

QUASI-DIAGONALITY AND THE FINITE SECTION METHOD

NATHANIAL P. BROWN

ABSTRACT. Quasi-diagonal operators on a Hilbert space are a large and important class (containing all self-adjoint operators for instance). They are also perfectly suited for study via the finite section method (a particular Galerkin method). Indeed, the very definition of quasi-diagonality yields finite sections with good convergence properties. Moreover, simple operator theory techniques yield estimates on certain rates of convergence. In the case of quasi-diagonal band operators both the finite sections and rates of convergence are explicitly given.

1. INTRODUCTION

This paper follows up on an observation from [5, Section 10]. Namely we see what happens when one takes the general theorems from [11] and applies them to the particular class of quasi-diagonal operators. The end results are quite satisfying both from a theoretical point of view and a practical computational point of view. We will see that the finite section method works very well on the class of quasi-diagonal operators in general and some explicit, and reasonable rates of convergence can be worked out in some interesting cases.

Let H be a separable, complex Hilbert space, let $B(H)$ denote the bounded, linear operators on H and let $T \in B(H)$ be given. An operator T is called *quasi-diagonal* if there exist finite rank orthogonal projections P_1, P_2, P_3, \dots such that $\|P_n(v) - v\| \rightarrow 0$ for all vectors $v \in H$ and which asymptotically commute with T ;

$$\|[T, P_n]\| = \|TP_n - P_nT\| \rightarrow 0.$$

This notion is due to Halmos (cf. [13], [4], [18]) and is inspired by the classical Weyl-von Neumann Theorem. If $S \in B(H)$ is a *self-adjoint* operator, then $S = D + k$, where D is a diagonal operator with respect to some orthonormal basis of H and k is a compact operator. Indeed, Halmos observed that $T \in B(H)$ is quasi-diagonal if and only if $T = B + k$, where B is a block diagonal operator with respect to some orthonormal basis of H and k is compact. (This means there exist projections, as in the definition of quasi-diagonality, but with the stronger properties that $P_n H \subset P_{n+1} H$ and $[P_n, B] = 0$ for all n , writing $H = P_1 H \oplus (P_2 - P_1)H \oplus (P_3 - P_2)H \oplus \dots$, the matrix of B with respect to this decomposition is easily seen to be block diagonal.) In particular, every self-adjoint operator is quasi-diagonal and, more generally, a result of Berg implies that every normal operator is also quasi-diagonal (cf. [2], [9, Corollary II.4.2]). The self-adjoints contain a large set

Received by the editor February 16, 2005 and, in revised form, November 9, 2005.

2000 *Mathematics Subject Classification.* Primary 65J10, 46N40.

The work of this author was partially supported by an NSF Postdoctoral Fellowship and DMS-0244807.

of interesting examples (almost Mathieu operators, discretized Hamiltonians, etc.), and there are plenty of interesting non-self-adjoint, but still normal, examples like difference operators coming from PDE's with constant coefficients. There are also interesting nonnormal, but still quasi-diagonal, operators such as certain weighted shifts and Toeplitz operators (not all of these are quasi-diagonal, but many examples are known to be quasi-diagonal) to which the ideas of this paper apply. In fact, employing a simple dilation trick, we can even apply our techniques to arbitrary banded Toeplitz operators (though many of these are not quasi-diagonal).

Section 2 of this note contains a rather extensive list of definitions and recalls some of the theorems from [11] that we will need. There is nothing new in this section, but we hope the unfamiliar reader will find it a convenient reference.

Sections 3, 4, 5, and 6 are devoted to some general theory. These sections are the theoretical foundation on which the actual numerical analysis will be based. That is, we make some very general observations which are of little computational value by themselves. In some sense, one can regard these results as analogues of an existence result for a large class of PDEs. When faced with an actual PDE it is, of course, good to know via some general theory that a solution exists and even better if one knows it is unique. However, the real goal of numerical analysis is to find explicit algorithms which are guaranteed to converge to a solution and, moreover, to study the rate of convergence. With this analogy in mind one can regard Sections 3 through 6 as proving existence and uniqueness results as well as giving general estimates on rates of convergence. Then in Sections 7 and 8 some explicit algorithms are described which put the general machine into motion in two special cases. Below we give a more detailed description of the contents of this paper.

In Section 3 we observe the *existence* of good finite sections for quasi-diagonal operators and the *uniqueness* of the limits of things like norms, condition numbers, spectral quantities, etc. of the corresponding finite sections.

Section 4 gives some simple estimates on rates of convergence for the finite section method. Though the rates of convergence are not always very good, they are the best possible (at this level of generality) as shown by simple examples. (However, in certain cases the rates of convergence can be quite good.) Section 5 simply takes the results of Sections 3 and 4 and specializes them to the cases of normal and self-adjoint operators where the rates of convergence have cleaner statements. In Section 6 we quickly point out that the resolution of a well-known problem in operator theory (Herrero's approximation problem) yields an improved existence result for many quasi-diagonal operators. Putting this result into practice in special cases seems quite difficult, but a related strategy, inspired by the observations of Section 6, will be the subject of another paper (cf. [6]).

Sections 7 and 8 are devoted to explicit descriptions of how to implement the quasi-diagonal approach in the cases of unilateral and, respectively, bilateral band operators. In the unilateral case we use canonical finite sections and translate our previous results into computable error estimates. In the bilateral case a famous technique of Berg turns out to work beautifully and, again, computable rates of convergence are given.

2. NOTATION, DEFINITIONS, AND KNOWN RESULTS

This section is devoted to setting notation, recalling a number of definitions and stating some theorems from [11] that will be required.

Let $T \in B(H)$ be a bounded operator on some complex Hilbert space H and let $P_1 \leq P_2 \leq \dots$ be a *filtration* (i.e., each $P_n = P_n^* = P_n^2$ is a finite rank, orthogonal projection, the range of P_n is contained in the range of P_{n+1} , and $\|P_n(v) - v\| \rightarrow 0$ for all vectors $v \in H$). If the rank of P_n is some integer $k(n)$, then we will often, but not always, identify $P_n B(H) P_n$ with the complex $k(n) \times k(n)$ matrices $M_{k(n)}(\mathbb{C})$. We will refer to the sequence of operators $(P_n T P_n)$ as the *finite sections* (of T with respect to $\{P_n\}$). The *finite section method* is simply to try and study the (infinite-dimensional) operator T via its finite sections with respect to a suitable filtration. For example, can one recover the spectrum of T as some sort of limit of the spectra of (the finite-dimensional matrices) $P_n T P_n$? Of course, the choice of filtration is very important, as it can easily happen that the finite sections do not approximate the original operator in any reasonable sense. A classic example is that of the unilateral shift S on $l^2(\mathbb{N})$. With respect to the canonical basis, the matrix of S is below.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Cutting S to the canonical filtration yields nilpotent matrices, and thus the spectra of the finite sections are all one-point sets $\{0\}$ which do not converge to the spectrum of S (which is the unit disc).

Some operator theory definitions.

Definition 2.1. Given an operator $T \in B(H)$:

- (1) For $\varepsilon > 0$, the ε -pseudospectrum of T is

$$\sigma^{(\varepsilon)}(T) = \{\lambda \in \mathbb{C} : \|(\lambda - T)^{-1}\| \geq \frac{1}{\varepsilon}\},$$

where $\lambda - T$ denotes the operator which maps $v \in B(H)$ to $\lambda v - T v$, and we will adopt the convention that if an operator $X \in B(H)$ is *not invertible*, then $\|X^{-1}\| = \infty$. In particular, the usual spectrum $\sigma(T)$ is contained in $\sigma^{(\varepsilon)}(T)$ for every $\varepsilon > 0$ and, moreover,

$$\sigma(T) = \bigcap_{\varepsilon > 0} \sigma^{(\varepsilon)}(T).$$

See the web site [10] for more, including software for computing pseudospectra.

- (2) The *singular values* of T are $\sigma_2(T) = \{\lambda > 0 : \lambda^2 \in \sigma(T^* T)\}$.
 (3) We let $C^*(T)$ denote the *unital, C^* -algebra generated by T* . It is the smallest unital C^* -subalgebra of $B(H)$ which contains T and can be realized as the norm closure of the set of all polynomials (where T^0 is defined to be the identity operator) in T and T^* . A *state* on $C^*(T)$ is a linear functional $\phi : C^*(T) \rightarrow \mathbb{C}$ such that $\|\phi\| = \phi(I) = 1$, where I is the identity operator on H . The set of all states is denoted $S(C^*(T))$.
 (4) The *spatial numerical range* of T is

$$SN(T) = \{\langle T v, v \rangle : v \in H, \|v\| = 1\},$$

and the *algebraic numerical range* of T is

$$AN(T) = \{\phi(T) : \phi \in S(C^*(T))\}.$$

Note that $SN(T) \subset AN(T)$.

- (5) T is called *Moore–Penrose invertible* if there exists an operator $T^+ \in B(H)$ such that $TT^+T = T$, $T^+TT^+ = T^+$, $(TT^+)^* = TT^+$ and $(T^+T)^* = T^+T$. If such T^+ exists, then it is unique and is called the Moore–Penrose inverse of T . It turns out that if T^+ exists, then $T^+ \in C^*(T)$ (cf. [11, Theorem 2.15]) and hence, for an abstract C^* -algebra the notion of Moore–Penrose invertible does not depend on the choice of faithful representation. Also, recall that T^+ determines *least square solutions* to operator equations. If $y \in H$, then $x = T^+y$ is the least square solution of the equation $Tx = y$ (i.e., x is the unique vector such that $\|Tx - y\| = \inf\{\|Tu - y\| : u \in H\}$ and $\|x\| = \inf\{\|u\| : \|Tu - y\| = \inf\{\|Tu - y\| : u \in H\}\}$).
- (6) If T is invertible, then the *condition number* of T is

$$\text{cond}(T) = \|T\|\|T^{-1}\|,$$

while the *generalized condition number* of a Moore–Penrose invertible operator T is the number $\|T\|\|T^+\|$.

We now need to recall some definitions motivated by the finite section method, but this will require introducing a bit more notation. If $P_1 \leq P_2 \leq \dots$ is a filtration of H and the rank of P_n is $k(n)$, then it is useful to build a C^* -universe where finite sections can be treated as a single element. Namely we let

$$\Pi M_{k(n)}(\mathbb{C}) = \{(x_n) : x_n \in M_{k(n)}(\mathbb{C}), \sup \|x_n\| < \infty\}$$

and

$$\bigoplus M_{k(n)}(\mathbb{C}) = \{(x_n) \in \Pi M_{k(n)}(\mathbb{C}) : \lim \|x_n\| = 0\}$$

denote the l^∞ and, respectively, c_0 direct sum of the $M_{k(n)}(\mathbb{C})$'s (where we identify $P_n B(H) P_n$ and $M_{k(n)}(\mathbb{C})$). Note that $\bigoplus M_{k(n)}(\mathbb{C})$ is a closed, two-sided ideal in $\Pi M_{k(n)}(\mathbb{C})$, and hence there is a canonical quotient mapping $\pi : \Pi M_{k(n)}(\mathbb{C}) \rightarrow \Pi M_{k(n)}(\mathbb{C}) / \bigoplus M_{k(n)}(\mathbb{C})$. This quotient algebra turns out to be the right place to study finite sections. Note that if $T \in B(H)$, then $\sup \|P_n T P_n\| < \infty$ (actually, $= \|T\|$) and hence the sequence of finite sections $(P_n T P_n)$ can be regarded as a single element in $\Pi M_{k(n)}(\mathbb{C})$. The real element of interest, however, turns out to be

$$\pi((P_n T P_n)) \in \frac{\Pi M_{k(n)}(\mathbb{C})}{\bigoplus M_{k(n)}(\mathbb{C})}.$$

Definitions arising from the finite section method.

Definition 2.2. Let $T \in B(H)$, $P_1 \leq P_2 \leq \dots$ be a filtration of H and let $(P_n T P_n)$ be the corresponding sequence of finite sections.

- (1) $(P_n T P_n)$ is said to be *applicable* if there exists an integer n_0 such that for every vector $y \in H$ and integer $n \geq n_0$ there exists a unique solution x_n to the equation $P_n T P_n x_n = y_n$ and, moreover, the vectors x_n converge (in norm) to a solution x of the equation $Tx = y$.
- (2) $(P_n T P_n)$ is said to be *stable* if there exists an integer n_0 such that for every integer $n \geq n_0$ the matrices $P_n T P_n$ are invertible and $\sup_{n \geq n_0} \|(P_n T P_n)^{-1}\| < \infty$. A basic result of Polski states that $(P_n T P_n)$ is applicable if and only if T is invertible and $(P_n T P_n)$ is stable (cf. [11, Theorem 1.4]).

- (3) The finite sections (P_nTP_n) are said to be *fractal* if they are “asymptotically isomorphic to any of their subsequences”. More precisely, for every subsequence $n_1 < n_2 < n_3 \cdots$ the C^* -algebras generated by the elements

$$\pi((P_nTP_n)) \in \frac{\Pi M_{k(n)}(\mathbb{C})}{\bigoplus M_{k(n)}(\mathbb{C})} \text{ and } \pi((P_{n_j}TP_{n_j})) \in \frac{\Pi M_{k(n_j)}(\mathbb{C})}{\bigoplus M_{k(n_j)}(\mathbb{C})}$$

should be isomorphic, and the isomorphism should map $\pi((P_nTP_n))$ to $\pi((P_{n_j}TP_{n_j}))$.

- (4) For a matrix $S \in M_k(\mathbb{C})$ and $\varepsilon > 0$, the ε -regularization of S is denoted S_ε and defined as follows. Let $S = UDV$ be the singular value decomposition of S (i.e., $U, V \in M_k(\mathbb{C})$ are unitaries and $D = \text{diag}(\lambda_1, \dots, \lambda_k)$, where $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k$ are the (ordered) singular values of S). Define $D_\varepsilon = \text{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_k)$, where $\tilde{\lambda}_j = \lambda_j$ if $\lambda_j > \varepsilon$ and $\tilde{\lambda}_j = 0$ otherwise. The ε -regularization of S is then defined to be $S_\varepsilon = UD_\varepsilon V$.
- (5) The finite sections (P_nTP_n) are said to be *stably regularizable* if there exists a number $\varepsilon > 0$ such that $\|(P_nTP_n)_\varepsilon - P_nTP_n\| \rightarrow 0$ and $\sup \|(P_nTP_n)_\varepsilon^+\| < \infty$. In this case a Polski-type theorem (cf. [11, Theorem 2.13]) states that for every vector $y \in H$ the least square solutions x_n to the (regularized) equations $(P_nTP_n)_\varepsilon x_n = P_n y$ converge in norm to the least square solutions of $Tx = y$ (i.e., T is Moore–Penrose invertible and $(P_nTP_n)_\varepsilon^+ y \rightarrow T^+ y$ in norm for every $y \in H$).

One of the main topics of this paper is convergence aspects of spectra. As such it will be convenient to introduce some set notation and recall Hausdorff convergence of sets.

Approximate inclusions and Hausdorff convergence of sets.

Definition 2.3. Let \mathbb{C} denote the complex plane.

- (1) For a set $\Sigma \subset \mathbb{C}$ and $\delta > 0$ we will let $N_\delta(\Sigma)$ denote the δ -neighborhood of Σ (i.e., the union of all δ -balls centered at points of Σ) and let $\overline{N_\delta(\Sigma)}$ be its closure.
- (2) Given two sets $\Sigma, \Lambda \subset \mathbb{C}$ we say Σ is δ -contained in Λ if $\Sigma \subset N_\delta(\Lambda)$. We prefer (and will use) the following notation in this situation; $\Sigma \subset^\delta \Lambda$.
- (3) Given two compact sets $\Sigma, \Lambda \subset \mathbb{C}$ their *Hausdorff distance* is

$$d_H(\Sigma, \Lambda) = \max\left\{\sup_{\sigma \in \Sigma} d(\sigma, \Lambda), \sup_{\lambda \in \Lambda} d(\lambda, \Sigma)\right\},$$

where $d(\sigma, \Lambda) = \inf_{\lambda \in \Lambda} |\sigma - \lambda|$. Note that $\delta > d_H(\Sigma, \Lambda)$ if and only if Σ is δ -contained in Λ and vice versa, i.e., $\Lambda \subset^\delta \Sigma$ and $\Sigma \subset^\delta \Lambda$.

- (4) Given a sequence of compact sets $\Sigma_n \subset \mathbb{C}$ and another compact $\Sigma \subset \mathbb{C}$ we will write

$$\Sigma_n \rightarrow \Sigma$$

if $d_H(\Sigma_n, \Sigma) \rightarrow 0$, i.e., if for each $\delta > 0$ we have that $\Lambda_n \subset^\delta \Lambda$ and $\Lambda \subset^\delta \Lambda_n$ for all sufficiently large n .

We close this section by stating all of the theorems from [11] that we will need. Roughly speaking these results say that all convergence questions for the finite section method have a C^* -formulation in $\Pi M_{k(n)}(\mathbb{C}) / \bigoplus M_{k(n)}(\mathbb{C})$.

Theorem 2.4 (Operator equations). (1) (Kozak's Theorem) (P_nTP_n) is stably if and only if $\pi((P_nTP_n))$ is an invertible element in $\Pi M_{k(n)}(\mathbb{C})/\bigoplus M_{k(n)}(\mathbb{C})$.
 (2) (P_nTP_n) is stably regularizable if and only if $\pi((P_nTP_n))$ is Moore–Penrose invertible in $\Pi M_{k(n)}(\mathbb{C})/\bigoplus M_{k(n)}(\mathbb{C})$.

See [11, Theorem 1.15] for a proof of the first statement above and [11, Theorem 2.19] for the second.

Theorem 2.5 ([11, Theorem 2.28]). If (P_nTP_n) is fractal and stably regularizable, then the generalized condition number of $\pi((P_nTP_n))$ is equal to $\lim \|P_nTP_n\| \|X_n\|$, where (X_n) is any sequence such that $\pi((X_n))$ is the Moore–Penrose inverse of $\pi((P_nTP_n))$.

Theorem 2.6 (Spectral approximations). If (P_nTP_n) is fractal, then the following hold:

- (1) For every $\varepsilon > 0$, $\sigma^{(\varepsilon)}(P_nTP_n) \rightarrow \sigma^{(\varepsilon)}(\pi((P_nTP_n)))$.
- (2) $\sigma_2(P_nTP_n) \rightarrow \sigma_2(\pi((P_nTP_n)))$.
- (3) If each P_nTP_n happens to be normal, then $\sigma(P_nTP_n) \rightarrow \sigma(\pi((P_nTP_n)))$.

The first statement above is a combination of Theorems 3.31, 3.33(b), and Proposition 3.6 in [11]. The second is a consequence of Theorems 3.22, 3.23, and Proposition 3.6 in [11]. The last statement follows from Corollary 3.18, Theorem 3.20, and Proposition 3.6 in [11].

3. EXISTENCE AND UNIQUENESS RESULTS

Let $P_1 \leq P_2 \leq \dots$ be a filtration of H . If the rank of P_n is some integer $k(n)$ and we identify $P_nB(H)P_n$ with the complex $k(n) \times k(n)$ matrices $M_{k(n)}(\mathbb{C})$, then we get a natural linear mapping $\phi_n : B(H) \rightarrow M_{k(n)}(\mathbb{C})$ defined by $\phi_n(X) = P_nXP_n$ for all $X \in B(H)$. We further define linear maps $\Phi : B(H) \rightarrow \Pi M_{k(n)}(\mathbb{C})$, $\Phi(X) = (\phi_n(X))$, and $\rho : B(H) \rightarrow \Pi M_{k(n)}(\mathbb{C})/\bigoplus M_{k(n)}(\mathbb{C})$, $\rho = \pi \circ \Phi$, where $\pi : \Pi M_{k(n)}(\mathbb{C}) \rightarrow \Pi M_{k(n)}(\mathbb{C})/\bigoplus M_{k(n)}(\mathbb{C})$ is the canonical quotient mapping.

The following lemma, which is well known to the operator algebra community, is really the key observation of this section. The proof is also very simple, so we include it for the reader's convenience.

Lemma 3.1. If $\|[T, P_n]\| \rightarrow 0$, then ρ restricts to an injective $*$ -homomorphism on the unital C^* -algebra, denoted $C^*(T)$, generated by T . In other words, $\rho|_{C^*(T)} : C^*(T) \rightarrow \Pi M_{k(n)}(\mathbb{C})/\bigoplus M_{k(n)}(\mathbb{C})$ is a $*$ -isomorphism onto its image.

Proof. First note that the projections P_n asymptotically commute with T^* as well (since $\|[T, P_n]\| = \|[T, P_n]^*\| = \|[T^*, P_n]\|$). A simple interpolation argument together with the triangle inequality shows that (P_n) asymptotically commute with any polynomial in T and T^* . It follows that the P_n 's asymptotically commute with every element in $C^*(T)$. Since $\|\rho(X)\| = \|X\|$ for all $X \in B(H)$ (in particular, ρ is injective), it only remains to show that ρ restricts to a multiplicative map on $C^*(T)$. This follows easily since for all $a, b \in C^*(T)$ we have $\|\phi_n(ab) - \phi_n(a)\phi_n(b)\| = \|P_nabP_n - P_naP_nbP_n\| = \|P_n(P_na - aP_n)bP_n\| \rightarrow 0$. \square

Lemma 3.2. Let $T \in B(H)$ be a quasi-diagonal operator and let $P_1 \leq P_2 \leq \dots$ be any increasing sequence of finite rank projections such that $\|[P_n, T]\| \rightarrow 0$ and $\|P_n(v) - v\| \rightarrow 0$, for all $v \in H$, as $n \rightarrow \infty$. Then (P_nTP_n) is fractal.

Proof. Every subsequence of $\{P_n\}$ also asymptotically commutes with T and converges strongly to the identity. Hence the corresponding map $\rho : C^*(T) \rightarrow \Pi M_{k(n_j)}(\mathbb{C}) / \bigoplus M_{k(n_j)}(\mathbb{C})$ is a $*$ -isomorphism. \square

As stated in the title of this section, we regard the following results as existence and uniqueness results. That is, for a quasi-diagonal operator one can always find a sequence of projections such that the finite sections converge where we want (existence) and, in fact, the finite sections arising from any other *asymptotically commuting* sequence will converge to the same thing (uniqueness).

Theorem 3.3 (Operator equations). *Let $T \in B(H)$ be a quasi-diagonal operator and let $P_1 \leq P_2 \leq \dots$ be any filtration of H which asymptotically commutes with T . Then the following statements hold:*

- (1) *The sequence $(P_n T P_n) \in \Pi M_{k(n)}(\mathbb{C})$ is stable if and only if T is invertible.*
- (2) *The sequence $(P_n T P_n) \in \Pi M_{k(n)}(\mathbb{C})$ is applicable if and only if T is invertible.*
- (3) *The sequence $(P_n T P_n)$ is stably regularizable if and only if T is Moore–Penrose invertible (i.e., the range of T is closed; see [11, Theorem 2.4]).*
- (4) *Assume that T has closed range, $\varepsilon > 0$ is some number strictly smaller than $\inf\{\lambda : 0 < \lambda \in \sigma_2(T)\}$, $y \in H$ is given, and $x \in H$ is the least square solution to the equation $Tx = y$. If $y_n = P_n y$ and x_n is the least square solution to the ε -regularized equation $(P_n T P_n)_\varepsilon x_n = y_n$, then $\|x_n - x\| \rightarrow 0$ as $n \rightarrow \infty$.*

Proof. (1) follows from Kozak’s Theorem and Lemma 3.1 since $\rho(T)$ is invertible if and only if T is invertible.

(2) follows immediately from (1) and Polski’s Theorem which was stated right after the definition of stability in the previous section.

(3) follows from the second part of Theorem 2.4 and Lemma 3.1.

(4) follows from (3) and the least square version of Polski’s Theorem stated after the definition of stably regularizable. Strictly speaking, [11, Theorem 2.13] only ensures the existence of an $\varepsilon > 0$ for which our claim holds true. However, the convergence of spectra results below will ensure that any $0 < \varepsilon < \inf\{\lambda : 0 < \lambda \in \sigma(T^*T)\}$ will do. \square

Theorem 3.4 (Norms and condition numbers). *Let $T \in B(H)$ be a quasi-diagonal operator and let $P_1 \leq P_2 \leq \dots$ be any filtration of H which asymptotically commutes with T . Then the following statements hold:*

- (1) $\|T\| = \lim_{n \rightarrow \infty} \|P_n T P_n\|$ and $\|T^{-1}\| = \lim_{n \rightarrow \infty} \|(P_n T P_n)^{-1}\|$.
- (2) If T is invertible, then $\text{cond}(T) = \lim_{n \rightarrow \infty} \text{cond}(P_n T P_n)$.
- (3) If T has closed range, $\varepsilon > 0$ is some number strictly smaller than $\inf\{\lambda : 0 < \lambda \in \sigma_2(T)\}$, and T^+ (resp. $(P_n T P_n)_\varepsilon^+$) denotes the Moore–Penrose inverse of T (resp. $(P_n T P_n)_\varepsilon$), then the generalized condition numbers of $(P_n T P_n)_\varepsilon$ converge to the generalized condition numbers of T , i.e., $\|T\| \|T^+\| = \lim_{n \rightarrow \infty} \|(P_n T P_n)_\varepsilon\| \|(P_n T P_n)_\varepsilon^+\|$.

Proof. The first part of statement (1) is trivial. In fact, for every $X \in B(H)$, the norms of $P_n X P_n$ increase up to $\|X\|$ (since $P_n P_{n+1} = P_n$ and thus $\|P_n X P_n\| = \|P_n (P_{n+1} X P_{n+1}) P_n\| \leq \|P_{n+1} X P_{n+1}\|$). For the inverse statement we first assume

that T is invertible and note that $\|T^{-1}\| = \lim \|P_n T^{-1} P_n\|$. Also, since $T^{-1} \in C^*(T)$, the fact that $\rho(T^{-1}) = \rho(T)^{-1}$ implies that $(P_n T^{-1} P_n) - ((P_n T^{-1} P_n)^{-1}) \in \bigoplus M_{k(n)}(\mathbb{C})$. It follows that $\|P_n T^{-1} P_n - (P_n T^{-1} P_n)^{-1}\| \rightarrow 0$ as $n \rightarrow \infty$ and hence $\|T^{-1}\| = \lim \|P_n T^{-1} P_n\| = \lim \|(P_n T P_n)^{-1}\|$. If T is not invertible but $\lim_{n \rightarrow \infty} \|(P_n T P_n)^{-1}\| \neq \infty$, then we can find a subsequence $\{n_j\}$ such that

$$\lim_{j \rightarrow \infty} \|(P_{n_j} T P_{n_j})^{-1}\| = c < \infty.$$

However, the projections $\{P_{n_j}\}$ also asymptotically commute with T , and the finite sections $(P_{n_j} T P_{n_j})$ are now stable. Thus, by Theorem 3.3(1), T is invertible—a contradiction.

(2) follows immediately from (1).

(3) follows from Theorem 2.5 and Lemmas 3.1 and 3.2. \square

Theorem 3.5 (Spectral approximations). *Let $T \in B(H)$ be a quasi-diagonal operator and let $P_1 \leq P_2 \leq \dots$ be any filtration of H which asymptotically commutes with T . Then the following statements hold:*

- (1) For every $\varepsilon > 0$, $\sigma^{(\varepsilon)}(P_n T P_n) \rightarrow \sigma^{(\varepsilon)}(T)$.
- (2) $\sigma(T) = \bigcap_{\varepsilon > 0} \left(\lim \sigma^{(\varepsilon)}(P_n T P_n) \right)$, where $\lim \sigma^{(\varepsilon)}(P_n T P_n) = \sigma^{(\varepsilon)}(T)$ by (1).
- (3) $\sigma_2(P_n T P_n) \rightarrow \sigma_2(T)$.

Proof. Given Lemmas 3.1 and 3.2, the proof is reduced to recalling Theorem 2.6 from the last section. \square

Remark 3.6. As mentioned earlier, a theorem of Berg states that every normal operator on a Hilbert space is quasi-diagonal, and hence the existence and uniqueness results above can be applied. Even in the case of a self-adjoint operator the *existence* of a filtration for which the finite sections are guaranteed to converge nicely appears to be a new result.

Remark 3.7. Many standard examples of operators are not quasi-diagonal. For example, any Fredholm operator with nonzero index is not quasi-diagonal. However, there are very nice characterizations of quasi-diagonality for essentially normal operators (cf. [9, Corollary IX.7.4, Section IX.8]) and weighted shift operators ([16], [15]).

Remark 3.8. Though many operators are not quasi-diagonal, a remarkable result of Voiculescu implies that if one starts with an arbitrary $T \in B(H)$, then it is possible to construct a quasi-diagonal operator out of T . Namely, let $\lambda_1, \lambda_2, \dots$ be any listing of all the rational numbers in the half open interval $(0, 1]$. If $T \in B(H)$ is arbitrary, then Voiculescu's homotopy invariance theorem (cf. [18]) implies that

$$\tilde{T} = \bigoplus_{i \in \mathbb{N}} \lambda_i T,$$

acting on $\bigoplus_{i \in \mathbb{N}} H$, is a quasi-diagonal operator.

We close this section with a result which has nothing to do with quasi-diagonality. Namely, we will consider the convergence of the algebraic numerical ranges of finite sections. It is well known that spatial numerical ranges are well behaved (cf. [11, Theorem 3.52]), but algebraic numerical ranges also behave nicely. Though we would be a bit surprised if this result is not already known, it is not mentioned in [11], and hence we include a proof.

Theorem 3.9 (Numerical ranges). *Let $T \in B(H)$ be arbitrary and let $P_1 \leq P_2 \leq \dots$ be any sequence of finite rank projections such that $\|P_n(v) - v\| \rightarrow 0$, as $n \rightarrow \infty$, for all $v \in H$. Then the following statements hold:*

- (1) *For every n , $SN(P_nTP_n) = AN(P_nTP_n) \subset SN(T) \subset AN(T)$.*
- (2) *$SN(P_nTP_n) = AN(P_nTP_n) \rightarrow AN(T) = \overline{SN(T)} = AN(T)$.*

Proof. It is immediate from the definitions that $SN(X) \subset AN(X)$ for any $X \in B(H)$. Moreover, it is also known that $AN(X) = \overline{SN(X)}$ (cf. [11, Theorem 3.45]). Since the unit ball of a finite-dimensional Hilbert space (e.g., P_nH) is (norm) compact, these two facts imply that $SN(P_nTP_n) = AN(P_nTP_n)$ (for a given state ϕ , choose unit vectors $v_k \in P_nH$ such that $\phi(P_nTP_n) = \lim \langle P_nTP_nv_k, v_k \rangle$ and any cluster point of $\{v_k\}$ will do the trick). The proof of (1) is then complete after we observe that $SN(P_nTP_n) \subset SN(T)$: for a unit vector $v \in P_nH$ we have $\langle Tv, v \rangle = \langle TP_nv, P_nv \rangle = \langle P_nTP_nv, v \rangle$.

Statement (2) follows from (1) and the general fact that $SN(P_nTP_n) \rightarrow \overline{SN(T)}$ (cf. [11, Theorem 3.52]). \square

4. COMMUTATORS AND RATES OF CONVERGENCE

In this section we observe that norms of the commutators $\|[P_n, T]\|$ provide a simple and natural upper bound for the rate of convergence of finite sections (P_nTP_n) . For spectral approximations this only yields “one sided” rates of convergence, but, sadly, a trivial example at the end of this section shows that this is best possible.

From a computational point of view this will, in general, not provide a very useful measurement. However, we will see in later sections that in some cases these commutator estimates yield computationally relevant rates of convergence.

As was the case for Lemma 3.1, the following (well-known) lemma is really the key remark of this section.

Lemma 4.1. *Let $T \in B(H)$ be arbitrary and let $P \in B(H)$ be a self-adjoint projection. Then*

$$\|T - (PTP + P^\perp TP^\perp)\| \leq \|PT - TP\|,$$

where $P^\perp = I - P$.

Proof. Since $T = PTP + PTP^\perp + P^\perp TP + P^\perp TP^\perp$ ($T = ITI$ and $I = P + P^\perp$) we have that $T - (PTP + P^\perp TP^\perp) = PTP^\perp + P^\perp TP$. Since the domains and ranges of the operators PTP^\perp and $P^\perp TP$ are orthogonal, we have $\|PTP^\perp + P^\perp TP\| = \max\{\|PTP^\perp\|, \|P^\perp TP\|\}$. However, each of the latter norms is dominated by $\|PT - TP\|$; for example, $\|P^\perp TP\| = \|P^\perp TP - P^\perp PT\| \leq \|P^\perp\| \|TP - PT\|$. \square

Though we prefer to state the results below in terms of norms of commutators, we should also point out that these norms are determined by the norms of finite dimensional matrices:

$$\begin{aligned} \|PT - TP\| &= \|PTP^\perp - P^\perp TP\| \\ &= \max\{\|PTP^\perp\|, \|P^\perp TP\|\} \\ &= \max\{\|P^\perp T^*P\|, \|P^\perp TP\|\} \\ &= \max\{\|PTP^\perp T^*P\|^{1/2}, \|PT^*P^\perp TP\|^{1/2}\}. \end{aligned}$$

Proposition 4.2 (Operator equations). *Let $T \in B(H)$ be a quasi-diagonal operator and let $P_1 \leq P_2 \leq \dots$ be any filtration such that $\|[P_n, T]\| \rightarrow 0$. Then the following statements hold:*

- (1) *If T is invertible and $\|[P_n, T]\| < \frac{1}{\|T^{-1}\|}$ then $P_n T P_n$ is also invertible.*
- (2) *If T is invertible, $\|[P_n, T]\| < \frac{1}{\|T^{-1}\|}$, $y \in H$ is given, $x \in H$ is the unique solution to $Tx = y$, $y_n = P_n y$, and x_n is the unique solution to $P_n T P_n x_n = y_n$, then*

$$\|x - x_n\| \leq \|T^{-1}\|(\|y - y_n\| + \|x_n\| \|[P_n, T]\|).$$

Proof. The first statement follows from the following inequalities:

$$\|P_n T P_n T^{-1} P_n - P_n\| = \|P_n (T P_n - P_n T) T^{-1} P_n\| \leq \|[P_n, T]\| \|T^{-1}\| < 1.$$

Indeed, this shows that $P_n T P_n T^{-1} P_n = (P_n T P_n)(P_n T^{-1} P_n)$ is invertible (we now regard P_n as the unit of the matrix algebra $P_n B(H) P_n$) and hence $P_n T P_n$ is right invertible. However, in finite dimensions right invertible implies invertible.

The second assertion is also straightforward:

$$\begin{aligned} \|x_n - x\| &= \|T^{-1}(Tx_n - Tx)\| \\ &\leq \|T^{-1}\| \|Tx_n - y\| \\ &= \|T^{-1}\| \|TP_n x_n - P_n T P_n x_n + y_n - y\| \\ &\leq \|T^{-1}\| (\|y_n - y\| + \|x_n\| \|(TP_n - P_n T)P_n\|) \\ &\leq \|T^{-1}\| (\|y - y_n\| + \|x_n\| \|[P_n, T]\|). \end{aligned}$$

□

Remark 4.3 (Computability). We regard the estimates above as reasonably computable objects (at least in concrete cases). Indeed, if T and $\{P_n\}$ are concretely given, then $\|[P_n, T]\|$ is computable in terms of two finite-dimensional matrices. (Warning: in this remark we take the point of view that anything in finite dimensions is “computable” and limits of finite-dimensional calculations are regarded as “reasonably computable”.) Also, note that $\|y - P_n y\| = \sqrt{\|y\|^2 - \|P_n y\|^2}$ is computable (at least if one knows $\|y\|$).

Recall that the polar decomposition implies that for any invertible operator S we have

$$\|S^{-1}\| = \sup_{\lambda \in \sigma_2(S)} \frac{1}{\lambda}.$$

Hence the quantities $\|T^{-1}\|$ in the estimates above are reasonably computable since we know the singular values of $P_n T P_n$ converge to the singular values of T . Also, $\|x_n\| \leq \|(P_n T P_n)^{-1}\| \|y_n\|$ and thus can be estimated in concrete cases.

Remark 4.4 (Least square estimates). We have been unable to find any “reasonable” estimates on the rate of convergence of least square solutions. Our feeling is that this may be possible, however we do not yet see how to avoid the Moore–Penrose projection (a certain spectral projection for T^*T ; see [11, Theorem 2.15]) which arises in the construction of T^+ .

Proposition 4.5 (Spectral approximations). *Let $T \in B(H)$ be a quasi-diagonal operator and let $P_1 \leq P_2 \leq \dots$ be any filtration such that $\|[P_n, T]\| \rightarrow 0$. Then the following statements hold:*

- (1) *For every $\varepsilon > 0$, $\sigma^{(\varepsilon)}(P_n T P_n) \subset \sigma^{(\varepsilon + \|[P_n, T]\|)}(T)$.*

(2) $\sigma_2(P_n T P_n) \subset^{\delta_n} \sigma_2(T)$, where $\delta_n = \sqrt{2\|T\|\| [P_n, T] \|}$.

Proof. We will need the following characterization of ε -pseudospectra (cf. [11, Theorem 3.27]): if $X \in B(H)$, then

$$\sigma^{(\varepsilon)}(X) = \bigcup_{Y \in B(H), \|Y\| \leq \varepsilon} \sigma(X + Y).$$

Given this fact, it is clear that if $\|X - Z\| \leq \delta$, then $\sigma^{(\varepsilon)}(X) \subset \sigma^{(\varepsilon+\delta)}(Z)$. Since $\sigma^{(\varepsilon)}(P_n T P_n) \subset \sigma^{(\varepsilon)}(P_n T P_n + P_n^\perp T P_n^\perp)$, statement (1) follows from Lemma 4.1 and the remarks above.

Part (2) requires the following well-known fact: if $X, Y \in B(H)$, X is self-adjoint (even normal will suffice) and $\|X - Y\| \leq \delta$, then $\sigma(Y) \subset^\delta \sigma(X)$. To complete the proof we first note that $\sigma(P_n T^* P_n T P_n) \subset \sigma(P_n T^* P_n T P_n + P_n^\perp T^* P_n^\perp T P_n^\perp) = \sigma((P_n T^* P_n + P_n^\perp T^* P_n^\perp)(P_n T P_n + P_n^\perp T P_n^\perp))$. However, Lemma 4.1 implies that $\|(P_n T^* P_n + P_n^\perp T^* P_n^\perp)(P_n T P_n + P_n^\perp T P_n^\perp) - T^* T\| \leq 2\|T\|\| [P_n, T] \|$. Combining all these facts we see that $\sigma(P_n T^* P_n T P_n) \subset^{2\|T\|\| [P_n, T] \|} \sigma(T^* T)$; i.e., if $\lambda \in \sigma_2(P_n T P_n)$, then there exists some point $\gamma \in \sigma_2(T)$ such that $|\lambda^2 - \gamma^2| \leq 2\|T\|\| [P_n, T] \|$. Since $(\lambda - \gamma)^2 \leq |\lambda^2 - \gamma^2|$, statement (2) follows. Note that for large singular values one does not have to take square roots, however the estimate $(\lambda - \gamma)^2 \leq |\lambda^2 - \gamma^2|$ becomes sharp at zero. \square

We end this section with a simple example which shows that there is no hope of giving rates of convergence of the quantities not mentioned in this section in terms of norms of commutators.

Let $j_1 < j_2 < j_3 \cdots$ be integers and consider the diagonal, self-adjoint operator on $l^2(\mathbb{N})$ whose matrix (w.r.t. the canonical orthonormal basis) is

$\text{diag}(2, \dots, 2, 1+1/2, 3-1/2, 2, \dots, 2, 1+1/3, 3-1/3, 2, \dots, 2, 1+1/4, 3-1/4, 2, \dots),$

where the first string of 2's has length j_1 , the second string of 2's has length j_2 , and so on. One should have in mind here that the j_k 's are growing like exponentials (or faster, if you like) in k . Now let P_n denote the orthogonal projection onto the span of the first n basis vectors of $l^2(\mathbb{N})$.

Note that $[P_n, T] = 0$ for all n . On the other hand, the norms of $P_n T P_n$ and its inverse approach the norms of T and its inverse as slowly (in n) as you like. Also, the condition numbers of $P_n T P_n$ are approaching the condition number of T as slowly as you like.

It would also be nice if we could approximately reverse the inclusions in Proposition 4.5; i.e., if we knew something like $\sigma(T) \subset \sigma^{(\varepsilon)}(P_n T P_n)$ for all n larger than some number which depended on commutators. This would amount to saying how quickly the pseudospectra of $P_n T P_n$ "fill up" the pseudospectrum of T . Unfortunately, the example above shows that this can happen as slowly as desired. Similarly the numerical ranges of $P_n T P_n$ will "fill up" the numerical range of T at as slow a rate as one wishes to prescribe.

Finally, it would be nice for solving operator equations if we could somehow relate $\|y - y_n\|$ and $\|[P_n, T]\|$, but it is not hard to see that this is also impossible to do in general.

5. NORMAL AND SELF-ADJOINT OPERATORS

Here we observe that some of the statements from the previous section can be cleaned up a bit in the case that T is a normal or, better yet, self-adjoint operator.

At the end of this section we also state a result from [5] which provides a very general *existence* result which is analogous to the Szegő-type results obtained by Bill Arveson for certain self-adjoints (cf. [11, Chapter 7]).

Proposition 5.1. *If $T \in B(H)$ is a normal operator and $P_1 \leq P_2 \leq \dots$ is any filtration such that $\|[P_n, T]\| \rightarrow 0$, then:*

- (1) *For every $\varepsilon > 0$, $\sigma^{(\varepsilon)}(P_n T P_n) \rightarrow \overline{N_\varepsilon(\sigma(T))}$.*
- (2) *For every $\varepsilon > 0$, $\sigma^{(\varepsilon)}(P_n T P_n) \subset^{\delta_n} \sigma(T)$, where $\delta_n = \varepsilon + \|[T, P_n]\|$.*

Proof. For a normal operator T we have $\overline{N_\varepsilon(\sigma(T))} = \sigma^{(\varepsilon)}(T)$. This follows from the fact that if T is normal and $\lambda \in \mathbb{C}$, then $\|(\lambda - T)^{-1}\| = \frac{1}{d(\lambda, \sigma(T))}$, where $d(\lambda, \sigma(T))$ is the distance from λ to $\sigma(T)$. Hence we have just reformulated the results of Sections 2 and 3. \square

Remark 5.2. On page 116 in [11] a general conjecture concerning convergence of spectra of “essentially normal” approximation sequences is formulated. Though we have been unable to resolve this conjecture, we do wish to point out that an affirmative answer would lead to a nice improvement in the previous result. Namely, it would follow that $\sigma(P_n T P_n) \rightarrow \sigma(T)$; i.e., we would not have to bother with pseudospectra in the result above.

Proposition 5.3. *If $T \in B(H)$ is a self-adjoint operator and $P_1 \leq P_2 \leq \dots$ is any filtration such that $\|[P_n, T]\| \rightarrow 0$, then:*

- (1) $\sigma(P_n T P_n) \rightarrow \sigma(T)$.
- (2) $\sigma(P_n T P_n) \subset^{\delta_n} \sigma(T)$, where $\delta_n = \|[T, P_n]\|$.

Proof. Assertion (1) follows from part (3) of Theorem 2.6 and Lemmas 3.1 and 3.2. The second part follows from Lemma 4.1, the fact that

$$\sigma(P_n T P_n) \subset \sigma(P_n T P_n + P_n^\perp T P_n^\perp),$$

and the first sentence in the proof of part (2) of Proposition 4.5. \square

Remark 5.4. Following Arveson’s ideas, it is possible to also compute the essential spectrum of a self-adjoint operator via quasi-diagonal finite sections (cf. [11, Chapter 7]). It does not seem easy, however, to develop any rate of convergence estimates in this case.

We conclude this section with an *existence* result which is analogous to Szegő’s celebrated limit theorem concerning spectral distributions of self-adjoint Toeplitz operators. In [5] we studied various notions of invariant means on C^* -algebras and observed in Theorem 10.4 that “quasi-diagonal” invariant means can be recovered following the ideas first laid out by Bill Arveson (as explained in [11, Section 7.2]). To avoid technicalities we will just restate the result in the case of normal operators.

Theorem 5.5. *Let $T \in B(H)$ be a normal operator. Then there exists a filtration $P_1 \leq P_2 \leq \dots$ such that $\|[P_n, T]\| \rightarrow 0$ and with the property that for every Borel, probability measure μ supported on the essential spectrum of T there exists a subsequence $\{P_{n_k}\}$ such that*

$$\text{tr}(P_{n_k} \cdot P_{n_k}) \rightarrow \mu$$

in the weak- $$ topology, where $\text{tr}(\cdot)$ is the (unique) tracial state on $P_{n_k} B(H) P_{n_k}$ and $\text{tr}(P_{n_k} \cdot P_{n_k})$ is the state on $C^*(T)$ defined by $x \mapsto \text{tr}(P_{n_k} x P_{n_k})$.*

Proof. Since normal operators generate exact C^* -algebras (even nuclear ones) and every trace on an abelian C^* -algebra is a quasi-diagonal invariant mean (cf. [5, Lemma 6.4]), this result is a special case of [5, Theorem 10.4]. \square

6. IMPROVED EXISTENCE RESULT WHEN $C^*(T)$ IS EXACT

Mainly due to the pioneering work of Kirchberg, *exact* C^* -algebras have been at the forefront of operator algebra theory for the past decade. (See [14], [20].) This is a technical notion which we do not wish to discuss here, however it turns out that most (but not all) of the basic examples of C^* -algebras enjoy this property (commutative algebras, AF algebras, irrational rotation algebras, Cuntz algebras, and all subalgebras of these examples). Similarly, it turns out that many standard examples of Hilbert space operators generate C^* -algebras with this property. For example, $C^*(T)$ is exact whenever $T \in B(H)$ is essentially normal (e.g., self-adjoint, unitary, or normal) or if T is a band operator (either unilateral or bilateral). Moreover, it follows from Arveson's extension theorem that this class of operator is closed under norm limits and hence band dominated operators (i.e., norm limits of band operators) also generate exact C^* -algebras.

In the early 80s Herrero asked whether every quasi-diagonal operator T was the norm limit of operators T_n such that $C^*(T_n)$ was finite-dimensional for each n . This turns out to be false (cf. [17], [19]), but the precise description of operators with this approximation property was discovered a couple of years ago (cf. [7], [8]).

Theorem 6.1 ([7, Theorem 1.1]). *Let $T \in B(H)$ be given. There exist operators $T_n \in B(H)$ such that $\|T - T_n\| \rightarrow 0$, and $C^*(T_n)$ is finite-dimensional (for all n) if and only if T is quasi-diagonal and $C^*(T)$ is exact.*

This result is relevant here because of the simple structure of finite-dimensional C^* -algebras: they are just finite direct sums of finite-dimensional matrix algebras. So if T is close to T_n and $C^*(T_n)$ is finite dimensional, then one immediately gets a matrix whose ε -pseudospectrum, for example, approximates that of T (since we may regard T_n as a direct sum of finite-dimensional matrices when viewed as an element of $C^*(T_n)$). In this setting rates of convergence are controlled by $\|T_n - T\|$ and estimates on the Hausdorff distance of spectra (in the normal case, at least) can be given. Moreover, standard representation theory of finite-dimensional C^* -algebras gives a decomposition of the underlying Hilbert space as a direct sum of finite-dimensional reducing subspaces for $C^*(T_n)$ and thus a way of getting appropriate finite sections. Of course, actually finding these approximating T_n 's and corresponding finite sections in specific examples does not appear to be easy and hence we have little more to say about it at this point. However, inspired by the observations of this section, an effective way of computing spectra, in irrational rotation algebras is worked out in [6].

7. QUASI-DIAGONAL UNILATERAL BAND OPERATORS

Though the results of the previous sections are quite satisfactory from a theoretical point of view, some practical numerical analytic questions have gone undressed. For example, if an explicitly given operator is known to be quasi-diagonal, how does one explicitly write down asymptotically commuting finite rank projections? Can one control the ranks of these projections? In the generality of the previous sections these questions cannot be answered (even for a general self-adjoint

operator this is probably hopeless). However, in special cases some progress can be made, and this is the subject of the final two sections of this paper.

In this section we consider certain band operators. As mentioned before, not all band operators are quasi-diagonal (e.g., unilateral shift), but a fairly broad class of band operators, described below, can be attacked with this technology.

Our Hilbert space will now be $l^2(\mathbb{N})$ with canonical orthonormal basis $\{\delta_i\}_{i \in \mathbb{N}}$. We will regard $l^\infty(\mathbb{N})$ as acting on $l^2(\mathbb{N})$ by multiplication operators (i.e., $f \in l^\infty(\mathbb{N})$ acts on $l^2(\mathbb{N})$ by sending δ_i to $f(i)\delta_i$). $T \in B(l^2(\mathbb{N}))$ will be a weighted shift operator with weight sequence $(\alpha_1, \alpha_2, \dots)$; i.e., $(\alpha_1, \alpha_2, \dots)$ is a bounded sequence of complex numbers and $T(\delta_i) = \alpha_i \delta_{i+1}$. Our basic assumption is:

assume the existence of a subsequence $\{i_k\}_{k \in \mathbb{N}}$ such that $|\alpha_{i_k}| \rightarrow 0$ as $k \rightarrow \infty$.

Note that a unilateral weighted shift is quasi-diagonal if and only if the assumption above holds. Below we will observe that this assumption implies quasi-diagonality, while the absence of such a subsequence implies that the weighted shift is a Fredholm operator with index -1 and, hence, is not quasi-diagonal.

The band operators considered will be of the form

$$S = f_{-m}(T^*)^m + f_{-m+1}(T^*)^{m-1} + \cdots + f_{-1}T^* + f_0 + f_1T \\ + \cdots + f_{m-1}T^{m-1} + f_mT^m,$$

where $f_j \in l^\infty(\mathbb{N})$, $-m \leq j \leq m$, and m is a positive integer. Note that S is a band operator whose main diagonal is just f_0 . The off diagonal entries of S are somewhat complicated to describe, but the “coefficients” f_j are arbitrary and hence operators of the form above are still are rather large class. The reader may find it useful to consider a weight sequence of only zeroes and ones to get a feel for the kinds of operators arising this way. In fact, any band operator whose matrix has “butterflies” tending to zero can be handled with the quasi-diagonal approach. In other words, any band operator which is a compact perturbation of a block diagonal band operator can be handled this way, but rates of convergence can be complicated to state precisely, so we stick to the situation described above.

For ease of notation we will let $q(x) = x^m + x^{m-1} + \cdots + x$ and let q' be the derivative of this polynomial. We will also let $M = \max_{-m \leq j \leq m} \{\|f_j\|\}$. Finally, we define finite rank projections P_k as the orthogonal projections onto the span of $\{\delta_1, \dots, \delta_{i_k}\}$.

Theorem 7.1. *Let T , S , and $\{P_k\}$ be as above. Then the following statements hold:*

- (1) *If S is invertible and $|\alpha_{i_k}| < \frac{1}{\|S^{-1}\|Mq'(\|T\|)}$, then P_kSP_k is invertible.*
- (2) *If S is invertible, $|\alpha_{i_k}| < \frac{1}{\|S^{-1}\|Mq'(\|T\|)}$, $y \in l^2(\mathbb{N})$ is given, $x \in l^2(\mathbb{N})$ is the unique solution to the equation $Sx = y$, $y_k = P_k y$, and x_k is the unique solution to $P_kSP_k x_k = y_k$, then*

$$\|x_k - x\| \leq \|S^{-1}\|(\|y - y_n\| + \|x_n\|Mq'(\|T\|)|\alpha_{i_k}|).$$

- (3) *For every $\varepsilon > 0$, the ε -pseudospectra of P_kSP_k converge to the ε -pseudospectrum of S and $\sigma^{(\varepsilon)}(P_nSP_n) \subset \sigma^{(\varepsilon + Mq'(\|T\|)|\alpha_{i_k}|)}(S)$.*
- (4) *The singular values of P_kSP_k converge to the singular values of S and $\sigma_2(P_nTP_n) \subset \delta_n \sigma_2(T)$, where $\delta_n = \sqrt{2\|T\|Mq'(\|T\|)|\alpha_{i_k}|}$.*

Proof. First note that $\|[P_k, T]\| = \|[P_k, T^*]\| = |\alpha_{i_k}| \rightarrow 0$ as $k \rightarrow \infty$. Using the estimates $\|[P, T^j]\| \leq \|PT^j - TPT^{j-1}\| + \|TPT^{j-1} - T^jP\| \leq \|[P, T]\| \|T\|^{j-1} + \|T\| \|[P, T^{j-1}]\|$, a simple induction argument shows that

$$\|[P_k, T^j]\| \leq j \|T\|^{j-1} \|[P_k, T]\| = j \|T\|^{j-1} |\alpha_{i_k}|.$$

Since S is a band operator whose lower diagonals are determined by the powers of T and upper diagonals are determined by the powers of T^* , we may use the fact that $\|[P, S]\| = \max\{\|PSP^\perp\|, \|P^\perp SP\|\}$ to arrive at the following inequality:

$$\|[P_k, S]\| \leq Mq'(\|T\|) |\alpha_{i_k}|.$$

With this inequality in hand, the various statements above follow from the corresponding results in Sections 3 and 4. \square

Remark 7.2. Note that the rank of P_k is i_k , and hence P_kSP_k is an $i_k \times i_k$ matrix. Thus the numerical efficiency in this setting is directly tied to the rate of growth of $\{i_k\}$ and the rate of decay of $\{\alpha_{i_k}\}$ (as functions of k).

Remark 7.3. Evidently the theorem above can be extended with no difficulty to the case of bilateral shifts on $l^2(\mathbb{Z})$, where one assumes the existence of subsequences tending to zero both in the positive and negative directions.

Remark 7.4. Of course, the numerical ranges of P_kSP_k also converge to the numerical range of S since the proof above shows that $\|[P_k, S]\| \rightarrow 0$. Note also that a similar result can be worked out for any element in $C^*(T)$, and (complicated) error estimates can be explicitly given for any polynomial in T and T^* .

Remark 7.5. Let $f \in C(\mathbb{T})$ be a continuous function on the unit circle and let T_f be the corresponding Toeplitz operator on $l^2(\mathbb{N})$. Of course, T_f is rarely a band operator, and the diagonals of the matrix of T_f never have subsequences tending to zero (since they are constant). However, it follows from BDF theory that T_f is quasi-diagonal if and only if the winding number of the curve $f(\mathbb{T})$ is zero about every point in the complement of $f(\mathbb{T})$ (cf. [9, Sections V.1 and IX.7]). In this case there are asymptotically commuting finite rank projections out there, and we feel it would be interesting (but perhaps hard?) to find them.

We conclude this section with an illustration of the usefulness of this method. Namely, we will give a new proof of the fact that the spectra of quasi-diagonal weighted shifts is as large as possible. It is well known that if T is any weighted shift and $\rho(T)$ is the spectral radius of T , then $\sigma(T) = \{\lambda \in \mathbb{C} : |\lambda| \leq \rho(T)\}$. While our technology will only handle the quasi-diagonal case, the argument is simple and cute, so we include a proof. It will be convenient to note that

$$\rho(T) = \lim_{l \rightarrow \infty} \|T^l\|^{1/l} = \lim_{l \rightarrow \infty} \left(\sup_{m \in \mathbb{N}} |\alpha_m \alpha_{m+1} \cdots \alpha_{m+l}| \right)^{1/l},$$

whenever T is a weighted shift with weight sequence $(\alpha_1, \alpha_2, \dots)$.

Proposition 7.6. *Let T be the weighted shift described at the beginning of this section. Then $\sigma(T) = \{\lambda \in \mathbb{C} : |\lambda| \leq \rho(T)\}$.*

Proof. Evidently we may assume that $\|T\| \leq 1$ by rescaling, if necessary.

Since the spectrum of T can be no larger than claimed, it suffices, thanks to part (2) of Theorem 3.5, to show that if

$$|\lambda| < \lim_{l \rightarrow \infty} \left(\sup_{m \in \mathbb{N}} |\alpha_m \alpha_{m+1} \cdots \alpha_{m+l}| \right)^{1/l},$$

then for every $\varepsilon > 0$, λ belongs to the ε -pseudospectra of $P_k T P_k$ for all sufficiently large k . In fact, we will show a bit more (as it requires no extra effort and simplifies notation). Let P_n be the projection onto the span of the first n basis vectors of $l^2(\mathbb{N})$, and we will show that λ belongs to the ε -pseudospectra of $P_n T P_n$ for all sufficiently large n .

First note that for all nonzero λ the matrix $\lambda - P_n T P_n$ is invertible. A straightforward computation shows that its inverse is given by the following matrix:

$$\begin{pmatrix} \frac{1}{\lambda} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \frac{\alpha_1}{\lambda^2} & \frac{1}{\lambda} & 0 & 0 & \cdots & 0 & 0 \\ \frac{\alpha_1 \alpha_2}{\lambda^3} & \frac{\alpha_2}{\lambda^2} & \frac{1}{\lambda} & 0 & \cdots & 0 & 0 \\ \frac{\alpha_1 \alpha_2 \alpha_3}{\lambda^4} & \frac{\alpha_2 \alpha_3}{\lambda^3} & \frac{\alpha_3}{\lambda^2} & \frac{1}{\lambda} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\prod_1^{n-1} \alpha_i}{\lambda^n} & \frac{\prod_2^{n-1} \alpha_i}{\lambda^{n-1}} & \frac{\prod_3^{n-1} \alpha_i}{\lambda^{n-2}} & \frac{\prod_4^{n-1} \alpha_i}{\lambda^{n-3}} & \cdots & \frac{\alpha_{n-1}}{\lambda^2} & \frac{1}{\lambda} \end{pmatrix}.$$

Our goal is to show that if $0 < |\lambda| < \lim_{l \rightarrow \infty} \left(\sup_{m \in \mathbb{N}} |\alpha_m \alpha_{m+1} \cdots \alpha_{m+l}| \right)^{1/l} \leq 1$, then for every $\varepsilon > 0$ the norms of the matrices above are eventually bigger than $1/\varepsilon$. To prove this it suffices to show that the modulus of some entry in the matrices above is greater than $1/\varepsilon$. This, however, is the case because if $|\lambda| < |\alpha_m \alpha_{m+1} \cdots \alpha_{m+l}|^{1/l} - \delta$, then

$$(1 + \delta)^l < \frac{|\alpha_m \alpha_{m+1} \cdots \alpha_{m+l}|}{|\lambda|^l} < \frac{|\alpha_m \alpha_{m+1} \cdots \alpha_{m+l}|}{|\lambda|^{l+1}}.$$

Hence, if $\varepsilon > 0$ is given, we choose l big enough that $(1 + \delta)^l > 1/\varepsilon$, and then we will be able to find a string $\alpha_m, \dots, \alpha_{m+l}$ such that

$$\frac{1}{\varepsilon} < \frac{|\alpha_m \alpha_{m+1} \cdots \alpha_{m+l}|}{|\lambda|^{l+1}}.$$

However, the numbers

$$\frac{\alpha_m \alpha_{m+1} \cdots \alpha_{m+l}}{\lambda^{l+1}}$$

are entries in all of the matrices above (for sufficiently large n). \square

Remark 7.7. Note that the proof above adapts easily to bilateral weighted shifts which have subsequences tending to zero in both the positive and negative directions.

Remark 7.8. If T denotes the bilateral shift on $l^2(\mathbb{Z})$ and P_n denotes the orthogonal projection onto the span of $\{\delta_{-n}, \dots, \delta_n\} \subset l^2(\mathbb{Z})$, then the finite sections $P_n T P_n$ are all nilpotent matrices and hence have spectrum $\{0\}$ for all n (which, in particular, do not converge to the spectrum of T). Moreover, the proof above shows that for every $\varepsilon > 0$, $\liminf \sigma^{(\varepsilon)}(P_n T P_n)$ contains the entire unit disc (see [11, Definition 3.1]). Hence we cannot recover the spectrum of T (i.e., the unit circle) by intersecting over $\varepsilon > 0$ as in part (2) of Theorem 3.5.

Note that S^t is unitarily equivalent to the transpose of S (reversing the order and identifying the positive integers with the negative ones); hence the spectrum of the dilation is the same as that of S . Evidently we could perform a similar trick with an arbitrary banded Toeplitz operator, and the resulting direct sum would have diagonals which are block balanced in the desired sense. Since this procedure does not change the spectrum and it is evident that linear systems (only involving the original Toeplitz operator) can be extended to the dilation, it follows that our technology can be applied to banded Toeplitz operators (even though these are often non-quasi-diagonal).

Let us come back to the general framework we were considering before. Since we have assumed $\{f_{-p}, \dots, f_{-1}, f_0, f_1, \dots, f_p\}$ to be block balanced we can, for any sequence $\varepsilon_k \rightarrow 0$ and natural numbers $s_k \rightarrow \infty$, find integers m_k, n_k such that $m_k + s_k \leq 0 < n_k$ and

$$|f_l(m_k + j) - f_l(n_k + j)| < \varepsilon_k,$$

for all $-p \leq l \leq p$ and $0 \leq j \leq s_k$. Define constants $\alpha_{j,k} = \frac{1}{2}(1 - \exp \frac{(j+1)\pi}{s_k+2})$ and $\beta_{j,k} = \frac{1}{2}(1 + \exp \frac{(j+1)\pi}{s_k+2})$ for $0 \leq j \leq s_k$. A calculation shows that the set of vectors $\{\alpha_{0,k}\delta_{m_k} + \beta_{0,k}\delta_{n_k}, \dots, \alpha_{s_k,k}\delta_{m_k+s_k} + \beta_{s_k,k}\delta_{n_k+s_k}\}$ is an orthonormal set. Hence the set

$$V_k = \{\delta_i : m_k + s_k < i < n_k\} \cup \{\alpha_{0,k}\delta_{m_k} + \beta_{0,k}\delta_{n_k}, \dots, \alpha_{s_k,k}\delta_{m_k+s_k} + \beta_{s_k,k}\delta_{n_k+s_k}\}$$

is also orthonormal.

Though it is not obvious, it turns out that the orthogonal projection onto the subspace spanned by V_k almost commutes with T . We will denote this projection by P_k .

Lemma 8.3. *Keeping notation as above we have*

$$\|[T, P_k]\| \leq \frac{(2p+1)\varepsilon_k}{2} + \frac{\pi}{s_k+2} \sum_{l=-p}^p |l| \|f_l\|.$$

Proof. Our first task will be to write down the 5×5 block matrix of P_k with respect to the decomposition $l^2(\mathbb{Z}) = H_1 \oplus H_2 \oplus H_3 \oplus H_4 \oplus H_5$, where $H_1 = \text{span}\{\delta_j : j < m_k\}$, $H_2 = \text{span}\{\delta_j : m_k \leq j \leq m_k + s_k\}$, $H_3 = \text{span}\{\delta_j : m_k + s_k < j < n_k\}$, $H_4 = \text{span}\{\delta_j : n_k \leq j \leq n_k + s_k\}$, and $H_5 = \text{span}\{\delta_j : n_k + s_k < j\}$.

A tedious but straightforward calculation shows that in this decomposition P_k becomes

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & A_k & 0 & C_k & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & C_k^* & 0 & B_k & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where I is the identity matrix,

$$A_k = \begin{pmatrix} |\alpha_0|^2 & 0 & \cdots & 0 \\ 0 & |\alpha_1|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |\alpha_{s_k}|^2 \end{pmatrix}, \quad B_k = \begin{pmatrix} |\beta_0|^2 & 0 & \cdots & 0 \\ 0 & |\beta_1|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |\beta_{s_k}|^2 \end{pmatrix},$$

and

$$C_k = \begin{pmatrix} \overline{\beta_0}\alpha_0 & 0 & \cdots & 0 \\ 0 & \overline{\beta_1}\alpha_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \overline{\beta_{s_k}}\alpha_{s_k} \end{pmatrix},$$

where $\overline{\beta_j}$ is the complex conjugate of β_j . Since the matrix of each f_l is also block diagonal with respect to the above decomposition we may write

$$f_l = \begin{pmatrix} F_1^l & 0 & 0 & 0 & 0 \\ 0 & F_2^l & 0 & 0 & 0 \\ 0 & 0 & F_3^l & 0 & 0 \\ 0 & 0 & 0 & F_4^l & 0 \\ 0 & 0 & 0 & 0 & F_5^l \end{pmatrix}.$$

Hence

$$\begin{aligned} \|[P_k, f_l]\| &= \max\{\|F_2^l C_k - F_4^l C_k\|, \|F_2^l C_k^* - F_4^l C_k^*\|\} \\ &\leq \varepsilon_k \max\{|\overline{\beta_0}\alpha_0|, \dots, |\overline{\beta_{s_k}}\alpha_{s_k}|\} \leq \frac{\varepsilon_k}{2}. \end{aligned}$$

Unfortunately, the matrix of the bilateral shift U is a bit messy with respect to the decomposition we have used above. However, conjugating by U (i.e., considering $U \cdot U^*$) just has the affect of shifting the matrix down the diagonal one entry. This remark, together with the fact that $\|[U, P_k]\| = \|UP_k U^* - P_k\|$, allows one to estimate $\|[U, P_k]\|$ above by

$$\max\{||\alpha_j|^2 - |\alpha_{j+1}|^2|, ||\beta_j|^2 - |\beta_{j+1}|^2|\} + \max\{|\overline{\beta_j}\alpha_j - \overline{\beta_{j+1}}\alpha_{j+1}|\}.$$

But $|\alpha_j|^2 = \frac{1}{2}(1 - \cos(\frac{j\pi}{s_k+2}))$, $|\beta_j|^2 = \frac{1}{2}(1 + \cos(\frac{j\pi}{s_k+2}))$, and $\overline{\beta_j}\alpha_j = -\frac{i}{2}\sin(\frac{j\pi}{s_k+2})$. Hence we get the inequality

$$\|[U, P_k]\| \leq \frac{\pi}{s_k+2}.$$

Having the bounds on $\|[P_k, f_l]\|$ and $\|[U, P_k]\|$ one completes the proof by a standard interpolation argument, similar to the proof of Theorem 7.1. \square

Remark 8.4. We leave it to the reader to check that $P_k(\delta_i) \rightarrow \delta_i$, as $k \rightarrow \infty$, for each $i \in \mathbb{Z}$. It will not always be the case that $P_k \leq P_{k+1}$, but the fact that $\|P_k(v) - v\| \rightarrow 0$ is good enough for all of the results in sections 3, 4, and 5 to hold (i.e., we really do not have to have an honest filtration in any of the proofs).

Remark 8.5. Note that the rank of P_k is $n_k - m_k$, and hence the numerical efficiency of this scheme will depend on how far out one must go in order to find long strings which are close together.

A number of interesting examples naturally give rise to tri-diagonal operators on $l^2(\mathbb{Z})$. For example, a natural discretization of $\frac{d}{dx}$ is $\frac{1}{2h}(U - U^*)$, where h is a fixed step size. A natural discretization of the one-dimensional differential operator $(-\Delta + c)f = \frac{d^2}{dx^2}f + cf$, for some constant c , would be $\frac{-1}{4h^2}(U + U^*) + (c + \frac{1}{2h^2})I$, where $I \in l^\infty(\mathbb{Z})$ is the constant function $I(n) = 1$ for all n . Also, Arveson's method of discretizing Schrödinger operators of one-dimensional quantum systems (cf. [1]) naturally leads to an analysis of operators of the form $U + U^* + f$, where $f(n) = v(\cos(n\theta))$ for some continuous function v on the interval $[-1, 1]$ and some fixed positive number θ . Note that if $\frac{\theta}{\pi} \in \mathbb{Q}$, then f will be a *periodic* function

while in general it will be “almost” periodic and hence block balanced. Hopefully these examples justify spending the remainder of this paper analyzing how to implement Berg’s technique on the class of tri-diagonal operators with periodic main diagonal and constant diagonals above and below the main diagonal. The interested reader should have no trouble now formulating a general result based on the previous lemma and the results of Section 4. However, while this method can be horribly inefficient in some cases, it works very well in the periodic tri-diagonal case. Moreover, for those interested in discretized Hamiltonians with $\frac{\theta}{\pi} \notin \mathbb{Q}$ we will describe how to use continued fractions to find good “blocks” at the very end. (See also [6] for improved spectral computations in this case.)

For the remainder of this paper T will be an operator of the form

$$T = c_1 U + c_{-1} U^* + f,$$

where $c_1, c_{-1} \in \mathbb{C}$ and $f \in l^\infty(\mathbb{Z})$ is a periodic function with period p (i.e., $f(n+p) = f(n)$, for all $n \in \mathbb{Z}$). Due to the periodicity of f and the constant coefficients on U and U^* we may take $\varepsilon_k = 0$, $s_k = pk - 1$, $m_k = -s_k$ and $n_k = 1$ (notation as in the setup for Lemma 8.3). Note that with these choices the “twisted” basis used in Berg’s technique is

$$V_k = \{\alpha_{0,k} \delta_{m_k} + \beta_{0,k} \delta_1, \dots, \alpha_{s_k,k} \delta_0 + \beta_{s_k,k} \delta_{s_k}\}.$$

A tedious but straightforward computation shows that if P_k is the orthogonal projection onto the span of V_k , then the matrix of $P_k T P_k$ with respect to Berg’s basis V_k is

$$P_k T P_k = \begin{pmatrix} \sigma_{0,0} & \sigma_{0,1} & 0 & 0 & \cdots & 0 & \sigma_{0,s_k} \\ \sigma_{1,0} & \sigma_{1,1} & \sigma_{1,2} & 0 & \cdots & 0 & 0 \\ 0 & \sigma_{2,1} & \sigma_{2,2} & \sigma_{2,3} & \cdots & 0 & 0 \\ 0 & 0 & \sigma_{3,2} & \sigma_{3,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \sigma_{s_k-1,s_k-1} & \sigma_{s_k-1,s_k} \\ \sigma_{s_k,0} & 0 & 0 & 0 & \cdots & \sigma_{s_k,s_k-1} & \sigma_{s_k,s_k} \end{pmatrix},$$

where

$$\begin{aligned} \sigma_{j,j} &= |\alpha_{j,k}|^2 f(m_k + j) + |\beta_{j,k}|^2 f(1 + j), \\ \sigma_{j,j+1} &= (\alpha_{j+1,k} \overline{\alpha_{j,k}} + \beta_{j+1,k} \overline{\beta_{j,k}}) c_{-1}, \\ \sigma_{j+1,j} &= (\alpha_{j,k} \overline{\alpha_{j+1,k}} + \beta_{j,k} \overline{\beta_{j+1,k}}) c_1, \\ \sigma_{0,s_k} &= \overline{\beta_{0,k}} \alpha_{s_k,k} c_1, \\ \sigma_{s_k,0} &= \beta_{0,k} \overline{\alpha_{s_k,k}} c_{-1}. \end{aligned} \tag{8.1}$$

While it is unfortunate that the entries of the matrix $P_k T P_k$ are more complicated than the entries of T , it is a significant feature of Berg’s technique that the property of being tri-diagonal (hence sparse) is almost preserved. Also note that $\|[T, P_k]\| \leq \frac{2\pi}{pk+1}$ in this setting while the rank of P_k is pk . Hence the error estimates are inversely proportional to the rank of the “almost” tri-diagonal matrix $P_k T P_k$.

We end this paper by putting all the error estimates together in the case we are considering. The proof is immediate from Lemma 8.3 and the results of Sections 3, 4, and 5.

Theorem 8.6. *Let $T, P_k \in B(l^2(\mathbb{Z}))$ be as above.*

(1) *Let $y \in l^2(\mathbb{Z})$ be given and assume that T is invertible. If*

$$k > \frac{1}{p}(2\pi\|T^{-1}\| - 1),$$

then the “almost” tri-diagonal matrix P_kTP_k is invertible, and if x_k is the unique solution to $P_kTP_kx_k = P_ky$, then

$$\|x - x_k\| \leq \|T^{-1}\|(\|y - P_ky\| + \frac{2\pi\|x_k\|}{pk+1}),$$

where x is the unique solution to $Tx = y$.

(2) *For every $\varepsilon > 0$, $\sigma^{(\varepsilon)}(P_kTP_k) \rightarrow \sigma^{(\varepsilon)}(T)$ and, moreover, $\sigma^{(\varepsilon)}(P_kTP_k) \subset \sigma^{(\varepsilon + \frac{2\pi}{pk+1})}(T)$.*

(3) *$\sigma_2(P_kTP_k) \rightarrow \sigma_2(T)$ and $\sigma_2(P_kTP_k) \subset \sqrt{\frac{4\pi\|T\|}{pk+1}} \sigma_2(T)$.*

(4) *If $T = T^*$, then $\sigma(P_kTP_k) \rightarrow \sigma(T)$ and $\sigma(P_kTP_k) \subset \frac{2\pi}{pk+1} \sigma(T)$.*

Remark 8.7. The theory of continued fractions gives a remarkable approximation property of irrational numbers (cf. [12]). Namely, if $\alpha \in (0, 1)$ is an irrational number, then there is a very simple and computable algorithm for finding coprime integers p_n, q_n such that $|\alpha - \frac{p_n}{q_n}| \leq \frac{1}{q_n^2}$. It follows that for any $0 \leq j \leq q_n$ we have the inequality $|j\alpha - j\frac{p_n}{q_n}| \leq \frac{j}{q_n} \leq \frac{1}{q_n}$. Hence we can apply Berg’s technique as above to a periodic approximation to a function of the form $f(m) = \cos(m\theta)$ and still have good control on the error estimates.

ACKNOWLEDGMENTS

I thank Bill Arveson and Marc Rieffel for sparking my interest in numerical analysis and, more generally, for providing more support and inspiration to me than they probably realize. I am also indebted to the authors of [11] and [3], as they have made a large body of work readily accessible to the C^* -community. Finally, I thank my colleague Ludmil Zikatanov for numerous helpful conversations.

REFERENCES

1. W. Arveson, *The role of C^* -algebras in infinite dimensional numerical linear algebra*, C^* -algebras: 1943-1993 (San Antonio, TX, 1993), 114–129, Contemp. Math. **167**, Amer. Math. Soc., Providence, RI, 1994. MR1292012 (95i:46084)
2. I.D. Berg, *An extension of the Weyl-von Neumann theorem to normal operators*, Trans. Amer. Math. Soc. **160** (1971), 365–371. MR0283610 (44:840)
3. A. Böttcher, *C^* -algebras in numerical analysis*, Irish Math. Soc. Bull. No. 45 (2000), 57–133. MR1832325 (2002b:46116)
4. N.P. Brown, *On quasi-diagonal C^* -algebras*, Adv. Stud. Pure Math., vol. 38, Math. Soc. Japan, Tokyo, 2004, pp. 19–64. MR2059800 (2005a:46107)
5. N.P. Brown, *Invariant means and finite representation theory of C^* -algebras*, Mem. Amer. Math. Soc. (accepted).
6. N.P. Brown, *AF embeddings and the numerical computation of spectra in irrational rotation algebras*, preprint.
7. N.P. Brown, *Herrero’s approximation problem for quasi-diagonal operators*, J. Funct. Anal. **186** (2001), 360–365. MR1864827 (2002j:47034)
8. M. Dadarlat, *On the approximation of quasi-diagonal C^* -algebras*, J. Funct. Anal. **167** (1999), 69–78. MR1710645 (2000f:46069)
9. K.R. Davidson, *C^* -algebras by example*, Fields Institute Monographs 6, American Mathematical Society, Providence, RI, 1996. MR1402012 (97i:46095)

10. M. Embree and L.N. Trefethen, *Pseudospectra Gateway*: <http://www.comlab.ox.ac.uk/pseudospectra>.
11. R. Hagen, S. Roch and B. Silbermann, *C*-algebras and numerical analysis*. Monographs and Textbooks in Pure and Applied Mathematics, 236. Marcel Dekker, Inc., New York, 2001. MR1792428 (2002g:46133)
12. G.H. Hardy and E.M. Wright, *An introduction to the theory of numbers*. Fifth edition. The Clarendon Press, Oxford University Press, New York, 1979. MR0568909 (81i:10002)
13. P.R. Halmos, *Ten problems in Hilbert space*, Bull. Amer. Math. Soc. **76** (1970), 887–933. MR0270173 (42:5066)
14. E. Kirchberg, *Exact C*-algebras, tensor products and the classification of purely infinite algebras*, Proceedings of the International Congress of Mathematicians, Vols. 1, 2 (Zurich, 1994), 943–954. MR1403994 (97g:46074)
15. S. Narayan, *Quasi-diagonality of direct sums of weighted shifts*, Trans. Amer. Math. Soc. **332** (1992), 757–774. MR1012511 (92j:47048)
16. R. Smucker, *Quasi-diagonal weighted shifts*, Pacific J. Math. **98** (1982), 173–182. MR0644948 (83c:47045)
17. S. Szarek, *An exotic quasi-diagonal operator*, J. Funct. Anal. **89** (1990), 274–290. MR1042211 (91e:47015)
18. D.V. Voiculescu, *Around quasi-diagonal operators*, Integr. Equ. and Op. Thy. **17** (1993), 137–149. MR1220578 (94e:47029)
19. D.V. Voiculescu, *A note on quasi-diagonal operators*, Operator Theory: Advances and Applications, Vol. 32, Birkhauser Verlag, Basel, 1988, 265–274. MR0951964 (89h:47029)
20. S. Wassermann, *Exact C*-algebras and related topics*, Lecture Notes Series no. 19, GARC, Seoul National University, 1994. MR1271145 (95b:46081)

DEPARTMENT OF MATHEMATICS, PENN STATE UNIVERSITY, STATE COLLEGE, PENNSYLVANIA 16802

E-mail address: `nbrown@math.psu.edu`