PROGRESS ON THE INVARIANT SUBSPACE PROBLEM

BEN GOLDMAN

Contents

Date: August, 2020.

1. Introduction

An important tool in working on any abstract space is to search for invariants, notions which do not change after a shift. The most elementary of these is the fixed point, an element which remains stationary after a transformation of space. The study of linear operators then introduces an invariant equipped with nonzero dimensionality, the eigenvector. In this paper, we endeavor to study how these simple principles can be extended to a more abstract notion, the invariant subspace. We begin with some very elementary, motivating examples.

1.1. **Finite Dimensions.** Let V, W be complex vector spaces.

Definition 1.1.1. Recall that an operator, $T : V \to W$, is linear if for all $v, \mu \in V$ and $\alpha, \beta \in \mathbb{R}$ we have that

$$T(\alpha v + \beta\mu) = \alpha T(v) + \beta T(\mu)$$

Definition 1.1.2. Recall that, v, is an eigenvector of a linear operator, T, with eigenvalue, $\lambda \in \mathbb{C}$, if $T(v) = \lambda v$.

Example 1. Consider the linear transformation, $T : \mathbb{R}^2 \to \mathbb{R}^2$, given by $T(x, y) = (2x, 3y)$. This has eigenvectors, $(1, 0)$ and $(0, 1)$ with real eigenvalues, 2 and 3, respectively.

Example 2. Now, consider the linear transformation, $T : \mathbb{R}^2 \to \mathbb{R}^2$, which rotates the plane by some angle, θ. It is fairly easy to observe that T has no real eigenvalues. However, one observes that the unit disk, $D = \{v \in \mathbb{R}^2 \mid \|v\| = 1\}$, is invariant in T (i.e. $T(D) = D$). This motivates our definition, below:

Definition 1.1.3. A subspace, $W \subseteq V$, is invariant in a linear operator, $T : V \to V$, if $T(W) \subseteq W$.

Remark. Granted, not all invariant subspaces are created equal. For instance, the origin in V is invariant under any linear transformation, but this does not tell us anything significant about the operator nor the space. Similarly, for any surjective linear operator, $T : V \to V$, we have that $T(V) = V$, which is another unimportant example. We dub these the *trivial* invariant subspaces for an operator.

Definition 1.1.4. The orbit of a vector, $v \in V$, with respect to a linear operator, $T : V \to V$, is given by

$$O_v = \{v, Tv, T^2v, T^3v, T^4v, \dots, T^nv, \dots\}$$

One notes that O_v is also an invariant subspace, but a fairly disappointing one.

Remark. In order to avoid the aforementioned simple cases, we state the Invariant Subspace Problem: Under which conditions does T have a nontrivial, closed invariant subspace?

It is fairly easy to prove this for the case of a finite dimensional complex vector space.

Theorem 1.1.5. *Any nonzero operator on a finite dimensional, complex vector space, V, admits an eigenvector.*

Proof. [A16] Let $n = \dim(V)$ and suppose $T : V \to V$ is a nonzero linear operator. Select nonzero $v \in V$ so that by finite dimensionality, we know that $O_v^n = \{v, Tv, T^2v, \ldots, T^n v\}$ must be a linearly dependent set. Thus,

$$\alpha_n T^n v + \alpha_{n-1} T^{n-1} v + \ldots \alpha_1 T v + \alpha_0 v = \eta (T - \lambda_1)(T - \lambda_2) \ldots (T - \lambda_n) v = 0$$

where we use $\lambda_i \in \mathbb{C}$ to denote $\lambda_i I$ (the identity map). At least one of these factors must have a nontrivial kernel which suffices to prove the theorem. One can adapt the proof above to real valued vector spaces, but the irreducible polynomials end up being quadratic instead. □

Corollary 1.1.6. *Any nonzero operator on a finite dimensional, complex vector space, V, admits a closed, nontrivial invariant subspace.*

2. Fundamentals of Functional Analysis

2.1. Banach Spaces.

Definition 2.1.1. Recall that a Banach Space is a complete Vector Space with norm, $\|.\|$.

Lemma 2.1.2 (Riesz). *Let $\mathcal{B} = \{x \in E, \|x\| = 1\}$ be the unit sphere in an infinite dimensional Banach Space, $(E, \|.\|)$. If $Y \subset E$ is a closed proper subspace of E then there exists $x^* = x^*(\varepsilon) \in \mathcal{B}$ such that*

$$\|y - x^*\| > \frac{1}{1+\varepsilon} \quad | \quad \forall y \in Y, \varepsilon > 0$$

Proof. Pick an arbitrary $x \in \mathcal{B}$ so that $d = \inf_{y \in Y} \big(\|x - y\| \big)$. Select y^* so that $\|x - y^*\| = d + d\varepsilon_0 < d + d\varepsilon$ and set

$$x^* \in \mathcal{B},\ x^* = \frac{1}{\|x - y^*\|}(x - y^*) = \frac{1}{d + d\varepsilon_0}(x - y^*)$$

We then have that

$$\|x^* - y\| = \frac{1}{d + d\varepsilon_0} \|x - y^* - (d + d\varepsilon_0) y\| > \frac{1}{d + d\varepsilon_0} \inf_{\hat{y} \in Y} (x - \hat{y}) > \frac{1}{1+\varepsilon}$$

where $\hat{y} = y^* + (d + d\varepsilon_0) y$ □

Theorem 2.1.3 (Riesz). *If $(E, \|.\|)$ is an infinite dimensional banach space then the unit sphere, $\mathcal{B}$, is not compact.*

Proof. Pick $x_1 \in \mathcal{B}$ and set $Y_1 = \overline{\text{span}}\{x_1\}$. Apply the previous lemma to yield x_2 such that $\inf_{y \in Y_1} \big(\|x_2 - y\| \big) > \frac{1}{2}$ and let $Y_2 = \overline{\text{span}}\{x_1, x_2\}$. Continue this process inductively to yield a sequence, $\{x_n\}$ and $\{Y_n\}$, such that x_n has no convergent subsequence. □

This stands in stark contrast to finite dimensional vector spaces which have compact unit spheres. We conclude this subsection with a fact about linear operators on Banach Spaces.

Definition 2.1.4. We define a norm on the space of linear operators between Banach spaces, $T : (X, \|.\|_X) \to (Y, \|.\|_Y)$, by

$$\|T\| = \sup_{x \in \mathcal{B}} \|T(x)\|_Y$$

One notes that there is not necessarily a vector, $x_0 \in \mathcal{B}$, for which $\|T(x_0)\| = \|T\|$ due to the loss of compactness.

Theorem 2.1.5. *If $T : (X, \|.\|_X) \to (Y, \|.\|_Y)$ is a linear map between Banach Spaces then it is continuous if and only if $\|T\| < +\infty$*

Proof. If $\|T\| < +\infty$ then for all $\varepsilon > 0$ if $\|x - y\|_X < \frac{\varepsilon}{\|T\|}$ then

$$\|T(x) - T(y)\|_Y = \|x - y\|_X \left\| T\left(\frac{1}{\|x - y\|_X}(x - y) \right) \right\|_Y \le \|T\| \, \|x - y\|_X < \varepsilon$$

We will use the manipulation above i.e. $\|T(x)\|_Y \le \|T\| \, \|x\|_X$ quite often. If $\|T\|$ is unbounded then for each $n \in \mathbb{N}$ select $x_n \in X$ such that $\|x_n\|_X = \frac{1}{n}$ and $\|T(x_n)\|_Y > n$. This produces a sequence, $x_n \to 0$ such that $\|T(x_n)\|_Y \ge 1$ for all n. And thus, $\limsup_{x \to 0} \|T(x)\|_Y \ge 1$, and so T is not continuous at the origin. □

2.2. The Open Mapping Theorem.

Theorem 2.2.1 (Baire Category Theorem). *Complete metric spaces are Baire spaces. More precisely, if (X, ρ) is a complete metric space and $\{\mathcal{O}_n\}_{n \in \mathbb{N}}$ is a countable collection of open dense subsets then the intersection, $\mathcal{I} = \bigcap_{n \in N} \mathcal{O}_n$, is also dense.* [R10]

We require the above theorem from non-functional analysis in order to justify the following results.

Theorem 2.2.2 (Open Mapping Theorem). *If $T : (X, \|.\|_X) \to (Y, \|.\|_Y)$ is a surjective, continuous linear map then T is an open map.*

Proof. [R10] Let $B_X = B_X(0,1)$ be the unit ball in X and B_Y be the unit ball in Y. Observe that $X = \bigcup_{r\in\mathbb{N}}(rB_X)$ and that $Y = \overline{\bigcup_{r\in\mathbb{N}} T(rB_X)}$. By the Baire Category Theorem, there exists $R \in \mathbb{N}$ such that $\overline{T(RB_X)} = \overline{T(B_X(0,R))}$ has nonempty interior.

Pick $B_Y(y_0, r_0) \subseteq \overline{T(B_X(0,R))}$ so that for all $y \in B_Y$ we have that $y_0 + r_0 y \in B_Y(y_0, r_0)$. Thus,

$$r_0 y \in \left(\overline{T(RB_X)} - \overline{T(RB_X)}\right) \subseteq \overline{T(2RB_X)}$$

and we thus have that $B_Y \subseteq \overline{T(R^* B_X)}$ where $R^* = 2R/r_0$.

For any $y \in Y$ and $\varepsilon > 0$, we find a point $\overline{x} \in (R^* B_X)$ such that $\|T(\overline{x}) - y\| < \varepsilon$. We can refine this approach to get a stronger result. Select $x_1 \in B\left(0, \frac{R^*}{2}\right)$ for which $\|Tx_1 - y\| < \frac{1}{2}$ and $x_n \in B\left(0, \frac{R^*}{2^n}\right)$ such that $\|T(x_1 + x_2 + \cdots + x_n) - y\| < \frac{1}{2^n}$. Letting $x^* = \sum_{n\in\mathbb{N}} x_n$ where $\|x^*\| \leq \sum_{n\in\mathbb{N}} \|x_n\| \leq 2R^*$ gives that $T(x^*) = y$. Thus, we have shown that $B_Y \subseteq T(2R^* B_X)$ and so $B_Y\left(0, \frac{1}{2R^*}\right) \subseteq T(B_X)$.

Finally, one shows that T is open. Since T is linear, it suffices to prove that the image of the unit ball is open (other open sets are achieved by translations/scalings). Take $y \in T(B_X)$ and select $x \in B_X$ such that $Tx = y$. Pick $\varepsilon > 0$ such that $B_X(x, \varepsilon) \subseteq B_X$ and pick r sufficiently small so that $B_Y(y, r) \subseteq T(B_X(x, \varepsilon))$ so that $B(y, r) \subseteq T(B_X)$. □

We finish this subsection with a handful of useful statements that are immediate consequences of the Open Mapping Theorem.

Corollary 2.2.3. *If $T : (X, \|.\|_X) \to (Y, \|.\|_Y)$ is a bijective, continuous operator between two banach spaces then T^{-1} is continuous.* [R10]

Corollary 2.2.4. *If $\|.\|_1$ and $\|.\|_2$ are two norms on a Banach space, X, such that $\|.\|_1 \leq C \|.\|_2$ then they are equivalent.* [R10]

2.3. **Compact Operators.**

Notation 2.3.1. *Let X be a norm space. We call $\mathcal{L}(X)$ the space of continuous linear operators mapping a space, X, into itself. If $T \in \mathcal{L}(X)$ then $T : X \to X$.*

For obvious reasons, we will be concentrating our efforts on these types of operators for the remainder of the paper. The loss of compactness in Banach spaces makes it difficult to study these linear operators to the same degree as one could in a Euclidean setting. However, we define a special class of operators which has very nice properties (and that we will see, later on, satisfies the invariant subspace problem).

Definition 2.3.2. An operator, $T \in \mathcal{L}(X)$, is compact if it maps the unit ball into a set with compact closure. In other words, $T(B_X)$ is a *relatively* compact set.

Remark. If $\{v_n\}$ is a bounded sequence in X then there exists a subsequence, v_{n_k}, such that $T(v_{n_k}) \to v$ for some $v \in X$.

Definition 2.3.3. Let $(X, \|.\|)$ be a norm space. A sequence, $\{v_n\} \subset X$, is said to converge weakly, $v_n \rightharpoonup v$, if $\varphi(v_n) \to \varphi(v)$ for every real-valued continuous linear map, $\varphi : X \to \mathbb{R}$. We alternately write that $\varphi \in X^*$, the dual of X.

Proposition 2.3.4. *If $v_n \to v$ then $v_n \rightharpoonup v$.*

Theorem 2.3.5. *Let X be a Banach space on which $x_n \rightharpoonup x$ is a weakly convergent sequence. It follows that $T(x_n) \to T(x)$ is strongly convergent if T is compact.*

Proof. If T is compact then for any subsequence x_{n_k}, there exists a further subsequence, $x_{n_{k_j}}$, such that $T(x_{n_{k_j}}) \to T(x)$. □

Remark. In general, solutions to the invariant subspace problem will require a broad construction of a suitable space, tailored to each type of operator. We will provide two different approaches to prove that compact operators have invariant subspaces which rely on this approach. We conclude this with a robust statement pertaining to compact operators due to Fredholm.

Theorem 2.3.6 (Fredholm)**.** *Let $(X, \|.\|)$ be a complex Banach space on which $T \in \mathcal{L}(X)$ is a compact operator. If $\lambda \in \mathbb{C}$ is not an eigenvalue then $(T - \lambda)$ has a bounded inverse.*

Our proof hinges on the lemma below.

Lemma 2.3.7. *If λ is not an eigenvalue then there exists a positive constant, $C > 0$, such that $\|(T - \lambda)(x)\| \geq C \|x\|$ for all $x \in X$.*

Proof. [T11] We will prove this via contradiction by showing that λ will otherwise be an eigenvalue. By assumption, there exists some sequence on the unit sphere, $\{x_n\}_{n \in \mathbb{N}} \subset \mathcal{B}$, such that $\|(T - \lambda)(x_n)\| \leq \frac{1}{n}$ for all n and thus, $\|(T - \lambda)(x_n)\| \to 0$. Moreover, we have that $\|T(x_n)\| \to |\lambda| > 0$ by assumption. Since T is compact, we can pass to a subsequence, x_{n_k}, such that $T(x_{n_k}) \to \overline{x}$ where $\overline{x}$ is nonzero and

$$(T - \lambda) \circ T(\overline{x}) = T \circ (T - \lambda)(\overline{x}) = 0 \implies T(\overline{x}) \in \ker(T - \lambda)$$

□

Proof. [T11] By Corollary 2.2.3, it suffices to prove that $(T - \lambda)$ is surjective. In other words, we need to show that $\operatorname{Im}(T - \lambda) = X$. We approach this via contradiction, supposing that there exists $\mu \in X$ such that $\mu \notin \mathrm{Im}(T-\lambda)$. Let $V_n = \mathrm{Im}\left([T-\lambda]^n\right) \subset X$ be a closed proper subspace.

It follows that $V_{n+1} \subset V_n$ is a closed proper subspace (in essence, that these sets are contracting). One proves this by induction, letting $S = (T-\lambda)$ be injective by assumption. Pick $x \notin \mathrm{Im}(S)$ so that if we assume $\mathrm{Im}(S) = \mathrm{Im}(S^2)$ then $Sx = S^2y$ for some y and thus $x = Sy$ gives a contradiction. The same holds for S^n, S^{n+1} for any $n \in N$.

We now apply the Riesz Lemma (2.1.2) to form a sequence,

$$\{x_n\}_{n\in N} \subseteq \mathcal{B}, \quad \inf_{v \in V_{n+1}} \|x_n - v\| > \frac{1}{2}$$

and so $\|Tx_n - Tx_m\| \geq \frac{|\lambda|}{2}$ for any $n, m \in \mathbb{N}$ which gives a sequence with no convergent subsequence. This contradicts the compactness of T. □

3. Spectral Theory and Linear Methods

This section will delve into some elementary functional calculus in order to better understand compact operators and derive some tenets of Spectral Theory. My goal is to avoid any strong complex analysis. From here, we will provide a proof of Gelfand's Spectral Radius Formula and solve the Invariant Subspace Problem for Compact Operators.

3.1. **Elementary Functional Calculus.**

Definition 3.1.1. Let $(X, \|.\|)$ be a Banach space on which $T : X \to X$ is a linear operator. The spectrum of the operator, $\sigma(T) \subseteq \mathbb{C}$, is defined as the set of complex values, $\lambda \in \sigma(T)$, for which $(T-\lambda)$ does not have a bounded inverse.

Definition 3.1.2. The *resolvent* function, $R(T,\lambda)$, is defined as this bounded inverse, $R(T,\lambda) = (T-\lambda)^{-1}$, when it exists. The domain of the resolvent function, the resolvent set $\rho(T)$, is the complement of $\sigma(T)$.

Proposition 3.1.3. *If $T \in \mathcal{L}(X)$ is compact and $\lambda \in \sigma(T)$ then λ is an eigenvalue.*

Theorem 3.1.4. *The spectrum, $\sigma(T)$, is non-empty, compact, and contained in the complex disk, $\mathcal{D} = \{z \in \mathbb{C} : |z| \leq \|T\|\}$.*

Proof. [B88] Define the map

$$g(\lambda) : X \to X, \quad g(\lambda) = \sum_{k=0}^{\infty} \frac{T^k}{\lambda^{k+1}}$$

where for convention we say $T^0 = I$. If $|\lambda| > \|T\|$ then

$$\left\| \sum_{k=0}^{\infty} \frac{T^k}{\lambda^{k+1}} \right\| \leq \sum_{k=0}^{\infty} \frac{\|T^k\|}{|\lambda|^k} < +\infty$$

and $g(\lambda)$ converges in norm. We can now extrapolate further,

$$(T-\lambda)g(\lambda) = (T-\lambda)\left(\frac{1}{\lambda}I + \frac{1}{\lambda^2}T + \frac{1}{\lambda^3}T^2 + \dots\right) = -I$$

which gives us that $g(\lambda) = -R(T,\lambda)$. We thus have that $g(\lambda) < +\infty$ if and only if $\lambda \in \rho(T)$ which gives us that $\sigma(T) \subseteq \mathcal{D}$.

The other claims follow from the maximum modulus principle. □

3.2. **Simple Invariant Subspace Case.**

Remark. Before seguing to the work of Gelfand, I have included a simple case of the invariant subspace problem that relies on these power series expansions. The motivation stems from the taylor expansion,

$$\sqrt{1-t} = 1 + \sum_{n=1}^{+\infty} \frac{(-1)^n t^n}{n!}\left(\prod_{k=1}^{n}\left(\frac{1}{2} - k + 1\right)\right) \tag{3.2.1}$$

Definition 3.2.2. A linear operator, $T : X \to X$, has a *hyperinvariant* subspace, $W \subset X$ if $S(W) \subseteq W$ for any operator, S, that commutes with T.

Theorem 3.2.3 (Beauzamy). *If $T : X \to X$ is a linear operator on X such that $\|T\|, \|T - I\| \geq 1/2$ and $\|T^2 - T\| \leq 1/4$ then T has a nontrivial, closed hyperinvariant subspace.*

Proof. [B88] Define the sum

$$S = I + \sum_{n=1}^{+\infty} \frac{(-4)^n}{n!}\left(\prod_{k=1}^{n}\left(\frac{1}{2} - k + 1\right)\right)\left[T - T^2\right]^n = I + K$$

It follows that S (and therefore K) is well defined due to (3.2.1), and that since for $|t| < 1$ we have $1 - \sqrt{1-t} < 1$ then we conclude $\|K\| < 1$. Further reference to the Taylor Expansion gives:

$$S^2 = I - 4(T - T^2) \implies (K + 2T) \circ (S + I - 2T) = 0 \tag{3.2.4}$$

Let $F = K + 2T$. We observe F is nonzero since

$$\|F\| = \|K + 2T\| \geq \|2T\| - \|K\| > 0$$

We turn to the subspace $U = \ker(F)$. To show that U contains a nonzero vector, we first claim that $S + I - 2T$ is nonzero since, if not, it would imply that $S - I = 2(T - I)$ which contradicts our assumption that $\|T - I\| \geq 1/2$. Since F commutes with T, U is a hyperinvariant subspace. □

Remark. This stronger notion of a hyperinvariant subspace will carry throughout the paper. As it turns out, these objects tend to arise more organically than the invariant subspaces, and we will often solve our simpler problem as a corollary of this broader notion.

3.3. **Gelfand's Spectral Radius Formula.**

Definition 3.3.1. Let $(X, \|.\|)$ be a Banach space. The spectral radius of an operator, $T : X \to X$, $r(T)$, is given by

$$r(T) = \max_{\lambda \in \sigma(T)} |\lambda|$$

Remark. We can further extrapolate about $r(T)$ using the function $g(\lambda)$ from before. The root test (from calculus) gives a further criteria for when $g(\lambda)$ is well defined, $|\lambda| > \limsup_{n\to\infty} \|T^n\|^{1/n}$. It remains to be shown that this limit is well defined.

Proposition 3.3.2. *The identity $\rho(T)^n = \rho(T^n)$ holds for $n \in \mathbb{N}$ where if $A \subseteq \mathbb{C}$ then $A^n = \{a^n, \ \forall a \in A\}$*

Proof. We may write

$$\begin{aligned}(T^n - \lambda^n) &= (T - \lambda) \circ (T^{n-1} - \lambda T^{n-2} + \dots \lambda^{n-1}) \\ &= (T - \gamma_1) \circ (T - \gamma_2) \circ \dots \circ (T - \gamma_{n-1}) \circ (T - \lambda)\end{aligned}$$

where $\gamma_1, \gamma_2, \dots, \gamma_{n-1} \in \mathbb{C}$. Observe that $\ker(T - \lambda) \subseteq \ker(T^n - \lambda^n)$. We have that $\lambda^n \in \rho(T^n)$ if and only if $T^n - \lambda^n$ has a bounded inverse, and therefore bijective. This occurs if and only if each of these factors is also bijective and bounded, and so Corollary 2.2.3 gives that each has a bounded inverse. Thus, if $\lambda^n \in \rho(T^n)$ then $\lambda \in \rho(T)$. Moreover, if $\lambda^n \notin \rho(T^n)$ then there exists some γ_i such that $\gamma_i \notin \rho(T)$. However, we deduce that $\gamma_i^n = \lambda^n$ and so $\lambda^n \notin \rho(T)^n$. □

Corollary 3.3.3. *It follows that $\sigma(T)^n = \sigma(T^n)$.*

Theorem 3.3.4 (Gelfand)**.** *The spectral radius of an operator, $r(T)$, is given by $r(T) \leq \limsup_{n\to\infty} \|T^n\|^{1/n} = \lim_{n\to\infty} \|T^n\|^{1/n}$*

Proof. [B11] It remains to be shown that $\lim_{n\to\infty} \|T^n\|^{1/n}$ exists. Define a sequence, $\{a_n\}_{n\in\mathbb{N}}$, by

$$a_n = \log\left(\|T^n\| \right)$$

Since T is bounded $\frac{a_n}{n}$ is non-increasing and bounded. Moreover, we use the inequalities below to deduce some important properties

$$\|T^{n+m}\| \le \|T^n\| \, \|T^m\| \implies a_{n+m} \le a_n + a_m$$

$$\|T^{nm}\| \le \|T^n\|^m \implies a_{nm} \le m a_n$$

By the Euclidean Algorithm, we deduce that for $m > n$ we have that

$$a_m \le q a_n + a_r \le \frac{m}{n} a_n + \alpha$$

where $q = \left\lfloor \frac{m}{n} \right\rfloor$, $0 \le r \le n$, and $\alpha = \max_{r\le n} \frac{a_r}{r}$. If $\eta = \inf_{n\in\mathbb{N}} \frac{a_n}{n}$ it suffices to show $\lim_{n\to\infty} a_n = \eta$. For each $\varepsilon > 0$ select $m > 0$ such that $\frac{a_m}{m} < \eta + \varepsilon$. For $n > m$, we have that $\frac{a_n}{n} \le \frac{a_m}{m} + \frac{\alpha}{m}$ by using the above decomposition. Select $N = \frac{\alpha}{\varepsilon}$ so that for $n > \max\{N, m\}$ we have that

$$\left| \frac{a_n}{n} - \eta \right| = \left| \frac{\alpha}{n} \right| < \varepsilon \implies \lim_{n\to\infty} \frac{a_n}{n} = \eta, \quad \lim_{n\to\infty} \|T^n\|^{1/n} = \lim_{n\to\infty} e^{\frac{a_n}{n}} = e^{\eta}$$

□

Corollary 3.3.5. *If $\sigma(T) = \{0\}$ then $T^n \to 0$ as $n \to \infty$.*

Remark. This above corollary is notable as it mirrors the statement for finite dimensional spaces where the spectrum reduces to the set of eigenvalues.

3.4. **Hilden's Method.**

Theorem 3.4.1 (Lomonosov, Hilden)**.** *Let $(X, \|.\|)$ be a Banach space on which $T \in \mathcal{L}(X)$ is a compact operator. Then T necessarily has a hyperinvariant subspace.*

Our paper will present two similar proofs, though the latter will be for a slightly more general statement. However, they both begin with a nearly identical construction. We assume T does not have an eigenvector.

Remark (Set Up)**.** [M77] Assume T does not have an eigenvector. Select $x^* \in X$ such that $\|T(x^*)\| > 1$ (and $x^* \notin \mathcal{B}$) and let $B = \overline{B(x^*, 1)}$ for ease of notation. By definition, $T(B)$ is relatively compact and is thus totally bounded. It also does not contain the origin. Let $y \in X$ be a nonzero vector and define $M_y \subset X$ by

$$M_y = \{Sy : S \text{ is a bounded operator that commutes with } T\}$$

so that M_y is hyperinvariant in T, as is $\overline{M_y}$. Since each M_y contains a nonzero vector (for nonzero y), it suffices to prove that there is some y for which M_y is not the entire space. We proceed by contradiction, assuming $M_y = X$ for all y. For each $S \in \mathcal{L}(X)$ let $\mathcal{U}(S) = \{y : Sy \in B\} \subseteq X$. Since M_y is dense in B we have

$$\bigcup_{ST=TS} \mathcal{U}(S) = X \sim \{0\}$$

Observe that the $\mathcal{U}(S)$ sets form an open cover of $\overline{T(B)}$, and so there exists a finite subcollection of operators, $\{S_i\}_{i=1}^n$ such that $T(B) \subseteq \bigcup_{i=1}^n \mathcal{U}(S_i)$. Finally, we are ready to present Hilden's method.

Proof (Hilden). [M77] We continue under the assumption that T has no eigenvector. By Theorem 2.3.6, Proposition 3.1.3, and Corollary 3.3.5 we know that $T^n \to 0$ (the same holds for any scalar multiple of T). By definition, there exists some i such that $T(x^*) \in \mathcal{U}(S_{i_1})$ and thus $S_{i_1}(T(x^*)) \in B$. Likewise, $T(S_{i_1}(T(x^*))) \in T(B)$ so we find by the same logic S_{i_2} such that $S_{i_2}TS_{i_1}T(x^*) \in B$ and thus extract an indefinite product for any j,

$$\left(\prod_{k=0}^{j-1} S_{i_{j-k}} \circ T\right)(x^*) \in B$$

Moreover, letting $c = \max_{1\leq i\leq n} \|S_i\|$ we get

$$\left(\prod_{k=0}^{j-1} \left(\frac{S_{i_{j-k}}}{c}\right)\right) \circ (c^jT^j)(x^*) \in B$$

but we know that $(c^jT^j) \to 0$ which implies 0 is a limit point of some sequence in B which is a contradiction. □

4. Lomonosov's Proof and Nonlinear Methods

Hilden's proof is easily the most accessible through traditional functional analysis, but Lomonosov's original derivation was equally (if not more) ingenious, and therefore deserving of a discussion in this paper. We require another major result from analysis in order to continue.

4.1. **Schauder's Theorem.**

Definition 4.1.1. Let $Y \subset X$ be a nonempty subset. We define $\text{conv}(Y)$ to be the *convex hull* of Y, the smallest convex subset containing Y.

Proposition 4.1.2. *The convex hull of a compact set is compact, and the convex hull of an open set is open.*

Theorem 4.1.3 (Brouwer's Fixed Point)**.** *Let F be a finite-dimensional space and $Q \subset F$ be a nonempty, compact convex set. If $f : Q \to Q$ is a continuous function then f has a fixed point.*

We can now generalize this result to a general Banach space. The proof below is my own work, but is assigned as an exercise in Brezis [B11].

Theorem 4.1.4. *Let E be a Banach Space where C is a nonempty, closed convex subset of E. Let $F : C \to C$ be a continuous function such that $F(C) \subset K$ where K is a compact subset of C. Then F has a fixed point in K.*

Proof. Pick $\varepsilon > 0$. Since K is compact (and therefore totally bounded), one extracts a finite $\frac{\varepsilon}{2}$-net with indices $i \in I$ so that

$$K \subset \bigcup_{i \in I} \overline{B\left(y_i, \frac{\varepsilon}{2}\right)} \;\middle|\; i \in I, y_i \in K$$

Set $q_i(x) = \max\left\{\varepsilon - \|F(x) - y_i\|, 0\right\}$. We define the continuous map, $q = \sum_{i \in I} q_i$, to be the sum of these distances.

Picking some arbitrary, $x \in C$, set $z = F(x)$ so that $z \in K$ and therefore it must be enclosed by at least one of the balls, $z \in B(y_i, \frac{\varepsilon}{2})$. Thus, $\|z - y_i\| \le \frac{\varepsilon}{2}$ and so $\varepsilon - \|z - y_i\| \ge \frac{\varepsilon}{2}$. Notice that q_i measures the 'nearness' of y_i to z, and so one approximates z by the y_i by taking a weighted average $z \approx \sum \left(\frac{q_i}{q}\right) y_i$. We have that $y_i \in I$ iff $\|z - y_i\| < \varepsilon$ by our definition. Finally, we can show that this approximation is ε fine:

$$\left\| z - \sum_{i \in I} \left(\frac{q_i}{q}\right) y_i \right\| \le \sum_{i \in I} \left\| \left(\frac{q_i}{q}\right)(z - y_i) \right\| \le \sum_{i \in I} \left(\frac{q_i}{q}\right) \varepsilon \le \varepsilon$$

We can now define a function, F_ε, that will approximate F in general

$$F_\varepsilon(x) = \frac{\sum_{i \in I} q_i(x) y_i}{q(x)}, \|F - F_\varepsilon\|_\infty = \sup_{x \in C} \|F - F_\varepsilon\| \le \varepsilon$$

where F_ε is continuous.

Let Q be the convex hull of $\{y_i\}_{i \in I}$. We know that Q is compact as it is a closed subset of K. Since the hull is generated by finitely many points, it has finite dimension. Our intent is to apply Brouwer's Theorem to $F_\varepsilon|_Q$, giving us a fixed point, $x_\varepsilon \in Q$. Note that F_ε is the convex combination of vectors in Q, and thus $F_\varepsilon(Q) \subseteq Q$.

For ease of notation, pick an arbitrary sequence, $\{\varepsilon_i\}$ descending to zero and set $x_{\varepsilon_m} = x_m, F_{\varepsilon_m} = F_m$ where $m \to \infty$ is equivalent to $\varepsilon \to 0$. We have that $F_m(x_m)$ is a sequence in K, and thus we can extract a convergent subsequence $F_{m_k}(x_{m_k}) \to \overline{x}$. For any $\eta > 0$ we have that

$$\|F(\overline{x}) - \overline{x}\| \le \|F(\overline{x}) - F_{m_k}(\overline{x})\| + \|F_{m_k}(\overline{x}) - F_{m_k}(x_{m_k})\| + \|x_{m_k} - \overline{x}\| < \eta$$

for k sufficiently large. Thus, $F(\overline{x}) = \overline{x}$ by continuity of F. □

4.2. **Lomonosov's Method.**

Remark. Before discussing this proof, it is worth understanding the extent to which Lomonosov's result was a revelation for its time. We state the results that were known up to this point.

Theorem 4.2.1 (Aronszajn and Smith, 1954)**.** *Any compact operator on a Banach space of dimension greater than one has an invariant subspace.* [A54]

The proof that they provide is fairly lengthy and relies heavily upon polynomial theory.

Definition 4.2.2. An operator, S, is *polynomially compact* if $S = p(T)$ where p is a finite degree polynomial and T is a compact operator.

Theorem 4.2.3 (Bernstein and Robinson, 1964)**.** *Any polynomially compact operator on a Hilbert space has an invariant subspace.* [L73]

This proof is fairly notorious as one of the few uses of non-standard analysis in the mathematical canon. However, the longwinded nature of these proofs made it all the more shocking when Lomonosov unveiled a one page proof of greater generality than any of its predecessors.

Theorem 4.2.4 (Lomonosov, 1973)**.** *Let $(X, \|.\|)$ be a Banach space on which $T \in \mathcal{L}(X)$ is a compact operator. Then T necessarily has a hyperinvariant subspace.*

Proof. [L73] We rely on the same setup in the lead up to Theorem 3.4.1., assuming that it does not have an eigenvector. Once again, we look at the finite covering given by $T(B) \subseteq \bigcup_{i=1}^{n} \mathcal{U}(S_i)$. For each $y \in T(B)$ let $q_i(y) = \max\{0, 1 - \|S_i(y) - x^*\|\}$ and $q(y) = \sum_{i=1}^{n} q_i(y)$ so that the weighted average

$$\Phi(y) = \sum_{i=1}^{n} \left(\frac{q_i(y)}{q(y)} \right) S_i(y)$$

is a map into the convex subset, B. Likewise, letting $\mathcal{C}$ be the convex hull of $(\Phi \circ T)(B)$ gives that $(\Phi \circ T)(\mathcal{C})$ maps into $\mathcal{C}$ (by the same logic we used in 4.1.4). Finally, Schauder's Fixed Point Theorem gives a vector, $x \in \mathcal{C}$, such that $(\Phi \circ T)(x) = x$ and thus $\sum_{i=1}^{n} \alpha_i S_i(T(x)) = x$ where $\alpha_i = \frac{q_i(y)}{q(y)}$. The subspace we look for is given by

$$W = \ker\left(\left(\sum_{i=1}^{n} \alpha_i S_i \circ T \right) - I \right)$$

and is hyperinvariant, which suffices to complete the proof. ☐

Lomonosov's proof is slightly more substantial as this method of approximating linear functions is very useful in resolving other issues in functional analysis.

5. The Counterexample

It is natural to ask, do all operators have an invariant subspace? Theorists initially questioned whether Lomonosov's criteria (i.e. that it commutes with a compact operator) extended to all operators on Banach Spaces. As it turns out, there are indeed counterexamples to each of these claims. The first was discovered by Per Enflo in 1977 before his methods were simplified by Bernard Beauzamy and Charles Reed [B88].

5.1. **Preliminaries.**

Definition 5.1.1. Let $(X, \|.\|)$ be a Banach space on which $T : X \to X$ is an operator. A point, $x \in X$, is said to be *cyclic* in T if

$$\overline{\text{span}}(x, Tx, T^2x, \dots, T^n x, \dots) = \overline{\text{span}}(O_x) = X$$

Proposition 5.1.2. *An operator, $T : X \to X$, does not have a non-trivial, closed invariant subspace if and only if every point is cyclic.*

Proof. If T has an invariant subspace, $W \subseteq X$, then W must contain O_x for any $x \in W$. For a non-cyclic point, $x \in X$, then W is a valid invariant subspace. ☐

This reformulation gives a far more concrete way to work with the invariant subspace problem. Two simple examples are listed below.

Example 3. Consider the right shift map, $T \in \mathcal{L}(l^2)$, given by $T(x_1, x_2, x_3, x_4, \dots) = (0, x_1, x_2, x_3, \dots)$. It is fairly easy to see that $e_2 = (0, 1, 0, 0, \dots) \in l^2$ is a non-cyclic vector, and thus T has a closed, nontrivial invariant subspace. To verify this claim, one can show that the closed unit ball, $\overline{B} = \{x \in X : \|x\| \le 1\}$, is such a subspace.

Corollary 5.1.3. *If $(X, \|.\|)$ is a non-separable Hilbert space (it does not have a countable, dense set), then a nonzero linear operator, $T : X \to X$, has a nontrivial, closed invariant subspace.*

Sketch. One shows that no point in such a space can be a cyclic point. Assuming the contrary (i.e. that $x \in X$ is cyclic), it suffices to show that $\text{span}\{O_x\}$ is separable since $\overline{\text{span}}\{O_x\} = X$. In a Hilbert Space, this is a known property, and therefore gives a contradiction. ☐

Proposition 5.1.4. *If x is a cyclic point then the nth order orbit, O_x^n, is a linearly independent set.*

Proof. Suppose that $T^n = \sum_{i<n} \alpha_i T^i$ for some $n \in \mathbb{N}$. For any $i > n$, one can express T^i in terms of these first n iterates. We prove this assertion by induction. For $i = n+1$ we have

$$T^{n+1} = T \circ \left(\sum_{i<n} \alpha_i T^i \right) = \sum_{i<n} \alpha_i T^{i+1}$$

which suffices. For $i = n+2$, we have

$$T^{n+2} = T \circ \sum_{i<n} \alpha_i T^{i+1} = T \circ \left(\alpha_{n-1} T^n + \sum_{i<n-1} \alpha_i T^{i+1} \right) = \alpha_{n-1} \sum_{i<n} \alpha_i T^{i+1} + \sum_{i<n-1} \alpha_i T^{i+2}$$

In general, if it holds for $i = n+j$ i.e. that $T^{n+j} = \sum_{i<n} \lambda_i T^i$ then for $i = n+j+1$ we have

$$T^{n+j+1} = T \circ \left(\sum_{i<n-1} \lambda_i T^i + \lambda_{n-1} T^{n-1} \right) = \sum_{i<n-1} \lambda_i T^{i+1} + \lambda_{n-1} \sum_{i<n} \lambda_i T^{i+1}$$

which implies that O_x is finite dimensional which contradicts our assumption that x is cyclic. □

We next need to think about the space that we will be using for our counterexample.

Definition 5.1.5. Our traditional norm, $\|.\|_1$, for a polynomial, p, is given by

$$\|p(z)\|_1 = \left\| \sum_{i \geq 0} a_i z^i \right\|_1 = \sum_{i \geq 0} |a_i|$$

Shortly, we will derive a different countable basis $\{f_0, f_1, f_2, f_3, \dots\}$ of vectors in the space of polynomials. The norm that we will be using will be given by

$$\|p(z)\| = \left\| \sum_{i \geq 0} \alpha_i f_i \right\| = \sum_{i \geq 0} |\alpha_i|$$

We define $\mathcal{P}_+$ to be the completion of the space of polynomials equipped with the norm, $\|.\|$.

Remark. Our goal is going to be show that each element of the completion is a cyclic vector. This is *not* merely the finite degree polynomials, but also those mixed in by taking the completion.

5.2. **Constructing the Norm.**

Remark. Before we construct the norm, it is worth establishing a bit more background on the space of polynomials.

Definition 5.2.1. Let $(\mathcal{P}_+, \|.\|)$ be (as before) the Banach space of polynomials with complex coefficients. The subset, $\mathcal{P}_n$, is defined as the collection of polynomials with degree not exceeding n.

Let $p(z) = \sum_{i\geq 0} a_i z^i$ be a polynomial in $\mathcal{P}_+$. We name the index of the first nonzero coefficient to be the *valuation* of the polynomial i.e. $\mathrm{val}(p) = \min\{j \mid a_j \neq 0\}$

Lemma 5.2.2. *Let $n \geq 0$ and $\varepsilon, \delta, M > 0$ and $g \in \mathcal{P}_n$ be a polynomial such that $\|g\|_1 \leq M$. There exists some constant $K > 0$ such that for any $0 \leq m \leq n$ if $\|P_m(g)\|_1 \geq \delta$ then there exists $q \in \mathcal{P}_n$ such that $\|P_n(gq) - z^m\|_1 < \varepsilon$ and $\|q\|_1 \leq K$.*

Proof. [B88] Define the compact subset (for each $0 < m \leq n$)

$$\Omega_m = \left\{p \in \mathcal{P}_n \middle| \, \|p\|_1 \leq M \text{ and } \|P_m(p)\|_1 \geq \delta\right\}$$

If $g \in \Omega_m$, then polynomial division gives a quotient and remainder, q, r, such that $z^m = gq + r$ and $\mathrm{val}(r) > n$. Thus $gq = z^m - r$ and $P_n(gq) = z^m$. We now look at the open cover of Ω_m given by neighborhoods of g, $N(g, \eta)$, that are sufficiently small ($\eta << 1$) so that if $p \in N(g, r)$ then $\|P_n(pq) - z^m\|_1 < \varepsilon$. One extracts a finite subcover, $\{U_{g_1}, U_{g_2}, \dots, U_{g_j}\}$, composed of polynomials where the remainder theorem gives respective quotients, $\{q_i\}_{i\in I} = \{q_1, q_2, q_3, \dots, q_j\}$. One bounds these quotients by $K_m = \max_{i\in I} \|q_i\|_1$. Repeating this process for each $m \leq n$ allows us to take $K = \max_{m\leq n} K_m$ which completes the proof. □

Now, we are ready to approach the construction. Each of the known counterexamples relies on some bookkeeping for a few (typically two) rapidly increasing sequences. I believe it is best to take Beauzamy's approach, holding off on writing exactly how fast each sequence must grow until we arrive at the parts of the proof that require these relationships. In the meantime, we will begin with some essentials.

Let $\{a_n\}_{n\in\mathbb{N}}, \{b_n\}_{n\in\mathbb{N}}$ be sequences in $\mathbb{N}$ such that

$$a_0 = b_0 = 1 < a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n < \dots$$

and let $\{v_n\}_{n\geq -1}$ be given by $v_n = (n-1)(a_n + b_n)$ and $v_{-1} = v_0 = 0$. We define the basis vectors, $\{f_0, f_1, f_2, f_3, \dots\} \subset \mathcal{P}_+$ from 5.1.5 as follows:

Let $f_0 = 1$, $f_1 = z - 1$, and define the function, $\Gamma_{k,r}(m,n) : \mathbb{R}^2 \to \mathbb{R}$ by

$$\Gamma_{k,r}(m,n) = \frac{\left(r - \frac{1}{2}\right)m - k}{n}$$

We begin with the case where $n \geq 2$ and $1 \leq r \leq n-1$. We require that a_n, b_n satisfy $(r-1)a_n + v_{n-r} < k \leq ra_n$ and that

$$f_k = a_{n-r}(z^k - z^{k-a_n}) \text{ if } ra_n \leq k \leq ra_n + v_{n-r-1} \tag{5.2.3}$$

$$f_k = 2^{\Gamma_{k,r}(a_n, b_{n-1})} z^k \text{ if } (r-1)a_n + v_{n-r} < k < ra_n \tag{5.2.4}$$

This allows us, for each $n \geq 1$, to define f_k for $k = v_{n-1}$ up to $k = a_n$. One can use the same logic to move upward to $k = a_n + v_{n-2}$ and then to $k = 2a_n$. One can complete this process inductively to define f for k up to $k = (n-1)a_n$. From here, we need to expand our definitions.

If $n \geq 2$ and $1 \leq r \leq n-1$ then we require that $(n-1)a_n + (r-1)b_n < r(a_n + b_n)$ and let

$$f_k = z^k - b_n z^{k-b_n} \text{ if } r(a_n + b_n) \leq k \leq (n-1)a_n + rb_n \tag{5.2.5}$$

$$f_k = 2^{\Gamma_{k,r}(b_n, a_n)} z^k \text{ if } (n-1)a_n + (r-1)b_n < k < r(a_n + b_n) \tag{5.2.6}$$

and if $n = 1$ then use (5.2.5) for $k = a_1 + b_1$ and (5.2.6) for $a_1 < k < a_1 + b_1$. Now, one can define f_k for $k = (n-1)a_n$ up to $k = a_n + b_n$, all the way to $k = v_n = (n-1)(a_n + b_n)$.

5.3. **The Remaining Lemmas.**

Remark. If (5.2.3) holds, then $v_{n-r} < a_n$ and $(n-2)(a_{n-1} + b_{n-1}) < a_n$ so that $\sum_{i=m}^{j} \frac{1}{a_i} \leq \frac{2}{a_m}$. Putting this all together, one arrives at the identities

$$\begin{aligned} z^k - z^{k-a_n} &= \frac{1}{a_{n-r}} f_k \\ z^{k-a_n} - z^{k-2a_n} &= \frac{1}{a_{n-r+1}} f_{k-a_n} \\ &\vdots \end{aligned}$$

$$z^k - z^{k-ra_n} = \sum_{i=0}^{r-1} \frac{1}{a_{n-(r-i)}} f_{k-ia_n} \implies \left\| z^k - z^{k-ra_n} \right\| = \sum_{i=0}^{r-1} \frac{1}{a_{n-(r-i)}} \leq \frac{2}{a_{n-r}}$$

One can show almost identically that when (5.2.5) holds,

$$\left\| z^k - b_n^r z^{k-rb_n} \right\| = \sum_{i=0}^{r-1} b_n^i \leq 2b^{r-1}$$

Theorem 5.3.1. *The operator, $T \in \mathcal{L}(\mathcal{P}_+)$, given by*

$$T\left(\sum_{i\geq 0} a_i z^i\right) = \sum_{i\geq 0} a_i z^{i+1}$$

does not have an invariant subspace.

Lemma 5.3.2. *If $\{a_n\}, \{b_n\}$ grow sufficiently rapidly then $\|T\| \leq 2$ and the three identities below hold for $n, r \in \mathbb{N}$*

$$\|Tf_{ra_n+v_{n-r-1}}\| \leq \frac{1}{a_{n-r}} \tag{5.3.3}$$

$$\|Tf_{ra_{n-1}}\| \leq \frac{1}{a_n} \tag{5.3.4}$$

$$\left\|Tf_{r(a_n+b_n)-1}\right\| \leq \frac{1}{b_n} \tag{5.3.5}$$

These bounds follow fairly easily from the definitions so long as we require that $\{a_n\}, \{b_n\}$ grow sufficiently rapidly. An exact proof is found in Beauzamy's paper.

Definition 5.3.6. We define $O_m \subseteq \mathbb{R}$ for each $m \in \mathbb{N}$ by

$$O_m = \{k \in \mathbb{N} : \exists n > m, (n-m)a_n \leq k \leq (n-m)a_n + v_{m-1}\} = \bigcup_{n>m} [(n-m)a_n, (n-m)a_n + v_{m-1}]$$

We are now prepared to present the next key lemma. I will refrain from proving this statement as the proof is very technical and does not shed much insight. It can be found in Beauzamy's paper.

Lemma 5.3.7. *If $\{a_n\}, \{b_n\}$ grows sufficiently rapidly for $m > 2$ and $k > (m-1)a_m, b_m + a_m \leq s \leq b_m + (m-1)a_m$ then*

$$\|T^s f_k\| \leq 4 \text{ if } k \notin O_m \tag{5.3.8}$$

$$\left\|T^s f_k + a_m z^{j+s}\right\| \leq 1 \text{ if } k \in O_m \text{ and we let } k = (n-m)a_n + j \tag{5.3.9}$$

Although I am omitting the justification, I am going to flesh out the utility of this statement.

Definition 5.3.10. Let $Q_m : \mathcal{P}_+ \to \mathcal{P}_{(m-1)a_m}$ be defined by

$$Q_m f_k = \begin{cases} f_k : & 0 \leq k \leq (m-1)a_m \\ 0 : & k > (m-1)a_m, k \notin O_m \\ -a_m z^j : & \text{else, where } j \text{ is as above} \end{cases}$$

be a linear map.

Proposition 5.3.11. *If $m > 2$ and $b_m + a_m \leq s \leq b_m + (m-1)a_m$ then $\|T^s f_k - T^s Q_m f_k\| \leq 4$ for $k \in \mathbb{N}$.*

Proof. [B88] Before delving into the more difficult cases, we have that if $k \le (m-1)a_m$ then $Q_m f_k = f_k$ gives $\|T^s f_k - T^s Q_m f_k\| = 0$

If $k \ge (m-1)a_m$ and $k \notin O_m$ then $Q_m f_k = 0$ so that Lemma 5.3.7 gives

$$\|T^s f_k - T^s Q_m f_k\| = \|T^s f_k\| \le 4$$

Otherwise, $Q_m f_k = -a_m z^{j+s}$ (where j is the same as in 5.3.7) which gives $\|T^s f_k - T^s Q_m f_k\| \le 1$ due to Lemma 5.3.7. □

Corollary 5.3.12. *If $g \in \mathcal{P}_+$ then $\|T^s g - T^s Q_m g\| \le 4\|g\|$.*

Proof. [B88] It suffices to prove that Q_m is bounded/continuous. Recall that $v_{m-1} < (m-1)a_m < v_m$ and that for each $k \le v_{m-1}$ one can find f_k in terms of $a_1, b_1, a_2, b_2, \dots, a_{m-1}, b_{m-1}, a_m$. If $0 \le j \le (m-1)a_m$ there exists a constant, $C_m = C_m(a_1, b_1, a_2, b_2, \dots, a_{m-1}, b_{m-1})$, depending on these sequence entries such that $\|f_j\|_1 \le C_m$ gives a valid upper bound. Moreover, we have that $\|Q_m f_k\| \le a_m C_m$ and $\|Q_m g\| \le a_m C_m \|g\|$ which completes the proof. □

Remark. It is useful to study the relationship between the two norms we have established. An important notion for the $\|.\|_1$ norm is that $\|z^\alpha p\|_1 = \|p\|_1$ for any polynomial, $p \in \mathcal{P}_+$, and any positive integer, $\alpha \in \mathbb{N}$. For $j \le (m-1)a_m$, one can express z^j in terms of $\{f_1, f_2, \dots, f_j\}$. Depending on whether or not $j \le v_{m-1}$, this linear combination will involve the elements of the sequence, $\{a_1, b_1, \dots, a_{m-1}, b_{m-1}\}$, or if larger, it will go up to $\{a_1, b_1, \dots, a_{m-1}, b_{m-1}, a_m\}$. This will allow us to define a constant, $C^* = C^*(a_1, b_1, \dots, a_{m-1}, b_{m-1}, a_m)$, such that $\|z^j\| \le C^*$. Moreover,

$$\|Q_m f_k\|_1 \le \max_{k \le (m-1)a_m} (\|f_k\|_1, a_m) \le C_m \implies \|Q_m g\|_1 \le C_m \|g\| \tag{5.3.13}$$

Corollary 2.2.4 gives that the norm, $\|.\|_*$, given by $\|g\|_* = \|Q_m g\|$ is equivalent to $\|.\|$.

Lemma 5.3.14. *If $g \in \mathcal{B}$ (the unit sphere) and there exists $m > 2, 1 \le r \le m-2$ for which*

$$\|(P_{ra_m} \circ Q_m)g\|_1 \ge \frac{1}{a_m} \tag{5.3.15}$$

then there exists some $\varphi \in \mathcal{P}_+$ such that

$$\|(\varphi \circ T)g - 1\| \le \frac{3}{a_{m-r-1}} \tag{5.3.16}$$

Proof. [B88] Fix m and r as above. Let $h = Q_m g$ so that the above remark gives $\|h\|_1 \le C_m$ so that Lemma 5.2.2 gives a positive constant, $K = K(C_m, a_m)$, and a polynomial q such that

$$\left\|P_{(m-2)a_m}(qh) - z^{ram}\right\|_1 < \frac{1}{a_m C_m}, \quad \|q\|_1 \le K$$

Set $\varphi = z^{a_m} q$ so that $\mathrm{val}(\varphi) \ge a_m$ and $\varphi \in \mathcal{P}_{(m-1)a_m}$ and

$$\left\|P_{(m-2)a_m}(\varphi h) - z^{(r+1)a_m}\right\|_1 < \frac{1}{a_m C_m}$$

and (5.3.13) gives that

$$\left\|P_{(m-2)a_m}(\varphi h) - z^{(r+1)a_m}\right\| < \frac{1}{a_m}$$

Set $\varphi_* = \frac{1}{b_m} z^{b_m} \varphi = \frac{1}{b_m} z^{a_m+b_m} q$ so that if $r(a_m + b_m) \le k \le (m-1)a_m + rb_m$ then 5.2.5 gives

$$\left\|\frac{1}{b_m} z^k - z^{k-b_m}\right\| = \frac{1}{b_m}, \quad \left\|\frac{1}{b_m} z^{k'} - z^{k'-b_m}\right\| = \frac{1}{b_m}$$

when $k' = k - b_m$. Next we define the constant given by

$$A = \left\|P_{(m-1)a_m b_m}(\varphi_* h) - P_{(m-1)a_m}(\varphi h)\right\| = \left\|P_{(m-1)a_m b_m}(\frac{1}{b_m} z^{b_m} \varphi h) - P_{(m-1)a_m}(\varphi h)\right\|$$

The standard-basis expansion of $\varphi h = \sum_{j \ge 0} \gamma_j z^j$ gives a string of inequalities

$$A \le \frac{1}{a_m} \sum_{j \ge 0} |\gamma_j| = \frac{1}{b_m} \|\varphi h\|_1 \le \frac{1}{b_m} \|\varphi\|_1 \|h\|_1$$

$$\le \frac{C_m}{b_m} \|\varphi\|_1 = \frac{C_m}{b_m} \|\varphi\|_1 = \frac{C_m}{b_m} \|q\|_1 \le \frac{KC_m}{b_m} \le \frac{1}{a_m}$$

where the final inequality emerges if b_m grows sufficiently larger than a_m and (5.2.4) gives (for b_m sufficiently large)

$$\left\|\varphi_* h - P_{(m-1)a_m+b_m} \varphi_* h\right\| \le 2^{\Gamma^*} \|\varphi h\|_1 \le \frac{KC_m}{b_m} 2^{\Gamma^*}, \quad \Gamma_* = \frac{2(m-1)a_m - 1/2b_m}{ma_m}$$

where $b_m \ge KC_m a_m$ gives $\left\|\varphi_* h - P_{(m-1)a_m+b_m}(\varphi_* h)\right\| \le \frac{1}{a_m}$ as an upper bound.

Since $\varphi_* = \frac{1}{b_m} z^{a_m+b_m} q$ we can expand in the standard basis $\varphi_* = \sum_{s=a_m+b_m}^{(m-1)a_m+b_m} \lambda_s z^s$ and note that $\|\varphi_*\|_1 \le \frac{K}{b_m}$ and thus $|\lambda_s| \le \frac{K}{b_m}$ for each (relevant) s. Lemma 5.3.2 gives that $\|T^s g - T^s Q_m g\| \le 4\|g\|$ and whenever $a_m + b_m \le s \le (m-1)a_m + b_m$ it follows that

$$\|T^s g - T^s Q_m g\| \le 4\|g\|$$

and therefore that

$$\|(\varphi_* \circ T)g - \varphi_* h\| = \left\|\sum_{s=a_m+b_m}^{(m-1)a_m+b_m} \lambda_s z^s g - \sum_{s=a_m+b_m}^{(m-1)a_m+b_m} \lambda_s z^s h\right\| \le \|g\| \frac{4K}{b_m} < \frac{1}{a_m}$$

where the final inequality holds for $b_m > 4ka_m$.

Finally, we can complete the proof. Recall at the beginning of this section where we discussed that for $1 \leq r \leq n-1$ and $ra_n \leq k \leq ra_n + v_{n-r-1}$ we have $\|z^k - z^{k-a_n}\| \leq \frac{2}{a_{n-r}}$ and that if $k = ra_n$ we get $\|z^{ra_n} - 1\| \leq \frac{1}{a_{n-r}}$. One arrives at the equivalent formulation, $\|z^{(r+1)a_m} - 1\| \leq \frac{2}{a_{m-r-1}}$. Applying the triangle inequality gives our desired identity:

$$\begin{aligned}\|(\varphi_* \circ T)g - 1\| &\leq \|(\varphi_* \circ T)g - \varphi_* h\| + \left\|\varphi_* h - P_{(m-1a_m+b_m}(\varphi_* h)\right\| \\ + \left\|P_{(m-1a_m+b_m}(\varphi_* h) - P_{(m-1)a_m}(\varphi h)\right\| &+ \left\|P_{(m-1)a_m}(\varphi h) - z^{(r+1)a_m}\right\| + \left\|z^{(r+1)a_m} - 1\right\| \\ &\leq \frac{4}{a_m} + \frac{2}{a_{m-r-1}} \leq \frac{3}{a_{m-r-1}}\end{aligned}$$

□

5.4. **The Proof.**

Remark. Recall our discussion at the beginning of this section. As long as $1 \in \mathcal{P}_+$ is in the orbit of each element of $\mathcal{P}_+$ (i.e. not just finite polynomials) then each vector, $p \in \mathcal{P}_+$, will by cyclic. This holds since $O_1 = \mathcal{P}$ and $1 \in O_p$ implies that $O_1 \subseteq O_p$.

The convenience of polynomials is that the elements formed by the span of any given nth orbit can be equivalently given by multiplying the polynomial by another polynomial. Of course, for the elements which are not finite degree polynomials it is more practical to use the completeness of the space. Thus, in order to finally prove Theorem 5.3.1, it suffices to show that for each $g \in \mathcal{B}$ and $\varepsilon > 0$ there exists some $\varphi_* = \varphi_*(\varepsilon)$ such that $\|\varphi_* Tg - 1\| < \varepsilon$.

Proof. [B88] Fix $\varepsilon > 0$ and select k such that $\frac{1}{a_k} < \frac{\varepsilon}{3}$. By Lemma 5.3.14, it suffices to find $m, r \in \mathbb{N}, m-r-1 \geq k$ such that $\|P_{ra_m}(Q_m g)\|_1 \geq \frac{1}{a_m}$. We proceed by contradiction, and assume that no such r, m exist such that $\|P_{ra_m}(Q_m g)\|_1 < \frac{1}{a_m}$ for all r, m. We note the elementary identity

$$\|P_{a_m}(Q_m g)\|_1 < \frac{1}{a_m} \tag{5.4.1}$$

For each m, let $D_m = D_m(a_1, b_1, a_2, b_2, \ldots, a_m, b_m)$ be defined such that $\|z^j\| \leq D_m$ for $0 \leq j \leq v_m$. Expand g in the $\{f_k\}$ basis, $g = \sum_{j\geq 0} \alpha_j f_j$ where $\sum_{j\geq 0} |\alpha_j| = 1$. For each $n > 0$, we divide g into two parts

$$g = \sum_{j=0}^{(n-1)a_n} \alpha_j f_j + \sum_{j>(n-1)a_n} \alpha_j f_j = \psi_n + \kappa_n, \quad \psi_n, \kappa_n \in \mathcal{P}_+$$

In terms of the standard basis, we have that

$$\psi_n = \sum_{j=0}^{(n-1)a_n} \beta_{n,j} \ , \quad Q_n(\kappa_n) = \sum_{j=0}^{v_{n-1}} \lambda_{n,j} z^j$$

where $\lambda_{n,j} = -a_n \sum_{m>n} \alpha_{j+(m-n)a_m}$. This follows directly from (5.3.10). Note that $Q_n(\psi_n) = \psi_n$, and so

$$Q_n(g) = \sum_{j=0}^{(n-1)a_n} \beta_{n,j} z^j + \sum_{j=0}^{v_{n-1}} \lambda_{n,j} z^j$$

If (5.2.3) holds then $\beta_{m,j} = a_{m-r}\alpha_j$ and our assumption gives that $\sum_{j\leq ra_m} |\beta_{n,j}| < \frac{1}{a_m}$ or likewise that $\sum_{ra_m+v_{m-r-2}}^{ra_m+v_{m-r-1}} |\alpha_j| < \frac{1}{a_m a_{m-r}}$. Moreover, for $n > k+2$ some further manipulation gives that

$$\sum_{j=v_{n-r-2}+1}^{v_{n-1}} |\lambda_{n,j}| = \sum_{j=v_{n-2}+1}^{v_{n-1}} |a_n| \sum_{m>n} |\alpha_j| = a_n \sum_{m>n} \sum_{j=v_{n-2}+1}^{v_{n-1}} |\alpha_{j+(m-n)a_m}| < \sum_{m>n} \frac{1}{a_m} < \frac{2}{a_{n+1}}$$

Applying (5.4.1) gives $\sum_{j=0}^{v_{n-1}} |\beta_{n,j}+\lambda_{n,j}| < \frac{1}{a_n}$ which gives immediately that $\sum_{j=v_{n-2}+1}^{v_{n-1}} |\beta_{n,j}| < \frac{2}{a_n}$.

From our original definitions, we know that if j satisfies (5.2.3) or (5.2.4) i.e. that $(r-1)a_n + v_{n-r} < j \leq ra_n$ then f_j will not contain a term of degree $i \in (v_{n-2}, v_{n-1}]$. Applying this to our situation gives,

$$\sum_{v_{n-2}+1}^{v_{n-1}} \alpha_j f_j - \sum_{n-2+1\leq v_j \leq n-1} \beta_{n,j} z_j \in \mathcal{P}_{v_{n-2}}$$

and thus

$$\sum_{j=v_{n-2}+1}^{v_{n-1}} |\alpha_j| \leq D_{n-1} \sum_{j=v_{n-2}+1}^{v_{n-1}} |\beta_{n,j}| < \frac{2D_{n-1}}{a_n} < \frac{1}{\sqrt{a_n}}$$

where the final inequality holds for a_n sufficiently large. Equivalently, we have that $\sum_{j=v_{m-1}+1}^{v_m} |\alpha_j| < \frac{1}{\sqrt{a_{m+1}}}$. Returning to one of our expansions from above i.e.

$$\sum_{j=v_{n-r-2}+1}^{v_{n-1}} |\lambda_{n,j}| = \sum_{j=v_{n-2}+1}^{v_{n-1}} |a_n| \sum_{m>n} |\alpha_j| \leq a_n \sum_{m>n} \frac{1}{\sqrt{a_{m+1}}} \leq \frac{1}{a_{n+1}}$$

Using that $\sum_{j=0}^{v_{n-1}} |\beta_{n,j}| = \left\| P_{v_{n-1}}\left(\sum_{j=0}^{(n-1)a_n} \alpha_j f_j \right) \right\|_1 < 2/a_n$ we can conclude that

$$\sum_{j=v_{n-1}+1}^{(n-1)a_n} |\alpha_j| \leq \frac{1}{\sqrt{a_{n+1}}} < \frac{1}{a_n}$$

If $v_{n-1} < j \le (n-1)a_n$ then our definitions yield

$$\|P_{v_{n-1}}(f_j)\|_1 \le a_{n-1} \implies \left\| P_{v_{n-1}}\left(\sum_{k=v_{n-1}+1}^{(n-1)a_n} \alpha_j f_j \right) \right\|_1 < \frac{a_{n-1}}{a_n}$$

which combines to give

$$\left\| \sum_{j=0}^{v_{n-1}} \alpha_j f_j \right\| < \frac{a_{n-1}+2}{a_n} \implies \sum_{j=0}^{v_{n-1}} |\alpha_j| \le D_{n-1} \frac{a_{n-1}+2}{a_n}$$

which gives that $\sum_{j\ge0} |\alpha_j| = 0$ as $n \to \infty$ which contradicts our assumption. $\square$

Corollary 5.4.2 (C. Read)**.** *There exists an operator on l^1 with no invariant subspace.* [R10]

It is not an immediate corollary, but it emerges due to the fact $l^1 \cong \mathcal{P}_+$. An actual construction can be found in Reed's paper.

6. Acknowledgements

It has been a pleasure to work on and write about this topic during this otherwise difficult time. I want to issue a special thank you to my advisor, Dolores Walton, for being a constant source of support and helping me refine several of the arguments that I presented here. I also express my gratitude to Peter May and the other volunteers who continued to run this program and transitioned to a remote setting.

Finally, I want to thank the professors and mentors who have gotten me to this point. Cheers to Andre Neves who introduced me to this area of study in the fall, as well as the remainder of the amazing 2019-2020 H. Analysis professors. And a final thank you to Adam Wagner for encouraging me to pursue mathematics, and for being a constant source of support.

References

[A54] N. Aronszajn and K. Smith, "Invariant subspaces of completely continuous operators," Ann. Math.,60, 345–350 (1954).

[A16] Axler, Sheldon J. Linear Algebra Done Right. New York: Springer, 2016. Print.

[B88] Beauzamy, B. (1988). Introduction to operator theory and invariant subspaces. Amsterdam: North-Holland.

[B11] Brézis, H. (2011). Functional analysis, sobolev spaces and partial differential equations. New York, NY: Springer.

[L73] Lomonosov, V.I. Invariant subspaces for the family of operators which commute with a completely continuous operator. Funct Anal Its Appl 7, 213–214 (1973). https://doi.org/10.1007/BF01080698

[M77] Michaels, A. J.. "Hilden's simple proof of Lomonosov's invariant subspace theorem." Advances in Mathematics 25 (1977): 56-58.

[R85] Read, C. "A Solution to the Invariant Subspace Problem on the Space l^1" The Bulletin of the London Mathematical Society, Vol 25 Issue 4 (1985): 305-317.

[R10] Royden, H. L., & Fitzpatrick, P. (2010). Section II. In Real Analysis (pp. 183-334). Boston, Mass.: Prentice Hall.

[T11] Tao, T. (2011, April 17). A proof of the Fredholm alternative. Retrieved August 10, 2020, from https://terrytao.wordpress.com/2011/04/10/a-proof-of-the-fredholm-alternative/