Reductions and Triangularizations of Sets of Matrices

by

Colin Davidson

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Mathematics
in
Pure Mathematics

Waterloo, Ontario, Canada, 2006

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Families of operators that are triangularizable must necessarily satisfy a number of spectral mapping properties. These necessary conditions are often sufficient as well. This thesis investigates such properties in finite dimensional and infinite dimensional Banach spaces. In addition, we investigate whether approximate spectral mapping conditions (being "close" in some sense) is similarly a sufficient condition.

Acknowledgements

I would like to thank my advisor, Heydar Radjavi, for providing me with excellent guidance, both in choosing a topic and in developing the results. His thoughtful advice has been of great help in completing this thesis. I would also like to thank the many professors and staff here at Waterloo who have helped to make my time here so enjoyable. In particular, I would like to acknowledge Brian Forrest, Laurent Marcoux, and Andu Nica who introduced me to analysis and revealed the wonderful world of mathematics.

To my friends, who had to put up with my ramblings: sorry. You've all been great. And to my family, who have always been there for me, a heartfelt thank you. I wouldn't be where I am today without you.

Contents

Chapter 1

Introduction

Every complex matrix in finite dimensions is similar to a matrix in upper triangular form. This leads to the question of whether two (or more) such matrices can be simultaneously placed in upper triangular form. An even simpler question is whether any two matrices have a common invariant subspace.

For algebras of complex matrices, the question often has an easy answer: Burnside's Theorem says that every proper subalgebra of finite dimensional operators has an invariant subspace. For semigroups, the situation is frequently more complex.

An area that has proved fertile for reducibility results on semigroups is that of partial spectral mapping conditions. Since the spectrum of a matrix in upper triangular form appears on its diagonal, simultaneous upper triangularization leads to several spectral mapping properties. These necessary conditions are often sufficient as well.

The question becomes: how many spectral mapping properties must be assumed before a semigroup becomes triangularizable? Do such results extend to infinite dimensions? Do we need to assume that the spectrum maps exactly, or is it enough that it's "close"?

This thesis will attempt to answer some of these questions.

In Chapter 2, we introduce the notion of simultaneous triangularization and touch on several important classical results, including Burnside's Theorem for algebras and Levitzki's Theorem for semigroups.

In Chapter 3, we investigate several necessary conditions for triangularizability to see if they are, in fact, sufficient conditions. We deal mostly with semigroups and partial spectral mapping conditions. For algebras, it is an easy consequence of Burnside's Theorem that, if $AB - BA$ is nilpotent for every pair $\{A, B\}$ in the algebra, the algebra is triangularizable. Chapter 3 culminates by extending this result to semigroups.

In Chapter 4, we discuss the concept of triangularizability in infinite dimensions. We extend many of our results from finite dimensions to compact operators on a Banach space and, in some cases, to bounded operators. In particular, we show that, if $AB - BA$ is quasinilpotent for a semigroup of compact operators, we have triangularizability.

In Chapter 5, we consider some recent work in the area of triangularizability. We show that positive results can be achieved even when $AB - BA$ is "small", but not necessarily nilpotent.

The majority of the results in this thesis come from *Simultaneous Triangularization* by Heydar Radjavi and Peter Rosenthal [7]. The material in Chapter 5 comes from a paper by Janez Bernik and Heydar Radjavi [1]. Material from other sources is cited where it appears.

Chapter 2

Definitions and Notation

In this thesis, we will be operating in the context of operators on a linear space over the complex numbers. In particular, all linear spans should be assumed to be over $\mathbb{C}$. Many of the results in this thesis extend, with a little caution, to algebraically closed finite fields with certain nonzero characteristics. However, that is beyond the scope of this work.

2.1 Triangularizability in finite dimensions

2.1.1 Definition

In this chapter, as well as Chapter 3 and most of Chapter 5, we will be working in the context of linear operators on finite dimensional normed linear spaces. For such a space $\mathcal{V}$, we denote the entire algebra of such operators by $\mathcal{B}(\mathcal{V})$ (an *algebra* is a family of operators that is closed under addition, multiplication, and scalar multiplication).

Note that if $\dim(\mathcal{V}) = n$ then $\mathcal{B}(\mathcal{V})$ may be identified with $M_n(\mathbb{C})$. We will use both notations throughout this paper, depending on the situation.

We let I be the identity in $\mathcal{B}(\mathcal{V})$. For simplicity, for a scalar λ and an operator A we use notation $A - \lambda$ as a short form of $A - \lambda I$. For an operator A in $M_n(\mathbb{C})$ we let A_{ij} be the entry in the i^{th} row and j^{th} column of A.

We denote the range of A by $\operatorname{ran}(A)$ or $A\mathcal{V}$ and its kernel by $\ker(A)$.

We use'$\subset$' to denote a proper subset. If a subset is not necessarily proper, we use '$\subseteq$'.

2.1.2 Definition

A semigroup is a family of operators that is closed under multiplication, but does not require the presence of a unit or inverses. A *group*, of course, is closed under multiplication and inverses and contains a unit.

2.1.3 Definition

For a semigroup $\mathcal{S}$, we say a subset $\mathcal{J}$ of $\mathcal{S}$ is an *ideal* of $\mathcal{S}$ if for every S in $\mathcal{S}$ and A in $\mathcal{J}$, AS and SA are in $\mathcal{J}$. We will often talk about the *rank* k *ideal* of $\mathcal{S}$. This is the ideal of $\mathcal{S}$ consisting of all elements of rank at most k.

2.1.4 Definition

We say that a subspace $\mathcal{M}$ of $\mathcal{V}$ is *invariant* for a family of operators $\mathcal{F}$ in $\mathcal{B}(\mathcal{V})$ if, for any A in $\mathcal{F}$ and x in $\mathcal{M}$, Ax is in $\mathcal{M}$. We say that $\mathcal{M}$ is *nontrivial* if it is neither $\{0\}$ nor $\mathcal{V}$. If such a nontrivial $\mathcal{M}$ exists for $\mathcal{F}$ we say that $\mathcal{F}$ is *reducible*. Otherwise, we say that $\mathcal{F}$ is *irreducible*.

2.1.5 Definition

For a subspace $\mathcal{M}$ we define its *perpendicular space* $\mathcal{M}^\perp$ to be the set of y in $\mathcal{V}$ such that for any x in $\mathcal{M}$, $\langle x, y\rangle = 0$ where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathcal{V}$.

For an operator A in $\mathcal{B}(\mathcal{V})$ we let its *adjoint* be the operator A^* such that

$$\langle Ax, y\rangle = \langle x, A^* y\rangle$$

for all x and y in $\mathcal{V}$. We say that an operator A is *self-adjoint* if $A = A^*$ and that a family $\mathcal{F}$ is self-adjoint if $\mathcal{F} = \mathcal{F}^*$ where

$$\mathcal{F}^* = \{A^* : A \in \mathcal{F}\}.$$

If $\mathcal{M}$ is invariant for a family $\mathcal{F}$, then for A in $\mathcal{F}$, x in $\mathcal{M}$ and y in $\mathcal{M}^{\perp}$,

$$\langle x, A^* y\rangle = \langle Ax, y\rangle = 0$$

so $\mathcal{M}^{\perp}$ is invariant for $\mathcal{F}^*$. The other direction also holds so $\mathcal{F}$ is irreducible if and only if $\mathcal{F}^*$ is.

2.1.6 Definition

We say that a family of operators $\mathcal{F}$ in $\mathcal{B}(\mathcal{V})$ is *triangularizable* if there is a basis for $\mathcal{V}$ relative to which every member of $\mathcal{F}$ is an upper triangular matrix. This is equivalent to the existence of a chain of invariant subspaces for $\mathcal{F}$

$$\{0\} = \mathcal{M}_0 \subset \mathcal{M}_1 \subset \cdots \subset \mathcal{M}_n = \mathcal{V},$$

where $\dim(\mathcal{M}_j) = j$ for each j. In fact, we can choose $\mathcal{M}_j = \operatorname{span}\{e_1, \ldots, e_j\}$ where $\{e_1, \ldots, e_n\}$ is a basis relative to which $\mathcal{F}$ is upper triangular. We call such a chain a *triangularizing chain* for $\mathcal{F}$.

Note that every collection of operators on a one dimensional space is triangularizable, however such a collection can never be reducible.

2.2 Triangularization and Irreducibility

Our first result, the Triangularization Lemma, is incredibly useful. It allows us to prove certain semigroups are triangularizable by showing reducibility, along with a set of inheritable properties. First we need to introduce quotient spaces.

2.2.1 Definition

For subspaces $\mathcal{N} \subset \mathcal{M}$ of $\mathcal{V}$ the quotient space $\mathcal{M}/\mathcal{N}$ is

$$\mathcal{M}/\mathcal{N} = \{[x] : x \in \mathcal{M}\},$$

where $[x] = x + \mathcal{N} = \{x + z : z \in \mathcal{N}\}$. For x and y in $\mathcal{M}$ and λ in $\mathbb{C}$, we define $[x] + [y] = [x + y]$ and $\lambda[x] = [\lambda x]$.

For an operator A in $\mathcal{B}(\mathcal{V})$ with invariant subspaces $\mathcal{N} \subset \mathcal{M}$, we define the *quotient operator* $\tilde{A}$ on $\mathcal{M}/\mathcal{N}$ by $\tilde{A}[x] = [Ax]$. Since $\mathcal{M}$ is invariant for A, Ax is in $\mathcal{M}$. Also, since $\mathcal{N}$ is invariant for A, if $[x] = [y]$ then $[Ax] = [Ay]$ so $\tilde{A}$ is well-defined.

In particular, the restriction of A to an invariant subspace $\mathcal{M}$, denoted by $A|_{\mathcal{M}}$, is the quotient operator for $\mathcal{M}$ and $\mathcal{N} = \{0\}$.

If $\mathcal{F}$ is a family of operators in $\mathcal{B}(\mathcal{V})$ with invariant subspaces $\mathcal{N} \subset \mathcal{M}$ then the quotient of $\mathcal{F}$ with respect to $\mathcal{M}$ and $\mathcal{N}$ is the family of quotients $\tilde{A}$ with respect to $\mathcal{M}/\mathcal{N}$ where A is in $\mathcal{F}$.

2.2.2 Definition

If $\mathcal{P}$ is a property of operators, we say it is *inherited by quotients* if, for every family of operators $\mathcal{F}$ in $\mathcal{B}(\mathcal{V})$ that satisfies $\mathcal{P}$, if $\mathcal{N} \subset \mathcal{M}$ are invariant subspaces for $\mathcal{F}$ then the quotient of $\mathcal{F}$ with respect to $\mathcal{M}/\mathcal{N}$ also satisfies $\mathcal{P}$.

We can now state and prove the Triangularization Lemma.

2.2.3 Lemma (Triangularization Lemma)

Let $\mathcal{P}$ be a set of properties, each of which is inherited by quotients. If every family of operators in $\mathcal{B}(\mathcal{V})$ with $\dim(\mathcal{V}) > 1$ that satisfies $\mathcal{P}$ is reducible, then every collection of transformations satisfying $\mathcal{P}$ is triangularizable.

Proof. Let $\mathcal{F}$ be a family of operators in $\mathcal{B}(\mathcal{V})$ that satisfies $\mathcal{P}$. Take a maximal chain of invariant subspaces for $\mathcal{F}$,

$$\{0\} = \mathcal{M}_0 \subset \mathcal{M}_1 \subset \cdots \subset \mathcal{M}_n = \mathcal{V}.$$

Denote this chain by $\mathcal{C}$ and assume that $\mathcal{C}$ is not a triangularizing chain. Then there must be a k such that $\mathcal{M}_k/\mathcal{M}_{k-1}$ has dimension at least 2. Then $\mathcal{F}|_{\mathcal{M}_k/\mathcal{M}_{k-1}}$ has property $\mathcal{P}$ and is reducible by the hypothesis. Therefore it has an invariant subspace $\mathcal{L}$.

Define $\mathcal{M} = \{x \in \mathcal{M}_k : [x] \in \mathcal{L}\}$. Since $\mathcal{L}$ is a proper subspace of $\mathcal{M}_k/\mathcal{M}_{k-1}$, $\mathcal{M}$ is properly between $\mathcal{M}_{k-1}$ and $\mathcal{M}_k$. Since $\mathcal{M}$ is an invariant subspace of a quotient by invariant subspaces of $\mathcal{F}$, it's an invariant subspace of $\mathcal{F}$. This contradicts the maximality of $\mathcal{C}$ so $\mathcal{C}$ must be a triangularizing chain.

$\square$

A simple and useful result of the above lemma is the following theorem.

2.2.4 Theorem

Every commutative family of operators in $\mathcal{B}(\mathcal{V})$ is triangularizable.

Proof. Let $\mathcal{F}$ be a commutative family of operators in $\mathcal{B}(\mathcal{V})$.

Since commutativity is a property inherited by quotients, we need only show $\mathcal{F}$ is reducible by the Triangularization Lemma (2.2.3).

If every element of $\mathcal{F}$ is a scalar then every subspace is invariant for $\mathcal{F}$ so it's triangularizable.

Otherwise, take a nonscalar A in $\mathcal{F}$. Let λ be an eigenvalue for A and let $\mathcal{M}$ be the corresponding eigenspace. Since A isn't scalar, $\mathcal{M}$ is nontrivial. For any B in $\mathcal{F}$ and x in $\mathcal{M}$,

$$ABx = BAx = B(\lambda x) = \lambda Bx.$$

Thus Bx is in $\mathcal{M}$ and $\mathcal{M}$ is an invariant subspace for $\mathcal{F}$. Therefore $\mathcal{F}$ is reducible and triangularizability follows.

$\square$

The following well-known result is an easy corollary.

2.2.5 Corollary (Schur's Theorem)

Every operator A in $\mathcal{B}(\mathcal{V})$ is triangularizable

Proof. The family $\{A\}$ is commutative so it's triangularizable by Theorem 2.2.4.

$\square$

When determining which families are triangularizable it helps to know which families are definitely not. Burnside's Theorem shows that, when we consider algebras, the question has a easy answer. To prove this, we'll use the following definition and result.

2.2.6 Definition

A family of operators $\mathcal{F}$ in $\mathcal{B}(\mathcal{V})$ is *transitive* if for every $x \neq 0$ and every y in $\mathcal{V}$, there is an F in $\mathcal{F}$ such that $Fx = y$.

2.2.7 Lemma

Let $\mathcal{A}$ be an algebra of operators in $\mathcal{B}(\mathcal{V})$ with $\dim(\mathcal{V}) \geq 2$. Then $\mathcal{A}$ is irreducible if and only if $\mathcal{A}$ is transitive.

Proof. Assume $\mathcal{A}$ is irreducible and take $x \neq 0$ from $\mathcal{V}$. Now, $\mathcal{A}x$ is an invariant subspace for $\mathcal{A}$ and $\mathcal{A}$ is irreducible, so $\mathcal{A}x$ is either $\{0\}$ or $\mathcal{V}$. But if $\mathcal{A}x = \{0\}$ then $\operatorname{span}\{x\}$ is a nontrivial invariant subspace. Therefore, $\mathcal{A}x = \mathcal{V}$ and there is an A in $\mathcal{A}$ such that $Ax = y$ so $\mathcal{A}$ is transitive.

Assume $\mathcal{A}$ is transitive and let $\mathcal{M} \neq \{0\}$ be an invariant subspace of $\mathcal{A}$. Take $x \neq 0$ in $\mathcal{M}$. Then, for every y in $\mathcal{B}(\mathcal{V})$, there is an A in $\mathcal{A}$ such that $Ax = y$. Since x is in $\mathcal{M}$, so is y so $\mathcal{M} = \mathcal{V}$. Therefore $\mathcal{A}$ is irreducible.

$\square$

2.2.8 Theorem (Burnside's Theorem)

If $\dim(\mathcal{V})$ is at least 2 then the only irreducible algebra of operators in $\mathcal{B}(\mathcal{V})$ is $\mathcal{B}(\mathcal{V})$. In other words, every proper subalgebra of $\mathcal{B}(\mathcal{V})$ is reducible.

Proof. The operators

$$\begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$$

are easily seen to have no common nontrivial invariant subspaces. Thus, $\mathcal{B}(\mathcal{V})$ is irreducible.

Let $\mathcal{A}$ be an irreducible algebra of operators in $\mathcal{B}(\mathcal{V})$. We know that the rank one operators span $\mathcal{B}(\mathcal{V})$ so we want to show that $\mathcal{A}$ contains them all.

First, we'll show that $\mathcal{A}$ contains a nonzero singular element K. Take any element A in $\mathcal{A}$. Let λ be an eigenvalue of A. Then $A-\lambda$ is singular and $A^2-\lambda A = A(A-\lambda)$ is also singular and is in $\mathcal{A}$. So, either $\mathcal{A}$ contains a nonzero singular element or every A in $\mathcal{A}$ is invertible with $A^2-\lambda A = 0$. But, A invertible and $A^2-\lambda A = 0$ implies $A = \lambda I$. Then $\mathcal{A}$ consists of scalars so $\mathcal{A}$ is triangularizable. However, $\mathcal{A}$ is irreducible and so it must contain a nonzero singular element, K.

We next show that $\mathcal{A}$ contains a rank one idempotent. We'll do this by induction on the dimension of $\mathcal{V}$. If $\dim(\mathcal{V}) = 2$ then K must have rank one since it's nonzero and singular. Triangularize K and let $\{e_1, e_2\}$ be the corresponding basis. Since it's rank one, we can assume it has the form

$$K = \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix}.$$

If $\alpha \neq 0$ then $\alpha^{-1}K$ is in $\mathcal{A}$ and is our desired idempotent. Otherwise, $\alpha = 0, \beta \neq 0$ so $\beta^{-1}K = E_{12}$ is in $\mathcal{A}$. Since $\mathcal{A}$ is transitive, there is an A in $\mathcal{A}$ such that $Ae_1 = e_2$. Then, $KAe_1 = Ke_2 = e_1$ so $\sigma(KA)$ contains 1, KA has rank one, and KA is in $\mathcal{A}$. This takes us to the case with $\alpha \neq 0$. Therefore, if $\dim(\mathcal{V}) = 2$, $\mathcal{A}$ has a rank one idempotent.

Let $\dim(\mathcal{V}) = n$ and assume all irreducible algebras of operators on spaces of dimension at most $n-1$ have a rank one idempotent. We have that $K\mathcal{A}$ is also an algebra and that $\mathcal{M} = K\mathcal{V}$ is invariant for $K\mathcal{A}$. In fact, $\mathrm{ran}(KA)$ is contained within $\mathrm{ran}(K)$ for every A in $\mathcal{A}$. Let $\mathcal{B} = K\mathcal{A}|_{K\mathcal{V}}$. We claim $\mathcal{B}$ is irreducible. By Lemma 2.2.7, we can check for transitivity instead. Take any $x \neq 0$ and y in $\mathcal{M}$. We have $y = Ky_0$ for some y_0 in $\mathcal{V}$. Since $\mathcal{A}$ is transitive, there is an A in $\mathcal{A}$ such that $Ax = y_0$. Then $KAx = Ky_0 = y$ so $\mathcal{B}$ is a transitive and irreducible subalgebra of $\mathcal{B}(K\mathcal{V})$.

Now, K is singular so $K\mathcal{V}$ has dimension less than $\mathcal{V}$. By induction, $\mathcal{B}$ contains a rank one idempotent, E. By construction, $E = KA|_{K\mathcal{V}}$ for some A in $\mathcal{A}$ and $\mathrm{ran}(KA)$ is contained in $\mathrm{ran}(K)$ so

$$F = KA = \begin{pmatrix} E & X \\ 0 & 0 \end{pmatrix}$$

is in $\mathcal{A}$. Since $E = E^2$,

$$F^2 = \begin{pmatrix} E & EX \\ 0 & 0 \end{pmatrix} \text{ and } F^4 = \begin{pmatrix} E & EX \\ 0 & 0 \end{pmatrix}.$$

Also, EX has rank at most one. Therefore F^2 is our rank one idempotent in $\mathcal{A}$ and the result follows by induction.

Take P to be our rank one idempotent in $\mathcal{A}$ and put it in Jordan normal form so $P = E_{11}$ with regards to the corresponding basis $\{e_1, \ldots, e_n\}$.

By transitivity, $\mathcal{A}e_1 = \mathcal{V}$. For any A in $\mathcal{A}$, $AP = \begin{pmatrix} Ae_1 & 0 \end{pmatrix}$ so every rank one operator with its last $n-1$ columns zero is in $\mathcal{A}$.

Now, $PA = \begin{pmatrix} e_1^* A \\ 0 \end{pmatrix}$. Assume $\mathcal{M} = \{(e_1^* A)^* : A \in \mathcal{A}\} \neq \mathcal{V}$. $\mathcal{A}$ is an algebra so $\mathcal{M}$ is a subspace and there there is a $x \neq 0$ in $\mathcal{V}$ such that x is in $\mathcal{M}^{\perp}$. Then $PAx = 0$ so $Ax \neq e_1$ for any A in $\mathcal{A}$. But that contradicts that $\mathcal{A}$ is irreducible and thus transitive. Therefore $\mathcal{M} = \mathcal{V}$ and $\mathcal{A}$ contains every matrix with its last $n-1$ rows zero.

But an arbitrary rank one operator in $\mathcal{B}(\mathcal{V})$ is simply xy^* with x and y in $\mathcal{V}$. And

$$x^* y = \begin{pmatrix} x & 0 \end{pmatrix} \begin{pmatrix} y^* \\ 0 \end{pmatrix},$$

so $\mathcal{A}$ contains all rank one operators. Therefore $\mathcal{A} = \mathcal{B}(\mathcal{V})$.

$\square$

The algebra $\mathcal{A}$ generated by a semigroup $\mathcal{S}$ is simply the linear span of elements of $\mathcal{S}$. Using this fact and Burnside's Theorem (2.2.8), we derive a number of sufficient conditions for reducibility of algebras and semigroups of operators in $\mathcal{B}(\mathcal{V})$. Our first result deals only with algebras, but will eventually be extended to semigroups (Theorem 3.4.15).

2.2.9 Lemma

If $\mathcal{A}$ is an algebra of operators in $\mathcal{B}(\mathcal{V})$ then $\mathcal{A}$ is triangularizable if and only if $AB - BA$ is nilpotent for every A and B in $\mathcal{A}$.

Proof. If $\mathcal{A}$ is triangularizable, then the diagonals of its operators commute so $AB - BA$ is nilpotent. This is seen in more detail in the Spectral Mapping Theorem (2.4.4).

For the converse, note that nilpotent commutators are inherited by quotients so we need only show reducibility by the Triangularization Lemma (2.2.3). Note that the operators

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

have nonnilpotent commutators as

$$AB - BA = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad (AB - BA)^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

so on spaces of dimension at least 2, not every pair of operators has nilpotent commutators (on spaces of dimension three or more simply add a direct summand of zero to extend A and B). Therefore $\mathcal{A} \neq \mathcal{B}(\mathcal{V})$ so by Burnside's Theorem (2.2.8), $\mathcal{A}$ is reducible and is therefore triangularizable.

□

2.2.10 Lemma

Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{B}(\mathcal{V})$ and let ϕ be a linear functional on $\mathcal{B}(\mathcal{V})$. If ϕ is nonzero, but $\phi|_{\mathcal{S}} = 0$ then $\mathcal{S}$ is reducible.

Proof. Let $\mathcal{A}$ be the algebra generated by $\mathcal{S}$. $\mathcal{A}$ consists of linear combinations of members of $\mathcal{S}$ so $\phi|_{\mathcal{A}} = 0$.

Assume $\mathcal{S}$ is irreducible. Then $\mathcal{A}$ is irreducible and, by Burnside's Theorem (2.2.8), $\mathcal{A} = \mathcal{B}(\mathcal{V})$. Then $\phi = 0$ which is a contradiction so $\mathcal{S}$ is reducible.

□

2.2.11 Theorem (Levitzki's Theorem)

Every semigroup of nilpotent operators in $\mathcal{B}(\mathcal{V})$ is triangularizable.

Proof. Let $\mathcal{S}$ be such a semigroup. Since nilpotence is a property inherited by quotients, it sufficies by the Triangularization Lemma (2.2.3) to show $\mathcal{S}$ is reducible.

For any element A in $\mathcal{S}$, $\operatorname{tr}(A) = 0$ since the only eigenvalues of a nilpotent operator are zero. Therefore tr is a nonzero functional on $\mathcal{B}(\mathcal{V})$ that is zero on $\mathcal{S}$ so $\mathcal{S}$ is reducible by Lemma 2.2.10.

□

2.2.12 Lemma

If a semigroup of operators $\mathcal{S}$ in $\mathcal{B}(\mathcal{V})$ has a nonzero reducible ideal then $\mathcal{S}$ is reducible. In other words, a nonzero ideal of an irreducible semigroup of operators in $\mathcal{B}(\mathcal{V})$ is irreducible.

Proof. Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{B}(\mathcal{V})$. Let $\mathcal{J} \neq 0$ be an ideal of $\mathcal{S}$ and assume $\mathcal{J}$ is reducible. Let $\mathcal{M}$ be a nontrivial invariant subspace for $\mathcal{J}$.

Let $\mathcal{M}_1 = \operatorname{span}\{J\mathcal{M} : J \in \mathcal{J}\}$ and $\mathcal{M}_2 = \cap\{\ker(J) : J \in \mathcal{J}\}$. For S in $\mathcal{S}$, J in $\mathcal{J}$ and x in $\mathcal{M}$, SJ is in $\mathcal{J}$ and SJx is in $\mathcal{M}_1$ so $\mathcal{M}_1$ is invariant for $\mathcal{S}$. If x is in $\mathcal{M}_2$ then JS is in $\mathcal{J}$ so $J(Sx) = (JS)x = 0$ and $\mathcal{M}_2$ is invariant for $\mathcal{S}$. We need only show one of them is nontrivial.

Since $\mathcal{M}$ is invariant for $\mathcal{J}$, $\mathcal{M}_1$ is contained in $\mathcal{M}$. Therefore $\mathcal{M}_1 \neq \mathcal{V}$ and thus, if it's a trivial invariant subspace, $\mathcal{M}_1 = \{0\}$. In this case, $\mathcal{J}\mathcal{M} = \{0\}$, so $\mathcal{M}_2$ contains $\mathcal{M}$ and is not $\{0\}$. However, $\mathcal{J} \neq 0$ so $\mathcal{M}_2 \neq \mathcal{V}$. Therefore $\mathcal{M}_2$ is nontrivial.

Therefore $\mathcal{S}$ has a nontrivial invariant subspace, so $\mathcal{S}$ is reducible.

$\square$

2.2.13 Lemma

Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{B}(\mathcal{V})$ and E an idempotent of rank at least 2, not necessarily in $\mathcal{S}$. If the collection $\mathcal{S}_0 = E\mathcal{S}E|_{E\mathcal{V}}$ is reducible then so is $\mathcal{S}$.

Proof. Let $\mathcal{M}$ be a nontrivial invariant subspace for $\mathcal{S}_0$. Take $x \neq 0$ in $\mathcal{M}$. Since $\mathcal{M} \subset E\mathcal{V}$, $Ex = x$. Let f be a nonzero linear functional on $E\mathcal{V}$ with $f(\mathcal{M}) = 0$ (which exists as $\mathcal{M}$ is a proper subspace of $E\mathcal{V}$). Define a functional ϕ on $\mathcal{B}(\mathcal{V})$ by $\phi(T) = f(ETEx)$ for all T in $\mathcal{B}(\mathcal{V})$.

Now, for S in $\mathcal{S}$,

$$\phi(S) = f(ESEx) = 0,$$

since $ESEx$ is in $\mathcal{M}$. However, f is nontrivial on $E\mathcal{V}$ so there is a y in $\mathcal{V}$ such that $f(Ey) \neq 0$. As $Ex = x \neq 0$, there is a T in $\mathcal{B}(\mathcal{V})$ such that $TEx = y$. Then

$$\phi(T) = f(ETEx) = f(Ey) \neq 0,$$

so ϕ is nontrivial. Therefore $\mathcal{S}$ is reducible by Lemma 2.2.10.

$\square$

2.2.14 Lemma

Let $\mathcal{S}$ be an irreducible semigroup of operators in $\mathcal{B}(\mathcal{V})$ and let

$$m = \min\{\text{rank}(S) : 0 \neq S \in \mathcal{S}\}.$$

Then there exists an element of the form $S_0 \oplus 0$ in $\mathcal{S}$, with respect to a suitable basis, where S_0 is invertible and has rank m.

Proof. Let A be in $\mathcal{S}$ with rank m. If $m = \dim(\mathcal{V})$ then A is invertible and we're done. Otherwise, if $A\mathcal{S} = \{0\}$ then $\ker(A)$ is an invariant subspace for $\mathcal{S}$ and A is neither 0 nor invertible so $\ker(A)$ is a nontrivial subspace. But $\mathcal{S}$ is irreducible so this can't happen and hence $A\mathcal{S} \neq \{0\}$.

Take $B \neq 0$ in $A\mathcal{S}$. Then B has rank m. If $\mathcal{S}B = \{0\}$ then $\text{ran}(B)$ is a nontrivial invariant subspace for $\mathcal{S}$ which can't happen. Therefore $\mathcal{J} = \mathcal{S}A\mathcal{S} \neq \{0\}$ and is a nonzero ideal of $\mathcal{S}$. Therefore $\mathcal{J}$ is irreducible by Lemma 2.2.12.

By Theorem 2.2.11, $\mathcal{J}$ must have non-nilpotent elements. Take such an element, B. By the minimality of m, B^k has rank m for all k in $\mathbb{N}$. Therefore, putting B in Jordan form gives us the desired operator.

□

The last result of this section is a boundedness result that follows from Levitzki's Theorem (2.2.11).

2.2.15 Lemma

An irreducible semigroup $\mathcal{S}$ of operators in $\mathcal{B}(\mathcal{V})$ is bounded if and only if the spectral radius is bounded on $\mathcal{S}$ (We denote the spectral radius by ρ as defined in Definition 2.4.2).

Proof. Since $\rho(A) \leq \|A\|$, boundedness clearly implies bounded spectral radius. For the converse, assume that $\mathcal{S}$ is not bounded. Since ρ is continuous, we can assume $\mathcal{S}$ is closed. We can also assume $\mathcal{S} = \gamma\mathcal{S}$ where $\gamma \in [0, 1]$ as this does not affect the boundedness of the spectral radius.

Let $\{S_n\}$ be a sequence in $\mathcal{S}$ with $\lim\limits_{n \longrightarrow \infty} \|S_n\| = \infty$. Then $\{S_n / \|S_n\|\}$ is a bounded sequence, so we can restrict it to a subsequence converging to some S in $\mathcal{S}$. Then

$$\rho(S) = \lim_{n \longrightarrow \infty} \frac{\rho(S_n)}{\|S_n\|} = 0.$$

Also, for any T in $\mathcal{S}$, $\rho(S_n T)$ is bounded so

$$\rho(ST) = \lim_{n \longrightarrow \infty} \frac{\rho(S_n T)}{\|S_n\|} = 0.$$

Thus the ideal $\mathcal{J}$ generated by S consists of nilpotents. By Levitzki's Theorem (2.2.11), $\mathcal{J}$ is reducible. By Lemma 2.2.12, $\mathcal{S}$ is reducible. This is a contradiction, so $\mathcal{S}$ must be bounded.

□

2.3 Reduction to Groups

These results help to reduce questions about semigroups to questions about unitary groups.

2.3.1 Theorem

Every bounded group $\mathcal{G}$ of operators in $\mathcal{B}(\mathcal{V})$ is simultaneously similar to a group of unitary operators.

Proof. We can assume $\mathcal{G} = \overline{\mathcal{G}}$, so $\mathcal{G}$ is compact. Let μ be the Haar measure on $\mathcal{G}$ (The existence of the Haar measure and its properties is discussed in many texts, including [9, p. 128]). That is, μ is a positive regular Borel measure on $\mathcal{G}$ with $\mu(\mathcal{G}) = 1$ and

$$\int_{\mathcal{G}} f(GG_0)d\mu(G) = \int_{\mathcal{G}} f(G)d\mu(G)$$

for all G_0 in $\mathcal{G}$ and measurable f.

Let $(\cdot, \cdot)$ be the standard inner product on $\mathcal{V}$ and define a new inner product by

$$\langle x, y \rangle = \int_{\mathcal{G}} (Gx, Gy)d\mu(G)$$

for all x and y in $\mathcal{V}$. Then linearity, sesquilinearity, and conjugate symmetry follow from the same properties of the inner product and the linearity of the integral. If

$x \neq 0$ then $(Gx, Gx) > 0$, so it's a positive function on the entire set $\mathcal{G}$ of nonzero measure. Thus $\langle x, x\rangle > 0$.

Take any G_0 in $\mathcal{G}$. Then

$$\langle G_0 x, G_0 y\rangle = \int_{\mathcal{G}} (GG_0 x, GG_0 y) d\mu(G) = \int_{\mathcal{G}} (Gx, Gy) d\mu(G) = \langle x, y\rangle\,,$$

so G_0 is unitary with respect to the new inner product. Therefore $\mathcal{G}$ is similar to a unitary group.

□

The following lemma will be of use in proving the final result of this section.

2.3.2 Lemma

Let U be a unitary operator in $\mathcal{B}(\mathcal{V})$. Then $\{U^n : n \in \mathbb{N}\}$ has subsequences coverging to I and U^{-1} respectively.

Proof. Since U is unitary, we have that

$$\|U\| = \sup_{x\in\mathcal{V}_1} \langle Ux, Ux\rangle = \sup_{x\in\mathcal{V}_1} \langle x, U^* Ux\rangle = \sup_{x\in\mathcal{V}_1} \|x\| = 1,$$

where $\mathcal{V}_1$ is the unit ball of $\mathcal{V}$. Also, for a unitary operator U and any operator T in $\mathcal{B}(\mathcal{V})$ we have that

$$\|UT\| = \sup_{x\in\mathcal{V}_1} \langle UTx, UTx\rangle = \sup_{x\in\mathcal{V}_1} \langle Tx, U^* UTx\rangle = \sup_{x\in\mathcal{V}_1} \|Tx\| = \|T\|\,.$$

Thus the sequence $(u_n)_{n=1}^{\infty}$ defined by $u_n = U^n$ is contained within the unit ball of $\mathcal{B}(\mathcal{V})$. Since the unit ball of $\mathcal{B}(\mathcal{V})$ is compact, the sequence must have a convergent subsequence $(u_m)_{m=1}^{\infty}$ where $u_m = U^{n_m}$ and if $m_1 < m_2$ then $n_{m_1} < n_{m_2}$.

Since this subsequence is convergent it must be Cauchy. Fix $\varepsilon > 0$. Then there is an $M > 0$ such that if $i, j \geq M$ then $\|u_i - u_j\| < \varepsilon$. In particular, $\|u_{M+1} - u_M\| < \varepsilon$. We also have that

$$\begin{aligned}\|u_{M+1} - u_M\| &= \|U^{n_{M+1}} - U^{n_M}\| \\ &= \left\|U^{n_M}(U^{n_{M+1}-n_M} - I)\right\| \\ &= \left\|U^{n_{M+1}-n_M} - I\right\| \\ &< \varepsilon,\end{aligned}$$

where the last equality follows as U is unitary.

Thus I is in the closure of $\{U, U^2, \ldots, U^n, \ldots\}$ since for any $\varepsilon > 0$ we can find an element of the set within ε of I. And using these elements we find for ε values of $1, \frac{1}{2}, \frac{1}{3}, \ldots$ we can build a subsequence converging to the identity.

Finding a subsequence convergent to U^{-1} follows immediately by multiplying every term in the sequence by U^{-1} (If the first term was U (and is now I) we drop it) and using the fact that $\|UT\| = \|T\|$.

$\square$

2.3.3 Lemma

Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{B}(\mathcal{V})$ satisfying $\mathcal{S} = \overline{\mathbb{R}^+\mathcal{S}}$, where $\mathbb{R}^+$ is the set of positive real numbers. Let m be the minimal rank of nonzero members of $\mathcal{S}$.

(i) If E is an idempotent in $\mathcal{S}$ of rank m, then the restriction of $E\mathcal{S}E\backslash\{0\}$ to $E\mathcal{V}$ is a group $\mathcal{G}$

(ii) Up to a simultaneous similarity, each such group $\mathcal{G}$ is contained in $\mathbb{R}^+\mathcal{U}$ where $\mathcal{U}$ is the group of unitaries in $\mathcal{B}(\mathcal{V})$.

(iii) If $\mathcal{S}$ is irreducible, then it contains idempotents of rank m, and, for each such idempotent, the corresponding group $\mathcal{G}$ is irreducible.

Proof. (i) Since $E\mathcal{S}E = \overline{\mathbb{R}^+E\mathcal{S}E}$ we can assume that $E = I$, $E\mathcal{S}E = \mathcal{S}$, and that $m = \dim(\mathcal{V})$.

By the minimality of m, every element in $\mathcal{S}$ is either 0 or invertible. Let S be a nonzero element in $\mathcal{S}$. First, we want to show that S is a scalar multiple of a unitary. We know $\mathbb{R}^+\mathcal{S} = \mathcal{S}$, so we'll assume $\rho(S) = 1$.

We can express S as

$$S = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix},$$

where $\sigma(B)$ is on the unit circle and $\rho(C) < 1$. Further, we can assume B is in Jordan form so $B = U + N$ where U is unitary, N is nilpotent, and $NU = UN$. Since

$$\rho(C) = \lim_{n \longrightarrow \infty} \|C^n\|^{\frac{1}{n}}$$

and $\rho(C) < 1$, we know that $\lim_{n \longrightarrow \infty} \|C^n\| = 0$.

We want to show that $N = 0$ and C acts on a zero dimensional space as then $S = U$. Take $k \geq 0$ such that $N^k \neq 0$ and $N^{k+1} = 0$. If $k = 0$ then $N = 0$. Otherwise, for any $n \geq k$,

$$(U+N)^n = U^n + \binom{n}{1} U^{n-1} N + \cdots + \binom{n}{k} U^{n-k} N^k$$

and, since U is unitary, Lemma 2.3.2 gives us a sequence of powers of U converging to I. In particular, take $\{n_j\}$ such that

$$\lim_{j \longrightarrow \infty} U^{n_j - k} = I.$$

For large enough n, the $\binom{n}{k}$ is the dominant coefficient of the expansion of $(U+N)^n$. Therefore

$$\lim_{j \longrightarrow \infty} \frac{(U+N)^{n_j}}{\binom{n_j}{k}} = \lim_{j \longrightarrow \infty} (U^{n_j - k} N^k) = N^k.$$

Then, as $\lim_{n \longrightarrow \infty} \|C^n\| = 0$,

$$\lim_{j \longrightarrow \infty} \frac{S^{n_j}}{\binom{n_j}{k}} = \begin{pmatrix} N^k & 0 \\ 0 & 0 \end{pmatrix},$$

which is then an element of $\mathcal{S}$ as $\mathcal{S}$ is closed. Howevever, $N^k \neq 0$, but it has rank less than m since it's not invertible. This contradicts the minimality of m so $k = 0$ and $N = 0$.

So $B = U$ and therefore B^{n_j} converges to I. Therefore

$$\lim_{j \longrightarrow \infty} S^{n_j} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix},$$

which will be in $\mathcal{S}$. But, if C acts on a space of positive dimension, then this element will have rank less than m and be nonzero. This would contradict the minimality of m so C must act on a zero dimensional space. Therefore S is a multiple of a unitary, U.

By Lemma 2.3.2, there is also a sequence of powers of U that converges to U^{-1}. Therefore $\mathcal{S}$ contains all the inverses of its nonzero elements so $\mathcal{G} = \mathcal{S} \backslash \{0\}$ is a group.

(ii) From proving (i), we know that $G/\rho(G)$ is similar to a unitary matrix for every G in $\mathcal{G}$, but we need to show a simultaneous similarity. Let $\mathcal{G}_0 = \{G/\rho(G) : G \in \mathcal{G}\}$.

For every G in $\mathcal{G}$, G is a multiple of a matrix similar to a unitary so the eigenvalues of G have constant modulus. Therefore $\rho(G)^m = |\det(G)|$ and

$$\rho(G_1)\rho(G_2) = |\det(G_1)\det(G_2)|^{1/n} = |\det(G_1G_2)|^{1/n} = \rho(G_1G_2),$$

so $\mathcal{G}_0$ is a group.

As ρ is continuous, $\mathcal{G}_0$ is closed. If $\mathcal{G}_0$ is bounded then it's simultaneously similar to a unitary group by Lemma 2.3.1. Assume otherwise and take $\{G_n\}$ in $\mathcal{G}_0$ with $\lim_{n\longrightarrow\infty} \|G_n\| = \infty$. Then the sequence $\{G_n/\|G_n\|\}$ is a bounded sequence in a compact space so it has a subsequence converging to some A in $\mathcal{G}_0$. Now $\rho(A) = 0$ since $\rho(G_n) = 1$ for all n. But this means that A is nilpotent and thus has rank less than m. However, $\|A\| = 1$ so $A \neq 0$ which contradicts the minimality of m. Therefore $\mathcal{G}_0$ must be bounded and is simultaneously similar to a unitary group.

(iii) By Lemma 2.2.14, there is an element $A_0 \oplus 0$ in $\mathcal{S}$ of rank m and A_0 invertible. Since $\mathcal{S} = \mathbb{R}^+\mathcal{S}$ we can assume $\rho(A_0) = 1$. From the proof of (i), we see that A_0 is similar to a unitary matrix. By Lemma 2.3.2, we have a sequence of powers of A_0 converging to I. Therefore $E = I \oplus 0$ is in $\mathcal{S}$ and is an idempotent of rank m. And $E\mathcal{S}E$ is irreducible by Lemma 2.2.13.

$\square$

2.4 The Spectrum in Finite Dimensions

The spectrum of an operator A will play a large role in many of our results.

2.4.1 Definition

The *spectrum* of an operator A in $\mathcal{B}(\mathcal{V})$ is the set

$$\{\lambda \in \mathbb{C} : A - \lambda \text{ is not invertible}\}.$$

We use $\sigma(A)$ to denote the spectrum of A.

In finite dimensions, this is just the eigenvalues of A. To see this, note that if A is in upper triangular form then its eigenvalues are the entries on its main diagonal. Then $A - \lambda$ has full rank (and is thus invertible) if and only if λ does not appear on the diagonal of A.

2.4.2 Definition

The *spectral radius* of an operator A in $\mathcal{B}(\mathcal{V})$ is

$$\rho(A) = \{|\lambda| : \lambda \in \sigma(A)\}.$$

2.4.3 Definition

A *word* in a family of operators, $\mathcal{F}$, in $\mathcal{B}(\mathcal{V})$ is a finite expression $F_1 F_2 \cdots F_k$ with F_i in $\mathcal{F}$. The F_i's need not be distinct.

A *noncommutative polynomial* in operators $\{A_1, \ldots, A_k\}$ in $\mathcal{B}(\mathcal{V})$ is any linear combination of words in the operators.

These definitions lead easily to the following result.

2.4.4 Theorem (Spectral Mapping Theorem)

If $\{A_1, \ldots, A_k\}$ is a triangularizable collection of linear transformations, and if p is any noncommutative polynomial in $\{A_1, \ldots, A_k\}$, then

$$\sigma(p(A_1, \ldots, A_k)) \subseteq p(\sigma(A_1), \ldots, \sigma(A_k)).$$

Proof. The A_i's are simultaneously triangularizable, so we'll assume they are in upper triangular form. Then the eigenvalues of A_i appear on its main diagonal.

The diagonal entries of a product of upper triangular matrices are the product of the diagonal entries of those matrices. Therefore

$$\sigma(A_{i_1} A_{i_2} \cdots A_{i_k}) \subseteq \sigma(A_{i_1})\sigma(A_{i_2}) \cdots \sigma(A_{i_k})$$

for any word in the A_i's.

Similarly, the diagonal entries of a linear combination of upper triangular matrices are a linear combination of the diagonal entries of the matrices. Therefore the result holds.

□

The next result deals with convergent sequences of operators in $\mathcal{B}(\mathcal{V})$. Since $\mathcal{V}$ is finite dimensional, all norms are equivalent and we don't need to specify a norm under which the sequence converges.

2.4.5 Lemma

If $\{A_n\}$ is a sequence of operators in $\mathcal{B}(\mathcal{V})$ with $A = \lim_{n \longrightarrow \infty} A_n$ then

$$\sigma(A) = \lim_{n \longrightarrow \infty} \{\sigma(A_n)\} = \{\lambda : \lambda = \lim_{n \longrightarrow \infty} \lambda_n, \lambda_n \in \sigma(A_n)\}.$$

Proof. Assume λ_n is in $\sigma(A_n)$ and $\lambda = \lim_{n \longrightarrow \infty} \lambda_n$. Since λ_n is an eigenvalue of A_n, $A_n - \lambda$ has zero as an eigenvalue and $\det(A_n - \lambda_n) = 0$. Since det is a continuous function, $\det(A - \lambda) = 0$, zero is an eigenvalue for $A - \lambda$, and λ is in $\sigma(A)$.

For the converse, take λ is in $\sigma(A)$ and assume it is not the limit of $\{\lambda_n\}$ where λ_n is in $\sigma(A_n)$. In other words, there is a closed disc, $\mathcal{D}$, around λ such that for any N there is an $n \geq N$ such that $\sigma(A_n) \cap \mathcal{D} = \emptyset$. Since $\sigma(A)$ is finite, we can also assume $\sigma(A) \cap \mathcal{D} = \{\lambda\}$. Take a strictly increasing sequence of integers n_j with $\sigma(A_{n_j}) \cap D = \emptyset$.

We can define polynomials $f_j(z) = \det(A_{n_j} - z)$ and $f(z) = \det(A - z)$. We want to show that f_j converges uniformly to f so fix $\varepsilon > 0$. Since det is a continuous function, we can find a $\delta > 0$ such that, if

$$\left\|A_{N_j} - A\right\| = \left\|(A_{N_j} - z) - (A - z)\right\| < \delta,$$

then

$$\|f_j(z) - f(z)\| = \left\|\det(A_{N_j} - z) - \det(A - z)\right\| < \varepsilon.$$

Since $\{A_{N_j}\}$ converges to A, we can find a J such that if $j > J$ then $\left\|A_{N_j} - A\right\| < \delta$. Since this J doesn't depend on z, f is the uniform limit of f_j.

As $\sigma(A) \cap \mathcal{D} = \{\lambda\}$, $f(z)$ is bounded away from zero on the boundary of $\mathcal{D}$, a compact set. Since f is the uniform limit of f_j, $|f_j(z)| \geq \varepsilon > 0$ for all $j > J$ for some $J > 0$. We can remove the smaller indices and assume this relation holds for all j.

Since $\sigma(A_{n_j}) \cap \mathcal{D} = \emptyset$, $f_{n_j}(z) \neq 0$ for every z in $\mathcal{D}$. So $1/f_j$ is analytic on $\mathcal{D}$ for all j and we can apply the maximum modulus priciple to determine that the maximum of $|1/f_j|$. Thus the minimum of $|f_j|$, occurs on the boundary of $\mathcal{D}$. Therefore $f_j(z) \geq$

$\varepsilon > 0$ for all z in $\mathcal{D}$. But this means that $|f(\lambda)| = \lim_{j \longrightarrow \infty} |f_j(\lambda)| \geq \varepsilon > 0$ which contradicts $f(\lambda) = 0$ and the result holds.

$\square$

2.5 Field and Ring Automorphisms

The following simple results and definitions will be of use in our discussion of the Finiteness Lemma (3.2.2) and subsequent results.

2.5.1 Lemma

Assume that the $2n$ numbers $\{\alpha_1, \ldots, \alpha_n, \beta_1, \ldots \beta_n\}$ are algebraically independent over $\mathbb{Q}$ (i.e. there is no nontrivial polynomial p in $2n$ indeterminates over $\mathbb{Q}$ with $p(\alpha_1, \ldots, \beta_n) = 0$). Then there exists a field automorphism ϕ of $\mathbb{C}$ such that $\phi(\alpha_i) = \beta_i$ and $\phi(\beta_i) = \alpha_i$ for every i.

Proof. Let $\mathbb{F}$ be the extension field $\mathbb{Q}(\alpha_1, \ldots, \beta_n)$. We can define a map $\phi : \mathbb{F} \longrightarrow \mathbb{F}$ by $\phi(f(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots \beta_n)) = f(\beta_1, \ldots, \beta_n, \alpha_1, \ldots, \alpha_n)$ for every rational function f. ϕ is well-defined since the $2n$ numbers are algebraically independent over $\mathbb{Q}$ and ϕ is an automorphism. By Zorn's Lemma, we can extend ϕ to be an automorphism of $\mathbb{C}$.

$\square$

2.5.2 Definition

Let ϕ be any field automorphism of $\mathbb{C}$. The map $\Phi : M_n(\mathbb{C}) \longrightarrow M_n(\mathbb{C})$ defined by

$$(\Phi(A))_{ij} = \phi(A_{ij})$$

for all i and j is a ring automorphism. We call it the *automorphism of $M_n(\mathbb{C})$ induced by ϕ*. Generally, we will use the notation Φ for all values of n and for the induced isomorphism from a semigroup of operators $\mathcal{S}$ in $M_n(\mathbb{C})$ to $\Phi(\mathcal{S})$.

2.5.3 Lemma

If ϕ is an automorphism of $\mathbb{C}$ and Φ is the induced automorphism, then $\sigma(\Phi(A)) = \phi(\sigma(A))$.

Proof. Since

$$\det(\Phi(A) - \phi(\lambda)) = \phi(\det(A - \lambda)),$$

λ is an eigenvalue of A ($\det(A - \lambda) = 0$) if and only if $\phi(\lambda)$ is an eigenvalue of $\Phi(A)$.

$\square$

Chapter 3

Finite Dimensions

In this chapter, we discuss a number of sufficient conditions for triangularizability in finite dimensions. Specifically, we look at permutability of the trace, sublinearity and subadditivity of the spectrum, and the nilpotence of the semigroup under certain polynomials.

3.1 Permutability of the Trace

We saw in Lemma 2.2.10 that a nonzero functional annihilating a semigroup was sufficient for reducibility. In this chapter, we consider a generalization of this condition, permutability. When permutability for an arbitrary functional proves insufficient, we consider the permutability of the trace.

3.1.1 Definition

Let ϕ be a linear functional on $\mathcal{B}(\mathcal{V})$. We say that ϕ is *permutable* on a family $\mathcal{F}$ of operators in $\mathcal{B}(\mathcal{V})$ if, for any $A_1, \ldots, A_n$ in $\mathcal{F}$ and any permutation τ of $\{1, \ldots, n\}$, we have

$$\phi(A_1 A_2 \cdots A_n) = \phi(A_{\tau(1)} A_{\tau(2)} \cdots A_{\tau(n)}).$$

We say that ϕ is *multiplicative* on $\mathcal{F}$ if $\phi(AB) = \phi(A)\phi(B)$ for all A and B in $\mathcal{F}$. Clearly, if ϕ is multiplicative on $\mathcal{F}$ then ϕ is permutable on $\mathcal{F}$.

3.1.2 Lemma

Let ϕ be a linear functional on $\mathcal{B}(\mathcal{V})$. If $\mathcal{S}$ is a semigroup of operators in $\mathcal{B}(\mathcal{V})$ then ϕ is permutable on $\mathcal{S}$ if and only if

(i) $\phi(AB) = \phi(BA)$, *and*

(ii) $\phi(ABC) = \phi(BAC)$

for all A, B and C in $\mathcal{S}$.

Proof. Properties (i) and (ii) are clearly implied by permutability. For the other direction, let ϕ be a linear functional satisfying (i) and (ii). We'll prove that ϕ is permutable on $\mathcal{S}$ by induction on the number of letters.

By using (i), we have that $\phi(ABC) = \phi(CAB) = \phi(BCA)$ and (ii) allows us to rearrange the first two letters. Thus ϕ is permutable on three letters from $\mathcal{S}$.

Assume that ϕ is permutable on fewer than n letters of $\mathcal{S}$. We want to show that $\phi(A_{\tau(1)} \cdots A_{\tau(n)}) = \phi(A_1 \cdots A_n)$ for all $A_1, \ldots, A_n$ in $\mathcal{S}$ and any permutation τ.

$\mathcal{S}$ is a semigroup so products of A_i's are still in $\mathcal{S}$. We have

$$\begin{aligned}
\phi(A_{\tau(1)} \cdots A_{\tau(n)}) &= \phi((A_* \cdots A_n)(A_* \cdots A_*)) \\
&= \phi((A_* \cdots A_*)(A_* \cdots A_n)) \\
&= \phi((A_* \cdots A_{n-1})(A_* \cdots A_*)(A_n)) \\
&= \phi((A_* \cdots A_*)(A_* \cdots A_{n-1})A_n),
\end{aligned}$$

where the first and third equality follow as τ is a permutation, the second follows from (i), and the fourth comes from (ii).

Since $A_{n-1}A_n$ is in $\mathcal{S}$, showing that the last line is equal to $\phi(A_1 \cdots A_{n-1}A_n)$ reduces to the case on $n-1$ elements and the result is proved.

$\square$

3.1.3 Lemma

Let $\mathcal{F}$ be a family of operators in $\mathcal{B}(\mathcal{V})$ and let ϕ be a nonzero linear functional on $\mathcal{B}(\mathcal{V})$. If ϕ is permutable on $\mathcal{F}$ then $\mathcal{F}$ is reducible.

Proof.

Let $\mathcal{S}$ and $\mathcal{A}$ be, respectively, the semigroup and the algebra generated by $\mathcal{F}$ and note that $\mathcal{A} = \mathrm{span}(\mathcal{S})$. As ϕ is permutable on $\mathcal{F}$ and $\mathcal{S}$ consists of products of members of $\mathcal{F}$, ϕ is permutable on $\mathcal{S}$. For any operators A, B and C in $\mathcal{A}$ we can write

$$A = \textstyle\sum_{i=1}^{n} \alpha_i A_i, \quad B = \sum_{j=1}^{m} \beta_j B_j, \quad C = \sum_{k=1}^{l} \gamma_k C_k,$$

with the A_i, B_j, and C_k in $\mathcal{S}$ and α_i, β_j, and γ_k in $\mathbb{C}$. Then

$$\begin{aligned}
\phi(ABC) &= \phi((\sum_{i=1}^{n} \alpha_i A_i)(\sum_{j=1}^{m} \beta_j B_j)(\sum_{k=1}^{l} \gamma_k C_k)) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l} \alpha_i \beta_j \gamma_k \phi(A_i B_j C_k) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l} \alpha_i \beta_j \gamma_k \phi(B_j A_i C_k) \\
&= \phi(BAC)
\end{aligned}$$

by using the linearity of ϕ and its permutability on $\mathcal{S}$. Similarly, $\phi(AB) = \phi(BA)$. By Lemma 3.1.2, ϕ is permutable on $\mathcal{A}$.

Assume $\mathcal{F}$ is irreducible. Then $\mathcal{A} = \mathcal{B}(\mathcal{V})$ by Burnside's Theorem (2.2.8). Therefore we can take A and B in $\mathcal{A}$ such that $AB - BA \neq 0$. Let $\mathcal{J} \neq \{0\}$ be the semigroup ideal of $\mathcal{A}$ generated by $AB - BA$. For any X and Y in $\mathcal{A}$ we have that $\phi(X(AB - BA)Y) = \phi(XABY) - \phi(XBAY) = 0$ since ϕ is permutable on $\mathcal{A}$. Therefore $\phi|_{\mathcal{J}} = 0$ and by Lemma 2.2.10, $\mathcal{J}$ is reducible. Then by Lemma 2.2.12, $\mathcal{A}$ is reducible. This is a contradiction, so $\mathcal{F}$ is reducible.

□

In certain situations, this result gives us triangularizability.

3.1.4 Theorem (Kolchin's Theorem)

If every member of a semigroup $\mathcal{S}$ of operators in $\mathcal{B}(\mathcal{V})$ is unipotent (i.e. every element S of $\mathcal{S}$ has $\sigma(S) = \{1\}$) then $\mathcal{S}$ is triangularizable

Proof. Let S be in $\mathcal{S}$. Then we can triangularize S and, since it's unipotent, its diagonal consists entirely of 1's. If $n = \dim(\mathcal{V})$ then $\mathrm{tr}(S) = n$ so the trace is constant on $\mathcal{S}$. Constancy is a special case of permutability so tr is a nonzero linear functional which is permutable on $\mathcal{S}$. By Lemma 3.1.3, $\mathcal{S}$ is reducible. The property of being unipotent is inherited by quotients, so $\mathcal{S}$ is triangularizable by the Triangularization Lemma (2.2.3).

□

In general, a permutable functional is doesn't give triangularizability. For instance,

$$\mathcal{A} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} : A \in M_n(\mathbb{C}) \right\}$$

is not triangularizable since $M_n(\mathbb{C})$ is irreducible. However, the functional ϕ on $M_{n+1}(\mathbb{C})$ where $\phi(A) = A_{11}$ is permutable (constant even) on $\mathcal{A}$.

However, if the trace is permutable on a family of operators in $\mathcal{B}(\mathcal{V})$ then we get triangularizability. Additionally, since $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for all operators A and B, Lemma 3.1.2 tells us that the trace is permutable on a family $\mathcal{F}$ if and only if $\mathrm{tr}(ABC) = \mathrm{tr}(BAC)$ for all A, B and C in $\mathcal{F}$.

In order to prove this, we'll use the following two technical results.

3.1.5 Lemma

Let $\{\alpha_1, \ldots, \alpha_n\}$ and $\{\beta_1, \ldots, \beta_n\}$ be in $\mathbb{C}^n$.

(i) If $\sum_{i=1}^n \alpha_i^k = \sum_{i=1}^n \beta_i^k$ for $k = 1, \ldots, n$, then there is a permutation τ on n letters such that $\beta_i = \alpha_{\tau(i)}$ for all i.

(ii) If $\sum_{i=1}^n \alpha_i^k = 0$ for $k = 1, \ldots, n$ then $\alpha_i = 0$ for all i.

(iii) If $\sum_{i=1}^n \alpha_i^k = c$ with c fixed for $k = 1, \ldots, n+1$, then c is an integer and each α_i is either 0 or 1.

Proof. (i) For each k, define symmetric polynomials T_k by

$$T_k(x_1, \ldots, x_n) = \sum_{i=1}^n x_i^k,$$

and elementary symmetric polynomials S_k of degree k (i.e. S_k is the sum of all products of k variables) so

$$S_k(x_1, \ldots, x_n) = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k}.$$

The initial conditions then become $T_k(\alpha_1, \ldots, \alpha_n) = T_k(\beta_1, \ldots, \beta_n)$ for $k = 1, \ldots, n$.

We can algebraically verify that, for every k,

$$T_k - T_{k-1}S_1 + T_{k-2}S_2 - \cdots + (-1)^{k-1}T_1 S_{k-1} + (-1)^k k S_k = 0.$$

We then claim that $S_k(\alpha_1, \ldots, \alpha_n) = S_k(\beta_1, \ldots, \beta_n)$ and prove it by induction. Since $S_1 = T_1$ the initial hypothesis proves the base case. Assume it holds for values less than k. Then S_k can be expressed in terms of $S_1, \ldots, S_{k-1}, T_1, \ldots, T_k$ by the above equation. By induction, we know that equality holds on all of those elements so equality holds for S_k.

It can easily be seen that

$$(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) = x^n + S_1(\alpha_1, \ldots, \alpha_n)x^{n-1} + \cdots + S_n(\alpha_1, \ldots, \alpha_n).$$

And similarly for the β_i. But then these two polynomials agree on all their coefficients so they're equal and have the same roots. But their roots are exactly $\{\alpha_1, \ldots, \alpha_n\}$ and $\{\beta_1, \ldots, \beta_n\}$, respectively and the result holds.

(ii) This is a special case of (i) with $\beta_i = 0$ for all i.

(iii) If $c = 0$ we're done by (ii). Otherwise, we can permute the α_i and assume $\alpha_1, \ldots, \alpha_m$ are nonzero while the rest are zero. Since zeroes don't affect the sum, we can assume $m = n$. We want to show that $c = n$ and that $\alpha_i = 1$ for all i.

Calculation gives

$$T_{n+1} = T_n S_1 - T_{n-1} S_2 + \cdots + (-1)^{n-1} T_1 S_n.$$

Apply this result to the specific case of the α_i's, denote $S_k(\alpha_1, \ldots, \alpha_n)$ by s_k, and recall that $T_k(\alpha_1, \ldots, \alpha_n) = c$. After dividing both sides by c and rearranging we have

$$1 - s_1 + s_2 + \cdots + (-1)^{n-1} s_{n-1} = (-1)^{n-1} s_n,$$

while the recursive equation in (i) gives

$$c(1 - s_1 + s_2 + \cdots + (-1)^{n-1} s_{n-1}) + (-1)^n n s_n = 0.$$

Combining these equations gives

$$(-1)^{n-1}cs_n + (-1)^n ns_n = 0,$$

so $c = n$. Then, by using (i) with $\beta_i = 1$ for all i we get $\alpha_i = 1$ for all i as required.

$\square$

3.1.6 Lemma

Let A and B be operators in $\mathcal{B}(\mathcal{V})$. If $\operatorname{tr}(A^k) = \operatorname{tr}(B^k)$ for $k = 1, \ldots, n$ then A and B have the same eigenvalues, counting multiplicity. In particular, if $\operatorname{tr}(A^k) = 0$ for $k = 1, \ldots, n$ then A is nilpotent.

Proof. We can triangularize A without affecting the trace. Let the diagonal of A be $\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$. Then the diagonal of A^k is $\operatorname{diag}(\alpha_1^k, \ldots, \alpha_n^k)$ and $\operatorname{tr}(A^k) = \sum_{i=1}^n \alpha_i^k$. Similarly, $\operatorname{tr}(B) = \sum_{i=1}^n \beta_i^k$ where the diagonal of B is $\operatorname{diag}(\beta_1, \ldots, \beta_n)$.

Using the fact that the eigenvalues of A are exactly its diagonal entries when it's in upper triangular form, the first part follows from Lemma 3.1.5 (i) while the second part follows from Lemma 3.1.5 (ii).

$\square$

We're now ready to show that if the trace is permutable on $\mathcal{S}$ then $\mathcal{S}$ is triangularizable.

3.1.7 Theorem

Let $\mathcal{F}$ be a family of operators in $\mathcal{B}(\mathcal{V})$. Then $\mathcal{F}$ is triangularizable if and only if trace is permutable on $\mathcal{F}$.

Proof. If $\mathcal{F}$ is triangularizable, then for any A, B and C in $\mathcal{F}$

$$(ABC)_{ii} = A_{ii}B_{ii}C_{ii} = B_{ii}A_{ii}C_{ii} = (BAC)_{ii}.$$

Therefore $\operatorname{tr}(ABC) = \operatorname{tr}(BAC)$ and trace is permutable on $\mathcal{F}$ by Lemma 3.1.2.

Assume trace is permutable on $\mathcal{F}$. As we saw in Lemma 3.1.2, trace is permutable on the semigroup $\mathcal{S}$ and the algebra $\mathcal{A}$ which are generated by $\mathcal{F}$. So for any A, B

and C in $\mathcal{A}$

$$tr((AB - BA)C) = \mathrm{tr}(ABC) - \mathrm{tr}(BAC) = 0.$$

Since $\mathcal{A}$ is an algebra, $(AB - BA)^k$ is in $\mathcal{A}$ for any $k \in \mathbb{N}$. If $C = (AB - BA)^k$ we have $\mathrm{tr}((AB - BA)^{k+1}) = 0$ for any natural number k. Also $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for any A and B. Therefore $AB - BA$ is nilpotent for all operators A and B in $\mathcal{A}$ by Lemma 3.1.6. Then $\mathcal{A}$ is triangularizable by Lemma 2.2.9. Therefore $\mathcal{F}$ is triangularizable.

□

3.1.8 Corollary (Kaplansky's Theorem)

Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{B}(\mathcal{V})$. If trace is constant on $\mathcal{S}$ then $\mathcal{S}$ is triangularizable. Moreover, every diagonal entry in a triangularization of such a semigroup is constantly zero or constantly one.

Proof. Since trace is constant on $\mathcal{S}$ it's permutable on $\mathcal{S}$. Therefore $\mathcal{S}$ is triangularizable by Theorem 3.1.7.

Triangularize $\mathcal{S}$ and consider any A in $\mathcal{S}$. Trace is unchanged by similarity so $\mathrm{tr}(A^k) = \sum_{i=1}^{n} A_{ii}^k = c$ for all $k \in \mathbb{N}$ and some constant c. By Lemma 3.1.5, each A_{ii} is either zero or one.

Now, if A and B are in $\mathcal{S}$ then $(AB)_{ii}$ is one if and only if both A_{ii} and B_{ii} are one. Since trace is constant on $\mathcal{S}$, A, B and AB must have exactly the same number of ones on their diagonal. Thus, for each i, either both of A_{ii} and B_{ii} are zero or both are one.

□

3.1.9 Corollary

Let $\mathcal{G}$ be a group of operators in $\mathcal{B}(\mathcal{V})$ and let $\mathcal{H}$ denote its commutator subgroup (the normal subgroup of $\mathcal{G}$ generated by all elements $A^{-1}B^{-1}AB$ with A and B in $\mathcal{G}$). Then the following are equivalent:

(i) $\mathcal{G}$ is triangularizable.

(ii) Trace is constant on each coset of $\mathcal{G}$ relative to $\mathcal{H}$.

(iii) Trace is constant on $\mathcal{H}$.

(iv) $\mathcal{H}$ consists of unipotent operators.

Proof. If $\mathcal{G}$ is triangularizable then trace is permutable on $\mathcal{G}$ by Theorem 3.1.7. For any A, B, and G in $\mathcal{G}$, $\mathrm{tr}(G(A^{-1}B^{-1}AB)) = \mathrm{tr}(GI) = \mathrm{tr}(G)$. Therefore for any H in $\mathcal{H}$ we have $\mathrm{tr}(GH) = \mathrm{tr}(G)$ so trace is constant on $G\mathcal{H}$ and (i) implies (ii).

As $\mathcal{H}$ is a coset of $\mathcal{G}$ relative to itself, (ii) implies (iii) trivially.

If the trace is constant on $\mathcal{H}$, $\mathcal{H}$ is triangularizable by Kaplansky's Theorem (3.1.8). Further, it tells us that $\sigma(H) \subseteq \{0, 1\}$ for every H in $\mathcal{H}$. Since $\mathcal{H}$ is a group, H is invertible and 0 cannot be in $\sigma(H)$. Therefore $\sigma(H) = \{1\}$ and (iii) implies (iv).

Finally, assume that $\mathcal{H}$ consists of unipotent operators. If $\mathcal{H} = \{I\}$ then $\mathcal{G}$ is commutative and $\mathcal{G}$ is triangularizable by Theorem 2.2.4.

Consider $\mathcal{H} \neq \{I\}$. The commutator subgroup of a quotient is the quotient of the original commutator subgroup. Also, $\sigma(H|_{\mathcal{M}/\mathcal{N}}) \subseteq \sigma(H) = \{1\}$ for H in $\mathcal{H}$ with invariant subspaces $\mathcal{N} \subset \mathcal{M}$. Since a commutator subgroup consisting of unipotents is inherited by quotients, it's sufficient to show that $\mathcal{G}$ is reducible by the Triangularization Lemma (2.2.3).

By Kolchin's Theorem (3.1.4), $\mathcal{H}$ is triangularizable. Take a triangularizing chain $\{0\} = V_0 \subset \cdots \subset V_{n-1} \subset V_n = \mathcal{V}$ for $\mathcal{H}$. We claim

$$\mathcal{M} = \mathrm{span}\{(H - I)\mathcal{V} : H \in \mathcal{H}\}$$

is a nontrivial invariant subspace for $\mathcal{G}$.

Since $\mathcal{H} \neq \{I\}$, $H - I \neq 0$ for some H in $\mathcal{H}$ and $\mathcal{M} \neq \{0\}$. Since, with respect to the V_i's, $\mathcal{H}$ consists of upper triangular unipotent operators, the last row of $H - I$ for every H in $\mathcal{H}$ is zero. Therefore $(H - I)\mathcal{V} \subseteq V_{n-1} \neq \mathcal{V}$. Hence $\mathcal{M}$ is nontrivial.

For any G in $\mathcal{G}$ and H in $\mathcal{H}$, $G(H - I) = (GHG^{-1} - I)G$, so

$$G(H - I)\mathcal{V} = (GHG^{-1} - I)G\mathcal{V} = (GHG^{-1} - I)\mathcal{V}$$

and $\mathcal{H}$ is normal so $GHG^{-1} \in \mathcal{H}$. Therefore $\mathcal{G}\mathcal{M} \subseteq \mathcal{M}$.

As $\mathcal{M}$ is a nontrivial invariant subspace for $\mathcal{G}$, $\mathcal{G}$ is reducible and thus triangularizable.

$\square$

3.1.10 Corollary

Let $\mathcal{F}$ be a self-adjoint family of operators in $\mathcal{B}(\mathcal{V})$. Then $\mathcal{F}$ is commutative if and only if trace is permutable on $\mathcal{F}$.

Proof. If $\mathcal{F}$ is commutative then trace is clearly permutable on $\mathcal{F}$.

Assume that trace is permutable on $\mathcal{F}$. By Theorem 3.1.7, $\mathcal{F}$ is triangularizable. Let $\{0\} = \mathcal{M}_0 \subset \mathcal{M}_1 \subset \cdots \subset \mathcal{M}_n = \mathcal{V}$ be a triangularizing chain for $\mathcal{F}$. Take $\mathcal{V}_j$ to be the orthogonal complement of $\mathcal{M}_{j-1}$ relative to $\mathcal{M}_j$ for $j = 1, \ldots, n$. Take a unit vector e_i from each V_i. Then the set $\{e_1, \ldots, e_n\}$ is an orthonormal basis for $\mathcal{V}$ with $\mathcal{M}_i$ spanned by $\{e_1, \ldots, e_i\}$.

Let $1 \leq i < j \leq n$. For any F in $\mathcal{F}$, $Fe_i \subseteq \mathcal{M}_i \subseteq \mathcal{M}_{j-1}$. By definition, $\langle Fe_i, e_j \rangle = 0$ as $e_j \in \mathcal{V}_j$. Additionally, $\mathcal{F}$ is self-adjoint so $F^* \in \mathcal{F}$. Therefore $\langle Fe_j, e_i \rangle = \langle e_j, F^* e_i \rangle = 0$. Therefore $Fe_i \perp e_j$ for $i \neq j$ so F is diagonal with respect to $\{e_1, \ldots, e_n\}$.

Therefore $\mathcal{F}$ is commutative as it's diagonal with respect to the e_i's.

□

3.2 The Finiteness Lemma

Our goal in this section is to prove a result that will allow us to reduce questions about certain semigroups of operators in $\mathcal{B}(\mathcal{V})$ to questions about finite groups. We are interested in semigroups with a property $\mathcal{P}$ that is stable under a number of conditions. For instance, if $\mathcal{S}$ has property $\mathcal{P}$ then we require that both $\overline{\mathcal{S}}$ and $\mathbb{C}\mathcal{S}$ also have property $\mathcal{P}$.

We'd also like to require that for any field automorphism ϕ and the induced automorphism Φ, $\Phi(\mathcal{S})$ have property $\mathcal{P}$. However, this would not allow us to consider semigroups that satisfy polynomial equations with non-rational coefficients as such polynomials are not stable under all field automorphisms of $\mathbb{C}$. Therefore, we ask that if $\mathcal{S}$ has property $\mathcal{P}$ then so does $\Phi^{-1}(\overline{\mathbb{C}\Phi(\mathcal{S})})$.

The following preliminary result from group theory will be necessary to prove our major result.

3.2.1 Lemma

Let $\mathcal{G}$ be a compact unitary group in $\mathcal{B}(\mathcal{V})$. If $\mathcal{G}$ is a torsion group (a group in which every element has finite order), then $\mathcal{G}$ is a finite group.

Proof.

First we'll show that there is a k in $\mathbb{N}$ such that $G^k = I$ for every G in $\mathcal{G}$. Consider an arbitrary G in $\mathcal{G}$. It's unitary so it can be diagonalized with its diagonal consisting of its spectrum which is a subset of $\mathbb{T}$. Since it's a torsion group, its spectrum consists of roots of unity. If every such λ was at most a k^{th} root of unity then $\lambda^{k!} = 1$ for all λ and $G^{k!} = I$ for all $G \in \mathcal{G}$.

Assume there is no k as above. Then for every m, there must be a G_m in $\mathcal{G}$ with a λ_m in its spectrum with λ_m at least an m^{th} root of unity. Then $\{\lambda_m^k : m, k \in \mathbb{N}\}$ is dense in $\mathbb{T}$, the unit circle in $\mathbb{C}$. Choose a λ in $\mathbb{T}$ which is not a root of unity. Take a sequence of powers of λ_m's converging to λ. Since $\mathcal{G}$ is compact, we can take a subsequence of the corresponding powers of G_m's that converges to a G in $\mathcal{G}$. But then λ is in $\sigma(G)$ by Lemma 2.4.5 which contradicts that the spectrum of elements of $\mathcal{G}$ consists of roots of unity. Therefore, there is a k such that $G^k = I$ for all G in $\mathcal{G}$.

Now we need to show that this means $\mathcal{G}$ is finite. Let $0 < \varepsilon < \left|1 - e^{2\pi i/k}\right|$ and define a neighbourhood of I in G by

$$\mathcal{N}_I = \{G \in \mathcal{G} : \|I - G\| < \varepsilon\}.$$

Let G be in $\mathcal{N}_I$. We can diagonalize G without affecting I. Then, $\|I - G\| < \varepsilon$ means that every element of $\sigma(G)$ is within ε of 1. But any root of unity (except 1) that close to 1 is at least a $(k+1)^{\text{st}}$ root of unity which would contradict $G^k = I$. Therefore, $\sigma(G) = \{1\}$ so $G = I$. Therefore $\mathcal{N}_I = \{I\}$.

For each G_0 in $\mathcal{G}$ we can define

$$\mathcal{N}_{G_0} = \{G \in \mathcal{G} : \|G_0 - G\| < \varepsilon\},$$

but if $G \in \mathcal{N}_{G_0}$ then

$$\left\|I - G_0^{-1}G\right\| = \left\|G_O^{-1}(G_0 - G)\right\| \leq \left\|G_0^{-1}\right\| \|G_0 - G\| = \|G_0 - G\| < \varepsilon,$$

so $G_0^{-1}G$ is in $\mathcal{N}$ so $G = G_0$ and $\mathcal{N}_{G_0} = \{G_0\}$.

Now, $\{\mathcal{N}_G\}_{G \in \mathcal{G}}$ is a cover for $\mathcal{G}$ and $\mathcal{G}$ is compact so there is a finite subcover. But each of those neighbourhoods consists of a single element so $\mathcal{G}$ is finite.

$\square$

We can now proceed to prove the major result of this section.

3.2.2 Lemma (The Finiteness Lemma)

Let $\mathcal{P}$ be a property defined for semigroups of operators in $\mathcal{B}(\mathcal{V})$ such that, whenever $\mathcal{S}$ has the property, so does the semigroup $\Phi^{-1}(\overline{\mathbb{C}\Phi(\mathcal{S})})$ for every choice of ring automorphism Φ induced by a field automorphism of $\mathbb{C}$. Let $\mathcal{S}$ be a maximal semigroup in $\mathcal{B}(\mathcal{V})$ with property $\mathcal{P}$. Denote the minimal nonzero rank in $\mathcal{S}$ by m. If E is an idempotent of rank m in $\mathcal{S}$, then $ESE|_{E\mathcal{V}}$ is of the form $\mathbb{C}\mathcal{G}$, where $\mathcal{G}$ is a finite group (similar to a unitary group in $M_m(\mathbb{C})$).

Proof. For any automorphism Φ, $\mathcal{S}$ is contained within $\Phi^{-1}(\overline{\mathbb{C}\Phi(\mathcal{S})})$ so, by maximality, $\mathcal{S} = \Phi^{-1}(\overline{\mathbb{C}\Phi(\mathcal{S})})$. In particular, the identity field automorphism induces the identity ring automorphism so $\mathcal{S} = \overline{\mathbb{C}\mathcal{S}}$ by substituting the identity for Φ.

By Lemma 2.3.3, we can assume $\mathcal{S}_0 = E\mathcal{S}E|_{E\mathcal{V}}$ is contained in $\mathbb{R}^+\mathcal{U}$ where $\mathcal{U}$ is the set of unitaries in $M_m(\mathbb{C})$. So, every element in $\mathcal{S}_0$ is of the form rU where r is a nonnegative number and U is a unitary. Since $\mathcal{S} = \mathbb{C}\mathcal{S}$, we have that U is in $\mathcal{S}_0$. Let $\mathcal{G} = \{S \in \mathcal{S}_0 : \det(S) = 1\}$ so, since $\mathcal{S}_0$ is contained in $\mathbb{R}^+\mathcal{U}$, $\mathcal{G}$ is a set of unitaries in $\mathcal{S}_0$. If we can show that $\mathcal{G}$ is finite then we're done.

First we'll show that every element of $\mathcal{G}$ has finite order, in other words, that $\mathcal{G}$ is a torsion group. Assume that $\mathcal{G}$ contains a member, A, with infinite order. If $A^r = \lambda I$ for some λ and positive integer r, then $A^{r^2} = I$ since $\lambda^r = \det(A^r) = (\det(A))^r = 1$ which would contradict A having infinite order.

Let α_0 be an eigenvalue (so $|\alpha_0| = 1$) of A and define $B = A/\alpha_0$. Then 1 is in $\sigma(B)$ and no power of B can be a scalar either. Since B is unitary, we can assume it's diagonal. If every eigenvalue, α_i, of B had a positive integer r_i such that $\alpha_i^{r_i} = 1$ then $B^{\mathrm{lcm}(r_1,\ldots,r_m)} = I$ which is impossible. Therefore, there must be some eigenvalue, α, of B such that $\alpha^r \neq 1$ for all r.

Then $\{\alpha^r : r \in \mathbb{N}\}$ is dense in the unit circle, $\mathbb{T}$. Let λ be a transcendental number in $\mathbb{T}$ and choose a sequence $\{\alpha^r\}$ converging to λ. Since $\mathcal{S} = \overline{\mathcal{S}}$, $\mathcal{G}$ is closed and it's a unitary group so it's bounded. Therefore $\mathcal{G}$ is compact so we can take a subsequence $\{B^{r_i}\}$ which converges to B_0.

We know 1 and λ are in $\sigma(B_0)$ by Lemma 2.4.5. Let μ be algebraically independent of λ with $|\mu| < 1$. Using Lemma 2.5.1, let ϕ be a field automorphism on $\mathbb{C}$ with $\phi(\lambda) = \mu$ and $\phi(\mu) = \lambda$. Let Φ be the ring automorphism induced by ϕ.

Since E is in $\mathcal{S}$, so is EB_0E. We know EB_0E has rank m (as it's a unitary on $E\mathcal{V}$) and we can assume it's diagonal as its restriction is a unitary. Since λ and 1 are in $\sigma(EB_0E)$, they appear on the diagonal of EB_0E. Then $\Phi(EB_0E)$ is also diagonal, has 1 and μ on its diagonal and has rank m.

As $\frac{\Phi(EB_0E)}{\rho(\Phi(EB_0E))}$ is diagonal with entries of modulus at most one,

$$F_n = \left(\frac{\Phi(EB_0E)}{\rho(\Phi(EB_0E))}\right)^n$$

is bounded. We can thus take a convergent subsequence, F_{n_i}, with limit F. Since $|\mu| < 1$ and $\rho(\Phi(EB_0E)) \geq 1$, $(\frac{\mu}{\rho(\Phi(EB_0E))})^{n_i}$ converges to zero. Therefore, F has rank at most $m-1$. Also, since at least one of the diagonal entries must have the same modulus as $\rho(\Phi(EB_0E))$, F has rank at least 1.

Now, F is in $\overline{\mathbb{C}\Phi(\mathcal{S})}$ so $\Phi^{-1}(F)$ is in $\hat{\mathcal{S}} = \Phi^{-1}(\overline{\mathbb{C}\Phi(\mathcal{S})})$ which contains $\mathcal{S}$. By maximality, we should have $\mathcal{S} = \hat{\mathcal{S}}$. However, $\Phi^{-1}(F)$ has the same rank as F, specifically, no more than $m-1$ and at least 1. But then $\Phi^{-1}(F)$ can't be in $\mathcal{S}$ as m was the minimal nonzero rank in $\mathcal{S}$. This contradicts maximality so $\mathcal{G}$ must be a torsion group.

By Lemma 3.2.1, $\mathcal{G}$ is finite and we're done.

□

The Finiteness Lemma can then be used to address issues of reducibility. To do so, we make use of the following technical lemma.

3.2.3 Lemma

Let $\mathcal{P}$ be a property satisfying the hypotheses of the Finiteness Lemma. Assume, furthermore, that

(i) if $\mathcal{S}$ is a semigroup with property $\mathcal{P}$, $\mathcal{J}$ is an ideal of $\mathcal{S}$, and E a minimal nonzero idempotent in $\mathcal{S}$, then both $\mathcal{J}$ and $ESE|_{E\mathcal{V}}$ have property $\mathcal{P}$,

(ii) every finite group with property $\mathcal{P}$ is reducible, and

(iii) every semigroup of operators of rank at most one with property $\mathcal{P}$ is reducible.

Then every semigroup $\mathcal{S}$ with property $\mathcal{P}$ is reducible.

Proof. We can assume without loss of generality that $\mathcal{S}$ is maximal with property $\mathcal{P}$ as if a maximal example is reducible then all semigroups with property $\mathcal{P}$ are reducible.

Assume that $\mathcal{S}$ is irreducible. Let m be the minimal nonzero rank in $\mathcal{S}$. By Lemma 2.3.3, there is an idempotent E of rank m in $\mathcal{S}$.

If $m \geq 2$, by the Finiteness Lemma, $E\mathcal{S}E|_{E\mathcal{V}} = \mathbb{C}\mathcal{G}$ where $\mathcal{G}$ is a finite group. By (i), $\mathcal{G}$ has property $\mathcal{P}$. By (ii), $\mathcal{G}$ is reducible so $E\mathcal{S}E|_{E\mathcal{V}}$ is reducible and $\mathcal{S}$ is reducible by Lemma 2.2.13.

If $m = 1$ then the ideal $\mathcal{J}$ of rank at most one operators in $\mathcal{S}$ is nontrivial. By (i), $\mathcal{J}$ has property $\mathcal{P}$. By (iii), $\mathcal{J}$ is reducible, so $\mathcal{S}$ is reducible by Lemma 2.2.12.

Either way, we reach a contradiction, so $\mathcal{S}$ must be reducible.

□

In the following two sections, we will consider partial spectral mapping properties and use the Finiteness Lemma to prove a number of reducibility results.

In particular, we will consider weakenings of the following property. If A and B are operators in $\mathcal{B}(\mathcal{V})$ and we can order the eigenvalues of A as $\{\alpha_1, \ldots, \alpha_n\}$ and those of B as $\{\beta_1, \ldots, \beta_n\}$ such that, for any polynomial p in two variables, the eigenvalues of $p(A, B)$ are precisely $p(\alpha_i, \beta_i)$ for $1 \leq i \leq n$ we say that the pair $\{A, B\}$ has property P. A family of operators has property P if every pair of operators in it has the property.

Property P is clearly necessary for triangularizability and we'll show that various weakenings of it are sufficient as well. Specifically, instead of requiring an ordering of the eigenvalues, we'll only require that

$$\sigma(p(A, B)) \subseteq \{p(\alpha, \beta) : \alpha \in \sigma(A), \beta \in \sigma(B)\}$$

We'll weaken the property further in each section.

3.3 Subadditive and Sublinear Spectra

In this section, we investigate two partial spectral mapping properties, sublinear and subadditive spectra, to see if they are sufficient conditions for triangularizability (like property P, they are clearly necessary).

3.3.1 Definition

The spectrum is said to be *sublinear* on two operators, A and B, if for every λ in $\mathbb{C}$

$$\sigma(A + \lambda B) \subseteq \sigma(A) + \lambda\sigma(B).$$

It is said to be *subadditive* if the inclusion holds for at least $\lambda = 1$.

Sublinearity and subadditivity of the spectrum on a family of operators $\mathcal{F}$ in $\mathcal{B}(\mathcal{V})$ means that the inclusion holds for every pair of operators A and B in $\mathcal{F}$.

Note that subadditivity is a weakened version of property P based on a single polynomial, $p(x, y) = x + y$, and sublinearity is a weakening by restriction to linear polynomials.

We will show that sublinearity of the spectrum is inherited by quotients, satisfies the properties of the Finiteness Lemma, and is sufficient for triangularizability. Subadditivity will also be sufficient in certain circumstances.

First we want to show that if the sublinearity condition holds for enough values of λ then it holds for all values of λ.

3.3.2 Lemma

Let A and B be operators in $\mathcal{B}(\mathcal{V})$. If

$$\sigma(A + \lambda B) \subseteq \sigma(A) + \lambda\sigma(B)$$

for each λ in an infinite set Λ then the spectrum is sublinear on A and B. In fact, we need only have $|\Lambda| > n^{2n}/(n-1)!$.

Proof. Let $n = \dim(\mathcal{V})$. For each λ in Λ define the eigenvalue set of $A + \lambda B$ to be

$$\mathcal{F}_\lambda = \{(\alpha, \beta, k) : \alpha + \lambda\beta \in \sigma(A + \lambda B), \alpha \in \sigma(A), \beta \in \sigma(B)\},$$

where k is the multiplicity of $\alpha + \lambda\beta$ as an eigenvalue. Now, α and β come from the finite sets $\sigma(A)$ and $\sigma(B)$ respectively and k can be at most the dimension of the space, n. Therefore there are a finite number of possible values for $\mathcal{F}_\lambda$, but an infinite number of λ's. Therefore there is an infinite subset Λ' of Λ so that every λ in Λ' has the same eigenvalue set. Without loss of generality, we take $\Lambda = \Lambda'$ and we call the single remaining eigenvalue set $\mathcal{F}$.

We want to show that $\mathcal{F}$ is the eigenvalue set for every λ in $\mathbb{C}$. Consider the function

$$f(\lambda, x) = \det(A - \lambda B - x).$$

Now f is a polynomial in x and λ. Therefore, f is analytic.

The eigenvalues of multiplicity one of $A + \lambda B$ are those values of x which are solutions to $f(\lambda, x) = 0$, but not $\frac{df}{dx}(\lambda, x) = 0$. The eigenvalues of multiplicity two of $A+\lambda B$ are those which are roots of $f(\lambda, x) = 0$ and $\frac{df}{dx}(\lambda, x) = 0$, but not $\frac{d^2 f}{dx^2}(\lambda, x) = 0$. This holds greater multiplicities in a similar manner.

Take (α, β, k) in $\mathcal{F}$. Then

$$g(\lambda) = f(\lambda, \alpha + \lambda\beta) = \det(A - \lambda B - (\alpha + \lambda\beta))$$

is a polynomial of degree at most n in λ. Since each λ in Λ is a root of g, it has infinitely many roots and is therefore the zero polynomial. Therefore every λ in $\mathbb{C}$ is a root for g. This means that $\alpha + \lambda\beta$ is an eigenvalue of multiplicity at least one for $A + \lambda B$ for all λ in $\mathbb{C}$.

Now, if $k > 1$, consider

$$h(\lambda) = \frac{df}{dx}(\lambda, \alpha + \lambda\beta).$$

Taking a derivative by x won't increase the degree of λ so h is still a polynomial in λ of degree at most n. Since $\alpha + \lambda\beta$ has multiplicity k as an eigenvalue of $A + \lambda B$ for λ in Λ, every such λ is a root of h. h then has infinitely many roots so it's the zero polynomial and $\alpha + \lambda\beta$ is a root for all λ in $\mathbb{C}$.

Repeating this argument we can see that $\alpha + \lambda\beta$ is an eigenvalue of multiplicity at least k for $A + \lambda B$ for all λ in $\mathbb{C}$. As this is true for an arbitrary element of $\mathcal{F}$ and $\mathcal{F}$ consists of n eigenvalues (counting multiplicity), $\mathcal{F}$ is the eigenvalue set for every λ in $\mathbb{C}$.

Note the only time we use the size of Λ is when we claim that $g(\lambda)$, $h(\lambda)$, and any further derivatives of f are actually zero polynomials. Each such polynomial has at most degree n so we need only have $n+1$ roots in order to conclude it's the zero polynomial. The number of roots is determined by how many values of λ share the same eigenvalue set.

First, how many possible eigenvalue sets are there? A has at most n eigenvalues and B has at most n eigenvalues. Therefore there are at most n^2 pairs $\{\alpha, \beta\}$. Then there are at most $\binom{n^2}{n}$ possible ways of selecting n (not necessarily distinct) eigenvalues and this selection determines the k in the eigenvalue set.

In order to have $n+1$ λ guaranteed to share the same eigenvalue set, we must have more than n times the number of possible eigenvalue sets. Calculation gives

$$\begin{aligned} n\binom{n^2}{n} &= \frac{nn^2!}{n!(n^2-n)!} \\ &= \frac{n^2(n^2-1)\cdots(n^2-n+1)}{(n-1)!}n^2(n^2-1)\cdots(n^2-n+1) \\ &\leq \frac{(n^2)^n}{(n-1)!} = \frac{n^{2n}}{(n-1)!}. \end{aligned}$$

Thus having more than $n^{2n}/(n-1)!$ elements in Λ is sufficient to guarantee sublinearity of spectrum.

$\square$

3.3.3 Theorem

Let A and B be operators with a common invariant subspace $\mathcal{M}$ and let $A_0 = A|_{\mathcal{M}}$ and $B_0 = B|_{\mathcal{M}}$. If the spectrum is sublinear on A and B then it is sublinear on A_0 and B_0. If the spectrum is subadditive on A and B and they both have rank at most one then the spectrum is subadditive on A_0 and B_0.

Proof. Assume the spectrum is sublinear on A and B. By the sublinearity of the spectrum, every eigenvalue of $A + \lambda B$ is of the form $\alpha + \lambda\beta$ for α in $\sigma(A)$ and β in $\sigma(B)$. Since $\mathcal{M}$ is a common invariant subspace for A and B, each eigenvalue of $C_\lambda = A_0 + \lambda B_0$ is also of this form. We want to show that for every eigenvalue of C_λ, α is in $\sigma(A_0)$ and β is in $\sigma(B_0)$.

Thanks to Lemma 3.3.2, it suffices to show the sublinearity condition on an infinite set. We define eigenvalue sets

$$\mathcal{E}_\lambda = \{(\alpha, \beta) : \alpha + \lambda\beta \in \sigma(C_\lambda), \alpha \in \sigma(A), \beta \in \sigma(B)\}.$$

Note that we don't have to worry about the multiplicity here. We want to show that $\mathcal{E}_\lambda \subseteq \sigma(A_0) \times \sigma(B_0)$ for infinitely many λ.

As $\mathcal{E}_\lambda$ is a subset of the finite set $\sigma(A) \times \sigma(B)$ there are only finitely many distinct $\mathcal{E}_\lambda$. Since $\mathbb{C}$ is infinite there must be an infinite subset Λ of $\mathbb{C}$ that shares the same $\mathcal{E}_\lambda$. We'll denote this shared set $\mathcal{E}$.

Let $m = \dim(\mathcal{M})$. Then for a fixed (α, β) in $\mathcal{E}$ define

$$p(\lambda) = \det((A_0 - \alpha) + \lambda(B_0 - \beta)) = \det(C_\lambda - (\alpha + \lambda\beta)),$$

which is a polynomial of degree at most m in λ. Since $\alpha + \lambda\beta$ is in $\sigma(C_\lambda)$ for every λ in Λ, all such λ are roots of p. Therefore p has infinitely many roots and is the zero polynomial. In particular, the coefficients of λ^0 and λ^m are zero. We claim these coefficients are $\det(A_0 - \alpha)$ and $\det(B_0 - \beta)$, respectively.

To see this, consider $C_\lambda - (\alpha + \lambda\beta)$ under a basis that makes $A_0 - \alpha$ upper triangular. We now calculate the determinant of $C_\lambda - (\alpha + \lambda\beta)$ using a cofactor expansion. Expand along the first column and, since $A_0 - \alpha$ is upper triangular and every entry of $\lambda(B_0 - \beta)$ contains λ, everything except for the $(1, 1)$ entry will produce only nonconstant terms which we can ignore when looking for the constant coefficient. The portion of the $(1, 1)$ term contributed by $\lambda(B_0 - \beta)$ also contains λ and won't contribute to the constant coefficient. We're left with the $(1, 1)$ entry of $A_0 - \alpha$ times the cofactor of $C_\lambda - (\alpha + \lambda\beta)$ with the first row and column removed. By repeating this argument, we get that the constant coefficient of $p(\lambda)$ is the product of the diagonal entries of $A_0 - \alpha$ or, in other words, $\det(A_0 - \alpha)$. Similarly, the coefficient of λ^m is $\det(B_0 - \beta)$.

Therefore $\det(A_0 - \alpha) = 0$ and $\det(B_0 - \beta) = 0$ so (α, β) is in $\sigma(A_0) \times \sigma(B_0)$. Therefore $\mathcal{E}$ is in $\sigma(A_0) \times \sigma(B_0)$ so the sublinearity condition holds for an infinite number of λ on A_0 and B_0. Therefore the spectrum is sublinear on A_0 and B_0 by Lemma 3.3.2.

Now, assume the spectrum is subadditive on A and B and that A and B have rank one. If $A_0 = 0$ then $\sigma(A_0) = \{0\}$ and $\sigma(A_0 + B_0) = \sigma(B_0) = \sigma(A_0) + \sigma(B_0)$ and similarly for $B_0 = 0$. Also, if $m \leq 1$ then A_0 and B_0 are at most one by one matrices and the result is obvious.

Assume $m \geq 2$ and $A_0, B_0 \neq 0$. Since A has rank one,

$$A = \begin{pmatrix} A_0 & A_1 \\ 0 & 0 \end{pmatrix},$$

so $\sigma(A) = \sigma(A_0)$ as $\sigma(A_0)$ contains zero as A_0 is rank one on a space of dimension at least 2. Similarly, $\sigma(B_0) = \sigma(B)$. Therefore

$$\sigma(A_0 + B_0) \subseteq \sigma(A + B) \subseteq \sigma(A) + \sigma(B) = \sigma(A_0) + \sigma(B_0),$$

so the result is proved.

□

3.3.4 Corollary

Sublinearity of spectrum is inherited by quotients. Subadditivity is inherited by quotients if the operators have rank at most one.

Proof. Let A and B be operators in $\mathcal{B}(\mathcal{V})$ with sublinear spectrum. Let $\mathcal{N}$ and $\mathcal{M}$ be invariant subspaces for A and B with $\mathcal{N}$ properly contained within $\mathcal{M}$. By Theorem 3.3.3, $A|_{\mathcal{M}}$ and $B|_{\mathcal{M}}$ have sublinear spectrum so we can assume $\mathcal{M} = \mathcal{V}$ and decompose A and B with respect to $\mathcal{N}$ and $\mathcal{N}^{\perp}$ as

$$A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} \qquad B = \begin{pmatrix} B_1 & B_2 \\ 0 & B_3 \end{pmatrix}.$$

Since

$$\sigma(A^* + \lambda B^*) = \overline{\sigma(A + \overline{\lambda} B)} \subseteq \overline{\sigma(A)} + \lambda\overline{\sigma(B)} = \sigma(A^*) + \lambda\sigma(B^*),$$

A^* and B^* have sublinear spectrum. Then $A_3^* = (A^*)|_{\mathcal{N}^{\perp}}$ and $B_3^* = (B^*)|_{\mathcal{N}^{\perp}}$ so, by Theorem 3.3.3, A_3^* and B_3^* have sublinear spectrum. But then A_3 and B_3 have sublinear spectrum by reversing the argument for A^* and B^*.

The same argument works for subadditivity by replacing λ with 1, so long as A and B have rank at most one.

□

Having shown inheritability by quotients, we turn to the case of operators of rank at most one.

3.3.5 Lemma

Let $\mathcal{S}$ be an irreducible semigroup in $M_n(\mathbb{C})$ consisting of operators of rank at most one. Then

(i) There exist two bases, $\{e_i\}$ and $\{f_j\}$, of column vectors for $\mathbb{C}^n$ such that the basis

$$\{e_i f_j^* : i, j = 1, \ldots, n\}$$

of $M_n(\mathbb{C})$ is contained in $\mathcal{S}$.

(ii) For each $k \leq n$, there exists a k-dimensional subspace $\mathcal{M}$ of $\mathbb{C}^n$ and a subsemigroup $\mathcal{S}_0$ of $\mathcal{S}$ leaving $\mathcal{M}$ invariant such that $\mathcal{S}_0|_{\mathcal{M}}$ is irreducible.

(iii) In particular, if $k = 2$, there exist numbers $\alpha, \beta, \gamma, \delta$ with $\alpha\delta - \beta\gamma \neq 0$ and $\beta\gamma \neq 0$ such that $\mathcal{S}_0|_{\mathcal{M}}$ is generated by

$$\begin{pmatrix} \alpha & 0 \\ \beta & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & \gamma \\ 0 & \delta \end{pmatrix}$$

with respect to an appropriate basis.

Proof. (i) Since $\mathcal{S}$ is irreducible, it contains a nonzero operator S. Since S has rank one, $S = ef^*$ for some nonzero column vectors e and f in $\mathbb{C}^n$. As $\mathcal{S}$ is irreducible, $\mathcal{S}e$ must contain a basis $\{e_i\}$ for $\mathbb{C}^n$. Also, $\mathcal{S}^*$ is irreducible, so $\mathcal{S}^* f$ must contain a basis $\{f_j\}$ for $\mathbb{C}^n$. Then $\mathcal{S}$ contains every operator of the form $\mathcal{S}ef^*\mathcal{S}$ which includes every $e_i f_j^*$.

(ii) Let $\mathcal{M} = \operatorname{span}\{e_1, \ldots, e_k\}$ and let $\mathcal{S}_0 = \{e_i f_j^* : i = 1, \ldots, k, j = 1, \ldots, n\}$. Then for any S in $\mathcal{S}_0$, $\operatorname{ran}(\mathcal{S})$ is contained in $\mathcal{M}$ so $\mathcal{M}$ is an invariant subspace for $\mathcal{S}_0$. We'll show $\mathcal{S}_0|_{\mathcal{M}}$ is irreducible by showing it contains a basis for $M_k(\mathbb{C})$.

Every f_j can be written as $f_j = g_j + h_j$ with g_j in $\mathcal{M}$ and h_j in $\mathcal{M}^{\perp}$. Since $\{f_j\}$ is a basis for $\mathbb{C}^n$, $\{g_j\}$ must be a spanning set for $\mathcal{M}$. Write $\mathcal{S}_0$ with respect to the basis $\{e_i\}$. Then for $e_i f_j^*$ in $\mathcal{S}_0$, $e_i f_j^*|_{\mathcal{M}} = e_i g_j^*|_{\mathcal{M}}$. As $\{e_i\}$ and $\{g_j\}$ are bases for $\mathcal{M}$, $e_i g_j^*|_{\mathcal{M}}$ is a basis for $M_k(\mathbb{C})$ contained in $\mathcal{S}_0|_{\mathcal{M}}$. Therefore, $\mathcal{S}_0|_{\mathcal{M}}$ is irreducible.

(iii) If $k = 2$ then $\mathcal{S}_0$ contains $\{e_i f_j^* : i, j = 1, 2\}$. Let $e_1 = \binom{\alpha}{\beta}$ and $e_2 = \binom{\gamma}{\delta}$ with respect to $\{f_1, f_2\}$. Since e_1 and e_2 are linearly independent $\alpha\delta - \beta\gamma \neq 0$. We can assume $\beta\gamma \neq 0$ since otherwise $\alpha\delta \neq 0$ and we can simply reverse the roles of e_1 and e_2. Then we know $\mathcal{S}_0$ contains

$$S = \begin{pmatrix} \alpha & 0 \\ \beta & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 0 & \gamma \\ 0 & \delta \end{pmatrix}.$$

Let $\mathcal{S}_1$ be the subsemigroup of $\mathcal{S}_0$ generated by S and T. Then

$$\delta S - TS = \begin{pmatrix} \alpha\delta - \beta\gamma & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \alpha T - ST = \begin{pmatrix} 0 & 0 \\ 0 & \alpha\delta - \beta\gamma \end{pmatrix}.$$

Since $\beta \neq 0$, $\gamma \neq 0$, and $\alpha\delta - \beta\gamma \neq 0$ we have that S, T, TS, and ST span $\mathcal{M}_2(\mathbb{C})$. Therefore $\mathcal{S}_1$ is an irreducible subsemigroup of $\mathcal{S}$.

□

3.3.6 Theorem

Let $\mathcal{S}$ be a semigroup of operators of rank at most one with subadditive spectrum. Then $\mathcal{S}$ is triangularizable.

Proof. By Corollary 3.3.4 and the Triangularization Lemma (2.2.3), it's enough to show that $\mathcal{S}$ is reducible. Assume $\mathcal{S}$ is irreducible. By Lemma 3.3.5, we can find a subsemigroup $\mathcal{S}_0$ of $\mathcal{S}$ and an $\mathcal{S}_0$-invariant subspace, $\mathcal{M}$, of dimension 2 such that $\mathcal{S}_0|_{\mathcal{M}}$ is irreducible and generated by

$$S = \begin{pmatrix} \alpha & 0 \\ \beta & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 0 & \gamma \\ 0 & \delta \end{pmatrix},$$

with $\alpha\delta - \beta\gamma \neq 0$ and $\beta\gamma \neq 0$.

Then subadditivity applies to S and T so $\sigma(S+T)$ must be contained in $\{0, \alpha, \delta, \alpha+\delta\}$. Now, the characteristic equation of $S+T$ is

$$\det(\lambda - (S+T)) = \lambda^2 - (\alpha+\delta)\lambda + \alpha\delta - \beta\gamma = 0.$$

Substituting α or δ for λ gives $\beta\gamma = 0$ while substituting 0 or $\alpha + \delta$ for λ gives $\alpha\delta - \beta\gamma = 0$. These are both contradictions so $\mathcal{S}$ is reducible and thus triangularizable.

□

3.3.7 Example

The assertion of Theorem 3.3.6 does not hold without the restriction on rank.

Proof. Take $\{E_{ij}\}$ to be the standard basis for $\mathcal{M}_2(\mathbb{C})$ and take $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Then the following matrices on $\mathbb{C}^4$ form a semigroup $\mathcal{S}$:

$$E_{11} \oplus I, \quad E_{22} \oplus I, \quad O \oplus I, \quad E_{12} \oplus J, \quad E_{21} \oplus J \quad , 0 \oplus J$$

Let $\mathcal{M}$ be the space acted on by the E_{ij}'s above. Then $\mathcal{M}$ is clearly invariant for $\mathcal{S}$ and as $\mathcal{S}|_{\mathcal{M}}$ contains the standard basis for $\mathcal{M}_2(\mathbb{C})$, it is irreducible. Therefore $\mathcal{S}$ is not triangularizable.

Since triangularizability implies subadditivity of spectrum and since every commuting pair is triangularizable (2.2.4), we need only consider the noncommuting pairs, of which there are five.

We have $\sigma(E_{11} \oplus I) = \sigma(E_{22} \oplus I) = \{0, 1\}$ while $\sigma(E_{12} \oplus J) = \sigma(E_{21} \oplus J) = \{0, 1, -1\}$. Now,

$$\sigma(E_{11} \oplus I + E_{12} \oplus J) = \{0, 1, 2\} \subset \sigma(E_{11} \oplus I) + \sigma(E_{12} \oplus J) = \{0, 1, -1, 2\},$$
$$\sigma(E_{11} \oplus I + E_{21} \oplus J) = \{0, 1, 2\} \subset \sigma(E_{11} \oplus I) + \sigma(E_{21} \oplus J = \{0, 1, -1, 2\},$$

so the spectrum is subadditive on these pairs. Similarly, the two pairs involving $E_{22} \oplus I$ have subadditive spectrum. Finally,

$$\sigma(E_{12} \oplus J + E_{21} \oplus J) = \{1, -1, 2, -2\} \subset \sigma(E_{12} \oplus J) + \sigma(E_{21} \oplus J) = \{0, 1, -1, 2, -2\},$$

so $\mathcal{S}$ has subadditive spectrum.

□

This concludes our discussion of subadditivity. For sublinearity, we require the following result from group theory.

3.3.8 Lemma

Every minimal nonabelian finite group $\mathcal{G}$ (i.e., group such that every proper subgroup is abelian) is solvable. In particular, such a group contains a normal subgroup of prime index.

Proof. Assume that the theorem is not true and let $\mathcal{G}$ be a counterexample of minimal order. We make a few preliminary claims.

First, we claim $\mathcal{G}$ is simple. Assume otherwise. Then there is a nontrivial normal subgroup $\mathcal{H}$ of $\mathcal{G}$. As $\mathcal{G}$ is minimal nonabelian, $\mathcal{H}$ is abelian and is thus solvable.

Every maximal subgroup of $\mathcal{G}/\mathcal{H}$ is of the form $\mathcal{M}/\mathcal{H}$ where $\mathcal{M}$ is a maximal subgroup of $\mathcal{G}$. Since $\mathcal{G}$ is minimal nonabelian we have that all such $\mathcal{M}$ are abelian and thus $\mathcal{M}/\mathcal{H}$ is abelian. Therefore $\mathcal{G}/\mathcal{H}$ is either abelian (and thus solvable) or minimal nonabelian and thus solvable since $\mathcal{G}$ was chosen to be the counterexample of minimal order. Then $\mathcal{H}$ and $\mathcal{G}/\mathcal{H}$ are solvable so $\mathcal{G}$ is solvable. This is a contradiction so $\mathcal{G}$ must be simple.

Next, we claim that if $\mathcal{M}_1 \neq \mathcal{M}_2$ are maximal subgroups of $\mathcal{G}$ then $\mathcal{M}_1 \cap \mathcal{M}_2 = \{I\}$. Assume otherwise. Then let $\mathcal{R} = \mathcal{M}_1 \cap \mathcal{M}_2 \neq \{I\}$. Consider the normalizer of $\mathcal{R}$ in $\mathcal{G}$:

$$\mathcal{N} = \{G \in \mathcal{G} \mid G^{-1}\mathcal{R}G = \mathcal{R}\}.$$

Since $\mathcal{G}$ is minimal nonabelian, $\mathcal{M}_1$ and $\mathcal{M}_2$ are abelian. Then $\mathcal{M}_1$ and $\mathcal{M}_2$ commute with $\mathcal{R}$ so $\mathcal{M}_1 \cup \mathcal{M}_2 \subseteq \mathcal{N}$. Since $\mathcal{G}$ is simple and $\mathcal{R} \neq \{I\}$, $\mathcal{R}$ can't be normal so $\mathcal{N} \neq \mathcal{G}$. Since $\mathcal{N}$ is a proper subgroup, it must be abelian so $\mathcal{M}_1 \cup \mathcal{M}_2$ must be abelian. As $\mathcal{M}_1$ and $\mathcal{M}_2$ are distinct maximal subgroups, they generate $\mathcal{G}$ so $\mathcal{G}$ is abelian, which is a contradiction. Therefore $\mathcal{M}_1 \cap \mathcal{M}_2 = \{I\}$.

We claim $\mathcal{G}$ has at least two non-conjugate maximal subgroups. Let $|\mathcal{G}| = p_1^{r_1} p_2^{r_2} \dots p_t^{r_t}$. Since $\mathcal{G}$ isn't solvable it isn't a p-group so $t \geq 2$ and every Sylow p_i-group of $\mathcal{G}$ is contained in a maximal subgroup of $\mathcal{G}$. If a maximal subgroup $\mathcal{M}$ contains a Sylow p_i-subgroup (which has order $p_i^{r_i}$) then $p_i^{r_i}$ divides $|\mathcal{M}|$. Thus such an $\mathcal{M}$ couldn't contain a p_i-subgroup for all $1 \leq i \leq t$ as it would then be the entire group. Therefore $\mathcal{G}$ must have at least two non-conjugate maximal subgroups.

Let $\{\mathcal{M}_1, \dots, \mathcal{M}_k\}$ be a maximal set of mutually nonconjugate maximal subgroups of $\mathcal{G}$. Fix $1 \leq i \leq k$ and consider the groups conjugate to $\mathcal{M}_i$. They will also be maximal subgroups of $\mathcal{G}$. If $x, y \in \mathcal{G}$ and $x^{-1}\mathcal{M}_i x = y^{-1}\mathcal{M}_i y$ then $yx^{-1}\mathcal{M}_i xy^{-1} = \mathcal{M}_i$. Let $z = xy^{-1}$ so $z^{-1}\mathcal{M}_i z = \mathcal{M}_i$.

If the subgroup generated by $\mathcal{M}_i$ and z were all of $\mathcal{G}$ then $\mathcal{M}_i$ would be normal, but $\mathcal{G}$ is simple so this can't happen. Since $\mathcal{M}_i$ is maximal, $\mathcal{M}_i$ and z must generate $\mathcal{M}_i$ so $z \in \mathcal{M}_i$. Then, as $z = xy^{-1}$, x and y are in the same coset of $\mathcal{M}_i$. So $x^{-1}\mathcal{M}_i x = y^{-1}\mathcal{M}_i y$ if and only if x and y are in the same coset of $\mathcal{M}_i$. Therefore the number of distinct maximal subgroups conjugate to $\mathcal{M}_i$ is $|\mathcal{G}| \,/\, |\mathcal{M}_i|$.

Such maximal subgroups must cover all of $\mathcal{G}$ and since maximal subgroups intersect trivially we have that

$$|\mathcal{G}| = 1 + \sum_{i=1}^{k} (|\mathcal{M}_i| - 1) \frac{|\mathcal{G}|}{|\mathcal{M}_i|} = 1 + k\,|\mathcal{G}| - \sum_{i=1}^{k} \frac{|\mathcal{G}|}{|\mathcal{M}_i|}.$$

The $\mathcal{M}_i$ are maximal so they're nontrivial. Therefore $|\mathcal{M}_i| \geq 2$ and $|\mathcal{G}| \geq 1 + \frac{k}{2}\,|\mathcal{G}|$. Thus $k < 2$ which is a contradiction as there are at least two non-conjugate maximal subgroups. Therefore all such groups are solvable as originally desired. Thus the commutator subgroup $\mathcal{G}$ is proper. Let $\mathcal{G}_0$ be a maximal abelian subgroup containing the commutator subgroup of $\mathcal{G}$. It's automatically normal so we need only show it has prime index. Let $x \in \mathcal{G} \backslash \mathcal{G}_0$. By the minimality of $\mathcal{G}$, $\mathcal{G}$ is generated by x and $\mathcal{G}_0$. Let k be the smallest power of x such that $x^k \in \mathcal{G}_0$. Again, since $\mathcal{G}$ was chosen to be minimal, k is a prime. (If $k = pn$ then x^n and $\mathcal{G}_0$ would generate a strictly smaller group that x and $\mathcal{G}_0$.)

$\square$

3.3.9 Lemma

Let p be a prime. If $A, B \in M_p(\mathbb{C})$,

$$A = \begin{pmatrix} 0 & & & 1 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix},$$

and B is nonscalar and diagonal, then the pair $\{A, B\}$ is irreducible.

Proof. As $\{A, B\}$ is reducible if and only if the algebra $\mathcal{A}$ generated by A and B is reducible, we'll show that $\mathcal{A}$ is irreducible.

If $\lambda \neq 0$ is a diagonal entry of B then $B - \frac{1}{\lambda} B^2$ will have the same zeroes on the diagonal as B, plus all those entries of B that were λ are now zero. All other nonzero entries will still be nonzero. By repeating this argument, we can create a scalar multiple of a nontrivial diagonal projection and by normalizing we get a nontrivial diagonal projection in $\mathcal{A}$.

We claim that if $\mathcal{A}$ contains a diagonal rank one projection E then $\mathcal{A}$ is irreducible. Let $\{e_i\}$ be the basis for $\mathbb{C}^p$ relative to which A and B have the above forms. Say E is the projection onto the span of e_k for some $1 \leq k \leq p$. Consider an element of the form $A^{-i}EA^i$ with $1 \leq i \leq p$. It's the projection onto e_l where $l + i \equiv k \mod p$ and $1 \leq l \leq p$.

The invariant subspaces of a projection on e_j are those spaces that contain e_j and those spaces that are subsets of $\{e_j\}^\perp$. Thus the invariant subspaces of

$$\mathcal{B} = \{A^{-i}EA^i : 1 \leq i \leq p\} \subseteq \mathcal{A}$$

are precisely those spaces that contain some subset of the e_j's and are perpendicular to all the others. However, if $\mathcal{M}$ is an invariant subspace for A and $\mathcal{M}$ contains e_j then $\mathcal{M}$ will contain $e_1, \ldots, e_p$ and thus $\mathcal{M} = \mathbb{C}^p$. Then the only common subspaces between A and $\mathcal{B}$ are the trivial subspaces so $\mathcal{A}$ is irreducible.

We want to show that we have such a rank one diagonal projection in $\mathcal{A}$. Let E be a diagonal projection in $\mathcal{A}$ of minimal positive rank r. Since B isn't scalar, we know $r < p$.

For any i and j, we have that $A^{-i}EA^i$ and $A^{-j}EA^j$ will still be diagonal projections of rank r. $(A^{-i}EA^i)(A^{-j}EA^j)$ will also be a diagonal projection and thus of rank either 0 or r by minimality. The only way that two diagonal projections of rank r can multiply to make another rank r projection is if they're equal.

If $r > 1$ then it's impossible for there to be p diagonal projections of rank r on a p dimensional space that are mutually orthogonal. Therfore there must be an $i \neq j$ such that $A^{-i}EA^i = A^{-j}EA^j$. Since $A^p = I$ there is a $1 \leq k < p$ such that $A^{-k}EA^k = E$. Thus $A^{-sk}EA^{sk} = E$ for all integers s. As p is prime, there is an integer s such that $sk \equiv 1 \mod p$ and therefore $A^{-1}EA = E$. But this forces E to be either 0 or I and E has postive rank $r < p$. This is a contradiction, so $r = 1$ and $\{A, B\}$ is irreducible as claimed

$\square$

3.3.10 Lemma

Let $\mathcal{G}$ be a minimal nonabelian finite group of operators on $\mathbb{C}^n$. Then there exist primes p and q, not necessarily distinct, and a p-dimensional subspace $\mathcal{M}$ of $\mathbb{C}^n$

invariant under $\mathcal{G}$ such that $\mathcal{G}|_{\mathcal{M}}$ is, after a similarity, generated by two operators of the form

$$A = \alpha \begin{pmatrix} 0 & & & 1 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix} \qquad B = \beta \begin{pmatrix} \theta_1 & & & \\ & \theta_2 & & \\ & & \ddots & \\ & & & \theta_p \end{pmatrix},$$

where B is nonscalar and $\theta_i^q = 1$ for all i. *Furthermore, $\alpha^{p^r} = \beta^{q^s} = 1$ for some nonnegative integers r and s.*

Proof. Since $\mathcal{G}$ is finite we can assume it's a unitary group by Theorem 2.3.1. Let $\mathcal{H}$ be the normal subgroup of index p from Lemma 3.3.8. Then $\mathcal{H}$ is a commutative group of unitaries as $\mathcal{G}$ is minimal nonabelian. As $\mathcal{H}$ is commutative, it's triangularizable and similar to a unitary group. Therefore, it's self-adjoint and, as we saw in Corollary 3.1.10, we can assume $\mathcal{H}$ is diagonal. Since $\mathcal{H}$ has index p in $\mathcal{G}$, we can take $G \in \mathcal{G}/\mathcal{H}$ such that $G^p \in \mathcal{H}$.

We can then decompose $\mathbb{C}^n = \mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_r$ where each $\mathcal{M}_i$ is a maximal subspace of $\mathbb{C}^n$ invariant under $\mathcal{H}$ such that $\mathcal{H}|_{\mathcal{M}_i}$ consists of scalars. Since $\mathcal{H}$ doesn't consist entirely of scalars (if it did then $\mathcal{G}$ would be commutative) we have that $r \geq 2$.

Fix H in $\mathcal{H}$. Since $\mathcal{H}$ is normal, $G\mathcal{H} = \mathcal{H}G$ and there is an H' in $\mathcal{H}$ such that $GH = H'G$. For any x in $\mathcal{M}_i$,

$$HGx = GH'x = G(\lambda x) = \lambda Gx,$$

where $H'|_{\mathcal{M}_i} = \lambda I|_{\mathcal{M}_i}$. Therefore $G\mathcal{M}_i$ is invariant for $\mathcal{H}$ and $\mathcal{H}|_{G\mathcal{M}_i}$ consists of scalars.

By definition, $G\mathcal{M}_i$ must be contained within some $\mathcal{M}_j$. If $G\mathcal{M}_i$ isn't maximal with the scalar property, then it is contained inside a larger subspace $\mathcal{N}$ with the scalar property. By a similar argument, $G^{-1}\mathcal{N}$ is a subspace with the scalar property that properly contains $\mathcal{M}_i$. This contradicts the maximality of $\mathcal{M}_i$. Therefore $G\mathcal{M}_i = \mathcal{M}_j$. Since G is invertible, distinct i's produce distinct j's. Therefore G induces a permutation τ such that $G\mathcal{M}_i = \mathcal{M}_{\tau(i)}$ for all i. Since $\mathcal{G}$ isn't commutative and is generated by $\mathcal{H}$ and G, there is an i such that $i \neq \tau(i)$.

Let $x \in \mathcal{M}_i$ and consider the subspace, $\mathcal{M}$, of $\mathbb{C}^n$ spanned by $\{x, Gx, \ldots, G^{p-1}x\}$. Then $\mathcal{M}$ is invariant under $\mathcal{H}$ as each $G^j x \in \mathcal{M}_{\tau^j(i)}$ is an eigenvector for every element of $\mathcal{H}$. Let $A = G|_{\mathcal{M}}$ and $\mathcal{H}_0 = \mathcal{H}|_{\mathcal{M}}$.

Since x and Gx come from different subspaces, $\mathcal{M}_i$ and $\mathcal{M}_{\tau(i)}$, at least one element of $\mathcal{H}$ isn't a scalar on $\mathcal{M}$ as this would contradict the maximality of the $\mathcal{M}_i$. Also, the vectors $\{x, Gx, \ldots, G^{p-1}x\}$ are a basis as the subspaces $G^j\mathcal{M}_i$ must be distinct. If they weren't, we'd have $G^k\mathcal{M}_i = \mathcal{M}_i$ for some k. Then $G^{ks}\mathcal{M}_i = \mathcal{M}_i$ for all $s \in \mathbb{N}$. As p is prime there is an s such that $ks \equiv 1 \mod p$ and, as $G^p \in \mathcal{H}$ and $\mathcal{M}_i$ is invariant for $\mathcal{H}$, $G^k\mathcal{M}_i = \mathcal{M}_i$. This is a contradiction, so the spaces are distinct.

Now each G^jx is an eigenvector for every element of $\mathcal{H}$ so $\mathcal{H}_0$ is diagonal with respect to this basis. And

$$A = \begin{pmatrix} 0 & & & \lambda \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}$$

as $Ax = Gx, \ldots, A^{p-1}x = G^{p-1}x$ and $G^p \in \mathcal{H}$. Therefore $A^p \in \mathcal{H}_0$ and A^p acts as a scalar on $x \in \mathcal{M}_i$.

$\mathcal{H}_0$ has nonscalar members as $\mathcal{M}$ was constructed from at least two distinct maximal scalar subspaces. Let B be a nonscalar in $\mathcal{H}_0$. Since $\mathcal{H}_0$ is a finite group, $B^n = I$ for some n. Let q be the smallest power such that B^q is scalar (such a q exists and is at most n). By taking powers of B, we can assume q is a prime. So $B^q = \mu I$ and, by taking $\beta = \mu$, B has the appropriate form.

Since $\mathcal{G}$ is a group, A is invertible and $\lambda \neq 0$. Take α to be a p^{th} root of λ. Then A is similar to αP where P is the invertible right shift. The similarity is give by $\operatorname{diag}(1, \alpha, \ldots, \alpha^{p-1})$, so it doesn't change B. Therefore A and B are as required. The group $\mathcal{G}_0$ generated by A and B is irreducible by Lemma 3.3.9. As $\mathcal{G}$ is minimal we get $\mathcal{G}_0 = \mathcal{G}|_{\mathcal{M}}$.

Finally, let the order of α be mp^r where p doesn't divide m. As p is prime, there is an integer t such that $mt \equiv 1 \mod p$. Then A^{mt} has the same form as A except that α has been replaced by $\alpha_1 = \alpha^{mt}$ and thus $\alpha_1^{p^r} = 1$. By minimality, $\alpha_1 = \alpha$ and $m = 1$. We can apply a similar argument to β.

$\square$

3.3.11 Theorem

A finite group of matrices with sublinear spectrum is abelian (and thus diagonalizable)

Proof. Assume that there are finite groups with sublinear spectrum that are not abelian. Consider a minimal counterexample. It's a minimal nonabelian finite group so, by Lemma 3.3.10, it has a restriction to a group $\mathcal{G}$ generated by operators A and B as in the lemma and acting on a space of dimension p, a prime.

By Theorem 3.3.3, the spectrum is sublinear on $\mathcal{G}$. Then the spectrum is sublinear on $\mathbb{C}\mathcal{G}$ as well. Therefore we can take $\alpha = \beta = 1$ without loss of generality.

If $p = q = 2$ then, since B is not scalar, we have, possibly after scaling, that

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad\qquad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

so $\sigma(A) = \{1, -1\}$ and $\sigma(B) = \{1, -1\}$. By sublinearity, $\sigma(A + \lambda B)$ is contained in $\{1 + \lambda, 1 - \lambda, -1 + \lambda, -1 - \lambda\}$. However,

$$xI - (A + \lambda B) = \begin{pmatrix} x - \lambda & -1 \\ -1 & x + \lambda \end{pmatrix},$$

so its characteristic equation is $x^2 - (\lambda^2 + 1) = 0$. But letting x take on any of the four values from sublinearity gives $\pm 2\lambda = 0$ which only holds for $\lambda = 0$ which contradicts sublinearity.

Assume at least one of p or q is not 2. We can scale B so that $\det(B) = 1$. Then $\det(AB) = \det(A)$ and

$$AB = \begin{pmatrix} 0 & & & \theta_p \\ \theta_1 & \ddots & & \\ & \ddots & \ddots & \\ & & \theta_{p-1} & 0 \end{pmatrix}.$$

By applying a diagonal similarity, we can show AB is similar to A. So $\sigma(AB) = \sigma(A) = \{z : z^p = 1\}$. Also

$$xI - (AB + \lambda A) = \begin{pmatrix} x & & & -(\lambda + \theta_p) \\ -(\lambda + \theta_1) & \ddots & & \\ & \ddots & \ddots & \\ & & -(\lambda + \theta_{p-1}) & x \end{pmatrix}.$$

By performing a sequence of row reductions we get that

$$xI - (AB + \lambda A) \sim \begin{pmatrix} x & & & & -(\lambda+\theta_p) \\ & x^2 & & & (-1)^3(\lambda+\theta_p)(\lambda+\theta_1) \\ & & \ddots & & \vdots \\ & & & x^{p-1} & (-1)^{2p-3}(\lambda+\theta_p)\prod_{i=1}^{p-2}(\lambda+\theta_i) \\ & & & & x^p + (-1)^{2p-1}\prod_{i=1}^{p}(\lambda+\theta_i) \end{pmatrix}.$$

The characteristic equation of $AB + \lambda A$ is thus

$$x^p + (-1)^{2p-1}\prod_{i=1}^{p}(\lambda+\theta_i) = 0,$$

since we can ignore the excess multiples of x (as 0 is in the spectrum of $AB + \lambda A$ if and only if it is not invertible which happens if and only if some $\lambda + \theta_i = 0$ in which case 0 will be the only root of the above equation). So the spectrum of $AB + \lambda A$ consists of the p^{th} roots of $\prod_{i=1}^{p}(\lambda+\theta_i)$.

But sublinearity says that $\sigma(AB+\lambda A)$ is contained within $\{\psi+\lambda\phi : \psi^p = \phi^p = 1\}$. There are finitely many p^{th} roots of unity so there are a finite number of elements of the form $\psi + \lambda\phi$ with $\psi^p = \phi^p = 1$. Therefore there is an element $\psi + \lambda\phi$ that is in $\sigma(AB + \lambda A)$ for infinitely many λ. This value is a p^{th} root for $\prod_{i=1}^{p}(\lambda+\theta_i)$ for all these λ, so

$$\prod_{i=1}^{p}(\lambda+\theta_i) = (\psi+\lambda\phi)^p = (\frac{\psi}{\phi}+\lambda)^p$$

for infinitely many λ. But that means that a polynomial in λ of degree p has infinitely many roots, so it must be the zero polynomial and therfore 0 on all values of λ. In particular, for every i, $\lambda = -\theta_i$ we see

$$0 = (\frac{\psi}{\phi} - \theta_i)^p,$$

so $\theta_i = \frac{\psi}{\phi}$ for all i which contradicts that B isn't scalar. Therefore no such counterexample exists so such a group is abelian.

Since it's a finite group, it's bounded and simultaneously similar to a unitary group by Theorem 2.3.1. Unitary groups are self-adjoint, abelian groups are triangularizable, and triangularizable self-adjoint groups are diagonalizable by Corollary 3.1.10.

□

We're now ready to show that sublinearity of the spectrum is a sufficient condition for triangularizability.

3.3.12 Theorem

Every semigroup of matrices with sublinear spectrum is triangularizable.

Proof.

By Corollary 3.3.4, sublinearity of spectrum is inherited by quotients so, by the Triangularization Lemma (2.2.3), we need only show reducibility.

If $\mathcal{S}$ has sublinear spectrum then so does $\mathbb{C}\mathcal{S}$. Also, for any A and B in $\overline{\mathcal{S}}$ and $\{A_n\}$ and $\{B_n\}$, sequences in $\mathcal{S}$ converging to A and B respectively,

$$\sigma(A+\lambda B) = \lim_{n\longrightarrow\infty} \sigma(A_n+\lambda B_n) \subseteq \lim_{n\longrightarrow\infty} \sigma(A_n) + \lambda \lim_{n\longrightarrow\infty} \sigma(B_n) = \sigma(A) + \lambda\sigma(B)$$

by Lemma 2.4.5. Thus $\overline{\mathcal{S}}$ has sublinear spectrum. Finally, for any ring automorphism Φ induced by a field automorphism ϕ and A and B in $\mathcal{S}$, we see that

$$\begin{aligned}
\sigma(\Phi(A)+\lambda\Phi(B)) &= \sigma(\Phi(A+\phi^{-1}(\lambda)B)) \\
&= \phi(\sigma(A+\phi^{-1}(\lambda)B)) \\
&\subseteq \phi(\sigma(A)+\phi^{-1}(\lambda)\sigma(B)) \\
&= \phi(\sigma(A)) + \lambda\phi(\sigma(B)) \\
&= \sigma(\Phi(A)) + \lambda\sigma(\Phi(B))
\end{aligned}$$

by Lemma 2.5.3. Therefore $\Phi(\mathcal{S})$ also has sublinear spectrum and sublinear spectrum satisfies the requirements of the Finiteness Lemma (3.2.2).

In order to show reducibility we need only show that sublinearity of spectrum meets the requirements of Lemma 3.2.3.

For (i), any ideal of a semigroup $\mathcal{S}$ with sublinear spectrum is a subset of $\mathcal{S}$ and thus has sublinear spectrum. For any minimal nonzero idempotent E in $\mathcal{S}$ and A and B in $\mathcal{S}$, EAE and EBE are in $\mathcal{S}$ so the spectrum is sublinear on them. By Theorem 3.3.3, the spectrum is sublinear on $EAE|_{E\mathcal{V}}$ and $EBE|_{E\mathcal{V}}$ since $E\mathcal{V}$ is an invariant subspace for EAE and EBE. Therefore, spectrum is sublinear on $E\mathcal{S}E|_{E\mathcal{V}}$.

(ii) is proved by Theorem 3.3.11 and (iii) is proved by Theorem 3.3.6 as sublinearity implies subadditivity. Therefore the requirements of Lemma 3.2.3 are met and $\mathcal{S}$ is reducible. Therefore $\mathcal{S}$ is triangularizable.

□

3.3.13 Corollary

If $\mathcal{S}$ is a self-adjoint semigroup of matrices with sublinear spectrum, then $\mathcal{S}$ is diagonalizable (and thus abelian).

Proof. By Theorem 3.3.12, $\mathcal{S}$ is triangularizable. As we saw in Corollary 3.1.10, a triangularizable, self-adjoint family is diagonalizable.

□

3.3.14 Corollary

If $\mathcal{G}$ is a unitary group with sublinear spectrum then $\mathcal{G}$ is abelian

Proof. For any A in $\mathcal{G}$, $A^* = A^{-1}$ which is in $\mathcal{G}$ as $\mathcal{G}$ is a group. Therefore $\mathcal{G}$ is abelian by Corollary 3.3.13.

□

3.3.15 Corollary

If every pair of operators in a semigroup $\mathcal{S}$ is triangularizable, then so is $\mathcal{S}$ itself.

Proof. For every A and B in $\mathcal{S}$, $\{A, B\}$ is triangularizable. Therefore spectrum is sublinear on every pair A and B from $\mathcal{S}$ so spectrum is sublinear on $\mathcal{S}$. Therefore $\mathcal{S}$ is triangularizable by Theorem 3.3.12.

□

3.3.16 Corollary

The following conditions are mutually equivalent for a semigroup $\mathcal{S}$ of operators in $\mathcal{B}(\mathcal{V})$:

(i) $\mathcal{S}$ is triangularizable.

(ii) for all integers m, scalars $\lambda_1, \ldots, \lambda_m$ and members $S_1, \ldots, S_m$ of $\mathcal{S}$,

$$\sigma(\lambda_1 S_1 + \cdots + \lambda_m S_m) \subseteq \lambda_1 \sigma(S_1) + \cdots + \lambda_m \sigma(S_m).$$

(iii) $\mathcal{S}$ has sublinear spectrum.

(iv) $\sigma(A + \lambda B) \subseteq \sigma(A) + \lambda\sigma(B)$ for all integers λ and all pairs A and B in $\mathcal{S}$.

(v) for every pair A and B in $\mathcal{S}$, there are infinitely many values of λ for which

$$\sigma(A + \lambda B) \subseteq \sigma(A) + \lambda\sigma(B).$$

(vi) for $n = \dim(\mathcal{V})$ and for every pair A and B in $\mathcal{S}$, there are more than $L = n^{2n}/(n-1)!$ values of λ for which

$$\sigma(A + \lambda B) \subseteq \sigma(A) + \lambda\sigma(B).$$

Proof. (i) clearly implies (ii) and (ii) implies (iii) by taking $m = 2$ and $\lambda_1 = 1$. (iii) implies (iv) as taking λ to be an integer is a restriction on the sublinearity condition. (iv) implies (v) as there are infinitely many integers and (v) obviously implies (vi). Also, (iii) implies (i) by Theorem 3.3.12.

To complete the proof, we'll show (vi) implies (iii). But this follows directly from the second part of Lemma 3.3.2, so we're done.

□

3.4 Polynomial Conditions on Spectra

In this section we consider the weakening of property P when it holds for a single polynomial $p(x, y)$ which vanishes whenever x and y commute. If A and B are simultaneously upper triangularizable then $p(A, B)$ would necessarily be nilpotent.

We investigate whether this nilpotence is a sufficient condition when we restrict p to being linear in the second variable.

3.4.1 Definition

Let $g(x) = \sum_{j=0}^{m} a_j x^j$. Then we define a noncommutative, homogeneous polynomial f_g by $f_g(x,y) = \sum_{j=0}^{m} a_j x^j y x^{m-j}$. When we define g as above, there is an implicit assumption that $a_m \neq 0$.

3.4.2 Definition

Given a polynomial $g(x)$, we say that f_g is *nilpotent* on a family $\mathcal{F}$ of operators in $\mathcal{B}(\mathcal{V})$ if $f_g(S,T)$ is nilpotent for every S and T in $\mathcal{F}$.

Unfortunately, there are many polynomials that are nilpotent, or even vanish, on irreducible groups. For instance, there are irreducible groups whose elements have order 1 or p. Any polynomial divisible by $x^p - 1$ will vanish on such groups. The results in this section show that this is the only real obstacle to triangularizability.

3.4.3 Lemma

Let p be a prime number and let $\{e_1, \ldots, e_p\}$ be a basis for $\mathbb{C}^p$. Let T be the invertible right shift defined by

$$Te_i = e_{i+1} \text{ for } i < p \text{ and } Te_p = e_1.$$

If $\mathcal{E}$ is any proper, nonempty subset of $\{1, \ldots, p\}$ then $e = \sum_{j \in \mathcal{E}} e_j$ is a cyclic vector for T (e is a cyclic vector if $\{T^n e : n \in \mathbb{N}\}$ spans $\mathbb{C}^p$).

Proof. Assume e is not a cyclic vector. Then the set $\{Te, T^2e, \ldots, T^pe\}$ must be linearly dependent as otherwise it would form a basis for $\mathbb{C}^p$. Note that $T^p = I$ so $T^pe = e$. By linear dependence, there exist $\alpha_0, \ldots, \alpha_{p-1} \in \mathbb{C}$, not all zero such that $\alpha_0 e + \alpha_1 Te + \cdots + \alpha_{n-1} T^{p-1} e = 0$. Then the polynomial $\phi(x) = \sum_{i=0}^{p-1} a_i x^j$ is such that $\phi(T)e = 0$, but ϕ is not divisible by the minimal polynomial of T, namely $x^p - 1$, since the degree of ϕ is less than p.

Therefore, in order to show that e is a cyclic vector for T, we'll show that any polynomial ϕ with $\phi(T)e = 0$ is divisible by $x^p - 1$, the minimal polynomial of T.

By performing a cyclic permutation on the basis, we can assume that $1 \notin \mathcal{E}$. Let the elements of $\mathcal{E}$ be $2 \leq r_1 < r_2 < \cdots < r_s$ and define $\psi(x) = x^{r_1 - 1} + x^{r_2 - 1} + \cdots +$

x^{r_s-1}. Then

$$\psi(T)e_1 = T^{r_1-1}e_1 + T^{r_2-1}e_1 + \cdots + T^{r_s-1}e_1 = e_{r_1} + e_{r_2} + \cdots + e_{r_s} = e.$$

Now, for any polynomial ϕ, $\phi(T)e = \phi(T)\psi(T)e_1$.

Assume $\phi(T)e = \phi(T)\psi(T)e_1 = 0$. For any $k \in \mathbb{N}$,

$$(\phi\psi)(T)T^k e_1 = T^k(\phi\psi)(T)e_1 = 0.$$

Since e_1 is cyclic for T, $(\phi\psi)(T)$ is 0 on a spanning set of $\mathbb{C}^p$. Therefore $(\phi\psi)(T)$ is zero, so by definition $\phi(x)\psi(x)$ is divisible by the minimal polynomial of T, namely $x^p - 1$.

Let $\delta(x)$ be the greatest common divisor of $\psi(x)$ and $x^p - 1$. Since $x^p - 1$ divides $\phi(x)\psi(x)$, if we can show that $\delta(x)$ is a constant then $x^p - 1$ must divide $\phi(x)$.

Since ψ and $x^p - 1$ are polynomials over the rationals, the division algorithm for polynomials tells us that the coefficients of $\delta(x)$ are rational. However, since p is prime, $(x-1)(x^{p-1} + \cdots + 1)$ is an irreducible factorization of $x^p - 1$ over the rationals. Since $\psi(1) = |\mathcal{E}| \neq 0$, $(x-1)$ doesn't divide $\psi(x)$. Therefore $\delta(x)$ must divide $x^{p-1} + \cdots + 1$ and since $\delta(x)$ has rational coefficients it is either constant or $x^{p-1} + \cdots + 1$. But $\delta(x)$ divides $\psi(x)$ which has at most $p-1$ terms of degree at most $p-1$. Therefore $x^{p-1} + \cdots + 1$ can't divide $\psi(x)$ so $\delta(x)$ is constant as required.

□

3.4.4 Theorem

Let $g(x)$ be a polynomial that is not divisible by $x^p - 1$ for any prime p. If $\mathcal{G}$ is a finite group of operators in $\mathcal{B}(\mathcal{V})$ on which f_g is nilpotent then $\mathcal{G}$ is abelian.

Proof. Let $\mathcal{G}$ be a minimal counterexample. By minimality, $\mathcal{G}$ must be a minimal nonabelian group as f_g is nilpotent on any subgroup of $\mathcal{G}$. Therefore Lemma 3.3.10 applies and we have $\mathcal{M}, A$, and B as in that lemma.

Since $\mathcal{M}$ is an invariant subspace, for any S and T in $\mathcal{G}$ if $f_g(S,T)$ is nilpotent then so is $f_g(S,T)|_{\mathcal{M}} = f_g(S|_{\mathcal{M}}, T|_{\mathcal{M}})$. Therefore f_g is nilpotent on $\mathcal{G}|_{\mathcal{M}}$. Also, A and B don't commute so $\mathcal{G}|_{\mathcal{M}}$ is nonabelian and we can assume $\mathcal{M}$ is the entire space.

By definition, f_g is homogeneous in x and y. Therefore every term in the polynomial will contain the same powers of α and β, so we can factor them out. Since nonzero scalar multiples don't affect nilpotence, we can assume that $\alpha = \beta = 1$.

Let $g(x) = \sum_{j=0}^{m} a_j x^j$ and $f_g(x, y) = \sum_{j=0}^{m} a_j x^j y x^{m-j}$. Since f_g is nilpotent on $\mathcal{G}$, $f_g(A, B^k A^{-m})$ is nilpotent for every $k \in \mathbb{N} \cup \{0\}$. Now,

$$f_g(A, B^k A^{-m}) = a_0 B^k + a_1 A B^k A^{-1} + \cdots + a_m A^m B^k A^{-m}$$

is a diagonal matrix as B is diagonal and conjugation by A doesn't change that. However, the only nilpotent diagonal matrix is 0 so $f_g(A, B^k A^{-m}) = 0$. By the linearity in y of $f_g(x, y)$ and the above statement holding for all $k \in \mathbb{N} \cup \{0\}$ we see that

$$a_0 h(B) + a_1 A h(B) A^{-1} + \cdots + a_m A^m h(B) A^{-m} = 0$$

for any polynomial h.

We want an $h(x)$ such that $h(B)$ is a nontrivial diagonal idempotent. The diagonal of B consists of q^{th} roots of unity where q is a prime. If $\mu, \lambda \neq 1$ are q^{th} roots of unity then the sets $\{\lambda, \ldots, \lambda^{q-1}\}$ and $\{\mu, \ldots, \mu^{q-1}\}$ are equal since all roots of unity for a prime q are primitive, except for 1. Now, $\theta_1^{-1} B$ has a 1 in the $(1, 1)$ position and q^{th} roots of unity on the rest of the diagonal. As B isn't scalar, there is at least one entry on the diagonal that isn't 1. Then $B + B^2 + \cdots + B^{q-1}$ has exactly two entries on its diagonal: q and $\lambda + \lambda^2 + \cdots + \lambda^{q-1}$. To get the required $h(x)$, subtract off qI and rescale so that the remaining nonzero entries of $h(B)$ are 1.

Let $E = h(B)$ and let u be the column vector whose components are all one. Let $e = Eu$ and note that e is a cyclic vector for A by Lemma 3.4.3. Also, $A^{-j} u = u$ for all $j \in \mathbb{N}$ since all of u's components are identical. However

$$g(A)e = \sum_{j=0}^{m} a_j A^j E u = \sum_{j=0}^{m} a_j A^j E A^{-j} u = (\sum_{j=0}^{m} a_j A^j h(B) A^{-j}) u = 0.$$

So $g(A) = 0$ since it is 0 on a cyclic vector. Therefore, $g(x)$ is divisible by the minimal polynomial of A, namely $x^p - 1$, which is a contradiction.

$\square$

We can get an affirmative result in certain limited situations. Note that if $g(x) = x - 1$ then $f_g(A, B) = AB - BA$. This special case is the subject of our next result. While we are currently limited to operators of rank one, we will eventually extend this result to operators of all ranks.

3.4.5 Theorem

Let $\mathcal{S}$ be a semigroup of operators of rank at most one in $\mathcal{B}(\mathcal{V})$ such that $AB - BA$ is nilpotent for all A and B in $\mathcal{S}$. Then $\mathcal{S}$ is triangularizable.

Proof. Nilpotent commutators are inhertited by quotients so we need only show $\mathcal{S}$ is reducible by the Triangularization Lemma (2.2.3).

Assume $\mathcal{S}$ is irreducible. Then by Lemma 3.3.5 there is a subsemigroup $\mathcal{S}_0$ of $\mathcal{S}$ and an invariant subspace $\mathcal{M}$ of $\mathcal{S}_0$ such that $\mathcal{S}_0|_{\mathcal{M}}$ contains

$$S = \begin{pmatrix} \alpha & 0 \\ \beta & 0 \end{pmatrix} \text{ and } T = \begin{pmatrix} 0 & \gamma \\ 0 & \delta \end{pmatrix}$$

with $\beta\gamma(\alpha\delta - \beta\gamma) \neq 0$. Then

$$\det(ST - TS) = \det \begin{pmatrix} -\beta\gamma & \alpha\gamma \\ -\beta\delta & \beta\gamma \end{pmatrix} = \beta\gamma(\alpha\delta - \beta\gamma).$$

However, S and T are in $\mathcal{S}_0|_{\mathcal{M}}$ so $ST - TS$ is nilpotent and thus $\beta\gamma(\alpha\delta - \beta\gamma) = \det(ST - TS) = 0$ which is a contradiction. Therefore $\mathcal{S}$ is reducible and thus triangularizable.

□

The next result shows that nilpotence on one of a large family of polynomials is sufficient for triangularizability.

3.4.6 Theorem

Let $g(x) = \sum_{j=0}^{m} a_j x^j$ such that $g(1) = \sum_{j=0}^{m} a_j \neq 0$. If $\mathcal{S}$ is a semigroup of operators in $\mathcal{B}(\mathcal{V})$ such that f_g is nilpotent on $\mathcal{S}$ then $\mathcal{S}$ is triangularizable.

Proof. Take any element A in $\mathcal{S}$. Then

$$f_g(A, A) = \sum_{j=0}^{m} a_j A^j A A^{m-j} = (\sum_{j=0}^{m} a_j) A^{m+1}$$

is nilpotent. Since $\sum_{j=1}^{n} a_j \neq 0$ this means that A^{m+1}, and thus A, is nilpotent. Therefore $\mathcal{S}$ consists of nilpotents and is triangularizable by Levitzki's Theorem (2.2.11).

$\square$

We must now consider what happens when $g(1)$ is zero. First, the following calculation will be of some use.

3.4.7 Lemma

Let $g = \sum_{j=0}^{m} a_j x^j$. For operators A and B in $\mathcal{B}(\mathcal{V})$

(i) If $A = A^2$, then

$$f_g(A,B) = a_0 AB(I-A) + (\sum_{j=0}^{m} a_j) ABA + a_m (I-A)BA.$$

(ii) If $A^2 = 0$ and $m \geq 3$ then $f_g(A,B) = 0$.

(iii) If $A^2 = 0$ and $m = 2$ then $f_g(A,B) = a_1 ABA$.

Proof. Let $A = A^2$. Then

$$\begin{aligned} f_g(A,B) &= \sum_{j=0}^{m} a_j A^j B A^{m-j} \\ &= a_0 AB + \sum_{j=1}^{m-1} a_j ABA + a_m BA \\ &= a_0 AB(I-A) + \sum_{j=0}^{m} a_j ABA + a_m (I-A)BA. \end{aligned}$$

Let $A^2 = 0$ and $m \geq 3$. Then $f_g(A,B) = \sum_{j=0}^{m} a_j A^j B A^{m-j}$ and either j or $m-j$ is at least 2 so either A^j or A^{m-j} is 0 and $f_g(A,B) = 0$. Finally, if $m = 2$ then

$$f_g(A,B) = a_0 A^2 B + a_1 ABA + a_2 BA^2 = a_1 ABA.$$

$\square$

The next two examples reveal some difficulties with nonlinear polynomials.

3.4.8 Example

There exists an irreducible semigroup on which f_g is nilpotent for every polynomial $g(x) = \sum_{j=0}^{m} a_j x^j$ so long as $m \geq 2$ and $g(1) = \sum_{j=0}^{m} a_j = 0$.

Proof. Let $\mathcal{S} = \{E_{ij} : 1 \leq i, j \leq n\} \cup \{0\}$. $\mathcal{S}$ where the E_{ij} are the standard matrix units in $M_n(\mathbb{C})$. Then $\mathcal{S}$ is a semigroup as the product of any two E_{ij}'s is another E_{ij} or 0. Also, since the E_{ij}'s are a basis for $M_n(\mathbb{C})$, $\mathcal{S}$ is irreducible.

For any A in $\mathcal{S}$ either $A^2 = A$ or $A^2 = 0$. We want to show that $f_g(A, B)$ is nilpotent for every A and B in $\mathcal{S}$.

Assume $A^2 = 0$. If $m \geq 3$ then by Lemma 3.4.7, $f_g(A, B) = 0$. If $m = 2$ then $f_g(A, B) = a_1 ABA$ so $(f_g(A, B))^2 = a_1^2 ABA^2BA = 0$. Either way, $f_g(A, B)$ is nilpotent.

Now assume $A^2 = A \neq 0$. By Lemma 3.4.7,

$$f_g(A, B) = a_0 AB(I - A) + a_m (I - A)BA$$

since $g(1) = 0$. Therefore $A = E_{ii}$ for some i. If $B = A$ or $B = 0$ then $f_g(A, B) = 0$. If $B = E_{jj}$ for $i \neq j$ then $AB = BA = 0$ so $f_g(A, B) = 0$. Finally, if $B = E_{kl}$ for $k \neq l$ then B is strictly upper or lower triangular. Since A and $I - A$ are diagonal, $f_g(A, B)$ will be either strictly upper or lower triangular so it's nilpotent. Therefore f_g is nilpotent on $\mathcal{S}$.

This can be extended to include operators of all ranks. If we adjoin all diagonal idempotents to $\mathcal{S}$ this creates another semigroup, $\mathcal{S}_1$. Take A and B from $\mathcal{S}_1$. If $A^2 = 0$ then $f_g(A, B)$ is nilpotent as before. If $A^2 = A$ and B is in $\mathcal{S}_0$ then $f_g(A, B)$ is nilpotent as before since A is diagonal. Finally, if $A = A^2$ and B is not from $\mathcal{S}_0$ then A and B are both diagonal so they commute and $f_g(A, B) = 0$. Therefore $f_g(A, B)$ is nilpotent on $\mathcal{S}_1$.

$\square$

3.4.9 Example

Let $g(x) = \sum_{j=0}^{m} a_j x^j$ with $m \geq 2$ and $g(0) = g(1) = 0$. Then f_g is nilpotent on the entire semigroup of operators of rank at most 1.

Proof. Let A and B be rank 1 operators. We need to show that $f_g(A, B)$ is nilpotent.

Since A is rank 1, either $A^2 = 0$ or $A^2 = \lambda A$ where λ is the single nonzero eigenvalue of A. In this second case, since $f_g(A, B)$ is homogeneous in A, we can factor out λ^{-m} from each term and can assume that $A^2 = A$ without affecting nilpotence.

Recall that $a_0 = g(0) = 0$ and $\sum_{j=0}^{m} a_j = g(1) = 0$. If $A^2 = A$ then Lemma 3.4.7 gives us that $f_g(A, B) = a_m(I - A)BA$. Therefore $(f_g(A, B))^2 = a_m^2(I - A)B(A - A^2)BA = 0$.

If $A^2 = 0$ and $m \geq 2$, $f_g(A, B) = 0$ by Lemma 3.4.7. If $A^2 = 0$ and $m = 2$, $f_g(A, B) = a_1 ABA$ by Lemma 3.4.7. In the second case, $(f_g(A, B))^2 = a_1^2 ABA^2BA = 0$.

Therefore f_g is nilpotent on all operators of rank at most one.

$\square$

Polynomials with $g(0) = g(1) = 0$ are thus of no use in showing triangularizability. We will therefore consider polynomials g with $g(0) \neq 0$. As for Example 3.4.8, the next result shows it is, up to similarity, the only irreducible semigroup of operators of rank at most one on which f_g, with sufficient restrictions, is nilpotent.

3.4.10 Theorem

Let $g = \sum_{j=0}^{m} a_j x^j$ with $m \geq 2$, $g(0) = a_0 \neq 0$ and $g(1) = \sum_{j=0}^{m} a_j = 0$. Let $\mathcal{S}$ be an irreducible semigroup of operators of rank at most one in $\mathcal{B}(\mathcal{V})$. If f_g is nilpotent on $\mathcal{S}$ then $\mathbb{C}\mathcal{S}$ is simultaneously similar to

$$\mathbb{C}\{e_i e_j^* : 1 \leq i, j \leq n\},$$

where $\{e_i\}$ is the standard basis of column vectors and $n = \dim(\mathcal{V})$.

Proof. Assume that $\mathcal{S} = \overline{\mathbb{C}\mathcal{S}}$ since f_g will still be nilpotent on this (possibly) larger set by linearity and continuity. We claim that, if E and F in $\mathcal{S}$ are distinct nonzero idempotents, then $EF = FE = 0$.

First we check that E and F must have either distinct ranges or distinct kernels. Since E is an idempotent, $E(I - E)x = 0$ for every $x \in \mathbb{C}^n$. Since $x = Ex + (I - E)x$ and $Ex \in \operatorname{ran}(E)$ and $(I - E)x \in \ker(E)$ we see that $\operatorname{ran}(E) + \ker(E) = \mathbb{C}^n$ and similarly for F. Therefore, if E and F share the same kernel and the same range then $E = F$ as they agree on a basis (for any $x \in \operatorname{ran}(E) = \operatorname{ran}(F)$, $Ex = x = Fx$ and for any $x \in \ker(E) = \ker(F)$, $Ex = 0 = Fx$). So, if E and F are distinct they must have either distinct ranges or distinct kernels.

We want them to have distinct ranges. Assume E and F have distinct kernels. We claim that E^* and F^* have distinct ranges. Let $x \in \operatorname{ran}(E^*)^\perp$. Then $\langle x, E^* z\rangle = 0$ for all z in $\mathbb{C}^n$. This is true if and only if $\langle Ex, z\rangle = 0$ for all z. In other words, if and only if $Ex = 0$ so $x \in \ker(E)$. Therefore, $\operatorname{ran}(E^*)^\perp = \ker(E)$. Since we're in finite dimensions

$$\operatorname{ran}(E^*) = \overline{\operatorname{ran}(E^*)} = (\operatorname{ran}(E^*)^\perp)^\perp = \ker(E)^\perp.$$

The same is true for F. As $\ker(E)$ and $\ker(F)$ are distinct subspaces, they have distinct perpendicular spaces. Therefore E^* and F^* have distinct ranges.

If $h = \sum_{j=0}^m \bar{a}_{m-j} x^j$ then

$$\begin{aligned}(f_h(A^*, B^*))^* &= (\sum_{j=0}^m \bar{a}_{m-j}(A^*)^j B^* (A^*)^{m-j})^* \\ &= \sum_{j=0}^m a_{m-j} A^{m-j} B A^j \\ &= \sum_{j=0}^m a_j A^j B A^{m-j} = f_g(A, B),\end{aligned}$$

so h is nilpotent on $\mathcal{S}^*$, has rank m since $a_0 \neq 0$, has $h(0) = \bar{a}_m \neq 0$ and has $h(1) = \overline{(g(1))} = 0$. Therefore $\mathcal{S}^*$ has the same properties as $\mathcal{S}$ and note that $EF = FE = 0$ if and only if $E^*F^* = F^*E^* = 0$. Therefore, passing to $\mathcal{S}^*$ if necessary, we may assume that E and F have distinct ranges.

Take e and f, nonzero vectors in the ranges of E and F respectively. Since E and F are rank one and have distinct ranges, e and f are linearly independent and $\operatorname{ran}(E) = \operatorname{span}\{e\}, \operatorname{ran}(F) = \operatorname{span}\{f\}$. Let $\mathcal{M} = \operatorname{span}\{e, f\}$. Then $\mathcal{M}$ contains the ranges of E and F so it is invariant under them. With respect to $\mathcal{M}$ we have

$$E = \begin{pmatrix} A & S \\ 0 & 0 \end{pmatrix} \qquad\qquad F = \begin{pmatrix} B & T \\ 0 & 0 \end{pmatrix},$$

where

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix} \qquad S = \begin{pmatrix} s_1 & \dots & s_{n-2} \\ 0 & \dots & 0 \end{pmatrix}$$
$$B = \begin{pmatrix} 0 & 0 \\ \beta & 1 \end{pmatrix} \qquad T = \begin{pmatrix} 0 & \dots & 0 \\ t_1 & \dots & t_{n-2} \end{pmatrix}.$$

We claim that $EF = FE = 0$ if and only if $AB = BA = 0$. If $AB = 0$ then $\alpha = 0$. If $BA = 0$ then $\beta = 0$. If $\alpha = \beta = 0$ then $AT = BS = 0$ so $EF = FE = 0$. The other direction is trivial. Therefore, in order to show that $EF = FE = 0$ we need only check that $AB = BA = 0$ or equivalently that $\alpha = \beta = 0$.

We have that $A^2 = A$ and $g(1) = 0$, so, by Lemma 3.4.7,

$$f_g(A, B) = a_0 AB(I - A) + a_m(I - A)BA.$$

By calculation, $\det(f_g(A, B)) = a_0 a_m \alpha\beta(\alpha\beta - 1)$. But f_g is nilpotent on $\mathcal{S}$, so $f_g(E, F)$ is nilpotent. Since nilpotence is preserved by quotients, $f_g(A, B)$ is nilpotent, so its determinant is 0. Therefore, as $a_0, a_m \neq 0$, $\alpha\beta(\alpha\beta - 1) = 0$.

As $\mathcal{S}$ is irreducible, it spans $M_n(\mathbb{C})$ by Burnside's Theorem (2.2.8). Now, $EM_n(\mathbb{C})$ contains all rank one operators with range equal to $\operatorname{span}\{e\}$, so it has dimension n. As $\mathcal{S}$ spans $M_n(\mathbb{C})$, $E\mathcal{S}$ must have dimension n and $E\mathcal{S}|_{\mathcal{M}}$ must have dimension two. Therefore $E\mathcal{S}|_{\mathcal{M}}$ must contain a member

$$C = \begin{pmatrix} \gamma & \delta \\ 0 & 0 \end{pmatrix},$$

which is linearly independent of A. In other words, $\alpha\gamma - \delta \neq 0$. Since $\mathcal{S} = \mathbb{C}\mathcal{S}$ we can assume that γ is either 1 or 0 so C^2 is either C or 0.

As A and B are both idempotent, Lemma 3.4.7 tells us that

$$f_g(B, C) = a_0 BC(I - B) + a_m(I - B)CB,$$
$$f_g(A, BC) = a_0 ABC(I - A) + a_m(I - A)BCA,$$

and calculation together with the nilpotence of f_g on $\mathcal{S}$ gives us that

$$\det(f_g(B, C)) = a_0 a_m \beta\delta(\gamma - \delta\beta) = 0,$$
$$\det(f_g(A, BC)) = a_0 a_m \alpha\beta^2\gamma(\alpha\gamma - \delta) = 0.$$

We have the following four relations:

(i) $\alpha\beta(\alpha\beta - 1) = 0$,

(ii) $\alpha\gamma - \delta \neq 0$,

(iii) $\beta\delta(\gamma - \delta\beta) = 0$, and

(iv) $\alpha\beta^2\gamma(\alpha\gamma - \delta) = 0$.

Now, if neither α nor β is 0 then combining (ii) and (iv) gives us that $\gamma = 0$. Combining this with (iii) then gives us that $\delta = 0$. But then $C = 0$ which is a contradiction. Therefore, either α or β is 0. Since reordering the basis of $\mathcal{M}$ switches the roles of A and B we can assume that $\alpha = 0$.

By (ii), $\delta \neq 0$. Now, if $\beta \neq 0$ then (iii) gives us that $\gamma = \delta\beta$. In particular, $\gamma \neq 0$ so, as previously mentioned, we can assume that $\gamma = 1$ so C is idempotent. Therefore, using Lemma 3.4.7,

$$f(C, BA) = a_0 CBA(I - C) + a_m(I - C)BAC.$$

Then calculation plus the nilpotence of f_g on $\mathcal{S}$ gives us that

$$\det(f_g(C, BA)) = a_0 a_m \beta^2 \delta^2 = 0,$$

which is impossible as $a_0, a_m, \beta, \delta \neq 0$. Hence $\alpha = \beta = 0$ and $EF = FE = 0$.

To complete the proof, we use Lemma 3.3.5 to obtain bases $\{e_i\}$ and $\{f_i\}$ such that $e_i f_j^* \in \mathcal{S}$ for every i and j. Since $\mathcal{S} = \mathbb{C}\mathcal{S}$, we can perform a similarity and scale so that $\{e_i\}$ coincides with the standard basis.

Since the f_i's form a basis there must be some j_i such that $f_{j_i}^* e^i = \langle f_{j_i}, e_i \rangle \neq 0$. Then $\operatorname{tr}(e_i f_{j_i}^*) = \operatorname{tr}(f_{j_i}^* e_i) \neq 0$ so $T_i = e_i f_{j_i}^* \neq 0$. Now, T_i has exactly one nonzero row, the i^{th} row, and nonzero trace means that its (i, i) entry is nonzero. Therefore, T_i is a multiple of an idempotent with range e_i and kernel $\{f_{j_i}\}^\perp$.

We claim that this j_i is unique. Assume there was a $k_i \neq j_i$ with the same properties as j_i and let $S_i = e_i f_{k_i}^*$. Then S_i is also a scalar multiple of an idempotent with the same range as T_i, but a different kernel as f_{j_i} and f_{k_i} are linearly independent, so they have distinct perpendicular spaces. By our first claim, $S_i T_i = 0$. But

$$\operatorname{tr}(S_i T_i) = \operatorname{tr}(e_i f_{j_i}^* e_i f_{k_i}^*) = \operatorname{tr}(f_{j_i}^* e_i f_{k_i}^* e_i) = \langle f_{j_i}, e_i \rangle \langle f_{k_i}, e_i \rangle \neq 0,$$

which is a contradiction so the j_i is unique.

We also claim that for $i \neq k$, $j_i \neq j_k$. If $j_i = j_k$ then $T_i = e_i f_{j_i}^*$ and $T_k = e_k f_{j_i}^*$ are multiples of idempotents with the same kernel, but different ranges. However, their product has nonzero trace so $T_i T_k \neq 0$ which contradicts our first claim.

Therefore there is a bijective map $i \mapsto j_i$. By reordering we can assume $i = j_i$. Then $\langle e_i, f_j \rangle$ is nonzero if $i = j$ and zero otherwise. By rescaling the f_i's, we can assume $f_i = e_i$ for every i. Therefore $\mathcal{S}$ contains $\mathbb{C}\{e_i e_j^* : 1 \leq i, j \leq n\}$.

Let A be any element in $\mathcal{S}$. Assume A is a multiple of an idempotent, but is not one of the $e_i e_i^*$'s. By our claim about distinct idempotents, $A e_i e_i^* = 0 = e_i e_i^* A$ for all i. But then A commutes with a spanning set of $\mathcal{B}(\mathcal{V})$ so $A = 0$. Therefore the only multiples of idempotents in $\mathcal{S}$ are the $e_i e_i^*$'s.

If A is not a multiple of an idempotent, then A is nilpotent since it's rank one. Then A must have an off diagonal entry. By a permutation of the basis and scaling A, we can assume that A's $(1,2)$ entry is one.

Now, $A e_2 e_1^*$ is in $\mathcal{S}$. Its first column is the second column of A and all its other columns are zero. Also, its $(1,1)$ entry is 1. Therefore, $A e_2 e_1^*$ is an idempotent. Since the only idempotents in $\mathcal{S}$ are the $e_i e_i^*$'s, $A e_2 e_1^* = e_1 e_1^*$. Therefore, the second column of A must have only one nonzero entry.

Similarly, $e_2 e_1^* A$ is in $\mathcal{S}$, has its second row equal to the first row of A, and has all other rows being zero. Its $(2,2)$ entry is 1, so it is an idempotent. We must then have that $e_2 e_1^* A = e_2 e_2^*$, so the first row of A must have only one nonzero entry.

Since A has rank one, all of its nonzero columns must be multiples of one another. Since the second column is the only one with a nonzero entry on the first row, all the other columns must be zero. Therefore $A = e_1 e_2^*$ and the result is proved.

□

3.4.11 Corollary

Let $g = \sum_{j=0}^{m} a_j x^j$ with $g(0) = a_0 \neq 0$ and $g(1) = \sum_{j=0}^{m} a_j = 0$. Let $\mathcal{S}$ be an irreducible semigroup of operators in $\mathcal{B}(\mathcal{V})$ that contains a rank one operator. If f_g is nilpotent on $\mathcal{S}$ then $\mathcal{S}$ has a matrix representation in which every member has at most one nonzero entry in each row and in each column.

Proof. We can assume $\mathcal{S} = \mathbb{C}\overline{\mathcal{S}}$ as $\mathcal{S}$ will have the required property if this potentially larger group does. Let $\mathcal{J} \neq \{0\}$ be the ideal of $\mathcal{S}$ consisting of operators of rank one at most one. Then $\mathcal{J}$ is irreducible by Lemma 2.2.12 and coincides with $\mathbb{C}\{e_i e_j^*\}$ by Theorem 3.4.10.

Let S be in $\mathcal{S}$. Then $Se_ie_i^*$ has range equal to the span of the Se_i and kernel $\{e_i\}^\perp$. Also, $Se_ie_i^*$ is an element of $\mathcal{J}$ so it is equal to $\lambda e_j e_i^*$ for some j which has range equal to the span of λe_j. Therefore Se_i, which is the i^{th} column of S, is in the span of λe_j so it has at most one nonzero entry, the j^{th}.

Next, $e_ie_i^*S$ has range contained in the span of e_i (it could have zero range, depending on S). Its kernel is $\{e_i^*S\}^\perp$, which is the perpendicular space of the i^{th} row of S. Since $e_ie_i^*S$ is in $\mathcal{J}$, it is equal to $\lambda e_i e_j$ for some j and this operator has kernel $\{e_j\}^\perp$. By taking the perpendicular space of each of these equal kernels, we see that the i^{th} row of S is contained within the span e_j and therefore has at most one nonzero entry.

$\square$

We can now achieve a positive result for full-rank operators in $\mathcal{B}(\mathcal{V})$.

3.4.12 Theorem

Let $g = \sum_{j=0}^m a_j x^j$ with $g(0) = a_0 \neq 0$ and g not divisible by $x^p - 1$ for any prime p. Let $\mathcal{S}$ be a semigroup of invertible operators in $\mathcal{B}(\mathcal{V})$. If f_g is nilpotent on $\mathcal{S}$ then $\mathcal{S}$ is triangularizable.

Proof. If $g(1) \neq 0$ then $\mathcal{S}$ is triangularizable by Theorem 3.4.6. Otherwise, if $m = 0$ then $f_g(A, B) = a_0 B$ so nilpotence of f_g on $\mathcal{S}$ implies nilpotence of every element of $\mathcal{S}$. But nilpotent elements aren't invertible so this is a contradiction so $m \geq 1$.

Since a nilpotent f_g extends to quotients we need only show reducibility by the Triangularization Lemma (2.2.3).

Let Φ be a ring automorphism of $\mathcal{B}(\mathcal{V})$ induced by the field automorphism ϕ. By Lemma 2.5.3, for any A and B in $\mathcal{S}$

$$\sigma(\Phi(f_g(A, B))) = \phi\left(\sigma(f_g(A, B))\right) = \phi(\{0\}) = \{0\}$$

since $f_g(A,B)$ is nilpotent. Define $h(x) = \sum_{j=0}^{m} \phi(a_j)x^j$. Since

$$\Phi(f_g(A,B)) = \Phi(\sum_{j=0}^{m} a_j A^j B A^{m-j}) = \sum_{j=0}^{m} \phi(a_j)\Phi(A)^j\Phi(B)\Phi(A)^{m-j},$$

f_h is nilpotent on $\Phi(\mathcal{S})$. By continuity and linearity, f_h is nilpotent on $\overline{\mathbb{C}\Phi(\mathcal{S})}$. Finally, Φ^{-1} is the field automorphism induced by ϕ^{-1} so, for any A and B in $\Phi(\mathcal{S})$,

$$\begin{aligned}
\Phi^{-1}(f_h(A,B)) &= \Phi^{-1}\left(\sum_{j=0}^{m} \phi(a_j)A^j B A^{m-j}\right) \\
&= \sum_{j=0}^{m} a_j \Phi^{-}1(A)^j \Phi^{-1}(B)\Phi^{-1}(A)^{m-j} \\
&= f_g(\Phi^{-1}(A), \Phi^{-1}(B)).
\end{aligned}$$

Thus

$$\sigma(f_g(\Phi^{-1}(A), \Phi^{-1}(B))) = \sigma(\Phi^{-1}(f_h(A,B))) = \phi(\sigma(f_h(A,B))) = \{0\},$$

so f_g is nilpotent on $\Phi^{-1}(\overline{\mathbb{C}\Phi(\mathcal{S})})$. Therefore the conditions of the Finiteness Lemma (3.2.2) are met for the property $\mathcal{P}$ of f_g being nilpotent.

If $m = 1$ then $a_1 = -a_0$. Since multiplication by a scalar doesn't affect nilpotence, we can assume that $g(x) = x - 1$. Therefore $f_g(x,y) = xy - yx$. We want to show that the conditions of Lemma 3.2.3 apply. Since $\mathcal{S}$ consists of invertible elements, the only nonzero idempotent in $\mathcal{S}$ is I and $I\mathcal{S}I|_{I\mathcal{V}} = \mathcal{S}$, on which f_g is nilpotent. If $\mathcal{J}$ is an ideal in $\mathcal{S}$ then f_g is nilpotent on $\mathcal{J}$ since $\mathcal{J}$ is a subset of $\mathcal{S}$. Every finite group with f_g nilpotent is abelian, and therefore reducible, by Theorem 3.4.4. And $f_g(A,B) = AB - BA$ and f_g nilpotent gives $AB - BA$ nilpotent so every such semigroup of operators of rank at most one is reducible by Theorem 3.4.5. So Lemma 3.2.3 applies and $\mathcal{S}$ is reducible.

Let $m \geq 2$ and assume $\mathcal{S}$ is irreducible. Let $\hat{\mathcal{S}}$ be the maximal semigroup containing $\mathcal{S}$ with property $\mathcal{P}$. By Lemma 2.3.3, $\hat{\mathcal{S}}$ contains a minimal rank idempotent E. The Finiteness Lemma (3.2.2) gives us that $E\hat{\mathcal{S}}E|_{E\mathcal{V}}$ is contained within multiples of a finite group. So Lemma 3.4.4 tells us that $E\hat{\mathcal{S}}E|_{E\mathcal{V}}$ is abelian. If E has rank at least 2 then $E\hat{\mathcal{S}}E|_{E\mathcal{V}}$ is reducible so $\hat{\mathcal{S}}$ is reducible by Lemma 2.2.13.

All that remains is the case where the rank of E is one. Then $\hat{\mathcal{S}}$ contains E, a rank one operator, so Corollary 3.4.11 applies and $\hat{\mathcal{S}}$ has a matrix representation with each member having at most one nonzero entry in each column and in each row. Let $\{e_i\}$ be the basis with respect to which $\hat{\mathcal{S}}$ has this form.

Let S be a member of $\hat{\mathcal{S}}$. We claim that, after a permutation of the basis, S is the direct sum of cyclic operators. Fix any i and consider the action of S on e_i. If the i^{th} column of S is all zeroes then $Se_i = 0$ and S is cyclic on $\text{span}(e_i)$.

Otherwise, the i^{th} column of S contains a single nonzero entry, λ_i. If λ_i occurs in the i^{th} row then $Se_i = \lambda_i e_i$ so S is cyclic on $\text{span}(e_1)$.

Finally, if λ_i occurs in the j^{th} row of S then $Se_i = \lambda_i e_j$. We repeat the above argument for the j^{th} column of S. If $Se_j = 0$ then S is cyclic on $\text{span}\{e_i, e_j\}$. If $Se_j = \lambda_j e_i$ for some $\lambda_j \neq 0$ then S is again cyclic on $\text{span}\{e_i, e_j\}$. Otherwise, the j^{th} column of S has a nonzero entry, λ_j, at the k^{th} row where $k \neq i$. Note also that $k \neq j$ since S has exactly one nonzero entry on its j^{th} row and that's λ_i. Continuing this argument we get a sequence of distinct basis elements. Eventually, either S will have a zero column or, since $\mathcal{V}$ is finite, S will map the most recent basis vector to $\text{span}(e_i)$. Either way, S is cyclic on the span of these vectors.

By permuting the basis, S becomes a direct sum of each of these cyclic operators. Let A be a nondiagonal cyclic operator. So

$$A = \begin{pmatrix} 0 & & & \lambda_n \\ \lambda_1 & \ddots & & \\ & \ddots & \ddots & \\ & & \lambda_{n-1} & 0 \end{pmatrix}.$$

By construction, all the λ_i's are nonzero, with the possible exception of λ_n. If $\det(A) = 0$ leave A as is. Otherwise, since $\hat{\mathcal{S}} = \mathbb{C}\hat{\mathcal{S}}$ by maximality, we can replace A with $(\det(A))^{-1/n}$. Therefore $\det(A)$ is either zero or one.

We can then apply a diagonal similarity, which won't change the form of A, to guarantee that each nonzero λ_i is one. Then

$$A = \begin{pmatrix} 0 & & & \alpha \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix},$$

where $\alpha = \det(A)$ is either zero or one.

Then $\mathcal{S}$ is a subset of $\hat{\mathcal{S}}$, so each member of $\mathcal{S}$ has this form. Since $\mathcal{S}$ is invertible it has full rank so every column and row must have at least one nonzero entry. Therefore, if A is in the direct sum of cyclic operators for some S in $\mathcal{S}$ then A is invertible and $\alpha = 1$.

Since $\mathcal{S}$ is irreducible it can't be diagonalizable so there is an S in $\mathcal{S}$ with such an A. Let $\{e_1, \ldots, e_n\}$ be the basis for A. Now, $A^n = I$. Let p be a prime that divides n and replace A with $A^{n/p}$ which is still a restriction of an element of $\hat{\mathcal{S}}$ since $\hat{\mathcal{S}}$ is a semigroup. Then A consists of $\frac{n}{p}$ cyclic permutations $e_i \mapsto e_{\frac{n}{p}+i} \mapsto \ldots \mapsto e_{(p-1)\frac{n}{p}+i} \mapsto e_i$ for $1 \leq i \leq \frac{n}{p}$. Further restrict A to one of these cycles so that A is a cyclic permutation on a p dimensional space $\mathcal{M}$.

Additionally, since $A^p = I$, $A^{p-1} = A^{-1}$ so A^{-1} is the restriction of S^{p-1} which is in $\hat{\mathcal{S}}$.

Let $\mathcal{J}$ be the ideal of rank at most one operators in $\hat{\mathcal{S}}$. Since $\hat{\mathcal{S}}$ has a rank one idempotent $\mathcal{J} \neq \{0\}$. Since $\hat{\mathcal{S}}$ is irreducible, $\mathcal{J}$ is irreducible by Lemma 2.2.12. By Theorem 3.4.10 and how Corollary 3.4.11 was constructed, $\mathcal{J}$ must contain $T = e_1 e_1^*$. Let B be the restriction of T to $\operatorname{span}\{e_1, \ldots, e_n\}$.

Then, since S is the direct sum of operators, and since T is 0 everywhere except B, as $f_g(S, TS^m)$ is nilpotent, so is $f_g(A, BA^{-m})$. But

$$\begin{aligned} f_g(A, BA^{-m}) &= \sum_{j=0}^{m} a_j A^j (BA^{-m}) A^{m-j} \\ &= \sum_{j=0}^{m} a_j A^j B A^{-j} \\ &= \operatorname{diag}(b_0, b_1, \ldots, b_{p-1}), \end{aligned}$$

where $b_i = a_i + a_{i+p} + \ldots$. This last equality is true since the only nonzero entry of $A^j BA^{-j}$ is in the (k,k) position where $k \equiv j \mod p$ and $0 \leq k \leq p-1$. Nilpotence implies that $b_0 = b_1 = \cdots = b_{p-1}$.

Now,

$$g(x) = \sum_{j=0}^{m} a_j x^j = \sum_{i=0}^{p-1} x^i (a_i + a_{i+p} x^p + \ldots),$$

and for any p^{th} root of unity λ,

$$\begin{aligned} g(\lambda) &= \sum_{i=0}^{p-1} \lambda^i (a_i + a_{i+p}\lambda^p + \dots) \\ &= \sum_{i=0}^{p-1} \lambda^i (a_i + a_{i+p} + \dots) \\ &= \sum_{i=0}^{p-1} \lambda^i b_i = 0, \end{aligned}$$

so $x^p - 1$ divides $g(x)$ which is a contradiction. Therefore $\mathcal{S}$ is reducible and thus triangularizable.

□

3.4.13 Corollary

Let $\{a_0, \dots, a_k\}$ be any scalars such that $\sum_{j=0}^{k} a_j x^j$ is not divisible by $x^p - 1$ for any prime p. Let $\mathcal{G}$ be a group of operators in $\mathcal{B}(\mathcal{V})$ such that $\sum_{j=0}^{k} a_j A^j B A^{k-j}$ is nilpotent for all A and B in $\mathcal{G}$. Then $\mathcal{G}$ is triangularizable and, in particular, if such a $\mathcal{G}$ consists of unitary operators then it's commutative.

Proof. Let r and t be such that a_r is the first nonzero a_i and a_t is the last nonzero a_i. Let $m = t - r$ and $g = \sum_{j=0}^{m} a_{r+j} x^j$. Then $x^r g(x) = \sum_{j=0}^{m} a_j x^j$ so, since $x^p - 1$ doesn't divide $\sum_{j=0}^{k} a_j x^j$ for any prime p, it doesn't divide $g(x)$ either. Also, $g(0) = a_r \neq 0$.

For any A and B in $\mathcal{G}$,

$$\begin{aligned} \sum_{j=0}^{m} a_j A^j B A^{k-j} &= \sum_{j=r}^{t} a_j A^j B A^{k-j} \\ &= \sum_{j=r}^{t} a_j A^{j-r} (A^r B A^{k-t}) A^{k-j-(k-t)} \\ &= \sum_{j=0}^{m} a_{j+r} A^j (A^r B A^{k-t}) A^{m-j} \\ &= f_g(A, A^r B A^{k-t}). \end{aligned}$$

Therefore $f_g(A, A^r B A^{k-t})$ is nilpotent. Since $\mathcal{G}$ is a group, $A^{-r} B A^{-(k-t)}$ is an element of $\mathcal{G}$. By replacing B with $A^{-r} B A^{-(k-t)}$ we see that $f_g(A, B)$ is nilpotent. By Theorem 3.4.12, $\mathcal{G}$ is triangularizable.

If $\mathcal{G}$ consists of unitaries then $\mathcal{G}$ is self-adjoint so it's diagonalizable and thus commutative (3.3.13).

□

3.4.14 Corollary

Let g be a polynomial that is not divisible by $x^p - 1$ for any prime p. Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{B}(\mathcal{V})$ on which f_g is nilpotent. Then either of the following conditions imply reducibility:

(i) $\mathcal{S}$ does not contain a rank-one operator and $g(0) \neq 0$.

(ii) $\overline{\mathbb{C}\mathcal{S}}$ does not contain a rank-one operator.

Proof. (i) Assume $\mathcal{S}$ is irreducible. Let m be the minimal rank in $\mathcal{S}$ and let A be a nonzero member of $\mathcal{S}$ with rank m. By putting A in Jordan form and taking powers of A we can assume $A = A_0 \oplus 0$ for some invertible A_0 acting on the m dimensional space $A\mathcal{V}$.

Consider $\mathcal{S}_0 = A\mathcal{S}A|_{A\mathcal{V}}$ and let $E = I_{A\mathcal{V}} \oplus 0$. Since $\mathcal{S}$ is irreducible, it must span $\mathcal{B}(\mathcal{V})$. Therefore the set $E\mathcal{S}E|_{A\mathcal{V}}$ must span $\mathcal{B}(A\mathcal{V})$ and therefore has dimension m^2. For any S in $\mathcal{S}$, $ASA = AESEA$ so $\mathcal{S}_0 = A_0 E\mathcal{S}EA_0|_{A\mathcal{V}}$. Since A_0 is invertible, the dimension of $\mathcal{S}_0$ is m^2. Therefore $\mathcal{S}_0$ spans $\mathcal{B}(A\mathcal{V})$, so $\mathcal{S}_0$ is irreducible. But for any S in $\mathcal{S}$,

$$ASA = \begin{pmatrix} A_0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A_0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A_0 S_{11} A_0 & 0 \\ 0 & 0 \end{pmatrix},$$

so the rank of ASA is at most m, the rank of A_0. By minimality of rank in $\mathcal{S}$, rank ASA is m. Therefore $A_0 S_{11} A_0$ has rank m and acts on an m dimensional space so it is invertible. So $\mathcal{S}_0$ consists of invertible operators. As f_g is nilpotent on $\mathcal{S}$, it is also nilpotent on $\mathcal{S}_0$. By Theorem 3.4.12, $\mathcal{S}_0$ is reducible.

That's a contradiction, so $\mathcal{S}$ is reducible.

(ii) Assume $\mathcal{S}$ is irreducible. Then $\overline{\mathbb{C}\mathcal{S}}$ is irreducible, so it contains an idempotent E of minimal rank by Lemma 2.3.3. The same lemma says that $E\overline{\mathbb{C}\mathcal{S}}E|_{E\mathcal{V}}$ is a group. By Corollary 3.4.13, $E\overline{\mathbb{C}\mathcal{S}}E|_{E\mathcal{V}}$ is triangularizable. But by Lemma 2.2.13, $E\overline{\mathbb{C}\mathcal{S}}E|_{E\mathcal{V}}$ is irreducible. This is a contradiction, so $\mathcal{S}$ is reducible.

□

We can now extend Theorem 3.4.5 to operators of arbitrary rank.

3.4.15 Theorem

Let $\mathcal{S}$ be a semigroup of operators such that $AB - BA$ is nilpotent for every pair $\{A, B\}$ in $\mathcal{S}$. Then $\mathcal{S}$ is triangularizable.

Proof. Let $g(x) = x - 1$. Then $f_g(A, B) = AB - BA$, so f_g is nilpotent on $\mathcal{S}$. Since nilpotence of commutators extends to quotients, we need only show reducibility by the Triangularization Lemma (2.2.3).

If $\mathcal{S}$ contains no rank one operators then as $g(0) \neq 0$, $\mathcal{S}$ is reducible by Corollary 3.4.14.

If $\mathcal{S}$ contains a rank one operator then the ideal of rank at most one operators in $\mathcal{S}$ is nonzero and reducible by Theorem 3.4.5. Then $\mathcal{S}$ is reducible by Lemma 2.2.12.

□

This concludes our discussion of the nilpotence of polynomials. We now consider extending our results to infinite dimensions. In Chapter 5, we return to Theorem 3.4.15 and consider if the condition of nilpotence can be weakened.

Chapter 4

Infinite Dimensions

In this chapter, we investigate extensions of the results in the previous chapter to infinite dimensional spaces. We will mostly be interested in $\mathcal{K}(\mathcal{X})$, the set of compact operators on a Banach space $\mathcal{X}$. Occasionally, we will restrict our results to $\mathcal{K}(\mathcal{H})$, the set of compact operators on a Hilbert space $\mathcal{H}$. We will also consider some extensions to the set of bounded operators $\mathcal{B}(\mathcal{X})$.

Many of the finite dimensional proofs require no modification. We will deal only with those results that require significant modification.

4.1 Definitions and Notation

In this chapter, linear subspaces are assumed to be closed and span should be read as the closed linear span. The concept of reducibility remains otherwise unchanged, but we need to extend the concept of triangularizability for Definition 2.1.6 to infinite dimensions. The following definition reduces to the previous case if $\mathcal{X}$ is finite dimensional.

4.1.1 Definition

A family of operators $\mathcal{F}$ in $\mathcal{B}(\mathcal{X})$ is said to be *triangularizable* if there is a chain $\mathcal{C}$ of subspaces of $\mathcal{X}$ which is maximal (as a chain of subspaces) and if $\mathcal{M}$ is in $\mathcal{C}$ then $\mathcal{M}$ is an invariant subspace for $\mathcal{F}$. Such a chain is called a *triangularizing chain.*

Triangularizing chains in infinite dimensions don't always look like they do in finite dimensions. See Example 4.1.5.

4.1.2 Definition

A chain of subspaces is called *complete* if it is closed under arbitrary intersections and closed spans.

If an element of a complete chain is not the span of its predecessors we can define its immediate predecessor as follows.

4.1.3 Definition

If $\mathcal{C}$ is a chain of subspaces and $\mathcal{M} \in \mathcal{C}$ then $\mathcal{M}_-$ is defined as

$$\mathcal{M}_- = \overline{\operatorname{span}\{\mathcal{N} \in \mathcal{C} : \mathcal{N} \subset \mathcal{M}\}}.$$

If $\mathcal{M}_- \neq \mathcal{M}$, then $\mathcal{M}_-$ is the *immediate predecessor* of $\mathcal{M}$ in $\mathcal{C}$.

This allows us to characterize maximal subspace chains in $\mathcal{X}$.

4.1.4 Theorem

A chain of subspaces of a Banach space $\mathcal{X}$ is maximal as a subspace chain if and only if it satisfies the following three conditions:

(i) it contains $\{0\}$ and $\mathcal{X}$,

(ii) it is complete, and

(iii) if $\mathcal{M}$ is in the chain and $\mathcal{M}_- \neq \mathcal{M}$, then the quotient space $\mathcal{M}/\mathcal{M}_-$ is one dimensional.

Proof. Let $\mathcal{C}$ be a maximal subspace chain. Then $\{0\}$ and $\mathcal{X}$ are clearly in $\mathcal{C}$ as they are comparable with every subspace of $\mathcal{X}$. Intersections and spans of elements of a chain are also comparable with every element in that chain so they must be in $\mathcal{C}$ by maximality. Finally, if $\mathcal{M}/\mathcal{M}_-$ has dimension at least 2 then the subspace $\mathcal{N}$ taken to be the span of $\mathcal{M}_-$ and an element in $\mathcal{M}$ that is not in $\mathcal{M}_-$ is properly between the two spaces. By maximality, it must be in $\mathcal{C}$, which is a contradiction as then $\mathcal{N}$

should be contained within $\mathcal{M}_-$. Therefore a maximal subspace chain satisfies all three properties.

Let $\mathcal{C}$ be a chain satisfying all three properties and let $\mathcal{M}$ be a subspace comparable with every element of $\mathcal{C}$. We want to show that $\mathcal{M}$ is in $\mathcal{C}$. Assume otherwise and define

$$\mathcal{M}_0 = \overline{\text{span}\{\mathcal{N} \in \mathcal{C} : \mathcal{N} \subset \mathcal{M}\}} \qquad\qquad \mathcal{M}_1 = \cap\{\mathcal{N} \in \mathcal{C} : \mathcal{N} \supset \mathcal{M}\}.$$

Since $\mathcal{C}$ is a complete chain, $\mathcal{M}_0$ and $\mathcal{M}_1$ are in $\mathcal{C}$. And $\mathcal{M}_0 \subseteq \mathcal{M} \subseteq \mathcal{M}_1$. Since $\mathcal{M}$ is not in $\mathcal{C}$, both inclusions are proper. But this means the gap between $\mathcal{M}_0$ and $\mathcal{M}_1$ has dimension at least 2.

Now, every proper subset of $\mathcal{M}_1$ in $\mathcal{C}$ is contained in $\mathcal{M}_0$ by definition. But this means that $\mathcal{M}_{1-}$ is contained in $\mathcal{M}_0$. However, this contradicts (iii). Therefore $\mathcal{M}$ must be in $\mathcal{C}$ and therefore $\mathcal{C}$ is maximal.

$\square$

In finite dimensions, $\mathcal{M}_- \neq \mathcal{M}$ for any $\mathcal{M} \neq \{0\}$ in a triangularizing chain. The following example shows this isn't necessarily true in infinite dimensions.

4.1.5 Example

For each t in $[0,1]$ let $\mathcal{M}_t = \{f \in \mathcal{L}^2(0,1) : f\chi_{(0,t)} = f\}$. Then the chain $\mathcal{C}$ consisting of $\{\mathcal{M}_t\}_{t\in[0,1]}$ is a triangularizing chain for $\mathcal{A} = \text{Alg}(\{\mathcal{M}_t\}_{t\in[0,1]})$ (the set of operators that leaves each $\mathcal{M}_t$ invariant) and $\mathcal{M}_- = \mathcal{M}$ for all $\mathcal{M}$ in $\mathcal{C}$.

Proof. By definition, $\mathcal{C}$ consists of invariant subspaces for $\mathcal{A}$. Also, $\mathcal{A}$ isn't trivial as the operators taking f to $\hat{f}$ where $\hat{f}(t) = \int_{[t,1]} f$ are in $\mathcal{A}$. We therefore need only show that $\mathcal{C}$ satisfies the conditions of Theorem 4.1.4.

Since $\mathcal{M}_0 = \{0\}$ and $\mathcal{M}_1 = \mathcal{L}^2(0,1)$, (i) is satisfied. For any subset Λ of $[0,1]$,

$$\bigcap_{t\in\Lambda} \mathcal{M}_t = \mathcal{M}_{\inf(\Lambda)},$$

so $\mathcal{C}$ is closed under arbitrary intersections. Now, let $t_0 = \sup \Lambda$. Since $\mathcal{M}_{t_0}$ is a closed linear subspace and $\mathcal{M}_t \subseteq \mathcal{M}_{t_0}$ for t in Λ, the span of the $\mathcal{M}_t$ with t from Λ is contained within $\mathcal{M}_{t_0}$. We claim the span is actually equal to $\mathcal{M}_{t_0}$. Let f be in $\mathcal{M}_{t_0}$.

Then for any t in Λ, $f_t = f\chi_{[0,t]}$ is in $\mathcal{M}_t$. Since $t_0 = \sup(\Lambda)$, we have f_t converges to f as t in Λ approaches t. Therefore f is in the span of the $\mathcal{M}_t$ with t in Λ. Therefore the span is $\mathcal{M}_{t_0}$ and $\mathcal{C}$ is closed under arbitrary spans so (ii) is satisfied.

For (iii), note that $(\mathcal{M}_{t_0})_-$ is the span of the $\mathcal{M}_t$ with t from $\Lambda = [0, t_0)$. Therefore $(\mathcal{M}_{t_0})_- = \mathcal{M}_{t_0}$ by the previous paragraph.

Therefore $\mathcal{C}$ is a maximal subspace chain so it's a triangularizing chain and $\mathcal{M}_- = \mathcal{M}$ for all $\mathcal{M}$ in $\mathcal{C}$.

□

We can now prove the infinite dimensional version of the Triangularization Lemma.

4.1.6 Lemma (Triangularization Lemma)

If $\mathcal{P}$ is a property of families of operators in $\mathcal{B}(\mathcal{X})$ that is inherited by quotients, and if every family on a space of dimension at least 2 satisfying $\mathcal{P}$ is reducible, then every family satisfying $\mathcal{P}$ is triangularizable.

Proof. Let $\mathcal{F}$ be a family satisfying $\mathcal{P}$ and let $\mathcal{C}$ be a chain of invariant subspaces for $\mathcal{F}$ that is maximal as a chain of invariant subspaces. We need to show that $\mathcal{C}$ is maximal. We'll do this by showing it satisfies the three properties of Theorem 4.1.4.

As $\{0\}$ and $\mathcal{X}$ are always invariant subspaces, (i) is clear. Spans and intersections of invariant subspaces also always produce invariant subspaces.

Assume that $\mathcal{C}$ does not satisfy (iii). Let $\mathcal{M}$ be in $\mathcal{C}$ with dimension of $\mathcal{M}/\mathcal{M}_-$ at least 2. Consider $\hat{\mathcal{F}}$, the set of quotients of $\mathcal{F}$ in $\mathcal{M}/\mathcal{M}_-$. $\hat{\mathcal{F}}$ then has property $\mathcal{P}$ and acts on a space of dimension at least 2. Therefore it is reducible and has an invariant subspace $\mathcal{L}$. Take $\mathcal{N} = \{x \in \mathcal{M} : [x] \in \mathcal{L}\}$. Then $\mathcal{N}$ is an invariant subspace for $\mathcal{F}$ that lies properly between $\mathcal{M}$ and $\mathcal{M}_-$, so, by maximality as a chain of invariant subspaces, $\mathcal{N}$ must be in $\mathcal{C}$. This is a contradiction as $\mathcal{M}_-$ is the immediate predecessor of $\mathcal{M}$ in $\mathcal{C}$.

Therefore $\mathcal{C}$ satisfies the three properties and is a triangularizing chain.

□

Every compact operator is triangularizable. In fact, we can extend this result to commutative families of operators.

4.1.7 Theorem

Every commutative family of compact operators is triangularizable.

Proof. See [7, Theorem 7.2.1].

□

We can define diagonal coefficients for compact operators that allow us to read off the spectrum of a compact operators from its triangular form.

4.1.8 Definition

Let $\mathcal{C}$ be any triangularizing chain for K in $\mathcal{K}(\mathcal{X})$. For each $\mathcal{M}$ in $\mathcal{C}$ we define the diagonal coefficient of K corresponding to $\mathcal{M}$ (denoted $\lambda_{\mathcal{M}}$) as follows: If $\mathcal{M} = \mathcal{M}_-$ then $\lambda_{\mathcal{M}} = 0$. Otherwise, $\lambda_{\mathcal{M}}$ is the lone element in the spectrum of the one dimensional operator $K|_{(\mathcal{M}/\mathcal{M}_-)}$. (It's the unique number such that $(K - \lambda_{\mathcal{M}} I)\mathcal{M} \subseteq \mathcal{M}_-$.)

4.1.9 Theorem (Ringrose's Theorem)

If K is in $\mathcal{K}(\mathcal{X})$, $\mathcal{X}$ is infinite dimensional, and $\mathcal{C}$ is a triangularizing chain for K then

$$\sigma(K) = \{0\} \cup \{\lambda_{\mathcal{M}} : \mathcal{M} \in \mathcal{C}\}.$$

Proof. See [7, Theorem 7.2.3].

□

4.1.10 Theorem

The diagonal multiplicity of each nonzero eigenvalue with respect to any triangularizing chain of an operator in $\mathcal{K}(\mathcal{X})$ is equal to its algebraic multiplicity.

Proof. See [7, Theorem 7.2.9].

□

4.1.11 Theorem

Let K be a compact operator on an infinite-dimensional space and let $\mathcal{C}$ be a complete chain of invariant subspaces of K. For each $\mathcal{M} \in \mathcal{C}$ for which $\mathcal{M}_- \neq \mathcal{M}$, define $K_{\mathcal{M}}$ to be the quotient operator on $\mathcal{M}/\mathcal{M}_-$ induced by K. Then

$$\sigma(K) = \{0\} \cup \{\sigma(K_{\mathcal{M}}) : \mathcal{M} \in \mathcal{C} \text{ and } \mathcal{M}_- \neq \mathcal{M}\}.$$

Proof. See [7, Theorem 7.2.7].

□

A number of finite dimensional results extend to compact operators.

4.1.12 Theorem (Spectral Mapping Theorem)

If $\{K_1, \ldots, K_n\}$ is a triangularizable family of operators in $\mathcal{K}(\mathcal{X})$ and p is any polynomial in n (possibly noncommuting) variables, then

$$\sigma(p(K_1, \ldots, K_n)) \subseteq p(\sigma(K_1), \ldots, \sigma(K_n)).$$

Proof. Adding multiples of the identity shifts both sides of the inclusion equally, so we can assume the constant term of p is zero. Therefore $p(K_1, \ldots, K_n)$ is in $\mathcal{K}(\mathcal{X})$. Take a triangularizing chain $\mathcal{C}$ for the family of operators. Then $\mathcal{C}$ also triangularizes $p(K_1, \ldots, K_n)$.

By Ringrose's Theorem (4.1.9), we have to check that $\lambda_{\mathcal{M}}$ (for $p(K_1, \ldots, K_n)$) is in the right hand side for every $\mathcal{M}$ in $\mathcal{C}$. Let $\mathcal{M}$ be in $\mathcal{C}$. If $\mathcal{M} = \mathcal{M}_-$ then $\lambda_{\mathcal{M}} = 0$ which is clearly in the right hand side since 0 is in the spectrum of every compact operator. Otherwise, let $\hat{K}_i$ be the quotient of K_i on the space $\mathcal{M}/\mathcal{M}_-$. Then $p(\hat{K}_1, \ldots, \hat{K}_n)[f] = \lambda_{\mathcal{M}}[f]$ for every $[f]$ in $\mathcal{M}/\mathcal{M}_-$. For each i, since $\mathcal{M}$ is invariant for K_i, there is a λ_i such that $\hat{K}_i[f] = \lambda_i[f]$ for each $[f]$ in $\mathcal{M}/\mathcal{M}_-$. That $\lambda_{\mathcal{M}} = p(\lambda_1, \ldots, \lambda_n)$ is clear. Finally, each λ_i is a diagonal coefficient for K_i, so $\lambda_i \in \sigma(K_i)$ and the inclusion holds.

□

4.1.13 Lemma

If $\{K_n\}$ is a sequence of compact operators converging to K in norm then

$$\sigma(K) = \{\lambda : \lambda = \lim_{n \longrightarrow \infty} \lambda_n, \lambda_n \in \sigma(K_n) \text{ for all } n\}.$$

Proof. See [7, Theorem 7.2.13].

$\square$

4.2 The Downsizing Lemma

The goal of this section is to reduce problems about compact operators to problems about finite dimensional operators. We need a number of preliminary results.

4.2.1 Lemma

If $\mathcal{S}$ is a semigroup of operators in $\mathcal{K}(\mathcal{X})$ with $\mathcal{S} = \overline{\mathbb{R}^+\mathcal{S}}$ and if $\mathcal{S}$ contains non-quasinilpotent operators then $\mathcal{S}$ contains a finite-rank operator other than zero that is nilpotent or idempotent

Proof. See [7, Lemma 7.4.5].

$\square$

4.2.2 Theorem (Turovskii's Theorem)

A semigroup of compact quasinilpotent operators on a Banach space is triangularizable.

Proof. See [7, Theorem 8.1.11].

$\square$

4.2.3 Lemma

If a semigroup of operators $\mathcal{S}$ in $\mathcal{B}(\mathcal{X})$ has a reducible nonzero ideal then $\mathcal{S}$ is reducible. In other words, a nonzero ideal of an irreducible semigroup of operators in $\mathcal{B}(\mathcal{X})$ is irreducible.

Proof. The proof is the same as for Lemma 2.2.12, except that we deal here with closed linear spans.

□

4.2.4 Definition

A linear functional ϕ on a linear space $\mathcal{L}$ in $\mathcal{B}(\mathcal{X})$ is called a *coordinate functional* if there is a nonzero x in $\mathcal{X}$ and a nonzero linear functional f on $\mathcal{X}$ such that $\phi(L) = f(Lx)$ for every L in $\mathcal{L}$.

Note that such a ϕ is continuous on $\mathcal{B}(\mathcal{X})$ if and only if f is continuous. Also, if $\mathcal{L}$ separates the points of $\mathcal{X}$ then ϕ is obviously nonzero. In particular, any coordinate functional on $\mathcal{B}(\mathcal{X})$ is nonzero.

4.2.5 Lemma

Let $\mathcal{S}$ be an arbitrary semigroup in $\mathcal{B}(\mathcal{X})$ and ϕ a continuous coordinate functional on $\mathcal{B}(\mathcal{X})$. Then $\mathcal{S}$ is reducible if ϕ is constant on $\mathcal{S}$.

Proof. Let $\phi(S) = f(Sx)$ for all S in $\mathcal{S}$, where x is in $\mathcal{X}$ and f is in $\mathcal{X}^*$.

If ϕ is zero on $\mathcal{S}$, take $\mathcal{M}$ to be the closed linear span of $\mathcal{S}x$. So $\mathcal{M}$ is invariant under $\mathcal{S}$. We know $\mathcal{M} \neq \{0\}$ as $\mathcal{S}$ is not the zero semigroup (if it is, it's reducible). Also, f is not zero (since ϕ is a coordinate functional), but $f|_{\mathcal{M}} = 0$, so $\mathcal{M} \neq \mathcal{X}$. Therefore $\mathcal{M}$ is a nontrivial invariant subspace for $\mathcal{S}$.

Assume instead that $\phi(S) = \lambda \neq 0$ for all S in $\mathcal{S}$. If $\mathcal{S}$ is a singleton the it's a commutative family and thus triangularizable by Theorem 4.1.7. Otherwise, take $A \neq B$ in $\mathcal{S}$. Take $\mathcal{A}$ to be the algebra spanned by $\mathcal{S}$ and $\mathcal{J} \neq \{0\}$ to be the ideal generated by $A - B$. For any S, T in $\mathcal{S}$,

$$\phi(S(A-B)T) = \phi(SAT) - \phi(SBT) = 0.$$

Therefore $\mathcal{J}$ is reducible by the first paragraph and $\mathcal{S}$ is reducible by Lemma 4.2.3.

$\square$

4.2.6 Lemma

Let P be a projection on $\mathcal{X}$ and $\mathcal{S}$ any semigroup in $\mathcal{B}(\mathcal{X})$. Let $\mathcal{C}$ be a chain of invariant subspaces for $P\mathcal{S}|_{P\mathcal{X}}$. Then there is a one-to-one, order preserving map from $\mathcal{C}$ into the lattice of invariant subspaces of $\mathcal{S}$.

Proof. For every $\mathcal{M}$ in $\mathcal{C}$ we define $\mathcal{M}_1$ to be the closed linear span of $\mathcal{M} \cup \mathcal{S}\mathcal{M}$. Then $\mathcal{M}_1$ is an invariant subspace for $\mathcal{S}$. Now take $\mathcal{N}$ in $\mathcal{C}$ with $\mathcal{M} \subset \mathcal{N}$. Clearly $\mathcal{M}_1 \subseteq \mathcal{N}_1$. We want to show that this inclusion is proper.

Take x from $\mathcal{N}/\mathcal{M}$. Then x is in $\mathcal{N}_1$. Assume that x is in $\mathcal{M}_1$. Now, x is in $\mathcal{N}$, so it's in $P\mathcal{X}$. Therefore $Px = x$. Also, $\mathcal{M}$ is a subset of $P\mathcal{X}$ so $P\mathcal{M} = \mathcal{M}$. Finally, since $\mathcal{M}$ is in $\mathcal{C}$, $P\mathcal{S}\mathcal{M} \subseteq \mathcal{M}$. Then

$$x = Px \in P\overline{\mathrm{span}(\mathcal{M} \cup \mathcal{S}\mathcal{M})} = \overline{\mathrm{span}(P\mathcal{M} \cup P\mathcal{S}\mathcal{M})} = \overline{\mathrm{span}(\mathcal{M} \cup \mathcal{M})} = \mathcal{M},$$

which is a contradiction, so $x \notin \mathcal{M}_1$ and the inclusion is proper.

$\square$

4.2.7 Lemma

Let $\mathcal{S}$ be an irreducible semigroup of operators of rank at most one in $\mathcal{B}(\mathcal{X})$, with $\mathcal{X}$ infinite dimensional.

(i) For each positive integer k, there is a k-dimensional subspace $\mathcal{M}$ of $\mathcal{X}$ and a subsemigroup $\mathcal{S}_0$ of $\mathcal{S}$ leaving $\mathcal{M}$ invariant such that $\mathcal{S}_0|_{\mathcal{M}}$ is irreducible.

(ii) $\mathcal{S}$ contains members A and B with independent ranges R_1 and R_2 such that the restrictions of A and B to $R_1 + R_2$ are simultaneously similar to

$$A_0 = \begin{pmatrix} \alpha & 0 \\ \beta & 0 \end{pmatrix} \qquad B_0 = \begin{pmatrix} 0 & \gamma \\ 0 & \delta \end{pmatrix}$$

with $\alpha\delta - \beta\gamma \neq 0$ and $\beta\gamma \neq 0$.

Proof. (ii) follows directly from (i) and Lemma 3.3.5.

(i) We first show that for any f in $\mathcal{X}^*$, $\mathcal{S}^*f$ is infinite dimensional. Assume otherwise. Then there are $S_1, \ldots, S_n$ in $\mathcal{S}$ such that $S_1^*f, \ldots, S_n^*f$ span $\mathcal{S}^*f$. Each S_i is rank one and $\mathcal{X}$ is infinite dimensional, so $\cap\{\ker(S_i) : 1 \leq i \leq n\}$ is nontrivial. Take x from this intersection. Then, for any S in $\mathcal{S}$, $f(Sx) = (S^*f)(x)$. But S^*f is in $\mathcal{S}^*f$ and $S_i^*f(x) = f(S_ix) = 0$. Therefore $f(Sx) = 0$ and $\mathcal{S}$ is reducible by Lemma 4.2.5. This is a contradiction, so $\mathcal{S}^*f$ is infinite dimensional.

Fix a positive integer k and take a nonzero $K = x \otimes f$ from $\mathcal{S}$ with x in $\mathcal{X}$ and f in $\mathcal{X}^*$. Since $\mathcal{S}^*f$ is infinite dimensional it contains k linearly independent functionals $f_j = T_j^*f$ for $1 \leq j \leq k$. Take $\mathcal{N} = \cap\{\ker(f_j) : 1 \leq j \leq k\}$ and note that $\mathcal{N}$ will have codimension k in $\mathcal{X}$ since every T_j is rank one. Since $\mathcal{S}$ is irreducible, $\mathcal{S}x$ spans $\mathcal{X}$ (otherwise, it's a nontrivial invariant subspace). We can therefore find elements $S_1, \ldots, S_k$ so that $S_1x, \ldots, S_kx$ are linearly independent and span a complement of $\mathcal{N}$. Call this complement $\mathcal{M}$ and let $x_i = S_ix$ for $1 \leq i \leq k$.

Let $\mathcal{S}_0$ be the subsemigroup of $\mathcal{S}$ generated by the elements

$$S_iKT_j = (S_ix) \otimes (T_j^*f) = x_i \otimes f_j$$

for $1 \leq i, j \leq k$. Since $\mathcal{M}$ includes all the x_i's, $\mathcal{M}$ is invariant for $\mathcal{S}_0$. Since the x_i's and f_j's are linearly independent, the k^2 operators that generate $\mathcal{S}_0$ are linearly independent. Also, they are zero on $\mathcal{N}$ by definition so they are of the form $A_{ij} \oplus 0$ relative to $\mathcal{M} \oplus \mathcal{N}$. Therefore $\mathcal{S}_0|_{\mathcal{M}}$ must span a space of dimension k^2. Therefore it contains a basis for $\mathcal{B}(\mathcal{M})$ and is irreducible.

$\square$

We can now prove the major result of this section.

4.2.8 Lemma (The Downsizing Lemma)

Let $\mathcal{P}$ be a property defined for semigroups in $\mathcal{K}(\mathcal{X})$. Assume that whenever the semigroup $\mathcal{S}$ has property $\mathcal{P}$, so do

(i) every subsemigroup of $\mathcal{S}$,

(ii) $\mathcal{S}|_{\mathcal{X}_0}$, where $\mathcal{X}_0 = \operatorname{span}\{\operatorname{ran}(S) : S \in \mathcal{S}\}$, and

(iii) the semigroup $\overline{\mathbb{R}^+\mathcal{S}}$.

Let $\mathcal{S}$ *be an irreducible semigroup in* $\mathcal{K}(\mathcal{X})$ *with property* $\mathcal{P}$. *Then there is an integer* $k \geq 2$ *and an idempotent* E *of rank* k *on* $\mathcal{X}$ *such that* $\mathcal{S}$ *contains a subsemigroup* $\mathcal{S}_0 = E\mathcal{S}_0$, *where* $\mathcal{S}_0|_{\mathrm{ran}(E)}$ *is an irreducible semigroup in* $M_k(\mathbb{C})$ *with property* $\mathcal{P}$. *Moreover,* E *can be chosen from* $\overline{\mathbb{R}^+\mathcal{S}}$ *if the minimal positive rank in* $\overline{\mathbb{R}^+\mathcal{S}}$ *is greater than* 1.

Proof. By (iii), we can assume that $\mathcal{S} = \overline{\mathbb{R}^+\mathcal{S}}$. By Turovskii's Theorem (4.2.2), $\mathcal{S}$ does not consist entirely of quasinilpotent operators. By Lemma 4.2.1, $\mathcal{S}$ contains a finite rank operator that is either nilpotent or idempotent.

If $\mathcal{S}$ contains a rank one operator then let $\mathcal{J} \neq \{0\}$ be the ideal of operators in $\mathcal{S}$ of rank at most one. Then $\mathcal{J}$ is irreducible by Lemma 4.2.3 and has property $\mathcal{P}$ by (i). By Lemma 4.2.7, there is a 2-dimensional subspace $\mathcal{M}$ and a subsemigroup $\mathcal{S}_0$ of $\mathcal{J}$ leaving $\mathcal{M}$ invariant such that $\mathcal{S}_0|_{\mathcal{M}}$ is irreducible. Taking E to be an idempotent with range $\mathcal{M}$ completes the proof.

Now, let $k \geq 2$ be the smallest nonzero rank in $\mathcal{S}$. Let F be a member in $\mathcal{S}$ of rank k. We claim that $\mathcal{S}_1 = F\mathcal{S}|_{\mathrm{ran}(F)}$ is an irreducible semigroup in $M_k(\mathbb{C})$. Since $\mathcal{S}$ is irreducible, $\{Sx : S \in \mathcal{S}\}$ spans $\mathcal{X}$ for any $x \neq 0$. Specifically, this is true for any nonzero x in $\mathrm{ran}(F)$. Therefore $\{FSx : S \in \mathcal{S}\}$ spans $\mathrm{ran}(F)$ for any nonzero x in $\mathrm{ran}(F)$. This shows that $\mathcal{S}_0$ is nonzero and is, in fact, an irreducible semigroup.

Our next claim is that every element of $\mathcal{S}_1$ has rank either zero or k. Obviously, every element has rank at most k. Assume $FS|_{\mathrm{ran}(F)}$ has nonzero rank less than k. Then it has zero in its spectrum. Then FSF is a nonzero element of $\mathcal{S}$ whose range is exactly $FS(\mathrm{ran}(F))$ which has dimension at most $k-1$ since $FS|_{\mathrm{ran}(F)}$ has rank less than k. But FSF is nonzero element of rank less than k in $\mathcal{S}$, which is a contradiction.

By Lemma 2.3.3, since $\mathcal{S}_1$ is irreducible with minimum rank k, it contains an idempotent of rank k which must be $I_{\mathrm{ran}F}$. By the construction of $\mathcal{S}_1$, $\mathcal{S}$ contains an element of the form

$$E = \begin{pmatrix} I & C \\ 0 & 0 \end{pmatrix},$$

relative to a decomposition of $\mathcal{X}$ using $\mathrm{ran}(F)$ and one of its complements. Then E is an idempotent in $\mathcal{S}$ of rank k with $\mathrm{ran}(E) = \mathrm{ran}(F)$. Then $EF = F$, so if $\mathcal{S}_0 = E\mathcal{S}$ then $\mathcal{S}_0|_{\mathrm{ran}(E)}$ contains $\mathcal{S}_1$. Thus $\mathcal{S}_0$ is irreducible and has property $\mathcal{P}$ by (i) and (ii).

□

4.3 Subadditive and Sublinear Spectra

The definitions of subadditivity and sublinearity of the spectrum remain the same as in finite dimensions (Definition 3.3.1). Our first new proof is an analogue of Theorem 3.3.3.

4.3.1 Theorem

Let A and B be operators in $\mathcal{K}(\mathcal{X})$ with a common invariant subspace $\mathcal{M}$. If the spectrum is sublinear on A and B then it is sublinear on $A|_{\mathcal{M}}$ and $B|_{\mathcal{M}}$. The same holds for subadditivity if A and B are at most rank one.

Proof. Assume the spectrum is sublinear on A and B. By the sublinearity of the spectrum, every eigenvalue of $A + \lambda B$ is of the form $\alpha + \lambda\beta$ for α in $\sigma(A)$ and β in $\sigma(B)$. Since $\mathcal{M}$ is a common invariant subspace for A and B, each eigenvalue of $C_\lambda = A_0 + \lambda B_0$ is also of this form. We want to show that for every eigenvalue of C_λ, α is in $\sigma(A_0)$ and each β is in $\sigma(B_0)$.

For every (α, β) in $(\sigma(A) \times \sigma(B))\backslash(\sigma(A_0) \times \sigma(B_0))$ define

$$\mathcal{F}_{(\alpha,\beta)} = \{\lambda : \alpha + \lambda\beta \in \sigma(C_\lambda)\}.$$

Showing that the spectrum is sublinear amounts to showing that every such $\mathcal{F}_{(\alpha,\beta)}$ is empty.

We claim that each $\mathcal{F}_{(\alpha,\beta)}$ is closed and nowhere dense. Let $\{\lambda_n\}$ be a sequence in $\mathcal{F}_{(\alpha,\beta)}$ converging to λ. Then $\{C_{\lambda_n}\}$ must converge to C_λ. By Lemma 4.1.13, $\alpha + \lambda\beta$ is in $\sigma(C_\lambda)$ so λ is in $\mathcal{F}_{(\alpha,\beta)}$, so the set is closed.

Assume that $\mathcal{F}_{(\alpha,\beta)}$ is not nowhere dense. Then it's uncountable, so C_λ is not invertible for uncountably many λ. Since we chose $(\sigma(A) \times \sigma(B))\backslash(\sigma(A_0) \times \sigma(B_0))$ either α is not in $\sigma(A_0)$ or β is not in $\sigma(B_0)$. If α is not in $\sigma(A_0)$ then $A_0 - \alpha$ is invertible so

$$\lambda^{-1}C_\lambda(A_0 - \alpha)^{-1} = \lambda^{-1} + (B_0 - \beta)(A_0 - \alpha)^{-1}$$

is not invertible for an uncountable number of λ's. But it's the translate of the compact operator $(B_0 - \beta)(A_0 - \alpha)^{-1}$ so it should have at most countable spectrum.

A similar contradiction is reached if β is not in $\sigma(B_0)$. Therefore $\mathcal{F}_{(\alpha,\beta)}$ is nowhere dense.

Let $\mathcal{F}$ be the union of all the $\mathcal{F}_{(\alpha,\beta)}$ and take $\mathcal{E} = \mathbb{C}\backslash\mathcal{F}$. Then $\mathcal{F}$ is the finite union of nowhere dense sets so $\mathcal{E}$ is a dense G_δ set by the Baire Category Theorem. If $\mathcal{E} = \mathbb{C}$ then each $\mathcal{F}_{(\alpha,\beta)}$ is empty and the spectrum is sublinear on A_0 and B_0. Since $\mathcal{E}$ is dense we need only show that it is closed.

Let $\{\lambda_n\}$ be a sequence in $\mathcal{E}$ converging to λ. Then $\{C_{\lambda_n}\}$ converges to C_λ. By Lemma 4.1.13

$$\sigma(C_\lambda) = \lim_{n\longrightarrow\infty} \sigma(C_{\lambda_n}).$$

Take a convergent sequence $\{\alpha_n+\lambda_n\beta_n\}$ with $\alpha_n+\lambda_n\beta_n$ in $\sigma(C_{\lambda_n})$. Then (α_n, β_n) is in $\sigma(A_0)\times\sigma(B_0)$ since λ_n is in $\mathcal{E}$. As $\sigma(A_0)$ and $\sigma(B_0)$ are finite, by taking subsequences

$$\alpha = \lim_{n\longrightarrow\infty} \alpha_n \qquad\qquad \beta = \lim_{n\longrightarrow\infty} \beta_n$$

and α is in $\sigma(A_0)$ and β is in $\sigma(B_0)$. Then

$$\lim_{n\longrightarrow\infty} \alpha_n + \lambda_n\beta_n = \alpha + \lambda\beta,$$

so $\sigma(C_\lambda)$ is contained in $\sigma(A_0) + \lambda\sigma(B_0)$. Therefore λ is not in $\mathcal{F}$, so λ is in $\mathcal{E}$. Therefore $\mathcal{E} = \mathbb{C}$ and the spectrum is sublinear on A_0 and B_0.

The proof for subadditivity on rank one operators is unchanged from the finite dimensional case.

$\square$

We next need to show inheritability by quotients in infinite dimensions (an analogue to Corollary 3.3.4).

4.3.2 Corollary

Sublinearity of spectrum for compact operators is inherited by quotients. Subadditivity is inherited by quotients if the operators have rank at most one.

Proof. Let A and B be operators in $\mathcal{K}(\mathcal{X})$ with sublinear spectrum. Due to Theorem 4.3.1 we need only show that, for an invariant subspace $\mathcal{M}$, $\hat{A}$ and $\hat{B}$, the induced operators on $\mathcal{X}/\mathcal{M}$, have sublinear spectrum.

Let $\mathcal{M}^\perp$ denote the annihilator of $\mathcal{M}$, the set of elements in $\mathcal{X}^*$ that vanish on $\mathcal{M}$ (if $\mathcal{X}$ is a Hilbert space this is simply the perpendicular space of $\mathcal{M}$). Then there is an isometric isomorphism from $\mathcal{M}^\perp$ onto $(\mathcal{X}/\mathcal{M})^*$ (see Rudin [8, p. 96]).

As A and B leave $\mathcal{M}$ invariant, A^* and B^* leave $\mathcal{M}^\perp$ invariant. Since $\sigma(K^*) = \sigma(K)$ for any K in $\mathcal{K}(\mathcal{X})$, A^* and B^* have sublinear spectrum. By Lemma 4.3.2, $A^*|_{\mathcal{M}^\perp}$ and $B^*|_{\mathcal{M}^\perp}$ have sublinear spectrum.

We can then use the isometric isomorphism to identify $A^*|_{\mathcal{M}^\perp}$ with $A^*|_{(\mathcal{X}/\mathcal{M})^*}$ so they share the same spectrum. And $\sigma(\hat{K}) = \sigma(K^*|_{(\mathcal{X}/\mathcal{M})^*})$ for any K in $\mathcal{K}(\mathcal{X})$ invariant on $\mathcal{M}$. Therefore $A|_{\mathcal{X}/\mathcal{M}}$ and $B|_{\mathcal{X}/\mathcal{M}}$ have sublinear spectrum.

The proof for subadditivity is unchanged from the finite dimensional case.

□

Subadditivity remains sufficient for triangularizability of rank one operators.

4.3.3 Theorem

Let $\mathcal{S}$ be a semigroup of operators of rank at most one in $\mathcal{B}(\mathcal{X})$ with subadditive spectrum. Then $\mathcal{S}$ is triangularizable.

Proof. By Lemma 4.3.2 and the Triangularization Lemma (4.1.6), it suffices to show $\mathcal{S}$ is reducible. Assume $\mathcal{S}$ is irreducible.

By Lemma 4.2.7 (i), we have a subsemigroup $\mathcal{S}_0$ of $\mathcal{S}$ and a two dimensional subspace $\mathcal{M}$ which is invariant for $\mathcal{S}_0$ such that $\mathcal{S}_0|_{\mathcal{M}}$ is irreducible. Then $\mathcal{S}_0$ consists of operators of rank at most one with subadditive spectrum. But by Theorem 3.3.6, $\mathcal{S}_0$ is triangularizable which contradicts irreducibility. Therefore $\mathcal{S}$ is reducible.

□

We can also extend Theorem 3.3.12.

4.3.4 Theorem

Every semigroup of operators in $\mathcal{K}(\mathcal{X})$ with sublinear spectrum is triangularizable.

Proof. By Corollary 4.3.2, sublinear spectrum is inherited by quotients so showing reducibility is sufficient by the Triangularization Lemma (4.1.6). Since Theorem 3.3.12

shows that sublinear spectrum is sufficient for triangularizability in finite dimensions, if we can show it satisfies the conditions of the Downsizing Lemma (4.2.8), then we're done.

For (i), every subset of a set with sublinear spectrum clearly has sublinear spectrum. For (ii), $\mathcal{X}_0$ is an invariant subspace for $\mathcal{S}$ so the property holds by Corollary 4.3.2. For (iii), if $\mathcal{S}$ has sublinear spectrum, then so does $\mathbb{R}^+\mathcal{S}$ as $\sigma(mA) = m\sigma(A)$ for any m in $\mathbb{R}^+$. Also, the spectrum is continuous by Lemma 4.1.13, so $\overline{\mathbb{R}^+\mathcal{S}}$ has sublinear spectrum.

Therefore the conditions of the Downsizing Lemma (4.2.8) are met and the result follows.

$\square$

As in finite dimensions, pairwise triangularizability is sufficient.

4.3.5 Corollary

If every pair of operators in a semigroup $\mathcal{S}$ in $\mathcal{K}(\mathcal{X})$ is triangularizable, then so is $\mathcal{S}$ itself.

Proof. If every pair is triangularizable then every pair has sublinear spectrum. Therefore the semigroup has sublinear spectrum so by Theorem 4.3.4 the entire semigroup is triangularizable.

$\square$

If we restrict ourselves to a Hilbert space $\mathcal{H}$ we can achieve diagonalizability of self-adjoint families.

4.3.6 Corollary

If $\mathcal{S}$ is a self-adjoint semigroup of operators in $\mathcal{K}(\mathcal{H})$ with sublinear spectrum, then $\mathcal{S}$ is abelian.

Proof. By Theorem 4.3.4, there is a triangularizing chain $\mathcal{C}$ for $\mathcal{S}$. If $\mathcal{M}$ is in $\mathcal{C}$ and $\mathcal{M}_- \neq \mathcal{M}$ then $\mathcal{M} \ominus \mathcal{M}_-$ is one dimensional. Since $\mathcal{S}$ is self-adjoint, it leaves invariant both $\mathcal{M} \ominus \mathcal{M}_-$ and its orthogonal complement.

Take $\mathcal{H}_0$ to be the direct sum of these one dimensional spaces and let $\mathcal{H}_1$ be its orthogonal complement. Then $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$ and $\mathcal{H}_0$ and $\mathcal{H}_1$ are both invariant under $\mathcal{S}$. Then $\mathcal{S}|_{\mathcal{H}_0}$ is diagonal since $\mathcal{S}$ was invariant on each of the one dimensional spaces and their orthogonal complements.

For S in $\mathcal{S}$, $S|_{\mathcal{H}_1}$ is quasinilpotent by Ringrose's Theorem (4.1.9) as all the one dimensional gaps are in $\mathcal{H}_0$. As $\mathcal{S}$ is self-adjoint, S^*S is in $\mathcal{S}$ and $S^*S|_{\mathcal{H}_1}$ is also quasinilpotent so $S|_{\mathcal{H}_1} = 0$.

Therefore $\mathcal{S}$ is diagonalizable and abelian.

$\square$

Sublinearity is actually stronger than necessary.

4.3.7 Theorem

The following conditions are mutually equivalent for a semigroup $\mathcal{S}$ of operators in $\mathcal{K}(\mathcal{X})$:

(i) $\mathcal{S}$ is triangularizable.

(ii) for all integers m, scalars $\lambda_1, \ldots, \lambda_m$ and members $S_1, \ldots, S_m$ of $\mathcal{S}$,

$$\sigma(\lambda_1 S_1 + \cdots + \lambda_m S_m) \subseteq \lambda_1 \sigma(S_1) + \cdots + \lambda_m \sigma(S_m).$$

(iii) $\mathcal{S}$ has sublinear spectrum.

(iv) $\mathcal{S}$ has real sublinear spectrum. That is, $\sigma(A + \lambda B) \subseteq \sigma(A) + \lambda\sigma(B)$ for all real numbers λ and all pairs A and B in $\mathcal{S}$.

Proof. (i) implies (ii) by the Spectral Mapping Theorem (4.1.12), (ii) implies (iii) by taking $m = 2$ and $\lambda_1 = 1$. (iii) clearly implies (iv).

The Baire-Category argument from Theorem 4.3.1 works for real sublinearity as well, so the analogous result to Corollary 4.3.2 holds and real sublinearity is inherited by quotients. It suffices by the Triangularization Lemma (4.1.6) to show that $\mathcal{S}$ is reducible. Since real sublinearity satisfies the conditions of the Downsizing Lemma in the same way as regular sublinearity, it suffices to show that a semigroup of finite dimensional operators with real sublinear spectrum is reducible. This is clear from Theorem 3.3.16(v).

□

In order to extend this result to bounded operators, we need the following definition.

4.3.8 Definition

An operator T in $\mathcal{B}(\mathcal{X})$ is called a *strong quasiaffinity* if $\overline{T\mathcal{M}} = \mathcal{M}$ for every invariant subspace of T.

4.3.9 Corollary

Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{B}(\mathcal{X})$ with sublinear spectrum. If S contains a nonzero compact K, then it is reducible. If K is a strong quasiaffinity then $\mathcal{S}$ is triangularizable.

Proof. Assume $\mathcal{S}$ is irreducible. The ideal of compact operators in $\mathcal{S}$ is nonzero since K is nonzero. By Lemma 4.2.3, it must be irreducible. But, by Theorem 4.3.4, the ideal must be reducible. This is a contradiction so $\mathcal{S}$ must be reducible.

For triangularizability, we want to show that the property of having a nonzero compact strong quasiaffinity is inherited by quotients. If $\mathcal{N} \subset \mathcal{M}$ are invariant subspaces for such a K we want to show that $K_0 = K|_{\mathcal{M}/\mathcal{N}}$ is a strong quasiaffinity.

Let $\mathcal{L}_0$ be a nontrivial invariant subspace for K_0. Then

$$\mathcal{L} = \{x \in \mathcal{M} : [x] \in \mathcal{L}_0\}$$

is a nontrivial invariant subspace for K. Then $\overline{K\mathcal{L}} = \mathcal{L}$, so $\overline{K_0\mathcal{L}_0} = \mathcal{L}_0$ and K_0 is a strong quasiaffinity. It's obviously compact and strong quasiaffinities are nonzero.

Therefore $\mathcal{S}$ is triangularizable.

□

4.4 Polynomial Conditions on Spectra

Many of the results from finite dimensions extend easily to infinite dimensions. We look at those proofs that require significant modification.

Recall that in Definition 3.4.1, for a polynomial $g(x) = \sum_{j=0}^{m} a_j x^j$ we defined a noncommutative, homogeneous polynomial f_g by $f_g(x, y) = \sum_{j=0}^{m} a_j x^j y x^{m-j}$. This definition is unchanged in infinite dimensions, however Definition 3.4.2 requires a slight modification.

4.4.1 Definition

Let $g(x)$ be a polynomial. We say f_g is *quasinilpotent* on a semigroup $\mathcal{S}$ if $f_g(A, B)$ is quasinilpotent for all A and B in $\mathcal{S}$.

Our first result addresses the case of $g(x) = x - 1$.

4.4.2 Theorem

Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{K}(\mathcal{X})$ such that $AB - BA$ is quasinilpotent for all A and B in $\mathcal{S}$. Then $\mathcal{S}$ is triangularizable.

Proof. By the Triangularization Lemma (4.1.6), we need only show reducibility as polynomial quasinilpotence extends to quotients.

Since Theorem 3.4.15 says any semigroup in finite dimensions with this property is triangularizable, it's sufficient to show that this property satisfies the conditions of the Downsizing Lemma (4.2.8). Every subsemigroup of $\mathcal{S}$ clearly satisfies the condition, as does any restriction to an invariant subspaces. Finally, polynomial quasinilpotence is unaffected by scalars and closure. Therefore the conditions are met and $\mathcal{S}$ is triangularizable.

$\square$

In Section 3.4, we saw that there are irreducible semigroups that are nilpotent on all polynomials of degree at least two (Example 3.4.8) and polynomials of degree at least two that are nilpotent on the entire semigroup of rank one operators (Example 3.4.9). These examples can easily be extended. Example 3.4.8 was already extended to arbitrary rank and the extension to infinite dimensions is analogous. Example 3.4.9 extends immediately.

In order to extend Theorem 3.4.12, we'll use the following result.

4.4.3 Lemma

Let $\mathcal{S}$ be an irreducible semigroup of operators in $\mathcal{B}(\mathcal{X})$ with rank at most one. Assume $\mathcal{S} = \mathbb{C}\mathcal{S}$. Let $g(x) = \sum_{i=0}^{m} a_0 x^m$ with $g(0) \neq 0$. If f_g is quasinilpotent on $\mathcal{S}$ then:

(i) the closed linear span of $\{E\mathcal{X} : E \in \mathcal{E}\}$ is $\mathcal{X}$,

(ii) the idempotents in $\mathcal{S}$ form an abelian semigroup $\mathcal{E}$, and

(iii) for each nonzero A in $\mathcal{S}$, there are unique E and F in $\mathcal{E}$ such that EAF is nonzero; moreover, $A = EAF$.

Proof. (i) Let $\mathcal{E}$ be the set of idempotents in $\mathcal{S}$. Let A be an arbitrary nonzero member of $\mathcal{S}$. If $A\mathcal{S}$ contains no nonzero idempotents then it consists entirely of nilpotents. But then, for any S and T in $\mathcal{S}$ and k in $\mathbb{N}$,

$$(SAT)^k = S(ATS)^{k-1}AT,$$

so the ideal of $\mathcal{S}$ generated by A consists of nilpotents. But then it's reducible by Turovskii's Theorem (4.2.2) and $\mathcal{S}$ is reducible by Lemma 4.2.3. This is a contradiction, so $A\mathcal{S}$ contains nonzero idempotents.

Let $E = AB$ be such an idempotent. Then $EA = ABA \neq 0$, so it has rank one. Since it has exactly the same kernel and range as A it must be equal to A. Since E is in $\mathcal{E}$, we see that $\mathcal{E}\mathcal{S} = \mathcal{S}$. Similarly $\mathcal{S}\mathcal{E} = \mathcal{S}$.

Since the closed linear span of $\mathcal{S}\mathcal{X}$ is invariant for $\mathcal{S}$ and $\mathcal{S}$ is irreducible, the closed linear span must be $\mathcal{X}$. But then the closed linear span of $\mathcal{E}\mathcal{X} \supset \mathcal{E}\mathcal{S}\mathcal{X} = \mathcal{S}\mathcal{X}$ must be $\mathcal{X}$ as well, so (i) is proved.

(ii) To prove (ii) we want to show for $E \neq F$ in $\mathcal{E}$ that $EF = 0$. First we'll show that $E\mathcal{X} \neq F\mathcal{X}$. Assume otherwise. Let $E\mathcal{X} = F\mathcal{X} = \mathbb{C}x_0$. Since $E \neq F$, they are rank one, and share the same range they must have distinct kernels as we saw in Theorem 3.4.10. Let ϕ and ψ be elements of $\mathcal{X}^*$ such that $E = x_0 \otimes \psi$ and $F = x_0 \otimes \phi$. Since E and F are idempotents, $\phi(x_0) = \psi(x_0) = 1$.

Let $\mathcal{N} = \ker(\phi) \cap \ker(\psi)$. Then $\mathcal{N}$ must have codimension 2 in $\mathcal{X}$. Since x_0 is not in their kernels, $\mathcal{N} \oplus \mathbb{C}x_0$ must have codimension 1 in $\mathcal{X}$. Take x_1 so that $\mathbb{C}x_1$ is not contained within $\mathcal{N} \oplus \mathbb{C}x_0$. Since $\mathcal{S}$ is irreducible, $\mathcal{S}\mathcal{X}$ must span $\mathcal{X}$ so $\mathcal{S}$ contains an element $T = x_1 \otimes \theta$.

Now, $\mathcal{S}$ is irreducible and $\mathcal{S}x_0 \neq \{0\}$. Since $\mathcal{S}x_0$ is invariant for $\mathcal{S}$, it must span $\mathcal{X}$. Therefore, there is an S in $\mathcal{S}$ such that $\theta(Sx_0) \neq 0$ so x_0 is not in the kernel of $S^*\theta$. Since $(x_1 \otimes \theta)S = x_1 \otimes S^*\theta$ we can assume x_0 is not in the kernel of θ.

Let $\mathcal{M} = \operatorname{span}\{x_0, x_1\}$. Then $\mathcal{M}$ is invariant for E, F, and T. Their restrictions to $\mathcal{M}$ are

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & \beta \\ 0 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 0 & 0 \\ \gamma & \delta \end{pmatrix}$$

respectively. Also $E = A \oplus 0$ and $F = B \oplus 0$ since their kernels have codimension one. Since $E \neq F$, $\alpha \neq \beta$. Since x_0 is not in the kernel of θ, $\gamma \neq 0$. Thus A, B, and C generate an irreducible semigroup $\mathcal{S}_0$ in $\mathcal{M}_2(\mathbb{C})$. However, f_g is nilpotent on $\mathcal{S}_0$ and the degree of g is at least 2 since linear polynomials are quasinilpotent only on reducible semigroups by Theorem 4.4.2. By Theorem 3.4.10, $AB = 0$ which is a clear contradiction. Therefore $E\mathcal{X} \neq F\mathcal{X}$.

Therefore $E = x \otimes \phi$ and $F = y \otimes \psi$ where $\mathbb{C}x \neq \mathbb{C}y$. Relative to the span of x and y we can write

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 \\ \beta & 1 \end{pmatrix}.$$

As $\mathcal{S}$ is irreducible it contains an element S with $\alpha x - y$ not in its kernel. Then $ES = x \otimes S^*\phi$ has a different kernel from E and has a restriction to the span of x and y of

$$C = \begin{pmatrix} \gamma & \delta \\ 0 & 0 \end{pmatrix},$$

which is linearly independent of A. The exact same calculations as in Theorem 3.4.10 apply and $EF = 0$.

(iii) When proving (i) we saw that $\mathcal{S} = \mathcal{E}\mathcal{S} = \mathcal{S}\mathcal{E}$ so $\mathcal{S} = \mathcal{E}\mathcal{S}\mathcal{E}$. This proves existence of E and F. For uniqueness, if $E_1AF_1 = A$ with $E_1 \neq E$ or $F_1 \neq F$ then, by what we just proved for (ii), either $E_1E = 0$ or $FF_1 = 0$. Either way, $A = E_1AF_1 = E_1EAFF_1 = 0$ which is a contradiction.

□

We can now extend Theorem 3.4.12 to a subset of $\mathcal{K}(\mathcal{X})$.

4.4.4 Theorem

Let $g = \sum_{j=0}^{m} a_j x^j$ with $g(0) \neq 0$ and g not divisible by $x^p - 1$ for any prime p. Let $\mathcal{S}$ be a semigroup of strong quasiaffinities in $\mathcal{K}(\mathcal{X})$. If f_g is nilpotent on $\mathcal{S}$ then $\mathcal{S}$ is triangularizable.

Proof. The quotients of strong quasiaffinities are strong quasiaffinities as seen in Corollary 4.3.9. Quasinilpotence of polynomials is also inherited by quotients. Therefore it's enough to show reducibility by the Triangularization Lemma (4.1.6). Assume $\mathcal{S}$ is irreducible.

Unfortunately, a semigroup consisting of strong quasiaffinities does not satisfy the conditions of the Downsizing Lemma (4.2.8) since limits of strong quasiaffinities can easily be 0. Quasinilpotence of polynomials does satisfy the conditions of the Downsizing Lemma though.

Since f_g is quasinilpotent on $\mathcal{S}$ it is also quasinilpotent on $\overline{\mathbb{C}\mathcal{S}}$ by Lemma 4.1.13. If $\overline{\mathbb{C}\mathcal{S}}$ does not contain rank one operators then the idempotent E of minimal rank $k \geq 2$ from the Downsizing Lemma (4.2.8) can be taken to be in $\overline{\mathbb{C}\mathcal{S}}$. Then everything in the irreducible subsemigroup $\mathcal{S}_0$ is either zero or has rank k. Therefore $\mathcal{S}_0|_{\mathcal{M}} \backslash \{0\}$ must consist of invertibles and is therefore reducible by Theorem 3.4.12. This contradicts the assumption that $\mathcal{S}$ is irreducible.

The only remaining case is when $\overline{\mathbb{C}\mathcal{S}}$ contains rank one operators. Then the ideal $\mathcal{J}$ of rank one operators in $\overline{\mathbb{C}\mathcal{S}}$ is irreducible by Lemma 4.2.3. Let $\mathcal{E}$ be the abelian semigroup of idempotents in $\mathcal{J}$ from Lemma 4.4.3. Then $\mathcal{S}\mathcal{E} = \mathcal{E}\mathcal{S} = \mathcal{J}$ since $\mathcal{J}\mathcal{E} = \mathcal{E}\mathcal{J} = \mathcal{J}$ and everything in $\mathcal{E}$ has rank at most one. Let $\mathcal{E}_1$ be the set of nonzero elements of $\mathcal{E}$ and take E_0 from $\mathcal{E}_1$.

As $\mathcal{S}$ is irreducible, $\{Sx : S \in \mathcal{S}\}$ spans $\mathcal{X}$ for any nonzero x in $\mathcal{X}$. Therefore there is a strong quasiaffinity T in $\mathcal{S}$ such that $TE_0 \neq E_0TE_0$. By Lemma 4.4.3, there are E_1, F in $\mathcal{E}_1$ such that $E_1TE_0F = TE_0$. By uniqueness, $F = E_0$ and $E_1TE_0 = TE_0$. Since T is a strong quasiaffinity, $TE_1 \neq 0$ and we can repeat this argument and get an E_2 in $\mathcal{E}_1$ such that $TE_1 = E_2TE_1 \neq 0$. We can continue this argument and get a sequence $\{E_n\}$ in $\mathcal{E}_1$ with $TE_n = E_{n+1}TE_n \neq 0$ for each n.

There are two cases: the sequence consists of distinct element or there are duplicates.

(1) Assume the sequence $\{E_n\}$ has distinct elements. Since $\mathcal{J}$ is irreducible we

can choose an S in $\mathcal{J}$ such that $R = E_0 S E_m \neq 0$. Since T is a strong quasiaffinity, $T^j R T^{m-j} \neq 0$ and therefore it must be rank one for all j. By repeated use of the property $TE_n = E_{n+1} T E_n$ for all n, we see that

$$T^j E_0 = E_j T^j E_0 \neq 0 \quad \text{and} \quad T^{m-j} E_j = E_m T^{m-j} E_j \neq 0 \ .$$

Combining this with the fact that $E_0 R E_m = R$, we see that

$$E_j T^j R T^{m-j} E_j = E_j T^j E_0 R E_m T^{m-j} E_j \neq 0.$$

By part (iii) of Lemma 4.4.3, $T^j R T^{m-j} = E_j T^j R T^{m-j} E_j$ so, since E_j is a rank one idempotent, $T^j R T^{m-j} = \mu_j E_j$ for some $\mu_j \neq 0$.

We can now apply the quasinilpotence of f_g. Since $\mathcal{E}_1$ is abelian, the sequence $\{E_n\}$ is triangularizable by Theorem 4.1.7. But

$$\sum_{j=0}^{m} a_j \mu_j E_j = \sum_{j=0}^{m} a_j T^j R T^{m-j},$$

with T and R in $\mathcal{S}$. Therefore the a_j must all be zero, which is a contradiction. Thus $\mathcal{S}$ must be reducible.

(2) The last case occurs if there exists positive integers $i < j$ with $E_i = E_j$ and we take the first such pair. As we saw in the first case, $T^k E_0 = E_k T^k E_0$ and T^k is a strong quasiaffinity so, by passing to a power of T, we can take $j = i + p$, p a prime. (We can't have $p = 1$. If it were, then $i \neq 0$ since $E_0 \neq E_1$ by definition. However, the two rank one operators $TE_i = E_i T E_i$ and $TE_{i-1} = E_i T E_{i-1}$ would have the same range. This is impossible as T is injective since it's a strong quasiaffinity and E_i and E_{i-1} rank one idempotents with distinct ranges.)

We can assume $i = 0$ and $j = p$ since $E_i \neq E_{i+1}$. Take $\mathcal{M}$ to be the span of the ranges of $E_0, \ldots, E_{p-1}$. Thus $\mathcal{M}$ is invariant for each of E_i's and also T since $T|_{E_k}$ has its range contained within the range of E_{k+1}. Let $\mathcal{S}_0$ be the semigroup of operators in $\mathcal{M}_p(\mathbb{C})$ generated by $T|_{\mathcal{M}}$ and $E_k|_{\mathcal{M}}$. Since the E_i's are idempotent, every invariant subspace of $\mathcal{S}_0$ would be a span of a subset of their ranges. However, T acts as a cyclical weighted permutation of their ranges. Therefore $\mathcal{S}_0$ is irreducible. Also, f_g is nilpotent on $\mathcal{S}_0$.

We can apply a diagonal similarity to $\mathcal{S}_0$, so that $A = T|_{\mathcal{M}}$ is a cyclical permutation of the ranges of the E_i's and $B = E_0|_{\mathcal{M}} = \mathrm{diag}(1, 0, \ldots, 0)$. The rest of the proof is as in Theorem 3.4.12. Moreover, $f_g(A, BA^{-m})$ is nilpotent so we get a contradiction and $\mathcal{S}$ is reducible.

$\square$

We can extend this result to bounded operators.

4.4.5 Corollary

Let g be as in Theorem 4.4.4 and let $\mathcal{S}$ be a semigroup of strong quasiaffinities in $\mathcal{B}(\mathcal{X})$. If f_g is quasinilpotent on $\mathcal{S}$ and $\mathcal{S}$ contains a nonzero compact operator then $\mathcal{S}$ is triangularizable.

Proof. We saw in Corollary 4.3.9 that having a strong quasiaffinity extends to quotients. Compactness also extends to quotients, as does quasinilpotence of polynomials. Therefore it is sufficient to show reducibility by the Triangularization Lemma (4.1.6).

The ideal of compact operators in $\mathcal{S}$ is nontrivial and it is reducible by Theorem 4.4.4. By Lemma 4.2.3, $\mathcal{S}$ is reducible.

$\square$

We can also extend the result to the following special case in $\mathcal{B}(\mathcal{X})$.

4.4.6 Theorem

Let g be as in Theorem 4.4.4. let $\mathcal{S} = \overline{\mathbb{R}^+\mathcal{S}}$ be any semigroup in $\mathcal{B}(\mathcal{X})$ with f_g quasinilpotent on $\mathcal{S}$. If the minimal rank in $\mathcal{S}$ is r and if $\mathcal{S}$ contains a finite rank idempotent E of rank r then $\mathcal{S}$ has distinct invariant subspaces

$$\{0\} = \mathcal{M}_0 \subset \mathcal{M}_1 \subset \cdots \subset \mathcal{M}_r.$$

Proof. Since r is minimal, $\mathcal{G} = E\mathcal{S}E|_{E\mathcal{X}}\backslash\{0\}$ is a group by Lemma 2.3.3. Now f_g is quasinilpotent on $\mathcal{G}$ so it's nilpotent as $\mathcal{G}$ operates on a finite dimensional space. Therefore $\mathcal{G}$ is triangularizable by Theorem 4.4.4. The result then follows from Lemma 4.2.6.

$\square$

4.5 Permutability of the Trace

In this section, we restrict ourselves to operators on a Hilbert space $\mathcal{H}$. More specifically, we restrict ourselves to trace class operators. The existence of such operators and their properties is not discussed here. For such a discussion, see [5, Chapter 1] or [7, Section 6.5].

4.5.1 Definition (Trace Class Operators)

The trace class operators are those compact operators K for which $\sigma(\sqrt{K^*K})$, listed according to multiplicity, is summable.

In other words, if $\{\lambda_n\}$ is the set of eigenvalues of K^*K then K is in the trace class if and only if $\sum_{n=1}^{\infty}\sqrt{\lambda_n} < \infty$ (The eigenvalues of K^*K are in $[0,\infty)$ for every compact operator K). In fact, this sum is a norm on the trace class operators known as the trace norm.

We start with a definition of the trace on the trace class.

4.5.2 Definition

For K in the trace class, we denote the trace of K by $\mathrm{tr}(K)$ and define it as

$$\mathrm{tr}(K) = \sum_{m=1}^{\infty} \langle K g_m, g_m \rangle$$

for any orthonormal basis $\{g_m\}$. This is well-defined ([7, Corollary 6.5.13]), $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, and it clearly reduces to the finite dimensional case if $\mathcal{H}$ is finite dimensional.

We want an analogue to Theorem 3.1.7 on the trace class. We will need the following two results.

4.5.3 Theorem (Lidskii's Theorem)

If K is in the trace class then $\mathrm{tr}(K)$ is the sum of the eigenvalues of K, counting multiplicity.

Proof. This result was originally proved by Lidskii [6].

□

4.5.4 Corollary

The trace of a trace-class operator is the sum of its diagonal coefficients relative to any triangularizing chain

Proof. This follows directly from Lidskii's Theorem (4.5.3) and Theorem 4.1.10.

□

We can now give an analogue for the trace class.

4.5.5 Theorem

Let $\mathcal{F}$ be a family of trace-class operators on a Hilbert space. Then $\mathcal{F}$ is triangularizable if and only if trace is permutable on $\mathcal{F}$.

Proof. If $\mathcal{F}$ is triangularizable then trace is permutable by Corollary 4.5.4. We now assume that trace is permutable.

The permutability of the trace will extend to the algebra generated by $\mathcal{F}$ so without loss of generality, $\mathcal{F}$ is an algebra. By Theorem 4.4.2, we need only show that $AB - BA$ is quasinilpotent for all A and B in $\mathcal{F}$.

We know $\operatorname{tr}(AB - BA) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0$. For $n \geq 2$, let $C = AB - BA$ and then

$$\operatorname{tr}(C^n) = \operatorname{tr}(ABC^{n-1}) - \operatorname{tr}(BAC^{n-1}) = 0$$

by permutability so $\operatorname{tr}((AB - BA)^n) = 0$ for all n in $\mathbb{N}$.

By Lidskii's Theorem (4.5.3), if we take the sequence of eigenvalues of $AB - BA$ to be $\{\lambda_i\}$ then

$$\sum_i \lambda_i^n = \operatorname{tr}((AB - BA)^n) = 0$$

for all n in $\mathbb{N}$. We claim that this means that $\lambda_i = 0$ for all i. Assume otherwise. We can assume that $|\lambda_1| \geq |\lambda_2| \geq \cdots$. In particular, since not all the λ_i are zero, $|\lambda_1| \neq 0$. We can then divide every λ_i by λ_1 (without affecting the sums) so that

$$1 = |\lambda_1| = |\lambda_2| = \cdots = |\lambda_k| > |\lambda_{k+1}| \geq \cdots .$$

Note that, since $AB - BA$ is trace class, $\sum_{i=1}^{\infty} |\lambda_i| < \infty$ ([5, Theorem 1.3] or [7, Lemma 6.5.10]).

We claim that $\lim\limits_{n\longrightarrow\infty} \sum_{i=k+1}^{\infty} \lambda_i^n = 0$. For any $\varepsilon > 0$, since the λ_i converge absolutely, we can find an $M > k$ such that $\sum_{i=M}^{\infty} |\lambda_i| < \frac{\varepsilon}{2}$. Since $|\lambda_i| < 1$ for $i \geq M$,

$$\sum_{i=M}^{\infty} \lambda_i^n \leq \sum_{i=M}^{\infty} |\lambda_i| < \frac{\varepsilon}{2}$$

for all n. And for each $k < i < M$, $|\lambda_i| < 1$ so λ_i^n converges to zero. Then for n large enough, we have $\sum_{i=k+1}^{M-1} \lambda_i^n < \frac{\varepsilon}{2}$. Therefore $\lim\limits_{n\longrightarrow\infty} \sum_{i=k+1}^{\infty} \lambda_i^n = 0$.

Since $\sum_i \lambda_i^n = 0$, we have $\lim\limits_{n\longrightarrow\infty} \lambda_1^n + \cdots + \lambda_k^n = 0$. Now, each λ_i^n is bounded so there is a subsequence n_j such that $\lambda_i^{n_j}$ converges to some μ_i for $1 \leq i \leq k$ with $|\mu| = 1$. And for any m in $\mathbb{N}$, $\lambda_i^{n_j m}$ converges to μ_i^m, so

$$\mu_1^m + \cdots + \mu_k^m = 0.$$

By Lemma 3.1.5, each $\mu_i = 0$ which contradicts that there are nonzero λ_i.

Therefore $AB - BA$ is quasinilpotent for every A and B in $\mathcal{F}$ so $\mathcal{F}$ is triangularizable.

$\square$

We can make this even simpler in the case of semigroups.

4.5.6 Lemma

Let ϕ be a linear functional on a semigroup $\mathcal{S}$ of operators in $\mathcal{B}(\mathcal{H})$. Then ϕ is permutable on $\mathcal{S}$ if and only if both

(i) $\phi(ST) = \phi(TS)$, *and*

(ii) $\phi(STR) = \phi(TSR)$

for all R, S, and T in $\mathcal{S}$.

Proof. The proof of this fact is exactly the same as Lemma 3.1.2.

$\square$

4.5.7 Corollary

A semigroup $\mathcal{S}$ of trace-class operators is triangularizable if and only if

$$\operatorname{tr}(ABC) = \operatorname{tr}(BAC)$$

for all A, B, and C in $\mathcal{S}$.

Proof. This follows directly from Theorem 4.5.5 and Lemma 4.5.6.

$\square$

Finally, we also have an analogue for Corollary 3.1.10.

4.5.8 Corollary

Let $\mathcal{F}$ be a self-adjoint family of trace-class operators. Then $\mathcal{F}$ is abelian if and only if trace is permutable on $\mathcal{F}$.

Proof. Commutativity clearly implies a permutable trace. For the other direction, we know that $\mathcal{F}$ is triangularizable by Theorem 4.5.5 and therefore so is $\mathcal{S}$, the semigroup generated by $\mathcal{F}$. Now $\mathcal{S}$ is also self-adjoint and, as it's triangularizable, it has sublinear spectrum. Therefore it is abelian by Corollary 4.3.6.

$\square$

Chapter 5

Commutators and Spectral Radius

We saw in Theorem 3.4.15 that if $AB - BA$ is nilpotent for every pair of operators A and B in a semigroup $\mathcal{S}$ then $\mathcal{S}$ is triangularizable. This is clearly a necessary condition for triangularizability and can be rephrased as $\rho(AB - BA) = 0$ for every A and B in $\mathcal{S}$.

In this section, we investigate whether we can loosen this condition and still retain triangularizability or, at least, reducibility. In particular, how small does $\rho(AB-BA)$ have to be relative to $\rho(A)$ and $\rho(B)$ before we get triangularizability (at which point, $AB - BA$ will automatically be nilpotent). These results were originally published in a paper by Janez Bernik and Heydar Radjavi [1].

We have also seen that a permutable trace is sufficient for triangularizability (Theorem 3.1.7 and Theorem 4.5.5). This condition can be similarly weakened so that an approximately permutable trace will still lead to triangularizability ([2]). However, this will not be dealt with here.

5.1 Compact Groups

We first consider the case of compact groups and discover that $\sqrt{3}$ is sufficiently small and, in fact, sharp. Note that compact groups of matrices are simultaneously similar to unitary groups (Theorem 2.3.1), so $\rho(A) = 1$ for every A in a compact group.

5.1.1 Theorem

Let $\mathcal{G}$ be a compact group of invertible operators in $\mathcal{B}(\mathcal{V})$ such that $\rho(AB-BA) < \sqrt{3}$ for every A and B in $\mathcal{G}$. Then $\mathcal{G}$ is abelian.

Proof. Assume that $\mathcal{G}$ is not abelian. Then $\mathcal{G}$ is a nonabelian compact group so, by [3], it contains a finite nonabelian group $\mathcal{H}$. We can assume, taking a subgroup if necessary, that $\mathcal{H}$ is a minimal nonabelian group. By Lemma 3.3.10, there is a prime p and a p-dimensional subspace $\mathcal{M}$ which is invariant under $\mathcal{H}$ such that $\mathcal{H}|_{\mathcal{M}}$ is generated by two elements of the form

$$A = \alpha \begin{pmatrix} 0 & & & 1 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix} \qquad B = \beta \begin{pmatrix} \theta_1 & & & \\ & \theta_2 & & \\ & & \ddots & \\ & & & \theta_p \end{pmatrix},$$

where α and β are roots of unity, the θ_i are q^{th} roots of unity, q a prime, and B is not scalar. If S and T are in $\mathcal{H}$ then $\rho(S|_{\mathcal{M}}T|_{\mathcal{M}} - T|_{\mathcal{M}}S|_{\mathcal{M}}) = \rho((ST-TS)|_{\mathcal{M}}) \leq \rho(ST-TS) < \sqrt{3}$ so the $\sqrt{3}$ condition holds for $\mathcal{H}|_{\mathcal{M}}$ as well.

For any X and Y in $\mathcal{B}(\mathcal{V})$ and μ, ν in $\mathbb{C}$ with $|\mu| = |\nu| = 1$, we have

$$\rho[(\mu X)(\nu Y) - (\nu Y)(\mu X)] = \rho(XY - YX),$$

so we can assume $\alpha = \beta = 1$ and that some $\theta_i = 1$. Since B isn't scalar, we can replace B with $A^j B A^{-j}$ for some j so that $\theta_1 = 1$ and $\theta_p \neq 1$. Then for every n in $\mathbb{N}$, $AB^nA^{-1} - B^n$ is a diagonal matrix with $\theta_p^n - 1$ as its first entry. Since θ_p is a primitive q^{th} root of unity, θ_p^n takes on all q^{th} roots of unity as n ranges over $\mathbb{N}$. Therefore there is a value of n such that $\left|\theta_p^n - 1\right| \geq \sqrt{3}$. For this n,

$$\rho((AB^n)A^{-1} - A^{-1}AB^n) = \rho(AB^nA^{-1} - B^n) \geq \left|\theta_p^n - 1\right| \geq \sqrt{3},$$

which contradicts that $\mathcal{H}|_{\mathcal{M}}$ satisfies the $\sqrt{3}$ inequality.

□

The $\sqrt{3}$ bound is sharp.

5.1.2 Example

The group $\mathcal{G}$ in $\mathcal{M}_3(\mathbb{C})$ generated by

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta & 0 \\ 0 & 0 & \zeta^2 \end{pmatrix},$$

where ζ is a primitive third root of unity, is irreducible and satisfies the relationship

$$\rho(CD - DC) \leq \sqrt{3}$$

for every C and D in $\mathcal{G}$.

Proof. Since B is nonscalar and diagonal, $\{A, B\}$ is irreducible by Lemma 3.3.9, so $\mathcal{G}$ is irreducible.

As for the spectral radius condition, note that every element of $\mathcal{G}$ turns out to be an assignment of values from the set $\{1, \zeta, \zeta^2\}$ to one of the three disjoint sets of entries (those corresponding to the nonzero entries of I, A, and A^2), along with the matrices I, A, and A^2. We can also check that if you take two matrices C and D of this form then $CD - DC = (1 - \zeta^n)CD$ where n is either $0, 1$, or 2. Since CD is a matrix in $\mathcal{G}$, $\sigma(CD)$ is either $\{1\}$ or $\{1, \zeta, \zeta^2\}$. Thus $\rho(CD) = 1$ and $\rho(CD - DC) = |1 - \zeta^n|$ which is either 0 or $\sqrt{3}$, depending on n.

$\square$

5.2 Semigroups

We now move on to semigroups and investigate the effect of the following condition.

5.2.1 Definition (The $\sqrt{3}$ Condition)

We say that a semigroup $\mathcal{S}$ of operators in $\mathcal{B}(\mathcal{V})$ satisfies the $\sqrt{3}$ condition if there is exists $0 < \varepsilon < \sqrt{3}$ such that

$$\rho(AB - BA) \leq \varepsilon \rho(A)\rho(B)$$

for every A, B in $\mathcal{S}$.

The following lemma will be useful.

5.2.2 Lemma

Let S be a nonnilpotent operator in $\mathcal{B}(\mathcal{V})$ and let r be the number of eigenvalues (counting multiplicities) of S with maximum modulus. Let $\mathcal{S}$ be the semigroup generated by S. Then $\overline{\mathbb{C}\mathcal{S}}$ contains either an idempotent of rank r or a nonzero nilpotent of rank strictly less than r.

Proof. This was actually shown during the proof of Lemma 2.3.3. Assume that $\overline{\mathbb{C}\mathcal{S}}$ contains no nonzero nilpotent of rank strictly less than r. We want to show that $\overline{\mathbb{C}\mathcal{S}}$ contains an idempotent of rank r. Since we are concerned with $\overline{\mathbb{C}\mathcal{S}}$, we may assume that $\rho(S) = 1$.

We can express S as

$$S = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix},$$

where $\sigma(B)$ is on the unit circle and $\rho(C) < 1$. Further, we can assume B is in Jordan form so $B = U + N$ where U is unitary, N is nilpotent, and $NU = UN$. Note that B acts on a space of dimension r since S has r eigenvalues (counting multiplicity) of maximum modulus.

As in Lemma 2.3.3, we can show that, if $N \neq 0$, then $\overline{\mathbb{C}\mathcal{S}}$ contains

$$\begin{pmatrix} N^k & 0 \\ 0 & 0 \end{pmatrix},$$

where $N^k \neq 0$, but $N^{k+1} = 0$. However, since N acts on a space of dimension r and is nilpotent, this operator has rank less than r. It is also nonzero and nilpotent. This contradicts our initial assumption so $N = 0$.

Then, as in Lemma 2.3.3, a sequence of powers of S converges to

$$\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix},$$

with I acting on the same space as B, which has dimension r and the result is proved.

5.2.3 Definition

A semigroup $\mathcal{S}$ of operators in $\mathcal{B}(\mathcal{V})$ is said to be *totally reducible* if $\mathcal{V}$ decomposes as $\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_m$ with each $\mathcal{V}_i$ invariant for $\mathcal{S}$ and each $\mathcal{S}|_{\mathcal{V}_i}$ irreducible. In particular, $\mathcal{S}$ has a block diagonal form with respect to the V_i's.

This condition is equivalent to the existence of mutually orthogonal idempotents $P_1, \ldots, P_m$ such that $\sum_{i=1}^m P_i = I$ with $\mathcal{S} = \sum_{i=1}^m P_i \mathcal{S} P_i$, and where $P_i \mathcal{S} P_i|_{\mathrm{ran}(P_i)}$ is irreducible. Note that these P_i commute with every S in $\mathcal{S}$. In particular, $P_i S P_i = P_i S = S P_i$.

We can now prove our first result on semigroups with the $\sqrt{3}$ condition.

5.2.4 Lemma

Let $\mathcal{S}$ be a totally reducible semigroup of operators in $\mathcal{B}(\mathcal{V})$ satisfying the $\sqrt{3}$ condition. Then $\mathcal{S}$ contains no nonzero nilpotents.

Proof. Let $P_1, \ldots, P_m$ be the complete set of mutually orthogonal projections onto minimal invariant subspaces for $\mathcal{S}$. Assume that $\mathcal{S}$ contains a nonzero nilpotent N. We can assume without loss of generality that $N^2 = 0$ as N^k is an element of $\mathcal{S}$ for every k in $\mathbb{N}$.

For any A in $\mathcal{S}$, $\rho(AN - NA) = 0$ by the $\sqrt{3}$ condition as $\rho(N) = 0$. Therefore $AN - NA$ is nilpotent. Also, there must exist a j such that $P_j N \neq 0$. When we pass to $P_j \mathcal{S} P_j$, $AN - NA$ remains nilpotent and N remains a nonzero nilpotent. We may therefore assume $\mathcal{S}$ is irreducible and $m = 1$.

Since N is a nonzero nilpotent with $N^2 = 0$, we may write it as

$$N = \begin{pmatrix} 0 & X \\ 0 & 0 \end{pmatrix},$$

with respect to the decomposition $\mathcal{V} = \ker(N) \oplus \ker(N)^{\perp}$ where $X \neq 0$. With respect to the same decomposition,

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}.$$

Therefore

$$AN - NA = \begin{pmatrix} -A_3 X & (A_1 - A_4) X \\ 0 & A_3 X \end{pmatrix},$$

so since $AN - NA$ is nilpotent, the only eigenvalues of $A_3 X$ are zero. Hence $\mathrm{tr}(A_3 X) = 0$. But $\phi(A) = \mathrm{tr}(A_3 X)$ is nonzero linear functional on $\mathcal{B}(\mathcal{V})$ which is zero on $\mathcal{S}$, so $\mathcal{S}$ is reducible by Lemma 2.2.10. This is a contradiction, so $\mathcal{S}$ contains no nonzero nilpotents.

$\square$

Unfortunately, even with arbitrarily small ε, there are examples of irreducible semigroups in $\mathcal{B}(\mathcal{V})$ that satisfy the $\sqrt{3}$ condition. Consider the following example.

5.2.5 Example

For $0 < \delta < 1$*, define*

$$\mathcal{S}_\delta = \left\{ c \begin{pmatrix} 1 & x^* \\ y & yx^* \end{pmatrix} : c \in \mathbb{C}^*, x, y \in \mathbb{C}^{n-1}, \|x\| \leq \delta, \|y\| \leq \delta \right\}.$$

Every $\mathcal{S}_\delta$ *is an irreducible semigroup and for any* $\varepsilon > 0$*, there is a delta such that* $\rho(AB - BA) < \varepsilon\rho(A)\rho(B)$ *for every* A *and* B *in* $\mathcal{S}_\delta$*.*

Proof. That $\mathcal{S}_\delta$ is a semigroup is immediate. To see that $\mathcal{S}_\delta$ is irreducible, consider the algebra $\mathcal{A}_\delta$ it generates. By taking A to be the matrix with $x = y = 0$ and B to be the matrix with $y = 0$ and x zero in every entry except the $(j-1)^{\text{st}}$, we can see that $A - B$ is a scalar multiple of E_{1j}. By using C as the matrix with $x = 0$ and y zero in every entry except the $(k-1)^{\text{st}}$, we can see that $A - C$ is a scalar multiple of E_{k1}. Therefore A_δ contains all the standard basis units so it is $\mathcal{B}(\mathcal{V})$ and $\mathcal{S}_\delta$ is irreducible.

Fix an $\varepsilon > 0$. We want to find an $\mathcal{S}_\delta$ that satisfies $\rho(AB - BA) < \varepsilon\rho(A)\rho(B)$ for every A and B in $\mathcal{S}_\delta$.

Let A be in $\mathcal{S}_\delta$ and calculate $\rho(A)$. $A = \begin{bmatrix} 1 \\ y \end{bmatrix} \begin{bmatrix} 1 & x^* \end{bmatrix}$ so

$$\begin{aligned} A^n &= \begin{bmatrix} 1 \\ y \end{bmatrix} \left(\begin{bmatrix} 1 & x^* \end{bmatrix} \begin{bmatrix} 1 \\ y \end{bmatrix} \right)^{n-1} \begin{bmatrix} 1 & x^* \end{bmatrix} \\ &= (1 + \langle x, y \rangle)^{n-1} \begin{bmatrix} 1 \\ y \end{bmatrix} \begin{bmatrix} 1 & x^* \end{bmatrix}. \end{aligned}$$

Therefore $(\|A^n\|)^{\frac{1}{n}}$ converges to $1 + \langle x, y \rangle$ so $\rho(A) = 1 + \langle x, y \rangle$. Since $\|x\| \leq \delta$ and $\|y\| \leq \delta$, $|\langle x, y \rangle| \leq \delta^2$ so $\rho(A)$ is in the set $[1 - \delta^2, 1 + \delta^2]$. The same, of course, is true for any B in $\mathcal{S}_\delta$.

We now need to look at $\rho(AB - BA)$. Since scalar multiples obviously cancel out of both sides of the $\sqrt{3}$ condition we can assume

$$A = \begin{pmatrix} 1 & x^* \\ y & yx^* \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & u^* \\ v & vu^* \end{pmatrix},$$

so

$$AB = \begin{bmatrix} 1 \\ y \end{bmatrix} [\ 1 \quad x^*\ ] \begin{bmatrix} 1 \\ v \end{bmatrix} [\ 1 \quad u^*\ ] = (1 + \langle x, v \rangle) \begin{bmatrix} 1 \\ y \end{bmatrix} [\ 1 \quad u^*\ ]$$

and

$$BA = (1 + \langle u, y \rangle) \begin{bmatrix} 1 \\ v \end{bmatrix} [\ 1 \quad x^*\ ].$$

Let $r = 1 + \langle x, v \rangle$ and $s = \langle u, y \rangle$. Then

$$AB - BA = \begin{pmatrix} r - s & rv^* - sx^* \\ ry - su & ryv^* - sux^* \end{pmatrix}.$$

We know the spectral radius of $AB - BA$ is less than its norm. All norms are equivalent in finite dimensions so we consider $\|AB - BA\|_1$. Now,

$$\begin{aligned} |r - s| &\le |\langle x, v \rangle| + |\langle u, y \rangle| \le 2\delta^2 \le 2(1 + \delta^2)\delta^2 \le 2(1 + \delta^2)\delta, \\ |rv^* - sx^*| &\le |r| \, \|v\| + |s| \, \|x\| \le 2(1 + \delta^2)\delta, \\ |ry - su| &\le |r| \, \|y\| + |s| \, \|u\| \le 2(1 + \delta^2)\delta, \\ |ryv^* - sux^*| &\le |r| \, \|y\| \, \|v\| + |s| \, \|u\| \, \|x\| \le 2(1 + \delta^2) \le 2(1 + \delta^2)\delta. \end{aligned}$$

Therefore $\|AB - BA\|_1 \le 8(1 + \delta^2)\delta$. Taking $C > 0$ to be the equivalence constant between $\|\cdot\|_1$ and the operator norm, $\rho(AB - BA) \le \|AB - BA\| \le 8C(1 + \delta^2)\delta$.

Given our estimates on spectral radii, we need to choose a δ such that

$$8C(1 + \delta^2)\delta < \varepsilon(1 - \delta^2)^2$$

or, by rearranging,

$$\frac{8C(1 + \delta^2)\delta}{(1 - \delta^2)^2} < \varepsilon.$$

Since the left side goes to zero as δ goes to zero we can find the required δ.

$\square$

Some positive results can be achieved. The semigroups in the above example are rank one. The following results shows that this is not accidental as, among other things, irreducible semigroups satisfying the $\sqrt{3}$ condition must contain rank one idempotents in their homogenized closure.

5.2.6 Theorem

Let $\mathcal{S}$ be a totally reducible semigroup of operators in $\mathcal{B}(\mathcal{V})$ satisfying the $\sqrt{3}$ condition with $P_1, \ldots, P_m$ denoting the complete set of mutually orthogonal idempotents to minimal invariant subspaces of $\mathcal{S}$. Then the following hold:

(i) There exist minimal idempotents in $\overline{\mathbb{C}\mathcal{S}}$.

(ii) Let E be any minimal idempotent in $\overline{\mathbb{C}\mathcal{S}}$. Then the rank of EP_i is either zero or one for all $i = 1, \ldots, m$. In particular, if $\mathcal{S}$ is irreducible, then there is a rank-one idempotent in $\overline{\mathbb{C}\mathcal{S}}$.

Proof. If $\mathcal{S}$ satisfies the $\sqrt{3}$ condition then so does $\overline{\mathbb{C}\mathcal{S}}$, so we can assume without loss of generality that $\mathcal{S} = \overline{\mathbb{C}\mathcal{S}}$. And as I commutes with everything we can also assume that I is in $\mathcal{S}$.

Take any nonzero element A in $\mathcal{S}$. By Lemma 5.2.4, it is nonnilpotent. By Lemma 5.2.2, $\mathcal{S}$ contains idempotents of rank at most equal to A or nonzero nilpotents of rank less than A. However, the second is impossible by Lemma 5.2.4, so $\mathcal{S}$ contains idempotents. By starting with an A of minimal rank, we get an idempotent E of rank at most equal to A. By minimality, E has rank equal to that of A and is an idempotent of minimal rank in $\mathcal{S}$. Therefore it must be minimal in $\mathcal{S}$. This proves (i).

Take any minimal idempotent E in $\mathcal{S}$. Then $E\mathcal{S}E$ is simultaneously similar to scalar multiples of unitaries by Lemma 2.3.3. By Theorem 5.1.1, it's abelian.

Assume P_j is such that $P_jE \neq 0$. By definition, $P_j\mathcal{S}P_j|_{\mathrm{ran}(P_j)}$ is irreducible. Since P_j commutes with $\mathcal{S}$, $P_jE\mathcal{S}EP_j|_{\mathrm{ran}(EP_j)}$ is irreducible by Lemma 2.2.13. But it's also abelian, so it must be on a space of dimension one or it would be reducible. Therefore P_jE has rank one.

$\square$

5.2.7 Corollary

Let $\mathcal{S}$ be a totally reducible semigroup of operators in $\mathcal{B}(\mathcal{V})$ satisfying the $\sqrt{3}$ condition with $P_1, \ldots, P_m$ denoting the complete set of mutually orthogonal idempotents to minimal invariant subspaces of $\mathcal{S}$. If $\overline{\mathbb{C}\mathcal{S}}$ contains a set $E_1, \ldots, E_l$ of mutually orthogonal minimal idempotents of ranks r_i respectively, then the lattice of invariant subspaces of $\mathcal{S}$ contains a chain of length at least $r_1 + \cdots + r_l$. In particular, if $E_1 + \cdots + E_l = I$, then $\mathcal{S}$ is diagonalizable.

Proof. We may assume without loss of generality that $\mathcal{S} = \overline{\mathbb{C}\mathcal{S}}$.

First we'll show that if $P_j E_i \neq 0$ then $P_j E_k = 0$ for all $k \neq i$. Assume otherwise and fix such i, j, and k. By definition, $P_j \mathcal{S} P_j|_{\mathrm{ran}(P_j)}$ is irreducible. Therefore $E_i P_j \mathcal{S} P_j E_k$ is nonzero since E_i and E_k are nonzero on $\mathrm{ran}(P_j)$ by our choice of i, j, and k. Since the P_j's commute with $\mathcal{S}$, $P_j E_i \mathcal{S} E_k P_j$ is nonzero, so $E_i \mathcal{S} E_k$ is nonzero. But each such element is in $\mathcal{S}$ and is nilpotent since $E_k E_i = 0$. By Lemma 5.2.4, $\mathcal{S}$ has no nonzero nilpotents, so this is a contradiction and no such i, j, and k exist.

Each E_i has rank r_i and each $P_j E_i$ has rank at most one. Since the ranges of the P_i span $\mathcal{V}$, there must be r_i such projections for E_i. By our first claim, the projections are unique for each E_i. Therefore we have $r_1 + \cdots + r_l$ distinct projections. This gives us the required chain of invariant subspaces. In particular, if the E_i's sum to the identity, their ranks sum to the dimension of $\mathcal{V}$. Thus each projection is rank one and the totally reducible semigroup is diagonalizable.

$\square$

We now drop the requirement of total reducibility for semigroups, but consider the special case of groups of invertible operators. In contrast to arbitrary semigroups, the $\sqrt{3}$ condition forces reducibility in such groups.

5.2.8 Theorem

Let $\mathcal{G}$ be a group of invertible operators in $\mathcal{B}(\mathcal{V})$ satisfying the $\sqrt{3}$ condition with $m = \dim(\mathcal{V})$. Then $\mathcal{G}$ is solvable and the following hold:

(i) If $m \leq 3$, then $\mathcal{G}$ is triangularizable and if $m \geq 4$, then the lattice of invariant subspaces of $\mathcal{G}$ contains a chain of length at least three.

(ii) The derived subgroup $\mathcal{G}'$ is triangularizable.

(iii) For each $A \in \mathcal{G}'$, we have $\sigma(A) \subset \{z \in \mathbb{C} : |z| = 1\}$.

(iv) If $\sigma(A) \subset \{z \in \mathbb{C} : |z| = \rho(A)\}$ for every A in $\mathcal{G}$, then $\mathcal{G}$ is triangularizable.

Proof. Multiplication by $\mathbb{C}^* = \mathbb{C}\backslash\{0\}$ does not affect the $\sqrt{3}$ condition so we can assume without loss of generality that $\mathcal{G} = \mathbb{C}^*\mathcal{G}$. We'll first deal with the case where $\mathcal{G}$ is totally reducible. Then the (totally reducible) semigroup $\mathcal{S} = \overline{\mathcal{G}}$ also satisfies the $\sqrt{3}$ condition and we let $P_1, \ldots, P_m$ be a complete set of mutually orthogonal idempotents onto minimal invariant subspaces of $\mathcal{G}$ (and $\mathcal{S}$).

By Theorem 5.2.6, $\mathcal{S}$ contains a minimal idempotent E. If $E = I$ then $\mathcal{S}$ (and $\mathcal{G}$) is diagonalizable by Corollary 5.2.7 and this result is trivial. Assume $E \neq I$ and consider a sequence $\{G_n\}$ in $\mathcal{G}$ converging to E. The sequence $\{G_n^{-1}/\, \|G_n^{-1}\|\}$ consists of elements of norm one so, by passing to a subsequence, we may assume that it converges to some $A \neq 0$ in $\mathcal{S}$.

Clearly, $AE = EA$ by construction. Since the norms of the G_n^{-1}'s are unbounded (otherwise E would be invertible and therefore equal to I), $AE = EA = 0$. Let $\mathcal{S}_1$ be the subsemigroup of $\mathcal{S}$ generated by A. As we saw in the proof of Theorem 5.2.6, $\overline{\mathbb{C}\mathcal{S}_1}$ must contain an idempotent, F. As $\mathcal{S}_1$ is generated by A, $EF = FE = 0$. Therefore there is an idempotent in $\mathcal{S}$ which is orthogonal to E so any maximal set of mutually orthogonal minimal idempotents in $\mathcal{S}$ must contain at least two elements.

Take $\{E_1, \ldots, E_k\}$ to be such a set. By Corollary 5.2.7, for each P_i, the rank of P_iE_j is either zero or one. If the rank is one, we claim that the rank of P_i is one as well. Assume otherwise.

We claim that E_j does not commute with everything of the form E_jB with B in $\mathcal{S}$. Assume otherwise. Then for every B in $\mathcal{S}$ and x in $\ker(E_j)$,

$$E_j(Bx) = E_j^2Bx = E_jBE_jx = 0,$$

so $\ker(E_j)$ is an invariant subspace of $\mathcal{S}$. Since P_iE_j is rank one and the rank of P_i is more than one, $\ker(E_j) \cap \operatorname{ran}(P_i)$ is a nontrivial invariant subspace for $P_i\mathcal{S}P_i|_{\operatorname{ran}(P_i)}$. But this is a contradiction as $P_i\mathcal{S}P_i|_{\operatorname{ran}(P_i)}$ is irreducible. Therefore the claim is proved and there is a B in $\mathcal{S}$ such that $C = E_jB$ does not commute with E_j. Thus $E_jC = C \neq CE_j$.

Take $\{G_n\}$ to be a sequence in $\mathcal{G}$ converging to E_j. Assume the sequence CG_n^{-1} was bounded. Then, by considering a subsequence, it converges to some T in $\mathcal{S}$. But then

$$C = CG_n^{-1}G_n \underset{n\longrightarrow\infty}{\longrightarrow} TE_j,$$

so $CE_j = C$. Therefore the sequence CG_n^{-1} cannot be bounded. Consider a subsequence so that $\|CG_n^{-1}\|$ converges to infinity. By taking yet another subsequence, we may also assume that the bounded sequence $\{CG_n^{-1}/\,\|CG_n^{-1}\|\}$ converges to some $S \neq 0$ in $\mathcal{S}$. Clearly, $E_jS = S$, while

$$SE_j = \lim_{n\longrightarrow\infty} \frac{CG_n^{-1}G_n}{\|CG_n^{-1}\|} = \lim_{n\longrightarrow\infty} \frac{C}{\|CG_n^{-1}\|} = 0.$$

Therefore $S^2 = SE_jS = 0$ so S is a nonzero nilpotent in S. But this is a contradiction by Lemma 5.2.4. Therefore P_i has rank one.

Let $\mathcal{M}$ be the span of the ranges of the E_j's. The fact we just proved shows that $\mathcal{S}$ is diagonal on $\mathcal{M}$ as each P_i whose range intersects $\mathcal{M}$ is rank one. We've also shown that there are at least two E_j's. If $m = 3$, the third projection must automatically be rank one as well, proving diagonalizability. The result for $m \geq 4$ is also clear. This proves (i) in the totally reducible case.

We now look at $\mathcal{G}'$. If $\mathcal{M} = \mathcal{V}$ then $\mathcal{G}$ is abelian and the entire result holds so assume $\mathcal{M} \neq \mathcal{V}$. We know $\mathcal{M}$ has dimension at least two as there are at least two E_j's. Since $\mathcal{G}$ is diagonal on $\mathcal{M}$, $\mathcal{G}'$ acts trivially on $\mathcal{M}$. Take $\mathcal{N} = (I - \oplus E_j)\mathcal{V}$ and consider the action of $\mathcal{G}'$ on $\mathcal{N}$.

We want to show that $\mathcal{G}'$ is simultaneously similar to a unitary group. Since $\mathcal{G}'$ acts trivially on $\mathcal{M}$, $\rho(A) \geq 1$ for all A in $\mathcal{G}'$. If $\rho(A) > 1$ for some A in $\mathcal{G}'$ then, using the technique from Lemma 5.2.2, we get an idempotent (since there are no nilpotents in $\mathcal{G}$ by Lemma 5.2.4) E in $\overline{\mathbb{C}\mathcal{G}'}$. Since $\rho(A|_{\mathcal{M}}) = 1$, $\mathcal{M}$ would be in $\ker(E)$. Also, $\operatorname{ran}(E)$ would be contained in $\mathcal{N}$ by the block diagonal nature of $\mathcal{G}'$ implied by total reducibility. This contradicts the maximality of the set $\{E_1, \ldots, E_k\}$. Therefore, $\rho(A) = 1$ for every A in $\mathcal{G}'$.

We claim that $\mathcal{G}'$ is totally reducible. Let $\mathcal{M}$ be a minimal invariant subspace for $\mathcal{G}'$ (no nontrivial subspace of $\mathcal{M}$ is invariant for $\mathcal{G}'$). We need to show that $\mathcal{M}$ has a complementary space $\mathcal{N}$ which is also invariant for $\mathcal{G}'$. Recall that $\mathcal{G}'$ is a normal subspace of $\mathcal{G}$. Now, for any G in $\mathcal{G}$,

$$\mathcal{G}'(G\mathcal{M}) = G(\mathcal{G}'\mathcal{M}) \subseteq G\mathcal{M},$$

so $G\mathcal{M}$ is invariant for $\mathcal{G}'$. Then $\mathcal{M} \cap G\mathcal{M}$ is invariant for $\mathcal{G}'$ and contained within $\mathcal{M}$. Thus $\mathcal{M}$ is minimal, the intersection must be either $\mathcal{M}$ or $\{0\}$.

If $\mathcal{M} \cap G\mathcal{M} = \mathcal{M}$ then, since $G\mathcal{M}$ is a minimal invariant subspace (as otherwise $G^{-1}\mathcal{M}$ would be a nontrivial invariant subspace properly contained within $\mathcal{M}$), we must have that $\mathcal{M} = G\mathcal{M}$. If $G\mathcal{M} = \mathcal{M}$ for all G in $\mathcal{G}$ then $\mathcal{M}$ is invariant for G and must be minimally, so as $\mathcal{G}' \subseteq \mathcal{G}$. Therefore, as $\mathcal{G}$ is totally reducible, $\mathcal{M}$ has a complementary subspace $\mathcal{N}$ that is invariant for $\mathcal{G}$ and thus $\mathcal{G}'$.

In the other case, there is some G such that $\mathcal{M} \cap G\mathcal{M} = \{0\}$ and $G\mathcal{M}$ is invariant for $\mathcal{G}'$. Let $\{G_1, \ldots, G_l\}$ be a minimal set with $\mathcal{M} \cap G_i\mathcal{M} = \{0\}$ such that for any H in $\mathcal{G}$, either $H\mathcal{M} = \mathcal{M}$ or $H\mathcal{M} \subseteq \operatorname{span}\{G_1\mathcal{M}, \ldots, G_l\mathcal{M}\}$. Let $\mathcal{L} = \operatorname{span}\{G_1\mathcal{M}, \ldots, G_l\mathcal{M}\}$. Obviously $\mathcal{N} \cap \mathcal{M} = \{0\}$. By definition, $\mathcal{M} + \mathcal{L}$ is an invariant subspace of $\mathcal{G}$. Since $\mathcal{G}$ is totally reducible, $\mathcal{M} + \mathcal{L}$ has a complementary invariant subspace $\mathcal{L}'$. Taking $\mathcal{N} = \mathcal{L} + \mathcal{L}'$ gives us the required complementary invariant subspace of $\mathcal{M}$ for $\mathcal{G}'$.

Now, $\mathcal{G}'$ is totally reducible and every element in $\mathcal{G}'$ has spectral radius one. $\mathcal{G}'$ is therefore the direct sum of irreducible groups, each of which has bounded spectral radius. By Lemma 2.2.15, each of these irreducible groups is bounded and therefore, $\mathcal{G}'$ is bounded so by Theorem 2.3.1 it is simultaneously similar to a unitary group. This proves (iii) for the totally reducible case. By Theorem 5.1.1, $\mathcal{G}'$ is abelian and triangularizable. This proves (ii) for the totally reducible case. If $\sigma(A) \subset \{z \in \mathbb{C} : |z| = \rho(A)\}$ then $\mathcal{G}_1 = \{A \in \mathcal{G} : \rho(A) = 1\}$ will be bounded and then abelian by Theorem 5.1.1. This proves (iv) for the totally reducible case. Finally, since $\mathcal{G}'$ is abelian, $\mathcal{G}$ is solvable.

We now consider the general case. Let $\mathcal{C}$ be a maximal chain of invariant subspaces of $\mathcal{G}$. Let $P_1, P_1 \oplus P_2, \ldots, P_1 \oplus \cdots \oplus P_l$ be the corresponding projections onto invariant subspaces. Consider the group

$$\mathcal{G}_s = \{P_1AP_1 \oplus P_2AP_2 \oplus \cdots \oplus P_lAP_l : A \in \mathcal{G}\}.$$

By the maximality of $\mathcal{C}$, $\mathcal{G}_s$ is totally reducible. Also, the map from $\mathcal{G}$ to $\mathcal{G}_s$ (block upper-triangular to block diagonal) preserves spectral radius. Therefore $\mathcal{G}_s$ has the $\sqrt{3}$ condition and the theorem holds for $\mathcal{G}_s$.

Results (i)-(iv) for $\mathcal{G}$ follow directly from the results for $\mathcal{G}_s$. Thus, we need only show that $\mathcal{G}$ is solvable. Take ϕ to be the map from $\mathcal{G}$ to $\mathcal{G}_s$. Then ϕ is a homomorphism and $\phi(G) = \mathcal{G}_s$ is solvable. Therefore, in order to show $\mathcal{G}$ is solvable, we need only show that $\ker(\phi)$ is solvable.

Now, $\ker(\phi)$ is a subset $\mathcal{F}$, the group of the upper triangular matrices with ones on the diagonal. The commutator subgroup $\mathcal{F}' = [\mathcal{F}, \mathcal{F}]$ is contained within those matrices that are block upper triangular with all blocks being 2×2 and whose diagonal blocks are the identity. Further, $[\mathcal{F}', \mathcal{F}']$ is contained within a similar set except that the blocks are 4×4. Repeating this $\lceil \log(\dim(V)) \rceil$ times, we are eventually left with $\{I\}$ so $\mathcal{F}$ is solvable. Since subgroups of solvable groups are solvable, $\ker(\phi)$ is solvable. Therefore $\mathcal{G}$ is solvable.

$\square$

For connected groups, the result is stronger.

5.2.9 Corollary

A connected group satisfying the $\sqrt{3}$ condition is triangularizable.

Proof. $\mathcal{G}$ is solvable so the claim follows from the Lie-Kolchin Theorem ([4, Corollary 1.5]) which says that a connected and solvable linear algebraic group is triangularizable.

$\square$

Without connectedness, we cannot hope for better as seen by the following example.

5.2.10 Example

Consider the group $\mathcal{G}$ in $\mathcal{M}_4(\mathbb{C})$ generated by the two elements

$$U = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

is a group that satisfies the $\sqrt{3}$ condition, but is not triangularizable.

Proof. Since $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ share no invariant subspaces, $\mathcal{G}$ is not triangularizable and its longest chain of invariant subspaces is of length three.

Note that $\mathcal{G}$ is commutative on the first two basis elements. On the third and fourth basis elements, it acts as one of $\pm E_{11} \pm E_{22}$ or $\pm E_{12} \pm E_{21}$, where E_{ij} are the standard matrix units on $\mathcal{M}_2(\mathbb{C})$. Consider the commutators of elements A and B of these two forms. If A and B are of the same type, (either both diagonal or both zero on the diagonal), then AB and BA are both diagonal matrices with entries of modulus 1. If A and B are different types, the AB and BA are both zero on the diagonal with entries of modulus 1 off it. In either case $AB - BA$ will have entries of modulus at most 2 either all on or all off the diagonal so $\rho(AB - BA) \leq 2$. Since $\mathcal{G}$ is commutative on the first two elements, $\rho(AB - BA) \leq 2$ for every element of $\mathcal{G}$.

Consider elements A and B in $\mathcal{G}$ and write each as a word in U and V. If the exponents of the U's in A add up to zero then A acts as one of $I, -I, E_{12} + E_{21}$ or $-(E_{12} + E_{21})$ on the third and fourth basis element. If the exponents of B similarly add up to zero then A and B will commute, so they satisfy the $\sqrt{3}$ condition.

Otherwise, either A or B has spectral radius at least 2 (due to the first or second basis element). The other will have spectral radius at least 1 since every element in $\mathcal{G}$ has that property. Therefore the $\sqrt{3}$ condition will be satisfied in this case as well.

$\square$

We now return to arbitrary semigroups for a few final results. We need the following definition.

5.2.11 Definition

Given a maximal chain $\mathcal{C}$ of invariant subspaces for a semigroup $\mathcal{S}$ of operators in $\mathcal{B}(\mathcal{V})$ with corresponding projections $P_1, P_1 \oplus P_2, \ldots, P_1 \oplus \cdots \oplus P_l$ we write the "block diagonal form" of $\mathcal{S}$ as

$$\mathcal{S}_s = \{P_1 S P_1 \oplus P_2 S P_2 \oplus \cdots \oplus P_l S P_l : S \in \mathcal{S}\}.$$

Then $\mathcal{S}_s$ is totally reducible.

5.2.12 Lemma

Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{B}(\mathcal{V})$ and assume S contains nonnilpotent elements. Let r be the minimal nonzero rank of elements of $\overline{\mathbb{C}\mathcal{S}}$ and let r_s be the minimal nonzero rank of element of $\overline{\mathbb{C}\mathcal{S}_s}$. Then $r_s \geq r$.

Proof. We first show that $\overline{\mathbb{C}\mathcal{S}_s}$ contains an idempotent E of rank r_s. Take $\mathcal{J}$ to be the ideal of elements of rank at most r_s in $\overline{\mathbb{C}\mathcal{S}_s}$. If $\mathcal{J}$ consists entirely of nilpotents, then it is triangularizable by Levitzki's Theorem (2.2.11). By total reducibility, $P_i\overline{\mathbb{C}\mathcal{S}_s}P_i$ is irreducible, so the ideal $P_i\mathcal{J}P_i$ should be irreducible by Lemma 2.2.12. Thus each P_i would have to operate on a one dimensional space but this means that $\mathcal{J}$, which consists entirely of nilpotents, is diagonal and thus the zero ideal. This contradicts that there are nonzero elements of rank r_s. Therefore, $\mathcal{J}$ contains a nonnilpotent operator. By Lemma 5.2.2, $\mathcal{J}$ contains either an idempotent E or a nilpotent of rank less than r_s. But r_s is the minimal rank, so such an idempotent E must exist.

Since E is a limit of elements in $\mathbb{C}\mathcal{S}_s$ and the spectrum is continuous (2.4.5), there must be an element A in $\mathcal{S}_s$ whose spectrum contains at most r_s elements (counting multiplicity) with maximal modulus. Let B be an element in $\mathcal{S}$ that maps to A under the obvious map from $\mathcal{S}$ to $\mathcal{S}_s$. Then B has exactly the same spectrum as A. By Lemma 5.2.2, $\overline{\mathbb{C}\mathcal{S}}$ contains an element of rank at most r_s so $r_s \geq r$.

□

We can now extend Theorem 5.1.1 from compact groups to semigroups.

5.2.13 Theorem

Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{B}(\mathcal{V})$ satisfying the $\sqrt{3}$ condition. Let m be the minimal nonzero rank in $\overline{\mathbb{C}\mathcal{S}}$. Then $\mathcal{S}$ has a chain of invariant subspaces of length at least m.

Proof. If $\mathcal{S}$ consists of nilpotents then it's triangularizable by Levitzki's Theorem (2.2.11). Otherwise, $\overline{\mathbb{C}\mathcal{S}_s}$ contains a minimal idempotent E of rank $r_s \geq m$. As $\mathcal{S}$ satisfies the $\sqrt{3}$ condition, so does $\mathcal{S}_s$. By Corollary 5.2.7, $\mathcal{S}_s$ has a chain of invariant subspaces of length at least m. This same chain is also a chain of invariant subspaces for $\mathcal{S}$, so the result is proved.

□

5.3 Infinite Dimensions

In infinite dimensions, there are few affirmative results, especially without any compactness assumptions. It is still an open problem whether a bounded operator T on a Hilbert space has invariant subspaces so we don't know if the semigroup generated by T is reducible. However, this semigroup is abelian and thus satisfies the $\sqrt{3}$ condition.

We will therefore restrict ourselves to semigroups of compact operators. We can extend Example 5.2.5 to infinite dimensions by replacing $\mathbb{C}^{n-1}$ with ℓ^2. Our one affirmative result in infinite dimensions is a partial analogue to Corollary 5.2.7.

5.3.1 Theorem

Let $\mathcal{S}$ be a semigroup of operators in $\mathcal{K}(\mathcal{X})$ which satisfies the $\sqrt{3}$ condition and let m be the minimal nonzero rank in $\overline{\mathbb{C}\mathcal{S}}$, which may be infinite. Then $\mathcal{S}$ has a chain of invariant subspaces of length at least m.

Proof. We can assume without loss of generality that $\mathcal{S} = \overline{\mathbb{C}\mathcal{S}}$. If m is infinite then $\mathcal{S}$ contains no finite rank operators. Therefore $\mathcal{S}$ consists entirely of quasinilpotents by Lemma 4.2.1. By Turovskii's Theorem (4.2.2), $\mathcal{S}$ is triangularizable. Therefore we can assume $m < \infty$.

Let $\mathcal{C}$ be a maximal chain of invariant subspaces of $\mathcal{S}$. We want to show that $\mathcal{C}$ has length at least m or, in other words, that it has at least $m+1$ elements including $\{0\}$ and $\mathcal{X}$. Assume it has $k+1 < m+1$ elements, then

$$\{0\} \subset \mathcal{M}_1 \subset \cdots \subset \mathcal{M}_k = \mathcal{X}.$$

For each $i \geq 1$, denote the quotient space $\mathcal{M}_i/\mathcal{M}_{i-1}$ by $\mathcal{X}_i$. We can then create a Banach space $\mathcal{Y} = \oplus_i \mathcal{X}_i$ where the norm of an element $y = (x_1, \ldots, x_k)$ in $\mathcal{Y}$ is $\|y\| = \max\{\|x_i\| : 1 \leq i \leq k\}$.

Consider the new semigroup $\mathcal{T}$ defined on $\mathcal{Y}$ where the elements of $\mathcal{T}$ are $\oplus_i A_i$ where A_i is the quotient operator on $\mathcal{X}_i$ induced by A for every A in $\mathcal{S}$. The homomorphism from $\mathcal{S}$ to $\mathcal{T}$ taking A to $\oplus_i A_i$ is a contraction and preserves spectrum, including multiplicity (Theorem 4.1.10 and Theorem 4.1.11).

Now, the proof of Lemma 5.2.12 works in infinite dimensions except that we consider quasinilpotents instead of nilpotents and use Turovskii's Theorem (4.2.2) instead of Levitzki's Theorem (2.2.11). Therefore there is a minimal idempotent E in $\overline{\mathbb{C}\mathcal{T}}$ of minimal rank $l \geq m$. The semigroup $E\overline{\mathbb{C}\mathcal{T}}E|_{E\mathcal{X}}$ is similar to scalar multiples of a unitary group by Lemma 2.3.3. Hence it is abelian by Theorem 5.1.1, is therefore diagonalizable, and thus has a chain of invariant subspaces of length l. By Lemma 4.2.6, $\mathcal{T}$ has a chain of invariant subspaces of length at least l which contradicts that such a chain for $\mathcal{T}$ has length at most k.

$\square$

Bibliography

[1] Janez Bernik and Heydar Radjavi, *How Small Can Nonzero Commutators Be?*, Indiana Univ. Math. J. **54** (2005), 309–320.

[2] Janez Bernik, R. Drnovsek, T. Košir, M. Mastnak, M. Omladič, and Heydar Radjavi, *Approximate Permutability of Trace on Operator Semigroups*, preprint.

[3] Janez Bernik, Robert M. Guralnick, and Mitja Mastnak, *Reduction theorems for groups of matrices*, Linear Algebra Appl. **383** (2004), 119–126.

[4] Armand Borel, *Linear Algebraic Groups*, Graduate Texts in Mathematics, vol. 126, Springer-Verlag, New York, 1991.

[5] Kenneth R. Davidson, *Nest Algebras*, Pitman Research Notes in Mathematics Series, vol. 191, Longman Scientific and Technical, Essex, 1988.

[6] V.B. Lidskiĭ, *Non-selfadjoint operators with a trace*, Dokl. Akad. Nauk SSSR **125** (1959), 485–487 (Russian) Transl. Amer. Math. Soc. (2) **47** (1965), 43–46, (English).

[7] Heydar Radjavi and Peter Rosenthal, *Simultaneous Triangularization*, Universitext, Springer-Verlag, New York, 2000.

[8] Walter Rudin, *Functional Analysis*, Second Edition. McGraw-Hill, New York, 1991.

[9] Walter Rudin, *Real and Complex Analysis*, Third Edition. McGraw-Hill, New York, 1987.