ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 4, Number 2, Spring 1974

VARIATIONAL APPROACH TO THE THEORY OF OPERATOR PADÉ APPROXIMANTS

J. D. BESSIS AND J. D. TALMAN

I. **Definition and Properties of the Operator Padé Approximants.** In many applications in theoretical physics, solutions of a problem appear as matrices or operators, and these operators are expressed as formal power series in some parameter (the coupling constant). In many situations the resulting power series are thought to be at best only asymptotic and also the coupling constant may be large, so that even if the series should converge, it may be of limited use. In these circumstances, it is natural to attempt to improve the validity or usefulness of the theory through the use of Padé approximants (P.A.s) even though no mathematical justification for such a step may exist.

We will be concerned with a formal power series

$$T(z) = \sum_{n=0}^{\infty} T_n z^n, \tag{1}$$

where the T_n are matrices or operators in some space. At least two different types of P.A. to $T(z)$ can be considered. The matrix elements of $T(z)$ in some basis may be regarded as a family of power series for each of which a P.A. can be constructed, or a P.A. which is a quotient of polynomial operators can be constructed. The latter approach seems more natural, since it is, as will be shown, basis independent, and it is this approach that will be considered here. In many practical examples, the spaces in question are direct products and intermediate P.A.s, scalar in some indices and operator in others, could be considered. It should be remarked that the individual matrix elements of an operator P.A. may be complicated functions of z; for example, the denominator of the $[M, N]$ P.A. to a matrix in n dimensions is a polynomial of degree nN.

We consider two types of P.A. to T, the right and the left, of the forms

$$\begin{aligned} [M, N]_T(z) &= P_M(z)[Q_N(z)]^{-1} \\ &= \left(\sum_{i=0}^{M} p_i z^i\right)\left(\sum_{j=0}^{N} q_j z^j\right)^{-1}, \end{aligned} \tag{2a}$$

$$[\tilde{M}, \tilde{N}]_T(z) = [\tilde{Q}_N(z)]^{-1}\tilde{P}_M(z)$$

Received by the editors February 8, 1973.

(2b)
$$= \left(\sum_{j=0}^{N} \tilde{q}_j z^j \right)^{-1} \left(\sum_{i=0}^{M} \tilde{p}_i z^i \right),$$

respectively. These are defined by the relations

(3a) $$[M, N]_T(z) - T(z) = O(z^{M+N+1}),$$

(3b) $$[\tilde{M}, \tilde{N}]_T(z) - T(z) = O(z^{M+N+1}),$$

or by the relations that will be assumed to be equivalent,

(4a) $$P_M(z) - T(z)Q_N(z) = O(z^{M+N+1})$$

(4b) $$\tilde{P}_M(z) - \tilde{Q}_N(z)T(z) = O(z^{M+N+1}).$$

It will be assumed throughout that the P.A.s exist and are uniquely determined.

Other P.A.s of the form e.g. $P_i(z)Q(z)^{-1}P_j(z)$, $i + j = M$, could be defined, but only those in which all the factors of the denominator are either to the left of, or the right of, those of the numerator will be considered here.

The coefficients of the P.A.s satisfy the equations

(5a) $$p_i = \sum_{j=0}^{(N,i)} T_{i-j} q_j, \quad i = 0, \cdots, M$$

(5b) $$0 = \sum_{j=0}^{(N,i)} T_{i-j} q_j, \quad i = M + 1, \cdots, M + N$$

where $(N, i) = \min(N, i)$. There is a similar set of equations for the $\tilde{p}_i$, $\tilde{q}_j$. If the P.A.s exist, these equations will have a solution. It is noted that p_i and q_j can be multiplied on the right by an arbitrary invertible matrix; it can then be assumed that $q_0 = I$.

The right and left P.A.s defined above are in fact equal. This follows since

$$P_M(z)[Q_N(z)]^{-1} - [\tilde{Q}_N(z)]^{-1}\tilde{P}_M(z) = O(z^{N+M+1})$$

which implies that

$$\tilde{Q}_N(z)P_M(z) - \tilde{P}_M(z)Q_N(z) = O(z^{N+M+1})$$

The left hand side is a polynomial of degree $N + M$ which must therefore vanish identically. It then follows that

$$P_M(z)[Q_N(z)]^{-1} = [\tilde{Q}_N(z)]^{-1}\tilde{P}_M(z).$$

It can be seen from any of the equation (3)–(5) that if $T(z)$ undergoes a similarity transformation, $P(z)$, $Q(z)$, p_i, q_j and $[M, N]_T(z)$ undergo the same similarity transformation, so that the operator P.A.s are basis independent. It is also noted that if $T(z)$ is multiplied on the left by A, p_i becomes Ap_i, while if A multiplies $T(z)$ on the right p_i and q_j become p_iA and $A^{-1}q_jA$.

The following properties follow readily from the definition of operator P.A.s and the fact that the left hand and right hand P.A.s are equal.

The P.A. to the direct sum of operators is the direct sum of the P.A.s; this is useful if $T(z)$ is reducible under a group of transformations.

If $[M, N](z)$ is the P.A. to $T(z)$, $[N, M](z)^{-1}$ is the P.A. to $T(z)^{-1}$, $[M, N](z^*)^\dagger$ is the PA. to $T(z^*)^\dagger$, $[M, N]^*(z^*)$ is the P.A. to $T^*(z^*)$, and $\widetilde{[M, N]}(z)$ is the P.A. to $\tilde{T}(z)$. If $M - N \geqq 0$, the $[M, N]$ P.A. to $T(z) + R(z)$, where $R(z)$ is a polynomial of degree less than or equal to $M - N$, is $[M, N] + R(z)$.

From the above facts, it can be proved that the diagonal ($M = N$) P.A. to the homographically transformed operator $(AT + B)(CT + D)^{-1}$ is $(A[N, N](z) + B)(C[N, N](z) + D)^{-1}$, provided C^{-1} exists. This follows by writing $(AT + B)(CT + D)^{-1} = AC^{-1} + (B - AC^{-1}D)(T - C^{-1}D)^{-1}C^{-1}$.

A unitary operator function is one that satisfies $S(z)S(z^*)^\dagger = I$; it is unitary for real z. It can be proved that the diagonal P.A.s to a unitary operator are themselves unitary. If $[N, N](z)$ is the P.A. to $S(z)$, $[N, N](z)^{-1}$ is the P.A. to S^{-1}, and $[N, N](z^*)^\dagger$ is the P.A. to $S(z^*)^\dagger$. Since $S(z)^{-1} = S(z^*)^\dagger$, it follows that $[N, N](z^*)^\dagger = [N, N](z)^{-1}$ and that the P.A. is unitary.

The interested reader will find in [1] more details on the transformation properties of Padé approximations in a non-commuting algebra.

II. **The Variational Principles.** Many problems of physical interest can be expressed in the form of a variational principle. One of these is known as the Schwinger variational principle in which the quantity of physical interest, the T matrix, is expressed as the stationary value of the functional [2], [1]

$$S = (\Psi_2^-, V\Phi_1)(\Psi_2^-, [V - zVG^+V]\Psi_1^+)^{-1}(\Phi_2, V\Psi_1^+) \tag{6}$$

where $G^\pm$ is an integral resolvent operator and V is a hermitian potential energy operator; $\Psi_i^\pm$ are solutions of the Lippmann-Schwinger equation [3]

$$\Psi_i^\pm = \Phi_i + zG^\pm V\Psi_i \tag{7}$$

and Φ_i are free particle plane wave functions.

Equation (7) is equivalent to the Schrodinger equation $(E - H_0 - zV)\Psi^{\pm} = 0$ together with the boundary condition that $\Psi^{\pm}$ is asymptotic to Φ plus an outgoing or incoming spherical wave.

It has been observed ([4], [5], [6]) that if the stationary value of S is calculated for approximate wave functions

$$\Psi_1{}^{+} = \sum_{i=1}^{N} z^i(G^+V)^i\Phi_1\mu_i \tag{8a}$$

$$\Psi_2{}^{-} = \sum_{j=1}^{N} z^j\lambda_j{}^*(G^-V)^j\Phi_2 \tag{8b}$$

with respect to variations in μ_i and λ_j, the result is the $[N, N+1]$ P.A. to the series expansion

$$(\Phi_2, T\Phi_1) = \sum_{n=0}^{\infty} z^n(\Phi_2, (VG^+)^nV\Phi_1) \tag{9}$$

to the T matrix.

It will be shown that this result applies also to the case in which Φ_i and Ψ_j are interpreted as vectors whose components are themselves wave functions; this situation arises in the theory of reaction processes. In this case μ_i and λ_j are matrices to be varied arbitrarily.

If the expressions (8) are substituted into S, it is found that

$$S = AB^{-1}\bar{A}, \tag{10}$$

where

$$A = \sum_{i=0}^{N} z^iT_i\mu_i \tag{11}$$

$$\bar{A} = \sum_{j=0}^{N} z^j\lambda_jT_j \tag{12}$$

$$B = \sum_{n=0}^{N}\sum_{m=0}^{N} z^{m+n}\lambda_m(T_{m+n} - zT_{m+n+1})\mu_n \tag{13}$$

and

$$T_i = (\Phi_2, V(G^+V)^i\Phi_1) \tag{14}$$

is the coefficient of z^i in the perturbation series expansion of the S matrix. In calculating S, it is necessary to use the physical properties that $V^\dagger = V$ and $(G^-)^\dagger = G^+$.

The mathematical problem can now be stated independently of its physical origin: given S defined by (10), (11), (12) and (13), show that the solution of $\delta S = 0$ is the $[N, N+1]$ operator P.A. constructed from the $T_j (j = 0, 1, \cdots, 2N+1)$.

The condition that S should be stationary with respect to variations in λ_j leads to the equations

$$0 = \frac{\delta S}{\delta \lambda} = \frac{\delta A}{\delta \lambda_j} B^{-1}\bar{A} - AB^{-1}\frac{\delta B}{\delta \lambda_j} B^{-1}\bar{A} + AB^{-1}\frac{\delta \bar{A}}{\delta \lambda_j}$$

$$= AB^{-1} \sum_{n=0}^{N} z^{j+n}(T_{j+n} - zT_{j+n+1})\,\mu_n B^{-1}\bar{A} + AB^{-1}z^j T_j$$

or

$$\sum_{n=0}^{N} z^n (T_{j+n} - zT_{j+n+1})\,\mu_n = T_j\bar{A}^{-1}B, \tag{15}$$

for $j = 0, 1, \cdots, N$. Similarly, the condition that S be stationary with respect to variations in the μ_i leads to the equations

$$\sum_{m=0}^{N} z^m \lambda_m (T_{m+i} - zT_{m+i+1}) = BA^{-1}T_i, \tag{16}$$

for $i = 0, 1, \cdots, N$.

It can be observed from the form of S that it is invariant under multiplications of λ_i on the left, and μ_j on the right by arbitrary matrices. From this it can be seen that $\bar{A}^{-1}B$ and BA^{-1} could be replaced by I in equations (15) and (16) respectively. This implies that the condition that S be stationary with respect to variations in λ_i and μ_j is equivalent to the condition that the bilinear functional

$$S' = A - B + \bar{A}$$

be stationary.

Equations (15) can also be written

$$T_j(\mu_0 - \bar{A}^{-1}B) + \sum_{n=1}^{N} T_{j+n}z^n(\mu_n - \mu_{n-1}) - z^{N+1}T_{j+N+1}\mu_N = 0.$$

This can be viewed as a set of $N+1$ equations in $N+2$ unknowns, $\mu_0 - \bar{A}^{-1}B$, $\mu_n - \mu_{n-1}$ for $n = 1, \cdots, N$, and $-\mu_N$. Comparison of these equations with (5b) in the case of the $[N, N+1]$ P.A. shows that the solution of these equations is related to the q_j of the P.A. by

$$\mu_0 - \bar{A}^{-1}B = q_{N+1}C \tag{17a}$$

$$z^{i+1}(\mu_{i+1} - \mu_i) = q_{N-i}C, \quad i = 0, \cdots, N-1 \tag{17b}$$

$$-z^{N+1}\mu_N = q_0 C \tag{17c}$$

where C is a constant matrix of proportionality.

Equations (17b) and (17c) can be readily solved for $\mu_N, \mu_{N-1}, \cdots, \mu_0$ by multiplying (17b) by z^{N-i} and adding the last n equations. It is then found that

$$\mu_{N-n} = -\left(\sum_{j=0}^{n} q_j z^j\right) C/z^{N+1}$$

for $n = 0, \cdots, N$. If μ_0 is substituted into equation (17a) it is found that

$$-z^{N+1}\bar{A}^{-1}B = \left(\sum_{j=0}^{N+1} q_j z^j\right) C$$

and hence that

$$C = -z^{N+1}Q_{N+1}(z)^{-1}\bar{A}^{-1}B,$$

where $Q_{N+1}(z)$ is the denominator of the $[N, N+1]$ P.A. It follows that

$$\mu_n = \left(\sum_{j=0}^{N-n} q_j z^j\right) Q_{N+1}(z)^{-1}\bar{A}^{-1}B. \tag{18}$$

It is now possible to calculate the stationary value of S, using (5a), (10), (11) and (18)

$$S = AB^{-1}\bar{A} = \sum_{i=0}^{N} z^i T_i \mu_i B^{-1}\bar{A} = \sum_{i=0}^{N} z^{i+j} T_i \mu_j Q_{N+1}(z)^{-1}$$

$$= \sum_{n=0}^{N} z^n \left(\sum_{j=0}^{n} T_{n-j}\mu_j\right) Q_{N+1}(z)^{-1} \tag{19}$$

$$= \sum_{n=0}^{N} p_n z^n Q_{N+1}(z)^{-1} = [N, N+1]_T(z).$$

It is evident that solution of equations (16) for λ_i will lead to the left hand P.A. $Q_{N+1}(z)^{-1}P_N(z)$ which we have remarked is the same as $P_N(z)Q_{N+1}(z)^{-1}$. This reflects the fact that equations (15) and (16) are not independent; in particular, if the T_n are hermitian, $\lambda_i = \mu_i{}^{\dagger}$.

A larger family of P.A.s is obtained if the first $p\,\lambda_i$ and μ_j are not varied, but fixed at the values BA^{-1} and $\bar{A}^{-1}B$ respectively. Equations (15) and (16) can then be shown to lead to the $[N + p, N - p + 1]$ P.A. In particular, for $p = N + 1$, there are no variation parameters and the result is the $(2N + 1)$ order perturbation approximation.

A second variational principle that is frequently used is the Kohn [7] variational principle in which the quantity varied is the functional

$$L = (\Psi_2{}^-, (E - H)\Psi_1{}^+) + 4\pi f_{21},$$

where f_{21} is the scattering amplitude (the amplitude of $\Psi_1{}^+$ in the direction of the plane wave part of $\Psi_2{}^-$). It is possible to write

$$4\pi f_{21} = -(\Phi_2, (E - H)\Psi_1) - z(\Phi_2|V|\Psi_1)$$

which leads to the expression

$$L = (\Psi_2 - \Phi_2, (E - H)\Psi_1) - (\Phi_2|V|\Psi_1). \tag{20}$$

If the Cini-Fubini approximation (8) [8] is substituted in (20) it is found that:

$$\begin{aligned} L = \sum_{n=1}^{N} \sum_{m=1}^{N} z^{m+n}\lambda_n(T_{m+n-1} - zT_{m+n})\mu_m \\ - \sum_{n=1}^{N} z^{n+1}\lambda_n T_n - \sum_{m=1}^{N} z^{m+1}T_m\mu_m - zT_0. \end{aligned} \tag{21}$$

An analysis similar to the preceding shows that the stationary value of L is $[N, N]_T(z)$. Also, if the first p coefficients λ_i and μ_j are fixed, the stationary value is $[N + p, N - p]_T(z)$.

Acknowledgments. One of us (D.B.) would like to thank Prof. J. W. McGowan, chairman of the Physics Department at the University of Western Ontario, for his very warm hospitality, while this paper was being written.

References

1. D. Bessis, *Topics in the theory of Padé approximants* in *Padé Approximants,* edited by P. R. Graves-Morris, The Institute of Physics, London, (1973), 19–44.

2. B. A. Lippmann and J. Schwinger, *Variational principles for Scattering Processes,* Physical Review 79 (1950), 469–480.

3. ——, *Variational principles for Scattering Processes,* Physical Review 79 (1950), 469–480.

4. L. Caneschi and R. Jengo, *On the Problem of Bound States in the Framework of Padé Approximants,* Nuovo Cimento **60A** (1969), 25–32.

5. D. Bessis and M. Pusterla, *Unitary Padé Approximants in Strong Coupling Field Theory and Application to the Calculation of the* ρ *and of meson Regge Trajectories,* Nuovo Cimento **54A** (1968), 243–294.

6. J. Nuttall, *The connection of Padé Approximant with Stationary Variational Principles and the Convergence of certain Padé Approximants,* in *The Padé Approximants in theoretical Physics,* G. A. Baker, Jr. and J. L. Gammel, editors, Academic Press, 1970, p. 219–229.

7. W. Kohn, *Two applications of the Variational Method to Quantum Mechanics,* Physical Review 71 (1947), 635–637.

8. M. Cini and S. Fubini, *Non perturbation treatment of scattering in Quantum Field Theory,* Nuovo Cimento **11** (1954), 142–152.

CENTRE D'ETUDES NUCLÉAIRES DE SACLAY, SACLAY, FRANCE, AND
UNIVERSITY OF WESTERN ONTARIO, LONDON, ONTARIO, CANADA