Proceedings of the Project Review, Geo-Mathematical Imaging Group (Purdue University, West Lafayette IN), Vol. 1 (2010) pp. 1-8.

A COMPLETE AAK-THEOREM FOR WEIGHTED SEQUENCE SPACES

MARCUS CARLSSON*

Abstract. We give a full extension of (one version of) the celebrated "AAK-theorem" on Hankel operators, to the case of weighted l^2-spaces with increasing weights. This theorem was conjectured in [3], and it improves earlier work by S. Treil and A. Volberg, [8]. We also show that the corresponding extension of the classical formulation of the "AAK-theorem" fails, and show that this is a consequence of the failure of a sharp version of Nehari's theorem in this setting. We also note that a weaker version of Nehari's theorem was proven in [8], and hence this theorem can not be improved.

1. Introduction. Let S denote the unilateral shift operator on $l^2 = l^2(\mathbb{N})$, i.e. $S((x_0, x_1, \ldots)) = (0, x_0, x_1, \ldots)$. Let $(e_m)_{m=0}^{\infty}$ be the standard basis for $l^2(\mathbb{N})$, i.e. $e_0 = (1, 0, 0, \ldots)$ and $e_m = S^m e_0$. Recall that an operator $\Gamma : l^2 \to l^2$ is Hankel if it satisfies

$$\Gamma S = S^* \Gamma,$$

and that S^* is the backward shift. This definition is equivalent to demanding that the matrix representation of Γ in the standard basis $(e_m)_{m=0}^{\infty}$ looks like a Hankel matrix

$$\Gamma \text{ " = " } \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots \\ \gamma_1 & \gamma_2 & \gamma_3 & \cdots \\ \gamma_2 & \gamma_3 & \gamma_4 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}, \quad \gamma_n \in \mathbb{C}. \tag{1.1}$$

Given a Hilbert space X we let $\mathcal{L}(X)$ denote the set of bounded operators on X. We define "generalized Hankel operators" as follows.

Definition 1.1. *Let X_1 and X_2 be Hilbert spaces and let $S \in \mathcal{L}(X_1)$ and $B \in \mathcal{L}(X_2)$ be given operators. A bounded operator $\Gamma : X_1 \to X_2$ will be called Hankel (with respect to S and B) if it satisfies*

$$\Gamma S = B\Gamma. \tag{1.2}$$

It is easy to see that this definition is equivalent with the one in [5], (Vol 1, Part B, Sec 1.7), and that it is slightly more general than the one used by S. Treil and A. Volberg in [8].

Example 1.2. *As an example, let $w = (w_m)_{m=0}^{\infty}$ be any positive sequence and define l^2_w as the set of sequences x that are finite with respect to the norm*

$$\|x\|^2 = \sum_{m=0}^{\infty} |x_m|^2 w_m.$$

Let w and v be two such sequences, let S be the shift on l^2_w and let B be the backward shift on l^2_v. Then $\Gamma : l^2_w \to l^2_v$ is Hankel if and only if its matrix representation (in the standard bases $(e_m)_{m=0}^{\infty}$ in l^2_w and l^2_v respectively) looks like (1.1). A concrete example would be to take $w(m) = v(m) = m+1$, in which case $l^2_w = l^2_v$ correspond to the Dirichlet space via the Fourier transform. For the unweighted space, (i.e. when $w(m) = 1$ for all m), we will continue to use the notation l^2.

We introduce more notation. Let $\Gamma : X_1, \to X_2$ be any bounded operator and recall that its singular values $\sigma_0, \sigma_1, \ldots$ are defined as

$$\sigma_n = \inf\{\|\Gamma|_{\mathcal{M}}\| : \mathcal{M} \leq X \text{ and } \mathsf{codim}\ \mathcal{M} = n\} = \inf\{\|\Gamma - K\| : \mathsf{Rank}\ K = n\}, \tag{1.3}$$

*Department of Mathematics, Purdue University, West Lafayette, IN, 47907.

where $\mathcal{M} \leq X$ means that $\mathcal{M}$ is a subspace and $\Gamma|_{\mathcal{M}}$ denotes the restriction of Γ to $\mathcal{M}$. A vector $u_n \in X$ will be called a σ_n-singular vector if $\|u_n\| = 1$ and

$$\sigma_n^2 u_n = \Gamma^* \Gamma u_n.$$

If $\sigma_n > \sigma_n > \sigma_{n+1}$ then u_n is unique up to multiplication by numbers, so by slight abuse of language we will refer to u_n as the singular vector for σ_n. (A very simple example of such vectors is computed in Section 3). Recall the celebrated result by Adamyan, Arov and Krein [1], known as the AAK-theorem and usually stated as follows.

THEOREM 1.3. *(AAK) Let $\Gamma : l^2 \to l^2$ be a Hankel operator and let σ_n be its n'th singular value. Then there is a rank n Hankel operator K such that*

$$\sigma_n = \|\Gamma - K\|.$$

However, the theorem is actually much stronger because its proof provides a way of actually calculating the best rank n Hankel approximation. This in turn is related to the curious fact that the Fourier series defined by the n:th singular vector has precisely n zeroes in the unit disc, (assuming that $\sigma_{n+1} < \sigma_n < \sigma_{n-1}$). We outline this in more detail below.

It is easy to see that a rank 1 Hankel operator necessarily has the following form

$$\Gamma_{(z_0)} = \begin{pmatrix} 1 & z_0 & z_0^2 & \cdots \\ z_0 & z_0^2 & z_0^3 & \cdots \\ z_0^2 & z_0^3 & z_0^4 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}, \quad |z_0| < 1, \tag{1.4}$$

but it is not true that any finite rank Hankel operator is a sum of such. In fact, a "symbol" for the above rank one Hankel operator is easily seen to be $(1 - z_0\overline{z})^{-1}$ and in general, any rank n Hankel operator has a symbol of the form $r(\overline{z})$ where r is a rational function with degree n and all poles lie in $\{z \in \mathbb{C} : |z| > 1\}$, (see e.g. [6]). In terms of applications, the power of the AAK-theorem comes from the fact that the location of these poles can be easily calculated using the singular vectors. For simplicity, let us assume that σ_n is distinct and denote the corresponding singular vector by u_n. Let $\mathcal{F}^{-1}(u_n) = \check{u}_n$ denote the analytic function in the unit disc defined by $\check{u}_n(z) = \sum_{m=0}^{\infty} u_n(m) z^m$. The proof of the AAK-theorem then shows that $\check{u}_n$ has precisely n roots $(z_j)_{j=1}^n$, counted with multiplicity, and that the poles of the rational symbol for K in the AAK-theorem are located at $(1/z_j)_{j=1}^n$, again counted with multiplicity. In particular, if $\check{u}_n$ has distinct zeroes, then the best rank n Hankel approximant of Γ is a sum of n matrices of the form (1.4) with z_0 replaced by z_j, $j = 1, \ldots, n$.

The subspace generated by taking the closure of the span of $\{S^m x : m \in \mathbb{N}\}$ will be denoted by $[x]_S$. Note that if $[x]_S$ has finite codimension n, then by Beurling's theorem, $\check{x}$ has precisely n zeroes (counted with multiplicity), and $\mathcal{F}^{-1}([x]_S)$ is the subspace of the Hardy space H^2 of analytic functions with zeroes precisely at the zeroes of $\check{x}$ (and with multiplicities greater than those of $\check{x}$). Using Beurling's and Nehari's theorem, a short argument shows that the AAK-theorem is equivalent with the following result, (compare with the two definitions of σ_n in (1.3)).

THEOREM 1.4. *(AAK*) Let $\Gamma : l^2 \to l^2$ be a Hankel operator and let σ_n be its n'th singular value. Then there is a singular vector u_n to σ_n such that* $\operatorname{codim}\, [u_n]_S = n$ *and* $\|\Gamma|_{[u_n]_S}\| = \sigma_n$.

We will now discuss S. Treil and A. Volberg extension of the AAK-theorem in [8]. Let X_1 and X_2 denote Hilbert spaces and let $\Gamma : X_1 \to X_2$ denote a Hankel operator with respect to some operators $S \in \mathcal{L}(X_1)$ and $B \in \mathcal{L}(X_2)$. Recall that S is called expansive if $\|Sx\| \geq \|x\|$ for all $x \in X_1$ and contractive if $\|S\| \leq 1$.

THEOREM 1.5. *(Treil, Volberg) Assume that S is expansive and that B is a contraction and let $\Gamma : X_1 \to X_2$ be a Hankel operator. Let σ_n be a fixed singular value of Γ. Then there exists an S-invariant subspace $\mathcal{M}$ with* $\operatorname{codim}\, \mathcal{M} = n$ *such that* $\|\Gamma|_{\mathcal{M}}\| = \sigma_n$.

Their proof relies on a fixed point lemma by Ky Fan and does not imply anything concerning the singular vectors. In particular, it is not clear whether

$$\mathcal{M} = [u_n]_S \tag{1.5}$$

holds or, which is a weaker statement, whether $\mathcal{M}$ is determined by the zeroes of $\check{u}_n$. Clearly (1.5) is not to be expected in the full generality of the above theorem. For instance, if $\mathsf{codim}\ S(X_1) > 1$ it is easy to see that $\mathsf{codim}\ [u]_S = \infty$ for all $u \in X_1$. However, the main result of this paper says that $\mathcal{M}$ is indeed determined by the zeroes of $\check{u}_n$ for the Hankel operators considered in Example 1.2. The conditions given in the above theorem on S (the shift on l^2_w) and B (the backward shift on l^2_v) are then equivalent with the weights w and v being increasing. We will assume that w is not "too increasing", so that $\check{u}$ is an analytic function on the unit disc $\mathbb{D}$ for all $u \in l^2_w$. One way to ensure this is to assume that

$$\lim_{k\to\infty} w_{k+1}/w_k = 1, \tag{1.6}$$

which we will do from now on. Given $z \in \mathbb{D}$, the functional $l^2_w \ni u \mapsto \check{u}(z)$ will be called the point evaluation at z. We denote the unit circle by $\mathbb{T}$. The main result is slightly complicated due to the fact that point evaluations might be bounded also for points in $\mathbb{T}$. We omit this case from the below statement of the main theorem, and refer to Theorem 2.4 for the full version.

THEOREM 1.6. *Let w, v be increasing weights of which at least one is strictly increasing, and let $\Gamma : l^2_w \to l^2_v$ be a Hankel operator. Assume that w is such that point evaluations are not bounded on $\mathbb{T}$. Then the singular values are strictly decreasing. Let u_n be the singular vector for σ_n, let $\lambda_j \in \mathbb{D}$ denote the zeroes of $\check{u}_n$ and let $s_j \in \mathbb{N}$ be the respective multiplicities. Then $\sum_j s_j = n$ and, setting*

$$\mathcal{M} = \{x \in l^2_w : \check{x} \textit{ has a zero at each } \lambda_j \textit{ of multiplicity} \geq s_j\},$$

we have

$$\|\Gamma|_{\mathcal{M}}\| = \sigma_n.$$

Note that we clearly have $\mathsf{codim}\ \mathcal{M} = n$ in the above theorem. Section 2 is devoted to the proof of the above result, which actually relies on the theorem by Treil and Volberg. In Section 3 we discuss questions naturally related to Theorem 1.6. In particular we show that the original version of the AAK-theorem does not extend to the weighted case, which in turn implies that a strong version of Nehari's theorem can not hold in this setting. A weaker version of Nehari's theorem has been obtained in [8]. We also point out, without giving examples, that Theorem 1.6 is false if either w or v are decreasing.

2. Main. Throughout this section let $w = (w_k)_{k=0}^{\infty}$ be a fixed positive increasing weight such that (1.6) holds. Given $u \in l^2_w$ we clearly have that $\check{u}(z) = \sum_{k=0}^{\infty} u_k z^k$ defines an analytic function in the unit disc $\mathbb{D}$. The space of such functions with $\|\check{u}\| = \|u\|_{l^2_w}$ will be denoted by H^2_w. Recall that $\mathcal{F}$ denotes the Fourier transform, i.e. the operator such that $\mathcal{F}(\check{u}) = u$ for all $u \in L^2_w$. We also write $\mathcal{F}(f) = \hat{f}$. Define the differentiation-operator D formally by $D(\sum_{k=0}^{\infty} u_k z^k) = \sum_{k=1}^{\infty} k u_k z^{k-1}$ and let $s_w \in \mathbb{N} \cup \{\infty\}$ be the smallest number such that

$$\sum_{k=0}^{\infty} \frac{k^{2s_w}}{w_k} = \infty. \tag{2.1}$$

Before getting to the main result, Theorem 2.4, we study basic properties of the spaces l^2_w in the coming four propositions.

PROPOSITION 2.1. *Given any $s \in \mathbb{N}$ and $\lambda \in \mathbb{D}$, the map $l^2_w \ni u \mapsto (D^s\check{u})(\lambda)$ is continuous. The same is true for $|\lambda| = 1$ if and only if $s < s_w$.*

Proof. We consider only the case $|\lambda|=1$, and moreover it suffices to consider $\lambda=1$ by the rotational symmetry. Assume first that $s<s_w$. Let $u\in L^2_w$ be an element with finite support. Then

$$|(D^s\check{u})(1)|\leq |\sum_{k=0}^{\infty}k^s u_k| = \sum_{k=0}^{\infty}|k^s u_k\frac{\sqrt{w_k}}{\sqrt{w_k}}|\leq \sqrt{\sum_{k=0}^{\infty}|u_k|^2 w_k}\sqrt{\sum_{k=0}^{\infty}\frac{k^{2s}}{w_k}}$$

by the Cauchy-Schwartz inequality, and the desired continuity follows immediately. Conversely, assume that the map is continuous for some fixed $s\in\mathbb{N}$. Assume that $s\geq 1$. Then there exists a $g\in l^2_w$ such that $(D^s\check{u})(1)=\langle u,g\rangle_w$. In particular, $\prod_{j=0}^{s-1}(k-j)=(D^s\check{e_k})(1)=\overline{g_k}w_k$ for all $k\geq s$, and since the sequence $(\prod_{j=0}^{s-1}(k-j)/k^s)_{k=s}^{\infty}$ is bounded we get that $(k^s/w_k)_{k=0}^{\infty}\in l^2_w$, and thus $s<s_k$. □ Whenever $l^2_w\ni u\mapsto (D^s\check{u})(\lambda)$ is continuous, we denote by $g^{\lambda,s}\in l^2_w$ the element such that $(D^s\check{u})(\lambda)=\langle u,g^{\lambda,s}\rangle$. For a given $u\in l^2_w$ we say that $\check{u}$ has a zero of multiplicity s at $\lambda\in\overline{\mathbb{D}}$ if $\langle u,g^{\lambda,t}\rangle=0$ for all $0\leq t<s$ but not for $t=s$. For λ in $\mathbb{D}$ this clearly coincides with the classical definition.

Proposition 2.2. *Let S denote the shift operator on l^2_w. Then $\sigma(S)=\overline{\mathbb{D}}$ and $S-\lambda$ is an injective Fredholm operator with*

$$\mathsf{ind}\ (S-\lambda)=-1$$

for all $\lambda\in\mathbb{D}$. Moreover, for $\lambda\in\mathbb{T}$ and any integer $s\geq 1$ we have

$$\mathsf{codim}\ (cl(\mathsf{Ran}\ (S-\lambda)^s))=\begin{cases} s \text{ if } s\leq s_w\\ s_w \text{ if } s>s_w\end{cases}$$

Proof. The spectral radius is given by $\overline{\lim}_{k\to\infty}\|S^k\|^{1/k}$. Clearly

$$\|S^k\|^{1/k}=\sup_{j\in\mathbb{N}}\left\{(w_{j+k}/w_j)^{1/k}\right\}=\sup_{j\in\mathbb{N}}\left\{(\prod_{l=0}^{k-1}w_{j+l+l}/w_{j+l})^{1/k}\right\}$$

and for each fixed j_0 we have

$$\overline{\lim}_{k\to\infty}\sup_{j\in\mathbb{N}}\left\{\left(\prod_{l=0}^{k}w_{l+l}/w_l\right)^{1/k}\right\}\leq\overline{\lim}_{k\to\infty}\left(\prod_{l=0}^{j_0}w_{l+1}/w_l\left(\sup_{j\geq j_0}\{w_{j+1}/w_j\}\right)^k\right)^{1/k}\leq$$
$$\leq\overline{\lim}_{k\to\infty}\sup_{j\geq j_0}\{w_{j+l}/w_j\}\leq\sup_{j\geq j_0}\{w_{j+l}/w_j\}.$$

Hence, by (1.6), the spectral radius is less than 1 so $\sigma(S)\subset\overline{\mathbb{D}}$. Moreover we clearly have

$$\mathsf{Ran}\ (S-\lambda)\subset\{g^{\lambda,0}\}^{\perp} \tag{2.2}$$

for all $\lambda\in\mathbb{D}$, so consequently $\sigma(S)=\overline{\mathbb{D}}$. For $\lambda\in\mathbb{D}$ we also have $\|(S-\lambda)u\|_{l^2_w}\geq(1-|\lambda|)\|u\|_{l^2_w}$, so $S-\lambda$ is injective with closed range. By this observation and (2.2), $\mathsf{ind}\ (S-\lambda)=-1$ follows if we show that

$$\mathsf{codim}\ \mathsf{Ran}\ (S-\lambda)\leq 1. \tag{2.3}$$

Let $l^2_{w,0}$ denote the set of elements with finite support in l^2_w and note that $\mathcal{F}^{-1}(l^2_{w,0})$ can be identified with the set of polynomials. It is easily seen that $\dim\left(l^2_{w,0}/\mathsf{Ran}\ (S-\lambda)|_{l^2_{w,0}}\right)=1$, and since $l^2_{w,0}$ is dense in l^2_w, (2.3) follows easily by standard arguments. (/ stands for the quotient space.)

It remains to investigate $|\lambda|=1$. Put $c=\mathsf{codim}\ (cl(\mathsf{Ran}\ (S-\lambda)^s))$. Like above we get that $0\leq c\leq s$, because the codimension of $\mathsf{Ran}\ (S-\lambda)^s|_{l^2_{w,0}}$ in $l^2_{w,0}$ is s and $l^2_{w,0}$ is dense in l^2_w. If

$s_w \geq s$, we clearly have $s = c$ because $g^{\lambda,0}, \ldots, g^{\lambda,s-1} \perp \mathsf{Ran}\ (S-\lambda)^s$. In a similar fashion we see that $s_w \leq c \leq s$ if $s_w < s$. Suppose that $c > s_w$. A short argument shows that if A is an injective operator with $\mathsf{codim}\ (cl(A)) = k$ for some $k \in \mathbb{N}$ then the function

$$t \mapsto \mathsf{codim}\ (cl(A^t))$$

is of the form $\min(kt, kq)$ for some $q \in \mathbb{N}$. Thus

$$cl(\mathsf{Ran}\ (S-\lambda)^{s_w}) \ominus cl(\mathsf{Ran}\ (S-\lambda)^{s_w+1}) \neq \{0\}. \tag{2.4}$$

Pick a non-zero $h \in cl(\mathsf{Ran}\ (S-\lambda)^{s_w}) \ominus cl(\mathsf{Ran}\ (S-\lambda)^{s_w+1})$, fix $m > s_w$ and let u be defined by the equation

$$z^m = \sum_{j=0}^{s_w} (D^j z^m)|_\lambda \frac{(z-\lambda)^j}{j!} + (z-\lambda)^{s_w+1} \check{u}.$$

Then

$$w_m \overline{h_m} = \sum_{j=0}^{s_w} (D^j z^m)|_\lambda \left\langle \frac{(z-\lambda)^j}{j!}, \check{h} \right\rangle_{H^2_w} = \mathcal{O}(m^{s_w-1}) + \lambda^{s_w} \frac{m!}{(m-s_w)! s_w!} \langle (z-\lambda)^{s_w}, \check{h} \rangle_{H^2_w},$$

where $\mathcal{O}(m^{s_w-1})$ stands for "big ordo". Moreover, $(z-\lambda)^{s_w}$ is clearly an element of $cl(\mathsf{Ran}\ (S-\lambda)^{s_w})$ but not of $cl(\mathsf{Ran}\ (S-\lambda)^{s_w+1})$, because this would imply $cl(\mathsf{Ran}\ (S-\lambda)^{s_w}) = cl(\mathsf{Ran}\ (S-\lambda)^{s_w+1})$ contrary to (2.4). Thus $\langle (z-\lambda)^{s_w}, \check{h} \rangle_{H^2_w} \neq 0$ and we conclude that

$$|h_m| = \frac{m^{s_w}(C + \mathcal{O}(m^{-1}))}{w_m}$$

where $C = |\langle (z-\lambda)^{s_w}, \check{h} \rangle_{H^2_w}| / s_w!$. The fact that $\|h\|_{l^2_w} < \infty$ now implies that $\sum_{m=0}^{\infty} \frac{m^{2s_w}}{w_m} < \infty$, which contradicts (2.1) and the proof is complete. ∎ The next proposition characterizes all S-invariant finite codimensional subspaces in l^2_w.

PROPOSITION 2.3. *Let $m \in \mathbb{N}$ and let $\mathcal{M} \subset l^2_w$ be an S-invariant subspace such that* $\dim(l^2_w/\mathcal{M}) = m$. *Then there are a finite number of points $\lambda_j \in \overline{\mathbb{D}}$ and integers $s_j \in \mathbb{N}$ such that $\sum_j s_j = m$, where $s_j \leq s_w$ if $|\lambda_j| = 1$, and*

$$\mathcal{M} = \left(\cup_j \left\{ g^{\lambda_j,t} \right\}_{t=0}^{s_j-1} \right)^{\perp} = cl \left(\mathsf{Ran}\ \prod_j (S-\lambda_j)^{s_j} \right).$$

Conversely, any set of this form is S-invariant with codimension m.

Proof. Denote by $S_{\mathcal{M}}$ the map on $l^2_w/\mathcal{M}$ induced by S. Let p denote the characteristic polynomial for $S_{\mathcal{M}}$ and recall that $p(S_M) = 0$, which in particular leads to $p(S) \subset \mathcal{M}$. Moreover, letting λ_j denote the zeroes of p, we have $p(z) = \prod_j (z-\lambda_j)^{s_j}$ for some s_j's such that $\sum_j s_j = m$. Given two injective operators A, B, it is easily seen that $\mathsf{codim}\ (cl(\mathsf{Ran}\ AB)) \leq \mathsf{codim}\ (cl(\mathsf{Ran}\ A)) + \mathsf{codim}\ (cl(\mathsf{Ran}\ B))$ so

$$m = \mathsf{codim}\ (\mathcal{M}) \leq \mathsf{codim}\ (cl(\mathsf{Ran}\ p(S))) \leq \sum_j \mathsf{codim}\ (cl(\mathsf{Ran}\ (S-\lambda_l)^{s_j})) \leq \sum_j s_j = m.$$

Thus the inequalities are equalities and in particular we have $\mathcal{M} = cl(\mathsf{Ran}\ p(S))$ and

$$\mathsf{codim}\ (cl(\mathsf{Ran}\ (S-\lambda_l)^{s_j})) = s_j$$

for all j. By Proposition 2.2 we conclude that $\lambda_j \in \overline{\mathbb{D}}$ for all j and moreover that $s_j \leq s_w$ whenever $|\lambda_j| = 1$. By Proposition 2.1 we get that $\left(\cup_j \left\{ g^{\lambda,t} \right\}_{t=0}^{s_j-1} \right)$ is a well defined set, which clearly has dimension m and is orthogonal to $\mathsf{Ran}\ (p(S))$. Thus $\left(\cup_j \left\{g^{\lambda,t}\right\}_{t=0}^{s_j-1}\right) \perp \mathcal{M}$ and since $\mathsf{codim}\ \mathcal{M} = m$ the desired conclusion follows. □

We are now ready for the generalized AAK-theorem. Let $\sigma_\infty(A)$ denote the essential norm of any operator A, and let $\mathcal{E}_A$ denote the projection-valued measure associated with $\sqrt{A^*A}$, as given by the spectral theorem, (see e.g. [4]).

THEOREM 2.4. *Let w, v be increasing weights of which at least one is strictly increasing, and let $\Gamma : l^2_w \to l^2_v$ be a Hankel operator with finitely many non-zero entries. Then the singular values are strictly decreasing. Let u_n be the singular vector for σ_n, let $\lambda_j \in \overline{\mathbb{D}}$ denote the zeroes of $\breve{u}$ and let $s_j \in \mathbb{N}$ be the respective multiplicities. Then $\sum_j s_j = n$ and if $\mathcal{M}$ is defined by these values as in Proposition 2.3, we have $\mathcal{M} = [u_n]_S$ and*

$$\|\Gamma|_{\mathcal{M}}\| = \sigma_n.$$

Proof. For all $k \in \mathbb{N}$ with $\sigma_k > \sigma_\infty$ pick $u_k \in \mathsf{Ran}\ \mathcal{E}_\Gamma(\{\sigma_k\})$ such that $\{u_k\}_k$ is an orthonormal basis for $\mathsf{Ran}\ \mathcal{E}_\Gamma((\sigma_\infty, \infty))$. Assume that n is such that $\sigma_{n-1} > \sigma_n$. By the Treil and Volberg theorem, there exists an S-invariant subspace of codimension n such that $\|\Gamma|_{\mathcal{M}}\| = \sigma_n$. Since $\mathsf{Span}\ \{u_k\}_{k=0}^n$ is $n+1$-dimensional, it has a non-zero intersection with $\mathcal{M}$, so there are $c_0, \ldots, c_n$ such that

$$\sum_{k=0}^n c_k u_k \in \mathcal{M}.$$

But then

$$\sigma_n^2(\sum_{k=0}^n |c_k|^2) \geq \|\sqrt{\Gamma^*\Gamma}(\sum_{k=0}^n c_k u_k)\|^2_{l^2_w} = \|\sum_{k=0}^n \sigma_k c_k u_k\|^2_{l^2_w} = \sum_{k=0}^n \sigma_k^2 |c_k|^2$$

which, since $\sigma_k > \sigma_n$ for $k < n$, is only possible if $c_k = 0$ for all $k < n$. We thus have $u_n \in \mathcal{M}$, which since u_n can be any vector in $\mathsf{Ran}\ \mathcal{E}_\Gamma(\{\sigma_n\})$, gives

$$\mathsf{Ran}\ \mathcal{E}_\Gamma(\{\sigma_n\}) \subset \mathcal{M}. \tag{2.5}$$

Let $s_j \in \mathbb{N}$ and $\lambda_j \in \overline{\mathbb{D}}$ characterize $\mathcal{M}$ as in Proposition 2.3. Assume that there exists a $u \in \mathsf{Ran}\ \mathcal{E}_\Gamma(\{\sigma_n\})$ such that $\breve{u}$ has a zero at some λ where either $\lambda \notin \{\lambda_j\}_j$ or where $\lambda = \lambda_{j_0}$ for some j_0 but the multiplicity of $\breve{u}$ at λ is greater than s_{j_0}. This will be possible if either of the two statements in the theorem are false. We will show that it leads to a contradiction.

Assume first that $|\lambda| < 1$. By Proposition 2.2 and 2.3 there is an $a \in \mathcal{M}$ such that $u = (S-\lambda)a$. Set $b = \Gamma(a)/\sigma_n$. The following inequality is immediate

$$\|Bb\|_{l^2_v} \leq \|b\|_{l^2_v} \leq \|a\|_{l^2_w} \leq \|Sa\|_{l^2_w}. \tag{2.6}$$

Also, note that at least one of the outer inequalities is strict, depending on wether w or v is strictly increasing. Moreover we have

$$\langle (B-\lambda)b, b\rangle_{l^2_v} = \left\langle (B-\lambda)\frac{\Gamma(a)}{\sigma_n}, \frac{\Gamma(a)}{\sigma_n}\right\rangle_{l^2_v} = \left\langle \frac{\Gamma^*\Gamma((S-\lambda)a)}{\sigma_n^2}, a\right\rangle_{l^2_w} = \langle u, a\rangle_{l^2_w} = \langle (S-\lambda)a, a\rangle_{l^2_w}$$

which implies that

$$|\lambda|^2\|a\|^2_{l^2_w} - |\lambda|^2\|b\|^2_{l^2_v} = \mathsf{Re}\ \overline{\lambda}\langle Sa, a\rangle_{l^2_w} - \mathsf{Re}\ \overline{\lambda}\langle Bb, b\rangle_{l^2_v} \tag{2.7}$$

This in turn gives

$$0 = \|u\|^2_{l^2_w} - \left\|\frac{\Gamma(u)}{\sigma_n}\right\|^2_{l^2_w} = \|(S-\lambda)a\|^2_{l^2_w} - \|(B-\lambda)b\|^2_{l^2_v} = \|Sa\|^2_{l^2_w} - 2\mathsf{Re}\ \overline{\lambda}\langle Sa, a\rangle_{l^2_w} + |\lambda|^2\|a\|^2_{l^2_w} -$$
$$- \left(\|Bb\|^2_{l^2_v} - 2\mathsf{Re}\ \overline{\lambda}\langle Bb, b\rangle_{l^2_v} + |\lambda|^2\|b\|^2_{l^2_v}\right) = \left(\|Sa\|^2_{l^2_w} - \|Bb\|^2_{l^2_v}\right) - |\lambda|^2\left(\|a\|^2_{l^2_w} - \|b\|^2_{l^2_v}\right).$$

On the other hand, (2.6) yields

$$\|Sa\|^2_{l^2_w} - \|Bb\|^2_{l^2_v} > \|a\|^2_{l^2_w} - \|b\|^2_{l^2_v} \geq |\lambda|^2 \left(\|a\|^2_{l^2_w} - \|b\|^2_{l^2_v}\right)$$

which is a contradiction. Note that this is enough to prove that $\sigma_n > \sigma_{n+1}$ for all $n \in \mathbb{N}$. Note also that Theorem 1.6 follows from the above. (Theorem 1.6 is stronger than Theorem 2.4 for the case $s_w = 0$, because it does not assume that Γ has finitely many non-zero entries.)

Assume now that $|\lambda| = 1$. By the assumption that Γ has finitely many non-zero entries, we get that $\check{u}$ is a polynomial and hence $u \in \mathsf{Ran}\ (S - \lambda)$. Again we let $a \in \mathcal{M}$ be such that $u = (S - \lambda)a$ and $b = \Gamma(a)/\sigma_n$. By the same calculations as earlier we then have

$$0 = \left(\|Sa\|^2_{l^2_w} - \|a\|^2_{l^2_w}\right) + \left(\|b\|^2_{l^2_v} - \|Bb\|^2_{l^2_v}\right),$$

which again contradicts that one of the inequalities in (2.6) is strict. Finally, since $\check{u}$ is a polynomial, standard functional analysis shows that $\mathsf{codim}\ [u]_S \leq n$ which combined with $[u]_S \subset \mathcal{M}$ yields that $\mathcal{M} = [u]_S$.
□

3. Counterexamples. In this section we provide a counterexample to two natural questions related to the material in the previous section. The first concerns the original version of the AAK-theorem in the weighted setting, the second concerns Nehari's theorem.

Given a sequence γ we will write Γ_γ for the Hankel operator defined via (1.1). Let us consider a Hankel operator $\Gamma_\gamma : l^2_v \to l^2_w$, where w, v are increasing sequences, and suppose for simplicity that $s_w = 0$. Then $\check{u}_n$ has precisely n zeroes in $\mathbb{D}$, by Theorem 1.6. We assume that these zeroes are distinct, (which generically is the case), and label them $z_1, \ldots, z_n$. Theorem AAK naturally leads to the question of whether

$$\inf\left\{\left\|\Gamma_\gamma - \sum_{j=1}^{n} c_j \Gamma_{(z_j^k)_{k=0}^\infty}\right\|_{l^2_v \to l^2_w} : c_1, \ldots, c_n \in \mathbb{C}\right\} = \sigma_n. \tag{3.1}$$

This is false, as the following example shows.

EXAMPLE 3.1. *It is no restriction to assume that $w_0 = v_0 = 1$. Take $\gamma = e_1$. Since* $\mathsf{Ran}\ S^2$ *is a reducing subspace for Γ_{e_1}, it suffices to consider*

$$\Gamma = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then

$$\Gamma^* = \begin{pmatrix} 1 & 0 \\ 0 & v_1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & w_1 \end{pmatrix}$$

so $\Gamma^*\Gamma = \begin{pmatrix} w_1 & 0 \\ 0 & v_1^{-1} \end{pmatrix}$ and therefore $\sigma_0 = \sqrt{w_1}$, $u_0 = e_0$, $\sigma_1 = 1/\sqrt{v_1}$, $u_1 = e_1/\sqrt{v_1}$, since $w_1, v_1 \geq 1$. We thus get $\check{u}_0(z) = 1$ and $\check{u}_1 = z/\sqrt{v_1}$. Note that $\check{u}_0$ has no zeroes in $\mathbb{D}$ whereas $\check{u}_1$ has 1 zero at 0, as well as that $\|\Gamma|_{[u_1]_S}\|$, in accordance with Theorem 2.4. The rank one Hankel operator corresponding to 0 is $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. If (3.1) were to hold, then we should have

$$\inf\left\{\left\|\begin{pmatrix} -c_1 & 1 \\ 1 & 0 \end{pmatrix}\right\|_{l^2_v \to l^2_w} : c \in \mathbb{C}\right\} = \sigma_1 = 1/\sqrt{v_1}. \tag{3.2}$$

(To be formally correct, we should introduce a new notation for the restriction of l_v^2 and l_w^2 to $\mathsf{Span}\,\{e_0, e_1\}$*, but we omit this technicality.) However,*

$$\left\| \begin{pmatrix} -c_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\|_{l_w^2} \Big/ \left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\|_{l_v^2} = \sqrt{c_1^2 + w_1},$$

which clearly is greater than $1/\sqrt{v_1}$ unless $c_1 = 0$ and $v_1 = w_1 = 1$.

Incidentally, this also provides a counterexample related to Nehari's theorem. Let us write $l^2(\mathbb{N})$ and $l^2(\mathbb{Z})$ for the standard unweighted l^2-spaces. Nehari's theorem says that an operator

$$\Gamma : l^2(\mathbb{N}) \to l^2(\mathbb{N})$$

that satisfies $B\Gamma = \Gamma S$ can be "lifted" to an operator $\tilde{\Gamma} : l^2(\mathbb{Z}) \to l^2(\mathbb{Z})$ such that $B\tilde{\Gamma} = \tilde{\Gamma} S$, $\|\tilde{\Gamma}\| = \|\Gamma\|$ and

$$\Gamma = P_{l^2(\mathbb{N})} \tilde{\Gamma}|_{l^2(\mathbb{N})}, \tag{3.3}$$

where $P_{l^2(\mathbb{N})} : l^2(\mathbb{Z}) \to l^2(\mathbb{N})$ denotes the orthogonal projection. Assume now that v, w are increasing positive sequences on $\mathbb{Z}$, and define $l_v^2(\mathbb{N})$, $l_v^2(\mathbb{Z})$, $l_w^2(\mathbb{N})$, $l_w^2(\mathbb{Z})$ and $P_{l^2(\mathbb{N})} : l_w^2(\mathbb{Z}) \to l_w^2(\mathbb{N})$ in the obvious way. The extension of Nehari's theorem by S. Treil and A. Volberg in [8] shows that any operator $\Gamma : l_v^2(\mathbb{N}) \to l_w^2(\mathbb{N})$ satisfying $B\Gamma = \Gamma S$ can be lifted to an operator $\tilde{\Gamma} : l_v^2(\mathbb{N}) \to l_w^2(\mathbb{Z})$ such that $B\tilde{\Gamma} = \tilde{\Gamma} S$, $\|\tilde{\Gamma}\| = \|\Gamma\|$ and

$$\Gamma = P_{l^2(\mathbb{N})} \tilde{\Gamma}.$$

However, Example 3.1 shows that it is not possible in general to find a $\tilde{\Gamma} : l_v^2(\mathbb{Z}) \to l_w^2(\mathbb{Z})$ with $B\tilde{\Gamma} = \tilde{\Gamma} S$, $\|\tilde{\Gamma}\| = \|\Gamma\|$ and

$$\Gamma = P_{l_v^2(\mathbb{N})} \tilde{\Gamma}|_{l_w^2(\mathbb{N})},$$

because this would imply that (3.2) holds.

Acknowledgments. This research was supported under NSF CMG grant DMS 0724644.

REFERENCES

[1] Adamjan, V. M.; Arov, D. Z.; Krein, M. G. *Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur-Takagi problem.* (Russian) Mat. Sb. (N.S.) 86(128) (1971), 34–75.

[2] Aleman, A. *Finite codimensional invariant subspaces in Hilbert spaces of analytic functions.* J. Funct. Anal. 119 (1994) no. 1, p.1-18.

[3] Carlsson, M. *AAK-theory on weighted spaces.* GMIG-project review report 2009, Purdue University.

[4] Conway, J. B. *A course in functional analysis.* Second edition. Graduate Texts in Mathematics, 96. Springer-Verlag, New York, 1990.

[5] Nikolski, N. K. *Operators, functions, and systems: an easy reading. Vol. 1. Hardy, Hankel, and Toeplitz.* Mathematical Surveys and Monographs, 92. American Mathematical Society, Providence, RI, 2002.

[6] Peller, V. V. *Hankel operators and their applications.* Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.

[7] Rochberg, R. *A new characterization of Dirichlet type spaces and applications.* Illinois J. Math. 37 (1993), no. 1, p. 101-122.

[8] Treil, S.; Volberg. A. *A fixed point approach to Nehari's problem and its applications.* 165–186, Oper. Theory Adv. Appl., 71, Birkhäuser, Basel, 1994.