Higher Order Variational Principles and Padé Approximants for Linear Functionals

R.W. Conn

May 1973

UWFDM-102

FUSION TECHNOLOGY INSTITUTE

UNIVERSITY OF WISCONSIN

MADISON WISCONSIN

Higher Order Variational Principles and Padé Approximants for Linear Functionals

R.W. Conn

Fusion Technology Institute
University of Wisconsin
1500 Engineering Drive
Madison, WI 53706

http://fti.neep.wisc.edu

May 1973

UWFDM-102

Higher Order Variational Principles and
Padé Approximants for Linear Functionals

by

Robert W. Conn

Nuclear Engineering Department
University of Wisconsin
Madison, Wisconsin 53706

FDM 102

Reprint: Nuclear Science and Engineering 55 (1974) 468.

Abstract

The relationship between higher order variational principles for linear functionals of the solution to an inhomogeneous equation and Padé approximants for the same functional is shown. This leads to a deeper understanding of these higher order principles. Further, it is noted that in certain cases, the Roussopoulos functional can yield divergent results yet using the Ritz procedure, shown to be equivalent to forming Padé approximants for the functional of interest, gives a generalized Schwinger normalization independent variational principle that can yield finite and convergent results.

I. Introduction

The Padé approximant[1] has proven to be a powerful method for extracting information of a nonperturbative nature when only an underlying perturbation series is available. In this paper, the connection between Padé approximants, as discussed by Nuttall[2], and higher order variational principles, as discussed by Kostin and Brooks[3] and by Pomraning[4], is studied. The results give a more complete understanding of such higher order principles. In addition, they clarify some distinguishing features between the Roussopoulous[5] and Schwinger[6] variational functionals, and illustrate a case where the Roussopoulos functional will yield diverging results while the Schwinger functional will remain finite.

II. The Padé Approximants

One is frequently faced with the problem of extracting information about a function or a functional from knowledge of an underlying infinite series, such as a Taylor or Neumann series. In particular, it is often of interest to have such knowledge outside the radius of convergence of the underlying series. The Padé approximant method is an approach for obtaining such information. To review briefly, it is a rational approximation which, as noted by Baker[7], can be viewed as an approximate analytic continuation beyond the radius of convergence of the defining series. This interpretation will be used shortly in discussing the relationship of the higher order variational functionals of Kostin and Brooks[3] and Pomraning[4] to Padé approximants[2,7].

The $[N,M](x)$ Padé approximant is the ratio of two polynominals,

$$[N,M](x) = \frac{P_M(x)}{Q_N(x)} \tag{1}$$

where

$$P_M(x) = \sum_{i=0}^{M} p_i x^i \tag{2}$$

and

$$Q_N(x) = \sum_{i=0}^{N} q_i x^i \quad . \tag{3}$$

The coefficients, p_i and q_i, are determined from the basic series

$$f(x) = \sum_{k=0}^{\infty} m(k) x^k, \tag{4}$$

by equating powers of x in the expression

$$f(x)Q_N(x) - P_M(x) = \alpha x^{N+M+1} + \beta x^{N+M+2} + \ldots \quad . \tag{5}$$

The exact radius of convergence of the [N,N] - Padé approximants is not known but it is not limited to the unit circle(7) and all indications are that it is much larger.

III. Higher Order Variational Principles and Padé Approximants

It is of general interest to estimate linear functionals of the solution to an inhomogeneous equation. For example, in reactor theory, we are often interested in reaction rates when the flux is determined from a solution to the transport equation. Kostin and Brooks(3) and Pomraning(4) have considered variational principles, particularly ones of higher order, which can be used to estimate such reaction rates. We shall show in this section the connection between these higher order principles and Padé approximants.

Suppose the functional of interest is $(S^{\dagger},\phi)$ where ϕ satisfies the equation (e.g., the transport equation)

$$L\phi = S. \tag{6}$$

Here, L can be the Boltzmann transport operator[7] and S is a source. The operator L can be decomposed and written as

$$L = L_o - \Delta L \tag{7}$$

where ΔL can represent, for example, some change in the transport operator resulting from some physical change in the system of interest. The basic equation can then be written as

$$L_o\phi = S + \Delta L\phi \tag{8}$$

and can be approximately solved by iteration. The N^{th}-order approximation to ϕ is

$$\phi_N = \sum_{n=1}^{N} L_o^{-1} (\Delta L\, L_o^{-1})^{(n-1)} S. \tag{9}$$

Defining $\varepsilon = ||L_o^{-1}\Delta L||$, this series converges to the solution of (6) only if $\varepsilon < 1$. Two questions which immediately arise are, first, can one accelerate the convergence of the above series and, second,,can one extract useful information from (9) even if the series diverges? The asnwer to these questions is often yes, and the Padé approximant provides a method for proceeding.

Our interest is in $(S^{\dagger},\phi)$ and an approximate value is obtained from $(S^{\dagger},\phi_N)$, namely,

$$(S^{\dagger},\phi_N) = \sum_{n=1}^{N} (S^{\dagger},\ L_o^{-1}[\Delta L\, L_o^{-1}]^{n-1}S). \tag{10}$$

Define

$$m(n) = (S^{\dagger}, L_o^{-1}(\Delta L\, L_o^{-1})^{n-1} S) \tag{11}$$

and define a functional, $F(\phi_N;x)$ by

$$F(\phi_N;x) = \sum_{n=1}^{N} m(n) x^{n-1}. \tag{12}$$

Clearly, $F(\phi_N;1) = (S^{\dagger}, \phi_N)$.

Padé approximants can be formed from equation (12) and, in particular, the [N,N-1]-Padé approximant can be formed from

$$F(\phi_{2N};x) = \sum_{n=1}^{2N} m(n) x^{n-1} \tag{13}$$

using Nuttall's compact formula[9] as

$$[N,N-1](x) = \underline{m}^T \cdot \underline{\underline{M}}^{-1}(x) \cdot \underline{m} \tag{14}$$

where $\underline{m}$ is a vector of dimension N with elements m(n), T denotes transpose, and $\underline{\underline{M}}(x)$ is an NxN matrix with elements $M_{ij}(x)$ given by

$$M_{ij}(x) = m(i + j - 1) - x m(i + j) \;. \tag{15}$$

Setting x = 1 gives a rational approximation to $(S^{\dagger}, \phi)$.

It can now be shown that equation (14) is identical to certain higher order variational principles obtained from the Roussoupoulos principle after a Ritz optimization procedure[10] is used to obtain a Schwinger-like variational expression.

Kostin and Brooks[3] and Pomraning[4] have obtained higher order variational principles by using iterative solutions to the basic equation, $L\phi = S$, and its adjoint equation

$$L^{\dagger} \phi^{\dagger} = S^{\dagger}. \tag{16}$$

An outline of their derivation follows[4]. The Roussopoulos functional

$$I_R[\phi^\dagger,\phi] = (S^\dagger,\phi) + (\phi^\dagger, S - L\phi) \tag{17}$$

is stationary about the functional of interest, with equations (6) and (16) the appropriate Euler equations. As in equation (7) for L, write $L^\dagger$ as $L_o^\dagger - \Delta L^\dagger$. Then the adjoint equation can also be written as

$$L_o^\dagger \phi^\dagger = S^\dagger + \Delta L^\dagger \phi^\dagger \quad . \tag{18}$$

Define $\varepsilon^\dagger = ||(L_o^\dagger)^{-1} \Delta L^\dagger||$. As with equation (9), one can safely construct iterative solutions to equations (18) if $\varepsilon^\dagger$ is sufficiently small. The N^{th}-order iterates, ϕ_N and $\phi_N^\dagger$, are therefore

$$\phi_N = \sum_{n=1}^{N} L_o^{-1}(\Delta L\ L_o^{-1})^{n-1} S \tag{19}$$

$$\phi_N^\dagger = \sum_{n=1}^{N} L_o^{\dagger-1}(\Delta L^\dagger L_o^{\dagger-1})^{n-1} S \quad . \tag{20}$$

Using ϕ_N and $\phi_N^\dagger$ as trial functions in $I_R[\phi^\dagger,\phi]$ leads to an estimate of $(S^\dagger,\phi)$, namely, $I_R[\phi_N^\dagger,\phi_N]$, that is in error by $(\varepsilon\varepsilon^\dagger)^N$ [3,4]. If $\varepsilon \sim \varepsilon^\dagger$, then $I_R[\phi_N^\dagger, \phi_N]$ has an error, ε^{2N}, or $I_R[\phi_N^\dagger, \phi_N]$ is a higher order variational functional. Also, one can readily show using only the property, $(\phi^\dagger, L\phi) = (L^\dagger\phi^\dagger,\phi)$, that

$$I_R[\phi_N^\dagger,\phi_N] = \sum_{n=1}^{2N} m(n) \tag{21}$$

where

$$m(n) = (S^\dagger, L_o^{-1}\ (\Delta L\ L_0^{-1})^{n-1} S). \tag{22}$$

Therefore, $I_R[\phi_N^\dagger,\phi_N]$ is identical to $F(\phi_N,1)$ obtained simply by using ϕ_{2N} as an approximation to ϕ in $(S^\dagger,\phi)$. Clearly, (21) is identical to $(S^\dagger,\phi_{2N})$ (see eqn. (13)) and both are thus accurate through order $(\varepsilon)^{2N}$.

If the iterative series, equations (19) and (20), are in fact divergent, the use of ϕ_N and $\phi_N^\dagger$ as trial functions yields a variational approximation of order 2N that is also diverging as N increases. In this case, the Roussopoulos functional will clearly yield poor results. It is nevertheless possible to construct meaningful results for $(S^\dagger,\phi)$, even under these extreme conditions,[11] through the use of Padé approximants. The result is equation (14), and one can show this will be identical to the higher order variational result obtained <u>after</u> applying the Ritz procedure to ϕ_N and $\phi_N^\dagger$. To prove this, use as trial functions in $I_R(\phi^\dagger,\phi)$ the expressions

$$\psi_N = \sum_{n=1}^{N} c_n L_o^{-1} (\Delta L\, L_o^{-1})^{n-1} S \tag{23}$$

and

$$\psi_N^\dagger = \sum_{n=1}^{N} d_n L_o^{\dagger -1} (\Delta L^\dagger L_o^{\dagger -1})^{n-1} S^\dagger \tag{24}$$

where the c_n and d_n are constants. Minimizing $I_R(\psi_N^\dagger, \psi_N)$ with respect to the coefficients, c_n and d_n, leads to the functional

$$I_S\ (\psi_N^\dagger,\ \psi_N) = \underline{m}^T \cdot \underline{\underline{M}}^{-1}(1) \cdot \underline{m}\ . \tag{25}$$

This expression is identical to $[N,N-1](x)$, eqn. (14), with $x = 1$. Also note that for $N = 1$, eqn. (25) is identical to the ordinary Schwinger variational principle with lowest order perturbation theory trial functions.

The main attribute previously assigned to equation (25) is that I_S is independent of the normalization of the trial functions.[3,4] In fact, we see now it is indeed much more. Through this connection to Padé approximants, one can view the Ritz procedure and the results, equations (14) or (25), as yielding an approximate analytic continuation[7] of the defining perturbation expansion, equation (13) or (21), beyond its radius of convergence. Hence one can now expect, on theoretical grounds, meaningful results from equation (22) even if equations (19) and (20) are either slowly convergent, or even divergent. (This is supported by a recent numerical calculation performed in another context.[11]) In addition, the Roussopoulos functional, $I_R(\phi_N^\dagger, \phi_N)$, is diverging when the partial sums, ϕ_N and $\phi_N^\dagger$, are diverging whereas the generalized Schwinger functional, eqn. (25), is rational in x and finite. Thus, the removal of the normalization dependence from the Roussopoulous principle is akin, in higher order principles, to an approximate resummation of the basic perturbation series to yield a meaningful result.

References

1. G. H. Baker and J. L. Gammel, eds., The Padé Approximant in Theoretical Physics (Academic Press, New York, 1970).

2. J. Nuttall, ibid., p 219.

3. M. D. Kostin and H. Brooks, J. Math. Phys. 5, 1691 (1965).

4. G. Pomraning, Nucl. Sci. Eng. 29, 220 (1967).

5. P. Roussopoulos, C. R. Acad. Sci. Paris 236, 1858 (1953).

6. B. A. Lippmann and J. Schwinger, Phys. Rev. 79, 469 (1950). See also, R. G. Newton, Scattering Theory of Particles and Waves (McGraw-Hill, New York, 1966) p. 319.

7. See, for example, article by G. A. Baker, in reference 1, p. 1.

8. G. I. Bell and S. Glasstone, Nuclear Reactor Theory (Van Nostrand Reinhold Co., New York, 1970).

9. J. Nuttall, Phys. Rev. 157, 1312 (1967).

10. R. Courand and D. Hilbert, Methods of Mathematical Physics (Interscience Publishers, Inc., New York, 1953) Chapter IV.

11. H. Rabitz and R. W. Conn, Phys. Rev. A7, 577 (1973).