Operators that Dominate Normal Operators

Jim Agler *
John E. McCarthy †
University of California at San Diego, La Jolla California 92093
Washington University, St. Louis, Missouri 63130

September 15 1997

0 Introduction

Let T be an operator on a Hilbert space (by *operator* we shall always mean a bounded linear operator). We say that T *dominates* an operator N if there is a non-zero operator L such that

$$LT = NL. \tag{0.1}$$

The order of this intertwining is important. We shall say that T dominates a normal operator if there is some normal N for which (0.1) holds. The object of this paper is to study operators that dominate normals.

The reason for the term "dominate" is that, under certain assumptions (A1 - A3 below), it follows from equation (0.1) that there is a vector γ and a constant c such that the inequality

$$\int |p|^2 d\mu \leq c\|p(T)\gamma\|^2$$

holds for all polynomials p, where μ is a scalar spectral measure of N.

There are two reasons to study these operators. The first is that they allow one to use what is known about the structure of normal operators to study a larger class of operators. One could then view this study as an extension of the theory of subnormal operators. (A subnormal operator is the restriction of a normal operator to an invariant subspace; it

*Partially supported by the National Science Foundation

†Partially supported by National Science Foundation grant DMS 9531967. The second author would like to acknowledge gratefully the hospitality of the mathematics department at UCSD, where most of this work was done.

corresponds in our setting to L being the inclusion map from an invariant subspace of X to the whole space on which X acts.) The theory of subnormal operators has been richly developed - see *e.g.* the book by Conway [3].

The second reason to study these operators comes from the first author's theory of states [1]. Operators that dominate normals correspond to non-extreme points of the state space, and so in some sense are generic.

We shall be mainly concerned with invariant subspaces of operators that dominate normals. (By an invariant subspace of T we shall always mean a closed, non-zero, proper subspace that T leaves invariant). If the operator L in (0.1) has a kernel, this will automatically be an invariant subspace of T; so we shall always assume that L has no kernel, and where convenient we shall take T to be cyclic. The closure of the range of L, call it $\mathcal{L}$, is an invariant subspace for N; the reducing subspace it generates, the closed linear span of $\{N^{*k}\xi : \xi \in \mathcal{L}\}$, we shall call $\mathcal{K}$. Replacing N by $N|_{\mathcal{K}}$ loses no information from (0.1), so we shall do this. We shall show (Proposition 1.1) that if $\mathcal{L}$ is not all of $\mathcal{K}$, then T has an invariant subspace. Thus, after Section 1, we shall make the following assumptions:

(A1) L has no kernel.

(A2) L has dense range.

(A3) T is cyclic.

If T is cyclic and dominates the operator N, then we show (Proposition 1.2) that there is another operator $\tilde{T}$ such that T is unitarily equivalent to the restriction of $\tilde{T} \oplus N$ to some invariant subspace $\mathcal{M}$:

$$T \;\cong\; \begin{pmatrix} \tilde{T} & 0 \\ 0 & N \end{pmatrix}\Bigg|_{\mathcal{M}} \tag{0.2}$$

Conversely, for any operator $\tilde{T}$ and any invariant subspace $\mathcal{M}$ of $\tilde{T} \oplus N$ that is not zero in the second slot, the restriction of $\tilde{T} \oplus N$ to $\mathcal{M}$ is an operator that dominates N. We study operators of this form. In Section 2, we consider the case $\tilde{T} = S^*$, where S is the unilateral shift. We give a complete description of the lattice of $S^* \oplus N$ for N a reductive cyclic normal operator, and show that if $\tilde{T}$ in (0.2) is the backward shift, then T has an invariant subspace.

In Section 3 we extend the invariant subspace result to the case that $\tilde{T} = S^{*(n)}$, where n is a positive integer. Unfortunately we cannot extend the result to the case $n = \infty$: if we could, it would follow that all operators that dominate normal operators have invariant subspaces.

In Section 4 we give examples of operators that dominate the operator of multiplication by x on $[0, 1]$, and show that three fundamentally different sorts of invariant subspaces can arise.

In Section 5 we collect some miscellaneous results.

1 Preliminaries

We show that if $\mathcal{L}$, the closure of the range of L, is not a reducing subspace for N, then T has an invariant subspace. We shall assume that N has been replaced, if necessary, by its restriction to $\mathcal{K}$, the reducing subspace generated by the range of L. Then Thomson's theorem on bounded point evaluations [12] will yield that either $N|_{\mathcal{L}}$ is normal (so $\mathcal{L} = \mathcal{K}$), or that bounded point evaluations pull back to T.

Proposition 1.1 *Suppose $LT = NL$ where N is normal and T is cyclic. Then either:*
(a) $N|_{\mathcal{L}}$ is normal.
or
(b) T has an invariant subspace of codimension one.

PROOF: Let $R = N|_{\mathcal{L}}$, and assume it is not normal. Let γ be a cyclic vector for T; then $L\gamma$ is cyclic for R. As R is subnormal, by Thomson's theorem [12] there is a complex number λ for which $(R-\lambda)\mathcal{L}$ is closed and of codimension one in $\mathcal{L}$. Let $\mathcal{M} = L^{-1}(R-\lambda)\mathcal{M}$.

It is easy to check that $\mathcal{M}$ is a closed subspace that is T-invariant, and it contains all vectors of the form $(T-\lambda)p(T)\gamma$, so is of codimension at most one. Let β be a non-zero vector in $\mathcal{L}$ that is orthogonal to $(R-\lambda)\mathcal{L}$. As

$$LL^{-1}(S-\lambda)\mathcal{L} \;\subseteq\; (S-\lambda)\mathcal{L},$$

it follows that $L^*\beta$ is orthogonal to $\mathcal{M}$; and

$$\langle L^*\beta, \gamma\rangle \;=\; \langle \beta, L\gamma\rangle \;\neq\; 0.$$

Therefore $\mathcal{M}$ is of codimension exactly one. $\square$

Next, we show that if R is an operator that dominates another operator X (normal or not), and T is the restriction of R to any cyclic subspace (*i.e.* a subspace of the form $cl(\{p(R)\xi : p \text{ a polynomial}\})$ for some vector ξ) then T has an extension to an operator of the form $\tilde{T} \oplus X$, where $\tilde{T}$ is similar to T. In other words, there is an invariant subspace $\mathcal{M}$ of $\tilde{T} \oplus X$ such that T is unitarily equivalent to $\tilde{T} \oplus X$ restricted to $\mathcal{M}$.

Proposition 1.2 *Suppose $LR = XL$ and L has no kernel. Let ξ be any non-zero vector, and let T be the restriction of R to the cyclic subspace generated by ξ. Then there is an* *operator $\tilde{T}$ similar to T such that T is unitarily equivalent to the restriction of $\tilde{T} \oplus X$ to some invariant subspace $\mathcal{M}$. Moreover, the condition number of the similarity between T and $\tilde{T}$ can be chosen arbitrarily close to 1.*

PROOF: Let $L\xi = \eta$, and choose ε so that $0 < \varepsilon < \frac{1}{\|L\|}$. Define an inner product $\langle .,. \rangle_1$ on polynomials by

$$\langle p, q \rangle_1 \ = \ \langle p(T)\xi, q(T)\xi \rangle - \varepsilon^2 \langle p(X)\eta, q(X)\eta \rangle.$$

Let $\tilde{\mathcal{H}}$ be the completion of the polynomials with respect to this sesquilinear form, and let $\tilde{T}$ be the operator of multiplication by the independent variable on $\tilde{\mathcal{H}}$.

Then $\tilde{T}$ is similar to T, and the condition number of the similarity is at most

$$\frac{1 + \varepsilon^2 \|L\|^2}{1 - \varepsilon^2 \|L\|^2},$$

which can be made arbitrarily close to 1.

Let the space on which T acts be called $\mathcal{H}$. Let $\mathcal{M}$ be the cyclic subspace of $\tilde{T} \oplus X$ generated by the vector $(1, \varepsilon\eta)$. Define $U : \mathcal{H} \to \mathcal{M}$ by $p(T)\xi \mapsto (p, \varepsilon p(X)\eta)$. This is an isometry on a dense set, so extends to a unitary onto its range, which by construction is all of $\mathcal{M}$. Moreover, it intertwines T and $\tilde{T} \oplus X$, as desired. $\square$

Note that although this construction yields a $\tilde{T}$ similar to T, if one starts with an operator $\tilde{T}$, and looks at $\tilde{T} \oplus N$ restricted to an invariant subspace whose projection onto the second component is not zero, then one gets an operator that dominates N but that is *not*, in general, similar to the operator $\tilde{T}$. Choosing $\tilde{T}$ and working backwards is what we do in the next two sections.

2 The case $\tilde{T} = S^*$

Let S be the unilateral shift on the Hardy space H^2, and let N be a contractive normal operator. In this section, we shall study operators of the form $S^* \oplus N$ restricted to a cyclic invariant subspace, as examples of operators that dominate N. Rather than directly characterizing the cyclic invariant subspaces of $S^* \oplus N$, we shall instead look at the invariant subspaces of the adjoint $S \oplus N^*$, and then take orthogonal complements.

By assumptions (A1) - (A3), we may assume that N is cyclic. The spectral theorem then yields that N^* is unitarily equivalent to some N_μ, the operator of multiplication by z on $L^2(\mu)$ for μ a finite compactly supported Borel measure on $\overline{\mathbb{D}}$ (see *e.g.* [2]).

First, we characterize the cyclic subspaces of $S \oplus N_\mu$. For any vector (g, v) in $H^2 \oplus L^2(\mu)$, let $[(g, v)]$ denote the cyclic subspace of $S \oplus N_\mu$ it generates. Let σ be normalized Lebesgue measure on the unit circle $\mathbb{T}$, and write μ as $\mu_1 + \mu_2$, where

$$\begin{aligned} \mu_1 &= \mu|_{\mathbb{D}} + \frac{d\mu}{d\sigma}\sigma \\ \mu_2 &= \mu|_{\mathbb{T}} - \frac{d\mu}{d\sigma}\sigma \end{aligned}$$

By Forelli's lemma [4, p. 43], there is a bounded sequence of polynomials that tends to 1 μ_2 a.e., and to 0 μ_1 a.e., so $[(g, v)]$ decomposes as $L^2(\mu_2) \oplus [(g, v\chi_F)]$, where F is a set of μ_2 measure zero. Thus the interesting part of $[(g, v)]$ corresponds to μ_1, that part of μ that lives on the open disk or is absolutely continuous with respect to Lebesgue measure on the circle.

The *Smirnov class* N^+ consists of those functions f analytic on the unit disk for which

$$\lim_{r \to 1} \frac{1}{2\pi} \int_{\mathbb{T}} log(1 + |f(re^{i\theta})|) d\theta = \frac{1}{2\pi} \int_{\mathbb{T}} log(1 + |f(e^{i\theta})|) d\theta < \infty.$$

(The boundedness of the integrals on the left implies that f is in the Nevanlinna class, and so has non-tangential limits almost everywhere on the circle. It is this boundary function that we mean in the second integral.) For basic facts about N^+ see *e.g.* [5] or [8]. They can be thought of as functions that are the ratio of two H^2 functions, where the denominator is outer.

We need a preliminary lemma.

Lemma 2.1 *Let f be an outer function in H^2 with $f(0) > 0$, and let r_n be a sequence of positive numbers that increase to 1. For each n, let h_n be the outer function that is positive at zero and satisfies*

$$|h_n(e^{i\theta})| = \min(|f(e^{i\theta})|, |f(r_n e^{i\theta})|).$$

Then h_n converges to f in H^2.

PROOF: As the norms of the functions h_n are bounded, they have a weak cluster point, h. By passing to a subsequence if necessary, assume that h_n converges weakly to h.

As h_n is outer,

$$\log |h_n(0)| = \int_{\{|f|, |f_{r_n}| > 1\}} \log^+(\min(|f|, |f_{r_n}|)) - \int_{\{\min(|f|, |f_{r_n}|) < 1\}} \log^-(\min(|f|, |f_{r_n}|)). \quad (2.2)$$

The first integral tends to $\int \log^+ |f|$ because f_{r_n} tends to f in measure and the family $\log^+ |f_{r_n}|$ is uniformly integrable [4, V.1.2]. The second integrand satisfies

$$\log^-(\min(|f|, |f_{r_n}|)) \leq \log^- |f| + \log^- |f_{r_n}|$$

Moreover,

$$\begin{aligned}
\lim_{n\to\infty} \int \log^- |f_{r_n}| &= \lim_{n\to\infty} [\int \log^+ |f_{r_n}| - \log |f_{r_n}(0)|] \\
&= \int \log^+ |f| - \log |f(0)| \\
&= \int \log^- |f|
\end{aligned}$$

So by the generalized Lebesgue Convergence Theorem [10, p.92], the second integral in (2.2) converges to $\int \log^- |f|$. Therefore

$$\log |h(0)| = \lim_{n\to\infty} \log |h_n(0)| = \log |f(0)|.$$

As f is outer, and $|h(z)| \leq |f(z)|$ for all z in $\mathbb{D}$, we get that h actually equals f. Moreover, as

$$\lim_{n\to\infty} \|h_n\| = \|f\|,$$

we get that h_n converges to f in norm. Finally, as we proved that any weak cluster point of the original sequence had to be f, we get that the original sequence must converge in norm to f, as desired. $\square$

We can now characterize the cyclic subspaces of $S \oplus N_\mu$.

Theorem 2.3 *Let μ be a measure on $\overline{\mathbb{D}}$ that is absolutely continuous with respect to σ on $\mathbb{T}$. Let (g, v) be in $H^2 \oplus L^2(\mu)$. Let $\mathcal{M}$ be the closed cyclic subspace of $S \oplus N_\mu$ generated by (g, v). Then*

$$\mathcal{M} = \{ \begin{pmatrix} fg \\ fv \end{pmatrix} : f \in N^+, \ fg \in H^2, \ fv \in L^2(\mu) \}.$$

Proof: ($\subseteq$): Let p_n be a sequence of polynomials such that $p_n g$ converges in H^2 to some function h, and $p_n v$ converges in $L^2(\mu)$ to a function v. Then the inner factor of g divides the inner factor of h, so $h/g := f$ is in N^+.

Moreover, by passing to a subsequence if necessary, $p_{n_k} g$ converges pointwise a.e. (σ) and everywhere on $\mathbb{D}$, so p_{n_k} converges μ a.e. to f, and therefore $p_{n_k} v$ converges to fv.

($\supseteq$) Let $A = S \oplus N_\mu$. We need to show that for any f in N^+, if fg is in H^2 and fv is in $L^2(\mu)$, then (fg, fv) is in $\mathcal{M}$. The proof shall be in three steps.

(i) Suppose first that f is in H^∞. Then $f(A)$ is in the σ-weakly closed algebra generated by A [3, II.11.4], and therefore in the weak operator topology closed algebra generated by A (which is actually the same thing [9]). Therefore $f(A)(g, v) = (fg, fv)$ is in the weak-closure of $\mathcal{M}$ which is $\mathcal{M}$.

(ii) Now suppose f is outer and $f(0) > 0$. Let h_n be as in Lemma 2.1. Each h_n is bounded, so $(h_n g, h_n v)$ is in $\mathcal{M}$. Moreover

$$\int |h_n g|^2 d\sigma + \int |h_n v|^2 d\mu \;\leq\; \int |fg|^2 d\sigma + \int |fv|^2 d\mu \;<\; \infty.$$

Therefore some subsequence of $(h_n g, h_n v)$ has a weak limit, (h, w) say, which is in $\mathcal{M}$; and by Lemma 2.1, h_n tends to f in H^2, and in particular almost uniformly on $\mathbb{D}$. By passing to convex combinations of the h_n , and then passing to a subsequence, we can assume that for some sequence of H^∞ functions k_n, $(k_n g, k_n v)$ tends to (h, w) in norm and almost everywhere. But $h/g = f$ on $\mathbb{D}$ and therefore σ a.e. on $\mathbb{T}$, so w/v also equals f μ a.e. Therefore (fg, fv) is in $\mathcal{M}$.

(iii) For a general f, factor it as IF where I is inner and F is outer, $F(0) > 0$. We would have (Fg, Fv) in $\mathcal{M}$ if Fv were in $L^2(\mu)$, but this need not be the case. So apply the argument in (ii) above to the vector (g, Iv) instead of (g, v). This argument yields a sequence of polynomials p_n such that $(p_n g, p_n Iv)$ tends in norm to (Fg, FIv), so

$$\int |p_n - F|^2 |g|^2 d\sigma + \int |p_n - F|^2 |Iv|^2 d\mu \to 0.$$

But then

$$\int |Ip_n - IF|^2 |g|^2 d\sigma + \int |p_n - F|^2 |Iv|^2 d\mu \to 0,$$

so (fg, fv) is the limit of $(Ip_n g, Ip_n v)$ and is therefore in $\mathcal{M}$. □

An operator is called *reductive* if every invariant subspace is also a reducing subspace. For the operator N_μ, being reductive is equivalent to the polynomials being weak-star dense in $L^\infty(\mu)$ [3, VII.1.3]. When N_μ is reductive, we can give a complete description of the lattice of $S \oplus N_\mu$.

Theorem 2.4 *Let μ be a measure on $\overline{\mathbb{D}}$ such that N_μ is reductive, and let $A = S \oplus N_\mu$. Let $\mathcal{M}$ be an invariant subspace of A. Then there are a Borel set E, a μ-measurable function w which is zero on E a.e.(μ) and also zero a.e. (μ_2), and a function I that is either inner or identically zero so that*

$$\mathcal{M} \;=\; \begin{pmatrix} 0 \\ L^2(\mu|_E) \end{pmatrix} \oplus \left\{ \begin{pmatrix} g \\ wg \end{pmatrix} :\; g \in IH^2,\; gw \in L^2(\mu) \right\}$$

Moreover any such choice of E, w and I gives rise to an invariant subspace.

PROOF: By the remarks at the beginning of the section, we can assume that μ_2 is zero, as it can just contribute an L^2 summand. Let P_1 and P_2 be the projections onto the first and second summands of $H^2 \oplus L^2(\mu)$.

Case (a): There is no vector ξ in $\mathcal{M}$ for which $P_1\xi = 0$ and $P_2\xi \neq 0$.

The space $P_1\mathcal{M}$ is invariant for S, so is of the form $I\mathcal{L}$, where I is inner and $\mathcal{L}$ is dense in H^2.

Claim: if (g_1, v_1) and (g_2, v_2) are in $\mathcal{M}$, then $g_1 v_2 = g_2 v_1$ a.e. (μ).

Proof of claim: Let g_1 and g_2 have inner-outer factorings I_1F_1 and I_2F_2 respectively. For any positive log-integrable function p on the unit circle, let $O\{p\}$ denote the outer function that is positive at 0 and has modulus p on the boundary. Then by Theorem 2.3

$$\begin{pmatrix} I_1F_1 \\ v_1 \end{pmatrix} \in \mathcal{M} \Rightarrow \begin{pmatrix} I_2O\{\min(1, \frac{1}{|F_1|}) \min(1, |F_2|)\}I_1F_1 \\ I_2O\{\min(1, \frac{1}{|F_1|}) \min(1, |F_2|)\}v_1 \end{pmatrix} \in \mathcal{M}$$

and similarly

$$\begin{pmatrix} I_1O\{\min(1, \frac{1}{|F_2|}) \min(1, |F_1|)\}I_2F_2 \\ I_1O\{\min(1, \frac{1}{|F_2|}) \min(1, |F_1|)\}v_2 \end{pmatrix} \in \mathcal{M}.$$

Subtracting these two vectors, one gets a vector whose first component vanishes, and the assumption of case (a) means the second component must also vanish, which yields the claim.

Let $\{(g_i, v_i)\}$ be a countable dense set in $\mathcal{M}$. Define w to be zero on the set where all the v_i vanish, and define it to be g_i/v_i whenever there is some v_i that is essentially non-zero. By the claim, w is well-defined, and $(g_i, v_i) = (g_i, wg_i)$. Therefore

$$\mathcal{M} = \left\{ \begin{pmatrix} g \\ wg \end{pmatrix} : g \in I\mathcal{L} \right\}.$$

As $\mathcal{M}$ is closed, by Lemma 2.5 below

$$\mathcal{M} = \left\{ \begin{pmatrix} g \\ wg \end{pmatrix} : g \in IH^2, \int |g|^2|w|^2 d\mu < \infty \right\},$$

as desired.

Case (b): Let

$$\mathcal{K}_1 = P_2 \left\{ \begin{pmatrix} 0 \\ f \end{pmatrix} : \begin{pmatrix} 0 \\ f \end{pmatrix} \in \mathcal{M} \right\}.$$

Then $\mathcal{K}_1$ is closed and N_μ invariant, so by the assumption that N_μ is reductive, $\mathcal{K}_1 = L^2(\mu|_E)$ for some E. Moreover, $\mathcal{M} \ominus (0, \mathcal{K}_1)$ is A-invariant, and reduces to case (a). Finally, w has to vanish on E because any vector of the form (g, wg) must be orthogonal to any vector of the form $(0, wg|_E)$. $\square$

We needed one lemma in the proof:

Lemma 2.5 *Let I be inner, and $\mathcal{L}$ be a dense subspace of H^2. Let*

$$\mathcal{N} \;=\; \left\{ \begin{pmatrix} g \\ wg \end{pmatrix} \;:\; g \in I\mathcal{L} \right\},$$

and assume $(S \oplus N_\mu)\mathcal{N} \subseteq \mathcal{N}$. Let $\mathcal{M}$ be the closure of $\mathcal{N}$. Then

$$\mathcal{M} \;=\; \left\{ \begin{pmatrix} f \\ wf \end{pmatrix} \;:\; f \in IH^2,\; \int |f|^2 |w|^2 d\mu < \infty \right\}.$$

PROOF: The right-hand side is closed and contains $\mathcal{N}$, so it suffices to prove that for any f in IH^2 with wf in $L^2(\mu)$, we have (f, wf) in $\mathcal{M}$.

Case (a): The inner function I is 1 and f is outer.

As $\mathcal{L}$ is dense, there are functions g_n (with inner-outer factorings $I_n F_n$) that converge to f in H^2; the problem is that $\{wg_n\}$ need not be bounded in $L^2(\mu)$. But by Theorem 2.3, we can multiply $(I_n F_n, I_n F_n w)$ by f/F_n to get that $(I_n f, I_n f w)$ is in $\mathcal{M}$. Thus we only need to show that I_n tends to 1 weakly.

By passing to a subsequence, if necessary, we can assume that g_n converges to f almost everywhere. Notice first that as $\lim_{n\to\infty} I_n(0)F_n(0) \;=\; f(0)$, we get

$$\begin{aligned} \liminf_{n\to\infty} \int \log^+|F_n| - \log^-|F_n| \;&=\; \liminf_{n\to\infty} \log|F_n(0)| \\ &\geq\; \log|f(0)| \\ &=\; \int \log^+|f| - \log^-|f| \end{aligned} \tag{2.6}$$

Because g_n tends to f in H^2, the set $\{\log^+|F_n|\}$ is uniformly integrable, so $\int \log^+|F_n|$ tends to $\int \log^+|f|$. Moreover, by Fatou's lemma,

$$\liminf_{n\to\infty} \int \log^-|F_n| \;\geq\; \int \log^-|f|.$$

Combining these facts, we get that the inequality in (2.6) is actually an equality, which means that $\lim_{n\to\infty} I_n(0) = 1$. Replacing Lebesgue measure by harmonic measure for other points in the disk, the above argument yields that the functions I_n converge to 1 pointwise on $\mathbb{D}$. Therefore the set $\{(I_n f, I_n f w)\}$ has (f, wf) as a weak cluster point, and so this vector is in $\mathcal{M}$ as desired.

Case (b): The function f is not outer.

Choose g_n in $\mathcal{L}$ as above converging to $O\{|f|\}$. Let J be the inner factor of f. Then $(Jg_n, Jg_n w)$ is in $\mathcal{M}$, so can be multiplied by $O\{|f|\}\,/\,F_n$, which as above will converge weakly to (f, wf). $\square$

By taking orthogonal complements, Theorem 2.4 also describes the lattice of $S^* \oplus N_\mu^*$, but the description is a little messier. For ν a measure on the closed disk let $\nu\hat{}$ be its balayage to the boundary, *i.e.* that measure on $\mathbb{T}$ whose integral against any continuous harmonic function is the same as that of ν. Let P_{H^2} and P_{IH^2} be projections onto H^2 and IH^2, respectively. Then the previous theorem becomes:

Corollary 2.7 *Let μ be a measure on $\overline{\mathbb{D}}$ such that N_μ is reductive, and let $\mathcal{N}$ be an invariant subspace of $S^* \oplus N_\mu^*$. Then there is a set F of μ_1 measure zero, and E, w and I as in Theorem 2.4, so that*

$$\mathcal{N} \;=\; \begin{pmatrix} 0 \\ L^2(\mu|_F) \end{pmatrix} \oplus \left\{ \begin{pmatrix} f \\ v \end{pmatrix} : \; v = 0 \text{ on } E, \; P_{IH^2} f = -I P_{H^2}[\overline{w}\overline{I} v \mu]\hat{} \right\}$$

It is possible to have A-invariant subspaces of the form

$$\mathcal{M} \;=\; \{ \begin{pmatrix} f \\ wf \end{pmatrix} \; : \; f \in H^2, \; \int |f|^2 (d\sigma + |w|^2 d\mu) < \infty \}$$

that are not cyclic, as the following example shows.

Example 2.8 Let α_n be an infinite Blaschke sequence on the positive real axis. Let $\beta_n = \alpha_n + \varepsilon_n$ be another, where $\alpha_n < \beta_n < \alpha_{n+1}$, and the ε_n's will be chosen later. Let I_1 be the Blaschke product with zeroes at the $\{\alpha_n\}$'s, and I_2 be the Blaschke product with zeroes at the $\{\beta_n\}$'s. Let μ be Lebesgue measure on $[0,1]$, and let $w = \min(1/|I_1|, 1/|I_2|)$. Let $\rho(x,y)$ denote the pseudo-hyperbolic distance from x to y.

Then for any function f,

$$\int_{\alpha_n}^{\beta_n} |f|^2 w^2 dx \;\geq\; \inf_{x \in [\alpha_n, \beta_n]} |f(x)|^2 \varepsilon_n \frac{1}{\rho(\alpha_n, \beta_n)^2} \tag{2.9}$$

As $\rho(\alpha_n, \beta_n) \leq \varepsilon_n/(1-\beta_n^2)$, if fw is in L^2, equation (2.9) yields that for all n

$$\inf_{x \in [\alpha_n, \beta_n]} |f(x)|^2 \;\leq\; \frac{\varepsilon_n}{(1-\beta_n^2)^2} \, \|fw\|_{L^2}^2 \tag{2.10}$$

But if f is outer, there is a constant c (the number $4 \int \log^- |f|$ will do) such that

$$|f(x)|^2 \;\geq\; e^{-c\frac{1}{1-\beta_n}} \qquad \text{for all } x \in [\alpha_n, \beta_n].$$

So if we choose

$$\varepsilon_n \;\leq\; (1-\alpha_{n+1}^2)^2 e^{-n\frac{1}{1-\alpha_{n+1}}},$$

then no outer function f can satisfy (2.10). Therefore if $\mathcal{M}$ were cyclic with cyclic vector (g,v), then by Theorem 2.3, the non-trivial inner factor of g would have to divide the first component of every vector in $\mathcal{M}$. But $\mathcal{M}$ contains $(I_1, I_1 w)$ and $(I_2, I_2 w)$ and I_1 and I_2 have greatest common divisor 1. Therefore $\mathcal{M}$ is not cyclic. $\square$

The space $\mathcal{M}$ constructed in Example 2.8 is 2-cyclic.

Question 2.11 *Is every invariant subspace of an operator of the form $S \oplus N_\mu$ at most 2-cyclic?*

Although we do not know as much about the lattice of A if N_μ is not reductive, we can prove that any restriction of $S^* \oplus N_\mu$ to an invariant subspace of dimension greater than one does itself have an invariant subspace.

Theorem 2.12 *Let N be a normal contraction, and let $\mathcal{N}$ be an invariant subspace of $S^* \oplus N$ of dimension greater than 1, and let $T = S^* \oplus N\,|_{\mathcal{N}}$. Then T has an invariant subspace.*

PROOF: We can assume N is star-cyclic, so writing $N = N_\mu^*$, we are in the situation we have been considering throughout this section. Let $\mathcal{M} = \mathcal{N}^\perp$. It suffices to show that $\mathcal{M}$ has a proper superspace that is A-invariant. If N_μ is reductive, the result follows immediately from Theorem 2.4 (just make E or I a little bigger). If N_μ is not reductive, then the conclusion of Theorem 2.4 would only fail to be true if the space $\mathcal{K}_1$ were invariant for N_μ but not for N_μ^*.

If the codimension of $\mathcal{K}_1$ in $L^2(\mu)$ is larger than 1, then N_μ^* restricted to $L^2(\mu) \ominus \mathcal{K}_1$ is subnormal, so has an invariant subspace $\mathcal{L}_1$. Add to $\mathcal{M}$ everything of the form

$$\left\{ \begin{pmatrix} 0 \\ f \end{pmatrix} : \ f \in L^2(\mu), \ f \perp \mathcal{L}_1 \right\}.$$

This superspace of $\mathcal{M}$ will be A invariant, and proper, since P_2 of it is just $\mathcal{L}_1^\perp$.

If the codimension of $\mathcal{K}_1$ is exactly 1, then everything in the first slot of $\mathcal{M}$ must have a common inner factor, because $\mathcal{M}$ has codimension larger than 1. Therefore enlarging $\mathcal{K}_1$ to all of $L^2(\mu)$ will still result in a proper subspace of $H^2 \oplus L^2(\mu)$. $\square$

3 The case $\tilde{T} = S^{*(n)}$

The general operator T that dominates X and satisfies assumptions (A1 - A3) is, by Proposition 1.2, the restriction of some operator of the form $\tilde{T} \oplus X$ to a cyclic invariant subspace $\mathcal{N}$, with cyclic vector (ξ_0, η_0) say. Moreover, by (A2), the vector η_0 is cyclic for X.

Let the space on which $\tilde{T}$ acts be called $\mathcal{H}$, and the space on which X acts be called $\mathcal{K}$. If we let $\mathcal{M}$ denote $\mathcal{N}^\perp$, then $\mathcal{M}$ cannot contain a vector with $\mathcal{H}$-component zero and

$\mathcal{K}$-component non-zero, because of the cyclicity of η_0. Therefore $\mathcal{M}$ can be represented as

$$\mathcal{M} \ = \ \left\{ \begin{pmatrix} \xi \\ \Lambda\xi \end{pmatrix} \right\}$$

where Λ is a closed, linear, densely defined operator from $\mathcal{H}$ to $\mathcal{K}$, with kernel the orthogonal complement of the $\tilde{T}^*$-invariant subspace generated by ξ_0. Moreover $\Lambda\tilde{T}^* \ = \ X^*\Lambda$.

Finding an invariant subspace for $\tilde{T} \oplus X|_{\mathcal{N}}$ is equivalent to finding a proper superspace of $\mathcal{M}$ invariant under $\tilde{T}^* \oplus X^*$. One way of doing this would be to find an invariant subspace of X^*, $\mathcal{K}_0$ say, and to look at

$$\mathcal{M}_1 \ = \ cl \left\{ \mathcal{M} + \begin{pmatrix} 0 \\ \mathcal{K}_0 \end{pmatrix} \right\}.$$

The only problem is that $\mathcal{M}_1$ might be all of $\mathcal{H} \oplus \mathcal{K}$. This will not occur if and only if there is a non-zero vector (ζ, η) in $\mathcal{M}^\perp = \mathcal{N}$ for which η is perpendicular to $\mathcal{K}_0$.

If we knew η, we could recover ζ because the fact that (ζ, η) is in $\mathcal{M}^\perp$ means

$$\langle \zeta, \xi \rangle \ = \ - \ \langle \eta, \Lambda\xi \rangle. \tag{3.1}$$

So starting with η in $\mathcal{K}$, we could find a ζ in $\mathcal{H}$ with (ζ, η) in $\mathcal{N}$, provided that

$$|\langle \Lambda\xi, \eta \rangle| \ \leq \ c\|\xi\|_{\mathcal{H}} \qquad \forall \xi \in \mathcal{H}, \tag{3.2}$$

because ζ would then be the unique vector in $\mathcal{H}$ whose inner products satisfy (3.1) as ξ ranges over a dense set. We could let $\mathcal{K}_0$ be the orthogonal complement, in $\mathcal{K}$, of the X-invariant subspace generated by η. The conclusion we reach is thus:

Proposition 3.3 *With the above notation, a sufficient condition for T to have an invariant subspace is the existence of a non-zero vector η in $\mathcal{K}$ that is not cyclic for X and that satisfies (3.2).*

Let us now specialize to the case where $X = N_\mu^*$ and $\tilde{T} = S^{*(n)}$ for some finite n. As in the previous section, we shall assume that μ_2 is zero, as the same argument we used in Section 2 shows that its contribution is uninteresting.

Let the domain of Λ be $\mathcal{L}$, where $\mathcal{L}$ is a dense subspace of $H^{2(n)}$. The map Λ has a nice form:

Proposition 3.4 *There are functions $u_1, \ldots, u_n$ so that the map $\Lambda : H^{2(n)} \supseteq \mathcal{L} \to L^2(\mu)$ sends $(f_1, \ldots, f_n)$ to $\sum_{i=1}^n u_i f_i$.*

PROOF: The idea of the proof is that, if n were 2 and $\mathcal{L}$ contained non-vanishing functions $(f,0)$ and $(0,g)$, then letting

$$u_1 = \frac{1}{f}\Lambda\begin{pmatrix} f \\ 0 \end{pmatrix}, \quad u_2 = \frac{1}{g}\Lambda\begin{pmatrix} 0 \\ g \end{pmatrix},$$

would work, because Λ intertwines $S^{(n)}$ and N_μ.

In general, let $(f_1, f_2, \ldots, f_n)$ and $(g_1, g_2, \ldots, g_n)$ be elements of $\mathcal{L}$ with $f_1 g_2 \neq f_2 g_1$. Let f_1 have inner-outer factoring IF, and g_1 have factoring JG.

As $\mathcal{L}$ is invariant under multiplication by H^∞ functions (by the same argument as in the proof of Theorem 2.3), we get that $\mathcal{L}$ contains

$$\begin{pmatrix} f_1 J\, O\{\min(1,\frac{1}{|F|})\min(1,|G|)\} \\ f_2 J\, O\{\min(1,\frac{1}{|F|})\min(1,|G|)\} \\ \vdots \\ f_n J\, O\{\min(1,\frac{1}{|F|})\min(1,|G|)\} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} g_1 I\, O\{\min(1,\frac{1}{|G|})\min(1,|F|)\} \\ g_2 I\, O\{\min(1,\frac{1}{|G|})\min(1,|F|)\} \\ \vdots \\ f_n I\, O\{\min(1,\frac{1}{|G|})\min(1,|F|)\} \end{pmatrix}. \tag{3.5}$$

The difference of these two vectors is a vector whose first component is zero, and whose second component is non-zero. By iterating this procedure, we can find n vectors $\xi_1, \ldots, \xi_n$ in $\mathcal{L}$ such that for each i, the i^{th} component of ξ_i is non-zero, and the other components are all zero. Define

$$u_i = \frac{1}{\xi_i}\Lambda(\xi_i). \tag{3.6}$$

The only problem with the definition is at atoms of μ where ξ_i vanishes.

Claim: for any choice of λ in $\mathbb{D}$, we can choose $\{\xi_i\}_{i=1}^n$ in $\mathcal{L}$ so that, for each i, all but the i^{th} component of ξ_i is identically zero, and the i^{th} component does not vanish at λ.

We shall let superscripts denote the components of a vector, so ξ_i^j is the j^{th} component function of the vector ξ_i. Without loss of generality, we can assume $i = n$.

The claim will follow from the following claim:

Claim: If $j < n$, there exist $(n-j)$ vectors $\{\eta_k\}_{k=j+1}^n$ such that $\eta_k^l \equiv 0$ for $1 \leq l \leq j$ and $\eta_k^l(\lambda) = \delta_{lk}$.

The proof is by induction on j. When $j = 0$, the claim is true because the density of $\mathcal{L}$ means that

$$\left\{ \begin{pmatrix} f_1(\lambda) \\ \vdots \\ f_n(\lambda) \end{pmatrix} : \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \in \mathcal{L} \right\}$$

is all of $\mathbb{C}^n$.

Assume now the claim has been proved for j, giving vectors $\{\eta_k\}_{k=j+1}^n$, and we want to prove the claim for $j+1<n$. Applying the construction in (3.5) to the pairs (η_j, η_l) as l ranges from $j+1$ to n, we get $n-j-1$ vectors such that the k^{th} component of the l^{th} vector is an outer function times $\eta_j^j\eta_l^k - \eta_j^k\eta_l^j$. After multiplying by a constant to take care of the outer factor, we get that these new vectors satisfy the claim.

Thus at every point of $\mathbb{D}$ we can choose ξ_i so that u_i are defined by (3.6).

Next we must show that they are well-defined, so suppose ξ and η are two vectors in $\mathcal{L}$, for which all but the i^{th} components vanish, and let $\Lambda(\xi) = v_1$, $\Lambda(\eta) = v_2$. Let $\xi = IF$ and $\eta = JG$ be the inner-outer factorings.

Then as

$$\xi J\, O\{\min(1, \frac{1}{|F|}) \min(1, |G|)\} \quad = \quad \eta I\, O\{\min(1, \frac{1}{|G|}) \min(1, |F|)\}$$

we get that

$$v_1 J\, O\{\min(1, \frac{1}{|F|}) \min(1, |G|)\} \quad = \quad v_2 I\, O\{\min(1, \frac{1}{|G|}) \min(1, |F|)\}$$

and hence

$$v_1\eta \quad = \quad v_2\xi,$$

so formula (3.6) is well-defined.

Finally we must prove that the formula in the statement of the proposition holds true for a general $(f_1, \ldots, f_n)$ in $\mathcal{L}$. Pick $\xi_1, \ldots, \xi_n$ as in the claim. Let J be the product of the inner factors of each ξ_i. Then there is an outer function F so that

$$\begin{pmatrix} JF & & \\ & \ddots & \\ & & JF \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = \sum_{i=1}^n \left(\frac{JFf_i}{\xi_i}\right) \xi_i.$$

Therefore

$$\begin{aligned} JF\, \Lambda \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} &= \Lambda\left(\begin{pmatrix} JF & & \\ & \ddots & \\ & & JF \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \right) \\ &= JF \sum_{i=1}^n u_i f_i \end{aligned}$$

If we divide both sides by JF, we get that the theorem holds except possibly at the zeroes of J. But let λ be such a zero. Then we can repeat the above argument with ξ_i chosen not to vanish at λ, and thus we get that $\Lambda(f_1, \ldots, f_n) = \sum_{i=1}^n u_i f_i$ almost everywhere (μ), as desired. □

It is now easy to prove the main result of this section.

Theorem 3.7 *Let N be a contractive normal operator, and let $\mathcal{N}$ be an invariant subspace of $S^{*(n)} \oplus N$ of dimension greater than 1. Let $T = S^{*(n)} \oplus N|_{\mathcal{N}}$. Then T has an invariant subspace.*

PROOF: We can assume that $\mathcal{N}$ is cyclic, and letting $N_\mu = N^*$ we are in the situation considered above. If Λ were zero, $\mathcal{N}$ would be contained in $H^{2^{(n)}}$, and the theorem would follow from the factoring of matrix-valued inner functions [6, Theorem 16].

If Λ is not zero, we can apply Propositions 3.3 and 3.4 to conclude that we need only find a vector η in $L^2(\mu)$, such that

$$(i) \qquad 0 < \mu\{\eta \neq 0\} < \|\mu\|$$

and

$$(ii) \qquad |\sum_{i=1}^{n} \langle u_i f_i, \eta \rangle_{L^2(\mu)}| \leq c \left\| \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \right\|_{H^{2(n)}}$$

The measure μ cannot consist of just one atom, because otherwise $\mathcal{M}$ would be of codimension 1. Therefore there is a set E such that $0 < \mu(E) < \|\mu\|$. If μ has mass on the open unit disk, choose E to be a subset of some set of the form

$$\{z : \max(|u_1(z)|, \ldots, |u_n(z)|) \frac{1}{\sqrt{1 - |z|^2}} \leq M\},$$

for some constant M; if μ lives solely on the circle, choose E to be a subset of

$$\{z : \max(|u_1(z)|, \ldots, |u_n(z)|) \frac{d\mu}{d\sigma}(z) \leq M\}.$$

In either case, if η is chosen to be the characteristic function of E, both (i) and (ii) will be satisfied. $\Box$

4 Examples

Throughout this section, μ will be Lebesgue measure on the unit interval $[0, 1]$, and $N = N_\mu$ will be the (self-adjoint) operator of multiplication by x on $L^2(\mu)$. We shall look at several different sorts of operator that dominate N.

Example 4.1 Let $M_n = (n+2)!(\log(n+2))^n$, and define a Hilbert space $\mathcal{H}$ by

$$\mathcal{H} \;=\; \left\{ f \in C^\infty[0,1] \,:\, \sum_{n=0}^{\infty} \|f^{(n)}\|^2 \frac{1}{M_n^2} \;<\; \infty \right\}.$$

Let T be multiplication by the independent variable on $\mathcal{H}$. Then:
(a) The spectrum of T is $[0,1]$.
(b) Evaluation at λ is continuous for all λ in $[0,1]$.
(c) $\mathcal{H}$ contains no functions that vanish on a set of positive measure.
(d) If $L : \mathcal{H} \to L^2(\mu)$ is the inclusion map, then $LT = NL$ and the range of L contains no non-cyclic vectors of N (except for 0).

Proof: (a): By Leibniz's formula,

$$\begin{aligned} \|(xf)^{(n)}\|_{L^2} &= \|xf^{(n)} + nf^{(n-1)}\|_{L^2} \\ &\leq \|f^{(n)}\|_{L^2} + n\|f^{(n-1)}\|_{L^2}, \end{aligned}$$

so

$$\begin{aligned} \|Tf\|^2_{\mathcal{H}} &\leq 2\sum_{n=0}^{\infty} \|f^{(n)}\|^2_{L^2} \left[\frac{1}{M_n^2} + \frac{(n+1)^2}{M_{n+1}^2} \right] \\ &\leq 4\|f\|^2_{\mathcal{H}}. \end{aligned}$$

Therefore T is bounded.

To prove (a), we shall estimate the norm of $f(x)/(\lambda - x)$ for λ not in $[0,1]$. Let $\varepsilon = dist(\lambda, [0,1])$. Then

$$\begin{aligned} \|D^n \left(\frac{f(x)}{\lambda - x} \right) \|^2_{L^2} &= \| \sum_{j=0}^{n} \binom{n}{j} f^{(j)}(x)(n-j+1)! \,\frac{1}{(\lambda - x)^{n-j+1}} \|^2_{L^2} \\ &\leq \frac{\pi^2}{6} \sum_{j=0}^{n} (n-j+1)^4 \left(\frac{n!}{j!} \right)^2 \varepsilon^{-2(n-j+1)} \|f^{(j)}\|^2_{L^2} \end{aligned}$$

where we used the fact that

$$|\sum_{j=0}^{n} a_j|^2 \leq \frac{\pi^2}{6} \sum_{j=0}^{n} (n-j+1)^2 |a_j|^2$$

(this follows from the Cauchy-Schwarz inequality). So

$$\begin{aligned} \|\frac{f(x)}{\lambda - x}\|^2_{\mathcal{H}} &\leq \frac{\pi^2}{6} \sum_{n=0}^{\infty} \frac{1}{M_n^2} \sum_{j=0}^{n} (n-j+1)^4 \left(\frac{n!}{j!} \right)^2 \varepsilon^{-2(n-j+1)} \|f^{(j)}\|^2_{L^2} \\ &\leq \frac{\pi^2}{6\varepsilon^2} \sum_{j=0}^{\infty} \frac{1}{M_j^2} \|f^{(j)}\|^2_{L^2} \sum_{n=j}^{\infty} (n-j+1)^4 \frac{n!^2}{M_n^2} \frac{M_j^2}{j!^2} \frac{\varepsilon^{2j}}{\varepsilon^{2n}} \end{aligned}$$

Substituting in for M_n and the second occurrence of M_j, this becomes

$$\|\frac{f(x)}{\lambda - x}\|_{\mathcal{H}}^2 \leq \frac{\pi^2}{6\varepsilon^2} \sum_{j=0}^{\infty} \frac{1}{M_j^2} \|f^{(j)}\|_{L^2}^2 \sum_{n=j}^{\infty} (n-j+1)^4 \frac{(j+1)^2}{(n+1)^2} \frac{[\varepsilon \log(j+2)]^{2j}}{[\varepsilon \log(n+2)]^{2n}} \tag{4.2}$$

Now

$$\begin{aligned} \sum_{n=j}^{\infty} (n-j+1)^4 \frac{(j+1)^2}{(n+1)^2} \frac{[\varepsilon \log(j+2)]^{2j}}{[\varepsilon \log(n+2)]^{2n}} \quad &\leq \quad \sum_{n=j}^{\infty} (n-j+1)^4 \frac{1}{[\varepsilon \log(n+2)]^{2(n-j)}} \\ &\leq \quad \sum_{n=1}^{\infty} n^4 \frac{1}{[\varepsilon \log(n+1)]^{2(n-1)}} \end{aligned}$$

and this sum is finite. Plugging this last inequality back into (4.2), we get that there is a constant c depending only on ε such that

$$\|\frac{f(x)}{\lambda - x}\|_{\mathcal{H}} \ \leq \ c \, \|f\|_{\mathcal{H}},$$

so $T - \lambda$ is invertible.

(b) This is obvious - indeed, evaluating any derivative of f at any point of $[0,1]$ is continuous, because the norm gives control over all the derivatives.

(c) In fact a much stronger conclusion holds: by the Denjoy-Carleman theorem (*e.g.* [7, V.2], the space $\mathcal{H}$ is quasi-analytic: no non-zero function in $\mathcal{H}$ can vanish to infinite order at any point in $[0,1]$. Therefore the zero set of any function in $\mathcal{H}$ must be finite.

(d) This follows from (c). $\square$

If Z is any finite subset of $[0,1]$, with multiplicities allowed, the set of all functions in $\mathcal{H}$ that vanish on Z to the prescribed orders will form an invariant subspace of T of finite codimension.

Question 4.3 *Is every invariant subspace of T of this form?*

Let us list four more examples of operators that dominate N, and then discuss them.

Example 4.4 Let $T = S \oplus N$ restricted to $[(1,1)]$. Then T is similar to S.

Example 4.5 Let ν be harmonic measure on the boundary of the unit square $[0,1] \times [0,1]$. Let $\mathcal{H}$ be the closure of the polynomials in $L^2(\nu)$, and T be multiplication by z on $\mathcal{H}$. Then T is just $\phi(S)$, where ϕ is the Riemann map from the disk to the square, and the invariant subspace lattice of T is described by Beurling's theorem.

Example 4.6 Let ξ be a cyclic vector for S^*, and let T be $S^* \oplus N$ restricted to $[(\xi, 1)]$. The lattice of T is described by Corollary 2.7, modulo identifying w and I. For the purpose of this example, let $w = 1$ and $I = 1$, which corresponds to choosing $\xi(z) = -\frac{1}{z}\log(1-z)$. Then $[(\xi, 1)] = \{(f, v) : P(v\mu)\hat{} = -f\}$. Both F and E are empty, and subspaces of $[(\xi, 1)]$ correspond to different sets E.

Example 4.7 Let $T = N$.

Notice that there are three qualitatively different behaviors for the invariant subspace lattice of T. In Examples 4.6 and 4.7, the invariant subspaces correspond precisely to functions vanishing on sets of positive μ-measure. Indeed there is a lattice anti-isomorphism between the invariant subspace lattice of T and the lattice of Borel subsets of $[0, 1]$ modulo μ-null sets.

In Examples 4.4 and 4.5 there is analytic structure, and one has analytic bounded point evaluations for T. In the former case, $[0, 1)$ is contained in the set of analytic bounded point evaluations, whereas in the latter there is only "one-sided" analyticity.

Finally, in Example 4.1, there is no analytic structure, yet there is a quasi-analytic structure that allows one to find bounded point evaluations and hence get invariant subspaces; but even though all the action is on the interval $[0, 1]$, these subspaces correspond to μ-null sets.

Given these three forms of behavior, it is hard to see in general how an operator that dominates N "inherits" an invariant subspace from it.

5 Miscellaneous Results

If both T and T^* dominate normals, it is tantamount to saying that T is quasi-similar to a normal, and indeed under hypotheses (A1 - A2), it is.

Theorem 5.1 *Suppose T and T^* both dominate normal operators. Then T has an invariant subspace.*

PROOF: By hypothesis, there are operators L_1 and L_2 and normal operators N_1 and N_2 so that $L_1 T = N_1 L_1$ and $L_2 T^* = N_2 L_2$. By the discussion in the introduction, we may assume that both L_1 and L_2 have no kernel and dense range, as otherwise T already has an invariant subspace.

As

$$N_1(L_1 L_2^*) \;=\; (L_1 L_2^*) N_2^*$$

the Fuglede-Putnam theorem (see *e.g.* [2]) gives

$$N_1^*(L_1 L_2^*) \;=\; (L_1 L_2^*) N_2.$$

So N_1 and N_2^* are quasi-similar, so T is quasi-similar to N_1, and hence has an invariant subspace. $\square$

Fix an operator T on $\mathcal{H}$. For any vector u in $\mathcal{H}$, let

$$E_u \;=\; \{\mu \,:\, \int |p|^2 d\mu \leq \|p(T)u\|^2 \text{ for all polynomials } p\}.$$

Define

$$\rho(u) \;=\; \sup\{\|\mu\|^{\frac{1}{2}} \,:\, \mu \in E_u\}.$$

For u in $\mathcal{H}$, and for each μ in E_u, define $L_{u,\mu}$ from the cyclic subspace generated by u to $P^2(\mu)$ (the closure of the polynomials in $L^2(\mu)$), to be that continuous operator whose action on a dense set is given by

$$L_{u,\mu} : \; p(T)u \mapsto p.$$

Let S_μ be multiplication by z on $P^2(\mu)$. Then $L_{u,\mu}T \;=\; S_\mu L_{u,\mu}$. If q is a rational function with poles off the spectrum of $T|_{[u]}$, then $L_{u,\mu}(q(T)u) \;=\; q$.

If T has no invariant subspaces, then ρ is a semi-norm, as a consequence of the following proposition.

Proposition 5.2 *Suppose T is transitive. Fix $u \neq 0$ in $\mathcal{H}$. Then, for all v in $\mathcal{H}$,*

$$E_v \;=\; \{|L_{u,\mu}(v)|^2 \mu \,:\, \mu \in E_u\}.$$

PROOF: Fix some $v \neq 0$.
($\supseteq$): There is a sequence of polynomials p_n so that $v = \lim_{n\to\infty} p_n(T)u$. For any μ in E_u, the sequence p_n is Cauchy in $L^2(\mu)$, so converges to some function, f say. As, for any polynomial p,

$$\begin{aligned}
\int |p|^2 |f|^2 d\mu &= \lim_{n\to\infty} \int |p|^2 |p_n|^2 d\mu \\
&\leq \lim_{n\to\infty} \|p(T)p_n(T)u\|^2 \\
&= \|p(T)v\|^2,
\end{aligned}$$

we get that $|L_{u,\mu}(v)|^2\mu$ is in E_v.
($\subseteq$): Fix $\nu \in E_v$. Let $\mu = |L_{v,\nu}(u)|^2\nu$, which is in E_u by the first part of the proof.
Claim: $|L_{u,\mu}(v)|^2\mu = \nu$.
Proof of claim: There are polynomials q_n so that $q_n(T)v$ converges to u. Let $g = L_{v,\nu}(u)$, so $\mu = |g|^2\nu$. Then

$$|L_{u,\mu}(v)|^2\mu = \lim_{k\to\infty} |p_k|^2(\lim_{n\to\infty} |q_n|^2)\nu.$$

Fix $\varepsilon > 0$. Choose K_0 so that

$$\|p_k(T)u - v\| \quad < \quad \frac{\varepsilon}{2} \text{ if } k \geq K_0,$$

and choose $N_0 = N_0(k)$ so that

$$\|q_n(T)v - u\| \quad < \quad \frac{\varepsilon}{2(1 + \|p_k(T)\|)} \text{if } n \geq N_0.$$

Then for $k \geq K_0$ and $n \geq N_0(k)$,

$$\begin{aligned}\left[\int |p_k q_n - 1|^2 d\nu\right]^{\frac{1}{2}} &\leq \left[\int |p_k g - 1|^2 d\nu\right]^{\frac{1}{2}} + \left[\int |p_k|^2 |q_n - g|^2 d\nu\right]^{\frac{1}{2}} \\ &< \varepsilon.\end{aligned}$$

As ε is arbitrary, it follows that $\lim_{k\to\infty}(\lim_{n\to\infty} p_k q_n) = 1$, so $|L_{u,\mu}(v)|^2\mu = \nu$, as desired.
□

Corollary 5.3 *If T is intransitive, $\rho(u + v) \leq \rho(u) + \rho(v)$.*

PROOF: By Proposition 5.2,

$$\begin{aligned}\rho(u+v) &= \sup\left\{\left[\int |L_{u,\mu}(u+v)|^2 d\mu\right]^{\frac{1}{2}} \ : \ \mu \in E_u\right\} \\ &\leq \sup\left\{\left[\int |L_{u,\mu}(u)|^2 d\mu\right]^{\frac{1}{2}} + \left[\int |L_{u,\mu}(v)|^2 d\mu\right]^{\frac{1}{2}} \ : \ \mu \in E_u\right\} \\ &\leq \rho(u) + \rho(v)\end{aligned}$$

□

Lemma 5.4 *If μ is in E_u, then the support of μ lies in the spectrum of $T|_{[u]}$.*

PROOF: Let $R = T|_{[u]}$, and let U be an open disk in the resolvent of R. For each λ in U

$$L_{u,\mu}((R - \lambda)^{-1}u) = \frac{1}{z - \lambda}$$

is in $L^2(\mu)$, and these functions have uniformly bounded norms. Therefore by Tonelli's theorem, $\mu(U) = 0$. □

Finally we show that if ρ is actually a complete norm, then T has an invariant subspace.

Theorem 5.5 *Suppose ρ is a complete norm on $\mathcal{H}$. Then T has an invariant subspace.*

PROOF: From the open mapping theorem, there must exist a constant $c > 0$ so that

$$c\|v\|_{\mathcal{H}} \ \leq \ \rho(v) \ \leq \ \|v\|_{\mathcal{H}}$$

for all vectors v in $\mathcal{H}$. Assume T is transitive, and fix some rational function q with poles off the spectrum of T.

By Proposition 5.2, for any vector $u \in \mathcal{H}$,

$$\begin{aligned} \rho(q(T)u)^2 &= \sup\{\int |L_{u,\mu}(q(T)u)|^2 d\mu \, : \, \mu \in E_u\} \\ &\leq \sup_{\mu \in E_u} \|q\|^2_{\infty, suppt(\mu)} \rho(u)^2. \end{aligned}$$

Let K be the closure of the union of the supports of all measures in E_u. Then we just showed that

$$\|q(T)\|_{(\mathcal{H},\rho)\to(\mathcal{H},\rho)} \leq \sup\{|q(z)| \, : \, z \in K\}.$$

As the ρ-norm is similar to the original norm, we have that K is an M-spectral set for T (with $M = 1/c$).

As K is contained in the spectrum of T by Lemma 5.4, it follows that the spectrum of T is an M-spectral set for T, so by a theorem of Stampfli [11] T would have to have an invariant subspace after all. □

References

[1] J. Agler. An abstract approach to model theory. In J.B. Conway and B.B. Morrel, editors, *Surveys of some recent results in Operator Theory, Vol. II*, pages 1–24. Longman, Harlow, 1988.

[2] J.B. Conway. *A course in Functional Analysis.* Springer-Verlag, New York, 1985.

[3] J.B. Conway. *The Theory of Subnormal Operators.* American Mathematical Society, Providence, 1991.

[4] T.W. Gamelin. *Uniform Algebras.* Chelsea, New York, 1984.

[5] John B. Garnett. *Bounded Analytic Functions.* Academic Press, New York, 1981.

[6] Henry Helson. *Lectures on Invariant Subspaces.* Academic Press, New York, 1964.

[7] Y. Katznelson. *Harmonic Analysis.* Dover, New York, 1968.

[8] N. K. Nikol'skiĭ. *Treatise on the Shift Operator: Spectral Function Theory*, volume 273 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, Berlin, 1985.

[9] R.F. Olin and J.E. Thomson. Algebras of subnormal operators. *J. Funct. Anal.*, 37:271–301, 1980.

[10] H.L. Royden. *Real Analysis.* Macmillan, New York, Third edition, 1988.

[11] J.G. Stampfli. An extension of Scott Browns invariant subspace theorem: k-spectral sets. *J. Operator Theory*, 3(1):3–21, 1980.

[12] J.E. Thomson. Approximation in the mean by polynomials. *Annals of Math.*, 133:477–507, 1991.