Padé approximation of matrix-valued series of Stieltjes

Björn von Sydow

1. Introduction

Let H be a finite-dimensional Hilbert space and let $L(H)$ denote the set of bounded linear operators on H. Let $A:[0, \infty[\to L(H)$ be a bounded, non-decreasing function (this is given a precise meaning in Section 2). Put $C_n=\int_0^\infty t^n dA(t)$ and suppose that $C_n \in L(H)$ for all $n \geqq 0$. Under these assumptions we say that $F(z)=\sum_{n=0}^\infty C_n(-z)^n$ is an $L(H)$-valued series of Stieltjes. This paper is concerned with the existence and convergence of Padé approximants to F. The investigation was motivated by two facts:

(1) When H is one-dimensional, i.e. for complex-valued series of Stieltjes a fairly complete convergence theory for Padé approximation exists (see e.g. Karlsson — von Sydow [2]);

(2) an analogous theory in the present setting would find applications in various branches of theoretical physics (see e.g. Zinn–Justin [8]). The convergence theory in the complex-valued case can be summarized as follows:

Theorem A. *Let $\alpha(t)$ be a real, bounded, non-decreasing function on $t \geqq 0$ such that $C_n=\int_0^\infty t^n d\alpha(t)$ are finite for all $n \geqq 0$. Put $f(z)=\sum_{n=0}^\infty C_n(-z)^n$ and let $f[m, n](z)$ denote the $[m, n]$ Padé approximant to f. Let $j \geqq -1$ be an integer. Then*

(a) *If $\sum_{n=0}^\infty C_n^{-1/2n}=\infty$ then*

$$f[n+j, n](z) \to \int_0^\infty \frac{d\alpha(t)}{1+zt} \quad \text{as} \quad n \to \infty$$

uniformly on compact subsets of the complex plane cut along the negative real axis.

(b) *If $\alpha(t)$ is constant for $t \geqq 1$ then for $z \notin]-\infty, -1]$*

$$\limsup_{n\to\infty} \left| f[n+j, n](z) - \int_0^1 \frac{d\alpha(t)}{1+zt} \right|^{1/2n} \leqq |(\sqrt{1+z}-1)/(\sqrt{1+z}+1)|$$

where $\sqrt{\ }$ denotes the principal branch of the square root. Furthermore, the expression to the right of the inequality sign cannot in general be replaced by a smaller number.

The principal result of this paper (Theorem 1) is that an exactly analogous theorem is true in the present setting.

The problem treated here has been discussed by Zinn–Justin [8] and Allen—Narcowich [1]. A related problem is treated in Mac Nerney [3]. All these authors make the assumption that the interval of integration is compact, i.e. the analogy of (*b*) above. None of them discusses degree of convergence. A detailed comparison with the work of Zinn–Justin is difficult to make since his assumptions on H are not explicitly stated and some calculations seem to be purely formal. However, in [8] Zinn–Justin states that convergence of the sequences of $[n-1, n]$ and $[n, n]$ Padé approximants can be proved. As to Allen—Narcowich, they state uniform convergence of some kind of (unspecified) mixed approximant with H a separable space. No proof is offered. Mac Nerney, finally, proves rigorously strong convergence of a continued fractions expansion in an arbitrary Hilbert space.

In Section 2 we give necessary definitions and state the main result. In Section 3 we cite the remarkable theorem by Naimark [4] in which $A(t)$ is exhibited as the projection onto H of an orthogonal resolution of the identity in a space containing H. Naimark's theorem is the main tool in the proof of Theorem 1 which is given in Section 5. The proof is preceded in Section 4 by some computational lemmas which have been separated from the proof Theorem 1 in order to keep the latter more easy to survey. In Section 6, finally, we discuss the difficulties that arise when H is not restricted to be finite-dimensional.

2. Definitions and the main result

H is a complex Hilbert space and $L(H)$ the set of bounded linear operators on H. With H' another Hilbert space $L(H, H')$ denotes the bounded linear transformations from H into H'. The inner product in H is denoted $\langle\ ,\ \rangle$. Norms, adjoints, self-adjoint and positive operators are defined as in any standard text on functional analysis (see e.g. Riesz—Sz. Nagy [5]). We shall also encounter positive, unbounded operators. Such an operator T cannot be defined on the whole space H and we shall denote its domain by $D(T)$. By a subspace H_1 of H we mean a closed subspace. The restriction of an operator T to H_1 is denoted $T|H_1$. If $\{H_i\}_{i=1}^n$ are pair-wise orthogonal subspaces, their direct sum is denoted by $\sum_{i=1}^n \oplus H_i$. Operators defined on a direct sum of subspaces will be displayed in matrix notation with obvious interpretation. The letter I will always denote the identity operator on a space which should be clear from the context. $R(T)$ and $N(T)$, finally, will denote the range and null space of the operator T.

With these conventions understood we are ready to define three concepts which will be basic to our subject.

Definition 1. *A bounded normalized $L(H)$-increasing function is a function $A:[0, \infty[\to L(H)$ such that $A(0)=0$, $A(t)$ is strongly continuous to the right, $\|A(t)\| \leqq M < \infty$ for some M and all t and $A(t_2)-A(t_1)$ is a non-negative operator if $t_2 \geqq t_1$.*

Definition 2. *An $L(H)$-valued series of Stieltjes is a formal power series $F(z)=\sum_{n=0}^{\infty} C_n(-z)^n$ where $C_n \in L(H)$ and $C_n=\int_0^\infty t^n dA(t)$ for some bounded normalized $L(H)$-increasing function A in the sense that $\langle C_n x, x\rangle=\int_0^\infty t^n d\langle A(t)x, x\rangle$ for all $x \in H$. The Stieltjes transform of A is the function $\tilde{F}(z)=\int_0^\infty (1+zt)^{-1} dA(t)$.*

Definition 3. *Let $G(z)=\sum_{n=0}^{\infty} T_n z^n$ where $T_n \in L(H)$ be a formal power series and m, n be non-negative integers. A right $[m, n]$ Padé approximant is a function $P_R Q_R^{-1}$ such that*

(i) *P_R and Q_R are polynomials of degree at most m and n with coefficients in $L(H)$,*

(ii) *$Q_R(z)$ is invertible for at least one z,*

(iii) *the formal expression $F(z)Q_R(z)-P_R(z)$ contains only terms of degree greater than $m+n$ when expanded in powers of z. A left approximant is defined analogously with the order of F and Q interchanged.*

In contrast to the complex-valued case, right and left Padé approximants to $L(H)$-valued power series do not always exist and if they do they need not be unique. A simple example is

$$F(z) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} z + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} z^2$$

which has no right [1, 1] approximant but infinitely many left [1, 1] approximants. If, however, $G(z)$ for some m and n has both a right approximant $P_R Q_R^{-1}$ and a left approximant $Q_L^{-1} P_L$ such that both these approximants are holomorphic in some open set Ω, then these must be identical and thus unique. This can be proved in the following way: (iii) implies

$$Q_L(z)F(z)Q_R(z)-Q_L(z)P_R(z) = O(z^{n+m+1}).$$

Subtract this from the analogous expression for the left approximant. The result is

$$Q_L(z)P_R(z)-P_L(z)Q_R(z) = O(z^{n+m+1}).$$

But the left part is a polynomial of degree less than or equal to $m+n$ and is thus identically zero. The equality of the right and left approximant now follows from the uniqueness theorem for holomorphic functions. We note that if H is finite-dimensional, each approximant is holomorphic in the complex plane except at finitely many points, so in this situation the existence of a left and right approximant implies their uniqueness. We reserve the notation $G[m, n](z)$ for such unique left and right approximants.

We shall not go into details of how to calculate these approximants but be content with the observation that if H is finite-dimensional with dimension N, a right $[n, n]$ approximant can be calculated by solving a set of linear equations systems requiring in the order of $N^3 \cdot \left(\frac{n^3}{3}+\frac{3}{2}n^2\right)$ arithmetical operations to be solved.

We are now in a position to state our main result.

Theorem 1. *Let H be a finite-dimensional Hilbert space. Let $A(t)$ be a bounded, normalized $L(H)$-increasing function and $F(z)=\sum_{n=0}^{\infty} C_n(-z)^n$ the corresponding $L(H)$-valued series of Stieltjes. Let $j \geqq -1$ be an integer. Then*

(a) *$F[n+j, n](z)$ exists for all $n \geqq 1$.*

(b) *If $\sum_{n=0}^{\infty} \|C_n\|^{-1/2n} = \infty$, then*

$$\|F[n+j, n](z)-\tilde{F}(z)\| \to 0 \quad \text{as} \quad n \to \infty,$$

uniformly on compact subsets of the complex plane cut along the negative real axis.

(c) *If A is constant for $t \geqq 1$ and $\tilde{F}(z)=\int_0^1 (1+zt)^{-1} dA(t)$, then for $z \notin]-\infty, -1]$*

$$\limsup_{n\to\infty} \|F[n+j, n](z)-\tilde{F}(z)\|^{1/2n} \leqq |(\sqrt{1+z}-1)/(\sqrt{1+z}+1)|,$$

where $\sqrt{}$ denotes the principal branch of the square root.

3. Naimark's theorem

The following remarkable result is due to Naimark [4] (a proof can also be found in Sz.-Nagy [7]).

Naimark's theorem. *Let H be a Hilbert space and $A(t)$ a normalized $L(H)$-increasing function satisfying $A(t) \to I$ strongly as $t \to \infty$. Then there exists a Hilbert space $K \supset H$ and an orthogonal resolution of the identity $\{E(t)\}_{t>0}$ on K such that*

$$\Pi_H E(t) | H = A(t) \quad \textit{for all} \quad t,$$

where Π_H denotes projection onto H. Furthermore, K and $\{E(t)\}$ can be chosen such that $E(t)$ is constant in any interval where $A(t)$ is constant.

This theorem is the main tool in the proof of Theorem 1, since it enables us to make use of the power of the spectral theorem for self-adjoint operators. Putting

$\mathbf{B}=\int_0^\infty t dE(t)$ it is seen that

$$\tilde{F}(z) = \Pi_H(I+\mathbf{B}z)^{-1}|\, H,$$

i.e. $\tilde{F}(z)$ is the projection of a rational function of type [0, 1] restricted to H. This relation will be exploited in the proof of Theorem 1.

4. Some lemmas

In this section we state and prove four lemmas which will be needed in the proof of Theorem 1. Each lemma is applied only once, but we feel that it is easier to maintain an over-all view of the proof of the Theorem 1 with these technicalities out of the way.

Lemma 1. *Let H_1 and H_2 be Hilbert spaces and let $T_{ij} \in L(H_j, H_i)$, $i, j=1, 2$. Define $T \in L(H_1 \oplus H_2)$ by*

$$T = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}.$$

Then, for small enough $|z|$,

$$\begin{pmatrix} I-T_{11}z & -T_{12}z \\ -T_{21}z & I-T_{22}z \end{pmatrix}^{-1} = \begin{pmatrix} S_{11}(z) & S_{12}(z) \\ S_{21}(z) & S_{22}(z) \end{pmatrix}$$

exists and is bounded. Furthermore,

$$S_{11}(z) = \left(I-T_{11}z-z^2T_{12}(I-T_{22}z)^{-1}T_{21}\right)^{-1}.$$

Proof: Since T_{ij} are bounded, so is T and for $|z|<\|T\|$

$$(I-Tz)^{-1} = \textstyle\sum_{k=0}^{\infty}(Tz)^k,$$

which proves the existence.

By performing the matrix multiplication $(I-Tz)\,(I-Tz)^{-1}$ we get

$$\begin{pmatrix} (I-T_{11}z)S_{11}-T_{12}z\cdot S_{21} & (I-T_{11}z)S_{12}-T_{12}z\cdot S_{21} \\ -T_{21}z\cdot S_{11}+(I-T_{22}z)S_{21} & -T_{21}z\cdot S_{12}+(I-T_{22}z)S_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

We shall need only the equality between the first columns of these matrices. The equality between the elements in position (1, 2) yields

$$S_{21} = (I-T_{22}z)^{-1}T_{21}S_{11}z.$$

Inserting this into the upper left-hand equality yields the desired expression for S_{11}. ∎

Our next lemma shows that if T_{21} in Lemma 1 is surjective, then $S_{11}(z)$ and analogously defined functions for direct sums of more than two spaces are actually rational functions.

Lemma 2. *Let $\{H_i\}_{i=1}^n$ be Hilbert spaces and $T_{ij}\in L(H_j, H_i)$, $i,j=1,2,\dots,n$, such that $T_{ij}=0$ if $|i-j|>1$ and $T_{i,i-1}$ is surjective for $i=2,3,\dots,n$. Define $T\in L(\sum_{i=1}^n \oplus H_i)$ by $\Pi_i T|H_j=T_{ij}$, where Π_i denotes projection onto H_i. Then $\Pi_1(I-Tz)^{-1}|H_1$ can be written in the form $R(z)\cdot(I-zQ(z))^{-1}$, where $R(z)$ and $Q(z)$ are polynomials of degree $\leqq n-1$.*

Proof: The proof is by induction on n. If $n=1$, choose $R\equiv I$, $Q\equiv T_{11}$. Suppose the lemma is true for $n<N$. Put $H'=\sum_2^N \oplus H_i$ and notice $\sum_1^N \oplus H_i=H_1\oplus H'$. With $T'=\Pi_{H'}T|H'$ we have

$$T=\left\{\begin{array}{c|c} T_{11} & T_{12}\ \ 0\ \ 0\dots 0 \\ \hline T_{21} & \\ 0 & \\ \cdot & T' \\ \cdot & \\ 0 & \end{array}\right\},$$

and by Lemma 1

$$\Pi_1(I-Tz)^{-1}|H_1=\left(I-T_{11}z-z^2T_{12}\Pi_2(I-zT')^{-1}|H_2T_{21}\right)^{-1}.$$

By the induction assumption $\Pi_2(I-zT')^{-1}|H_2=R(z)\cdot(I-zQ(z))^{-1}$ where R and Q have degree $\leqq N-2$.

Furthermore if we define $U\colon H_2\to H_1$, by $Ux=y$ where y is the unique element in $N(T_{21})^\perp$ such that $T_{21}y=x$, then U is bounded and $T_{21}U=I$, the identity operator on H_2. Thus

$$T_{21}\left(I-U\cdot zQ(z)T_{21}\right)^{-1}=T_{21}\cdot\sum_{k=0}^\infty U(zQ)^kT_{21}=\left(I-zQ(z)\right)^{-1}T_{21}.$$

This relation is exactly what is needed to complete the proof since it implies

$$\begin{aligned}&\left(I-T_{11}z-z^2T_{12}R(z)(I-zQ(z))^{-1}T_{21}\right)^{-1}\\ &=\left(I-zUQ(z)T_{21}\right)\left\{(I-T_{11}z)\left(I-zUQ(z)T_{21}\right)-\right.\\ &\qquad\left.-z^2T_{12}R(z)T_{21}\right\}^{-1},\end{aligned}$$

which is a rational function of the desired form. ∎

Our next lemma is a partial generalization of the well-known result that for complex-valued series of Stieltjes, the sequence of $[n-1,n]$ Padé approximants converge to the Stieltjes transform of the measure if the moments do not grow too rapidly (Theorem A).

Lemma 3. *Let* $\{M_n\}_{n=0}^\infty$ *be a sequence of positive numbers such that* $\sum_0^\infty M_n^{-1/2n} = \infty$. *Let* $d\alpha$ *and* $\{d\alpha_n\}_{n=0}^\infty$ *be positive measures on* $[0, \infty[$. *Put* $c_k = \int_0^\infty t^k d\alpha(t)$ *and suppose that* $c_k \leqq M_k$ *for all* k. *Suppose also that* $\int_0^\infty t^k d\alpha_n(t) = c_k$, $k = 0, 1, \ldots, 2n$. *Then for non-real* z *in the right half plane,*

$$\int_0^\infty \frac{d\alpha_n(t)}{1+zt} \to \int_0^\infty \frac{d\alpha(t)}{1+zt} \quad \text{as} \quad n \to \infty.$$

Remark. This is certainly not the sharpest possible result in this direction, but it will be sufficient for our purposes.

Proof: The proof depends on the following result by M. Riesz [6]: For non-real z

$$\left|\int_0^\infty \frac{d\alpha(t)}{1+zt} - \int_0^\infty \frac{d\alpha_n(t)}{1+zt}\right| \leqq \frac{|z|}{|\mathrm{Im}\, z|} \cdot \left(\sum_0^n \left|P_k\left(-\frac{1}{z}\right)\right|^2\right)^{-1},$$

where $P_k(w)$ is the k:th orthogonal polynomial with respect to $d\alpha$. That $\sum_0^n \left|P_k\left(-\frac{1}{z}\right)\right|^2 \to \infty$ if $\sum_0^\infty c_k^{-1/2k} = \infty$ is well-known in the theory of the moment problem. In the proof of Theorem 2 of [2] a proof for z on the positive real axis can be found. Since $\{P_k\}$ has all their zeros on the positive real axis this implies that $\sum_0^\infty \left|P_k\left(-\frac{1}{z}\right)\right|^2 = \infty$ for z in the right half plane. ∎

Our last lemma is a sharpening of the well-known result that if a sequence of Padé approximants is uniformly bounded in a region containing the origin, then the sequence converges.

Lemma 4. *Let* Ω *be a simply connected proper subset of the complex plane containing the origin. Let* $F(z)$ *and* $\{F_n(z)\}_{n=1}^\infty$ *be holomorphic in* Ω *with values in a Banach space* B. *Suppose that* $F_n(z) - F(z) = O(z^{2n})$ *as* $z \to 0$ $\forall n$ *and that* $\{F_n(z)\}_1^\infty$ *is uniformly bounded in any compact subset of* Ω. *Let* $\varphi(z)$ *be a conformal mapping of* Ω *onto the interior of the unit circle such that* $\varphi(0) = 0$. *Then for any* $z \in \Omega$

$$\limsup_{n\to\infty} \|F_n(z) - F(z)\|^{1/2n} \leqq |\varphi(z)|.$$

Proof: Put $\psi(z) = \varphi^{-1}(z)$ and define $G(z) = F(\psi(z))$, $G_n(z) = F_n(\psi(z))$. Then G and $\{G_n\}_1^\infty$ are holomorphic in the unit disc and uniformly bounded on compact subsets. Furthermore $G_n(z) - G(z) = O(z^{2n})$ as $z \to 0$ since $\varphi(0) = 0$. Now fix $z \in \Omega$ and choose $\varepsilon > 0$. Put $r = |\varphi(z)|$ and choose R, $r < R < 1$ such that $\frac{r}{R} < r + \varepsilon$.

Choose M such that $\|G_n(w)\| \leqq M$, $\|G(w)\| \leqq M$ for $|w|=r$. Then by a well-known argument, since $(G_n(w)-G(w))/w^{2n}$ is holomorphic in the unit disc

$$\|F_n(z)-F(z)\| = \|G_n(\varphi(z))-G(\varphi(z))\| = \frac{r^{2n}}{2\pi}\cdot\left\|\int_{|w|=R}\frac{G_n(t)-G(t)\,dt}{t^{2n}(t-w)}\right\| \leqq$$
$$\leqq \left(\frac{r}{R}\right)^{2n}\cdot\frac{2MR}{R-r}.$$

Taking $2n$-th roots we obtain

$$\limsup_{n\to\infty}\|F_n(z)-F(z)\|^{1/2n} \leqq \frac{r}{R} < |\varphi(z)|+\varepsilon.$$

Since ε is arbitrary, the lemma follows. ∎

5. Proof of Theorem 1

Theorem 1 is established by combining the statements of the following ten steps.

Step 1. If for any m and n a right $[m, n]$ Padé approximant to F exists, it is also a left approximant and is thus unique.

Proof: For any polynomial $T(z)=\sum_{k=0}^{i} T_k z^k$, put $T^+(z)=\sum_{k=0}^{i} T_k^* z^k$. Since C_k is selfadjoint for all k we have formally for real z

$$(F(z)Q(z)-P(z))^* = Q^+(z)F(z)-P^+(z) = O(z^{2n}).$$

This proves that $(Q^+)^{-1}P^+$ is a left $[m, n]$ Padé approximant. Uniqueness follows from the discussion in Section 2.

Step 2. We may assume without loss of generality that $C_0=I$.

Proof: Since C_0 is self-adjoint and has closed range, $H=R(C_0)\oplus N(C_0)$ where $R(C_0)$ and $N(C_0)$ reduce C_0. Since $A(t)$ is increasing it follows that $R(C_0)$ and $N(C_0)$ reduce $A(t)$ for all t and hence C_k for all k. We may thus restrict our attention to $R(C_0)$, or equivalently assume that $N(C_0)=\{0\}$. Then C_0 is positive and surjective and thus posseses a unique positive square-root with bounded inverse R^{-1}. Now put $A_1(t)=R^{-1}A(t)R^{-1}$. Then A_1 is increasing, $A_1(0)=0$ and $A_1(t)\to I$ as $t\to\infty$. Finally it is clear that if PQ^{-1} is a Padé approximant to the series of Stieltjes generated by A_1, then $RP(R^{-1}Q)^{-1}$ is a Padé approximant to F with the same convergence properties.

Step 3. There is a Hilbert space $K\supset H$ and a positive operator $\mathbf{B}: K\supset D(\mathbf{B})\to K$ such that $H\subset D(\mathbf{B}^n)$ for all n and

$$\tilde{F}(z) = \Pi_H(I+z\mathbf{B})^{-1}|\,H, \quad C_n = \Pi_H\mathbf{B}^n|\,H,$$

where Π_H denotes the projection onto H.

Proof: Naimark's theorem gives K and an orthogonal resolution of the identity $\{E(t)\}_{t>0}$ such that $\Pi_H E(t)|H=A(t)$. Put $\mathbf{B}=\int_0^\infty t\,dE(t)$. Then $\mathbf{B}^n=\int_0^\infty t^n dE(t)$ and $(I+z\mathbf{B})^{-1}=\int_0^\infty (1+zt)^{-1}dE(t)$ by the functional calculus for self-adjoint operators. Hence $C_n=\int_0^\infty t^n dA(t)=\Pi_H\mathbf{B}^n|H$ and similarly for $\tilde{F}$. Also, for any $x\in H$

$$\int_0^\infty t^{2n}d\langle E(t)x,x\rangle=\int_0^\infty t^{2n}d\langle A(t)x,x\rangle=\langle C_{2n}x,x\rangle<\infty$$

and thus $x\in D(\mathbf{B}^n)$.

Step 4. It is possible to choose pair-wise orthogonal subspaces $\{H_i\}_{i=0}^\infty$ with projections Π_i such that $H_0=H$, $H_i\subset D(\mathbf{B})$ for all i and $B_{ij}=\Pi_i\mathbf{B}|H_j$ satisfies

$$B_{ij}=0 \quad \text{if} \quad |i-j|>1 \quad \text{and} \quad R(B_{i+1,i})=H_{i+1}.$$

Proof: We put $H_0=H$ and define H_i inductively. Suppose $\{H_i\}_{i=0}^k$ have been chosen in such a way that every element x in H_i is a finite sum of the type $x=\sum_{\nu=0}^i \mathbf{B}^\nu x_\nu$ where $x_\nu\in H_0$. Put $Q_k=I-\sum_{j=0}^k \Pi_j$ and define

$$H_{k+1}=Q_k\mathbf{B}(H_k).$$

Then obviously every element of H_{k+1} can be written $y=\sum_{\nu=0}^{k+1}\mathbf{B}^\nu y_\nu$ where $y_\nu\in H_0$ and hence $H_{k+1}\subset D(\mathbf{B})$. Also, H_{k+1} is finite-dimensional and thus a subspace.

That $R(B_{i+1,i})=H_i$ follows from the construction as well as the fact that $B_{ij}=0$ if $i>j+1$. If $i<j-1$, pick $x\in H_i$, $y\in H_j$. Then

$$\langle B_{ij}y,x\rangle=\langle \Pi_i\mathbf{B}y,x\rangle=\langle \mathbf{B}y,x\rangle=\langle y,\mathbf{B}x\rangle=\langle y,B_{ji}x\rangle=0$$

whence $B_{ij}=0$.

Step 5. Define $\mathbf{B}_n\in L\left(\sum_0^n\oplus H_i\right)$ by

$$\mathbf{B}_n=\begin{Bmatrix} B_{00} & B_{01} & & & 0 \\ B_{10} & B_{11} & B_{12} & & \\ & B_{21} & \cdot & \cdot & \\ & & \cdot & \cdot & \cdot \\ 0 & & & \cdot & B_{nn} \end{Bmatrix}.$$

Put $F_n(z)=\Pi_0(I+z\mathbf{B}_n)^{-1}|H_0$. Then $F_n(z)$ is a rational function of type $[n,n+1]$.

Remark: Note that $\mathbf{B}_n=\left(\sum_{j=0}^n\Pi_j\right)\mathbf{B}|\sum_{j=0}^n\oplus H_j$.

Proof: Apply Lemma 2.

Step 6. $F(z)-F_n(z)=O(z^{2n+2}),\ z\to 0.$

Proof: From the construction in Step 4 follows that $\mathbf{B}^k(H_0)\subset\sum_{i=0}^k\oplus H_i$. Thus $C_k=\Pi_0\mathbf{B}^k|H_0=\Pi_0\mathbf{B}_k^k|H_0$ and we see that it is enough to prove that $\Pi_0\mathbf{B}_n^k|H_0=$ $=\Pi_0\mathbf{B}_{2n+1}^k|H_0$ for $k=0, 1, \ldots, 2n+1$. But for a fixed k we can calculate both sides of this proposed equality using matrix multiplication. The result is in both cases the sum of all possible products of the type

$$B_{0,m_1}\cdot B_{m_1,m_2}\cdot\ldots\cdot B_{m_{k-1},0},$$

where the second index of each factor equals the first ind ex of the next. There are more possible factors in $\mathbf{B}_{2n-1}$ than in $\mathbf{B}_n$ but since $B_{ij}=0$ for $|i-j|>1$ a non-zero product cannot contain a factor B_{ij} with i or j greater than $\left[\frac{k}{2}\right]$ in order to "be back in time for having last index 0". ([ ] of course denotes the integer part.) Thus for $k\leqq 2n+1$ no factors with an index greater than n occur which proves $\Pi_0\mathbf{B}_n^k|H_0=C_k$.

Remark: By Steps 5 and 6, $F_n(z)$ is the $[n, n+1]$ Padé approximant to F and from now on we use the notation $F[n, n+1]\,(z)$ instead of $F_n(z)$.

Step 7. $\|F[n, n+1]\,(z)\|\leqq\sup_{t\geqq 0}|1+zt|^{-1}$.

Proof: We prove the stronger statement $\|(I+z\mathbf{B}_n)^{-1}\|\leqq\sup_{t\geqq 0}|1+zt|^{-1}$.s This follows immediately from the spectral theorem. Since $\mathbf{B}$ is a non-negative operator, so is $\mathbf{B}_n$ and thus

$$(I+z\mathbf{B}_n)^{-1}=\int_0^\infty\frac{1}{1+zt}\,d\,E_n(t)$$

for some spectral family $\{E_n(t)\}_{t\geqq 0}$. Hence

$$\|(I+z\mathbf{B}_n)^{-1}\|\leqq\sup_{t\geqq 0}|1+zt|^{-1}$$

by the functional calculus.

Step 8. $\|F[n, n+1]\,(z)-\tilde{F}(z)\|\to 0$ as $n\to\infty$, uniformly on compact subset of $\mathbf{C}\setminus\mathbf{R}_-$.

Proof: By Step 7, $\{\|F[n, n+1]\,(z)\|\}_{n=0}^\infty$ is uniformly bounded on compact subsets of $\mathbf{C}\setminus\mathbf{R}_-$. Fix $x\in H$ with $\|x\|=1$. Then $\langle\tilde{F}(z)x, x\rangle$ is a complex Stieltjes transform with moments $c_k=\langle C_kx, x\rangle\leqq\|C_k\|$. Furthermore, $\langle F[n, n+1]\,(z)x, x\rangle$ is a complex Stieltjes transform with moments equal to c_k for $k\leqq 2n+1$. Thus by Lemma 3, with $M_k=\|C_k\|$, $\langle F[n, n+1]\,(z)x, x\rangle\to\langle\tilde{F}(z)x, x\rangle$ for non-real z in the right half-plane. By the polarization identity this yields $\langle F[n, n+1]\,(z)x, y\rangle\to$ $\langle\tilde{F}(z)x, y\rangle$ for all y or equivalently, since H is finite-dimensional, $\|F[n, n+1](z)-$ $-F(z)\|\to 0$. It remains to apply Vitali's theorem to get the desired conclusion.

Step 9. If $A(t)$ is constant for $t\geqq 1$ then

$$\limsup_{n\to\infty}\|F[n, n+1](z)-\tilde{F}(z)\|^{1/2n}\leqq\left|\frac{\sqrt{1+z}-1}{\sqrt{1+z}+1}\right|.$$

Proof: In this case, for $x \in \sum_0^n \oplus H_i$,

$$0 \leqq \langle \mathbf{B}_n x, x \rangle = \langle \mathbf{B}x, x \rangle = \int_0^1 t\, d\langle E(t)x, x \rangle \leqq 1$$

and the result of Step 7 can be sharpened to

$$\|F[n, n+1](z)\| \leqq \sup_{t \in [0,1]} |1+zt|^{-1}.$$

Thus the sequence of approximants is uniformly bounded in compact subsets of $\mathbf{C} \setminus]-\infty, -1]$ and the result follows from Lemma 4, noting that $\varphi(z) = (\sqrt{1+z}-1)/ /(\sqrt{1+z}+1)$ is a conformal mapping from $\mathbf{C} \setminus]-\infty, -1]$ onto the unit disc.

Step 10. F has a $[n+j, n]$ Padé approximant for all n and all $j \geqq -1$. The convergence properties of all sequences $\{F[n+j, n]\}_{n=0}^{\infty}$ are the same in the sense of (b) and (c).

Proof: Fix $j \geqq -1$ and put $G_j(z) = \sum_{k=0}^{\infty} C_{k+j+1}(-z)^k$. Then G_j is a series of Stieltjes generated by the increasing function $A_j(t) = \int_0^t u^{j+1} dA(u)$ and it is easy to prove that

$$F[n+j, n](z) = \sum_{k=0}^{j} C_k \cdot (-z)^k + (-z)^{j+1} \cdot G_j[n-1, n](z),$$

which proves the existence of $F[n+j, n]$. Furthermore

$$\tilde{F}(z) - F[n+j, n](z) = (-z)^{j+1}(\tilde{G}_j(z) - G_j[n-1, n](z)),$$

which proves that the convergence properties are the same. This completes the proof of Theorem 1.

6. The infinite-dimensional case

A natural question to ask at this point is whether Theorem 1 is true in an arbitrary Hilbert space. An inspection of the proof shows that for arbitrary H the result in Step 3 above is true with the same proof. If we suppose $A(t)$ to be constant for $t \geqq 1$, then $\mathbf{B}$ is bounded and we can carry out the construction in Step 4 in the same manner with the difference that we may have to take closures in defining the H_k. Then $B_{k,k+1}$ is not surjective and Step 5 fails. Steps 6 and 9 work, however, and thus $F_n(z) \to \tilde{F}(z)$ with the degree of convergence implied by (c). From Lemmas 1 and 2 it is seen that $F_n(z)$ can be viewed as a convergent in a continued fractions expansion. (This expansion is essentially the same as that obtained by Mac Nerney.) We feel that in the infinite-dimensional case this expansion may turn out to be a more natural interpolation procedure than the unsymmetrical Padé approximation. This feeling is supported by the following argument. In the infinite-dimensional case the region of holomorphy of a rational function may be any open set (since e.g. resolvents are rational functions).

It seems reasonable to demand of an approximant obtained by interpolation at the origin that the origin should be included in the set where it is holomorphic. This is also essential for the proof of Theorem 1(c). But it turns out that even for series of Stieltjes an $[n-1, n]$ Padé approximant $P\cdot Q^{-1}$ with $Q(0)$ having a bounded inverse need not exist. We illustrate this by the following simple example.

Let H be an infinite-dimensional Hilbert space and let A and T be positive, bounded, one-to-one operators on H. Let $\int_0^1 t\,dE(t)$ be the orthogonal spectral resolution of A. Suppose further that $A(H)=H$, $T(H)\neq H$ and $AT(H)\not\subset T(H)$. Put

$$F(z) = T(I+Az)^{-1}T = \int_0^1 \frac{dTE(t)T}{1+zt}.$$

Thus $F(z)$ is a series of Stieltjes. We shall try to define a right [0, 1] Padé approximant to F with $Q(0)$ having a bounded inverse or, equivalently, $Q(0)=I$. We must find P_0 and Q_1 such that $F(z)\,(I+Q_1z)-P_0=O(z^2)$. But

$$T(I+Az)^{-1}T(I+Q_1z)-P_0 = T^2-P_0+(T^2Q_1-TAT)z+O(z^2).$$

Thus we must have $T^2Q_1-TAT=0$, which forces $TQ_1=AT$ since T is one-to-one. But TQ_1 maps H into $T(H)$ while $AT(H)\not\subset T(H)$ by assumption. Hence no approximant of the desired type exists.

References

1. ALLEN, G. D., NARCOWICH, F. J., On the representation and approximation of a class of operator-valued analytic functions, *Bull. Amer. Math. Soc.* **81** (1975) 410—412.
2. KARLSSON, J., von SYDOW, B., The convergence of Padé approximants to series of Stieltjes, *Ark. Mat.* **14** (1976) 43—53.
3. MAC NERNEY, J. S., Hermitian moment sequences, *Trans. Amer. Math. Soc.* **103** (1962) 45—81.
4. NAIMARK, M. A. (Neumark, M. A.), On a representation of additive operator set functions, *Dokl. Akad. Nauk SSSR* **41** (1943) 359—361.
5. RIESZ, F., SZ.-NAGY, B., *Leçons d'Analyse Fonctionelle*, 3me ed., Academie des sciences de Hongrie, Szeged (1955).
6. RIESZ, M., Sur le problème des moments III, *Arkiv för Matematik, Astronomi och Fysik* **17** (1923) No 16, 1—52.
7. SZ.-NAGY, B., *Prolongements des transformations de l'espace de Hilbert qui sortent de cet espace*, Appendice aux [5].
8. ZINN-JUSTIN, J., Strong interaction dynamics with Padé approximants, *Phys. Rep.* **1** (1971) 55—102.

Received December 23, 1976

Björn von Sydow
Matematiska institutionen
Umeå universitet
S—90187 Umeå
Sweden