# N-VARIABLE RATIONAL APPROXIMANTS AND METHOD OF MOMENTS* <br> C. Alabiso $\dagger$ and P. Butera $\dagger \dagger$ <br> Stanford Linear Accelerator Center Stanford University, Stanford, California 94306 


#### Abstract

The method of moments is applied to pairs of linear permutable selfadjoint operators A and B in a Hilbert space $\mathscr{H}$. An approximate expression for the diagonal matrix elements of the operator $(1-w A-z B)^{-1}$, where $w, z$ are complex numbers, is taken as a guide to the definition of rational approximants from general formal power series in two variables. Starting from an operator convergence theorem in a certain Hilbert space, we prove the convergence of our approximants to analytic functions of two complex variables with the integral representation $G(w, z)=\iint \frac{d \sigma(\alpha, \beta)}{1-w \alpha-z \beta}$, under suitable restrictions on the positive measure $\sigma(\alpha, \beta)$. The same approximation scheme can also be applied to the diagonal matrix elements of the operator $[(1-\mathrm{wA})(1-\mathrm{z} \cdot \mathrm{B})]^{-1}$, leading to a different rational approximant which we prove to converge to functions with the integral representation $\widetilde{G}(w, z)=\iint \frac{d \sigma(\alpha, \beta)}{(1-w \alpha)(1-z \beta)}$. In both cases the convergence is uniform on appropriate compact subsets of $c^{2}$. The extension to the n-dimensional case is straightforward for both approximants. The connections with a standard variational principle are also briefly discussed.


[^0]
## I. INTRODUCTION

In the past few years the technique of Padé Approximants (PAs) ${ }^{1-3}$ for the approximate summation of power series of one complex variable has been looked at with some interest by physicists as an effective tool for many quantum mechanical and field-theoretic models whose solutions are only available in the form of a perturbative series. We recall that, given the formal power series $f(z) \simeq \sum_{n} f_{n} z^{n}$, the $[N / M](z) P A$ is the rational function $P_{N}(z) / Q_{M}(z)$, where $P_{N}(z)$ and $Q_{M}(z)$ are polynomials of degree $N$ and $M$ respectively, such that $P_{N}(z) / Q_{M}(z)=\sum_{n=0}^{N+M} f_{n} z^{n}+O\left(z^{N+M+1}\right)$. A simple closed expression is available for the $[\mathrm{N} / \mathrm{M}](\mathrm{z})$ PA and it can be shown that the PAs have some significant formal properties; e.g., if $N=M$ they are invariant under homographical transformations both of the variable and of the function. The PAs converge uniformly on compacts to extended Stieltjes functions, i.e., the functions $\mathrm{g}(\mathrm{z})$ of the form $g(z)=\int_{-\infty}^{\infty} \frac{d \sigma(t)}{1-z t}$ where $\sigma(t)$ is a positive measure with finite moments $\mu_{\mathrm{n}}=\int_{-\infty}^{\infty} \mathrm{t}^{\mathrm{n}} \mathrm{d} \sigma(\mathrm{t})$ not too fastly increasing with n ; moreover, in a suitably generalized sense, they converge to meromorphic functions. ${ }^{4}$ Unfortunately, the extension to the multidimensional case is not straightforward. In fact, the simplest generalization of the usual definition of the PAs, even in the case of two variables only, does not, in general, determine uniquely a rational approximant: additional constraints must be provided. To this problem, very interesting alternative solutions have been recently proposed. In one of these, ${ }^{5,6}$ the constraints are chosen in such a way that the many variable approximants retain the main formal properties of the usual PAs. In spite of this, the study of the convergence properties is not easy and, up to now, only generalizations of de Montessus theorem are available. ${ }^{7}$ For another kind of approximant ${ }^{8}$
the convergence to holomorphic functions has been proved under the stringent assumption of uniform boundedness of the approximants themselves.

In this paper, we would like to indicate a different approach to the construction of many-variable rational approximants starting from the following remarks. Consider a linear self-adjoint operator $A$ and a vector $f$ on a Hilbert space $\mathscr{H}$. Let $\sum_{n=0}^{\infty} z^{n}\left(f, A^{n} f\right)$ be the Neumann expansion (not necessarily convergent) of the diagonal matrix element (f, (1-zA) ${ }^{-1} \mathrm{f}$ ). Then, for every $N \geq 1$ the $[\mathrm{N}-1 / \mathrm{N}]$ ( z ) PA for this series coincides a) with the matrix element (f, $\left(1-z A_{N}\right)^{-1} f$ ) where $A_{N}$ is the $N$ rank operator obtained at the $N$-th order in the approximation scheme known as the method of moments; 9,10 b) with the stationary value of an appropriate functional on a certain finite dimensional subspace of $\mathscr{H} .^{11}$ Therefore, we suggest generalizing the PA to the multidimensional case by starting from the definition in terms of the method of moments rather than from the usual definition; more precisely, we suggest that the direct extension of the method of moments to the operator $(1-\mathrm{wA}-\mathrm{zB})^{-1}$ with A and B linear self-adjoint permutable operators should be taken as a guide to the definition and justification of two variable rational approximants. As a result, although some formal properties of the usual PAs are lacking, we still have the same connection with the method of moments (and the variational method). This enables us to give, for a relevant class of functions, a convergence proof which is both simple and of practical use since it involves only assumptions about the analytic properties of the functions to be approximated rather than about the behavior of the approximants themselves. Furthermore, our approximants have a simple explicit expression in any order of approximation.

We shall not study here any application of our approximation scheme, but let us just remark that a natural field of application should be the approximate summation of the perturbative solution of quantum mechanical and field theoretic models with more than one coupling constant. However, whether the physically interesting models fulfill all the requirements of our convergence theorem, is a question which requires further study. It is also worth mentioning that there are classical special functions which, for a particular choice of some of the defining parameters, have the integral representation required in our convergence proofs, i.e., the two-variable Appel hypergeometric functions and their $n$-variable generalizations, the Lauricella functions. ${ }^{12}$ The numerical computation of such functions is therefore another possible application of our approximation procedure which, in this case, provides a direct generalization of the classical Jacobi continued fraction expansion of the Gauss hypergeometric function ${ }_{2} F_{1}(1, \beta, \gamma, z)$.

In Section II we consider the method of moments for a pair of self-adjoint permutable operators A and B and we give the "approximate" expression for the matrix element $\left(f,(1-w A-z B)^{-1} f\right)$ where $f$ is a suitable vector of the Hilbert space. From this we obtain a rational expression which can be associated with any double power series. In Section III we prove a convergence theorem for operators in a Hilbert space and we use this result to state in Section IV a convergence theorem of our approximants to functions of two complex variables with a well defined analytic structure. In Section $V$ we present the trivial extension to the $n$-dimensional case; the connection with a standard variational principle; and another kind of approximant, suggested by the application of the method of moments to the operator $[(1-w A)(1-z B)]^{-1}$.

## II. THE METHOD OF MOMENTS

Let $A$ and $B$ be two linear self-adjoint permutable operators with domains $\mathscr{D}(\mathrm{A})$ and $\mathscr{D}(\mathrm{B})$ in the Hilbert space $\mathscr{H}$. Then there exists a dense subset $\mathscr{2}$ of vectors of $\mathscr{H}$ which are quasi-analytic ${ }^{13,14}$ for both A and B. Let $\mathrm{f} \in \mathscr{Q}^{15}$ be such that

$$
\begin{equation*}
f_{p-q, q}=A^{p-q_{B} q_{f}} \quad p=0, \ldots, N ; q=0, \ldots, p \tag{1}
\end{equation*}
$$

are linearly independent vectors for any $N$. Then, the vectors $\left\{\mathrm{f}_{\mathrm{r}, \mathrm{s}}\right\}$ generate a sequence of $\frac{\mathrm{N}(\mathrm{N}+1)}{2}$-dimensional Hilbert spaces $\mathscr{H}_{\mathrm{N}} \subseteq \mathscr{H}$ and the related orthogonal projection operators $\mathrm{P}_{\mathrm{N}}$. Let us consider the equation

$$
\begin{equation*}
(1-\mathrm{wA}-\mathrm{zB}) \psi=\mathrm{f} \tag{2}
\end{equation*}
$$

where $w$ and $z$ are complex numbers. For any $\{w, z\}$ such that the operator $R(w, z) \equiv(1-w A-z B)^{-1}$ exists and is bounded, the solution of Eq. (2) is

$$
\psi=R(w, z) \mathbf{f}
$$

In order to obtain an approximate solution of Eq. (2) let us consider the following equation in the finite-dimensional subspace $\mathscr{H}_{\mathrm{N}}$

$$
\begin{equation*}
\left(1-\mathrm{wA}_{N}-\mathrm{zB}_{\mathrm{N}}\right) \psi_{\mathrm{N}}=\mathrm{f} \tag{4}
\end{equation*}
$$

where

$$
\begin{equation*}
A_{N} \equiv P_{N} A P_{N}, \quad B_{N} \equiv P_{N} B P_{N} \tag{5}
\end{equation*}
$$

The solution of Eq. (4) is

$$
\begin{equation*}
\psi_{N}=\left(1-w A_{N}-z B_{N}\right)^{-1} f \equiv R_{N}(w, z) f \tag{6}
\end{equation*}
$$

for $\{w, z\}$ such that $R_{N}(w, z)$ exists and is bounded. Since $\psi_{N} \in \mathscr{H}{ }_{N}$ we can also solve Eq. (4) explicitly by expanding $\psi_{N}$ on the complete set $\left\{f_{r, s}\right\}$ :

$$
\begin{equation*}
\psi_{N}=\sum_{p=0}^{-N} \quad \underset{q=0}{p} a_{p, q}{ }_{p-q, q} \tag{7}
\end{equation*}
$$

We obtain

$$
\begin{equation*}
\psi_{N}=\sum_{p=0}^{N} \sum_{q=0}^{p} \sum_{r=0}^{N} \sum_{s=0}^{r}\left(M_{N}^{-1}\right)_{p, q ; r, s} F_{r-s, s} A^{p-q_{B} q_{f}}, \tag{8}
\end{equation*}
$$

where

$$
\begin{gather*}
F_{r-s, s} \equiv\left(f, A^{r-s} B^{s} f\right)  \tag{9}\\
\left(M_{N}\right)_{p, q ; r, s} \equiv F_{p+r-q-s, q+s}-w F_{p+r+1-q-s, q+s}-z_{p+r-q-s, q+s+1} \\
p, r=0, \ldots, N ; q=0, \ldots, p ; \quad s=0, \ldots, r
\end{gather*}
$$

If we project Eq. (8) on the vector f we obtain the simple expression

$$
\left(f, \psi_{N}\right)=\left(f, R_{N}(w, z) f\right)=\sum_{p=0}^{N}{\underset{q}{i}=0}_{p}^{N} \sum_{s=0}^{N} \sum_{p-q, Q^{N}}^{N}\left(M_{N}^{-1}\right)_{p, q ; r, s} F_{r-s, s} \equiv F_{N}^{T} M_{N}^{-1} F_{N}
$$

with obvious definitions for the column matrix $\mathrm{F}_{\mathrm{N}}$, its transposed $\mathrm{F}_{\mathrm{N}}^{\mathrm{T}}$ and the matrix $M_{N}$. The corresponding matrix element of the operator $R(w, z)$ has the integral representation

$$
\begin{equation*}
(f, R(w, z) f)=\iint \frac{\mathrm{d}(\mathrm{f}, \mathrm{E}(\alpha, \beta) \mathrm{f})}{1-\mathrm{w} \alpha-\mathrm{z} \beta} \tag{11}
\end{equation*}
$$

where $\mathrm{E}(\alpha, \beta)$ is the spectral family associated with the self-adjoint permutable operators A and B.

In Section III we prove the strong convergence of $R_{N}(w, z)$ to $R(w, z)$ in a subspace of $\mathscr{H}$ for $\{\mathrm{w}, \mathrm{z}\}$ in a suitable domain and, as a consequence, the convergence of $\psi_{N}$ to $\psi$ and of $\left(f, \psi_{N}\right)$ to ( $f, \psi$ ). Since in the latter case we have a rational approximation converging to an analytic function of two complex variables, we are naturally led to introduce for any formal double power series

$$
\begin{equation*}
G(w, z)=\sum_{m, n} w^{m} z^{n}\binom{m+n}{m} G_{m, n} \tag{12}
\end{equation*}
$$

the rational approximant $G_{N}(w, z)$ by the formula:

$$
\begin{equation*}
G_{N}(w, z)=G_{N}^{T} Q_{N}^{-1} G_{N} \tag{13}
\end{equation*}
$$

where $G_{N}, G_{N}^{T}$ and $Q_{N}$ are a column matrix, its transposed and a matrix defined in terms of the coefficients $G_{m, n}$ by

$$
\begin{gather*}
\left(G_{N}\right)_{p, q} \equiv G_{p-q, q} \\
\left(Q_{N}\right)_{p, q ; r, s} \equiv G_{p+r-q-s, q+s}-w G_{p+r+1-q-s, q+s}-z G_{p+r-q-s, q+s+1}  \tag{14}\\
p, r=0, \ldots, N ; \quad q=0, \ldots, p ; \quad s=0, \ldots, r
\end{gather*}
$$

Let us write explicitly $G_{1}(w, z)$ and $G_{2}(w, z)$ :

$$
\begin{aligned}
& G_{1}(w, z)=\frac{G_{0,0} G_{0,0}}{G_{0,0}-w G_{1,0}{ }^{-z G_{0,1}}},
\end{aligned}
$$

It is convenient to notice that this rational approximant has some obvious formal properties, e.g., $G_{N}(w, z)$ is real analytic if $G(w, z)$ is a real analytic function of $w$ and $z$ and it is symmetric in $w$ and $z$ if $G(w, z)$ is. One may also notice that the approximant to a factorized function does not in general factorize and that no simple analogue of the homographical covariance properties of the Padé approximant seems to hold.

## III. SOME RESULTS ON OPERATOR CONVERGENCE

In this section we shall extend some results of Refs. 9 and 10 where the case of a single selfadjoint operator has been studied. Let $\mathscr{L}_{\mathrm{f}}$ be the linear manifold of all finite linear combinations of the vectors $\left\{f_{r, s}\right\}$ defined in Section II. The closure of $\mathscr{L}_{\mathrm{f}}$ is a Hilbert space $\mathscr{H}_{\mathrm{f}} \subseteq \mathscr{H}{ }^{16}$ Consider now the restrictions $A^{\prime}$ and $B^{\prime}$ of the operators $A$ and $B$ to $\mathscr{L}_{f}$ and their closures $\bar{A}^{\prime}$ and $\bar{B}^{\prime}$. Since $f$ is assumed to be a quasi-analytic vector for both $A$ and $B$ then, by the theorems 4 and 6 of Ref. $13, \bar{A}^{\prime}$ and $\bar{B}^{\prime}$ are still selfadjoint permutable operators on $\mathscr{H}_{\mathrm{f}}$, and from now on we shall simply call them A and B. These operators and the related ones $A_{N} \equiv P_{N} A P_{N}$ and $B_{N} \equiv P_{N} B P_{N}$ define in $\mathscr{H}_{\mathrm{f}}$ the operators $\mathrm{T}(\mathrm{w}, \mathrm{z}) \equiv \mathrm{wA}+\mathrm{zB}, \mathrm{T}_{\mathrm{N}}(\mathrm{w}, \mathrm{z}) \equiv \mathrm{P}_{\mathrm{N}} \mathrm{T}(\mathrm{w}, \mathrm{z}) \mathrm{P}_{\mathrm{N}}, \mathrm{R}(\mathrm{w}, \mathrm{z}) \equiv(1-\mathrm{T}(\mathrm{w}, \mathrm{z}))^{-1}$, $R_{N}(w, z) \equiv\left(1-T_{N}(w, z)\right)^{-1}$ where $\{w, z\} \in C^{2}$ is a pair of complex numbers. For simplicity we shall occasionally drop the $\{w, z\}$ dependence from our operators. Let us also stress that, throughout the paper, by operator convergence we shall always mean strong operator convergence. T is a normal maximal operator and, since it is closed, ${ }^{18}$ it is the closure of the operator $\mathrm{wA}^{1}+\mathrm{zB}$. $\mathrm{T}_{\mathrm{N}}$ is a bounded operator and, in general, it is not normal.

In order to prove that $R_{N}(w, z)$ converges to $R(w, z)$ on $\mathscr{H}_{f}$ we need some information on the behavior of $T_{N}(\mathrm{w}, \mathrm{z})$ as $\mathrm{N} \rightarrow \infty$, which is given by the following

Theorem 1. $\mathrm{T}_{\mathrm{N}}(\mathrm{w}, \mathrm{z}) \rightarrow \mathrm{T}(\mathrm{w}, \mathrm{z})$ in $\mathscr{L}_{\mathrm{f}}$, uniformly with respect to $\{w, z\}$.
Proof. Any vector $g \in \mathscr{L}_{f}$ can be written as $g=\sum_{m=0}^{M} \sum_{n=0}^{m} a_{m n} A^{m-n} B^{n}$. If $N \geq M+1$ then $T_{N}(w, z) g \equiv P_{N}(w A+z B) P N_{N}=(w A+z B) g=T(w, z) g$. Of course the convergence is uniform with respect to $\{w, z\}$.

Let us now recall that $\overline{(0)(0)}$, the closure of the numerical range ${ }^{17}$ of a linear bounded operator $O$ is a convex set containing the spectrum $\sigma(\mathrm{O})$ of $0 .{ }^{19}$ If $O$ is a normal maximal operator (not necessarily bounded) $\overline{\operatorname{Or}(\mathrm{O})}$ is the convex hull of $\sigma(\mathrm{O})$, i.e., $\overline{\Theta(\mathrm{O})}$ is the smallest closed convex set containing $\sigma(\mathrm{O}) .{ }^{20}$ Theorem 1 and the following theorem enable us to prove that $R_{N}(w, z) \rightarrow R(w, z)$ in $\mathscr{L}_{\mathrm{f}}$.

Theorem 2. For all $\{w, z\}$ such that the point 1 is at a positive distance $d$ from $\overline{(\Theta)(T(w, z))}, R(w, z)$ and $R_{N}(w, z)$ exist as bounded operators and satisfy the bounds $\|R(w, z)\| \leq d^{-1},\left\|R_{N}(w, z)\right\| \leq \delta^{-1}$ where $\delta^{-1} \equiv \max \left\{1, \mathrm{~d}^{-1}\right\}$.

Proof. Since $\overline{(0)(\mathrm{T}(\mathrm{w}, \mathrm{z}))}$ is the closed convex hull of the spectrum of a normal maximal operator, the point 1 is at least at a distance $d$ from the spectrum itself, and ther efore the operator $R(w, z)$ exists and is bounded. If $g \neq 0$ then $R(w, z) g \neq 0$ and we can consider the normalized vector $h=\frac{R(w, z) g}{\|R(w, z) g\|} \cdot$ By assumption

$$
0<d \leq|(h, T(w, z) h)-1|=\frac{\mid R(w, z) g, g) \mid}{\|R(w, z) g\|^{2}} \leq \frac{\|g\|}{\|R(w, z) g\|}
$$

Therefore $\|R(w, z)\| \leq d^{-1}$. A similar result also obtains for $R_{N}(w, z)$ from the remark that $\left|\left(h, R_{N}(w, z) h\right)\right| \leq\left\|R_{N}^{\prime}(w, z)\right\|\left\|h_{N}\right\|^{2}+\left\|h_{\perp}\right\|^{2}$ where $R_{N}^{\prime}(\mathrm{w}, \mathrm{z})$ is the restriction of $\mathrm{R}_{\mathrm{N}}(\mathrm{w}, \mathrm{z})$ to $\mathscr{H}_{\mathrm{N}},\|\mathrm{h}\|=1$, $h_{N} \equiv P_{N} h$, and $h_{\perp} \equiv\left(1-P_{N}\right) h$. Since $\oplus_{N}\left(T_{N}\right) \subseteq(T)$, where $\Theta_{N}\left(T_{N}\right)$ is the numerical range of $T_{N}$ in $\mathscr{H}_{N}$, then $\left\|R_{N}^{\prime}(w, z)\right\| \leq d^{-1}$. Therefore $\left\|R_{N}(\mathrm{w}, \mathrm{z})\right\| \leq \delta^{-1}=\max \left\{1, \mathrm{~d}^{-1}\right\}$.

Next lemma will be used to extend the convergence from $\mathscr{L}_{\mathrm{f}}$ to the whole space $\mathscr{H}_{\mathbf{f}}$.

Lemma 1. Let $O(w, z)$ be a linear bounded operator defined on a Hilbert space $\mathscr{H}$ and depending on the two complex variables $\{\mathrm{w}, \mathrm{z}\}$. Let $\left\{\mathrm{O}_{\mathrm{N}}(\mathrm{w}, \mathrm{z})\right\}$ be a sequence of such operators, uniformly bounded with respect to $N$. If, for a given $\{w, z\}, O_{N}(w, z) \rightarrow O(w, z)$ on $\mathscr{S}(\mathrm{w}, \mathrm{z})$ where $\mathscr{S}(\mathrm{w}, \mathrm{z})$ is a dense subset of $\mathscr{H}$, then $\mathrm{O}_{\mathrm{N}}(\mathrm{w}, \mathrm{z}) \rightarrow \mathrm{O}(\mathrm{w}, \mathrm{z})$ also on $\mathscr{H}$. If, for all $\{\mathrm{w}, \mathrm{z}\}$ in a domain $\Delta \subset c^{2}$, a) $\mathrm{O}(\mathrm{w}, \mathrm{z})$ and $\left\{\mathrm{O}_{\mathrm{N}}(\mathrm{w}, \mathrm{z})\right\}$ are uniformly bounded, b) $\mathscr{F} \equiv \mathscr{P}(\mathrm{w}, \mathrm{z})$ does not depend on $\{\mathrm{w}, \mathrm{z}\}$, c) $\mathrm{O}_{\mathrm{N}}(\mathrm{w}, \mathrm{z}) \rightarrow \mathrm{O}(\mathrm{w}, \mathrm{z})$ on $\mathscr{S}$ uniformly in $\Delta$, then $\mathrm{O}_{\mathrm{N}}(\mathrm{w}, \mathrm{z}) \rightarrow \mathrm{O}(\mathrm{w}, \mathrm{z})$ on $\mathscr{H}$ uniformly in $\Delta$.
Proof. Consider a fixed $\{w, z\}$, then for all $g \in \mathscr{H}$ there exists a sequence $\left\{g_{n}\right\} \in \mathscr{S}(\mathrm{w}, \mathrm{z})$ such that $\mathrm{g}_{\mathrm{n}} \rightarrow \mathrm{g}$. Therefore

$$
\begin{aligned}
& \left\|\left(O_{N}(w, z)-O(w, z)\right) g\right\| \leq\left\|O_{N}(w, z)\left(g-g_{n}\right)\right\|+ \\
& \quad+\left\|O(w, z)\left(g-g_{n}\right)\right\|+\left\|\left(O_{N}(w, z)-O(w, z)\right) g_{n}\right\| \leq 2 M(w, z)\left\|g-g_{n}\right\|+ \\
& \quad+\left\|\left(O_{N}(w, z)-O(w, z)\right) g_{n}\right\|
\end{aligned}
$$

where $\|O(w, z)\| \leq M(w, z), \quad\left\|O_{N}(w, z)\right\| \leq M(w, z) \quad$ for all $N$.
Let us fix $n_{\epsilon}$ in such a way that $\left\|g-g_{n_{\epsilon}}\right\| \leq \frac{\epsilon}{4 M(W, z)}$. Since $\mathrm{g}_{\mathrm{n}} \in \mathscr{S}(\mathrm{w}, \mathrm{z})$ and $\mathrm{O}_{\mathrm{N}}(\mathrm{w}, \mathrm{z}) \rightarrow \mathrm{O}(\mathrm{w}, \mathrm{z})$ in $\mathscr{S}(\mathrm{w}, \mathrm{z})$, we can choose $N_{\epsilon}(w, z)$ such that $\left\|\left(O_{N}(w, z)-O(w, z)\right) g_{n_{\epsilon}}\right\|<\frac{\epsilon}{2}$ for all $N>N_{\epsilon}(w, z)$. Then the first part of the theorem follows. If, for $\{w, z\} \in \Delta$, the operators are uniformly bounded, the set $\mathscr{P} \equiv \mathscr{P}(\mathrm{w}, \mathrm{z})$ does not depend
on $\{\mathrm{w}, \mathrm{z}\}$ and the convergence is uniform on $\mathscr{S}$, then $\mathrm{M}, \mathrm{n}_{\epsilon}$ and $\mathrm{N}_{\epsilon}$ do not depend on $\{\mathrm{w}, \mathrm{z}\}$ and the convergence is uniform on $\Delta$.

Before applying lemma 1 to our case we need the following

Lemma 2. If $1 \in \rho(T(w, z))$, where $\rho(T(w, z))$ is the resolvent set of $\mathrm{T}(\mathrm{w}, \mathrm{z})$, then $\mathscr{S}_{\mathrm{f}}(\mathrm{w}, \mathrm{z}) \equiv(1-\mathrm{T}(\mathrm{w}, \mathrm{z})) \mathscr{L}_{\mathrm{f}}$ is a dense linear manifold of $\mathscr{H}$. Moreover, for any finite $\{w, z\}$ and $\left\{w^{\mathbf{+}}, z^{\prime}\right\}$ such that $1 \in \rho(T(w, z))$ and $1 \in \rho\left(T\left(w^{\prime}, z^{\prime}\right)\right), \mathscr{P}_{f}(w, z)$ coincides with $\mathscr{S}_{f}\left(w^{\prime}, z^{\prime}\right)$. Proof. By assumption ( $1-\mathrm{T}(\mathrm{w}, \mathrm{z}))^{-1}$ exists as a bounded operator on $\mathscr{H}_{\mathrm{f}}$. Therefore any vector $\mathrm{h} \in \mathscr{H}_{\mathrm{f}}$ can be written as $\mathrm{h}=(1-\mathrm{T}) \mathrm{g}$ with $\mathrm{g}=(1-\mathrm{T})^{-1} \mathrm{~h}$. Since $\mathscr{L}_{\mathrm{f}}$ is dense in $\mathscr{H}_{\mathrm{f}}$ and T is the closure of an operator with domain $\mathscr{L}_{\mathrm{f}}$, there exists a sequence $\mathrm{g}_{\mathrm{n}} \rightarrow \mathrm{g}$ with $g_{n} \in \mathscr{L}_{\mathrm{f}}$ such that $\mathrm{h}_{\mathrm{n}}=(1-\mathrm{T}) \mathrm{g}_{\mathrm{n}} \rightarrow \mathrm{h}$. The second part of the theorem is proved by a direct check that any vector in $\mathscr{S}_{\mathrm{f}}(\mathrm{w}, \mathrm{z})$ also belongs to $\mathscr{S}_{\mathbf{f}}\left(\mathrm{w}^{\prime}, \mathrm{z}^{\prime}\right)$ and vice versa.

We can now state the main theorem

Theorem 3. Let $\Delta$ be a domain of $C^{2}$ such that the point 1 is at a positive distance $d$ from $\times(T(w, z))$. Then, for $\{w, z\} \in \Delta, R_{N}(w, z)$ converges strongly to $R(w, z)$ on $\mathscr{F}_{f}$, uniformly on any bounded subset $\Gamma \subset \Delta$.

Proof. For a fixed $\{w, z\} \in \Delta$, let $R(w, z) g$ be in $\mathscr{L}_{f}$ :

$$
\begin{aligned}
\left.\| R_{N}(w, z)-R(w, z)\right) g \| & =\left\|R_{N}(w, z)\left(T_{N}(w, z)-T(w, z)\right) R(w, z) g\right\| \leq \\
& \leq\left\|R_{N}(w, z)\right\|\left\|\left(T_{N}(w, z)-T(w, z)\right) R(w, z) g\right\| .
\end{aligned}
$$

From Theorem 1 we have that $\mathrm{T}_{\mathrm{N}} \rightarrow \mathrm{T}$ in $\mathscr{L}_{\mathrm{f}}$ and from Theorem 2 that $\left\|R_{N}(\mathrm{w}, \mathrm{z})\right\| \leq \delta^{-1}$. Therefore $\mathrm{R}_{\mathrm{N}}(\mathrm{w}, \mathrm{z}) \rightarrow \mathrm{R}(\mathrm{w}, \mathrm{z})$ on $\mathscr{S}_{\mathrm{f}}(\mathrm{w}, \mathrm{z}) \equiv(1-\mathrm{T}(\mathrm{w}, \mathrm{z})) \mathscr{L}_{\mathrm{f}}$. Since, by lemma 2, $\mathscr{S}_{\mathrm{f}}(\mathrm{w}, \mathrm{z})$ is dense in $\mathscr{H}_{\mathrm{f}}$, it follows that $\mathrm{R}_{\mathrm{N}}(\mathrm{w}, \mathrm{z}) \rightarrow \mathrm{R}(\mathrm{w}, \mathrm{z})$ on $\mathscr{H}_{\mathrm{f}}$. To prove uniform convergence let us remark that $\mathrm{T}_{\mathrm{N}}(\mathrm{w}, \mathrm{z}) \rightarrow \mathrm{T}(\mathrm{w}, \mathrm{z})$ uniformly with respect to $\{\mathrm{w}, \mathrm{z}\}$ and that, for $\{\mathrm{w}, \mathrm{z}\} \in \Gamma, \mathscr{S}_{\mathrm{f}} \equiv \mathscr{S}(\mathrm{w}, \mathrm{z})$ does not depend on $\{w, z\}$. It follows that $R_{N}(w, z) \rightarrow R(w, z)$ in $\mathscr{S}_{f}$, uniformly in $\Gamma$. By lemma 1 we conclude that $R_{N}(\mathrm{w}, \mathrm{z}) \rightarrow \mathrm{R}(\mathrm{w}, \mathrm{z})$ on the whole $\mathscr{H}_{\mathrm{f}}$, uniformly in $\Gamma$.

## IV. CONVERGENCE OF APPROXIMANTS FOR DOUBLE POWER SERIES

In Section II convergence theorems have been formulated for operators in an abstract Hilbert space. Let us now turn our attention to the approximant $G_{N}(w, z)$ defined by Eq. (13) starting from the formal double power series (12) associated to a function of two complex variables $G(w, z)$. Under suitable hypotheses we can prove the convergence of $G_{N}(w, z)$ to $G(w, z)$, by reducing the problem to the Hilbert space problem considered in Section III.

For this purpose let us restrict to the class of functions with the following representation in some domain of $c^{2}$

$$
\begin{equation*}
\mathrm{G}(\mathrm{w}, \mathrm{z})=\int_{-\infty}^{\infty} \frac{\mathrm{d} \sigma(\alpha, \beta)}{1-\mathrm{w} \alpha-\mathrm{z} \beta} \tag{16}
\end{equation*}
$$

where $\sigma(\alpha, \beta)$ is a bounded positive Radon measure in $\boldsymbol{R}^{2}$ and the formal (not necessarily convergent) double power series expansion

$$
\begin{equation*}
G(w, z) \simeq \sum_{m, n} w^{m} z^{n}\binom{m+n}{n} G_{m, n} \tag{17}
\end{equation*}
$$

exists, i.e., the moments $G_{m, n}=j \alpha^{m_{\beta}} \mathrm{d} \sigma(\alpha, \beta)$ are finite. The class of functions with the integral representation (16) may be considered as one possible generalization of the extended Stieltjes functions to the case of two variables.

Let us define the subset $\sigma_{\sigma}(\mathrm{t})$ of the complex t-plane

$$
\begin{equation*}
\Theta_{\sigma}(t) \equiv\left\{t \equiv \mathrm{w} \alpha+\mathrm{z} \beta \mid\{\alpha, \beta\} \in \Sigma_{\sigma}\right\} \tag{18}
\end{equation*}
$$

where $\Sigma_{\sigma}$ is the convex hull of the support of $\sigma(\alpha, \beta)$ in $R^{2}$. We can state now the following convergence theorem for the approximants $G_{N}(w, z)$ defined by

Eq. (13):

Theorem 4. Let $\Delta$ be a domain of $C^{2}$ such that the point 1 is at positive distance d from $\overline{@_{\sigma^{(t)}}}$. If $\sum_{m=0}^{\infty}\left(G_{2 m, 0}\right)^{-1 / 2 m}=\infty, \sum_{n=0}^{\infty}\left(G_{0,2 n}\right)^{-1 / 2 n}=\infty$ and $\{w, z\} \in \Delta$, then $G_{N}(w, z)$ converges to $G(w, z)$ as $N \rightarrow \infty$. The convergence is uniform in any bounded subset $\Gamma \subset \Delta$. Proof. Let $\mathscr{L}_{2}\left(R^{2}, \sigma\right)$ be the Hilbert space of the functions on $R^{2}$, square integrable with the measure $\sigma(\alpha, \beta)$. Consider the multiplication operators $\hat{\alpha}$ and $\hat{\beta}$ defined by $\hat{\alpha} \mathrm{g}(\alpha, \beta) \equiv \alpha \mathrm{g}(\alpha, \beta)$ and $\hat{\beta} \mathrm{g}(\alpha, \beta) \equiv \beta \mathrm{g}(\alpha, \beta)$. They are selfadjoint permutable operators in $\mathscr{L}_{2}\left(R^{2}, \sigma\right)$ and the constant vector $u(\alpha, \beta) \equiv 1$ is quasi-analytic for both $\hat{\alpha}$ and $\hat{\beta}$ by assumption, since $\left\|\hat{\alpha}^{\mathrm{m}} \hat{\beta}^{\mathrm{n}} \mathrm{u}\right\|_{\sigma}^{2}=\iint \alpha^{2 \mathrm{~m}_{\beta}^{2 \mathrm{n}}} \mathrm{d} \sigma(\alpha, \beta)=\mathrm{G}_{2 \mathrm{~m}, 2 \mathrm{n}}$. Therefore the operators $\hat{\alpha}$ and $\hat{\beta}$ and the vector $u(\alpha, \beta)$ satisfy the same hypotheses as $\mathrm{A}, \mathrm{B}$ and f considered in Section III. Clearly $G(w, z)=\left(u,(1-w \hat{\alpha}-z \hat{\beta})^{-1} u\right)_{\sigma}$ and $G_{N}(w, z)=\left(u,\left(1-w \hat{\alpha}_{N}{ }^{-z} \hat{\beta}_{N}\right)^{-1} u\right)_{\sigma}$, where $\hat{\alpha}_{N}$ and $\hat{\beta}_{N}$ are defined like $A_{N}$ and $B_{N}$ by Eq. (5). Furthermore it is easy to see that $\overline{\Theta_{\sigma}(t)} \equiv \overline{\Theta(w \hat{\alpha}+\mathrm{z} \hat{\beta})}$ and the theorem follows from Theorem 3.

Instead of assuming the integral representation (16) we could as well start from the series (17). In this case sufficient conditions for the double sequence $\left\{G_{m, n}\right\}$ to be a determined moment double sequence have been given in Theorem 10 of Ref. 13. $\left\{G_{n, m}\right\}$ must satisfy a certain positivity condition and both the sequences $\left\{G_{m, 0}\right\}$ and $\left\{G_{0, n}\right\}$ must satisfy the Carleman criterion: $\sum_{m=0}^{\infty}\left(G_{2 m, 0}\right)^{-1 / 2 m}=\infty$ and $\sum_{n=0}^{\infty}\left(G_{0,2 n}\right)^{-1 / 2 n}=\infty$. Since the positivity condition is necessary for $\left\{G_{m, n}\right\}$ to be a moment sequence, the two starting points are equivalent.

## V. GENERALIZATIONS AND FINAL REMARKS

The extension of our results to any number of selfadjoint permutable operators $A_{1}, A_{2}, \ldots, A_{p}$ is straightforward. The operator $T^{(p)}=z_{1} A_{1}+z_{2} A_{2}+\ldots$ $+z_{p} A_{p}$ is still normal maximal and we can repeat all the considerations of Section III, ending up with convergence theorems which generalize Theorem 3. The structure of the matrix element (f, $R_{N}\left(z_{1}, z_{2}, \ldots, z_{p}\right) f$ ) still suggests an approximation scheme which can be used for any function of $p$ complex variables given by its formal multiple power series expansion

$$
\begin{equation*}
G\left(z_{1}, z_{2}, \ldots, z_{p}\right) \cong \sum_{\left\{n_{i}\right\}} \frac{\left(\sum_{i=1}^{p} n_{i}\right)!}{\prod_{i=1}^{p} n_{i}!} G_{n_{1}, n_{2}}, \ldots, n_{p} z_{1}^{n_{1}} z_{2}^{n_{2}} \ldots z_{p}^{n_{p}} \tag{19}
\end{equation*}
$$

In fact we can still write the expression

$$
\begin{equation*}
G_{N}\left(z_{1}, z_{2}, \ldots, z_{p}\right)=G_{N}^{T}\left(Q_{N}\right)^{-1} G_{N} \tag{20}
\end{equation*}
$$

where the vector $G_{N}$ and the matrix $Q_{N}$ are now defined by

$$
\begin{gather*}
\left(G_{N}\right)_{n_{1}}, \ldots, n_{p}=G_{n_{1}-n_{2}, n_{2}-n_{3}, \ldots, n_{p-1}-n_{p}, n_{p}} \\
\left\langle Q_{N}\right\rangle_{n_{1}}, \ldots, n_{p} ; m_{1}, \ldots, m_{p}=G_{n_{1}-n_{2}+m_{1}-m_{2}, n_{2}-n_{3}+m_{2}-m_{3}, \ldots, n_{p}+m_{p}} \\
-\sum_{i=1}^{p} z_{i} G_{n_{1}-n_{2}+m_{1}-m_{2}+\delta_{i 1}, n_{2}-n_{3}+m_{2}-m_{3}+\delta_{i 2}, \ldots, n_{s}-n_{s+1}+m_{s}-m_{s+1}+\delta_{i s}, \ldots, n_{p}+m_{p}+\delta_{i p}} \\
n_{1}=0, \ldots, N ; \quad n_{2}=0, \ldots, n_{1} ; \quad n_{p}=0, \ldots, n_{p-1} ;  \tag{21}\\
m_{1}=0, \ldots, N ; \quad m_{2}=0, \ldots, m_{1} ; \quad m_{p}=0, \ldots, m_{p-1}
\end{gather*}
$$

Like in the $\mathrm{p}=2$ case the convergence properties of $\mathrm{G}_{\mathrm{N}}\left(\mathrm{z}_{1}, \mathrm{z}_{2}, \ldots, \mathrm{z}_{\mathrm{p}}\right)$ to $G\left(z_{1}, z_{2}, \ldots, z_{p}\right)$ can be obtained from the study of ( $f, R_{N}\left(z_{1}, z_{2}, \ldots, z_{p}\right) f$ ).

Since $R_{N}(w, z)$ converges strongly to $R(w, z)$ on the Hilbert space $\mathscr{H}_{f}$, we can also apply the method of moments to the equation

$$
\begin{equation*}
(1-\mathrm{wA}-\mathrm{zB}) \psi=\mathrm{g} \tag{22}
\end{equation*}
$$

where $g$ is any vector in $\mathscr{H}_{\mathrm{f}}$. Then

$$
\begin{equation*}
\left(g, R_{N}(w, z) g\right)=\sum_{r=0}^{N} \sum_{s=0}^{r} \sum_{p=0}^{N} \sum_{q=0}^{p} E_{r-s, s}^{*}\left(M_{N}\right)_{r, s ; p, q}^{-1} E_{p-q ; q} \equiv E^{T *} M_{N}^{-1} E \tag{23}
\end{equation*}
$$

where the matrix $M$ is defined as in Eq. (10) in terms of the matrix elements (f, $A^{m} B_{f}$ ) only, while the column matrix $E_{m-n, n}$ is

$$
\begin{equation*}
E_{m-n, n} \equiv\left(f, A^{m-n} B^{n_{g}}\right) \tag{24}
\end{equation*}
$$

Although Eq. (23) does not seem relevant for the study of approximants to a general power series, the freedom in the choice of the generating vector $f$ can be used to improve the approximation in purely Hilbert space problems. In fact, a simple variational formulation is available for the approximation procedure we have been discussing. More precisely, consider the functional

$$
\begin{equation*}
J=(\mathrm{g}, \phi)+\left(\phi^{\prime}, \mathrm{g}\right)-\left(\phi^{\prime},(1-\mathrm{wA}-\mathrm{zB}) \phi\right) \tag{25}
\end{equation*}
$$

and choose the following natural ansatz

$$
\begin{align*}
& \phi=\sum_{m=0}^{N} \sum_{n=0}^{m} a_{m n} A^{m-n} B_{f}^{n}  \tag{26}\\
& \phi^{\prime}=\sum_{m=0}^{N} \sum_{n=0}^{m} a_{m n}^{\prime} A^{m-n} B_{f}^{n} \tag{26}
\end{align*}
$$

Then the stationary value $\bar{J}$ of $J$ with respect to the parameters $\left\{a_{m n}\right\}$ and $\left\{a_{m n}^{\prime}\right\}$ coincides with formula (23). $\bar{J}$ can still be made stationary even with respect to the choice of the vector f . ${ }^{(21)}$ The extension of these considerations to the n-dimensional case is immediate.

In Section II, instead of starting from Eq. (2), we could as well start from the equation:

$$
\begin{equation*}
\mathrm{f}=(1-\mathrm{wA})(1-\mathrm{zB}) \psi \tag{27}
\end{equation*}
$$

All convergence theorems of Section III hold with obvious modifications for the normal maximal operator $\widetilde{T}_{N}(w, z) \equiv P_{N}(W A+z B-W z A B) P_{N}$ and a simple sufficient condition for $1 \notin(\widetilde{T}(w, z))$ is in this case that both $\operatorname{Im} w \neq 0$ and $\operatorname{Im} \mathrm{z} \neq 0$. Therefore, we are led to consider functions of two complex variables with the following integral representation:

$$
\begin{equation*}
\widetilde{G}(w, z)=\iint \frac{d \sigma(\alpha, \beta)}{(1-w \alpha)(1-z \beta)} \simeq \sum_{m, n=0}^{\infty} w^{m} z^{n} G_{m, n} \tag{28}
\end{equation*}
$$

where, again, $\sigma(\alpha, \beta)$ is a positive bounded Radon measure in $R^{2}$ and $G_{m, n}$ are its moments. For these functions we introduce the approximants

$$
\begin{equation*}
\widetilde{G}_{N}(w, z) \equiv G_{N}^{T} \widetilde{Q}_{N}^{-1} G_{N} \tag{29}
\end{equation*}
$$

which differ from the approximants $G_{N}(w, z)$ defined in Eq. (13) only for the matrix $\widetilde{\mathrm{Q}}_{\mathrm{N}}$ which now reads:

$$
\begin{align*}
\left(\widetilde{Q}_{N}\right\rangle_{p, q ; r, s}=G_{p+r-q-s, q+s} & -w G_{p+r-q-s+1, q+s} \\
& -z G_{p+r-q-s, q+s+1}+w z G_{p+r-q-s+1, q+s+1} \tag{30}
\end{align*}
$$

If the Carleman condition is satisfied for both the sequences $\left\{G_{m, 0}\right\}$ and $\left\{G_{0, n}\right\}$, we can repeat the proof of Theorem 4 and conclude that $\widetilde{G}_{N}(w, z) \rightarrow \widetilde{G}(w, z)$ at least for both $\operatorname{Im} w \neq 0$ and $\operatorname{Im} z \neq 0$. Again the convergence is uniform in compact sets of $c^{2}$.

Also for these approximants the extension to the n-dimensional case as well as the variational formulation are straightforward.

## Acknowledgements

One of us (C.A.) is indebted to Prof. J. L. Mc Gregor for a very useful discussion. P. B. thanks Prof. M. L. Goldberger for his hospitality at the Department of Physics of Princeton University where part of this work has been done and both of us thank Prof. S. Drell for his hospitality at SLAC.

## REFERENCES

1. G. A. Baker Jr. in Advances in Theorectical Physics, K. A. Brueckner, ed. (Academic Press, New York, 1965); Vol. 1, p. 1.
2. G. A. Baker Jr. and J. L. Gammel, eds., Padé Approximant in Theoretical Physics (Academic Press, New York, 1970); Vol. 71.
3. P. R. Graves Morris, ed. Padé Approximants and Their Applications (Academic Press, London, 1973).
4. J. Nuttall, J. Math. Anal. and App. 31, 147 (1970).
5. J.S.R. Chisholm, Math. of Comp. 27, 841 (1973).
6. A. K. Common, P. R. Graves Morris, "Some properties of Chisholm approximants, " University of Canterbury preprint (1973); R. Hughes Jones, "General rational approximants in N variables," University of Canterbury preprint (1973).
P. R. Graves Morris, R. Hughes Jones, G. Makinson, "The calculation of some general approximants in two variables, " University of Canterbury preprint (1973).
7. J.S.R. Chisholm, P. R. Graves Morris, "Generalizations of the theorem of de Montessus to two-variable approximants," University of Canterbury preprint (1974).
8. C. H. Lutterodt, "A two-dimensional analogue of Pade approximant theory," University of Birmingham preprint (1973).
9. Y. V. Vorobyev, Method of Moments in Applied Mathematics (Gordon and Breach, New York, 1965). See also J. A. Shohat and J. D. Tamarkin, The Problem of Moments (American Methematical Society, New York, 1943), and N. I. Akhiezer, The Classical Moment Problem, and Some Related Questions in Analysis (Hafner, New York, 1965).
10. D. Masson, "Hilbert space and the Pade approximant," Padé Approximant in Theoretical Physics (Academic Press, New York, 1970); Vol. 71, p. 197.
11. J. Nuttal, "Variational principles and Pade approximants," and references therein, Padé Approximants and Their Applications (Academic Press, London, 1973).
12. A. Erdelyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, Higher Transcendental Functions (McGraw, New York, 1953); Vol. I.
13. A. E. Nussbaum, Arkiv för Mat. 6, 179 (1965).
14. A vector $f \in \underset{n>1}{\cap} \mathscr{D}\left(A^{n}\right)$ such that $\rangle_{n=1}^{\infty}\left\|A^{n_{f}}\right\|^{-1 / n}=\infty$ is called a quasianalytic vector for the operator A . Given the spectral family $\mathrm{E}(\alpha, \beta)$ associated with the permutable selfadjoint operators $A$ and $B$, the vectors $\mathrm{g}_{\Gamma}=\iint_{\Gamma} \mathrm{dE}(\alpha, \beta) \mathrm{g}$ where $\mathrm{g} \in \mathscr{H}$ and $\Gamma$ is any compact domain in $R^{2}$, are quasi-analytic vectors for both A and B and they are dense in $\mathscr{H}$.
15. The assumption that $f$ is a quasi-analytic vector is a sufficient condition in the convergence proof we shall give in Section III.
16. Even if the vectors $\left.{ }^{\prime} f_{r, s}\right\}$ are not all linearly independent, all following results will remain valid.
17. The set of complex numbers $\Theta(\mathrm{O})=\{(\mathrm{h}, \mathrm{Oh}) \mid \mathrm{h} \in \mathscr{D}(\mathrm{O}),\|\mathrm{h}\|=1\}$ is called the numerical range of the operator $O$.
18. M. Stone, Linear Transformations in Hilbert Space and Their Applications to Analysis (American Mathematical Society, New York, 1932), Theor. 8.8.
19. Ref. 18, Theor. 4. 20.
20. Ref. 18, Theor. 8. 14.
21. For applications of these ideas in the one-dimensional case, see C. Alabiso, P. Butera, G. M. Prosperi, Nucl. Phys. B42, 493 (1972); B46, 593 (1972).

[^0]:    *Work supported in part by the U. S. Atomic Energy Commission.
    $\dagger$ Consiglio Nazionale delle Ricerche fellow.
    $\dagger \dagger$ On leave of absence from Istituto Nazionale di Fisica Nucleare, sezione di Milano, Consiglio Nazionale delle Ricerche fellow.

