences $x = (x_n)_{n=0}^{\infty}$ in $\mathbb{R}$ or $\mathbb{C}$ of bounded modulus is denoted by ℓ_∞, with the usual operations and with the *uniform norm* or *supremum norm* $\|x\| = \sup_n\{|x_n|\}$;

4. Let $C[0,1]$ be the set of all real continuous functions on the interval $[0,1] \subset \mathbb{R}$. The set $C[0,1]$ becomes a Banach space under point wise addition and scalar multiplication and the supremum norm $\|f\| = \sup_{x\in[0,1]}\{f(x)\}$.

Definition B.28. A Banach space X is said to be *separable* if X has a countable dense subset.

Example B.29. The spaces $\mathbb{R}^n$, $\mathbb{C}^n$, and ℓ_p for $1 \leq p < \infty$ (see Example B.27) are separable spaces. This follows from two basic facts: (1) $\mathbb{Q}$ is dense in $\mathbb{R}$; and (2) The set of finitely non-zero sequences over $\mathbb{Q}$ is countable and dense in ℓ_p, $1 \leq p < \infty$.

However, the space ℓ_∞ is non-separable. To see this, consider the set A of sequences having coordinates equal to 0 or 1. Clearly every such sequence is bounded, and so we have $A \subseteq \ell_\infty$. The elements of A are in a 1-to-1 correspondence with the set of all subsets of $\mathbb{N}$, and therefore A is uncountable.[1] Also, any pair of distinct points $x, y \in A$ satisfies $\|x - y\| \geq 1$. Therefore, the sets of the collection $\mathscr{C} = \{B_{1/2}(x) : x \in A\}$ are mutually disjoint. Suppose that S is any dense subset of ℓ_∞. Then every set in $\mathscr{C}$ must contain an element of S, and therefore S is uncountable.

Definition B.30. A complete inner product space is called a *Hilbert space.*

Example B.31. The following are examples of Hilbert spaces:

1. $\mathbb{R}^n$ with the *standard inner product* (also called the *dot product*) $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ for $x, y \in \mathbb{R}^n$, which clearly induces the Euclidean norm;

2. $\mathbb{C}^n$ with the *standard inner product* $\langle x, y\rangle = \sum_{i=1}^n x_i \overline{y_i}$ for $x, y \in \mathbb{C}^n$, which again induces the Euclidean norm;

[1] The collection of subsets of a set B is known as the *power set* of B, denoted by $\mathscr{P}(B)$. A famous result of Cantor states that $|\mathscr{P}(B)| > |B|$ for any set B. Thus, $\mathscr{P}(\mathbb{N})$ is uncountable.

3. The vector space ℓ_2 with the inner product $\langle x, y\rangle = \sum_n x_n \overline{y_n}$, where $\overline{y_n}$ denotes the complex conjugate of y_n. This inner product induces the 2-norm from Example B.27.

Appendix C

Linear Operators

Loosely speaking, linear operators are functions between vector spaces which preserve the basic algebraic structure. One may think of linear operators as the natural categorical morphisms between vector spaces, similar to the homomorphisms in group theory, or continuous functions in topology.

Definition C.1. Given vector spaces V and W over a common field $\mathbb{F}$, a *linear operator* is a function $T : V \to W$ such that for all vectors $u, v \in V$ and scalars $a, b \in \mathbb{F}$ we have $T(au + bv) = aT(u) + bT(v)$. For $x \in V$ we often write Tx when referring to $T(x)$.

Definition C.2. Let V be a vector space. The operator $I_V : V \to V$ defined by $I_V x = x$ for all $x \in V$ is called the *identity operator*. We simply denote I_V by I when the space is clear from context.

Definition C.3. We let $\mathcal{L}(V, W)$ denote the set of all linear operators mapping V to W, or simply $\mathcal{L}(V)$ in the case that $W = V$.

Remark C.4. The set $\mathcal{L}(V, W)$ is a vector space under the usual point wise addition and scalar multiplication; ie. $(aT_1 + bT_2)x = aT_1x + bT_2x$ for all x.

Definition C.5. For a linear operator $T : V \to W$ we define the *kernel* (or *nullspace*) of T to be the set $ker(T) = \{x \in V : Tx = 0\}$. The *range* of T is defined by $T(V) = \{Tx : x \in V\}$.

Remark C.6. If $T : V \to W$ is a linear operator, then $ker(T)$ is a subspace of V and $T(V)$ is a subspace of W.

Definition C.7. Given a linear operator $T : V \to W$, we define the *rank* of T to be the dimension of $T(V)$. We say that T is a *finite-rank* operator if $T(V)$ is finite-dimensional.

The next definition introduces the important idea of a bounded operator, the main focus of this thesis.

Definition C.8. Given normed spaces V and W the *operator norm* is defined by $\|T\| = \sup\{\|Tx\| : \|x\| \leq 1\}$ for $T \in \mathcal{L}(V, W)$. An operator $T : V \to W$ is said to be *bounded* if $\|T\| < \infty$. The set of all bounded linear operators mapping V to W is denoted by $\mathcal{B}(V, W)$, or $\mathcal{B}(V)$ in the case that $W = V$.

Note. Be careful to recognize that there are in fact 3 different 'norms' in the above definition: the norms on V and W as well as the operator norm on $\mathcal{B}(V, W)$.

Remark C.9. The operator norm satisfies the conditions of a norm on $\mathcal{B}(V, W)$ (Definition B.1). If V and W are Banach spaces, then $\mathcal{B}(V, W)$ is a Banach space under the operator norm and the operations inherited from $\mathcal{L}(V, W)$.

While all linear operators preserve the algebraic structure of normed spaces, only certain operators preserve the topological structure as well. As it turns out, these are precisely the bounded operators.

Remark C.10. An operator $T : V \to W$ is bounded if, and only if, it is continuous.

Remark C.11. If T is a bounded operator and $(x_n)_{n=0}^{\infty}$ is a sequence such that $x_n \to x$, then $Tx_n \to Tx$.

Definition C.12. Given normed spaces V and W, an injective continuous linear operator $T : V \to W$ such that T^{-1} is continuous on the range of T is called an *isomorphism.*

Note that this definition of an isomorphism does not match the usual categorical definition, since isomorphisms are typically surjective. This definition more closely resembles a topological *imbedding*, or categorcial *section.* Nonetheless, this definition agrees with "normed space custom." [35, p. 31]

Remark C.13. If T is a linear operator on a normed space V, then for all $x \in V$ we have $\|Tx\| \leq \|T\| \, \|x\|$. In fact, if T is bounded, then $\|T\|$ is the smallest real number such that $\|Tx\| \leq \|T\| \, \|x\|$ for all $x \in V$.

Definition C.14. Given linear operators A and T on a vector space V, the operator AT on V is obtained by composing A with T. That is, $ATx = A(T(x))$ for $x \in V$. We define $T^0 = I$ and for $n \geq 1$ we let $T^n = TT^{n-1}$.

Remark C.15. For linear operators A and T on a normed space V, we have $\|AT\| \leq \|A\|\|T\|$.

The following result is a well-known fact about bounded operators on dense subspaces.

Proposition C.16. *Let Y be a dense subspace of a Banach space X and let $A : Y \to Y$ be a linear operator. If A is bounded, then A extends to an operator $T : X \to X$ such that $\|T\| = \|A\|$ and $T(y) = A(y)$ for all $y \in Y$.*

Appendix D

Polynomials

The study of polynomials and their roots is a fundamental topic in mathematics. In this thesis, we often encounter the idea of *polynomial combinations* of linear operators. Several basic facts about polynomials are required so that these encounters can go smoothly.

Definition D.1. Let $\mathbb{F}$ be a field. A *polynomial* over $\mathbb{F}$ in *variable* t is a function $p : \mathbb{F} \to \mathbb{F}$ of the form $p(t) = \sum_{n=0}^{\infty} a_n t^n$ where each $a_n \in \mathbb{F}$ and $a_n \neq 0$ for only finitely many n. We let $\mathbb{F}[t]$ denote the set of all such polynomials.[1]

Definition D.2. The *degree* of a non-zero polynomial $p(t) = \sum_{n=0}^{\infty} a_n t^n$ over $\mathbb{F}$ is defined to be the maximum index $n \geq 0$ such that $a_n \neq 0$, written $deg(p)$. Polynomials of degree 0 are said to be *constant*. By convention, the degree of the zero polynomial is undefined.

Definition D.3. We say $\lambda \in \mathbb{F}$ is a *root* of p if $p(\lambda) = 0$.

Theorem D.4 (Factor Theorem)**.** *Given a field $\mathbb{F}$, let $p \in \mathbb{F}[t]$. A scalar λ is a root of p if, and only if, p can be written as $p(t) = (t - \lambda)q(t)$ for some polynomial q over $\mathbb{F}$.*

[1] Technically, the set $\mathbb{F}[t]$ is a *ring* under the usual operations of addition and multiplication of polynomials.

Definition D.5. If a polynomial p over $\mathbb{F}$ can be written as $p(t) = r(t)q(t)$ for polynomials p and q, then we say that r *divides* p, written $r \mid p$.

Remark D.6. If $p(t) = q(t)r(t)$ where p, q and r are polynomials over $\mathbb{F}$, then $deg(p) = deg(r) + deg(q)$.

The field of complex numbers has the nifty property of being *algebraically closed*. That is, every polynomial over $\mathbb{C}$ has a complex root. This result is known as the Fundamental Theorem of Algebra.[2]

Theorem D.7 (Fundamental Theorem of Algebra)**.** *Every non-constant polynomial over* $\mathbb{C}$ *has a root in* $\mathbb{C}$*.*

These corollaries follow from the Fundamental Theorem of Algebra, Factor Theorem, and properties of real polynomials.

Corollary D.8. *If* p *is a polynomial over* $\mathbb{C}$*, then there is a polynomial* r *of degree* 1 *so that* $r \mid p$*.*

Corollary D.9. *If* p *is a polynomial over* $\mathbb{R}$*, then there is a polynomial* r *of degree either* 1 *or* 2 *so that* $r \mid p$*.*

Given a polynomial p and an operator $T : X \to X$, it is often useful to define an operator $p(T)$ based on p and T.

Definition D.10. Given a linear operator T on a vector space V over $\mathbb{F}$ and polynomial $p \in \mathbb{F}[t]$, we define an *associated operator* $p(T) = \sum_{n=0}^{\infty} a_n T^n$ using point wise addition and scalar multiplication, where T^n is defined as in Definition C.14.

[2]The Fundamental Theorem of Algebra has an interesting history. Leibniz and Nikolaus Bernoulli both believed that they had found counterexamples in the early 18th century, until Euler proved them wrong. Famous mathematicians including d'Alembert, Euler, de Foncenex, Lagrange, Laplace, Wood, and Gauss thought that he had proofs, but there were a gaps. Finally, a rigorous proof was published by Argand in 1806, settling any controversy.

Remark D.11. Let T be a linear operator on a normed space V over $\mathbb{F}$ and let p be a polynomial over $\mathbb{F}$. If T is bounded, then so is $p(T)$.

Appendix E

Properties of ℓ_1

The counterexample to the invariant subspace problem in Sections 4.2 - 4.9 require some special properties about the Banach space ℓ_1. We include this here.

Let $(f_i)_{i=0}^{\infty}$ be the sequence of standard unit vectors in ℓ_1. Every vector x in ℓ_1 can be represented uniquely as a sum $x = \sum_{i=0}^{\infty} \lambda_i f_i$ for scalars λ_i.[1] For each $m \geq 0$, we let F_m denote the finite-dimensional space $\langle \{f_i : i = 0, 1, \ldots, m\} \rangle$. We let F denote the set of all finitely non-zero sequences. That is, $F = \cup_m F_m$. Clearly F is a dense subspace of ℓ_1. The following gives a useful formula for calculating the norm of an operator on ℓ_1.

Proposition E.1. *If A is a linear operator on ℓ_1 or F, then $\|A\| = \sup\{\|Af_i\|\}_i$.*

Proof. Consider an operator A on ℓ_1 or F. The bound $\|A\| \geq \sup\{\|Af_i\|\}_i$ is clear by definition of the operator norm. For the converse, let $x = (x_0, x_1, \ldots)$ be any vector such that $\|x\| \leq 1$. Then we have the following bound

$$\|Ax\| = \left\| A \left(\sum_{i=0}^{\infty} x_i f_i \right) \right\| \leq \sum_{i=0}^{\infty} |x_i| \|Af_i\| \leq \sum_{i=0}^{\infty} |x_i| \sup\{\|Af_i\|\}$$

[1] This means that $(f_i)_i$ is a *Schauder basis* for ℓ_1. This is actually true for all sequence spaces ℓ_p, $1 \leq p \leq \infty$.

$$= \sup\{\|Af_i\|\} \sum_{i=0}^{\infty} |x_i| \leq \sup\{\|Af_i\|\} \text{ (since } \|x\| \leq 1).$$

The result follows. □

The standard norm on ℓ_1 simply calculates the sum of $|\lambda_i|$ where $x = \sum_{i=0}^{\infty} \lambda_i f_i$. We can also define a 1-norm with respect to a basis $(e_i)_{i=0}^{\infty}$ for F different from $(f_i)_i$.

Definition E.2. Let $(e_i)_{i=0}^{\infty}$ be a basis for F. We may a define norm $\|\cdot\|_e$ on F by

$$\|x\|_e = \sum_{i=0}^{n} |\lambda_i| \text{ where } x = \sum_{i=0}^{n} \lambda_i e_i \text{ for some } n.$$

For our purposes, this norm is referred to as the *e-norm*.

Note that $\|\cdot\|_e$ is defined on F rather than ℓ_1. This is sometimes necessary in order to guarantee convergence. Given such a sequence $(e_i)_{i=0}^{\infty}$ and $m \geq 0$, let $E_m = \langle\{e_0, e_1, \ldots, e_m\}\rangle$. The following lemma is fairly straightforward.

Lemma E.3. *Let $(e_i)_{i=0}^{\infty}$ be a sequence of vectors which span F. Suppose $m \geq 0$ so that $e_0, e_1, \ldots, e_m$ are linearly independent and let $J_m : E_m \to F_m$ be the isomorphism so that $J_m(e_i) = f_i$ for $1 \leq i \leq m$. Then, for each $y \in E_m$ we have*

$$\|y\|_e = \|J_m y\| \leq \|J_m\|\|y\|.$$

Proof. Let us write $y = \sum_{i=0}^{m} \lambda_i e_i$. We have

$$\|y\|_e = \sum_{i=0}^{m} |\lambda_i| = \left\| \sum_{i=0}^{m} \lambda_i f_i \right\| = \|J(y)\|.$$

The fact that $\|J(y)\| \leq \|J\|\|y\|$ is by Remark C.13. □

Bibliography

[1] C. Apostol, C. Foiaş, and D. Voiculescu, *Some results on non-quasitriangular operators. II*, Rev. Roumaine Math. Pures Appl. **18** (1973), 159–181. MR 0333785 (48 #12109a)

[2] ______, *Some results on non-quasitriangular operators. III*, Rev. Roumaine Math. Pures Appl. **18** (1973), 309–324. MR 0333785 (48 #12109a)

[3] ______, *Some results on non-quasitriangular operators. IV*, Rev. Roumaine Math. Pures Appl. **18** (1973), 487–514. MR 0333785 (48 #12109a)

[4] ______, *Some results on non-quasitriangular operators. V*, Rev. Roumaine Math. Pures Appl. **18** (1973), 1133–1149. MR 0333785 (48 #12109a)

[5] C. Apostol, C. Foiaş, and L. Zsidó, *Some results on non-quasitriangular operators*, Indiana Univ. Math. J. **22** (1972/73), 1151–1161. MR 0322537 (48 #899)

[6] S. A. Argyros and R. G. Haydon, *A hereditarily indecomposable L_∞-space that solves the scalar-plus-compact problem*, 2009, arXiv:0903.3921.

[7] N. Aronszajn and K. T. Smith, *Invariant subspaces of completely continuous operators*, Ann. of Math. (2) **60** (1954), 345–350. MR 0065807 (16,488b)

[8] W. B. Arveson and J. Feldman, *A note on invariant subspaces*, Michigan Math. J. **15** (1968), 61–64. MR 0223922 (36 #6969)

[9] S. Axler, *Linear algebra done right*, second ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1997. MR 1482226 (98i:15001)

[10] B. Beauzamy, *Un opérateur sans sous-espace invariant: simplification de l'exemple de P. Enflo*, Integral Equations Operator Theory **8** (1985), no. 3, 314–384. MR 792905 (88b:47011)

[11] ———, *Introduction to operator theory and invariant subspaces*, North-Holland Mathematical Library, vol. 42, North-Holland Publishing Co., Amsterdam, 1988. MR 967989 (90d:47001)

[12] H. Bercovici, *Review: Bernard Beauzamy, Introduction to operator theory and invariant subspaces*, Bull. Amer. Math. Soc. (N.S.) **22** (1990), no. 1, 148–152, Review of [11].

[13] A. R. Bernstein and A. Robinson, *Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos*, Pacific J. Math. **16** (1966), 421–431. MR 0193504 (33 #1724)

[14] S. W. Brown, *Some invariant subspaces for subnormal operators*, Integral Equations Operator Theory **1** (1978), no. 3, 310–333. MR 511974 (80c:47007)

[15] ———, *Hyponormal operators with thick spectra have invariant subspaces*, Ann. of Math. (2) **125** (1987), no. 1, 93–103. MR 873378 (88c:47010)

[16] C. C. Cowen, *An analytic Toeplitz operator that commutes with a compact operator and a related class of Toeplitz operators*, J. Funct. Anal. **36** (1980), no. 2, 169–184. MR 569252 (81d:47020)

[17] A. M. Davie, *Review: On the invariant subspace problem for Banach spaces*, Mathematical Review. American Mathematical Society. Review of [20].

[18] P. Enflo, *A counterexample to the approximation problem in Banach spaces*, Acta Math. **130** (1973), 309–317. MR 0402468 (53 #6288)

[19] ———, *On the invariant subspace problem in Banach spaces*, Séminaire Maurey–Schwartz (1975–1976) Espaces L^p, applications radonifiantes et géométrie des espaces de Banach, Exp. Nos. 14-15, Centre Math., École Polytech., Palaiseau, 1976, p. 7. MR 0473871 (57 #13530)

[20] ———, *On the invariant subspace problem for Banach spaces*, Acta Math. **158** (1987), no. 3-4, 213–313. MR 892591 (88j:47006)

[21] R. Fry, *Real Analysis Notes*, Unpublished lecture notes.

[22] B. Fuglede, *A commutativity theorem for normal operators*, Proc. Nat. Acad. Sci. U. S. A. **36** (1950), 35–40. MR 0032944 (11,371c)

[23] T. A. Gillespie, *An invariant subspace theorem of J. Feldman*, Pacific J. Math. **26** (1968), 67–72. MR 0231232 (37 #6787)

[24] D. W. Hadwin, E. A. Nordgren, H. Radjavi, and P. Rosenthal, *An operator not satisfying Lomonosov's hypothesis*, J. Funct. Anal. **38** (1980), no. 3, 410–415. MR 593088 (81m:47013)

[25] P. R. Halmos, *Normal dilations and extensions of operators*, Summa Brasil. Math. **2** (1950), 125–134. MR 0044036 (13,359b)

[26] ———, *Invariant subspaces of polynomially compact operators*, Pacific J. Math. **16** (1966), 433–437. MR 0193505 (33 #1725)

[27] ______, *Quasitriangular operators*, Acta Sci. Math. (Szeged) **29** (1968), 283–293. MR 0234310 (38 #2627)

[28] ______, *Ten problems in Hilbert space*, Bull. Amer. Math. Soc. **76** (1970), 887–933. MR 0270173 (42 #5066)

[29] ______, *Progress Reports: Invariant Subspaces*, Amer. Math. Monthly **85** (1978), no. 3, 182–183. MR 1538642

[30] ______, *Ten years in Hilbert space*, Integral Equations Operator Theory **2** (1979), no. 4, 529–564. MR 555777 (81c:47003)

[31] D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, IV, Nachr. Kgl. Gesells. Wiss. Göttingen Math.-Phys. Kl. (1906), 157–227.

[32] N. D. Hooker, *Lomonosov's hyperinvariant subspace theorem for real spaces*, Math. Proc. Cambridge Philos. Soc. **89** (1981), no. 1, 129–133. MR 591979 (84a:47009)

[33] B. Johnson (mathoverflow.net/users/2554), *Is the Invariant Subspace Problem interesting?*, MathOverflow, URL: http://mathoverflow.net/questions/48941 (version: 2010-12-10).

[34] V. I. Lomonosov, *Invariant subspaces of the family of operators that commute with a completely continuous operator*, Funkcional. Anal. i Priložen. **7** (1973), no. 3, 55–56. MR 0420305 (54 #8319)

[35] R. E. Megginson, *An introduction to Banach space theory*, Graduate Texts in Mathematics, vol. 183, Springer-Verlag, New York, 1998. MR 1650235 (99k:46002)

[36] P. Meyer-Nieberg, *Quasitriangulierbare Operatoren und Untervektorräume stetiger linearer Operatoren*, Arch. Math. (Basel) **22** (1971), 186–199. MR 0290156 (44 #7341)

[37] A. J. Michaels, *Hilden's simple proof of Lomonosov's invariant subspace theorem*, Adv. Math. **25** (1977), no. 1, 56–58. MR 0500214 (58 #17893)

[38] C. Pearcy and N. Salinas, *An invariant-subspace theorem*, Michigan Math. J. **20** (1973), 21–31. MR 0317075 (47 #5623)

[39] H. Radjavi and P. Rosenthal, *The invariant subspace problem*, Math. Intelligencer **4** (1982), no. 1, 33–37. MR 678734 (84i:47010)

[40] ______, *Invariant subspaces*, second ed., Dover Publications Inc., Mineola, NY, 2003. MR 2003221 (2004e:47010)

[41] C. J. Read, *A solution to the invariant subspace problem*, Bull. London Math. Soc. **16** (1984), no. 4, 337–401. MR 749447 (86f:47005)

[42] ______, *A solution to the invariant subspace problem on the space* l_1, Bull. London Math. Soc. **17** (1985), no. 4, 305–317. MR 806634 (87e:47013)

[43] ______, *A short proof concerning the invariant subspace problem*, J. London Math. Soc. (2) **34** (1986), no. 2, 335–348. MR 856516 (87m:47020)

[44] ______, *The invariant subspace problem for a class of Banach spaces. II. Hypercyclic operators*, Israel J. Math. **63** (1988), no. 1, 1–40. MR 959046 (90b:47013)

[45] ______, *Quasinilpotent operators and the invariant subspace problem*, J. London Math. Soc. (2) **56** (1997), no. 3, 595–606. MR 1610408 (98m:47004)

[46] F. Riesz, *Über lineare Funktionalgleichungen*, Acta Math. **41** (1916), no. 1, 71–98. MR 1555146

[47] J. R. Ringrose, *Super-diagonal forms for compact linear operators*, Proc. London Math. Soc. (3) **12** (1962), 367–384. MR 0136998 (25 #458)

[48] D. Sarason, *Weak-star density of polynomials*, J. Reine Angew. Math. **252** (1972), 1–15. MR 0295088 (45 #4156)

[49] V. S. Sunder, *Paul Halmos—expositor par excellence*, Gaṇita Bhāratī **28** (2006), no. 1-2, 193–199. MR 2468642

[50] V. G. Troitsky, *Lomonosov's theorem cannot be extended to chains of four operators*, Proc. Amer. Math. Soc. **128** (2000), no. 2, 521–525. MR 1641129 (2000c:47016)

[51] B. S. Yadav, *The present state and heritages of the invariant subspace problem*, Milan J. Math. **73** (2005), 289–316. MR 2175046 (2006d:47016)