THE APPROXIMATION OF FUNCTIONS WITH BRANCH POINTS.

LESLIE SHORT.

A dissertation submitted for the degree of Doctor of Philosophy at the University of Kent at Canterbury.

December 1977.

To my Parents,

In Gratitude for Everything.

TABLE OF CONTENTS

ABSTRACT

In recent years Padé approximants have proved to be one of the most useful computational tools in many areas of theoretical physics, most notably in statistical mechanics and strong interaction physics. The underlying reason for this is that very often the equations describing a physical process are so complicated that the simplest (if not the only) way of obtaining their solution is to perform a power series expansion in some parameters of the problem. Furthermore, the physical values of the parameters are often such that this perturbation expansion does not converge and is therefore only a formal solution to the problem; as such it cannot be used quantitatively. However, the relevant information is contained in the coefficients of the perturbation series and the Padé approximants provide a convenient mathematical technique for extracting this information in a convergent way. A major difficulty with these approximants is that their convergence is restricted to regions of the complex plane free from any branch cuts; for example, the (N/N+j) Padé approximants to a series of Stieltjes converge to an analytic function in the complex plane cut along the negative real axis. The central idea of the present work is to obtain convergence along these branch cuts by using approximants which themselves have branch points.

The ideas presented in this thesis are expected to be only the beginning of a large investigation into the use of multi-valued approximants as a practical method of approximation.

In Chapter 1 we shall see that such approximants arise as natural generalisations of Padé approximants and possess many of the properties of Padé approximants; in particular, the very important property of homographic covariance. We term these approximants 'algebraic' approximants (since they satisfy an algebraic equation) and we are mainly concerned with the 'simplest' of these approximants, the quadratic approximants of Shafer. Chapter 2 considers some of the known convergence results for Padé approximants to indicate the type of results we may reasonably expect to hold (and to be able to prove) for quadratic (and higher order) approximants. A discussion of various numerical examples is then given to illustrate the possible practical usefulness of these latter approximants.

A major application of all these approximants is discussed in Chapter 3, where the problem of evaluating Feynman matrix elements in the physical region is considered; in this case, the physical region is along branch cuts. Several simple Feynman diagrams are considered to illustrate (a) the potential usefulness of the calculational scheme presented and (b) the relative merits of rational (Padé), quadratic and cubic approximation schemes.

The success of these general approximation schemes in one variable (as exhibited by the results of Chapters 2 and 3) leads, in Chapter 4, to a consideration of the corresponding approximants in two variables. We shall see that the two variable scheme developed for rational approximants can be extended in a very natural way to define two variable "t-power" approximants. Numerical results are presented to indicate the usefulness of these schemes in practice.

A final application to strong interaction physics is given in Chapter 5, where the analytic continuation of Legendre series is considered. Such series arise in partial wave expansions of the scattering amplitude. We shall see that the Padé Legendre approximants of Fleischer and Common can be generalised to produce corresponding quadratic Legendre approximants; various examples are considered to illustrate the relative merits of these schemes.

ACKNOWLEDGEMENTS

I wish to express my thanks to my supervisor Professor Chisholm for suggesting many of the ideas expressed here and for his constant help and encouragement, and also to Dr. Roberts and Dr. Graves-Morris for many helpful discussions. I would also like to acknowledge a Science Research Council grant which made this work possible.

CHAPTER 1: PADÉ APPROXIMANTS AND THEIR ONE VARIABLE GENERALISATIONS

1. INTRODUCTION

The central theme of this thesis is the extraction of quantitative information, about the solution of certain field theory problems, from perturbation series expansions. The motivation for considering such expansions is that they often provide the only feasible method of solution of a problem. Such series expansions have at best only a limited region of convergence and, more often than not, the region of interest in a particular problem lies outside this convergence domain; in certain cases the generated series has a zero radius of convergence and can only be considered as a formal series. The problem considered here is that of obtaining useful information from such series expansions; we clearly have to analytically continue the function under discussion beyond the circle of convergence of its power series.

In principle such a process is relatively simple (1). We compute the value of the function and as many derivatives as necessary, and to as high a degree of accuracy as required, at a new point within the circle of convergence which is closer to the point of interest than the original origin. In this way we generate a new convergent series expansion of the function and, by sufficient repetition of this process, the desired function value at the point of interest can be obtained (provided this is a non-singular point). There are obvious practical objections to this process:

(a) the amount of computation involved may become prohibitive, especially if the point of interest lies far from the convergence region of the series; excessive computation may lead to loss of accuracy, especially if the power series coefficients are not known exactly (which is invariably the case in practice);

(b) having obtained a convergent power series near the point of interest, the rate of convergence of this series may be so slow as to be of little practical use. Even if the point of interest lies within the convergence region of the original series, this slow rate of convergence may still be a problem. In such a case we require a method of accelerating the rate of convergence within the region of convergence;

(c) if the power series has zero radius of convergence the above process is inapplicable.

Any method of continuing a Taylor series outside its circle of convergence can be regarded as a summation procedure for divergent series. Many methods are available for defining the sum of a divergent series (2); one of the more well known is that due to Borel (3), which uses the formal identity

$$\sum_{n=0}^{\infty} a_n z^n = \int_0^{\infty} e^{-x} \sum_{n=0}^{\infty} a_n \frac{(xz)^n}{n!}\, dx \qquad (1.1)$$

where the series on the right hand side may converge even if the left hand side series does not. However, there is no useful way of truncating the integrand on the right hand side and the previous objection (a) still holds to some extent; at each point of interest we may have to perform a large number of integrations.

The most fruitful method of summing divergent series is based upon the idea of Padé (rational) approximants (4,5,6); we shall give the definition of these approximants in the next section. We can motivate the use of these approximants by considering Euler's method (2), which represents one of the simplest methods of analytic continuation. We suppose that we have a function $f(z)$, analytic in the z-plane cut from $-\infty$ to -1 along the real axis (for example, $\ln(1+z)$), and we wish to evaluate $f(2)$. Since the Taylor series for $f(z)$ converges for the point $z=2$ lies outside the circle of convergence. We make the transformation

$$z = \frac{y}{1-y} \qquad (1.2)$$

which is an example of an Euler transformation. The cut is mapped onto $[1,\infty]$ and $z=2$ is mapped to $y=2/3$, which is inside the convergence circle of $g(y)=f(y/1-y)$. We can therefore use a convergent Taylor series to evaluate $f(2)$, provided we use the homographic change of variable (1.2). For the particular example of $\ln(1+z)$ we obtain $-\ln(1-y)$ and the result is trivial; however, the result holds for any analytic function cut on

By making the more general transformation

$$z = \frac{\lambda_1 y}{1 + \lambda_2 y} \tag{1.3}$$

we can map any part of the domain of analyticity into the convergence circle.

The major drawback with transformations such as (1.3) is that we do not in general know the location of all the singularities of the function being studied. What we seek to do is to generate a sequence of approximants to the function which are invariant under the (restricted) group of homographic transformations

$$z = \frac{Aw}{1 + Bw} \tag{1.4}$$

We expect that such a sequence will converge at least as well as the best power series obtainable by making transformations such as (1.2), and will hopefully accelerate the rate of convergence with the circle of convergence. More important, such a sequence should automatically produce the analytic continuation to any point not isolated from the origin by singularities of the function. This is in contrast to (1.1), where we have no "automatic" method of choosing n. We shall see (7) that Padé approximants, or at least a subsequence of them, have the required invariance property under the homographic transformations (1.4).

Now suppose that for the logarithmic function ln(1+z) we wish to evaluate f(-2), that is, along the branch cut. Then it is not possible to choose λ_1 and λ_2 in (1.3) such that z=-2 is mapped inside the convergence circle in the y-plane; for example, the transformation (1.2) maps z=-2 to y=2, which lies outside $|y| \leq 1$. The basic reason for this behaviour is that f(z) has a discontinuity at z=-2 and (1.3) represents a continuous transformation; (1.3) cannot map z=-2 into $|y| \leq 1$ where there is no discontinuity in g(y). This illustrates that, using Padé approximants, we can only obtain convergence in regions free of branch cuts. Furthermore, and more important, if we wish to obtain convergence along branch cuts we should use approximants with branch cuts "built-in" to their definition. We shall see (7) that such approximants arise as natural generalizations

of Padé approximants. The potential use of these approximants is enhanced by the fact that they also satisfy the property of homographic invariance (8) as well as the important properties of reciprocal convariance and unitarity (8) satisfied by the Pade approximants; these properties are discussed in section 3.

2. DEFINITION OF THE APPROXIMATION SCHEMES

In this section we define the Padé approximant and the generalizations designed to incorporate a branch cut structure into the approximation scheme.

We assume given a function f(z) with the (formal) Taylor series expansion

$$f(z) = \sum_{\alpha=0}^{\infty} C_\alpha z^\alpha \tag{2.1}$$

DEFINITION 2.1: The (N/M) Pade approximant to f(z), denoted by $f_{(N/M)}(z)$, is the ratio of the polynomials $P_N(z)$ and $Q_M(z)$, of degree at most N and M respectively, which has the same N+M first derivatives of f(z) at z=0:

$$f_{(N/M)}(z) \equiv \frac{P_N(z)}{Q_M(z)} = f(z) + O(z^{N+M+1}) \tag{2.2}$$

or equivalently

$$f_{(N/M)}(z) = \frac{P_N(z)}{Q_M(z)} \tag{2.3a}$$

where

$$f(z)\,Q_M(z) - P_N(z) = O(z^{N+M+1}) \tag{2.3b}$$

We can interpret (2.3b) as defining $f_{(N/M)}(z)$ as the solution of an equation linear in f(z). Shafer (7) has suggested using approximants which are solutions of higher order equations; he gives the following definitions:

DEFINITION 2.2: The (p/q/r) quadratic (Shafer) approximant to f(z), denoted by $f_{(p/q/r)}(z)$, is defined in terms of the polynomials P(z), Q(z) and R(z), of order at most p,q and r respectively, by the equation

$$P(z)\left[f_{(p/q/r)}(z)\right]^2 + Q(z)\,f_{(p/q/r)}(z) + R(z) = 0 \tag{2.4a}$$

where

$$P(z)\,f^2(z) + Q(z)\,f(z) + R(z) = O(z^{p+q+r+2}) \tag{2.4b}$$

DEFINITION 2.3: The (p/q/r/s) cubic approximant to f(z), denoted by $f_{(p/q/r/s)}(z)$, is defined in terms of the polynomials P(z), Q(z), R(z) and S(z), of degree at most p,q,r and s respectively, by the equation

$$P(z)\left[f_{(p/q/r/s)}(z)\right]^3+Q(z)\left[f_{(p/q/r/s)}(z)\right]^2+R(z)f_{(p/q/r/s)}(z)+S(z)=0 \tag{2.5a}$$

where

$$P(z)f^3(z)+Q(z)f^2(z)+R(z)f(z)+S(z)=O\left(z^{p+q+r+s+3}\right) \tag{2.5b}$$

The above two definitions involve $f^2(z)$ and $f^3(z)$; from (2.1) we see that

$$f^2(z)=\sum_{\alpha=0}^{\infty} b_\alpha z^\alpha \tag{2.6a}$$

where

$$b_\alpha=\sum_{\beta=0}^{\alpha} c_\beta c_{\alpha-\beta} \tag{2.6b}$$

together with a similar expressing defining $f^3(z)$. However, it should be remembered that since the power series expansion (2.1) of f(z) may only be formal, (2.6a) may only be a formal series; we have only performed a "formal multiplication" of the power series (2.1) with itself.

It is clear that we can define approximants which are solutions of equations of arbitrary degree in f(z); we give here only the definitions for the two simplest approximants (quadratic and cubic) since it is with these approximants that we shall be mainly concerned. We can see that these approximation schemes do indeed produce approximants with a branch cut structure built in; for example, from (2.4a),

$$f_{(p/q/r)}(z)=\left[-Q(z)\pm\left(Q^2(z)-4P(z)R(z)\right)^{1/2}\right]/2P(z) \tag{2.7}$$

so that quadratic approximants are two-valued functions with square root branch points, in addition to poles and zeros.

We can define other classes of approximants in a natural way. Suppose we have a function g(z) which is known to have a logarithmic branch point; then we can form the (formal) series exp(g(z)) and form Padé approximants, $g_p(z)$, (or quadratic or cubic approximants) to this series. We might then expect ln $g_p(z)$ to provide a good representation of the logarithmic branch point. This idea can be generalized by, instead of the exponential and

logarithmic functions, considering a general invertible operator O. If we denote by P the formation of a suitable approximant, then we can form the general class of approximants

$$O^{-1} P O f(z) \tag{2.8}$$

By choosing the operator O appropriately, we may hope to "build in" to our approximant whichever type of singularity seems suitable for the problem under consideration.

It is also possible, for example, to use approximation schemes of the form

$$P_L(z) \frac{d^2 f(z)}{dz^2} + Q_M(z) \frac{df(z)}{dz} + R_N(z) = O\left(z^{L+M+N+1}\right) \tag{2.9}$$

Approximants of this form have been used, for example, in critical phenomena (8).

We shall not be greatly concerned with some of these more general schemes and shall mainly consider the generalizations of Padé approximants, exemplified by quadratic and cubic approximants. For the class of approximants to which these latter two approximants belong we shall use the general term "algebraic approximant", since they are the solution of algebraic equations. The approximants defined by (2.9) are therefore not algebraic approximants.

3. BASIC PROPERTIES OF PADÉ AND QUADRATIC APPROXIMANTS

We now consider some of the basic properties of the (one-variable) Pade and quadratic approximants; we shall also indicate the extension of these properties to the higher order algebraic approximants.

(I) NORMALISATION CONDITIONS

(2.3b) produces a linear system of (N+M+1) equations to determine the (N+M+2) unknown coefficients of $P_N(z)$ and $Q_M(z)$. If we impose no extra conditions on this system (which leads to the approximants defined by Frobenius (5)) the following situation can occur (9). Consider the (1/1) Padé approximant to

$$f(z) = 1 + z^2 + O(z^3) \tag{3.1}$$

Using (2.3b) we find

$$P_1(z) = Q_1(z) = Az$$

where A is an arbitrary constant. Then,

$$f(z) - \frac{P_1(z)}{Q_1(z)} = O(z^2)$$

instead of $O(z^3)$; this occurs because $Q_1(0) = 0$

This undesireable feature is eliminated by imposing the normalisation condition (9)

$$Q_N(0) = 1 \tag{3.2}$$

for all non-negative integers N. The disadvantage of (3.2) is that not all entries in the Pade table need exist (10) (for example, the (1/1) Pade approximant to (3.1) does not exist according to this definition); using the Frobenius definition (that is, $Q_{11}(z) \neq 0$) all approximants must exist (5). However, for those entries which do exist

$$f(z) - \frac{P_M(z)}{Q_N(z)} = O(z^{M+N+1}) \tag{3.3}$$

We shall always use the condition (3.2).

The above situation is also encountered with quadratic approximants. (2.4b) produces a system of (p+q+r+2) linear equations for the (p+q+r+3) unknown coefficients of P(z), Q(z) and R(z); again we have the choice of imposing no extra conditions on this system or of adopting a normalisation condition of the form (3.2). If we choose the former alternative and consider the (1/1/1) quadratic approximant to

$$f(z) = 1 + z^2 + z^3 + O(z^5) \tag{3.4}$$

we find

$$P(z) = Az \; ; \; Q(z) = -2Az \; ; \; R(z) = Az$$

where A is again arbitrary. Then,

$$f(z) - f_{(1/1/1)}(z) = O(z^2)$$

instead of $O(z^5)$. This situation occurs because either P(0) = 0 or Q(0) = 0; in this particular case both vanish. We can overcome this difficulty by adopting one of the normalisation conditions

$$P_N(0) = 1 \tag{3.5a}$$

or

$$Q_N(0) = 1 \tag{3.5b}$$

for all non-negative integers N. Again, if we adopt (3.5a) or (3.5b) not all the quadratic approximants need exist; for example, the (1/1/1) approximant to (3.4) does not exist for either (3.5a) or (3.5b). However, for those approximants which do exist

$$f(z) - f_{(p/q/r)}(z) = O(z^{p+q+r+2}) \tag{3.6}$$

where $f_{(p/q/r)}(z)$ is defined by (2.7).

The fact that we have a choice of normalisation conditions can be useful. For example, the (1/1/1) quadratic approximant to ln(1-z) does not exist if (3.5a) is used but is well defined with (3.5b); in fact, for this particular example, the even order diagonal approximants only exist if (3.5a) is used and the odd order approximants exist only using (3.5b). We can now see explicitly that Padé approximants are a particular case of quadratic approximants (which justifies calling quadratic approximants generalizations of Padé approximants). For, using (3.5b) we have from (2.4b) with $P(z) \equiv 0$,

$$Q(z)f(z) + R(z) = O(z^{r+q+1})$$

which is the defining equation (apart from a minus sign) for the (r/q) Padé approximant (since $Q(0) = 1$). Also, using (3.5a), we have when $R(z) \equiv 0$ the equation (after cancelling a factor $f(z)$),

$$P(z)f(z) + Q(z) = O(z^{p+q+1})$$

which (again apart from a minus sign) defines the (q/p) Pade approximant.

It is clear that the considerations of this section extend to the higher order algebraic approximants.

(II) DETERMINANT FORMULAE

Provided the defining equations are not singular, we (11) can give an explicit determinantal representation for the (m/n) Padé approximant. With the notation

$$C_k(z) = \sum_{i=0}^{k} c_i z^i \tag{3.7a}$$

$$f(z) \equiv C(z) = \sum_{i=0}^{\infty} c_i z^i \tag{3.7b}$$

and

$$c_n \equiv 0 \quad \text{if } n < 0$$

(3.7c)

the representation is given by the following theorem.

THEOREM 3.1: The (m/n) Pade approximant, $f_{(m/n)}(z)$, has the (not necessarily reduced) representation

$$f_{(m/n)}(z) = \frac{P_m(z)}{Q_n(z)}$$

where

$$P_m(z) = \det \begin{vmatrix} C_m(z) & z\,C_{m-1}(z) & \cdots & z^n C_{m-n}(z) \\ c_{m+1} & c_m & \cdots & c_{m-n+1} \\ \vdots & \vdots & & \vdots \\ c_{m+n} & c_{m+n-1} & \cdots & c_m \end{vmatrix}$$

and

$$Q_n(z) = \det \begin{vmatrix} 1 & z & \cdots & z^n \\ c_{m+1} & c_m & \cdots & c_{m-n+1} \\ \vdots & \vdots & & \vdots \\ c_{m+n} & c_{m+n-1} & \cdots & c_m \end{vmatrix}$$

provided

$$\triangle = \det \begin{vmatrix} c_m & c_{m-1} & \cdots & c_{m-n+1} \\ c_{m+1} & c_m & \cdots & c_{m-n+2} \\ \vdots & \vdots & & \vdots \\ c_{m+n-1} & c_{m+n} & \cdots & c_m \end{vmatrix} \neq 0 \quad (3.8)$$

Moreover,

$$f(z)\,Q_n(z) - P_m(z) = (-1)^n \sum_{k=1}^{\infty} \det \begin{vmatrix} c_{m+1} & c_m & \cdots & c_{m-n+1} \\ \vdots & \vdots & & \vdots \\ c_{m+n} & c_{m+n-1} & \cdots & c_m \\ c_{m+n+k} & c_{m+n+k-1} & \cdots & c_{m+k} \end{vmatrix} z^{m+n+k}$$

The following theorem gives a similar characterization for quadratic approximants. Using the notation of (3.7) and letting

$$B_k(z) = \sum_{i=0}^{k} b_i z^i$$

where

$$f^2(z) \equiv B(z) = \sum_{i=0}^{\infty} b_i z^i$$

and

$$b_n \equiv 0 \quad \text{if } n<0$$

we have the following representation.

THEOREM 3.2: The (p/q/r) quadratic approximant, $f_{(p/q/r)}(z)$, has the (not necessarily reduced) representation

$$f_{(p/q/r)}(z) = \frac{-Q_q(z) \pm \sqrt{Q_q^2(z) - 4P_p(z)R_r(z)}}{2P_p(z)}$$

where

$$P_p(z) = \det \begin{vmatrix} 1 & \cdots & z^p & 0 & \cdots & 0 \\ b_{r+1} & \cdots & b_{r-p+1} & c_{r+1} & \cdots & c_{r-q+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & \cdots & c_{p+r+1} \end{vmatrix} \qquad (3.9a)$$

$$Q_q(z) = \det \begin{vmatrix} 0 & \cdots & 0 & 1 & \cdots & z^q \\ b_{r+1} & \cdots & b_{r-p+1} & c_{r+1} & \cdots & c_{r-q+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & \cdots & c_{p+r+1} \end{vmatrix} \qquad (3.9b)$$

and

$$R_r(z) = -\det \begin{vmatrix} B_r(z) & \cdots & z^r B_{r-p}(z) & C_r(z) & \cdots & z^r C_{r-q}(z) \\ b_{r+1} & \cdots & b_{r-p+1} & c_{r+1} & \cdots & c_{r-q+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & \cdots & c_{p+r+1} \end{vmatrix} \qquad (3.9c)$$

provided

$$\triangle_p = \det \begin{vmatrix} b_r & b_{r-1} & \cdots & b_{r-p+1} & c_{r+1} & c_r & \cdots & c_{r-q+1} \\ b_{r+1} & b_r & \cdots & b_{r-p+2} & c_{r+2} & c_{r+1} & \cdots & c_{r-q+2} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ b_{p+q+r} & b_{p+q+r-1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & c_{p+q+r} & \cdots & c_{p+r+1} \end{vmatrix} \qquad (3.9d)$$

and

$$\Delta_q = \det \begin{vmatrix} b_{r+1} & b_r & \cdots & b_{r-p+1} & c_r & c_{r-1} & \cdots & c_{r-q+1} \\ b_{r+2} & b_{r+1} & \cdots & b_{r-p+2} & c_{r+1} & c_r & \cdots & c_{r-q+2} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & b_{p+q+r} & \cdots & b_{q+r+1} & c_{p+q+r} & c_{p+q+r-1} & \cdots & c_{p+r+1} \end{vmatrix} \tag{3.9e}$$

are not both zero.

Furthermore,

$$B(z)P_p(z) + C(z)Q_q(z) + R_r(z)$$

$$= (-1)^{p+q+1} \sum_{k=1}^{\infty} \det \begin{vmatrix} b_{r+1} & \cdots & b_{r-p+1} & c_{r+1} & \cdots & c_{r-q+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & \cdots & c_{p+r+1} \\ b_{p+q+r+k+1} & \cdots & b_{q+r+k+1} & c_{p+q+r+k+1} & \cdots & c_{p+r+k+1} \end{vmatrix} \tag{3.9f}$$

PROOF: We show that (3.9a), (3.9b) and (3.9c) imply (3.9f). From (3.9a),

$$B(z)P_p(z) = \det \begin{vmatrix} B(z) & \cdots & z^p B(z) & 0 & \cdots & 0 \\ b_{r+1} & \cdots & b_{r-p+1} & c_{r+1} & \cdots & c_{r-q+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & \cdots & c_{p+r+1} \end{vmatrix} \tag{3.10a}$$

and from (3.9b)

$$C(z)Q_q(z) = \det \begin{vmatrix} 0 & \cdots & 0 & C(z) & \cdots & z^q C(z) \\ b_{r+1} & \cdots & b_{r-p+1} & c_{r+1} & \cdots & c_{r-q+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & \cdots & c_{p+r+1} \end{vmatrix} \tag{3.10b}$$

For i=1,2,....,(p+q) multiply the (i+1)st.row of (3.9c) by z^{r+i} and add to the 1st. row. This gives

$$R_r(z) = -\det \begin{vmatrix} B_{p+q+r+1}(z) & \cdots & z^p B_{r+q}(z) & C_{p+q+r+1}(z) & \cdots & z^q C_{p+r}(z) \\ b_{r+1} & \cdots & b_{r-p+1} & c_{r+1} & \cdots & c_{r-q+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & \cdots & c_{p+r+1} \end{vmatrix} \tag{3.10c}$$

Then, from (3.10a), (3.10b) and (3.10c), we obtain

$$B(z)P_p(z) + C(z)Q_q(z) + R_r(z)$$

$$= \det \begin{vmatrix} \sum_{k=p+q+r+2}^{\infty} b_k z^k & \cdots & \sum_{k=q+r+1}^{\infty} b_k z^{k+p} & \sum_{k=p+q+r+2}^{\infty} c_k z^k & \cdots & \sum_{k=p+r+1}^{\infty} c_k z^{k+q} \\ b_{r+1} & \cdots & b_{r-p+1} & c_{r+1} & \cdots & c_{r-q+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & \cdots & c_{p+r+1} \end{vmatrix}$$

$$= (-1)^{p+q+1} \det \begin{vmatrix} b_{r+1} & \cdots & b_{r-p+1} & c_{r+1} & \cdots & c_{r-q+1} \\ b_{r+2} & \cdots & b_{r-p+2} & c_{r+2} & \cdots & c_{r-q+2} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+1} & c_{p+q+r+1} & \cdots & c_{p+r+1} \\ \sum_{k=1}^{\infty} b_\sigma z^\sigma & & \sum_{k=1}^{\infty} b_{\sigma-p} z^\sigma & \sum_{k=1}^{\infty} c_\sigma z^\sigma & & \sum_{k=1}^{\infty} c_{\sigma-q} z^\sigma \end{vmatrix} \qquad (3.11)$$

where

$$\sigma = p+q+r+k+1$$

Clearly (3.11) is equivalent to (3.9f).

The non-vanishing of Δ_p and Δ_q, defined by (3.9d) and (3.9e), expresses the non-vanishing of $P_p(0)$ and $Q_q(0)$ respectively; this is clear from (3.9a) and (3.9b). Thus, if both Δ_p and Δ_q are zero, we cannot impose either of the normalization conditions (3.5a) and (3.5b) and so, according to our definition, the (p/q/r) quadratic approximant does not exist. If $\Delta_p = 0$ and $\Delta_q \neq 0$ then we must choose (3.5b) as the normalization condition and if $\Delta_q = 0$ with $\Delta_p \neq 0$ then (3.5a) must be chosen.

This completes the proof of the theorem.

An important point to note is that the proof of Theorem 3.2 does not make use of the relation between B(z) and C(z):

$$C(z) = f(z) \quad , \quad B(z) = f^2(z) \qquad (3.12)$$

The theorem will therefore apply to any class of approximants satisfying

$$P(z)B(z) + Q(z)C(z) + R(z) = O\left(\frac{1}{z}^{p+q+r+2}\right)$$

and not just for the choice of B(z) and C(z) given by (3.12). For example, we may choose

$$C(z) = f(z) \quad , \quad B(z) = \exp(f(z))$$

which, in our terminology, would define a class of "transcendental" approximants.

Theorem 3.2 extends in a straightforward way to the higher order approximants. For example, the polynomials defining cubic approximants will be represented by determinants containing coefficients of f, f^2 and f^3; explicitly, with the notation of Definition 2.3 and with

$$f^3(z) = \sum_{k=0}^{\infty} d_k z^k$$

we obtain, for example,

$$\begin{vmatrix} 1 & \cdots & z^p & 0 & \cdots & 0 & 0 & \cdots & 0 \\ b_{s+1} & \cdots & b_{s-p+1} & c_{s+1} & \cdots & c_{s-q+1} & d_{s+1} & \cdots & d_{s-r+1} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ b_{p+q+r+s+2} & \cdots & b_{q+r+s+2} & c_{p+q+r+s+2} & \cdots & c_{p+r+s+2} & d_{p+q+r+s+2} & \cdots & d_{p+q+s+2} \end{vmatrix}$$

(III) RECIPROCAL CONVARIANCE

Padé approximants satisfy the property of reciprocal convariance which ensures that if $f_{(m/n)}$ is the (m/n) Padé approximant to f(z) then, provided $f(0) \neq 0$, $1/f_{(m/n)}$ is the (n/m) Padé approximant to the reciprocal series 1/f(z) (12).

For quadratic approximants the corresponding convariance property states that if $f_{(p/q/r)}$ is the (p/q/r) quadratic approximant to f(z) then, provided $f(0) \neq 0$, $1/f_{(p/q/r)}$ (which we denote by $(f_{(p/q/r)})^{-1}$) is the (r/q/p) quadratic approximant to 1/f(z) (which we denote by $f^{-1}(z)$). To see this we have, by definition (and using the notation of Section 2),

$$f_{(p/q/r)}(z) = \frac{-Q \pm (Q^2 - 4PR)^{1/2}}{2P} \tag{3.13a}$$

where

$$Pf^2 + Qf + R = O(z^{p+q+r+2}) \tag{3.13b}$$

From the series (2.1) for f(z) we form the form the formal series for $f^{-1}(z)$, defined by

$$f^{-1}(z) = \sum_{\alpha=0}^{\infty} d_\alpha z^\alpha \tag{3.14}$$

where

$$d_0 = c_0^{-1}$$

and

$$\sum_{\beta=0}^{\delta} c_{\delta-\beta}\, d_\beta = 0 \qquad (0<\delta<\infty)$$

define the coefficients d_α; since, by assumption, $f(0) = c_0 \neq 0, d_0$ is well defined. We now multiply (3.13b) formally by $f^{-2}(z) = (f^{-1}(z))^2$ (obtained from (3.14) by formal multiplication) to give

$$R\left(f^{-1}(z)\right)^2 + Q f^{-1}(z) + P = O\left(z^{p+q+r+2}\right) \tag{3.15}$$

Thus, denoting by $f^{-1}_{(\alpha/\beta/\lambda)}$ the $(\alpha/\beta/\lambda)$ quadratic approximant formed from the reciprocal series $f^{-1}(z)$, we have from (3.15)

$$f^{-1}_{(r/q/p)} = \frac{-Q \mp (Q^2-4PR)^{1/2}}{2R} \tag{3.16}$$

Finally, from (3.13a) and (3.16), we have

$$\begin{aligned}\left[f_{(p/q/r)}(z)\right]^{-1} &= 2P \Big/ \left[-Q \pm (Q^2-4PR)^{1/2}\right] \\ &= \left[-Q \mp (Q^2-4PR)^{1/2}\right] \Big/ 2R \\ &= f^{-1}_{(r/q/p)}(z)\end{aligned}$$

which proves the result.

For higher order approximants the corresponding result is clear; for example, if $f_{(p/q/r/s)}$ is the (p/q/r/s) cubic approximant to f(z) then $(f_{(p/q/r/s)})^{-1}$ is the (s/r/q/p) cubic approximant to $f^{-1}(z)$.

Since the above approximants are many-valued we must interpret the reciprocal covariance property as being true provided we remain on a given Riemann sheet. Thus, for quadratic approximants, we choose a particular sheet of $f_{(p/q/r)}$ (determined by the requirement that $f_{(p/q/r)}$ and f(z) 'agree' on this sheet) and a particular sheet of $f^{-1}_{(r/q/p)}$ (chosen so that $f_{(r/q/p)}$ and $f^{-1}(z)$ 'agree' on this sheet); then, provided we identify $f_{(p/q/r)}$ and $f^{-1}_{(r/q/p)}$ with their appropriate values on these chosen sheets, the reciprocal covariance property is valid.

(IV) HOMOGRAPHIC COVARIANCE

For all $t>0$, the diagonal "t-power approximants" of Section 2 (where t=1,2 and 3 correspond to Padé, quadratic and cubic approximants respectively) are invariant under the homographic transformation

$$z = \frac{Aw}{1+Bw} \tag{3.17}$$

We shall prove the result for t=2 (8); the generalization to arbitrary t is immediate.

Consider the (N/N/N) quadratic approximant to f(z) and let

$$g(w) = f\left(\frac{Aw}{1+Bw}\right)$$

and denote by $A_N(f(z))$ the (N/N/N) quadratic approximant to f(z). Homographic invariance requires that

$$A_N(g(w)) = A_N(f(z))\Big|_{z=\frac{Aw}{1+Bw}} \tag{3.18}$$

Now A (f(z)) is defined by the equation

$$\left(\sum_{i=0}^{N} p_i z^i\right) f^2(z) + \left(\sum_{j=0}^{N} q_j z^j\right) f(z) + \sum_{k=0}^{N} r_k z^k = O(z^{3N+2}) \tag{3.19}$$

where p_i, q_i and r_i $(0 \leq i \leq N)$ denote the coefficients of P(z), Q(z) and R(z) respectively. Substituting for z from (3.17) into (3.19) and multiplying through formally by $(1+Bw)^N$ gives

$$(1+Bw)^N P\left(\frac{Aw}{1+Bw}\right) g^2(w) + (1+Bw)^N Q\left(\frac{Aw}{1+Bw}\right) g(w) + (1+Bw)^N R\left(\frac{Aw}{1+Bw}\right) = O(w^{3N+2})$$

which establishes (3.18) (since the coefficients of $g^2(w)$, $g(w)$ and 1 are polynomials of degree N).

It is precisely this homographic covariance property which we expect will make the "t-power approximants" of practical use, since this ensures the invariance of the approximant under the Euler transformations (which, as we have seen in Section I, provide a means of analytic continuation).

(V) UNITARITY

The application of the Padé approximant method in Quantum Field Theory is motivated by the fact that the diagonal Pade approximants to the S-matrix are unitary (13). In (8) the following result is proved for quadratic approximants: if f(z) is a unitary function (so that $f(z)f^*(z)=1$ for z real, where * denotes complex conjugation) then the (N/M/N) quadratic approximant to f(z) is unitary. The cited proof is very unclear and seems to omit (or at least to explain) several important steps; it is hoped that the following argument clarifies the proof.

Let $f_{(p/q/r)}$ dencte the (p/q/r) quadratic approximant to f(z); we

seek to prove that

$$f_{(p/q/r)}\left[f_{(p/q/r)}\right]^{*} = 1 \tag{3.20}$$

when p=r. By definition

$$f_{(p/q/r)} = \frac{-Q \pm (Q^2-4PR)^{1/2}}{2P} \tag{3.21a}$$

where

$$Pf^2 + Qf + R = O(z^{p+q+r+2}) \tag{3.21b}$$

Multiplying (3.21b) formally by $(f^*)^2$ and using the identity

$$f(z)f^*(z) = 1$$

we obtain

$$R(f^*)^2 + Qf^* + P = O(z^{p+q+r+2}) \tag{3.22a}$$

as the equation defining $f^*_{(r/q/p)}$. But we can also define $f^*_{(p/q/r)}$ by the equation

$$P^*(f^*)^2 + Q^* f^* + R^* = O(z^{p+q+r+2}) \tag{3.22b}$$

obtained by conjugation of (3.21b). If the quadratic approximant to $f^*(z)$ is to be uniquely determined, (3.22a) and (3.22b) imply that

$$P^* = R \quad \text{and} \quad Q^* = Q \tag{3.23}$$

Finally, from (3.21a),

$$f^*_{(p/q/r)} = \frac{-Q^* \mp (Q^{*2}-4P^*R^*)^{1/2}}{2P^*} \tag{3.24}$$

In (3.24) we have written $\mp$ as opposed to $\pm$ to indicate that, whatever sign is chosen in (3.21a), the opposite sign must be chosen in (3.24). This is necessitated by the requirement that both $f(z)$ and $f^*(z)$, and hence $f_{(p/q/r)}(z)$ and $f^*_{(p/q/r)}(z)$, tend to unity as $z \to 0$ (since $f(z)$ is unitary). The required result (3.20) now follows from (3.21a), (3.24) and (3.23).

Having discussed some of the more important properties common to Padé and quadratic approximants, we now illustrate some of the differences between the two approximation schemes.

(VI) THE C-TABLE

The determinant Δ defined by (3.8) is more usually written in the form

$$C(m/n) = \det \begin{vmatrix} c_{m-n+1} & c_{m-n+2} & \cdots & c_m \\ \vdots & \vdots & & \vdots \\ c_m & c_{m+1} & \cdots & c_{m+n-1} \end{vmatrix} \tag{3.25}$$

so that, from Theorem 3.1, we can see that $C(L/M)\neq 0$ is a sufficient condition for the existence of the (L/M) Padé approximant. The table built up from the determinants of (3.25) is called the 'C-table'; this table is useful in that it enables one to examine the general structure of the Pade table itself (14). The main feature of the C-table is that any zero entries occur in square blocks entirely surrounded by non-zero entries; this property enables one to prove the "block theorem" of Pade (14):

THEOREM 3.3 (PADÉ): The Padé table can be completely dissected into rxr blocks with horizontal and vertical sides, $r\geq 1$. If (λ/μ) denotes the unique minimal ($\lambda+\mu$ = minimum) member of a particular rxr block, then:

(i) The (λ/μ) exists and the numerator and denominator are of full nominal degree.

(ii) $(\lambda+p/\mu+q) = (\lambda/\mu)$ for $p+q\leq r-1, p\geq 0, q\geq 0$

(iii) $(\lambda+p/\mu+q)$ do not exist for $p+q\geq r, r-1\geq p\geq 1, r-1\geq q\geq 1$

(iv) The equations for the $(\lambda+p/\mu)$, $0\leq p\leq r-1$, and $(\lambda/\mu+q)$, $0\leq q\leq r-1$, are nonsingular, and those for the other block members are singular.

(v) $C(\lambda+p/\mu+q)=0$ for $1\leq p\leq r-1, 1\leq q\leq r-1$; and $C\neq 0$ otherwise.

The "block theorem" thus shows which entries in the Padé table are non-singular, which entries are equal and which entries have singular consistent/inconsistent equations associated with them.

We can define a similar table for quadratic approximants; in fact, since there are two possible normalisation conditions associated with quadratic approximants, we can define two tables. If we set

$$\Delta_p = C_Q(p/q/r) \tag{3.26}$$

where Δ_p is defined by (3.9d), then we can generate a "C_Q-table" (the subscript Q denoting that we are dealing with quadratic approximants). We might hope that the "C_Q-table" would have the property that any zero entry in the table occurs in a cubic block entirely surrounded by non-zero entries. Unfortunately this is not the case; if we consider Gragg's example (11)

$$R(z) = \frac{1-z+z^3}{1-2z+z^2}$$

then we find no discernable pattern amongst the zero entries of the "C_Q-table"

(except when r=0 when the zero entries occur in square blocks entirely surrounded by non-zero entries). This seems to indicate that the table of quadratic approximants will probably not have the block structure of the Padé table.

(VII) RECURSION RELATIONS

One of the basic starting points for deriving recursion relations between elements of the Pade table is the result (14)

$$C(L/M+1)\,C(L/M-1) = C(L+1/M)\,C(L-1/M) - [C(L/M)]^2 \qquad (3.27)$$

which follows from Sylvester's determinant identity (14)

$$\det|A|\,\det|A_{rs;pq}| = \det|A_{rp}|\,\det|A_{sq}| - \det|A_{rq}|\,\det|A_{sp}| \qquad (3.28)$$

where the subscripts denote the rows and columns to be deleted from the matrix A.

If we apply (3.28) to (3.26) we run into difficulties - essentially because the C_Q-determinants contain both b and c coefficients (the coefficients of f^2 and f) and the row and column deletions prescribed by (3.28) affect these sets of coefficients in an unsymmetrical way. As yet, it appears that no simple relation of the form (3.27) exists for the "C_Q-table".

(VIII) ALGORITHMS

For Padé approximants the existence of simple relations of the form (3.27) leads to the possibility of generating the approximants recursively. The most widely used algorithm in this context is probably the ϵ -algorithm ((15) and (16)) but many more algorithms are known ((11) and (17)). For quadratic approximants no such algorithms are yet known. Computationally, this presents no difficulties; indeed it is sometimes preferable to solve the system of linear equations directly and the algorithms used for Padé approximants can converge slowly for high order approximants.

The method employed here for calculating the quadratic approximants is to solve the system of linear equations (as opposed to using the determinant formalism)

$$A\,\underline{x} = \underline{b}$$

where

$$\underline{x}^T = [p_1, \ldots, p_p; q_1, \ldots, q_{q+1}; r_1, \ldots, r_{r+1}]$$

$$\underline{b}^T = [b_1, b_2, \ldots, b_{p+q+r+2}]$$

$$\text{A} = \left[\begin{array}{c|c|c} \begin{matrix} 0 & 0 & \\ b_1 & 0 & 0 \\ b_2 & b_1 & \\ & & \\ b_p & \cdots & b_1 \end{matrix} & \begin{matrix} c_1 & & \\ c_2 & c_1 & 0 \\ & & \\ & & \\ c_{q+1} & \cdots & c_1 \end{matrix} & \begin{matrix} 1 & & \\ 1 & & 0 \\ 0 & & \\ & & \\ & & 1 \end{matrix} \\ \hline \begin{matrix} b_{p+1} & \cdots & b_2 \\ \vdots & & \vdots \\ b_{p+q+r+1} & \cdots & b_{q+r+2} \end{matrix} & \begin{matrix} c_{q+2} & \cdots & c_2 \\ \vdots & & \vdots \\ c_{p+q+r+2} & \cdots & c_{p+r+2} \end{matrix} & \begin{matrix} \\ \\ 0 \end{matrix} \end{array}\right] \qquad (3.29)$$

(for definiteness we have assumed $p_o=1$).

Alternatively we can solve for the coefficients of P and Q and then determine the coefficients of R. Whichever method we adopt, the practical calculation of quadratic approximants is straightforward (since, as with Padé approximants, it is a linear system of equations that we have to solve) provided the order of the approximant is not "too large". This latter requirement is a common feature of all Padé type schemes and is attributable to loss of accuracy resulting from inverting the matrix A of (3.29).

Although lack of recursion relations for quadratic approximants presents no problems from a practical point of view, theoretically it is a great problem. The proofs of most of the theorems presently known for Padé approximants depend at some stage on the existence of relations of the form (3.27). It is hoped that the results presented in this thesis will encourage the search for such recursion relations and enable the formulation of convergence theorems for quadratic approximants.

REFERENCES

1. E.T. Copson, "An Introduction to the Theory of Functions of a Complex Variable", Oxford, Clarendon Press (1972), p 84-88.
2. G.H. Hardy, "Divergent Series", Oxford (1949).
3. E. Borel, "Leçons sur les Séries Divergentes", Gauthier-Villars, Paris (1928).
4. C.G.J. Jacobi, "Uber die Darstellung einer Reihe Gegebner Werthe durch eine Gerbrochne Rationale Function", J. Reine Angew Math.(Crelle)30, 127-156 (1846).
5. G. Frobenius, "Ueber Relationen Zwischen den Naherungsbruchen von Potenzreihen", J. für Math. (Crelle) 90, 1-17(1881).
6. H. Padé, (Thesis) "Sur la représentation approchee d'une fonction pour des fractions rationnelles", Ann.Sci.École Norm.Sup.Suppl(3),9, 1-93 (1892).
7. R.E. Shafer, Siam Jour. Num.Anal. 11, 447 (1974).
8. J.L.Gammel, "Review of Two Recent Generalizations of the Padé Approximant" in "Padé Approximants and their Applications" (ed. P.R. Graves-Morris) p.7 (1973).
9. G.A. Baker Jr., "Essentials of Padé Approximants", Academic Press (1975), p.6 - 8.
10. G.A. Baker Jr., J. Math.Anal. Appl., 43,498-528 (1973).
11. W.B. Gragg, Siam Review 14,1-62 (1972).
12. Reference 9,p 112.
13 J.L. Gammel and F.A. McDonald, Phys.Rev.142,
14. Reference 9, Chapter 2.
15. A.C. Genz, "Applications of the ϵ -algorithm to quadrature problems" in"Padé Approximants and their Applications" (ed.P.R. Graves-Morris), p.105-116.
16. A.C. Genz, "The ϵ -algorithm and some other applications of Pade approximants in numerical analysis", in "Padé Approximants" (ed. P.R. Graves-Morris), p112-125 (1973).
17. Reference 9, Chapter 6.

CHAPTER 2: CONVERGENCE THEOREMS AND NUMERICAL EXAMPLES.

1. CONVERGENCE THEORY FOR PADÉ APPROXIMANTS

A review of the known convergence theorems for Padé approximants is given in (1) and proofs of many of these results can be found in (2). We shall content ourselves with only the more important results in an attempt to indicate the type of convergence theorems we may reasonably expect to hold for quadratic (and higher order) approximants.

One of the earliest results (3) is

THEOREM 1.1 (de Montessus): Let $f(z)$ be regular inside the circle $|z|<R$, except for (possibly multiple) poles $z_1, z_2, \ldots, z_N$ inside the circle. Then

$$\lim_{m\to\infty} f_{(m/N)} \longrightarrow f(z)$$

uniformly on any compact subset of

$$\left\{z : |z|<R,\ z \neq z_i \ (i=1,\ldots,N)\right\}$$

Extensions of Theorem 1.1 have been given by Wilson (4), Saff (5) and Gragg (6).

When $f(z)$ is entire, Beardon (7) and Baker (8) have investigated the properties of rows of the Padé table.

THEOREM 1.2 (Beardon): If $f(z)$ is analytic in $(z)^{1/\rho}$, then there exists an infinite subsequence of (L/1) Padé approximants which converge uniformly in any disc $|z| \leq R$, for $R\rho>1$, to the function defined by the power series.

Further details can be found in (2).

The above results are only concerned with convergence of rows or columns of the Padé table. We now discuss the important diagonal and paradiagonal sequences of approximants (their importance being related to their invariance under homographic transformations). As yet, there is no theorem of the form

"The diagonal sequence of approximants to a function $f(z)$ converges to $f(z)$, with z in a domain D, if and only if..."

and Gammel (8) and Wallin (9) have produced counter examples to any straightforward general theorem. Essentially their examples show that an entire function can be constructed for which many diagonal Padé approximants

have poles in arbitrarily prescribed places. However, Baker (8) has shown that for these examples there exists an infinite subsequence (m/N) of Padé approximants ($N\to\infty$) which converge uniformly in any closed, bounded region of the complex plane to the entire function defined by the power series. These (and other) considerations have led to the formulation of the following conjecture (10); there are no known counter-examples nor any valid proof of it.

PADÉ CONJECTURE: If $f(z)$ is a power series representing a function which is regular for $|z|\leq 1$, except for m poles within this circle and except for $z=+1$, at which the function is assumed continuous as $z\to 1$ from the region $|z|\leq 1$, then at least a subsequence of the (N/N) Padé approximants converge uniformly to the function (as N tends to infinity) in the domain formed by removing the interiors of small circles with centres at these poles.

Certain rigorous convergence theorems can be obtained if we impose conditions on the function $f(z)$, or equivalently on the power series coefficients $\{c_k\}$. For series of Stieltjes (11) it is possible to completely characterize the location of the poles and zeros of the Padé approximants, to prove certain monotonicity properties and finally to establish convergence properties. Below we give a brief summary of these results, further details being obtainable in (2) and (12).

We define

$$f(z) = \sum_{j=0}^{\infty} f_j(-z)^j$$

to be a series of Stieltjes if and only if there is a bounded, non-decreasing function $\phi(u)$, taking on infinitely many values in the interval $0\leq u<\infty$ such that

$$f_j = \int_0^{\infty} u^j d\phi(u)$$

Then the following theorems hold:

THEOREM 1.3: If $\sum_{j=0}^{\infty} f_j(-z)^j$ is a series of Stieltjes, then the poles of the (N/N+j), $j\geq -1$, Padé approximants lie on the negative real axis. Furthermore, the poles of successive approximants interlace, and all the residues are positive.

The roots of the numerator also interlace those of the denominator.

THEOREM 1.4: Any sequence of (N/N+j) Pade approximants to a series of Stieltjes converges to an analytic function in the cut complex plane $(-\infty \le z \le 0)$. If, in addition,

$$\sum_{p=1}^{\infty} (f_p)^{-\frac{1}{(2p+1)}}$$

diverges, then all the sequences tend to a common limit. If the f_p are a convergent series with a radius of convergence R, then any (N/N+j) sequence converges in the cut plane $(-\infty \le z \le -R)$ to the analytic function defined by the power series.

The convergence of general sequences $f_{(M/N)}$ of Padé approximants is a much more difficult problem; the theorems so far proved in this context only establish weak types of convergence, that is convergence in measure and capacity. We quote the following two theorems (see (13) and (14) respectively) in this connection:

THEOREM 1.5 (Nuttall): Let $P_N(z)$ be the (N/N+j) Padé approximant to a meromorphic function F(z), and D be a closed, bounded region of the complex plane. Then, given any $\epsilon, \delta > 0$ there exists N_o such that, for all $N > N_o$,

$$\left| P_N(z) - F(z) \right| < \epsilon$$

for all $z \in \bar{D}_N$ where $\bar{D}_N \subset D$ and the measure of $D - \bar{D}_N$ is less than δ.

THEOREM 1.6 (Pommerenke): Let $E \subset \mathbb{C}$ be a compact set with cap E=0, and let f(z) be (single-valued and) analytic in the complement G of E. Then, for $\epsilon > 0, \eta > 0, r > 1, \lambda > 1$ there exists m_o such that

$$\left| P_{mn}(f(z)) - f(z) \right| < \epsilon^m \quad \left(m > m_o, \frac{1}{\lambda} \le \frac{n}{m} \le \lambda \right)$$

for $|z| \le r, z \notin E_{mn}$ where cap $E_{mn} < \eta$ (and $P_{mn}(f(z))$ is the (m/n) Padé approximant to f(z)).

The major limitation of the previous two theorems is that convergence in measure or in capacity does not exclude non-convergence at a countably infinite number of points; however, the study of a large number of examples has indicated that, in practice, points of non-convergence appear to be very limited. It appears that stronger forms of convergence can only be obtained

by imposing conditions on the power series coefficients.

2. SINGULARITY STRUCTURE OF PADÉ AND QUADRATIC APPROXIMANTS

The distribution of the poles and zeros of Padé approximants play an important role in determining the region of convergence of the approximants. From the study of a large number of examples, the following general picture has emerged (2). Given a particular sequence of Padé approximants the great majority of the poles and zeros, beyond those required to represent the poles and zeros of the function being approximated, tend to cluster along curves which cut the complex plane in such a way as to leave the function single valued in that part of the plane connected to the origin. For horizontal (or vertical) sequences of approximants the boundary curve is usually a circle centred at the origin; for diagonal sequences, the locations of the curves where the extraneous poles and zeros cluster is determined by the homographic invariance properties of the approximants. In the case of a function with a natural barrier, such as a branch cut, the poles and zeros of the diagonal approximants are clustered along the natural barrier.

We can illustrate these remarks, and indicate their possible extension to quadratic approximants, by considering the specific function

$$f(z) = \frac{\ln(1+z)}{z}$$

which is the function considered by Baker ((2),p.223). If we consider the (4M/M) Padé approximant then, from general considerations (see (2) p.223), we expect convergence of the approximants from $z < 4.5$; confirmation of this is provided by Table Ia. In Fig.Ia the extra zeros of the approximants, not associated with poles lying along the cut, are plotted; we can see that these zeros do appear to mark out the boundary of the region of convergence. To see if these considerations extend to quadratic approximants we have repeated the above procedure (with the results given in Table Ib and Fig.Ib) for the (M/O/4M) quadratic approximants. Again the region of convergence of the (quadratic) approximants does seem to be determined by the corresponding zeros of the approximant. This example also indicates (as is evident from Tables Ia and Ib) that it is only with the diagonal (and possibly paradiagonal) approximants that we can expect to obtain convergence in the whole complex

Fig. Ia: THE ZEROS (NOT ON THE NEGATIVE REAL AXIS) OF THE (4M/M) PADE APPROXIMANTS TO $z^{-1}\ln(1+z)$

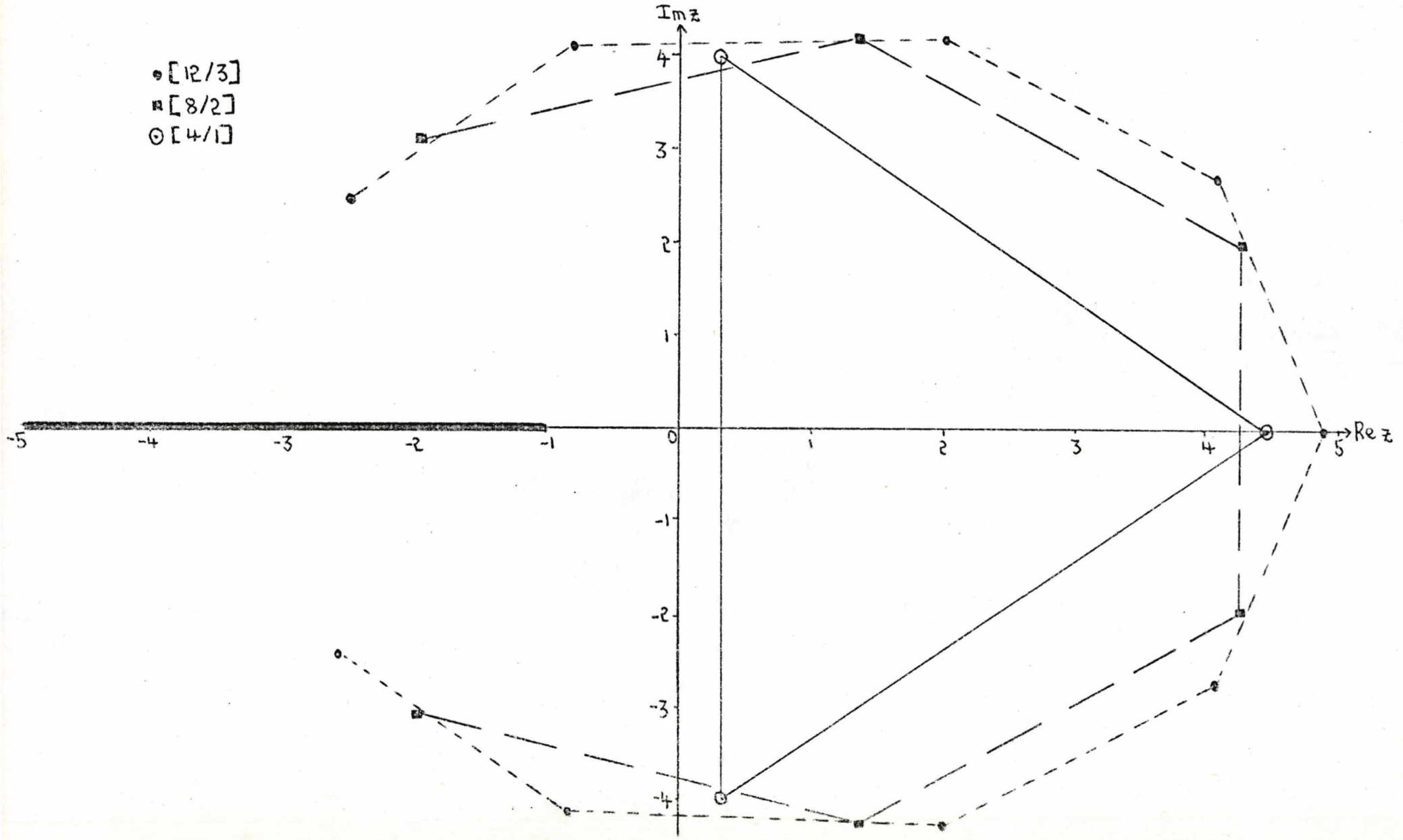


Fig. Ib: THE ZEROS (NOT ON THE NEGATIVE REAL AXIS OF THE $(M/0/4M)$ QUADRATIC APPROXIMANTS TO $z^{-1}\ln(1+z)$

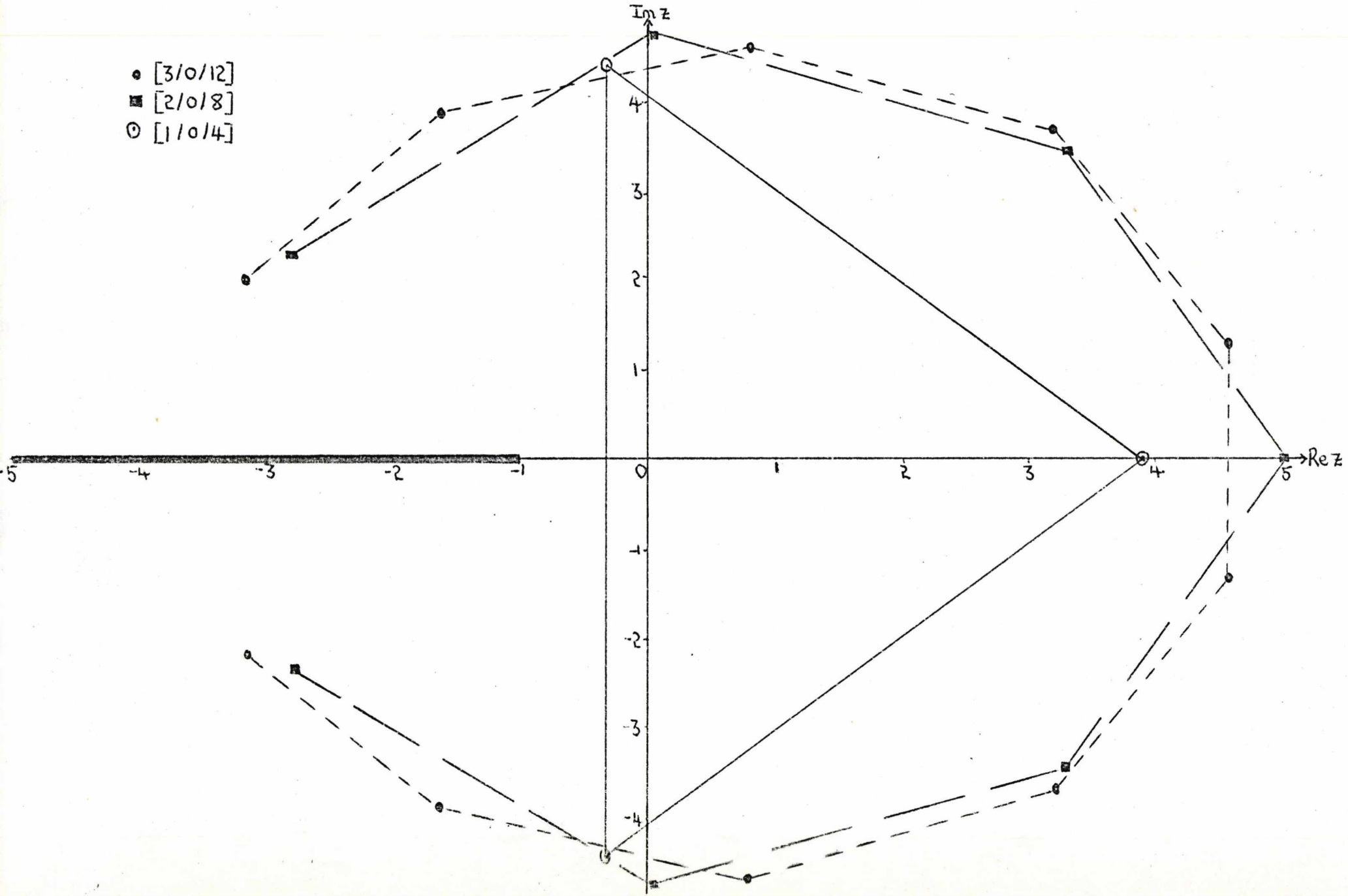


TABLE Ia: (4M/M) PADE APPROXIMANTS TO $z^{-1}\ln(1+z)$

z \ M	1	2	3	4	$z^{-1}\ln(1+z)$
1	0.6924	0.693149	0.69314718	0.69314718	0.69314718
6	-0.8	2.15	-2.6	5.1	0.32
9	-4.4	28.9	-177	1092	0.26

TABLE Ib: (M/O/4M) QUADRATIC APPROXIMANTS TO $z^{-1}\ln(1+z)$

z \ M	1	2	3	$z^{-1}\ln(1+z)$
-4	$0.08\pm1.3i$	$0.027\pm1.44i$	$-0.008\pm1.8i$	$-0.28\pm0.8i$
-1	13	8	40	00
1	0.49	0.68	0.67	0,693
6	0.44	$0.01\pm0.6i$	$\pm0.98i$	0.32
9	1.13	$0.003\pm2.7i$	$\pm7.3i$	0.26

plane (cut along $-\infty \leq z \leq -1$ for the Padé approximants); clearly, the above non-diagonal approximants do not even converge in measure outside the heart-shaped regions of Figs.Ia and Ib.

For certain functions with branch points Nuttall has recently obtained convergence theorems for the corresponding Padé approximants. In Chapter 3 we shall describe a method of evaluating Feynman integrals based on the idea of rotation of branch cuts; in this type of problem we are interested in the location of the poles of the resulting Padé approximants and Nuttall's results are of use in this connection. In fact, although the following two theorems have only been proved for a certain class of functions (see below), in Chapter 3 we will assume that the concepts contained in the theorems are true much more generally.

Nuttall's first result (15) is concerned with a function F(Z) with two branch points:

$$F(Z) = \int_a^b \frac{w(x)\,dx}{1 - Zx} \tag{2.1}$$

where a,b,w(z) are complex. Provided the weight function w(x) satisfies certain conditions, the following theorem holds:

THEOREM 2.1: The (N-1/N) Padé approximant to F(Z) (defined by (2.1)) converges uniformly, as $N \to \infty$, in any closed, bounded region of the Z-plane cut along the arc

$$Z = x^{-1} \quad , \quad x = \tfrac{1}{2}\left[(b-a)t + (a+b)\right] \quad , \quad -1 \leq t \leq 1$$

The second of Nuttall's results (16) deals with a function with an even number of branch points with principal singularities of square root type. The function, f(t), considered is defined by

$$f(t) = \int_S dt'\, X_+^{-\frac{1}{2}}(t')\,\sigma(t')\,(t'-t)^{-1} \tag{2.2}$$

where

(i) if $d_i \; (i = 1, 2, \ldots, 2L)$ denote the branch points of f(t)

$$X(t) = \prod_{i=1}^{2L} (t - d_i)$$

and $X_+^{-\frac{1}{2}}$ denotes the limit from a particular side of S and

(ii) the set S consists of L analytic Jordan arcs joining pairs of branch points; in general, the L arcs are non-intersecting.

Again, if the weight function $\sigma(t)$ satisfies certain conditions, the following theorem holds:

THEOREM 2.2: If S consists of L non-intersecting arcs, the sequence of (N/N) Pade approximants to f(t) (defined by (2.2)) converges in capacity to f(t), as $N\to\infty$, in any closed, bounded domain not intersecting S.

This theorem shows that the diagonal Padé approximants converge in capacity away from a set of arcs whose location is completely determined by the location of the branch points; the location of these arcs is related to the "principle of minimum logarithmic capacity" (13). In Figs.2a and 2b we illustrate the concepts contained in Nuttall's results by considering a function with two (B_1,B_2) and four (B_1,B_2,B_3,B_4) branch points respectively. In Fig.2a the expected region of non-convergence (of the diagonal Padé approximants) is the arc B_1B_2 of the circle passing through O,B_1 and B_2; in Fig.2b (where B_4 is at infinity) the non-convergence region is contained within the shaded area.

The behaviour of quadratic approximants in the presence of branch points is somewhat different. For example, for the function $f(z)=\ln(1-z)$, Padé approximants simulate the branch cut at z=1 by clustering poles along the real axis below z=1; quadratic approximants, however, do not need to simulate the branch cut in this way since they have an "in-built" (square root) branch point. The poles of these quadratic approximants are in this sense "redundant"; where then do the poles of the quadratic approximants lie? From (2.7) of Chapter I we see that the approximants have poles when P(z)=0 provided the negative square root is chosen. For the above example, whenever P(z)=0 it is the positive square root which is used in the approximant; in a sense, the approximants place the poles on a sheet of the function in which we are not interested. From a study of many examples it seems that this behaviour is a general feature; quadratic (and presumably the higher order) approximants do not produce "unwanted" poles.

3. NUMERICAL EXAMPLES

To illustrate the nature of the results we may expect to obtain with quadratic approximants, we consider the examples of Baker ((2) and (17)).

FIG.2a: EXPECTED NON-CONVERGENCE REGION FOR A FUNCTION WITH TWO BRANCH POINTS AT B_1 AND B_2.

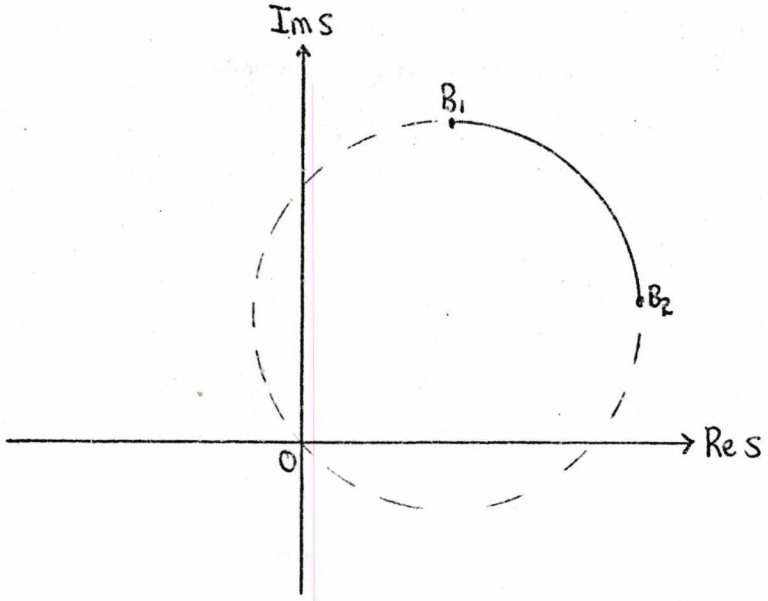
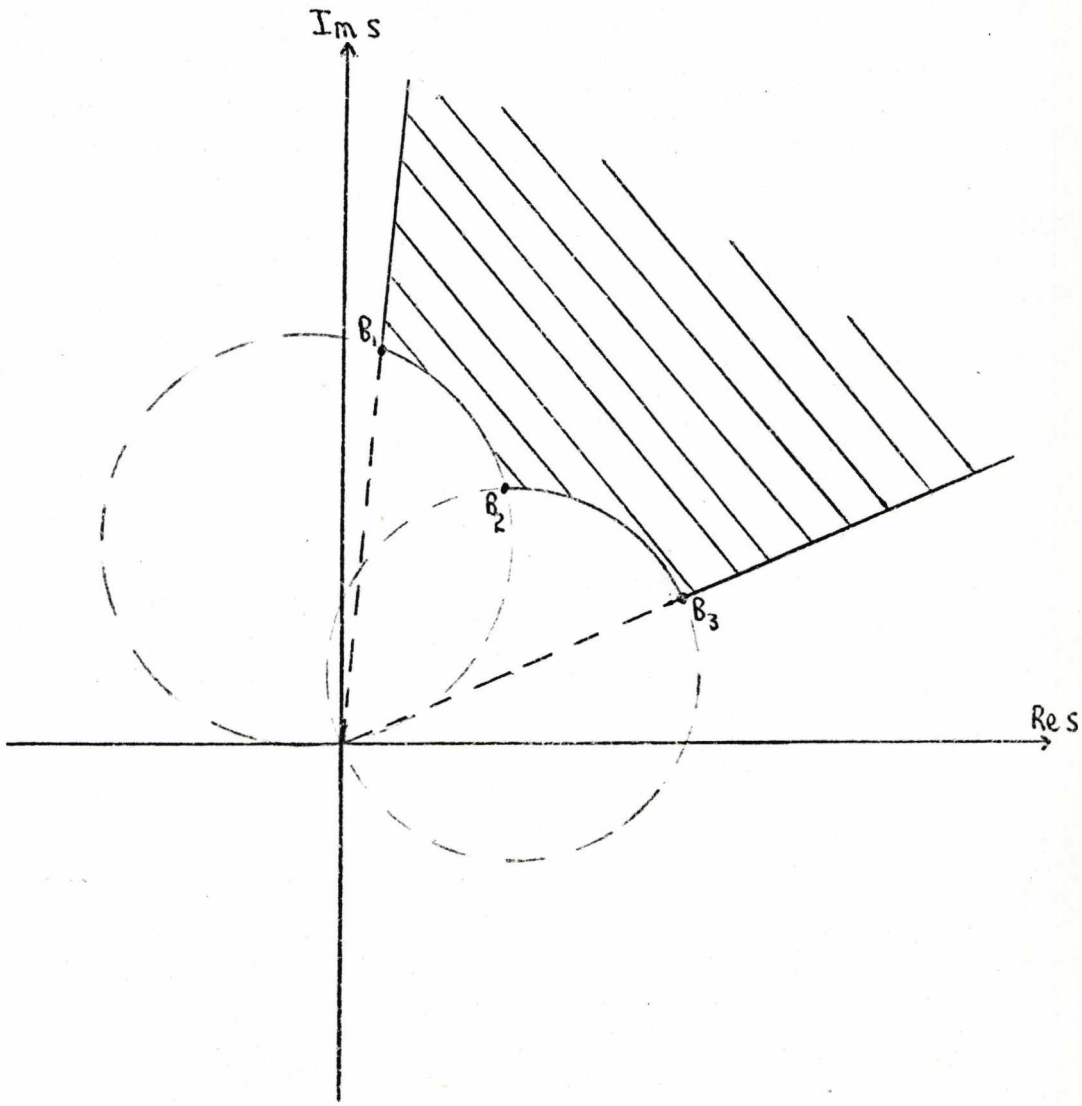


FIG.2b: EXPECTED NON-CONVERGENCE REGION FOR A FUNCTION WITH FOUR BRANCH POINTS AT B_1, B_2, B_3 AND $B_4=\infty$.



In Tables 2a and 2b we compare Padē and quadratic approximants to the following four functions:

(i) $f_1(z) = (1 - e^{-z})/z$

(ii) $f_2(z) = 1 - e^{-z}$

(iii) $f_3(z) = (1+z^2)^{1/2}/(1+z)$

and (iv) $f_4(z) = (\tan^{-1}\sqrt{z})^2$

In each case the approximant is evaluated at infinity. If we denote the (N/N) Padē and (N/N/N) quadratic approximants, evaluated at infinity, to the function $f_i(z)$ (i=1,2,3,4) by $f^{(i)}_{(N/N)}$ and $f^{(i)}_{(N/N/N)}$ respectively then we can make the following observations

(a) $f^{(1)}_{(N/N)}$ converges like 1/(N+1) whereas $f^{(1)}_{(N/N/N)}$ converges considerably faster.

(b) $f^{(2)}_{(N/N)}$ does not converge whereas $f^{(2)}_{(N/N/N)}$ converges fairly rapidly.

(c) $f^{(3)}_{(N/N)}$ does not converge and the remarkable convergence of $f^{(3)}_{(N/N/N)}$ is attributable to the fact that, for $N \geq 2$, $f^{(3)}_{(N/N)} = f_3(z)$; to see this we only need choose

$$Q(z) = 0, \quad P(z) = (1+z)^2, \quad R(z) = 1 + z^2$$

(d) The function $f_4(z)$ is regular in the complex plane cut from z=-1 to $z=-\infty$, and it is apparent that $f^{(4)}_{(N/N/N)}$ converges much more rapidly than $f^{(4)}_{(N/N)}$; the notation (a,b) of Table 2b means a+ib and we shall adopt this notation throughout the remainder of this thesis. This example also illustrates the convergence of the quadratic approximant along branch cuts; in Table 3 we give the errors in the real and imaginary parts of the (4/4/4) quadratic approximant for $z < -1$. In this region the Padē approximants, being the ratio of two real polynomials, must yield real (and hence non-convergent) results.

We thus see that, for this particular example, the quadratic approximants

(a) exhibit better convergence properties than the corresponding Padē approximants within the domain of convergence of the Padē approximants and

(b) converge outside the convergence domain of the Padē approximants

TABLE 2a: PADÉ AND QUADRATIC APPROXIMANTS TO $f_1(z)$ and $f_2(z)$, EVALUATED AT INFINITY

FUNCTION	FUNCTION VALUE AT INFINITY	N	(N/N) PADÉ APPROXIMANT	(N/N/N) QUADRATIC APPROXIMANT
$f_1(z)$	0	1	$-\frac{1}{2}$	-0.064
		2	1/3	0.042
		3	-1/4	-0.008
		4	1/5	0.006
		5	-1/6	
		6	1/7	
		7	-1/8	
		8	1/9	
$f_2(z)$	1	1	2	1.27
		2	0	0.87
		3	2	1.06
		4	0	1.02
		5	2	
		6	0	
		7	2	
		8	0	

TABLE 2b: PADÉ AND QUADRATIC APPROXIMANTS TO $f_3(z)$ AND $f_4(z)$, EVALUATED AT INFINITY.

FUNCTION	FUNCTION VALUE AT INFINITY	N	(N/N) PADÉ APPROXIMANT	(N/N/N) QUADRATIC APPROXIMANT
$f_3(z)$	1	1	0.33	0.61
		2	2.33	1
		3	0.412	
		4	2.412	
		5	0.414141	
		6	2.414141	
		7	0.414211	
		8	2.414211	
$f_4(z)$	2.46740	1	1.5	(2.39,0.27)
		2	1.87	(2.466,0.035)
		3	2.03	(2.46736,0.004)
		4	2.127	
		5	2.188	
		6	2.2299	
		7	2.2611	
		8	2.2850	

TABLE 3: ERRORS IN THE (4/4/4) QUADRATIC APPROXIMANT TO $f_4(z)$, FOR $z<-1$.

z	ERROR IN REAL PART OF (4/4/4) QUADRATIC APPROXIMANT	ERROR IN IMAGINARY PART OF (4/4/4) QUADRATIC APPROXIMANT
-10	6×10^{-6}	2×10^{-7}
-9	6×10^{-6}	1×10^{-6}
-9	5×10^{-6}	3×10^{-6}
-7	4×10^{-6}	5×10^{-6}
-6	1×10^{-6}	6×10^{-6}
-5	3×10^{-6}	5×10^{-6}
-4	7×10^{-6}	1×10^{-6}
-3	4×10^{-6}	9×10^{-6}
-2	2×10^{-5}	1×10^{-5}

and, most important, converge rapidly on the branch cut.

Throughout the remainder of this thesis we shall present results to suggest that these properties are a general feature of multi-valued approximants.

In Table 4 we list values of the limit as $z \to \infty$ of the (N/N) Padē and (N/N/N) quadratic approximants to the functions:

$$\text{(i)} \quad g_1(z) = \int_0^\infty \frac{e^{-t}}{1+zt}\, dt$$

$$\text{(ii)} \quad g_2(z) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{e^{-t^2}}{1+zt^2}\, dt$$

and

$$\text{(iii)} \quad g_3(z) = \int_0^\infty e^{-t} \left(\frac{1+2zt}{1+zt} \right)^{\frac{1}{2}} dt$$

These functions all have the property that they tend to a constant on every ray except the negative real axis as $z \to \infty$; $g_1(z)$ and $g_2(z)$ are both series of Stieltjes and hence the (N/N) Padē approximant must converge and $g_3(z)$ is the reciprocal of a series of Stieltjes. Although we cannot establish convergence of the quadratic approximants, it is evident from Table 4 that they are converging more rapidly than the Padē approximants. Furthermore, using the (6/6) Padē approximant, we calculate the value of Euler's divergent series (given by $g_1(1)$) to be 0.5968, compared with the exact value 0.5963; the (4/4/4) quadratic approximant yields the value 0.59636.

When a function does not tend at infinity to a constant times an integral power of z, then no Padē approximant can represent the function in the neighbourhood of infinity. Quadratic approximants can represent a larger class of functions than Padē approximants at infinity in that a function behaving at infinity like a constant times a half-integral power of z can be represented by a quadratic approximant.

To conclude this section we consider the two functions

$$\text{(i)} \quad h_1(z) = (1-z)^{-\frac{1}{3}}$$

and

$$\text{(ii)} \quad h_2(z) = (1-z)^{-\frac{1}{4}}$$

Using cubic and quartic approximants we can approximate $h_1(z)$ and $h_2(z)$ exactly; however, as the results of Table 5 indicate, we can obtain very good representations of these two functions using quadratic approximants. Again we note that, in the region of convergence of the Padē approximants to h_1 and h_2, the quadratic approximants still appear to converge more

TABLE 4: PADÉ AND QUADRATIC APPROXIMANTS TO $g_1(z)$, $g_2(z)$ AND $g_3(z)$, EVALUATED AT INFINITY

FUNCTION	FUNCTION VALUE AT INFINITY	N	(N/N) PADÉ APPROXIMANT	(N/N/N) QUADRATIC APPROXIMANT
$g_1(z)$	0	1	$\frac{1}{2}$	0.3
		2	1/3	0.23i
		3	1/4	0.065
		4	1/5	0.037
		5	1/6	
$g_2(z)$	0	1	2/3	0.5
		2	24/35	0.36
		3	0.46	0.19
		4	0.41	0.04
		5	0.37	
$g_3(z)$	1.41421	1	6/5	1.28
		2	376/297	1.36
		3	1.299	1.38
		4	1.319	1.3996
		5	1.333	

TABLE 5: PADÉ AND QUADRATIC APPROXIMANTS TO $h_1(z)$ AND $h_2(z)$. - INDICATES THE APPROXIMANT DOES NOT CONVERGE

FUNCTION	z	ERROR IN (6/6) PADÉ APPROXIMANT	ERROR(REAL AND IMAGINARY PART) IN(4/4/4) QUADRATIC APPROXIMANT
$h_1(z)$	-2	4×10^{-8}	2×10^{-10}
	2	———	$(8\times10^{-5}, 8\times10^{-5})$
	4	———	$(1\times10^{-4}, 5\times10^{-5})$
	6	———	$(1\times10^{-4}, 2\times10^{-4})$
	8	———	$(3\times10^{-4}, 2\times10^{-5})$
$h_2(z)$	-2	4×10^{-8}	10^{-10}
	2	———	$(8\times10^{-5}, 8\times10^{-5})$
	4	———	$(2\times10^{-4}, 5\times10^{-5})$
	6	———	$(1\times10^{-4}, 2\times10^{-4})$
	8	———	$(4\times10^{-4}, 5\times10^{-5})$

rapidly than the corresponding Padé approximants.

In practice, as well as calculating $f_{(p/q/r)}(z)$ from (2.6) (of Chapter I) it is sometimes advisable to also calculate $f_{(p/q/r)}(z)$ by

$$f_{(p/q/r)}(z) = -2R(z)\Big/\left[Q(z)-\left(Q^2(z)+4P(z)R(z)\right)^{\frac{1}{2}}\right] \qquad (2.3)$$

This can be important if, for example, $4P(z)\,R(z) \ll Q^2(z)$ and the positive square in (2.6) is required. Reformulation of (2.6) (of Chapter 1) as (2.3) ensures that no loss of accuracy occurs in this case (and also in the case when $2P(z) \ll -Q(z) \pm (Q^2(z)-4P(z)R(z))^{\frac{1}{2}}$).

4. THE AN-HARMONIC OSCILLATOR

As an indication of the range of applicability of the multi-valued approximants previously discussed, we consider the problem of determining the energy levels of the an-harmonic oscillator. The Hamiltonian for this system is

$$\mathcal{H} = p^2 + x^2 + \beta x^4$$

and the analytic properties of $\mathcal{H}$ are discussed in (18). For β real and positive $\mathcal{H}$ is well defined, and the energy levels are analytic functions of β in a neighbourhood of the positive real axis. The point $\beta=0$ is a singular point; it is in fact a limit point of singularities. Also

(a) the nth-energy level $E_n(\beta)$ has a "global" third-order branch point at $\beta=0$; by this we mean that any path of continuation which winds three times around $\beta=0$ and circles clockwise about all branch points, returns $E_n(\beta)$ to its starting point and a path that winds one (or two) times around does not and

(b) on the three-sheeted surface, $\beta=0$ is not an isolated singularity; there are therefore infinitely many singularities.

In view of (a) and (b) we may expect the multi-valued approximants to be of use in calculating $E_n(\beta)$. We write

$$E_n(\beta) = \sum_{p=0}^{\infty} a_p^{(n)} \beta^p$$

where the coefficients $a_p^{(n)}$ are tabulated in the literature; for the results quoted here we use the coefficients of (19). We tabulate the following:

(i) In Table 6 we compare the (10/10) Padé, (5/5/5) quadratic and

TABLE 6: PADÉ, QUADRATIC AND CUBIC APPROXIMANTS TO $E_N(\beta)$, THE NTH ENERGY LEVEL OF THE AN-HARMONIC OSCILLATOR, FOR VARIOUS N AND β (≤ 1).

N	β	ERROR IN (10/10) PADÉ APPROXIMANT	ERROR IN (5/5/5) QUADRATIC APPROXIMANT	ERROR IN (3/3/3/3) CUBIC APPROXIMANT
1	0.25	6×10^{-6}	5×10^{-8}	1×10^{-7}
	0.50	3×10^{-4}	4×10^{-6}	6×10^{-6}
	0.75	2×10^{-3}	1×10^{-5}	2×10^{-5}
	1.00	5×10^{-3}	7×10^{-5}	1×10^{-4}
3	0.25	2×10^{-4}	2×10^{-5}	7×10^{-6}
	0.50	6×10^{-3}	3×10^{-4}	1×10^{-4}
	0.75	3×10^{-2}	1×10^{-3}	1×10^{-3}
	1.00	6×10^{-2}	3×10^{-3}	2×10^{-3}
5	0.25	1×10^{-3}	7×10^{-4}	5×10^{-5}
	0.50	2×10^{-2}	4×10^{-3}	6×10^{-4}
	0.75	1×10^{-1}	1×10^{-2}	2×10^{-3}
	1.00	2×10^{-1} *	2×10^{-2} *	5×10^{-3} *
7	0.25	1×10^{-2}	2×10^{-4}	5×10^{-4}
	0.50	7×10^{-2}	2×10^{-3}	5×10^{-3}
	0.75	3×10^{-1}	1×10^{-3}	1×10^{-2}
	1.00	8×10^{-1}	2×10^{-2}	4×10^{-2}
9	0.25	3×10^{-2}	8×10^{-3}	2×10^{-3}
	0.50	2×10^{-1} *	8×10^{-2} *	2×10^{-2} *
	0.75	9×10^{-1} *	2×10^{-1} *	6×10^{-2} *
	1.00	1.9 *	0.5 *	0.1 *

TABLE 7: PADÉ AND QUADRATIC APPROXIMANTS TO $E_0(\beta)$ FOR $\beta \geq 1$.

FIGURES QUOTED IN BRACKETS HAVE NOT YET CONVERGED.

β	(20/20) PADÉ APPROXIMANT	(5/5/5) QUADRATIC APPROXIMANT
1	1.39234	1.3923(5)
2	1.6071	1.607(6)
3	1.767	1.76(98)
4	1.897	1.9(03)
5	2.00(5)	2.0(19)
6	2.10(0)	2.1(22)
7	2.18(2)	2.2 (16)
8	2.25(0)	2.3(0)
9	2.31(3)	2.3(8)
10	2.37(0)	2.4(5)

(3/3/3/3) cubic approximants (requiring respectively 21,17 and 15 coefficients $a_p^{(n)}$) for various values of n and β, with $\beta \leq 1$. The exact values of $E_n(\beta)$ used to calculate the tabulated errors are those given in (19); an entry with an asterisk denotes that the 'exact' value of $E_n(\beta)$ used is only an upper bound to the eigenvalue.

(ii) In Table 7 we quote results obtained for $\beta \geq 1$ for the ground state energy $E_o(\beta)$. The quoted values for the (20/20) Padé approximant are those given in (18); the results obtained for the (5/5/5) quadratic and (3/3/3/3) cubic approximants are almost identical and we only quote the former. Figures quoted in brackets must be regarded as unreliable, since they do not yet appear to have converged.

From these results we can see the usefulness of the quadratic and cubic approximants; for $\beta < 1$ they appear to give about two more significant figures than the corresponding Padé approximants. For $\beta > 1$ the approximants compare favourably with the Padé approximants, remembering that in Table 7 the quadratic approximants use less than half the number of terms required by the Padé approximants. In particular, for $\beta = 1$, the two approximation schemes give almost identical accuracy, the Padé approximants with 41 terms and the quadratic approximants with only 17 terms.

Finally, we remark that the method of Borel summability (discussed in Chapter I) can be applied to the multi-valued approximants to obtain improved convergence, as is the case with Padé approximants (18). The one disadvantage is that in order to perform the required numerical integration, we need to know which branch of the approximant to choose and there appears to be no general method of doing this.

5. SOME FURTHER APPLICATIONS OF QUADRATIC APPROXIMANTS

(i) SEQUENCE EXTRAPOLATION

As mentioned in Chapter I, even if a series converges it may converge very slowly and Padé (and hence quadratic) approximants can be used to accelerate convergence. Given a sequence $a_1, a_2, .., a_n$ (generated, for example, during an iterative procedure to solve an equation) we form the

series (20)

$$G(z) = a_1 + (a_2 - a_1)z + (a_3 - a_2)z^2 + \cdots\cdots \qquad (5.1)$$

where, by definition, the exact result is $U_\infty = G(1)$. Generally, provided G(z) is regular at z=1, Padē approximants to G(z) at z=1 will give a better estimate of the final answer than the last computed term a_n.

To compare the usefulness of Padē and quadratic approximants as sequence extrapolation devices, we give the following two examples:

(i) If $a_i = i^{-1}$, so that G(1)=0, we obtain

$$a_{14} = 0.07$$

$$G_{(7/7)}(1) = 0.02$$

$$G_{(4/4/4)}(1) = 0.006$$

(ii) A major application of the sequence extrapolation technique (using Padē approximants) is in numerical quadrature (21); here the sequence (a_n) is normally generated by successively halving the integration step length. For the function

$$F(z) = \int_0^1 \frac{dz}{z^{\frac{1}{2}} + z^{\frac{1}{3}}}$$

and using the mid-point trapezoidal rule, we obtain

$$\left| G_{(5/5)}(1) - F(z) \right| = 6 \times 10^{-6}$$

and

$$\left| G_{(3/3/3)}(1) - F(z) \right| = 4 \times 10^{-7}$$

However, for difference integrands extrapolation with Padē approximants is preferable; for the integrands considered in (21) the two extrapolation methods give comparable results.

(ii) APPROXIMANTS OF TYPE II

When G(z) in (5.1) is not regular at z=1 the following procedure can be used (20), which we illustrate for the series

$$S = \sum_{n=1}^{\infty} \frac{1}{n^2} \qquad (5.2)$$

Let

$$S_p = \sum_{n=1}^{p} \frac{1}{n^2} \qquad (5.3)$$

The series (5.2) converges slowly and we can accelerate the convergence using

the previously considered sequence extrapolation method; for example, the (1/1) Padé approximant gives the result 1.45, compared with the exact result $S=\pi^2/6 = 1.6449$. We can, however, improve upon this estimate whilst still only using the first 3 terms of (5.2). We consider the S_p as functions of the variable $1/p$, so that $S_p=F(1/p)$, with the exact sum as $F(0)$. This method gives, using only the terms S_1, S_2 and S_3 of (5.3) (which incorporates the same amount of information as the (1/1) Padé approximant)

$$G_{(1/1)}(0) = 1.65$$

We can obviously employ the same procedure with quadratic approximants; as a comparison we give the results

$$\left| G_{(2/2)}(0) - S \right| = 4\times10^{-5}$$

and

$$\left| G_{(1/1/1)}(0) - S \right| = 2\times10^{-5}$$

From this type of example we are led to the following definitions

DEFINITION 5.1 (23): Let $z_1,\ldots,z_p$ be p complex numbers. We define the Padé approximant of type II, $f^{II}_{(N/M)}(z)$, to the function $f(z)$ by

$$f^{II}_{(N/M)}(z) = \frac{P_N(z)}{Q_M(z)}$$

with

$$f^{II}_{(N/M)}(z_i) = f(z_i) \quad (i=1,2,\ldots,p)$$

and

$$p = N + M + 1$$

where $P_N(z)$ and $Q_M(z)$ are polynomials of maximum degree N and M respectively.

DEFINITION 5.2: Let $z_1,\ldots,z_p$ be p complex numbers and $P_L(z)$, $Q_M(z)$ and $R_N(z)$ polynomials of maximum degree L, M and N respectively. Then we define the quadratic approximant of type II, $f^{II}_{(L/M/N)}(z)$, to the function $f(z)$ by the equation

$$P(z)\left[f^{II}_{(L/M/N)}(z)\right]^2 + Q(z) f^{II}_{(L/M/N)}(z) + R(z) = 0$$

with

$$f^{II}_{(L/M/N)}(z_i) = f(z_i) \quad (i=1,2,\ldots,p)$$

and

$$p = L + M + N + 2 \ .$$

A practical application of type II Padé approximants is the Bulirsch-Stoer method in numerical quadrature (22). A sequence (S_p) of approximants, corresponding to division of the integration region into sub-intervals of length h_p, are generated (by, for example, the trapezoidal rule). Padé approximants of type II, denoted by R(h), are then formed according to

$$R(h_p) = S_p \qquad (p = 1, 2, \ldots\ldots, M+N+1)$$

where

$$R(h) = \frac{a_0 + a_1 h^2 + \ldots\ldots + a_N h^{2N}}{1 + b_1 h^2 + \ldots\ldots + b_M h^{2M}}$$

The integral is then approximated by

$$R(0) = a_0 .$$

Again we can use quadratic approximants in place of Padé approximants in this procedure; in practice, the two methods give similar results for the examples so far considered.

(iii) SINGULAR INTEGRALS

We have in mind the numerical evaluation of integrals, the integrands of which have a branch point within the interval of integration; for example,

$$\int_a^b (x - \alpha)^{\beta/2} \qquad (a < \alpha)$$

with β an integer. Generally, we consider the problem of evaluating

$$I = \int_a^b f(x)\, dx \tag{5.4}$$

Using the method more fully explained in Chapter 3, we expand f(x) in a Taylor series about a point x_0 in the complex plane:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \tag{5.5}$$

Then, from (5.4) and (5.5),

$$I = \sum_{n\ 0}^{\infty} \frac{f^{(n)}(x_0)}{(n+1)!} (x - x_0)^{n+1} \Bigg|_{x = a - x_0}^{x = b - x_0} \tag{5.6}$$

We now form approximants (Padé or quadratic) to the power series defined by

(5.6); denoting the approximant by G(x), we obtain as an approximation to I

$$I \approx G(b-x_0) - G(a-x_0)$$

To illustrate the use of this method, we consider the following three integrals:

(a) $$I_1 = \int_1^2 (x-\tfrac{3}{2})^{\frac{1}{2}}\, dx$$

(b) $$I_2 = \int_0^2 \ln(1-x)\, dx$$

and

(c) $$I_3 = \int_0^1 x^{\frac{1}{2}}\, dx$$

In Table 8 we compare the results obtained using Padé and quadratic approximants for these three examples; the method itself appears to work reasonably well and again the quadratic approximants provide a better method of approximation than the Padé approximants.

TABLE 8: PADÉ AND QUADRATIC APPROXIMANTS TO THE INTEGRALS I_1, I_2 AND I_3.

INTEGRAL AND EXACT VALUE	N	(N/N) PADÉ APPROXIMANT	N	(N/N/N) QUADRATIC APPROXIMANT	x_o
I_1=(0.23570226, 0.23570226)	6	$(6\times10^{-3}, 8\times10^{-2})$	3	$(10^{-8}, 10^{-8})$	-1+i
I_2=(-2, π)		$(4\times10^{-4}, 8\times10^{-3})$	5	$(2\times10^{-6}, 2\times10^{-6})$	1+i
I_3=0.66		$(3\times10^{-4}, 1\times10^{-3})$	4	$(1\times10^{-7}, 6\times10^{-8})$	1+i

1. J.S.R. Chisholm, "Convergence properties of Padé approximants" in "Padé Approximants and their Applications" (ed. P.R. Graves-Morris), p.11-21 (1973).

2. G.A. Baker, Jr., "Essentials of Padé Approximants", Chapters 11-14 (1975).

3. De Montessus de Ballore, Bull.Math.Soc.de France 30,28 (1902).

4. R. Wilson, Proc.London Math.Soc. 27, 497-512 (1928a). (A comprehensive list of related papers is given in (2)).

5. E.B. Saff, Jour.Approx.Theory 6, 63-67 (1972).

6. W.B. Gragg, Siam Review 14, 1-62 (1972).

7. A.F. Beardon, Jour. Math.Analysis and Applns. 21, 469 (1968b).

8. G.A. Baker, Jr., Jour. Math. Analysis and Applns.43, 498-528 (1973).

9. H. Wallin, "The convergence of Padé approximants and the size of the power series coefficients", Applicable Anal. 4, 235-251 (1975).

10. G.A. Baker, J.L. Gammel and J.G. Wills, Jour. Math. Analysis and Applns. 2, 405 (1961).

11. T.J. Stieltjes, Ann.Fac.Sci.Univ.Toulouse Sci. Math.Sci.Phys. 8,9,1 (1894).

12. G.A. Baker, Jr., "The theory and application of the Padé approximant method" in Advances in Theoretical Physics, (1965)

13. J. Nuttall, Jour. Math.Anal. and Appl. 31,147-153 (1970).

14. Ch.Pommerenke, Jour.Math.Anal. and Appl. 41,775-780 (1973).

15. J. Nuttall, "Orthogonal pdynomials for complex weight functions and the convergence of related Pade approximants", unpublished (1972).

16. J. Nuttall, "The convergence of Pade approximants to functions with branch points", Jour. Approx.Theory, to appear.

17. G.A. Baker,Jr., Advances in Theor.Phys. 1, 1-58 (1965).

18. B. Simon and A. Dicke, Annals of Physics 58,76-136 (1970).

19. C.E. Reid, International Jour. Quantum Chemistry 1, 521-534 (1967).

20. J.L. Basdevant, "The Padé approximant and its physical applications", Cem preprint, Ref.Th.1441 (December 1971).

21. J.S.R. Chisholm, A.C. Genz and G.E. Rowlands, Jour.Comp.Phys.10,284-307(1972)

22. R. Bulirsch and J. Stoer, Numer.Math.6, 413 (1964).

23. J. Zinn-Justin, Physics Reports 1, 55 (1971).

CHAPTER 3: CALCULATION OF FEYNMAN INTEGRALS IN THE PHYSICAL REGION.

1. REPRESENTATION OF FEYNMAN INTEGRALS

The purpose of this chapter is to extend the methods used in (1) to calculate Feynman integrals in the physical region. The general parametric representation of Feynman integrals is well known ((2) and (3)). For interacting scalar fields the representation is of the form

$$\lim_{\epsilon \to 0+} \int_0^1 \prod_{i=1}^{n} d\alpha_r \, \delta\left(1 - \sum_{r=1}^{n} \alpha_r\right) \frac{C^{n-2l-2}}{[D + i\epsilon]^{n-2l}} \qquad 1.1)$$

where C is a real polynomial in the Feynman parameters $\{\alpha_r\}$

D is a polynomial in $\{\alpha_r\}$ which is also a <u>linear</u> function of the invariants s,t,...of the physical process being considered (where, as usual $\sqrt{s}$ denotes the total energy and t the four-momentum transfer); r is the number of internal lines in the graph (which represents the particular physical process being considered);

l is the number of internal momenta in the graph and

n is the number of vertices in the graph.

We do not prove (1.1) but, by considering certain specific examples, we shall later illustrate the general features of the proof.

In general, the singularity structure of the integrands in (1.1) is very complicated, but they possess some standard features (3). A typical set of singularities in the s-plane is illustrated in Fig I, and consists of a pole P and a series of branch points $B_1, B_2, \ldots$ on the positive real axis; for clarity, in Fig I the branch cuts associated with these branch points have been displaced from the positive real axis, where standard techniques of calculation (3a) place them. The variable s is essentially the energy of the system, and we wish to calculate (1.1) for real positive values of s, exactly where the discontinuity in the function is normally placed. The limiting process $\epsilon \to 0+$ prescribes that we calculate (1.1) as s approaches the real axis from above, as indicated in Fig I; this procedure corresponds to calculating (1.1) in the "physical region".

The basic problem is therefore the calculation of a function, in this

FIG.1: TYPICAL SINGULARITY STRUCTURE OF (1.1) IN THE RIGHT-HALF OF THE s-PLANE.

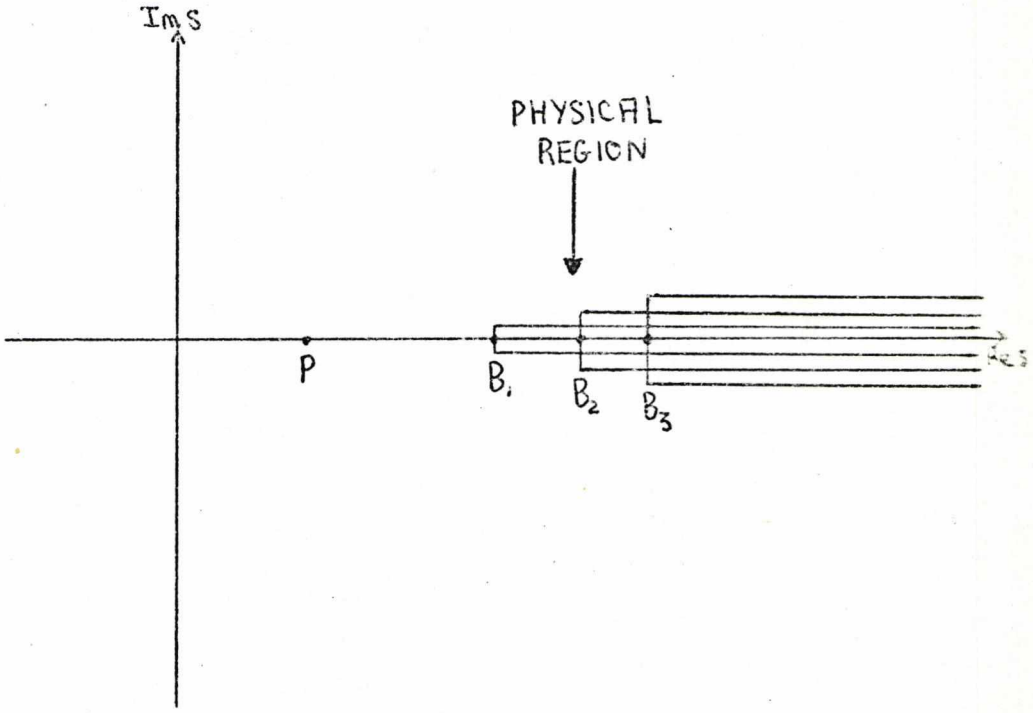
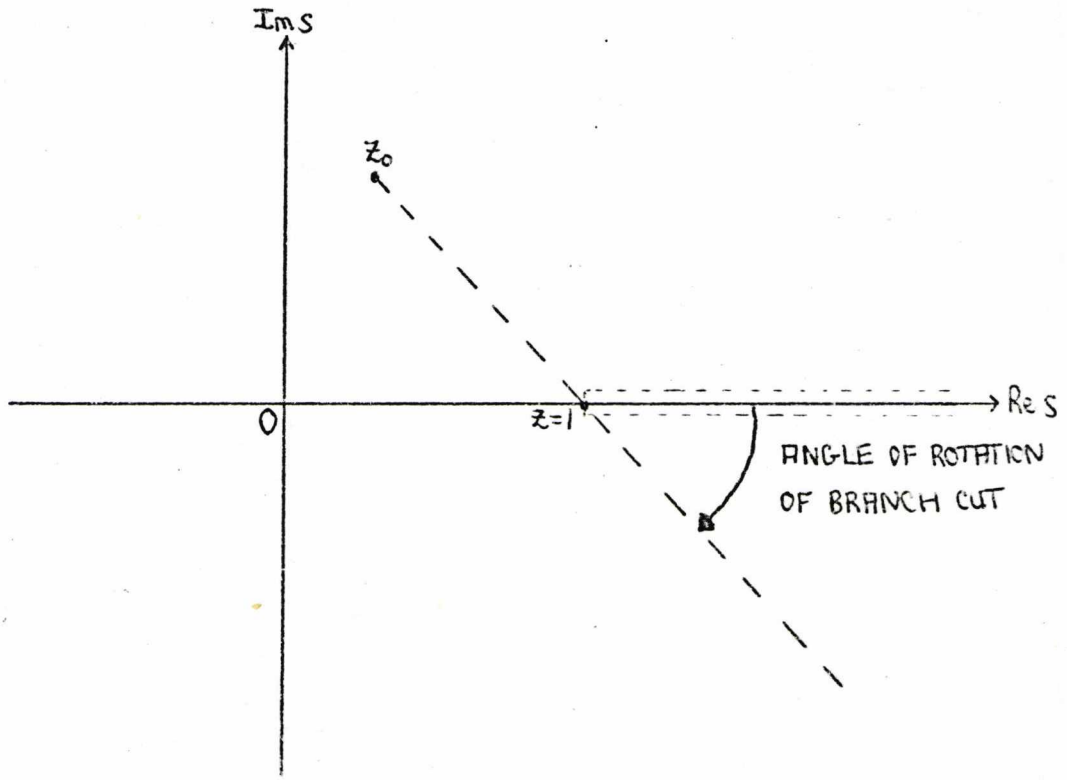


FIG.2: EFFECTIVE ROTATION OF THE BRANCH CUT FOR ln(1-z)



case that given by (1.1), in a region where the function has one or more branch points. Before discussing the method used for Feynman integrals we first illustrate the procedure applied to a 'simpler' function, which we choose as ln(1-z).

2. METHOD OF CALCULATION

The calculational procedure is based on the idea incorporated in the following theorem of Van Vleck (4).

THEOREM 2.1: Consider the S-fraction

$$\cfrac{a_0}{1-\cfrac{a_1 z}{1-\cfrac{a_2 z}{\ddots}}} \tag{2.1}$$

Then: (a) if $\lim_{p\to\infty} a_p = 0 \quad (a_p \neq 0)$

the S-fraction converges to a meromorphic function of z. The convergence is uniform over every closed, bounded region containing none of the poles of this function.

(b) if $\lim_{p\to\infty} a_p = a \neq 0$

the S-fraction converges in the domain exterior to the rectilinear cut running from $^1/_{4a}$ to ∞ in the direction of the vector from 0 to $^1/_{4a}$, to a function having at most polar singularities in this domain. The convergence is uniform on every closed, bounded region exterior to the cut which contains no poles of the function.

We can illustrate this theorem for the function ln(1-z), which has the continued fraction expansion (4)

$$\frac{\ln(1-z)}{z} = \cfrac{-1}{1-\cfrac{1^2 z}{2-\cfrac{1^2 z}{3-\cfrac{2^2 z}{\ddots}}}} \tag{2.2}$$

from which it follows that, in the notation of Theorem 1.1,

$$a_{2p+1} = \frac{(1+p)}{2(1+2p)} \quad , \quad a_{2p+2} = \frac{(1+p)}{2(3+2p)}$$

so that

$$\lim_{p\to\infty} a_p = \frac{1}{4}$$

Theorem 1.1 therefore shows that the S-fraction to $f(z)=\ln(1-z)$ converges uniformly in the entire z-plane cut from $z=1$ to $z=\infty$ along the real axis. Since the S-fraction (2.1) is equivalent to a diagonal and a paradiagonal sequence of Padē approximants (4), we are assured of uniform convergence of these approximants in the cut z-plane.

Now suppose we move our origin to the point z_o, which lies in the upper half of the z-plane. Then, according to Van Vleck's theorem, the S-fraction (2.2) will converge uniformly in the z-plane cut from $z=1$ to $z=\infty$ along the line joining $z=z_o$ and $z=1$. This situation is illustrated in Fig 2; the S-fraction (and the equivalent Padē approximant) now converges on the real z-axis (except possibly at $z=1$). In effect, we have rotated the cut by moving our origin from $z=0$ to $z=z_o$.

For the particular example $\ln(1-z)$ we can see this behaviour explicit Expansion about $z=0$ gives

$$f(z) = -\sum_{n=1}^{\infty} \frac{z^n}{n}$$

and, if we calculate the (m/m) and (m/m-1) Padē approximants to $f(z)$, the poles and zeros of the approximants will simulate the cut on the real axis (4). We move the origin to the point z_o by writing

$$\ln(1-z) = \ln(1-z_o) - \ln\left(1 - \frac{(z-z_o)}{(1-z_o)}\right) \qquad (2.3a)$$

$$= \ln(1-z_o) - \sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{z-z_o}{1-z_o}\right)^n \qquad (2.3b)$$

If we now calculate the Padē approximants to the series (2.3b) then, from (2.3a) the poles and zeros will now lie on the point set

$$\left\{ z - z_o = \alpha(1-z_o) \;;\; \alpha > 1 \right\}$$

and this set is just that part of the straight line through z_o and 1 which lies in the lower half-plane; the branch cut has thus been rotated into the lower half-plane.

Originally (1) Padē approximants to the series (2.3b) were used to

calculate f(z) at real positive values of z_1 both for z<1 and z>1; for z>1, the branch is defined by approaching the real axis from above, giving the value

$$\ln|-z| - i\pi$$

Other re-summing procedures can, however, be used. In Table I we compare the results obtained with the (8/8) Padé and (5/5/5) quadratic approximants (both requiring 17 terms of the series (2.3b)), and with the expansion point $z_o=1+i$. The main observation from Table I is that the quadratic approximants reduce the errors by roughly two orders of magnitude as compared with the Padé approximants; later results will show that this behaviour is by no means confined to this particular example. The increase in the error very close to the branch point occurs because the logarithm is unbounded at z=1 whereas the square root function is bounded in the neighbourhood of the branch point. We also note that reasonable results can be obtained even for $z_o=0$ when quadratic approximants are used; the discontinuity across the cut, now simulated on the real z-axis (z>1), can be represented by the square root term associated with the quadratic approximants. In contrast, the Padé scheme cannot work for $z_o=0$.

3. CALCULATIONAL PROCEDURE FOR FEYNMAN INTEGRALS

We adopt the following method of calculation (see (1)):

(i) For the graphs considered here, D (of (1.1)) is a function of at most the two invariants s and t (in the usual notation). We expand D about $(s-s_o,t)$ or $(s-s_o,t-t_o)$, where s_o and t_o lie in the upper half of the complex s and t plane; this expansion is easily performed since D is a linear function of its invariants.

(ii) Perform the α-integrations for a given number, N, terms of the series.

(iii) Approximate the sum of this series by a suitable type of approximant.

This scheme has the following features:

(a) If the concepts of Nuttall's theorems (see Chapter 2) hold more generally, then the simulated branch cuts for the singularities of Fig I will

TABLE 1: PADÉ AND QUADRATIC APPROXIMANTS TO $\ln(1-z)$, WITH $z_o = 1+i$.

z	ERROR IN (8/8) PADÉ APPROXIMANT	ERROR IN (5/5/5) QUADRATIC APPROXIMANT
0.1	$(2\times10^{-6}, 6\times10^{-7})$	$(5\times10^{-9}, 3\times10^{-9})$
0.3	$(3\times10^{-6}, 6\times10^{-8})$	$(3\times10^{-9}, 1\times10^{-8})$
0.5	$(1\times10^{-5}, 3\times10^{-7})$	$(1\times10^{-7}, 2\times10^{-8})$
0.7	$(4\times10^{-5}, 3\times10^{-5})$	$(1\times10^{-7}, 2\times10^{-7})$
0.9	$(8\times10^{-4}, 7\times10^{-3})$	$(4\times10^{-6}, 3\times10^{-5})$
1.1	$(2\times10^{-3}, 5\times10^{-3})$	$(3\times10^{-5}, 2\times10^{-5})$
1.3	$(1\times10^{-5}, 8\times10^{-5})$	$(7\times10^{-8}, 2\times10^{-8})$
1.5	$(7\times10^{-6}, 3\times10^{-6})$	$(2\times10^{-8}, 2\times10^{-8})$
1.7	$(3\times10^{-6}, 1\times10^{-6})$	$(5\times10^{-9}, 1\times10^{-8})$
1.9	$(1\times10^{-6}, 5\times10^{-7})$	$(2\times10^{-9}, 9\times10^{-9})$

lie within the shaded region of Fig 3.

(b) The new denominator D_o (where, for example, $D_o = D(s)$ $(s=s_o)$ is linear in the complex variable s_o, and is complex for real α. Hence, there are no singularities in the α-integrals.

(c) The same denominator D_o occurs in all the terms of the series, and need only be calculated once at each quadrature point; computationally this is a very important point.

(d) The final results should be independent of the choice of the expansion point s_o, thus giving a check on the calculations. The problem then arises of making the optimal choice for s_o; later in the chapter we shall discuss this problem more fully.

(e) Since it is assumed that the branch cuts have been moved into the lower half s-plane, the limit process $\epsilon \to 0+$ of (1.1) can be omitted, the integral being available directly on the real axis, except at P and possibly at the branch points.

To illustrate the nature of the results we may expect from this method, we now consider individual Feynman graphs. For each graph considered, the explicit form of (1.1) is derived in an appendix at the end of this chapter; the singularity structure of the graphs is also considered in the appendices.

4. SECOND ORDER SELF ENERGY GRAPH AND ZERO MOMENTUM VERTEX PART

(a) SECOND ORDER SELF ENERGY

We first consider the renormalized second order self energy of a scalar particle of mass m, emitting and absorbing a scalar particle of mass μ. The matrix element for the process is proportional to

$$F(s) = \int_0^1 \ln \left[\frac{\mu^2(1-\alpha) + m^2\alpha - \alpha(1-\alpha)s}{\mu^2(1-\alpha) + m^2\alpha^2} \right] d\alpha \qquad (4.1)$$

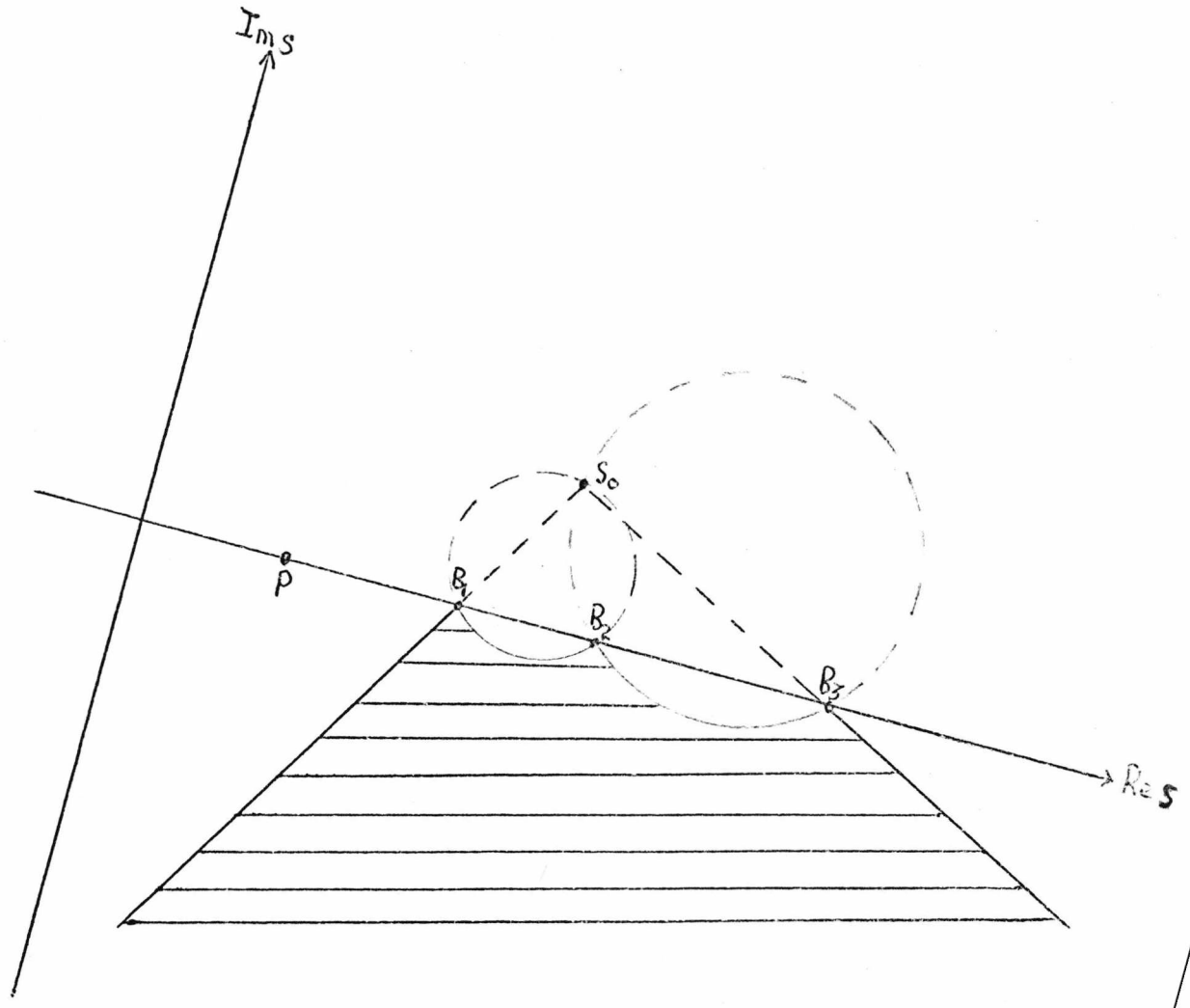
F(s) has a branch point at

$$s_1 = (m+\mu)^2$$

lying on the physical sheet (and corresponding to B_1 in Fig I) and an unphysical singularity (that is, one lying on an unphysical sheet) at

$$s_2 = (m-\mu)^2$$

FIG.3: REGION WHERE SIMULATED BRANCH CUTS ARE EXPECTED TO BE PLACED, WITH EXPANSION POINT s_o.



together with a 'second-type' singularity at s=0 (see Appendix I).

Before giving any numerical results for this graph, we discuss the integration method used. In the original paper (1) two quadrature methods were considered:

(i) the mid-point rule and

(ii) the composite two-point Gauss rule.

In all the graphs studied these two methods have been found to be not the most efficient, in the sense of accuracy obtained compared to the number of quadrature points used. In general, Gauss-Legendre quadrature appears to be the most efficient method and, unless explicitly stated otherwise, this is the quadrature rule used in all the following examples.

Table 2 gives a typical selection of the results obtained; here we compare the (6/6) Padé and (4/4/4) quadratic approximants obtained with the mass parameters m=1 and μ=1/6, which produces a branch point at s=1.36 on the physical sheet. The integration rule has a maximum of 23 points in each dimension. The improved convergence obtained with the quadratic approximants is evident; we again find that we obtain roughly two more significant figures compared with the Pade approximants. The optimal choice of s_0 seems to be given by

$$\mathrm{Re}\, s_0 \approx s_1 \,, \quad \mathrm{Im}\, s_0 \approx 1$$

Results of a similar nature are obtained with cubic approximants, as can be seen from Table 3; here we tabulate the (3/3/3/3) cubic approximant with s_0=1+i, again with the mass parameters m=1 and μ=1/6.

We now turn to a consideration of the un-physical sheets. The method of §3 is expected to produce results only on the physical sheet, since the power series expansion of (4.1) only incorporates information about F(s) on the physical sheet; using multi-valued approximants, however, we may hope to obtain convergence on more than one sheet. The self energy matrix element has infinitely many Riemann sheets, since the logarithm has, and essentially we have, for $s>s_1$,

$$F(s) = F_1(s) + (2n-1)i\pi F_2(s)$$

where n is an integer and F_1 and F_2 are suitable functions of s. The

TABLE 2: COMPARISON OF THE (6/6) PADÉ AND (4/4/4) QUADRATIC APPROXIMANTS TO THE SECOND ORDER RENORMALIZED SCALAR SELF ENERGY, FOR VARIOUS VALUES OF s_o.

s	s_o=1+i MODULI OF ERRORS	s_o=1.4+i MODULI OF ERRORS	s_o=1.8+i MODULI OF ERRORS	EXACT VALUE OF F(s).
1.00	P.A. $(1\times10^{-5},2\times10^{-5})$ Q.A. $(2\times10^{-7},1\times10^{-7})$	P.A. $(4\times10^{-4},4\times10^{-4})$ Q.A. $(9\times10^{-6},7\times10^{-6})$	P.A. $(9\times10^{-3},2\times10^{-3})$ Q.A. $(6\times10^{-5},3\times10^{-6})$	(0,0)
1.25	P.A. $(7\times10^{-4},1\times10^{-3})$ Q.A. $(3\times10^{-6},5\times10^{-6})$	P.A. $(4\times10^{-3},6\times10^{-3})$ Q.A. $(8\times10^{-5},5\times10^{-5})$	P.A. $(1\times10^{-2},3\times10^{-2})$ Q.A. $(1\times10^{-4},1\times10^{-4})$	(0.38520611,0)
1.50	P.A. $(8\times10^{-3},2\times10^{-4})$ Q.A. $(1\times10^{-5},2\times10^{-5})$	P.A. $(2\times10^{-4},5\times10^{-3})$ Q.A. $(3\times10^{-5},5\times10^{-5})$	P.A. $(6\times10^{-3},7\times10^{-4})$ Q.A. $(1\times10^{-4},2\times10^{-4})$	(0.87764128,0.70055158)
1.75	P.A. $(6\times10^{-4},1\times10^{-3})$ Q.A. $(3\times10^{-7},5\times10^{-6})$	P.A. $(6\times10^{-4},6\times10^{-7})$ Q.A. $(4\times10^{-6},3\times10^{-6})$	P.A. $(3\times10^{-4},2\times10^{-4})$ Q.A. $(1\times10^{-5},3\times10^{-5})$	(0.72505492.1.15017822)
2.00	P.A. $(4\times10^{-4},2\times10^{-4})$ Q.A. $(2\times10^{-6},4\times10^{-6})$	P.A. $(2\times10^{-4},6\times10^{-5})$ Q.A. $(3\times10^{-7},2\times10^{-6})$	P.A. $(1\times10^{-4},1\times10^{-4})$ Q.A. $(2\times10^{-5},2\times10^{-5})$	(0.57377476,1.43459801)
2.25	P.A. $(3\times10^{-4},3\times10^{-4})$ Q.A. $(4\times10^{-6},1\times10^{-6})$	P.A. $(2\times10^{-4},6\times10^{-5})$ Q.A. $(2\times10^{-6},3\times10^{-7})$	P.A. $(1\times10^{-4},2\times10^{-4})$ Q.A. $(3\times10^{-5},3\times10^{-5})$	(0.43055756,1.64185143)
2.50	P.A. $(1\times10^{-4},4\times10^{-4})$ Q.A. $(3\times10^{-6},7\times10^{-6})$	P.A. $(2\times10^{-5},2\times10^{-4})$ Q.A. $(1\times10^{-6},1\times10^{-6})$	P.A. $(7\times10^{-5},1\times10^{-4})$ Q.A. $(6\times10^{-5},4\times10^{-5})$	(0.29715650,1.80200491)
2.75	P.A. $(5\times10^{-4},6\times10^{-5})$ Q.A. $(6\times10^{-6},7\times10^{-6})$	P.A. $(3\times10^{-4},3\times10^{-5})$ Q.A. $(6\times10^{-7},2\times10^{-6})$	P.A. $(2\times10^{-4},1\times10^{-4})$ Q.A. $(3\times10^{-5},1\times10^{-4})$	(0.17353695,1.93025880)
3.00	P.A. $(2\times10^{-4},5\times10^{-4})$ Q.A. $(2\times10^{-5},4\times10^{-7})$	P.A. $(8\times10^{-5},4\times10^{-4})$ Q.A. $(1\times10^{-6},3\times10^{-6})$	P.A. $(9\times10^{-4},3\times10^{-4})$ Q.A. $(5\times10^{-5},1\times10^{-4})$	(0.05903536,2.03559403)

TABLE 3: (3/3/3/3) CUBIC APPROXIMANT TO THE SECOND ORDER RENORMALIZED SCALAR SELF ENERGY, WITH $s_o=1+i$.

s	ERROR IN (3/3/3/3) CUBIC APPROXIMANT
1.00	$(1x10^{-8}, 2x10^{-9})$
1.25	$(1x10^{-7}, 2x10^{-7})$
1.50	$(2x10^{-7}, 1x10^{-6})$
1.75	$(5x10^{-7}, 2x10^{-7})$
2.00	$(3x10^{-7}, 4x10^{-7})$

TABLE 4: QUADRATIC AND CUBIC APPROXIMANTS TO THE SECOND ORDER RENORMALIZED SCALAR SELF ENERGY ON THE UNPHYSICAL SHEET, WITH $s_o=1+i$.

s	ERROR IN (6/6/6) QUADRATIC APPROXIMANT	ERROR IN (3/3/3/3) CUBIC APPROXIMANT	EXACT VALUE
1.4	$(5x10^{-5}, 3x10^{-4})$	$(4x10^{-4}, 4x10^{-4})$	(0.93516,-0.3717)
1.5	$(9x10^{-4}, 4x10^{-4})$	$(5x10^{-4}, 2x10^{-4})$	(0.8777,-0.7006)
1.6	$(2x10^{-4}, 2x10^{-3})$	$(2x10^{-4}, 5x10^{-4})$	(0.81730,-0.9132)
1.7	$(3x10^{-3}, 2x10^{-3})$	$(1x10^{-4}, 7x10^{-4})$	(0.7559,-1.0788)
1.8	$(5x10^{-3}, 2x10^{-3})$	$(6x10^{-4}, 5x10^{-4})$	(0.6944,-1.2157)
1.9	$(4x10^{-3}, 6x10^{-3})$	$(1x10^{-3}, 2x10^{-4})$	(0.6335,-1.3327)
2.0	$(1x10^{-3}, 1x10^{-2})$	$(1x10^{-3}, 2x10^{-4})$	(0.574,-1.4346)

physical sheet is defined by n=1 and the first unphysical sheet by n=0; it is this unphysical sheet with which we are concerned. On this sheet we have three regions to consider:

(i) $s>s_1$: In this region both quadratic and cubic approximants give convergent results, as shown in Table 4 (which corresponds to the same parameters as Table 2). A noticeable feature is that the rate of convergence of the approximants is markedly less on the unphysical sheet than on the physical sheet. This is not surprising; it is rather more surprising that we obtain any convergence at all on the unphysical sheet. Although the values of ImF(s) on the physical and unphysical sheet differ only in sign, we cannot expect this sign difference to be exactly reproduced by the sign difference associated with the two branches of the square root term in the quadratic approximants. This is because all the power series coefficients of F(s) are complex; if, for example, we choose $s_o=0$ (thus making the coefficients real), then the quadratic approximants would produce exactly the same errors on the physical and unphysical sheet. It is the complex nature of the coefficients which make convergence on the unphysical sheet a non-trivial result.

(ii) $s_2<s<s_1$: In this region we obtain only very limited convergence; typically one or two significant figures, as compared with at least six figures on the physical sheet.

(iii) $s<s_2$: In this region the approximants do not converge and do not indicate the presence of the 'second-type' singularity at s=0.

Thus we see that, using multi-valued approximants, we can obtain a limited amount of information on the unphysical sheet together with very good results on the physical sheet.

This is a remarkable result; using only information about a function on one Riemann sheet, we have obtained (using the quadratic and cubic approximants) information on a second sheet of the function. Effectively, the quadratic (and cubic) approximants perform an analytic continuation of the function from one Riemann sheet to another; this is a very powerful form of analytic continuation.

(b) ZERO MOMENTUM VERTEX PART

The zero momentum vertex part is obtained by differentiating the unrenormalized scalar self energy graph with respect to m. (We need only consider the unrenormalized self energy graph since the ensuing differentiation is equivalent to the introduction of a zero momentum particle into the graph; the resulting graph is not divergent.) The matrix element is easily derived from F(s) of §4a, and is a multiple of

$$G(s) = \int_0^1 \frac{\alpha\, d\alpha}{\mu^2(1-\alpha) + m^2\alpha - \alpha(1-\alpha)s}$$

G(s) has the Taylor expansion (about s_0)

$$\sum_{n=0}^{\infty} (s-s_0)^n \int_0^1 \frac{\alpha^{n+1}(1-\alpha)^n\, d\alpha}{\left[\mu^2(1-\alpha) - m^2\alpha - \alpha(1-\alpha)s_0\right]^{n+1}}$$

As shown in Table 5 (again with m=1 and μ=1/6) the results obtained for G(s) exhibit the same features as those found for the self energy graph of §4a.

5. THREE POINT FUNCTIONS

5a. THE TRIANGLE GRAPH AND ANOMALOUS THRESHOLDS

With the notation of Appendix 2, the matrix element for this process is proportional to

$$M(s) = \int_0^1 du \int_0^1 \frac{u\, dv}{u^2 v(1-v)s + u(1-u)M^2 - m^2} \qquad (5a.1)$$

If we write

$$D(s) = u^2 v(1-v)s + u(1-u)M^2 - m^2$$

then the expansion of D(s) about s_0, a point in the upper half s-plane, is

$$D(s) = D_0 + u^2 v(1-v)(s-s_0)$$

where $D_0 = D(s_0)$; no higher order terms occur in this Taylor expansion since D(s) is linear in s. We then obtain the following expansion for M:

$$\sum_{r=0}^{\infty} (-1)^r (s-s_0)^r \int_0^1 du \int_0^1 \frac{u\left[u^2 v(1-v)\right]^r}{D_0^{r+1}}\, dv \qquad (5a.2)$$

The singularity structure of the triangle graph (see Appendix 2) shows that we have two cases to consider:

TABLE 5: COMPARISON OF THE (6/6) PADÉ AND (4/4/4) QUADRATIC APPROXIMANTS TO THE SECOND ORDER ZERO MOMENTUM VERTEX PART, FOR VARIOUS VALUES OF s_o.

s	s_o=1+i MODULI OF ERRORS	s_o=1.4+i MODULI OF ERRORS	s_o=1.8+i MODULI OF ERRORS	EXACT VALUE OF G(s)
1.00	P.A. $(2\times10^{-5},6\times10^{-5})$ Q.A. $(1\times10^{-6},7\times10^{-6})$	P.A. $(4\times10^{-3},4\times10^{-3})$ Q.A. $(1\times10^{-4},2\times10^{-6})$	P.A. $(2\times10^{-2},1\times10^{-3})$ Q.A. $(6\times10^{-4},5\times10^{-3})$	(1.91613928,0)
1.25	P.A. $(1\times10^{-3},1\times10^{-2})$ Q.A. $(1\times10^{-3},8\times10^{-4})$	P.A. $(1\times10^{-2},1\times10^{-1})$ Q.A. $(3\times10^{-3},1\times10^{-3})$	P.A. $(2\times10^{-1},7\times10^{-2})$ Q.A. $(8\times10^{-3},7\times10^{-2})$	(3.49105878,0)
1.50	P.A. $(7\times10^{-2},4\times10^{-2})$ Q.A. $(2\times10^{-3},1\times10^{-3})$	P.A. $(6\times10^{-2},2\times10^{-2})$ Q.A. $(1\times10^{-4},3\times10^{-3})$	P.A. $(2\times10^{-1},1\times10^{-1})$ Q.A. $(2\times10^{-2},6\times10^{-3})$	(0.26480516,3.30467088)
1.75	P.A. $(3\times10^{-3},7\times10^{-3})$ Q.A. $(7\times10^{-5},4\times10^{-5})$	P.A. $(2\times10^{-3},4\times10^{-3})$ Q.A. $(5\times10^{-5},2\times10^{-4})$	P.A. $(5\times10^{-3},2\times10^{-3})$ Q.A. $(1\times10^{-4},6\times10^{-4})$	(0.04698211,2.17928505
2.00	P.A. $(7\times10^{-4},3\times10^{-3})$ Q.A. $(3\times10^{-5},3\times10^{-5})$	P.A. $(1\times10^{-3},1\times10^{-3})$ Q.A. $(3\times10^{-5},6\times10^{-5})$	P.A. $(1\times10^{-3},3\times10^{-4})$ Q.A. $(1\times10^{-5},3\times10^{-4})$	(-0.07913401,1.76770078)
2.25	P.A. $(3\times10^{-3},2\times10^{-3})$ Q.A. $(1\times10^{-5},1\times10^{-5})$	P.A. $(1\times10^{-3},2\times10^{-4})$ Q.A. $(5\times10^{-5},5\times10^{-6})$	P.A. $(7\times10^{-4},4\times10^{-4})$ Q.A. $(1\times10^{-4},1\times10^{-5})$	(-0.15667384,1.51724664)
2.50	P.A. $(2\times10^{-3},2\times10^{-3})$ Q.A. $(1\times10^{-5},1\times10^{-5})$	P.A. $(1\times10^{-4},8\times10^{-4})$ Q.A. $(1\times10^{-5},7\times10^{-5})$	P.A. $(2\times10^{-4},1\times10^{-4})$ Q.A. $(2\times10^{-5},3\times10^{-4})$	(-0.20613282,1.33882541)
2.75	P.A. $(2\times10^{-3},2\times10^{-3})$ Q.A. $(2\times10^{-5},3\times10^{-5})$	P.A. $(1\times10^{-3},3\times10^{-4})$ Q.A. $(4\times10^{-5},4\times10^{-5})$	P.A. $(1\times10^{-3},5\times10^{-4})$ Q.A. $(4\times10^{-4},1\times10^{-4})$	(-0.23833397,1.20197737)
3.00	P.A. $(1\times10^{-3},3\times10^{-3})$ Q.A. $(4\times10^{-5},1\times10^{-5})$	P.A. $(1\times10^{-3},6\times10^{-4})$ Q.A. $(6\times10^{-5},1\times10^{-5})$	P.A. $(2\times10^{-3},1\times10^{-3})$ Q.A. $(6\times10^{-4},1\times10^{-4})$	(-0.25943861,1.09241191)

(i) $M^2>2m^2$: There is an anomalous threshold on the physical sheet at $s=4m^2(1-r^2)$, where $r=(2m^2-M^2)/2m^2$, lying below the normal threshold at $s=4m^2$.

In Table 6 we compare the (7/7) Padé, (4/4/4) quadratic and (3/3/3/3) cubic approximants to (5a.1), using the series expansion (5a.2). The mass parameters used are M=1 and m=0.6, producing the normal threshold at s=1.44 and an anomalous threshold at s=1.22. Since we have no exact result for (5a.1), the number of figures quoted in the results indicate the agreement between successive approximants. The expansion parameter used is $s_o=1+i$, and for all the results quoted in this (and the following section) the quadrature rule used in (5a.2) had a maximum of 25 integration points in each dimension.

The results show that, compared to the Padé approximants, the quadratic (and cubic) approximants give roughly two more significant figures below threshold; above threshold roughly one more significant figures is obtained.

(ii) $M^2<2m^2$: The anomalous threshold lies on the unphysical sheet reached by passing through the normal threshold branch point and, for $M^2=2m^2$, it emerges through the normal threshold branch point.

As with the self-energy graph of §4 the multi-variable approximants can be used to obtain information on the unphysical sheet, and in particular to trace the path of the anomalous threshold. Tables 7 and 8 illustrate the situation for the mass values M=1 and m=0.9 (still with $s_o=1+i$); this produces the normal threshold at s=3.24 and the anomalous threshold on the unphysical sheet at s=2.77. With reference to Table 7, the most striking feature of the results is the tremendous improvement in convergence obtained with the quadratic and cubic approximants, especially above threshold. Again the number of figures quoted in the results indicate the agreement between successive approximant; for all the examples of this chapter, for which we have no exact result available, we shall adopt this presentation of the results. The results of Table 8 illustrate the nature of the accuracy obtainable on the unphysical sheet; we note that the branch cut starting at $s=4m^2(1-r)$ is placed by the approximants from $-\infty$ to $4m^2(1-r)$. Along the

TABLE 6: (7/7) PADÉ,(4/4/4) QUADRATIC AND (3/3/3/3) CUBIC APPROXIMANTS TO THE TRIANGLE GRAPH, ON THE PHYSICAL SHEET, IN THE PRESENCE OF AN ANOMALOUS THRESHOLD LYING ON THE PHYSICAL SHEET, WITH EXPANSION POINT $s_o = 1+i$.

s	(7/7) PADÉ APPROXIMANT	(4/4/4) QUADRATIC APPROXIMANT	(3/3/3/3) CUBIC APPROXIMANT
0.4	(-3.60755,0.00002)	(-3.607531,-0.0000005)	(-3.60753103,-0.00000009)
0.6	(-4.11176,-0.00007)	(-4.111773,-0.000004)	(-4.1117728,0.0000005)
0.8	(-4.8918,0.0004)	(-4.891841,-0.000008)	(-4.8918399,-0.000002)
1.0	(-6.379,-0.006)	(-6.37889,-0.0001)	(-6.37891,-0.000006)
1.2	(15,-1)	(-13.1,0.10)	(-13.2,0.09)
1.4	(-3.54,-10.26)	(-3.99,-10.8)	(-3.90,-9.57)
1.6	(-1.067,-6.894)	(-1.095,-6.874)	(-1.105,-6.892)
1.8	(-0.586,-5.651)	(-0.5819,-5.6576)	(-0.578,-5.6562)
2.0	(-0.2759,-4.939)	(-0.2743,-4.9350)	(-0.2753,-4.9337)

TABLE 7: (8/8) PADÉ, (5/5/5) QUADRATIC AND (3/3/3/3) CUBIC APPROXIMANTS TO THE TRIANGLE GRAPH, IN THE PHYSICAL REGION, WHEN THE ANOMALOUS THRESHOLD LIES ON THE UNPHYSICAL SHEET, WITH $s_0=1+i$.

s	(8/8) PADÉ APPROXIMANT	(5/5/5) QUADRATIC APPROXIMANT	(3/3/3/3) CUBIC APPROXIMANT
2.3	$(-1.1952361,-0.4\times10^{-10})$*	$(-1.195236126,-0.2\times10^{-12})$*	$(-1.195236126,0.1\times10^{-13})$*
2.5	$(-1.2769717,0.2\times10^{-8})$*	$(-1.276971668,-0.7\times10^{-11})$*	$(-1.276971668,-0.8\times10^{-13})$*
2.7	$(-1.3836861,0.6\times10^{-7})$	$(-1.383686029,-0.2\times10^{-10})$	$(-1.383686029,-0.3\times10^{-11})$*
2.9	$(-1.53588,-0.4\times10^{-6})$	$(-1.535875788,-0.1\times10^{-9})$	$(-1.535875788,-0.4\times10^{-10})$
3.1	$(-1.80144,-0.3\times10^{-3})$	$(-1.80136557,0.4\times10^{-8})$	$(-1.80136557,0.2\times10^{-9})$
3.3	(-2.8,-1)	(-2.39689,-0.64639)	(-2.39688,-0.64639)
3.5	(-2.24,-1)	(-1.976224,-1.120416)	(-1.976225,-1.120416)
3.7	(-1.6,-1.1)	(-1.677079,-1.2830256)	(-1.677079,-1.2830257)
3.9	(-1.6,-1.4)	(-1.4503496,-1.352197)	(-1.4503495,-1.352197)
4.1	(-1.2,-1.5)	(-1.271293,-1.379320)	(-1.2712933,-1.379320)

TABLE 8: QUADRATIC AND CUBIC APPROXIMANTS TO THE TRIANGLE GRAPH, ON THE UNPHYSICAL SHEET, WHEN THE ANOMALOUS THRESHOLD LIES ON THE UNPHYSICAL SHEET; WITH $s_o = 1+i$.

——— INDICATES THE APPROXIMANT HAS NOT CONVERGED.

s	(5/5/5) QUADRATIC APPROXIMANT	(3/3/3/3) CUBIC APPROXIMANT
2.5	———	———
2.7	———	———
2.9	(-7.30,0.004)	(-7.36,-0.01)
3.1	(-4.318,-0.0001)	(-4.319,-0.0003)
3.3	(-2.39689,0.64639)	(-2.39689,0.64639)
3.5	(-1.97622,1.12042)	(-1.97622,1.12041)
3.7	(-1.67701,1.28302	(-1.67708,1.28303)
3.9	(-1.45035,1.35220)	(-1.4503509,1.35220)
4.1	(-1.27129,1.3793236)	(-1.27129,1.37932)

cut generated by the anomalous threshold convergence is not yet attained, in the region between the two branch points three to four significant figures are obtained and along the normal threshold branch cut roughly six significant figures are available.

5b. A THREE POINT PRODUCTION PROCESS

The matrix element (apart from constant factors) for the three point production process under consideration is given in Appendix 2b. In this section we are only concerned with the one variable situation in which t is fixed and an expansion is formed in $(s-s_0)$; the two variable case will be treated in Chapter 4. In the present case, from (A2.11), the matrix element has the expansion

$$M(s-s_0) = \sum_{j=0}^{\infty} (-1)^j (s-s_0)^j \int_0^1 u\,du \int_0^1 \frac{[u^2 v(1-v)]^j}{D_0^{j+1}}$$

where

$$D_0 = u^2 v(1-v)s_0 + uv(1-u)t + u(1-u)(1-v)M^2 - m^2$$

The singularity surface is given by (A2.12) and Fig 4 (of Appendix 2) indicates the singularities lying on the physical sheet. We can again use the multi-valued approximants to simulate the unphysical sheet singularities. For the results quoted in Tables 9,10 and 11 we have used the mass values M=m=1 and expansion point $s_0=1+i$; in Tables 9 and 11 we have, with t=1, the anomalous thresholds at s=0 and s=3 on the unphysical sheet and the normal threshold at s=4 on the physical sheet. We make the following observations:

(i) In Table 9 we again find that the Padé approximants do not converge above threshold (cf. §5a) whereas the quadratic and cubic approximants produce roughly seven significant figures in this region.

(ii) In Table 10 we see that the position of the anomalous threshold on the physical sheet is reasonably well approximated by all the approximation schemes. However, as in (i), the Padé scheme is not converging above threshold in contrast to the quadratic and cubic approximants which converge well in this region.

(iii) In Table 11, with t=1, the unphysical singularities occur at s=0 and s=3 and it is only the singularity at s=3 which is simulated by the

TABLE 9: PADÉ, QUADRATIC AND CUBIC APPROXIMANTS TO THE THREE POINT PRODUCTION PROCESS, ON THE PHYSICAL SHEET, WHEN THE ANOMALOUS THRESHOLDS LIE ON THE UNPHYSICAL SHEETS; WITH $t=1$ and $s_o=1+i$.

s	(8/8) PADÉ APPROXIMANT	(5/5/5) QUADRATIC APPROXIMANT	(3/3/3/3) CUBIC APPROXIMANT
0	$(-0.604599788,-0.2\times10^{-14})^*$	$(-0.604599788,-0.4\times10^{-14})^*$	$(-0.604599788,0.3\times10^{-15})^*$
1	$(-0.6712531057,0.1\times10^{-14})^*$	$(-0.6712531057,-0.9\times10^{-13})^*$	$(-0.6712531057,0.5\times10^{-16})^*$
2	$(-0.768471355,0.4\times10^{-13})^*$	$(-0.768471355,0)^*$	$(-0.768471355,-0.2\times10^{-15})^*$
3	$(-0.93779031,0.8\times10^{-9})^*$	$(-0.9377903074,-0.3\times10^{-12})^*$	$(-0.9377903074,-0.2\times10^{-13})^*$
3.5	$(-1.101658,-0.4\times10^{-6})$	$(-1.101657276,0.9\times10^{-11})^*$	$(-1.101657276,0.9\times10^{-11})^*$
4	$(-1.66,-0.22)$	$(-1.8144,0.0008)$	$(-1.8144,0.0002)$
4.5	$(-1.1,-0.9)$	$(-1.3413969,-0.849204318)$	$(-1.3413971,-0.8492042)$
5	$(-1.1,-1)$	$(-1.0467752,-0.97384596)$	$(-1.0467750,-0.97384605)$
5.5	$(-0.9,-0.9)$	$(-0.843689823,-1.0022043)$	$(-0.84368981,-1.0022041)$
6	$(-0.8,-1)$	$(-0.6949563,-0.99633121)$	$(-0.6949565,-0.99633118)$

TABLE 10: PADÉ, QUADRATIC AND CUBIC APPROXIMANTS TO THE THREE POINT PRODUCTION PROCESS, ON THE PHYSICAL SHEET, WHEN THERE IS AN ANOMALOUS THRESHOLD (AT s=3.66) ON THE PHYSICAL SHEET; WITH t=3.8 AND s_o=1+i.

——— INDICATES THE ENTRY HAS NOT CONVERGED.

s	(8/8) PADÉ APPROXIMANT	(5/5/5) QUADRATIC APPROXIMANT	(3/3/3/3) CUBIC APPROXIMANT
2	$(-1.6376068296, 0.1\times10^{-11})$*	$(-1.6376068296, 0.1\times10^{-11})$*	$(-1.6376068296, -0.1\times10^{-11})$
3	$(-2.4992187, 0.2\times10^{-6})$	$(-2.49921828, 0.4\times10^{-8})$	$(-2.49921837, 0.3\times10^{-7})$
3.5	(-4.320,-0.003)	(-4.3201,-0.0004)	(-4.3209,0.0002)
3.6	(-5.83,-0.06)	(-5.876,-0.01)	(-5.896,0.004)
3.7	(-1.25,-10)	(-5.7,-5.2)	(-6.3,-5.24)
4.1	———	(-0.796,-3.997)	(-0.736,-4.03)
4.6	(———,3.19)	(-3.180,-2.7634)	(-3.21,-2.761)
5.1	(,-2.01)	(-1.238,-2.2728)	(-1.242,-2.2740)
5.6	(———,-2)	(-0.0155,-1.970765)	(-0.0149,-1.9712)
6.1	(———,-1.7)	(0.05222,-1.75765)	(0.0527,-1.7575)

TABLE 11: QUADRATIC AND CUBIC APPROXIMANTS TO THE THREE POINT PRODUCTION PROCESS ON THE PHYSICAL AND UNPHYSICAL SHEETS, WITH t=1; ON THE UNPHYSICAL SHEET THERE IS AN ANOMALOUS THRESHOLD AT s=3.

s	(5/5/5) QUADRATIC APPROXIMANT ON PHYSICAL SHEET	(3/3/3/3) CUBIC APPROXIMANT ON PHYSICAL SHEET	(5/5/5) QUADRATIC APPROXIMANT ON UNPHYSICAL SHEET	(3/3/3/3) CUBIC APPROXIMANT ON UNPHYSICAL SHEET
3	$(-0.9377903074, -0.3\times10^{-12})$	$(-0.9377903074, -0.2\times10^{-13})$	(-12.36,-0.4)	(-14,-2)
3.5	$(-1.101657276, 0.9\times10^{-11})$	$(-1.101657276, -0.1\times10^{-10})$	(-4.3896,-0.003)	(-4.394,-0.002)
4	(-1.8131,0.0008)	(-1.8132,0.0002)	(-1.8145,-0.0008)	(-1.8144,-0.0002)
4.5	(-1.3413969,-0.849204318)	(-1.3413970,-0.8492042)	(-1.3413968,0.849203)	(-1.341398,0.8492033)
5	(-1.0467752,-0.97384596)	(-1.0467751,-0.9738461)	(-1.0467761,0.9738452)	(-1.046787,0.973851)

approximants. As Table 11 indicates, we can obtain remarkably accurate results on the unphysical sheet for $s>3$; in contrast, for $s<3$, convergent results cannot be obtained and for this reason we do not tabulate any results for $s<3$.

We again conclude that useful information on the unphysical sheet can be obtained with the quadratic and cubic approximants.

6. FOURTH ORDER SCALAR BOX GRAPH

The fourth order scalar box graph is a function of the two invariants s and t (see Appendix 3) and, since we are only concerned with one variable approximation schemes in this chapter (two variable schemes being considered in the following chapter), we consider t to be fixed and consider expansions in $(s-s_0)$. From (A3.6) we then obtain the following expansion for the matrix element M:

$$\sum_{j=0}^{\infty}(j+1)(-1)^j(s-s_0)^j\int_0^1 du\int_0^1 dv\int_0^1 dw\ \frac{u(1-u)[u^2v(1-v)]^j}{D_0^{j+1}}$$

where $$D_0(s_0) = u^2v(1-v)s_0+(1-u)^2w(1-w)t-(1-u)^2M^2-um^2$$

Because of the complicated nature of the singularities of this graph (3a) we do not attempt to produce results on the unphysical sheets and we only consider the physical sheet containing the normal threshold at $s=4M^2$.

Some typical results are shown in Tables 12 and 13, where we compare the Padé, quadratic and cubic approximation schemes for $t=0$ and $t=-1$, with expansion point $s_0=3+3i$. The mass values used are $M=m=1$ (producing the normal threshold at $s=4$) and the quadrature rule places a maximum of 25 points in each dimension. The exact values quoted in the tables are derived from (A3.7) and (A3.8); for $t=-1$ no exact values are quoted above threshold and in this case the number of figures quoted for each approximant is based on the values of the next lowest order approximant. Again the most noticeable feature of the results is the tremendous improvement in convergence above threshold obtained with the quadratic and cubic approximants.

We can obtain an idea of the class of functions for which the rotation method of this chapter will be of value by looking at the poles of the Padé

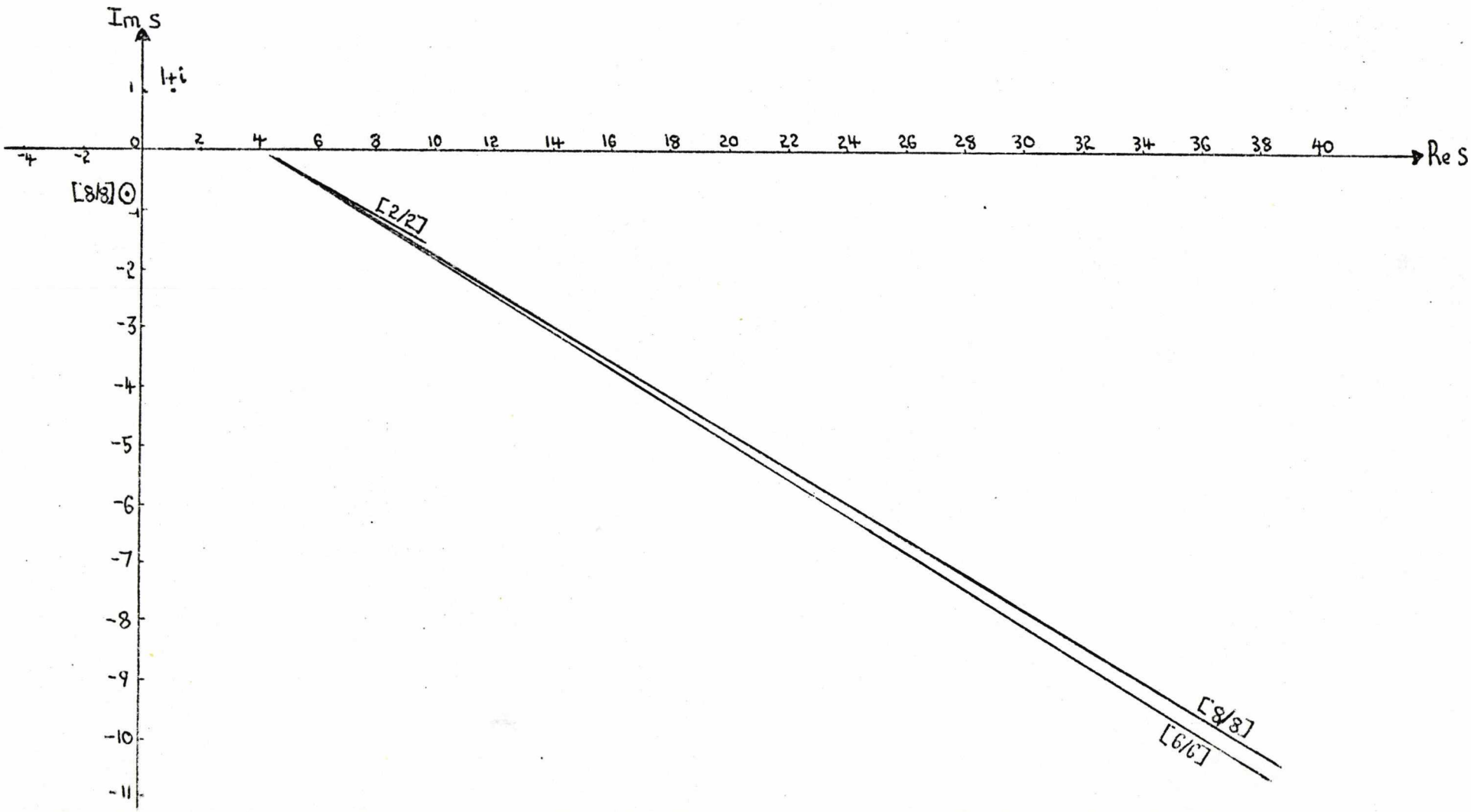
TABLE 12: PADÉ, QUADRATIC AND CUBIC APPROXIMANTS TO THE FOURTH ORDER SCALAR BOX GRAPH WITH t=0 and $s_o=3+3i$.

z	ERROR IN (7/7) PADÉ APPROXIMANT	ERROR IN (4/4/4) QUADRATIC APPROXIMANT	ERROR IN (3/3/3/3) CUBIC APPROXIMANT	EXACT VALUE
0	$(1\times10^{-9},3\times10^{-7})$	$(10^{-9},-8\times10^{-9})$	$(7\times10^{-9},6\times10^{-9})$	(0.263600141,0)
1	$(9\times10^{-8},2\times10^{-7})$	$(4\times10^{-9},2\times10^{-9})$	$(5\times10^{-9},3\times10^{-9})$	(0.3022998940,0)
2	$(4\times10^{-7},-1\times10^{-7})$	$(5\times10^{-9},2\times10^{-9})$	$(4\times10^{-9},4\times10^{-9})$	(0.361596751,0)
3	$(5\times10^{-6},5\times10^{-6})$	$(9\times10^{-9},2\times10^{-8})$	$(2\times10^{-8},-5\times10^{-9})$	(0.472799717,0)
4	(0.16,0.1)	$(6\times10^{-3},8\times10^{-3})$	$(3\times10^{-3},8\times10^{-3})$	(1.2092,0)
5	$(1\times10^{-3},1\times10^{-4})$	$(6\times10^{-7},7\times10^{-7})$	$(1\times10^{-6},5\times10^{-7})$	(0.3893953,0.7024815)
6	$(1\times10^{-4},2\times10^{-4})$	$(4\times10^{-7},1\times10^{-7})$	$(6\times10^{-7},1\times10^{-6})$	(0.1496179,0.6045998)
7	$(1\times10^{-4},1\times10^{-4})$	$(3\times10^{-8},3\times10^{-7})$	$(1\times10^{-6},5\times10^{-9})$	(0.04587218,0.514163791)
8	$(3\times10^{-5},1\times10^{-4})$	$(2\times10^{-6},4\times10^{-8})$	$(8\times10^{-8},1\times10^{-6})$	(-0.00745018,0.44428830)
9	$(1\times10^{-4},6\times10^{-5})$	$(1\times10^{-7},4\times10^{-7})$	$(2\times10^{-6},6\times10^{-7})$	(-0.03758282,0.3902675)

TABLE 13: PADÉ, QUADRATIC AND CUBIC APPROXIMANTS TO THE FOURTH ORDER SCALAR BOX GRAPH WITH $t=-1$ AND $s_o=3+3i$.

s	(7/7) PADÉ APPROXIMANT	(4/4/4) QUADRATIC APPROXIMANT	(3/3/3/3) CUBIC APPROXIMANT	EXACT VALUE
0	$(0.2357113, 0.3\times10^{-8})$	$(0.235711282, -0.2\times10^{-8})$	$(0.235711282, -0.6\times10^{-12})$	$(0.235711282, 0)$
1	$(0.271009560, 0.2\times10^{-9})$	$(0.271009560, 0.2\times10^{-10})$	$(0.271009560, -0.9\times10^{-11})$	$(0.271009560, 0)$
2	$(0.325389313, -0.4\times10^{-10})$	$(0.325389313, 0.2\times10^{-10})$	$(0.325389313, -0.1\times10^{-11})$	$(0.325389313, 0)$
3	$(0.428303305, 0.8\times10^{-10})$	$(0.428303305, -0.3\times10^{-10})$	$(0.428303305, -0.7\times10^{-12})$	$(0.428303305, 0)$
4	(1.03, 0.04)	(1.1382, -0.0003)	(1.1390, -0.0007)	
5	(0.33, 0.647)	(0.32802186, 0.64923075)	(0.3280217, 0.6492306)	
6	(0.116, 0.54)	(0.1143077, 0.5452639)	(0.1143079, 0.545264147)	
7	(0.023, 0.46)	(0.0256293, 0.4582049)	(0.0256287, 0.4582052)	
8	(0.019, 0.389)	(0.0186903, 0.39314875)	(0.0186909, 0.3931477)	
9	(-0.038, 0.34)	(-0.04312853, 0.3437376)	(0.04312767, 0.343736)	

FIG.4: LOCATION OF THE POLES OF THE DIAGONAL PADÉ APPROXIMANTS TO THE FOURTH ORDER SCALAR BOX GRAPH, WITH EXPANSION POINT $s_o = 1+i$.



approximants. Regarding the box graph as a function of s, we do not know, for example, whether it is a series of Stieltjes and so we cannot be certain of the location of the poles of the corresponding Padé approximants (see Theorem 1.3 of Chapter 2). So, after rotation, we cannot be sure that the poles lie in the region indicated by Van Vleck's theorem (and indeed we are not sure that Van Vleck's theorem is valid for the function representing the box graph). We can therefore regard the location of the poles of the (diagonal and paradiagonal) Padé approximants as an indication of the validity of Van Vleck's theorem for functions whose S-fraction expansion is not readily (if at all) computable. In Fig 4 we have plotted the positions of the poles of the diagonal Pade approximants obtained with expansion point $s_o=1+i$. The great majority of the poles lie almost exactly on straight lines, in the lower half s-plane, passing through s_o and the normal threshold branch point, as indicated by Van Vleck's theorem. Certain poles, especially of the higher order approximants, do not conform to this scheme; however, the residue at these poles is extremely small. For example, the residue at the "well-behaved" poles of the (9/9) Padé approximant is of the order of 10^{-2}, whereas for the remaining "badly-behaved" poles the residue is about 10^{-8}. For the box graph this behaviour of the poles occurs for all the values of s_o so far investigated; in practice, however, these anomalous poles do not cause any trouble.

If we consider the multi-valued approximants the situation is rather different. In contrast with Padé approximants, the approximants do not need to simulate the branch cut by a sequence of poles; in general, the poles of these multi-valued approximants are placed on one of the 'unphysical' sheets of the approximant (as explained in § 2 of Chapter 2) and hence cause no difficulties. This behaviour is clearly a desirable feature for the type of calculations discussed in this chapter.

7. FOURTH ORDER SCALAR SELF ENERGY GRAPH

From (A4.1) (of Appendix 4) we obtain the following expansion for the fourth order scalar self energy graph:

TABLE 14: PADÉ, QUADRATIC AND CUBIC AND LOGARITHMIC QUADRATIC APPROXIMANTS TO THE FOURTH ORDER SCALAR SELF ENERGY GRAPH, WITH $s_o=3+3i$.

z	(8/8) PADÉ APPROXIMANT	(5/5/5) QUADRATIC APPROXIMANT	(4/4/4/4) CUBIC APPROXIMANT	(5/5/5) LOGARITHMIC QUADRATIC APPROXIMANT
0	$(-0.781284, 0.5\times10^{-5})$	$(-0.781275, 0.3\times10^{-4})$	$(-0.781263, 0.2\times10^{-4})$	$(-0.781270, 0.2\times10^{-4})$
1	$(-0.92362, 0.1\times10^{-4})$	$(-0.92364, 0.6\times10^{-5})$	$(-0.923637, 0.2\times10^{-4})$	$(-0.9236341, 0.2\times10^{-4})$
2	$(-1.146811, 0.1\times10^{-4})$	$(-1.146793, 0.3\times10^{-5})$	$(-1.1467992, 0.5\times10^{-5})$	$(-1.1468017, 0.8\times10^{-6})$
3	$(-1.5756, 0.1\times10^{-4})$	$(-1.5755, 0.3\times10^{-4})$	$(-1.57553, 0.1\times10^{-4})$	$(-1.57548, 0.2\times10^{-4})$
3.5	$(-2.0315, 0.6\times10^{-4})$	$(-2.0328, 0.1\times10^{-3})$	$(-2.0325, 0.4\times10^{-3})$	$(-2.0331, 0.3\times10^{-3})$
4	(-3.8, 0.39)	(-3.967, -0.27)	(-4.0, -0.002)	(-3.81, -0.20)
4.5	(-2.38, -2.779)	(-2.3877, -2.8040)	(-2.3899, -2.8080)	(-2.3889, -2.8338)
5	(-1.247, -2.949)	(-1.2506, -2.9435)	(-1.2501, -2.9419)	(-1.2577, -2.9313)
6	(-0.085, -2.521)	(-0.0832, -2.5194)	(-0.08214, -2.5203)	(-0.076, -2.5188)
7	(0.433, -2.025)	(0.4317, -2.0260)	(+0.4310, -2.0241)	(+0.428, -2022)
8	(0.662, -1.6091)	(0.6640, -1.6050)	(0.6609, -1.6064)	(0.6607, -1.607)
9	(0.750, -1.274)	(0.7494, -1.2710)	(0.7492, -1.2757)	(0.7488, -1.2748)
10	(0.76, -1.017)	(0.759, -1.017)	(0.7633, -1.021)	(0.761, -1.021)
15	(0.56, -0.42)	(0.566, -0.4280)	(0.569, -0.420)	(0.577, -0.420)

$$\sum_{r=0}^{\infty}(-1)^r(s-s_c)^r\int_0^1 du\int_0^1 dv\int_0^1 dw\int_0^1 dz\ \frac{u^2v(1-u)\sum^r}{C\,D_o^{r+1}} \qquad (7.1)$$

where $\sum$, C and $D_o=D(s=s_o)$ are functions of the integration variables u,v,w and z and are explicitly defined in Appendix 4. As in the case of the box graph of 6, we confine our attention to the physical sheet containing the normal threshold.

The results of Table 14 correspond to the mass values M=m=1, so that the normal threshold is at s=4; the quadrature rule used places a maximum of 23 points in each dimension. We have compared the Padé, quadratic, cubic and also the logarithmic quadratic approximants (see (2.8) of Chapter I); the corresponding logarithmic Padé approximants have also been used but these produce relatively poor results and we do not tabulate them here. The basic trend of the results is as before, with the multi-valued approximants producing more accurate results than the single-valued Padé approximants. A comparison with the previous tables shows that the results obtained for this fourth order graph are noticeably less accurate than the results for the diagrams previously considered. This loss of accuracy occurs because the power series coefficients (given by the four dimensional integrals of (7.1)) are determined considerably less accurately than for the previous graphs. Improved results could possibly be obtained by employing more sophisticated integration routines (see § 8).

8. CONCLUSIONS

From the examples considered in this chapter we draw the following conclusions:

(i) The proposed method of calculation (described in § 3) of Feynman matrix elements in the physical region certainly appears to be a feasible method of calculation. For higher order graphs certain difficulties in performing the numerical integration over the Feynman parameters are encountered, as exemplified by the self energy graph of § 7. Also, for the box graph of § 6, difficulties arise in integrations when the external particles are off the mass shell and it has not yet been possible to reproduce the anomalous thresholds which arise in this case (11). However,

in such instances it is not the basic idea of the method which is at fault but rather that the numerical integration routines used are not sufficiently reliable. It may be that adaptive integration routines (12), where the assignment of quadrature points is based upon the behaviour of the integrand, will prove to be of use in this context. Another possibility is to integrate directly over the simplex, instead of introducing the Feynman parameters to convert the problem to one over the pypercube (and then using the standard Gauss-Legendre quadrature). In this context, the integration formulae discussed in (13) have been applied to the triangle graph of §5a, though at present without any real success.

(ii) The use of multi-valued approximants is to be recommended. Naively, we would expect to be able to approximate a function with a branch point by an approximant which itself possesses a branch point. We can interpret the results of this chapter with this point of view. In Appendix 5 the nature of the leading singularities of the Feynman graphs discussed in this chapter are given. In particular, the fourth order box graph of §6 possesses a leading square root singularity and, in fact, it was this observation (14) which produced the original use of quadratic approximants in this context. For this graph the quadratic approximants (especially above the normal threshold) produce dramatically better results than the Padé approximants; essentially, we can regard the quadratic approximants as approximating a square root branch point by a square root branch point. Continuing with this line of reasoning, we would expect the logarithmic Padé and quadratic approximants to work well for the triangle graph, whose leading singularity is logarithmic. However, the numerical results obtained with these approximation schemes are disappointing; the logarithmic Padé approximant is not noticeably better than the ordinary Padé approximant and the logarithmic quadratic approximant is definitely worse than the ordinary quadratic approximant. The explanation of this behaviour is not clear and it appears that more numerical study of these approximation schemes is required.

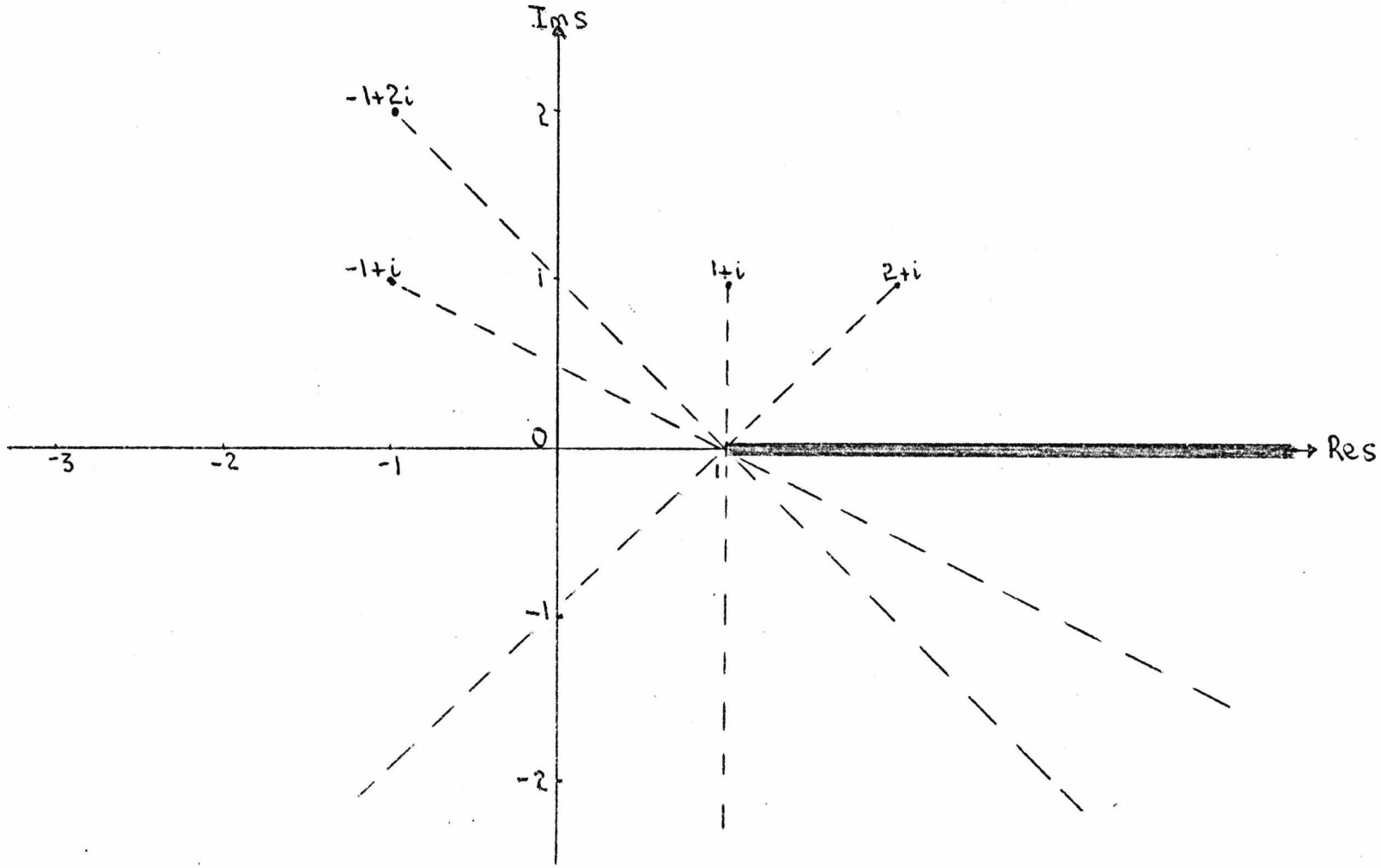
(iii) A feature of the methods discussed in this chapter is the arbitrary

nature of the expansion parameters s_o, which allows a consistency check to be made on the results. The question of the optimal value of this parameter then naturally arises; here we attempt only to give a qualitative discussion of the problem. In Fig 5 we illustrate the expected rotation of the branch cut, for various values of s_o, of, for example, the function $f(s) = \ln(1-s)$. If we wish to calculate values of $f(s)$ both for $s<1$ and $s>1$, then we would expect the optimum value of s_o to be directly above the branch point $s=1$; in practice, $s_o=1+i$ seems to provide an optimum (or at least a near optimum) choice. If, however, we wish only to calculate $f(s)$ for $s>1$ then an optimum value of s_o with $\mathrm{Re}\, s_o > 1$ would seem more reasonable (although the actual value of s_o is not at all obvious). Thus, even with the 'simple' function $\ln(1-s)$, the optimum value of s_o is dependent upon exactly where we are interested in approximating the function.

For the Feynman diagrams the situation is very much more complicated. Now the value of s_o not only determines the angle through which the branch cut is rotated, but also determines the accuracy to which the multi-dimensional integrals can be evaluated. However, the exact relation between the accuracy obtainable and the value of s_o is not at all clear. In practice, if we wish to evaluate the Feynman matrix element both above and below the normal threshold (at $s=s_n$, say), we have found that choosing $\mathrm{Re}\, s_o$ slightly less than s_n (and $\mathrm{Im}\, s_o \approx \mathrm{Re}\, s_o$) generally produces reasonable good (though undoubtedly not optimum) results. For example, for the box graph of 6 with $s_n=4$, we choose $s_o=3+3i$ and Tables 12 and 13 indicate the type of results this choice produces. The actual determination of an optimum value for s_o is clearly a very difficult problem.

(iv) A priori, the only information we have about the particular graph under discussion is that the imaginary part of the matrix element is zero along the real axis below the normal threshold (or perhaps below an anomalous threshold). It should be possible to incorporate this information into the method of calculation. For example, we can modify the equations defining the Padé approximant so as to make the imaginary part of the approximant zero at certain pre-selected values below threshold. However,

FIG. 5 EXPECTED ROTATION OF THE BRANCH CUT OF ln(1-s) FOR VARIOUS VALUES OF s_o.



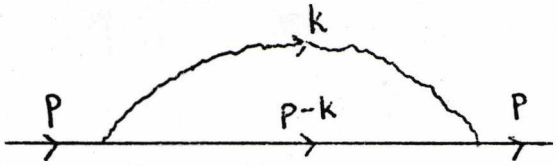
in practice, such alterations seem to destroy the homographic invariance properties of the approximants and lead to relatively poor results. Nevertheless, it does seem desirable to incorporate any available information into the computational scheme.

(v) The graphs of sections 4 and 5 illustrate the very powerful method of analytic continuation exhibited by the quadratic (and cubic) approximants; namely, the continuation of a function from one Riemann sheet to another. In this connection we expect the multi-valued approximants considered here to prove very useful in the many areas of mathematics where multi-valued functions commonly arise.

APPENDIX I: SECOND ORDER RENORMALIZED SCALAR SELF ENERGY

(a) EXPRESSION FOR THE MATRIX ELEMENT

FIG I.



The Feynman graph (5) for this process is given in Fig.I, the incoming particle (of mass m) having four-momentum p and the emitted particle (of mass μ) four-momentum k. The matrix element for this process is given by the usual Feynman rules (5) and is proportional to

$$f(s) = \int_{-\infty}^{\infty} \frac{d^4k}{(k^2-\mu^2+i\epsilon)((p-k)^2-m^2+i\epsilon)} \qquad (A1.1)$$

To the mass of each particle occuring in an internal line of the graph a small negative imaginary part is added; this is the Feynman prescription (5) for obtaining physical amplitudes. It is equivalent to the limit obtained by approaching the branch cut of F(s) (see below for the definition of F(s)) from the upper half s-plane; the Feynman prescription allows (A1.1) to be evaluated with real external four-momenta (that is, with s real).

To evaluate (A1.1) we combine the denominators using the Feynman formula (5)

$$\frac{1}{ab} = \int_0^1 \frac{d\alpha}{[\alpha a + (1-\alpha)b]^2}$$

to give

$$f(s) = f(p^2) = \int_0^1 d\alpha \left[k^2 - 2kp\alpha \quad (p^2-m^2)\alpha - (1-\alpha)\mu^2 - i\epsilon\right]^{-2}$$

We diagonalize the denominator by introducing the variable k′:

$$k' = k - \alpha p$$

to give

$$f(p^2) = \int_0^1 d\alpha \int_{-\infty}^{\infty} d^4k' \left(k'^2 - L + i\epsilon\right)^{-2} \qquad (A1.2)$$

where

$$L = p^2(\alpha^2-\alpha) + \mu^2 + (m^2-\mu^2)\alpha$$

To evaluate (A1.2) consider

$$J = \int_{-\infty}^{\infty} d^4k \, (k^2 - L + i\epsilon)^{-3}$$

The path of the k_o-integration is along the real k_o axis, the poles having been displaced by the Feynman prescription. We can rotate the integration path by $\pi/2$ in the complex plane so the integration is along the imaginary axis from $-i\infty$ to $+i\infty$. This procedure is permissible since, during the rotation, we never cross any singularities (these being located above the negative real axis and below the positive real axis) (6). We therefore introduce the integration variable k_4:

$$k_o = ik_4$$

Then (using the metric of (6))

$$k^2 = k_o^2 - \underline{k}^2 = -(k_4^2 + \underline{k}^2)$$

so that the integral now becomes one over Euclidean four-space:

$$J = -i \int_{-\infty}^{\infty} \frac{dk_4 \, d^3k}{(k_4^2 + k_1^2 + k_2^2 + k_3^2 + L)^3}$$

Introducing four-dimensional spherical polar co-ordinates, we find (6)

$$J = -2\pi^2 i \int_0^{\infty} \frac{x \, dx}{(x+L)^3} = -\frac{i\pi^2}{2L} \qquad \text{(A1.3)}$$

Since f(s) is logarithmically divergent we must consider the renormalized second order (scalar) self energy F(s), defined by

$$F(s) = f(s) - f(m^2)$$

$$= \int_0^1 d\alpha \int_{-\infty}^{\infty} d^4k' \left\{ \frac{1}{(k'^2 - L + i\epsilon)^2} - \frac{1}{(k'^2 - L_o + i\epsilon)^2} \right\}$$

where

$$L_o = L(p^2 = m^2) = \mu^2(1-\alpha) + m^2\alpha^2$$

Noting that

$$F(s) = -2 \int_0^1 d\alpha \int_{L_o}^{L} J \, dL$$

and using (A1.3) we obtain

$$F(s) = i \int_0^1 d\alpha \, \ln \left[\frac{\mu^2(1-\alpha) + m^2\alpha - \alpha(1-\alpha)s}{\mu^2(1-\alpha) + m^2\alpha^2} \right] \qquad \text{(A1.4)}$$

(b) SINGULARITY STRUCTURE OF F(s)

Although we can determine the singularity structure of F(s) directly from (A1.4) we shall find it more convenient to employ the Landau-Cutkosky

rules ((7) and (8)); these rules are very useful for determining the singularities of a graph (especially for complicated graphs), and provide a simple method for deciding which of the singularities lie on the physical sheet.

To define these rules we write the Feynman integral (after introducing Feynman parameters α_i) in the form

$$I = \int_0^1 \prod_{i=1}^{r} d\alpha_i \int_{-\infty}^{\infty} \prod_{j=1}^{L} d^4 q_j \frac{\delta\left(1 - \sum_{i=1}^{r} \alpha_i\right)}{\left[\sum_{i=1}^{r} \alpha_i (k_i^2 - m_i^2)\right]^r} \tag{A1.5}$$

where L and r are defined in §1, and the q_j are the independent "loop momenta". The k_i are internal momenta, and are determined by the momentum conservation delta functions as linear functions (with coefficients +1, -1 or 0) of the q_j and the external momenta (and the $i\epsilon$ has been absorbed into the m_i). The Landau-Cutkosky rules then give the following conditions for I to have a singularity:

$$\text{(i)}\quad k_i^2 = m_i^2 \quad (i = 1, \ldots, r)$$

and

$$\text{(ii)}\quad \sum \pm \alpha_i k_i = 0 \quad (i = 1, \ldots, r)$$

for each independent loop of the diagram. The sign is plus or minus according to whether the direction of k_i is with or against a chosen direction around the loop. The leading singularity of the graph (that is, a singularity not shared by the contracted graphs) is then given by a solution of (i) and (ii) with $\alpha_i \neq 0$ for all i; furthermore, the singularity lies on the physical sheet provided $\alpha_i > 0$ for all i (3a).

For the graph of Fig I, writing $k_1 = k$ and $k_2 = p - k$, conditions (i) and (ii) give

$$\begin{aligned} k_1^2 &= \mu^2 \\ k_2^2 &= m^2 \end{aligned} \tag{A1.6a}$$

and

$$\alpha_1 k_1 - \alpha_2 k_2 = 0 \tag{A1.6b}$$

Multiplying (A1.6b) in turn by k_1 and k_2 we obtain the system of equations

$$\begin{aligned} \alpha_1 k_1^2 - \alpha_2 k_1 k_2 &= 0 \\ \alpha_1 k_1 k_2 - \alpha_2 k_2^2 &= 0 \end{aligned} \tag{A1.7}$$

which have a non-trivial solution if and only if

$$k_1^2 k_2^2 - (k_1 k_2)^2 = 0$$

Using the relation

$$s = p^2 = (k_1 + k_2)^2 = m^2 + \mu^2 + 2k_1 k_2$$

(obtained from (A1.6a)) this condition reduces to

$$s = (m \pm \mu)^2 \qquad \text{(A1.8)}$$

To determine which of the singularities given by (A1.8) lie on the physical sheet, we have from (A1.7),

$$2\alpha_1 \mu^2 = \alpha_2 \left(s - (m^2 + \mu^2)\right)$$

For $s=(m-\mu)^2$ this gives

$$2\mu^2 \alpha_1 = -2m\mu\alpha_2 ,$$

so that α_1 and α_2 cannot both be positive (as they can for $s=(m+\mu)^2$). Hence, only the singularity at $s=(m+\mu)^2$ lies on the physical sheet.

Finally, we note that, in order to obtain all the singularities of a Feynman integral, we must permit 'solutions' of the Landau-Cutkosky equations in which the momentum vectors may become infinite (8). Such solutions give rise to "second-type singularities", and it can be shown (3a) that the second order self energy graph has such a singularity at s=0.

APPENDIX 2: THREE POINT FUNCTIONS (9)

2a THE TRIANGLE GRAPH AND ANOMALOUS THRESHOLDS

(i) FEYNMAN REPRESENTATION

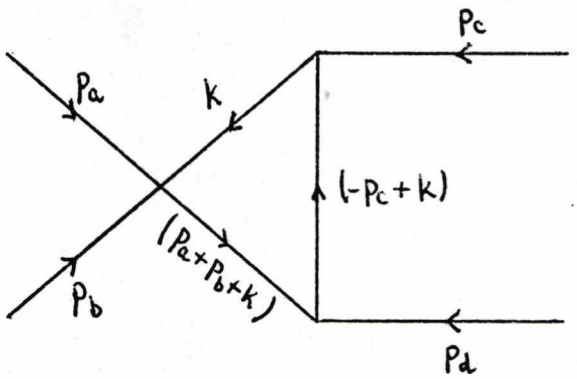


FIG.I

The triangle graph (or three point function) matrix element is, from Fig. I, proportional to

$$I = \int_{-\infty}^{\infty} \frac{d^4k}{(k^2-m^2)\left[(p_a+p_b+k)^2-m^2\right]\left[(-p_c+k)^2-m^2\right]} \tag{A2.1}$$

where p_a, p_b, p_c and p_d are the external four-momenta and k is the internal loop momenta. For simplicity, we take all internal masses to be m and all external masses to be M and we incorporated all $i\epsilon$ factors into the mass factors; again, we only consider scalar particles. We therefore have

$$p_a^2 = p_b^2 = p_c^2 = p_d^2 = M^2 \tag{A2.2}$$

$$p_a + p_b + p_c + p_d = 0$$

and

$$S = (p_a+p_b)^2$$

To evaluate I we proceed as in Appendix I; we combine the denominators in (A2.1) using the Feynman formula

$$\prod_{i=1}^{3} A_i^{-1} = 2! \int_0^1 \prod_{i=1}^{3} d\alpha_i\, \delta\left(1 - \sum_{i=1}^{3} \alpha_i\right) S^{-3}$$

where

$$S = \sum_{i=1}^{3} \alpha_i A_i$$

and then diagonalize the resulting denominator with the substitution

$$k' = k + (p_a+p_b)\alpha_2 - p_c\alpha_3$$

Using (A2.2) and (A1.3) we arrive at the result

$$I = i\pi^2 \int_0^1 \prod_{i=1}^{3} d\alpha_i\, \delta\left(\sum_{i=1}^{3} \alpha_i - 1\right) D^{-1} \tag{A2.3}$$

where

$$D(s) = \alpha_1\alpha_2 s + \alpha_3(\alpha_1+\alpha_2)M^2 - m^2$$

We remove the δ-function from (A2.3) by introducing Feynman parameters (5) u and v, where

$$\alpha_1 = uv$$
$$\alpha_2 = u(1-v)$$

and

$$\alpha_3 = 1-u$$

Then I is proportional to

$$\int_0^1 u\,du \int_0^1 \frac{dv}{u^2 v(1-v)s + u(1-u)M^2 - m^2} \qquad \text{(A2.4)}$$

(ii) SINGULARITY STRUCTURE

With the notation

$$k_1 = -k,\ k_2 = p_c - k,\ k_3 = -(p_a+p_b+k) \qquad \text{(A2.5a)}$$

the Landau-Cutkosky rules, applied to Fig I, give

$$k_1^2 = k_2^2 = k_3^2 = m^2 \qquad \text{(A2.5b)}$$

and

$$\alpha_1 k_1 + \alpha_2 k_2 + \alpha_3 k_3 = 0 \qquad \text{(A2.5c)}$$

Multiplying (A2.5c) in turn by k_1, k_2 and k_3 we obtain the system

$$\begin{aligned} \alpha_1 m^2 + \alpha_2 k_1k_2 + \alpha_3 k_1k_3 &= 0 \\ \alpha_1 k_1k_2 + \alpha_2 m^2 + \alpha_3 k_2k_3 &= 0 \\ \alpha_1 k_1k_3 + \alpha_2 k_2k_3 + \alpha_3 m^2 &= 0 \end{aligned} \qquad \text{(A2.6)}$$

which has a non-trivial solution if and only if

$$\det \begin{vmatrix} m^2 & k_1k_2 & k_1k_3 \\ k_1k_2 & m^2 & k_2k_3 \\ k_1k_3 & k_2k_3 & m^2 \end{vmatrix} = 0 \qquad \text{(A2.7)}$$

Using the momentum conservation conditions

$$p_a + p_b + k_3 = k_1$$

$$p_d + k_2 = k_3$$

and

$$p_c + k_1 = k_2$$

(A2.7) reduces to

$$\det \begin{vmatrix} 1 & r & -\frac{s}{2m^2} \\ r & 1 & r \\ -\frac{s}{2m^2} & r & 1 \end{vmatrix} = 0 \qquad \text{(A2.8)}$$

where

$$r = \frac{2m^2 - M^2}{2m^2}$$

The stability of the particles requires that

$$M^2 < 4m^2$$

so that

$$-1 < r < 1$$

On expansion, (A2.8) reduces to the two sclutions

$$s = 0 \qquad \text{(A2.9a)}$$

and

$$s = 4m^2(1 - r^2) \qquad \text{(A2.9b)}$$

From (A2.6), the solution s=0 leads to

$$\alpha_1 + \alpha_3 = 0$$

so that both α_1 and α_3 cannot be positive; the s=0 singularity thus does not lie on the physical sheet. The solution (A2.9b) leads, from (A2.6), to

$$2r\alpha_1 = -\alpha_2 \quad , \quad \alpha_1 = \alpha_3$$

and so the α's are all positive if and only if

$$r < 0$$

which is equivalent to

$$M^2 > 2m^2 \qquad \text{(A2.10)}$$

Thus, when (A2.10) holds, we have an additional singularity on the physical sheet lying on the real axis below the normal threshold at $4m^2$. This is referred to as an anomalous threshold since the position of the singularity does not correspond to the mass of any physical intermediate state.

We can follow the path of the anomalous threshold singularity as the external mass M is varied; when $M^2<2m^2$ the anomalous threshold lies on the unphysical sheet reached by going through the normal threshold cut and it emerges through the normal threshold branch point at $M^2=2m^2$. This situation is shown in Fig 3, where the path on the unphysical sheet is indicated by a

FIG.2: THE PATH OF THE ANOMALOUS THRESHOLD OF THE TRIANGLE GRAPH AS THE EXTERNAL MASS INCREASES THROUGH $\sqrt{2m^2}$.

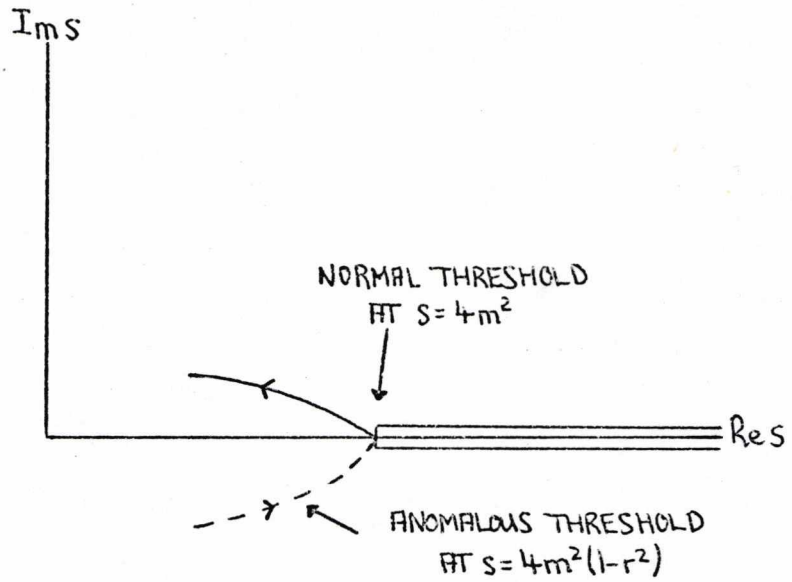
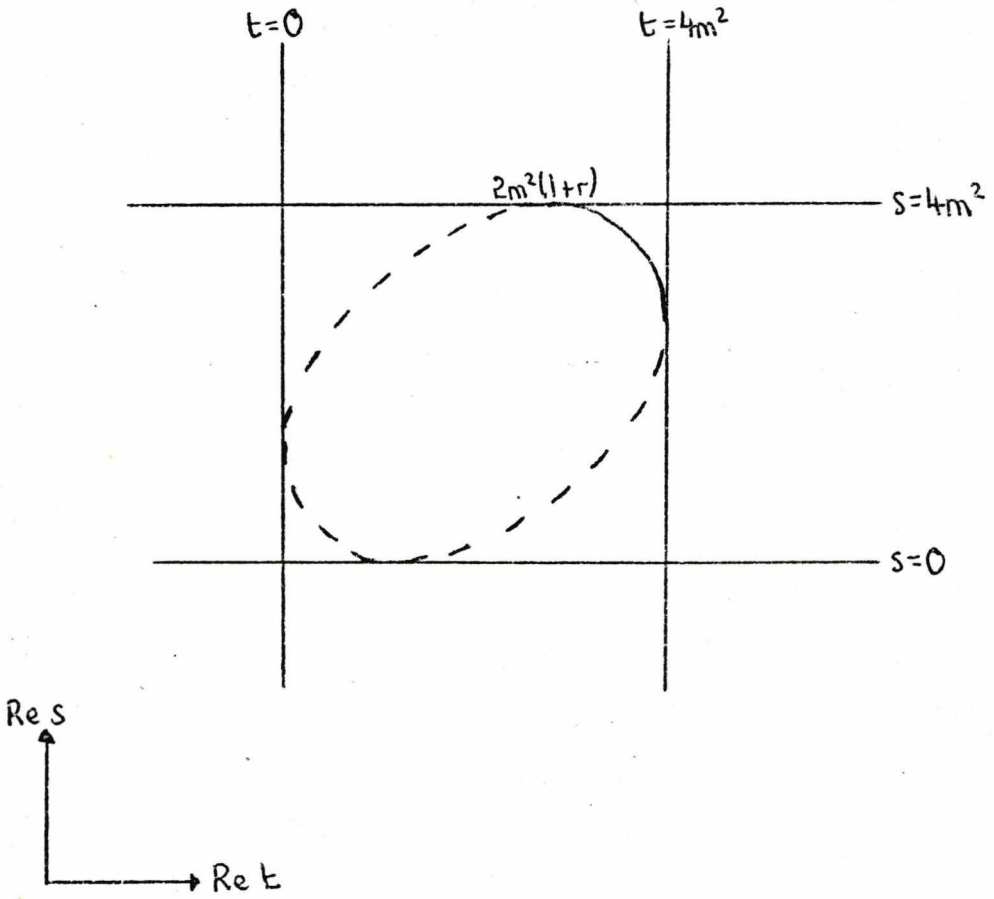


FIG.4: THE REAL SECTION OF THE LEADING LANDAU CURVE FOR THE THREE POINT PRODUCTION PROCESS. THE UNBROKEN LINE SHOWS THE SINGULARITIES LYING ON THE PHYSICAL SHEET.



broken line and the anomalous threshold is displaced off the axis for clarity.

2b. A THREE POINT PRODUCTION PROCESS

(i) FEYNMAN REPRESENTATION

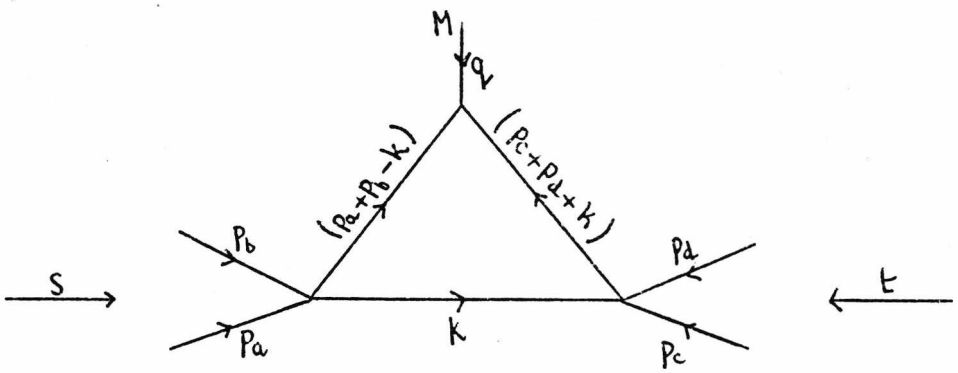


FIG.2

The Feynman diagram for the process is shown in Fig 2, and the matrix element is proportional to

$$I = \int_{-\infty}^{\infty} \frac{d^4k}{(k^2-m^2)\left[(p_a+p_b-k)^2-m^2\right]\left[(p_c+p_d+k)^2-m^2\right]}$$

I can be evaluated using the methods of Appendices I and 2a, and we obtain

$$I = \int_0^1 u\,du \int_0^1 \frac{dv}{u^2v(1-v)s + uv(1-u)t + u(1-u)(1-v)M^2 - m^2} \qquad \text{(A2.11)}$$

where

$$s = (p_a+p_b)^2$$

$$t = (p_c+p_d)^2$$

and the Feynman parameters u and v are introduced according to

$$\alpha_1 = uv$$

$$\alpha_2 = u(1-v)$$

and

$$\alpha_3 = 1-u$$

(ii) SINGULARITY STRUCTURE

For this process the equation corresponding to (A2.8) is

$$\det \begin{vmatrix} 1 & -\frac{t}{2m^2} & -\frac{s}{2m^2} \\ -\frac{t}{2m^2} & 1 & r \\ -\frac{s}{2m^2} & r & 1 \end{vmatrix} = 0 \qquad \text{(A2.12)}$$

which, provided r<1, is the equation of an ellipse and is illustrated in

FIG.5: THE PATH OF THE ANOMALOUS THRESHOLD OF FIG.3 AS t IS INCREASED THROUGH REAL VALUES.

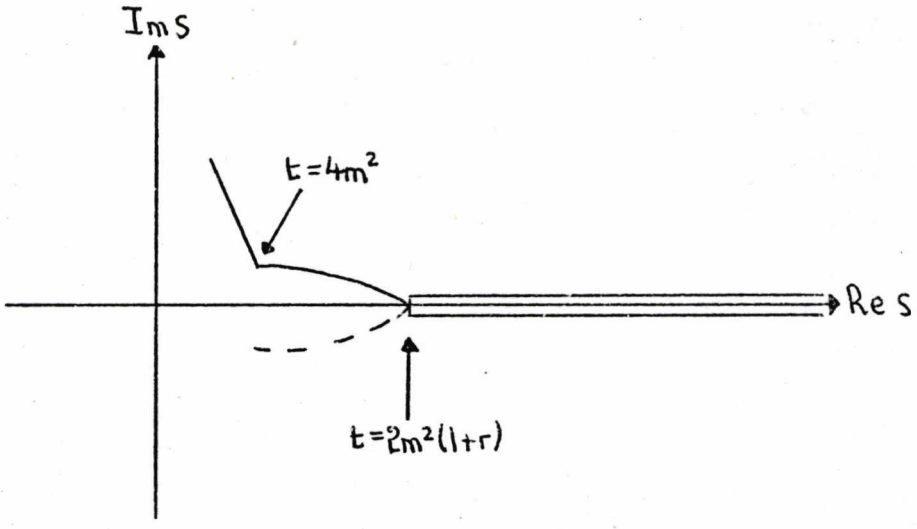


Fig 4. By using the criterion that all the α's be positive, we can show that only the part of the ellipse shown by a full line in Fig 4 is singular on the physical sheet.

The path of the anomalous threshold as t is increased through real values is clear from Fig.4, and is shown in Fig.5. The part of the Landau-Cutkosky curve which goes into the complex plane and is attached to the singular arc of Fig.4 gives rise to a complex singularity for real positive $t>4m^2$.

APPENDIX 3: FOURTH ORDER SCALAR BOX GRAPH

(i) FEYNMAN REPRESENTATION

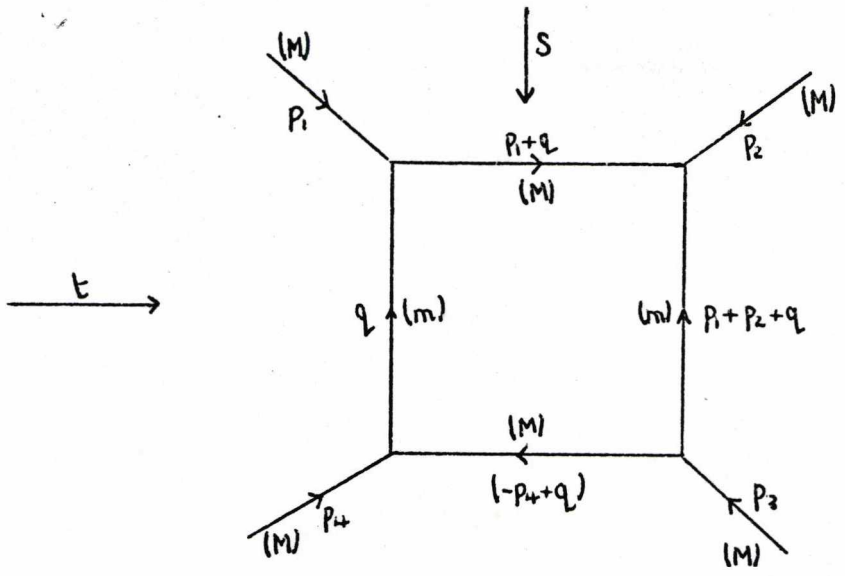


FIG.I

The Feynman diagram for the fourth order box graph is given in Fig.I; the mass of the particle associated with each line is given in brackets next to the line. For the external particles on the mass shell,

$$p_i^2 = M^2 \quad (i=1,\ldots,4)$$

and we also have

$$\sum_{i=1}^{4} p_i = 0$$
$$S = (p_1+p_2)^2 \qquad \text{(A3.1)}$$

and

$$t = (p_1+p_4)^2$$

The matrix element is proportional to

$$I = \int_{-\infty}^{\infty} \frac{d^4q}{(q^2-m^2)\left[(p_1+q)^2-M^2\right]\left[(p_1+p_2+q)^2-m^2\right]\left[(-p_4+q)^2-M^2\right]} \qquad \text{(A3.2)}$$

The process of evaluation of (A3.2) by diagonalization of the denominator and subsequent use of the Feynman formulae to perform the integrations of the loop momenta, as described in Appendices 1 and 2, can be formulated for an arbitrary Feynman graph (2). For a graph with L internal (loop) momenta, denoted by $k_1,\ldots,k_L$ and r internal lines, the integration over the internal momenta gives

$$I = (i\pi^2)^L \frac{(r-2L-1)!}{(r-1)!} \int_0^1 \prod_{i=1}^{r} d\alpha_i\, \delta\left(1-\sum_{i=1}^{r}\alpha_i\right) \frac{C^{r-2L-2}}{D^{r-2L}} \qquad \text{(A3.3)}$$

where

$$Q(k_1,\ldots,k_L)=\sum A_{st}k_sk_t+\sum A_sk_s+\tilde{C} \tag{A3.4a}$$

is the denominator obtained by combining the propagators, and

$$C=\det\begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1L}\\ A_{21} & A_{22} & \cdots & A_{2L}\\ \vdots & \vdots & & \vdots\\ A_{L1} & A_{L2} & \cdots & A_{LL}\end{vmatrix} \tag{A3.4b}$$

and

$$D=\det\begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1L} & A_1\\ A_{21} & A_{22} & \cdots & A_{2L} & A_2\\ \vdots & \vdots & & \vdots & \vdots\\ A_{L1} & A_{L2} & \cdots & A_{LL} & A_L\\ A_1 & A_2 & \cdots & A_L & \tilde{C}\end{vmatrix} \tag{A3.4c}$$

For the box graph r=4 and L=1 so that, neglecting constant factors, (A3.3) gives

$$I=\int_0^1\prod_{i=1}^{4}d\alpha_i\,\delta\Big(1-\sum_{i=1}^{4}\alpha_i\Big)D^{-2} \tag{A3.5a}$$

where

$$\begin{aligned}Q(q)&=\alpha_1(q^2-m^2)+\alpha_2\left[(p_1+q)^2-M^2\right]+\alpha_3\left[(p_1+p_2+q)^2-m^2\right]+\alpha_4\left[(-p_4+q)^2-M^2\right]\\&=q^2-2q\left[\alpha_2p_1+\alpha_3(p_1+p_2)-\alpha_4p_4\right]+\left[-\alpha_1m^2+\alpha_3(s-m^2)\right]\end{aligned}$$

using (A3.1). Thus, from (A3.4c) and (A3.1),

$$\begin{aligned}D&=\det\begin{vmatrix} 1 & \alpha_2p_1+\alpha_3(p_1+p_2)-\alpha_4p_4\\ \alpha_2p_1+\alpha_3(p_1+p_2)-\alpha_4p_4 & \alpha_1m^2+\alpha_3(s-m^2)\end{vmatrix}\\&=\alpha_1\alpha_3s+\alpha_2\alpha_4t+M^2(\alpha_2+\alpha_4)^2-(\alpha_1+\alpha_3)m^2\end{aligned} \tag{A3.5b}$$

Introducing Feynman parameters u,v and w by

$$\alpha_1=uv$$

$$\alpha_3=u(1-v)$$

$$\alpha_2=(1-u)w$$

and

$$\alpha_4=(1-u)(1-w)$$

we obtain, from (A3.5a) and (A3.5b),

$$I=\int_0^1 u(1-u)du\int_0^1 dv\int_0^1 \frac{dw}{\left[u^2v(1-v)s-(1-u)^2w(1-w)t-(1-u)^2M^2-um^2\right]^2} \qquad (A3.6)$$

The singularity structure of the box graph is complicated and details can be found in (3a).

(ii) DISPERSION RELATION FOR BOX GRAPH

For $s>4M^2$, the following dispersion relation (due to Mandelstam (10)) is valid for the box graph of Fig.I (11):

$$T_B(s,t)=\int_{4M^2}^{\infty}\frac{ds'}{(s'-s)}\int_{4m^2}^{\infty}\frac{dt'}{(t'-t)}\;\frac{\theta(K)}{[s't'K]^{1/2}} \qquad (A3.7)$$

where

$$K(s',t')=(s'-4M^2)(t'-4m^2)-4m^4$$

and

$$\theta(x)=\begin{cases}1 & ,x>0\\ 0 & ,x<0\end{cases}$$

For the special case of forward scattering we (11) can obtain an explicit representation for T_B. In this case

$$\mathrm{Im}\,T_B(s,t=0)=\frac{\pi}{2}\sqrt{\frac{s-4M^2}{s}}\cdot\frac{1}{m^2(s+m^2-2M^2)}$$

and we can evaluate the single variable dispersion relation (t fixed)

$$T_B(s,t=0)=\frac{1}{\pi}\int_{4M^2}^{\infty}\frac{ds'}{(s'-s)}\,\mathrm{Im}\,T_B(s',t'=0)$$

to give, for $s>4M^2$,

$$T_B(s,t=0)=\frac{1}{(m^2s+\Delta)}\left\{i\frac{A\pi}{2}+\frac{\tan^{-1}B}{B}-A\ln\left(\sqrt{\frac{s}{4M^2}}+\sqrt{\frac{s}{4M^2}-1}\right)\right\} \qquad (A3.8)$$

where

$$\Delta=m^2(m^2-4M^2)$$

$$A=\sqrt{\frac{s-4M^2}{s}}$$

and

$$B = \sqrt{\frac{-\Delta}{4M^2m^2 + \Delta}}$$

For $s<4M^2$, we can obtain $T_B(s,t=0)$ by analytically continuing (A3.8) in s.

APPENDIX 4: FOURTH ORDER SCALAR SELF ENERGY GRAPH

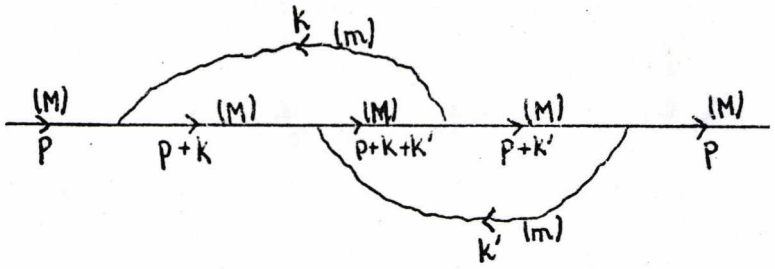


FIG.I

The Feynman diagram for this process is given in Fig.I, where p is the external four-momenta and k, k' the internal loop momenta; for each line the mass of the corresponding particle is given in brackets by the line. Evaluating the matrix element I using the method of Appendix 3 we obtain, neglecting constant factors,

$$I = \int_0^1 u^2(1-u)du \int_0^1 v\,dv \int_0^1 dw \int_0^1 \frac{dz}{CD} \qquad (A4.1)$$

where

$$C = AB - H^2$$

$$D = \sum s + C\tilde{C}$$

$$\sum = Cu + 2FGH - AF^2 - BG^2$$

and

$$G = u(1-v+vw)$$

$$F = u(1-vw)$$

$$A = G + (1-u)z$$

$$B = F + (1-u)(1-z)$$

$$C = su + \tilde{C}$$

$$\tilde{C} = -M^2u - m^2(1-u)$$

and

$$H = u(1-v)$$

APPENDIX 5: NATURE OF THE LEADING SINGULARITY OF FEYNMAN GRAPHS

The discussion here follows that of (3a). For a Feynman graph G with N internal lines, the result of integrating over the L independent loop momenta is, apart from constant multiplicative factors,

$$I = \int_0^1 \prod_{i=1}^{N} d\alpha_i \, \delta\left(\sum_{i=1}^{N} \alpha_i - 1\right) \frac{C^{N-2L-2}}{D^{N-2L}} \tag{A5.1}$$

(see (A3.3)). Consider the singularity corresponding to v contractions ($C \leq v < N-1$) of G. The Landau-Cutkosky rules can then be written as

$$D = 0$$

and

$$\alpha_i = 0 \qquad i = 0, 1, \ldots, v \tag{A5.2}$$

or

$$\frac{\partial D}{\partial \alpha_i} = 0 \qquad i = v+1, v+2, \ldots, N$$

Performing the α_N-integration in (A5.1) has the effect of changing $D(z; \alpha_1, \ldots \alpha_N)$ to

$$D(z; \alpha_1, \ldots, \alpha_{N-1}, 1 - \alpha_1 - \ldots - \alpha_{N-1}) = D'(z; \alpha_1, \ldots, \alpha_{N-1})$$

(where z represents the invariants s,t and so on of G). Then (A5.2) becomes

$$\begin{aligned} &(a) && D' = 0 \\ &(b) && \alpha_i = 0 \qquad i = 0, 1, \ldots, v \\ &(c) && \frac{\partial D'}{\partial \alpha_i} = 0 \qquad i = v+1, \ldots, N-1 \end{aligned} \tag{A5.3}$$

If (a) and (b) have solution, for a given z_r,

$$\alpha_i = \bar{\alpha}_i(z_r)$$

then (a) gives the singularity surface of (A5.1) as

$$D'(z_r; \bar{\alpha}_i(z_r)) = 0 \tag{A5.4}$$

In the neighbourhood of z_r (satisfying (A5.4)) we can expand D by Taylor's theorem; retaining only the lowest order terms we have

$$D'(z_r; \alpha_i) = D'(z_r; \bar{\alpha}_i) + \sum_{i=1}^{V} (\alpha_i - \bar{\alpha}_i) \left.\frac{\partial D'}{\partial \alpha_i}\right|_{\alpha_i = \bar{\alpha}_i} + \frac{1}{2} \sum_{j,k=v+1}^{N-1} (\alpha_j - \alpha_j)(\alpha_k - \alpha_k) \left.\frac{\partial^2 D'}{\partial \alpha_j \partial \alpha_k}\right|_{\alpha = \bar{\alpha}}$$

Although we are only concerned with a finite segment of hypercontour in the neighbourhood of $\alpha = \bar{\alpha}$, the singular part will be unaffected if we let each α-integration extend from $-\infty$ to $+\infty$, provided the power (N-2L) of the denominator in (A5.1) is sufficiently large. Explicit integration then shows that the singularity is the same as that in

$$\left[D'(z_r; \bar{\alpha}_i(z_r))\right]^{-\gamma} \tag{A5.5}$$

provided

$$\gamma = \tfrac{1}{2}[N-v+1] - 2L > 0$$

If $\gamma \leq 0$ the infinite extension of the hypercontour is invalid. In this case we replace D' by $D'+\eta$ and differentiate the integral with respect to η sufficiently so that the infinite hypercontour extension becomes valid. Then (A5.5) is replaced by

$$\left[D'(z_r; \bar{\alpha}_i(z_r))\right]^{|\gamma|} \log D' \tag{A5.6}$$

where γ is a negative integer. When γ is half-integral and negative (A5.5) again holds.

APPLICATION TO THE GRAPHS OF CHAPTER 3

We can apply these general results to determine the nature of the leading singularity of the following graphs discussed in Chapter 3:

(i) SECOND ORDER SELF ENERGY GRAPH: In this case, and in all the following, we have v=0. Here, N=2 and L=1 so

$$\gamma = \tfrac{1}{2}(N-v-1) + 2L = -\frac{1}{2}$$

Then (A5.5) shows that the leading singularity is a square root singularity.

(ii) TRIANGLE GRAPH: Here N=3 and L=1 so

$$\gamma = 0$$

and (A5.6) shows that the leading singularity is a logarithmic singularity.

(iii) FOURTH ORDER BOX GRAPH: Here N=4 and L=1, giving

$$\gamma = \frac{1}{2}$$

and hence a leading square root singularity.

(iv) FOURTH ORDER SELF ENERGY GRAPH: Here N=5 and L=2, so that

$$\gamma = -1$$

and hence a leading logarithmic singularity.

REFERENCES

1. J.S.R. Chisholm, A.C. Genz and M. Pusterla, "A method for computing Feynman amplitudes with branch cuts", Jour.Comp.Appl.Phys.2,73-76 (1976).
2. J.S.R. Chisholm, Proc.Camb.Phil.Soc. 48,300 (1952).
3. (a) R.J. Eden, P.V. Landshaff, D.I. Olive and J.C. Polkinghorne, "The analytic S-matrix", C.U.P. (1966).
   (b) R.C. Hwa and V.L. Teplitz, "Homology and Feynman integrals", Benjamin (1966).
4. H.S. Wall, "Analytic theory of continued fractions", Van Nostrand (1948).
5. R.P. Feynman, Phys.Rev.76,749 and 769 (1949).
6. S.S. Schweber, "An introduction to relativistic quantum field theory", (1962).
7. L.Landau, Nuclear Physics 13,181 (1959).
8. R.E. Cutkosky, J. Math. Phys. 1,429 (1960).
9. E.J. Squires, "An Introduction to relativistic S-matrix theory", in "Strong Interactions and high energy physics" (Scottish Universities' Summer School, 1963), ed. R.G. Moorhouse.
10. S. Mandelstam, Phys.Rev. 115, 1741 (1959).
11. I.T. Todorov, "Analytic properties of Feynman diagrams in quantum field theory", (Pergamon, 1971).
12. A.C. Genz, (Ph.D. Thesis) "The Approximate calculation of multi-dimensional integrals using extrapolation methods", (unpublished ,1975).
13. A. Grundmann and H.M. Moller, "Invariant integration formulas for the n-simplex by combinatorial methods", (unpublished, 1977).
14. D.E. Roberts, private communication.

CHAPTER 4: MULTIVARIATE APPROXIMANTS

1. TWO VARIABLE DIAGONAL CHISHOLM RATIONAL APPROXIMANTS

In the previous chapters we have considered certain types of multi-valued approximants in one variable; the success of these approximation schemes makes it worthwhile to investigate the possible extension of the one variable schemes to $N \geq 2$ variables. In this chapter we shall concentrate on the case $N=2$ and, at the end of the chapter, we shall indicate the possible extension to three and more variables. In analogy with the one variable case, these two variable (multi-valued) approximants are defined by extending the method used to define the corresponding two variable rational approximants; for this reason we first consider these latter approximants. It is also convenient to consider diagonal approximants initially because

(a) by analogy with the one variable approximants, it is the diagonal approximants which we expect to be of greatest use computationally, and

(b) the necessary generalizations to multi-valued approximants are most easily seen in terms of the diagonal approximants. We shall later see that the off-diagonal approximants can be defined in a natural way from the diagonal approximants.

Given a two variable power series

$$f(\underline{z}) = \sum_{\underline{\alpha}=\underline{0}}^{\underline{\infty}} c_{\underline{\alpha}} \underline{z}^{\underline{\alpha}} \tag{1.1}$$

where we use the notation

$$\underline{z} = (z_1, z_2), \quad \underline{\alpha} = (\alpha_1, \alpha_2), \quad \underline{0} = (0,0), \quad \underline{\infty} = (\infty, \infty)$$

and

$$\underline{z}^{\underline{\alpha}} = z_1^{\alpha_1} z_2^{\alpha_2}$$

Chisholm (1) has defined two variable diagonal rational approximants in the following manner. The (m/m) diagonal approximant (where $\underline{m}=(m_1,m_2)$) to $f(z)$, written $f_{(\underline{m}/\underline{m})}(\underline{z})$, is the rational function

$$f_{(\underline{m}/\underline{m})}(\underline{z}) = \frac{\sum_{\underline{\alpha}=\underline{0}}^{\underline{m}} a_{\underline{\alpha}} \underline{z}^{\underline{\alpha}}}{\sum_{\underline{\beta}=\underline{0}}^{\underline{m}} b_{\underline{\beta}} \underline{z}^{\underline{\beta}}} \tag{1.2}$$

If we adopt the conventional normalisation

$$b_{00} = 1$$

then, to define $f_{(\underline{m}/\underline{m})}(z)$, we need to give a prescription for calculating the $2(m+1)^2-1$ unknown coefficients $\{a_\alpha\}$, $\{b_\beta\}$ in (1.2). These coefficients are determined by requiring that they satisfy $2(m+1)^2-1$ homogeneous linear equations, formed by equating to zero certain linear combinations of coefficients in

$$E(z) = \sum_{\underline{\epsilon}=\underline{0}}^{\infty} e_{\underline{\epsilon}} z^{\underline{\epsilon}} = \left[\sum_{\underline{\beta}=\underline{0}}^{\underline{m}} b_{\underline{\beta}} z^{\underline{\beta}}\right]\left[\sum_{\underline{\lambda}=\underline{0}}^{\infty} c_{\underline{\lambda}} z^{\underline{\lambda}}\right] - \sum_{\underline{\alpha}=\underline{0}}^{\infty} a_{\underline{\alpha}} z^{\underline{\alpha}} \tag{1.3}$$

With this notation, the linear system of equations becomes

$$e_{\underline{\lambda}} = 0 \qquad (0 \le \lambda_1, \lambda_2 \le m) \tag{1.4}$$

$$e_{\underline{\lambda}} = 0 \qquad (0 \le \lambda_1 + \lambda_2 \le 2m\ ;\ \lambda_1 > m \text{ or } \lambda_2 > m) \tag{1.5}$$

$$w_{\lambda;1}\, e_{\lambda, 2m+1-\lambda} + w_{\lambda;2}\, e_{2m+1-\lambda, \lambda} = 0 \quad (\lambda = 1, 2, \ldots, m) \tag{1.6}$$

The relative weighting of the symmetrizing equations (1.6) is determined by the ratio $w_{\lambda;1} : w_{\lambda;2}$. The original choice (1) was $w_{\lambda;1} = w_{\lambda;2} = 1$ and, in order to avoid the complication of introducing specific weights, we adopt this choice. We shall later consider possible schemes for choosing the ratio $w_{\lambda;1} : w_{\lambda;2}$.

We can understand equations (1.4)-(1.6), and their method of solution, in terms of a lattice space diagram in the $\underline{\alpha} = (\alpha_1, \alpha_2)$ plane (see Fig.1). This diagram indicates the terms $z_1^{\alpha_1} z_2^{\alpha_2}$ of (1.1) and (1.2) which are to be matched; matching occurs on the triangle $\alpha_1 \ge 0, \alpha_2 \ge 0, \alpha_1 + \alpha_2 \le 2m$, together with the symmetrization of (1.6) which is represented in Fig.1 by the numbered crosses along the line $\alpha_1 + \alpha_2 = 2m+1$. An important point is that (1.5) and (1.6) can be solved independently of (1.4); we first solve (1.5) and (1.6) for the denominator coefficients $\{b_\sigma\}$ and we can then solve (1.4) for the numerator coefficients $\{a_\sigma\}$. In the lattice space diagram this corresponds to matching terms in the triangles T_1 and T_2 to determine $\{b_\sigma\}$ and then matching terms in the square S to determine $\{a_\sigma\}$.

The equations (1.5) and (1.6) are solved by the so-called "prong method"

FIG.I: LATTICE SPACE FOR EQUATIONS (1.4)-(1.6); THE NUMBERED CROSSES DENOTE SYMMETRISATION AND PRONG j IS REPRESENTED BY THE VERTICAL AND HORIZONTAL LINES MARKED j.

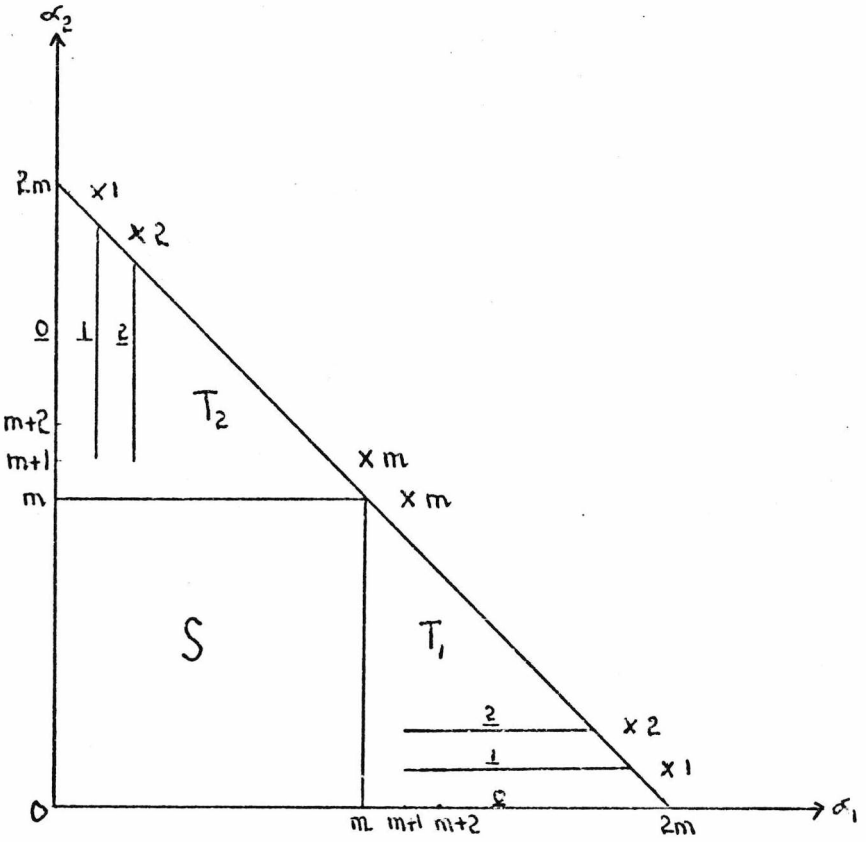
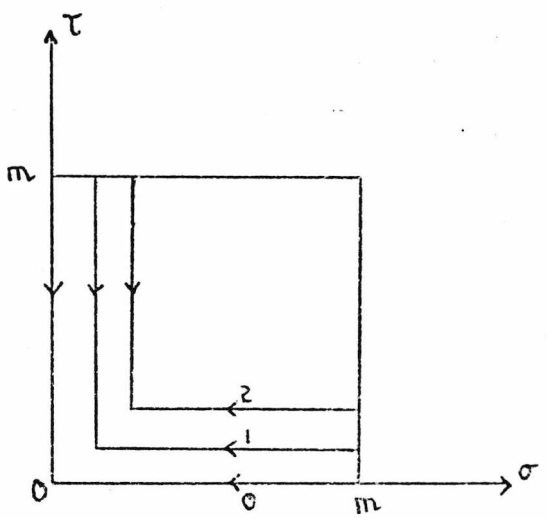


FIG. 2: ORDERING FOR $\{b_{\sigma\tau}\}$.



(2); this is essentially a block by block method of solution and, especially for large m, greatly reduces the amount of computation and storage required. Define, for $0 \le j < m$,

$$P_j = \left\{(\alpha_1, j): m+1 \le \alpha_1 \le 2m-j\right\} \cup \left\{(j, \alpha_2): m+1 \le \alpha_2 \le 2m-j\right\} \tag{1.7}$$

Then P_j is referred to as "prong j". The matching of coefficients along prong zero produces the linear system

$$G_{m+1}\, \underline{b}_0 = \underline{u}_m \tag{1.8a}$$

where

$$G_{m+1} = \left[\begin{array}{c|c|c} \begin{matrix} c_{1,0} & \cdots & c_{m,0} \\ \vdots & & \vdots \\ c_{m,0} & \cdots & c_{2m-1,0} \end{matrix} & \underline{0} & \begin{matrix} c_{m+1,0} \\ \vdots \\ c_{2m,0} \end{matrix} \\ \hline \underline{0} & \begin{matrix} c_{0,1} & \cdots & c_{0,m} \\ \vdots & & \vdots \\ c_{0m} & \cdots & c_{0,2m-1} \end{matrix} & \begin{matrix} c_{0,m+1} \\ \vdots \\ c_{0,2m} \end{matrix} \\ \hline 0 \cdots 0 & 0 \cdots 0 & 1 \end{array}\right] \tag{1.8b}$$

$$\underline{b}_0^T = \left[b_{m,0} \cdots b_{1,0}\,;\; b_{0,m} \cdots b_{0,1}\,;\; b_{0,0}\right] \tag{1.8c}$$

and

$$\underline{u}_m^T = \left[0 \cdots 0\,;\; 0 \cdots 0\,;\; 1\right] \tag{1.8d}$$

If we represent the coefficients $\{b_{\sigma\tau}\}$ on the square $0 \le \sigma \le m, 0 \le \tau \le m$ (see Fig.2), then the ordering adopted is that indicated by the L-shaped line marked zero (which is, in effect, prong zero of Fig.I). The important point is that the (2m-1) coefficients $\{b_{\sigma,\tau}\}$, represented by the segment labelled I in Fig.2, can be expressed in terms of the $\{b_{\sigma\tau}\}$ represented by the segment labelled 0 (which have already been determined). The equations come from matching terms on prong I together with the symmetrised equation obtained from the points (1,2m) and (2m,1). We can write the equations on prong I in the form

$$R_m\, \underline{b}_0 + D_m\, \underline{b}_1 = \underline{0} \tag{1.9a}$$

where

$$D_m = \begin{bmatrix} c_{1,0} & \cdots & c_{m-1,0} & & \underline{0} & & c_{m,0} \\ \vdots & & \vdots & & & & \vdots \\ c_{m-1,0} & \cdots & c_{2m-3,0} & & & & c_{2m-2,0} \\ \hline & \underline{0} & & c_{0,1} & \cdots & c_{0,m-1} & c_{0,m} \\ & & & \vdots & & \vdots & \vdots \\ & & & c_{0,m-1} & \cdots & c_{0,2m-3} & c_{0,2m-2} \\ \hline c_{m,0} & \cdots & c_{2m-2,0} & c_{0,m} & \cdots & c_{0,2m-2} & c_{0,2m-1} + c_{2m-1,0} \end{bmatrix} \qquad (1.9b)$$

and

$$\underline{b}_1^T = \left[b_{m,1} \cdots b_{2,1} \; ; \; b_{1,m} \cdots b_{1,2} \; ; \; b_{1,1} \right] \qquad (1.9c)$$

Continuing the process of matching terms, the $\{b_{\sigma\tau}\}$ represented by the segment labelled 2 can be expressed in terms of the (already determined) $\{b_{\sigma\tau}\}$ on the segments 0 and 1. This prong structure illustrates the important fact that, at each stage, the number of equations in (1.5) and (1.6) introduced exactly matches the number of new unknowns. It is now clear that the equations (1.5) and (1.6) produce the following block system of linear equations:

$$\begin{bmatrix} G_{m+1} & & & \\ R_m & D_m & & \underline{0} \\ R_{m-1} & & D_{m-1} & \\ & \ddots & & \ddots \\ R_1 & & & D_1 \end{bmatrix} \begin{bmatrix} \underline{b}_0 \\ \underline{b}_1 \\ \vdots \\ \underline{b}_m \end{bmatrix} = \begin{bmatrix} \underline{u}_m \\ \underline{0} \\ \vdots \\ \underline{0} \end{bmatrix} \qquad (1.10)$$

where the D_i are square, of order 2i-1, and have the same structure as D_m; the b_i have dimension 2m-2i+1 and their general form follows from that of b_o and b_1 given by (1.8c) and (1.9c).

By writing (1.5) and (1.6) in the form (1.10), we can see that the $\{b_{\sigma\tau}\}$ can be found by a block by block process provided none of the D_i (i=1,2,...,m), nor G_{m+1}, is singular. The inversion process then produces

$$\begin{aligned} \underline{b}_0 &= G_{m+1}^{-1}\, \underline{u}_m \\ \underline{b}_1 &= -D_m^{-1} R_m \underline{b}_0 \\ &\vdots \\ \underline{b}_m &= -D_1^{-1} R_1 \begin{bmatrix} \underline{b}_0 \\ \vdots \\ \underline{b}_{m-1} \end{bmatrix} \end{aligned} \tag{1.11}$$

To make a comparison of the two variable approximation schemes considered in this chapter, we illustrate the appropriate prong structures collectively at the end of the chapter; the prong structure for the diagonal rational approximants of this section is given in Fig.3.

Having defined the Chisholm rational approximants and indicated how they can be determined efficiently in practice, we now discuss the main properties of the approximants. The system of equations (1.4)-(1.6) are chosen because they produce approximants satisfying a number of properties which are considered desirable. These properties (proofs of which can be found in (1)) are:

P1: SYMMETRY PROPERTY

The approximants are symmetric in z_1 and z_2 (provided we choose the weights $w_{\lambda;1}=w_{\lambda;2}=1$ in (1.6)).

P2: PROJECTION PROPERTY

If $z_1=0$ or $z_2=0$, the approximants reduce to one variable Padé approximants in the other variable.

P3: RECIPROCAL COVARIANCE

The $(\underline{m}/\underline{m})$ approximant derived from the formal reciprocal of the series (1.1) is

$$\left[F_{(\underline{m}/\underline{m})}(\underline{z})\right]^{-1}$$

P4: HOMOGRAPHIC COVARIANCE

The (m/m) approximant is invariant under the transformation

$$z_r = \frac{A w_r}{1 - B_r w_r} \qquad (r=1,2) \tag{1.12}$$

where $A(\neq 0), B_1$ and B_2 are complex numbers.

For a certain choice of weights (3), P4 can be extended to include the relative scale transformations

$$w_r = K_r w_r \quad (r=1,2\ ;\ K_1K_2=1) \tag{1.13}$$

Properties P1-P4 extend to the N-variable diagonal approximants (4), and properties P1-P3 hold for the N-variable off-diagonal generalisations (5) of the Chisholm approximants. In defining the two variable multi-valued approximants we seek to preserve properties P1-P4, together with the prong method of solution.

2. DIAGONAL TWO VARIABLE QUADRATIC APPROXIMANTS

(a) DEFINITION OF THE APPROXIMANTS

The extension of rational two variable diagonal approximants to two variable diagonal approximants with branch cuts has been given by Chisholm (6). The general type of branch cut we are trying to build into the approximants is a $t^{th.}$ root ($t=2,3,\ldots$); for the one variable situation, we have already considered in detail the cases $t=2$ and $t=3$, which give rise to quadratic and cubic approximants respectively. In general, for approximants with a $t^{th.}$ root branch point, we refer to "t-power approximants".

Before defining two variable "t-power approximants" we have to make an important (though seemingly trivial) alteration to the definition of the corresponding one variable approximants. The one variable "t-power approximants" satisfy an equation of the form

$$\sum_{k=0}^{t} P_k(z) f^k(z) = O(z^\sigma) \tag{2.1}$$

where

$$\sigma = \sum_{k=0}^{t} p_k + t \tag{2.2}$$

and p_k is the degree of $P_k(z)$. The required modification is to let

$$\sigma = \sum_{k=0}^{t} p_k + 1 \tag{2.3}$$

There are thus (t-1) coefficients which must be determined initially; we shall later consider methods for doing this.

We now consider the specific case of two variable quadratic approximants, as these approximants illustrate the basic features common to all "t-power approximants". In this case we write (2.1), taking account of (2.3), in the more familiar form

$$P(\underline{z})f^2(\underline{z}) + Q(\underline{z})f(\underline{z}) + R(\underline{z}) = O(\underline{z}^{3m+1}) \tag{2.4}$$

where P,Q and R are all of degree m and q_{00} is assumed known. Given $f(\underline{z})$ defined by (1.1) we form the formal square of the power series

$$f^2(\underline{z}) = \sum_{\underline{\alpha}=\underline{0}}^{\infty} \overline{c}_{\underline{\alpha}}\, \underline{z}^{\underline{\alpha}} \tag{2.5}$$

where

$$\overline{c}_{\underline{\alpha}} = \sum_{\underline{\lambda}+\underline{\beta}=\underline{\alpha}} c_{\underline{\lambda}}\, c_{\underline{\beta}} \tag{2.6}$$

We define the polynomials

$$P(\underline{z}) = \sum_{\underline{\alpha}=\underline{0}}^{m} p_{\underline{\alpha}}\, \underline{z}^{\underline{\alpha}}$$

$$Q(\underline{z}) = \sum_{\underline{\alpha}=\underline{0}}^{m} q_{\underline{\alpha}}\, \underline{z}^{\underline{\alpha}}$$

and

$$R(\underline{z}) = \sum_{\underline{\alpha}=\underline{0}}^{m} r_{\underline{\alpha}}\, \underline{z}^{\underline{\alpha}}$$

with coefficients defined over the square $(\underline{0},\underline{m})$ of Fig.4 (at the end of this chapter), by equating to zero a certain set of coefficients in

$$E(\underline{z}) = \sum_{\underline{\epsilon}=\underline{0}}^{\infty} e_{\underline{\epsilon}}\, \underline{z}^{\underline{\epsilon}} = P(\underline{z})f^2(\underline{z}) + Q(\underline{z})f(\underline{z}) + R(\underline{z}) \tag{2.7}$$

To satisfy the projection property P2 (of §1) to the one variable approximants (2.4), prong 0 must consist of the points

$$\underline{\epsilon} = (1,0), (2,0), \ldots\ldots\ldots, (3m,0)$$

and

$$\underline{\epsilon} = (0,1), (0,2), \ldots\ldots\ldots, (0,3m)$$

With reference to Fig.4, the set of equations corresponding to the solid sections of prong 0 will then normally determine the sets

$$(p_{1,0}, \ldots\ldots, p_{m,0}\,;\, q_{1,0}, \ldots\ldots, q_{m,0})$$

and

$$(p_{0,1}, \ldots\ldots, p_{0,m}\,;\, q_{0,1}, \ldots\ldots, q_{0,m})$$

uniquely in terms of p_{oo} and q_{oo}, where q_{oo} is assumed known and p_{oo} is fixed by the normalisation condition

$$p_{oo} = 1$$

It is now clear why Shafer's definition needed to be modified; extension of prong 0 to the points $(3m+1,0)$ and $(0,3m+1)$, necessary to satisfy the projection property to Shafer's approximants, would produce an extra equation and hence two (probably different) values of q_{oo}.

The broken part of prong 1 corresponds to $(2m-1)$ new $\{p_{\underline{\alpha}}\}$ and $(2m-1)$ new $\{q_{\underline{\alpha}}\}$

$$\left(p_{1,1}, \ldots\ldots, p_{m,1} \; ; \; p_{1,2}, \ldots\ldots\ldots, p_{1,m}\right)$$

and

$$\left(q_{1,1}, \ldots\ldots, q_{m,1} \; ; \; q_{1,2}, \ldots\ldots\ldots, q_{1,m}\right)$$

The solid section of prong 1 must thus consist of $2(2m-1)$ points, and hence must correspond to vanishing $e_{\underline{\epsilon}}$ with

$$\underline{\epsilon} = (m+1,1), \ldots\ldots\ldots\ldots\ldots, (3m-1,1)$$

and

$$\underline{\epsilon} = (1,m+1), \ldots\ldots\ldots\ldots\ldots, (1,3m-1)$$

Then clearly prong 2 will consist of $(2m-3)$ points in each direction, and so on, until $(p_{m-1,m-1}, p_{m,m-1}, p_{m-1,m}; q_{m-1,m-1}, q_{m-1,m})$ are determined with

$$\underline{\epsilon} = (m-1,m+1), (m-1,m+2), (m-1,m+3)$$

and

$$\underline{\epsilon} = (m+1,m-1), (m+2,m-1), (m+3,m-1)$$

On the final prong, prong m, we have a slight difficulty. To determine $p_{m,m}$ and $q_{m,m}$ we would expect to use the points $(m+1,m)$ and $(m,m+1)$. However, at these points, the parts of $e_{\underline{\epsilon}}$ containing $p_{m,m}$ and $q_{m,m}$ are

$$\bar{c}_{1,0}\, p_{m,m} + c_{1,0}\, q_{m,m} = c_{1,0}\left(2c_{0,0}\, p_{m,m} + q_{m,m}\right)$$

and

$$\bar{c}_{0,1}\, p_{m,m} + c_{0,1}\, q_{m,m} = c_{0,1}\left(2c_{0,0}\, p_{m,m} + q_{m,m}\right)$$

which obviously produce linearly dependent equations and, in general, will give rise to an inconsistency. Three procedures for avoiding this inconsistency, whilst preserving symmetry, have been suggested;

(i) Omit prong m; this is equivalent to setting

$$p_{m,m} = q_{m,m} = r_{m,m} = 0 \tag{2.8a}$$

In this case we must omit the lattice point (m,m) at the corner of the square in Fig.4.

(ii) Set $q_{m,m}=0$ and use the equations

$$e_{m,m} = e_{m+1,m} + e_{m,m+1} = 0 \tag{2.8b}$$

to determine $p_{m,m}$ and $r_{m,m}$.

(iii) Use the double symmetrisation scheme

$$e_{m+1,m} + e_{m,m+1} = e_{m+2,m} + e_{m,m+2} = 0 \tag{2.8c}$$

to determine $p_{m,m}$ and $q_{m,m}$.

The relative merits of these schemes will be discussed in §2(b). The prong structure is illustrated in Fig.4, where crosses denote the points contributing to the symmetrised equations in (2.8b) and (2.8c). We note that, in contrast to the approximants of §1, no symmetrised equations are required on the internal prongs (prongs 1 to m-1). This occurs because of the coincidence of the power t(=2) with the number of variables N; in general, the existence of symmetrised equations depends upon the relation between t and N.

The preceding system of equations, over the lattice of Fig.4, will normally determine $P(\underline{z}),Q(\underline{z})$ and $R(\underline{z})$ uniquely, once p_{00} and q_{00} are given. The diagonal two variable quadratic approximant, $f_{(\underline{m}/\underline{m}/\underline{m})}(\underline{z})$, is then defined by

$$P(\underline{z})\left[f_{(\underline{m}/\underline{m}/\underline{m})}(\underline{z})\right]^2 + Q(\underline{z}) f_{(\underline{m}/\underline{m}/\underline{m})}(\underline{z}) + R(\underline{z}) = 0 \tag{2.9}$$

As can be seen from the above prong structure, the method of determining the diagonal quadratic approximants is very similar to the procedure adopted for the diagonal Chisholm approximants of §1. In fact equations (1.8)-(1.11) hold for the quadratic approximants, where (adopting (2.8c) on prong m for definiteness):

$$\underline{b}_0^T = \left[p_{m,0}, \ldots, p_{1,0}; q_{m,0}, \ldots, q_{1,0}; p_{0,m}, \ldots, p_{0,1}; q_{0,m}, \ldots, q_{0,1}; p_{0,0}, q_{0,0}\right]$$

$$\underline{u}_m^T = \left[0, \ldots, 0; 0, \ldots, 0; 1, 0\right]$$

$$G_{m+1} = \begin{bmatrix} b_{1,0} \cdots b_{m,0} \; c_{1,0} \cdots c_{m,0} & \underline{0} & b_{m+1,0} \; c_{m+1,0} \\ \vdots \qquad \vdots \quad \vdots \qquad \vdots & & \vdots \qquad \vdots \\ b_{2m,0} \cdots b_{3m-1,0} \; c_{2m,0} \cdots c_{3m-1,0} & & b_{3m,0} \; c_{3m,0} \\ \underline{0} & b_{0,1} \cdots b_{0,m} \; c_{0,1} \cdots c_{0,m} & b_{0,m+1} \; c_{0,m+1} \\ & \vdots \qquad \vdots \quad \vdots \qquad \vdots & \vdots \qquad \vdots \\ & b_{0,2m} \cdots b_{0,3m-1} \; c_{0,2m} \cdots c_{0,3m-1} & b_{0,3m} \; c_{0,3m} \\ 0 \cdots 0 \; 0 \cdots 0 & 0 \cdots 0 \; 0 \cdots 0 & 1 \quad 0 \\ b_{2m+1,0} \cdots b_{3m,0} \; c_{2m+1,0} \cdots c_{3m,0} & b_{0,2m+1} \cdots b_{0,3m} \; c_{0,2m+1} \cdots c_{0,3m} & b_{3m+1,0}+b_{0,3m+1} \; c_{3m+1,0}+c_{0,3m+1} \end{bmatrix} \tag{2.10}$$

$$D_m = \begin{bmatrix} b_{1,0} \cdots b_{m-1,0} \; c_{1,0} \cdots c_{m-1,0} & \underline{0} & b_{m,0} \; c_{m,0} \\ \vdots \qquad \vdots \quad \vdots \qquad \vdots & & \vdots \qquad \vdots \\ b_{2m-1,0} \cdots b_{3m-3,0} \; c_{2m-1,0} \cdots c_{3m-3,0} & & b_{3m-2,0} \; c_{3m-2,0} \\ \underline{0} & b_{0,1} \cdots b_{0,m-1} \; c_{0,1} \cdots c_{0,m-1} & b_{0,m} \; c_{0,m} \\ & \vdots \qquad \vdots \quad \vdots \qquad \vdots & \vdots \qquad \vdots \\ & b_{0,2m-1} \cdots b_{0,3m-3} \; c_{0,2m-1} \cdots c_{0,3m-3} & b_{0,3m-2} \; c_{0,3m-2} \end{bmatrix} \tag{2.11}$$

$$\underline{b}_i^T = \left[p_{m,i}, \ldots, p_{2,i} ; q_{m,i}, \ldots, q_{2,i} ; p_{i,m}, \ldots, p_{i,2} ; q_{i,m}, \ldots, q_{i,2} ; p_{i,i}, q_{i,i} \right]$$

and the D_i are square of order $4i-2$ and the $\underline{b}_i$ have dimension $4m-4i+2$; the coefficients b_i occuring in the above formulae are the $\overline{c}_{\underline{\alpha}}$ of (2.5) and (2.6).

We again have the result that the $\{p_{\sigma,\tau} ; q_{\sigma,\tau}\}$ can be found by a block by block process provided that none of the $D_i (i=1,\ldots,m)$, nor G_{m+1}, is singular. The only minor modifications necessary are:

(a) the precise form of D_1 depends upon the scheme adopted on the final prong, and

(b) the precise form of G_{m+1} depends upon the method used for determining q_{00}; the form of G_{m+1} given by (2.10) corresponds to determining q_{00} by the symmetrisation

$$e_{3m+1,0} + e_{0,3m+1} = 0$$

We shall discuss this scheme in § 2(b).

(b) PROPERTIES OF THE APPROXIMANTS

The approximants of §2(a) were defined in such a way as to preserve, to as large an extent as possible, the properties P1-P4 of §1; we now examine the degree to which these properties have been preserved (6).

By their very definition, the approximants satisfy P1 and P2. To discuss P3 (reciprocal covariance) we assume that $c_{00} \neq 0$, so that we can define a unique formal reciprocal $f^{-1}(\underline{z})$ of (1.1), and a corresponding unique formal reciprocal $f^{-2}(\underline{z})$ of (2.5). Multiplying (2.7) formally by $f^{-2}(\underline{z})$, we obtain the expression

$$G(\underline{z}) \equiv R(\underline{z}) f^{-2}(\underline{z}) + Q(\underline{z}) f^{-1}(\underline{z}) + P(\underline{z}) \qquad (2.12)$$

Now we make use of the rectangle rule ((1),(4) and (5)); this rule imposes a geometrical condition on the lattice space (on which the approximant is defined) which is equivalent to the covariance properties P3 and P4. Essentially, the rule states that if $\underline{\alpha}$ is a symmetrised point (that is, a point contributing to a symmetrised equation), then no other points in the rectangle, with diagonal joining $\underline{0}$ and $\underline{\alpha}$, can be symmetrised points; if this condition holds, then P3 and P4 are valid. This rule ensures that coefficients in $G(\underline{z})$ corresponding to a point on the lattice of Fig.4, will depend linearly on the coefficients $e_{\underline{\xi}}$ in (2.7) with $\xi_r \leq \alpha_r$ $(r=1,2)$. We have two cases to consider:

(a) For the choices (i) and (ii) (equations (2.8a) and (2.8b)) on prong m, the system of equations $e_{\underline{\xi}}=0$ and (2.8a) (or (2.8b)) applied to $E(\underline{z})$ and $G(\underline{z})$, produce equivalent sets of equations for $\{p_{\underline{\alpha}}\}, \{q_{\underline{\alpha}}\}$ and $\{r_{\underline{\alpha}}\}$. It is also clear that $f^{-1}_{(\underline{m}/\underline{m}/\underline{m})}$ is defined through (2.12) from $f^{-1}(\underline{z})$ in exactly the same way as $f_{(\underline{m}/\underline{m}/\underline{m})}$ is defined from $f(\underline{z})$. The approximants will thus satisfy reciprocal covariance, provided we can choose q_{00} in a consistent way. To preserve symmetry, the most natural choice for q_{00} is given by the symmetrisation

$$e_{3m+1,0} + e_{0,3m+1} = 0 \qquad (2.13)$$

This scheme produces the circled points of Fig.4. A disadvantage of (2.13) is that the projection property P2 no longer holds; despite this, the

choice (2.13) seems a natural one and, for the numerical results quoted later in the chapter, this is the method used for determining q_{00}.

(b) For the choice (iii) (equation (2.8c)) on prong m, the equations arising from $E(\underline{z})$ and $G(\underline{z})$ are equivalent only if

$$c_{0,1} = c_{1,0} \tag{2.14}$$

To see this, let

$$f^{-1}(\underline{z}) = \sum_{\underline{\alpha}=\underline{0}}^{\infty} d_{\underline{\alpha}}\, \underline{z}^{\underline{\alpha}}$$

where

$$d_{00} = c_{00}^{-1}$$

and

$$\sum_{\underline{\alpha}=\underline{0}}^{\underline{\beta}} c_{\underline{\beta}-\underline{\alpha}}\, d_{\underline{\alpha}} = 0 \tag{2.15}$$

the summation extending over the lattice rectangle with diagonal joining 0 and $\underline{\beta}$. If we write

$$f^{-2}(\underline{z}) = \sum_{\underline{\alpha}=\underline{0}}^{\infty} \bar{d}_{\underline{\alpha}}\, \underline{z}^{\underline{\alpha}} \tag{2.16}$$

then, multiplying $E(\underline{z})$ formally by (2.16),

$$E'(\underline{z}) = \sum_{\underline{\epsilon}=\underline{0}}^{\infty} e'_{\underline{\epsilon}}\, \underline{z}^{\underline{\epsilon}} = R(\underline{z}) f^{-2}(\underline{z}) + Q(\underline{z}) f^{-1}(\underline{z}) + P(\underline{z}) \tag{2.17}$$

The rectangle rule now shows that

$$e'_{\underline{\epsilon}} = \bar{d}_{0,0}\, e_{\underline{\epsilon}} + \bar{d}_{0,1}\, e_{\epsilon_1,\epsilon_2-1} + \bar{d}_{1,0}\, e_{\epsilon_1-1,\epsilon_2} + \cdots\cdots \tag{2.18}$$

is a linear combination of $\{e_{\underline{\alpha}}\}$, with $\underline{\alpha}$ restricted to the rectangle $(\underline{0},\underline{\epsilon})$. Hence, for all unsymmetrised points in Fig.4,

$$e_{\underline{\epsilon}} = 0$$

implies

$$e'_{\underline{\epsilon}} = 0 \tag{2.19}$$

at the same points. For the first pair of symmetrised points in (2.8c), using (2.18) and (2.19),

$$e'_{m+1,m} + e'_{m,m+1} = \bar{d}_{0,0}\left(e_{m+1,m} + e_{m,m+1}\right) = 0$$

For the second pair of symmetrised points in (2.8c),

$$e'_{m+2,m} + e'_{m,m+2} = \bar{d}_{0,0}\left(e_{m+2,m} + e_{m,m+2}\right) + \bar{d}_{1,0}\, e_{m+1,m} + \bar{d}_{0,1}\, e_{m,m+1}$$

which is zero if

$$\bar{d}_{1,0} = \bar{d}_{0,1} \tag{2.20}$$

But, by definition,

$$\bar{d}_{1,0} = 2d_{0,0}d_{1,0} \; ; \; \bar{d}_{0,1} = 2d_{0,0}d_{0,1}$$

and

$$c_{0,1}d_{1,0} = c_{1,0}d_{0,1}$$

Thus (2.20) implies

$$c_{1,0} = c_{0,1}$$

which is (2.14), as required.

The condition (2.14) is equivalent to a choice of the relative scale of the variables z_1 and z_2, and does not seem an unreasonable choice to make. We shall see in §2(c) that other choices of the scale are possible.

We now discuss the homographic covariance property P4, defined by (1.12). Define polynomials $P'(\underline{w}), Q'(\underline{w})$ and $R'(\underline{w})$ by

$$P'(\underline{w}) = \prod_{r=1}^{2} (1-B_r w_r)\, P\left[\frac{A_1 w_1}{1-B_1 w_1}, \frac{A_2 w_2}{1-B_2 w_2}\right] \tag{2.21}$$

with two similar equations for $Q'(\underline{w})$ and $R'(\underline{w})$. Substituting (1.12) into (2.7), multiplying by $\prod_{r=1}^{2}(1-B_r w_r)^m$ and formally expanding the inverse powers of $(1-B_r w_r)$, we arrive at

$$E'(\underline{w}) = P'(\underline{w}) f'^2(\underline{w}) + Q'(\underline{w}) f'(\underline{w}) + R'(\underline{w}) \tag{2.22}$$

where $f'(\underline{w})$ is the formal expansion of

$$F\left[\frac{A_1 w_1}{1-B_1 w_1}, \frac{A_2 w_2}{1-B_2 w_2}\right]$$

We now have three possible cases to consider, depending upon the scheme adopted on prong m:

(a) With choice (i), the rectangle rule again ensures that the equations derived from $E(\underline{z})$ and $E'(\underline{w})$ are equivalent. Thus, from (2.9) and (2.22),

$$P'(\underline{w})\left[f'_{(\underline{m}/\underline{m}/\underline{m})}(\underline{w})\right]^2 + Q'(\underline{w}) f'_{(\underline{m}/\underline{m}/\underline{m})}(\underline{w}) + R'(\underline{w}) = 0$$

with

$$F'_{(m/m/m)}(\underline{w}) = F_{(m/m/m)}\left(\frac{A_1 w_1}{1-B_1 w_1}, \frac{A_2 w_2}{1-B_2 w_2}\right) \tag{2.23}$$

which establishes homographic covariance.

(b) With choice (ii), covariance under (1.12) is only obtained with

$$A_1 = A_2 \tag{2.24}$$

This is because the inclusion of relative scale transformations (1.13) in the group (1.12) depends upon the absence of symmetrised equations (1). We shall prove (2.24) in (c) below.

(c) With choice (iii), the covariance group is even further restricted; we now require

$$A_1 = A_2 \; ; \; B_1 = B_2 \tag{2.25}$$

To see this, we have

$$E'(\underline{w}) = \prod_{r=1}^{2} (1-B_r w_r)^m E\left(\frac{A_r w_r}{1-B_r w_r}\right)$$

$$= \prod_{r=1}^{2} (1-B_r w_r)^m \sum_{\underline{\alpha}=\underline{0}}^{\infty} e_{\underline{\alpha}} \prod_{r=1}^{2} \left(\frac{A_r w_r}{1-B_r w_r}\right)^{\alpha_r}$$

$$= \sum_{\underline{\alpha}=\underline{0}}^{\infty} e_{\underline{\alpha}} \prod_{r=1}^{2} (A_r w_r)^{\alpha_r} (1-B_r w_r)^{m-\alpha_r}$$

$$= \sum_{\underline{\alpha}=\underline{0}}^{\infty} e_{\underline{\alpha}} \prod_{r=1}^{2} A_r^{\alpha_r} \sum_{\beta_r=\alpha_r}^{m-\alpha_r} (-B_r)^{\beta_r-\alpha_r} w_r^{\beta_r} \binom{m-\alpha_r}{\beta_r-\alpha_r}$$

In this expansion, the coefficient of $\underline{w}^{\underline{\epsilon}}$ is

$$\sum_{\underline{\alpha}=\underline{0}}^{\underline{\epsilon}} e_{\underline{\alpha}} \prod_{r=1}^{2} A_r^{\alpha_r} (-B_r)^{\epsilon_r-\alpha_r} \binom{m-\alpha_r}{\epsilon_r-\alpha_r}$$

Remembering that

$$E'(\underline{w}) = \sum_{\underline{\alpha}=\underline{0}}^{\infty} e'_{\underline{\alpha}} \underline{w}^{\underline{\alpha}}$$

we see that

$$e'_{\underline{\epsilon}} = \sum_{\underline{\alpha}=\underline{0}}^{\underline{\epsilon}} e_{\underline{\alpha}} \prod_{r=1}^{2} A_r^{\alpha_r} (-B_r)^{\epsilon_r-\alpha_r} \binom{m-\alpha_r}{\epsilon_r-\alpha_r} \tag{2.26}$$

Clearly $e_{\underline{\epsilon}}=0$ implies $e'_{\underline{\epsilon}}=0$, and

$$\epsilon'_{m+1,m} + \epsilon'_{m,m+1} = (A_1 A_2)^m \left[A_1 \epsilon_{m+1,m} + A_2 \epsilon_{m,m+1} \right] \qquad (2.27)$$

since only the terms with $(\alpha=)\underline{\epsilon}=(m+1,m)$ and $\underline{\epsilon}=(m+1,m)$ contribute to the sum in (2.26). Clearly (2.27) vanishes if $A_1=A_2$ (which verifies (2.24)).

Also

$$e'_{m+2,m} + e'_{m,m+2} = -(A_1 A_2)^m \left[A_1 B_1 \binom{-1}{1} e_{m+1,m} + A_2 B_2 \binom{-1}{1} e_{m,m+1} \right]$$
$$+ (A_1 A_2)^m \left[A_1^2 \epsilon_{m+2,m} + A_2^2 \epsilon_{m,m+2} \right]$$
$$= 0$$

if

$$B_1 = B_2 \quad \text{and} \quad A_1 = A_2$$

This verifies (2.25).

The additional symmetrised equation (2.13) is also covariant under the group with restriction (2.24). Thus, with (2.13) included, the choices (i) and (ii) on prong m satisfy reciprocal and homographic covariance under the restricted group (1.12). The introduction of weighting factors into (2.13) provides a possible method of ensuring relative scale covariance (3); in the following section we shall consider this possibility.

We conclude this section by suggesting another method of choosing q_{00} (6). Our initial power series expansion (1.1) corresponds to function values defined on a single Riemann sheet. If we have information concerning the analytic continuation of the function onto other sheets, we may fix q_{00} by requiring that the function takes the correct value at the origin on another Riemann sheet. The advantage of this procedure is that the projection property is preserved and, with choice (i) on prong m, we have covariance under the full group (1.13). We then have an approximant which incorporates information about the analytic continuation of (1.1). Another point in favour of this scheme is that it extends to arbitrary "t-power approximants". The possibility of using several terms of the series on a second sheet, to define an approximant incorporating information on more than one sheet, also suggests itself. Although these ideas may prove to be useful, we have not yet investigated any of these methods from a practical point of view.

(c) CHOICE OF WEIGHT FACTORS

The symmetrised equations on prongs 0 and m allow the possibility of introducing weighting factors to, for example, try and ensure relative scale covariance of the approximants (3). Since prong m is not expected to be too important in determining the behaviour of the approximants (since prong m only determines the higher order coefficients), we shall mainly be concerned with prong 0. If we adopt the symmetrisation scheme to determine q_{00}, then we can write (2.13) more generally as

$$\lambda_0 e_{3m+1,0} + \mu_0 e_{0,3m+1} = 0 \tag{2.28}$$

The method of choosing a suitable ratio $\lambda_0 : \mu_0$ depends upon being able to give an explicit representation for the determinant of G_{m+1} of (2.10) (with suitable account being taken of (2.28)). The matrix, D_{00}, occuring on prong 0 has the form

$$D_{00} = \begin{bmatrix} E_0^{(1)} & 0 & X_0^{(1)} \\ 0 & E_0^{(2)} & X_0^{(2)} \\ \lambda_0 U_0^{(1)} & \mu_0 U_0^{(2)} & S_0 \end{bmatrix} \tag{2.29}$$

where

$$E_0^{(1)} = \begin{bmatrix} b_{1,0} & \cdots & b_{m,0} & c_{1,0} & \cdots & c_{m,0} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{2m,0} & \cdots & b_{3m-1,0} & c_{2m,0} & \cdots & c_{3m-1,0} \end{bmatrix} \tag{2.29a}$$

$$X_0^{(1)T} = \left[c_{m+1,0}, c_{m+2,0}, \cdots, c_{3m,0} \right] \tag{2.29b}$$

$$U_0^{(1)} = \left[b_{2m+1,0}, \cdots, b_{3m,0}; c_{2m+1,0}, \cdots, c_{3m,0} \right] \tag{2.29c}$$

$$S_0 = \lambda_0 c_{3m+1,0} + \mu_0 c_{0,3m+1} \tag{2.29d}$$

and the corresponding entries with superscript 2 are obtained from those with superscript 1 by interchanging the indices on the b and c coefficients.

Following (7), we now factorize D_{00} as

$$D_{00}^{\rho\sigma} = \sum_{r=1}^{2} S^{\rho r}(\lambda_0, \mu_0) E^{r\sigma} \tag{2.30}$$

where ρ and σ refer to the block indices of (2.29). In block form the matrices S and E are

$$S(\lambda_o, \mu_o) = \begin{bmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \\ S^{31} & S^{32} \end{bmatrix} \tag{2.31}$$

with

$$S^{11} = S^{22} = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & 1 \; 0 \end{bmatrix}$$

$$S^{31} = (0, 0, \ldots\ldots, 0, \lambda_o)$$

$$S^{32} = (0, 0, \ldots\ldots, 0, \mu_o)$$

$$S^{12} = S^{21} = 0$$

and

$$E = \begin{bmatrix} E^{11} & E^{12} & E^{13} \\ E^{21} & E^{22} & E^{23} \end{bmatrix} \tag{2.32}$$

with

$$E^{11} = \begin{bmatrix} b_{1,0} & \cdots & b_{m,0} & c_{1,0} & \cdots & c_{m,0} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{2m,0} & \cdots & b_{3m-1,0} & c_{2m,0} & \cdots & c_{3m-1,0} \\ b_{2m+1,0} & \cdots & b_{3m,0} & c_{2m+1,0} & \cdots & c_{3m,0} \end{bmatrix}$$

$$E^{13} = \left[c_{m+1,0}, \ldots\ldots\ldots, c_{3m+1,0} \right]^T$$

$$E^{23} = \left[c_{0,m+1}, \ldots\ldots\ldots, c_{0,3m+1} \right]^T$$

$$E^{12} = E^{21} = 0$$

and E^{22} being derived from E^{11} by interchanging the indices on the b and c coefficients.

From (2.30) we have, using the Cauchy-Binet theorem (see Appendix I), the determinant result

$$\det D_{oo} = \lambda_o \det F_o^{(1)} \det E_o^{(2)} + \mu_o \det E_o^{(1)} \det F_o^{(2)} \tag{2.33}$$

where

$$F_0^{(i)} = \begin{bmatrix} E_0^{(i)} & X_0^{(i)} \\ u_0^{(i)} & S_0^{(i)} \end{bmatrix} \qquad (i=1,2)$$

and

$$S_0^{(1)} = c_{3m+1,0} \quad , \quad S_0^{(2)} = c_{0,3m+1}.$$

We can use (2.33) in two possible ways:

(i) Following (9), we note that we can always choose the ratio $\lambda_0 : \mu_0$ so as to make prong 0 of any approximant indeterminate; by this we mean that in (2.33) there is a unique value of the ratio $\lambda_0 : \mu_0$ which causes $\det D_{00}=0$, and results in degenerate equations. We seek to choose $\lambda_0 : \mu_0$ so that the opposite situation is produced and $|\det D_{00}|$ is a maximum; in this way we hope to move any spurious singularities of the approximants as far from the origin as possible. Maximising $|\det D_{00}|$ subject to

$$|\lambda_0|^2 + |\mu_0|^2 = 1$$

leads to the choice

$$\frac{\lambda_0}{\mu_0} = \frac{\det F_0^{(1)} \det E_0^{(2)}}{\det F_0^{(2)} \det E_0^{(1)}} \qquad (2.34)$$

For symmetric functions this reduces to $\lambda_0 = \mu_0$ as expected; the ratio $\lambda_0 : \mu_0$ for antisymmetric functions, however, is unclear.

(ii) Following (3) we seek to choose $\lambda_0 : \mu_0$ in order to produce scale covariance of the approximants. Let

$$f(\underline{z}) = \sum_{\underline{\nu}=\underline{0}}^{\underline{m}} c_{\underline{\nu}} \underline{z}^{\underline{\nu}}$$

and

$$f^2(\underline{z}) = \sum_{\underline{\nu}=\underline{0}}^{\underline{m}} b_{\underline{\nu}} \underline{z}^{\underline{\nu}}$$

We make the scale transformations

$$z_i = K_i Z_i \qquad (i=1,2) \qquad (2.35)$$

with

$$K_1 K_2 \neq 0$$

Define

$$g(Z_1, Z_2) = f(K_1 Z_1, K_2 Z_2)$$

$$= \sum_{\underline{\gamma}=\underline{0}}^{\underline{m}} \bar{c}_{\underline{\gamma}} \underline{Z}^{\underline{\gamma}}$$

where

$$\bar{c}_{\underline{\gamma}} = K_1^{\gamma_1} K_2^{\gamma_2} c_{\underline{\gamma}} \tag{2.36a}$$

Similarly,

$$g^2(\underline{Z}) = \sum_{\underline{\gamma}=\underline{0}}^{\underline{m}} \bar{b}_{\underline{\gamma}} \underline{Z}^{\underline{\gamma}}$$

with

$$\bar{b}_{\underline{\gamma}} = K_1^{\gamma_1} K_2^{\gamma_2} b_{\underline{\gamma}} \tag{2.36b}$$

It is clear that any unsymmetrised equations are scale covariant. For example, within the square lattice of Fig.4, we have the following system of equations:

$$\sum_{\sigma=0}^{\alpha} \sum_{\tau=0}^{\beta} \left[p_{\sigma\tau} b_{\alpha-\sigma,\beta-\tau} + q_{\sigma\tau} c_{\alpha-\sigma,\beta-\tau} \right] = -r_{\alpha\beta} \quad (0 \le \alpha,\beta \le m) \tag{2.37}$$

Under the scale transformation (2.35), the approximant will be invariant if the polynomials P, Q and R are invariant. Let

$$\bar{P}(Z_1,Z_2) = P(K_1 Z_1, K_2 Z_2) = \sum_{\underline{\alpha}=\underline{0}}^{\underline{m}} \bar{p}_{\underline{\alpha}} \underline{Z}^{\underline{\alpha}}$$

where

$$\bar{p}_{\underline{\alpha}} = K_1^{\alpha_1} K_2^{\alpha_2} p_{\underline{\alpha}} \tag{2.38}$$

together with similar relations for $\bar{q}_{\underline{\alpha}}$ and $\bar{r}_{\underline{\alpha}}$. Scale covariance requires

$$\sum_{\sigma=0}^{\alpha} \sum_{\tau=0}^{\beta} \left[\bar{p}_{\sigma\tau} \bar{b}_{\alpha-\sigma,\beta-\tau} + \bar{q}_{\sigma\tau} \bar{c}_{\alpha-\sigma,\beta-\tau} \right] = -\bar{r}_{\alpha\beta}$$

Using (2.38) this reduces to

$$\sum_{\sigma=0}^{\alpha} \sum_{\tau=0}^{\beta} \left[K_1^{\sigma} K_2^{\tau} K_1^{\alpha-\sigma} K_2^{\beta-\tau} p_{\sigma\tau} b_{\alpha-\sigma,\beta-\tau} + K_1^{\sigma} K_2^{\tau} K_1^{\alpha-\sigma} K_2^{\beta-\tau} q_{\sigma\tau} c_{\alpha-\sigma,\beta-\tau} \right] = -K_1^{\alpha} K_2^{\beta} r_{\alpha\beta}$$

which is just (2.37).

It is therefore only the symmetrised equations which prevent scale covariance. On prong O, the symmetrised equation determining q_{00} is

$$\lambda_0 \sum_{\beta_1=0}^{m} \left[p_{\beta_1,0} b_{3m+1-\beta_1,0} + q_{\beta_1,0} c_{3m+1-\beta_1,0} \right] + \mu_0 \sum_{\beta_2=0}^{m} \left[p_{0,\beta_2} b_{0,3m+1-\beta_2} + q_{0,\beta_2} c_{0,3m+1-\beta_2} \right] = 0 \tag{2.39}$$

In the scaled co-ordinates $\underline{Z}$, (2.39) becomes

$$\bar{\lambda}_0 K_1^{3m+1} \sum_{\beta_1=0}^{m} \left[p_{\beta_1,0}\, b_{3m+1-\beta_1,0} + q_{\beta_1,0}\, c_{3m+1-\beta_1,0} \right] + \bar{\mu}_0 K_2^{3m+1} \sum_{\beta_2=0}^{m} \Big[\tag{2.40}$$

$$p_{0,\beta_2}\, b_{0,3m+1-\beta_2} + q_{0,\beta_2}\, c_{0,3m+1-\beta_2} \Big] = 0$$

with the appropriate weight factors $\bar{\lambda}_0, \bar{\mu}_0$. From (2.39) and (2.40), scale covariance requires

$$\frac{\bar{\lambda}_0 K_1^{3m+1}}{\bar{\mu}_0 K_2^{3m+1}} = \frac{\lambda_0}{\mu_0} \tag{2.41}$$

We can guarantee that (2.41) holds by choosing $\lambda_0 : \mu_0$ such that the two terms in (2.33) are equal:

$$\frac{\lambda_0}{\mu_0} = \frac{\det E_0^{(1)} \det F_0^{(2)}}{\det E_0^{(2)} \det F_0^{(1)}} \tag{2.42}$$

(which is the reciprocal of the choice (2.34)). To see that (2.42) does indeed imply (2.41) we must look at the structure of the matrices $F_0^{(i)}$ and $E_0^{(i)}$ (i=1.2). Suppose that, under (2.35), $E_0^{(i)}$, $F_0^{(i)}$ and $S_0^{(i)}$ transform to $\bar{E}_0^{(i)}$, $\bar{F}_0^{(i)}$ and $\bar{S}_0^{(i)}$ respectively (i=1,2). The matrix structure then implies

$$\frac{\det \bar{F}_0^{(i)}}{\det \bar{E}_0^{(i)}} = \frac{\bar{S}_0^{(i)} \det F_0^{(i)}}{S_0^{(i)} \det E_0^{(i)}}$$

$$= K_i^{3m+1} \frac{\det F_0^{(i)}}{\det E_0^{(i)}} \tag{2.43}$$

In the scaled co-ordinates (2.42) becomes

$$\frac{\bar{\lambda}_0}{\bar{\mu}_0} = \frac{\det \bar{E}_0^{(1)} \det \bar{F}_0^{(2)}}{\det \bar{E}_0^{(2)} \det \bar{F}_0^{(1)}} \tag{2.44}$$

Thus, from (2.42)-(2.44),

$$\frac{\bar{\lambda}_0}{\bar{\mu}_0} = \frac{\det \bar{F}_0^{(2)}}{\det \bar{E}_0^{(2)}} \cdot \frac{\det \bar{E}_0^{(1)}}{\det \bar{F}_0^{(1)}}$$

$$= \frac{K_2^{3m+1}}{K_1^{3m+1}} \cdot \frac{\det F_0^{(2)}}{\det E_0^{(2)}} \cdot \frac{\det E_0^{(1)}}{\det F_0^{(1)}}$$

$$= \left(\frac{K_2}{K_1} \right)^{3m+1} \frac{\lambda_0}{\mu_0}$$

which is (2.41). Again the choice (2.42) reduces to $\lambda_0 = \mu_0$ for symmetric functions.

We finally give a brief discussion of the corresponding situation on prong m. For choice (i) (equation (2.8a)), since no symmetrisation is required, we can produce overall scale covariance with the choice (2.42). For choice (ii), we can write (2.8b) more generally as

$$e_{m,m} = 0 \quad \text{and} \quad \lambda_m e_{m+1,m} + \mu_m e_{m,m+1} = 0 \tag{2.45}$$

This gives

$$\begin{aligned} D_{mm} &= \lambda_m b_{1,0} + \mu_m b_{0,1} \\ &= 2c_{0,0}\left(\lambda_m c_{1,0} + \mu_m c_{0,1}\right) \end{aligned}$$

Maximizing $\left|\det D_{mm}\right|$ produces the choice

$$\frac{\lambda_m}{\mu_m} = \frac{c_{1,0}}{c_{0,1}} \tag{2.45a}$$

whilst scale covariance requires

$$\frac{\lambda_m}{\mu_m} = \frac{c_{0,1}}{c_{1,0}} \tag{2.45b}$$

For symmetric and antisymmetric functions (2.45a) and (2.45b) reduce to the expected results with $\lambda_m = \mu_m$ and $\lambda_m = -\mu_m$ respectively.

For choice (iii), we write (2.8c) more generally as

$$\lambda_1 e_{m+1,m} + \mu_1 e_{m,m+1} = 0 \tag{2.46}$$

and

$$\lambda_2 e_{m+2,m} + \mu_2 e_{m,m+2} = 0$$

This gives

$$\det D_{mm} = -\left(\lambda_1 c_{1,0} + \mu_1 c_{0,1}\right)\left(\lambda_2 c_{1,0}^2 + \mu_2 c_{0,1}^2\right)$$

If we choose

$$\lambda_2 = \mu_2 = 1$$

then the ratio $\lambda_1 : \mu_1$ reduces to (2.45a) or (2.45b). Choosing

$$\lambda_1 = \mu_1 = 1$$

we obtain as the maximised determinant choice

$$\frac{\lambda_2}{\mu_2} = \left(\frac{c_{1,0}}{c_{0,1}}\right)^2 \tag{2.46a}$$

and

$$\frac{\lambda_2}{\mu_2} = \left(\frac{c_{0,1}}{c_{1,0}}\right)^2 \tag{2.46b}$$

as the scale covariant choice. We note that neither (2.46a) nor (2.46b) give the expected result for antisymmetric functions.

3. DETERMINANT FORMULAE FOR, AND DEGENERACY OF, THE TWO VARIABLE DIAGONAL QUADRATIC APPROXIMANTS

(a) DETERMINANT FORMULAE

We consider the diagonal approximants of §2 subject to the conditions:

(i) $P_{0,0} = 1$

(ii) $\lambda_0 e_{3m+1,0} + \mu_0 e_{0,3m+1} = 0$ on prong 0, and

(iii) on prong m, the double symmetrisation

$$e_{m+1,m} + e_{m,m+1} = 0$$

and

$$\lambda_m e_{m+2,m} + \mu_m e_{m,m+2} = 0$$

For other choices of (i), (ii) and (iii) we obtain minor differences in the following results.

Define, for $0 \le \alpha \le m$ (where m denotes the order of the approximant)

$$\underline{d}_\alpha^T = \left(p_{m,\alpha}, \ldots, p_{\alpha+1,\alpha}; q_{m,\alpha}, \ldots, q_{\alpha+1,\alpha}; p_{\alpha,m}, \ldots, p_{\alpha,\alpha+1}; q_{\alpha,m}, \ldots, q_{\alpha,\alpha+1}; p_{\alpha,\alpha}, q_{\alpha,\alpha} \right) \quad (3.1)$$

Then the equations defining the approximants take the form

PRONG 0: $$D_{00}\,\underline{d}_0 = (0,0,\ldots,1,0,\ldots,0)^T \quad (3.2)$$

PRONG α: $$D_{\alpha\alpha}\,\underline{d}_\alpha + \sum_{\beta=0}^{\alpha-1} G_{\alpha\beta}\,\underline{d}_\beta = 0 \quad (0 < \alpha \le m) \quad (3.3)$$

where

(i) in (3.2) the entry 1 on the right-hand side occurs in the (2m+1)st. position, and

$$D_{00} = \begin{bmatrix} E_0^{(1)} & 0 & X_0^{(1)} & F_0^{(1)} & 0 & Y_0^{(1)} \\ 0 & E_0^{(2)} & X_0^{(2)} & 0 & F_0^{(2)} & Y_0^{(2)} \\ U_0^{(1)} & U_0^{(2)} & S_0^{(1)} & V_0^{(1)} & V_0^{(2)} & S_0^{(2)} \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$E_\alpha^{(1)} = \begin{bmatrix} b_{1,0} & \cdots & b_{m-\alpha,0} \\ \vdots & & \vdots \\ b_{2m-2\alpha+\theta(\alpha),0} & \cdots & b_{3m-3\alpha+\theta(\alpha)-1,0} \end{bmatrix}$$

$$F_\alpha^{(1)} = \begin{bmatrix} c_{1,0} & \cdots & c_{m-\alpha,0} \\ \vdots & & \vdots \\ c_{2m-2\alpha+\theta(\alpha),0} & \cdots & c_{3m-3\alpha+\theta(\alpha)-1,0} \end{bmatrix}$$

$$X_\alpha^{(1)T} = [b_{m+1-\alpha,0}, \ldots, b_{3m-3\alpha+1-\theta(1-\alpha),0}]$$

$$Y_\alpha^{(1)T} = [c_{m+1-\alpha,0}, \ldots, c_{3m-3\alpha+1-\theta(1-\alpha),0}]$$

$$U_0^{(1)} = \lambda_0 [b_{2m+1,0}, \ldots, b_{3m,0}]$$

$$U_0^{(2)} = \mu_0 [b_{0,2m+1}, \ldots, b_{0,3m}]$$

$$V_0^{(1)} = \lambda_0 [c_{2m+1,0}, \ldots, c_{3m,0}]$$

$$V_0^{(2)} = \mu_0 [c_{0,2m+1}, \ldots, c_{0,3m}]$$

$$S_0^{(1)} = \lambda_0 b_{3m+1,0} + \mu_0 b_{0,3m+1}$$

$$S_0^{(2)} = \lambda_0 c_{3m+1,0} + \mu_0 c_{0,3m+1}$$

and the remaining entries with superscript 2 and obtained from the corresponding entries with superscript 1 by interchanging the indices on the b and c coefficients,

(ii) $$D_{\alpha\alpha} = \begin{bmatrix} E_\alpha^{(1)} & 0 & X_\alpha^{(1)} & F_\alpha^{(1)} & 0 & Y_\alpha^{(1)} \\ 0 & E_\alpha^{(2)} & X_\alpha^{(2)} & 0 & F_\alpha^{(2)} & Y_\alpha^{(2)} \end{bmatrix}$$

The matrix $G_{\alpha\beta}$ is a little difficult to define concisely; we write it in the block form

$$G_{\alpha\beta} = \begin{bmatrix} G_{\alpha\beta}^{11} & H_{\alpha\beta}^{11} & G_{\alpha\beta}^{12} & H_{\alpha\beta}^{12} & G_{\alpha\beta}^{13} & H_{\alpha\beta}^{13} \\ G_{\alpha\beta}^{21} & H_{\alpha\beta}^{21} & G_{\alpha\beta}^{22} & H_{\alpha\beta}^{22} & G_{\alpha\beta}^{23} & H_{\alpha\beta}^{23} \end{bmatrix}$$

If we denote by $\mathfrak{M}(r\times s)$ the class of matrices with r rows and s columns, and label the elements of a matrix by i=1,2,..,r and j=1,2,..,s, then

$$G_{\alpha\beta}^{11} \in \mathfrak{M}(2m-2\alpha+1 \times m-\beta) \qquad (G_{\alpha\beta}^{11})_{ij} = b_{i+j-1,\alpha-\beta}$$

$$G_{\alpha\beta}^{12} \in \mathfrak{M}(2m-2\alpha+1 \times m-\beta) \qquad (G_{\alpha\beta}^{12})_{ij} = b_{m-\beta+i,\alpha-m-1+j}$$

$$G_{\alpha\beta}^{13} \in \mathfrak{M}(2m-2\alpha+1 \times 1) \qquad (G_{\alpha\beta}^{13})_{ij} = b_{m-\beta+i,\alpha-\beta}$$

$$G_{\alpha\beta}^{21} \in \mathfrak{M}(2m-2\alpha+1 \times m-\beta) \qquad (G_{\alpha\beta}^{21})_{ij} = b_{\alpha-m-1+j,m-\beta+i}$$

$$G_{\alpha\beta}^{22} \in \mathfrak{M}(2m-2\alpha+1 \times m-\beta) \qquad (G_{\alpha\beta}^{22})_{ij} = b_{\alpha-\beta,i+j-1}$$

$$G_{\alpha\beta}^{23} \in \mathfrak{M}(2m-2\alpha+1 \times 1) \qquad (G_{\alpha\beta}^{23})_{ij} = b_{\alpha-\beta,m-\beta+i}$$

$H_{\alpha\beta}^{kl}$ is obtained from $G_{\alpha\beta}^{kl}$ by replacing the b coefficients by c coefficients, and

(iii)

$$D_{mm} = \begin{bmatrix} b_{1,0} + b_{0,1} & c_{1,0} + c_{0,1} \\ \lambda_m b_{2,0} + \mu_m b_{0,2} & \lambda_m c_{2,0} + \mu_m c_{0,2} \end{bmatrix}$$

$$G_{m\beta} = \begin{bmatrix} G^{11}_{m\beta} + G^{21}_{m\beta} & H^{11}_{m\beta} + H^{21}_{m\beta} & G^{12}_{m\beta} + G^{22}_{m\beta} & H^{12}_{m\beta} + H^{22}_{m\beta} \\ \lambda_m \bar{G}^{11}_{m\beta} + \mu_m \bar{\bar{G}}^{21}_{m\beta} & \lambda_m \bar{H}^{11}_{m\beta} + \mu_m \bar{\bar{H}}^{21}_{m\beta} & \lambda_m \bar{G}^{12}_{m\beta} + \mu_m \bar{\bar{G}}^{22}_{m\beta} & \lambda_m \bar{H}^{12}_{m\beta} + \mu_m \bar{\bar{H}}^{22}_{m\beta} \\ & & G^{13}_{m\beta} + G^{23}_{m\beta} & H^{13}_{m\beta} + H^{23}_{m\beta} \\ & & \lambda_m \bar{G}^{13}_{m\beta} + \mu_m \bar{\bar{G}}^{23}_{m\beta} & \lambda_m \bar{H}^{13}_{m\beta} + \mu_m \bar{\bar{H}}^{23}_{m\beta} \end{bmatrix}$$

where $\bar{G}$) indicates that +1 should be added to the 1st.) index (subscript)
$\bar{\bar{G}}$) 2nd.)
G)

of G.

For choices other than (iii) (of §3(a)) on prong m, $G_{m\beta}$ assumes a considerably simpler form.

To give an explicit determinant representation for P,Q and R we must first separate the contributions arising from the b and c coefficients. We use, for example, the notation $D_{\alpha\alpha}(b)$ to represent that part of $D_{\alpha\alpha}$ containing only the b coefficients:

$$D_{\alpha\alpha}(b) = \begin{bmatrix} E^{(1)}_\alpha & 0 & X^{(1)}_\alpha \\ 0 & E^{(2)}_\alpha & X^{(2)}_\alpha \end{bmatrix} \qquad (0 < \alpha < m)$$

and we also define $\tilde{D}_{00}$ to be D_{00} minus the last row. We can then write the system to be solved in the matrix form

$$\begin{bmatrix} \tilde{D}_{00}(b) & & 0 & \tilde{D}_{0c}(c) & & 0 \\ G_{10}(b) & D_{11}(b) & & G_{10}(c) & D_{11}(c) & \\ \vdots & \ddots & & \vdots & \ddots & \\ G_{m0}(b) & G_{m1}(b) & D_{mm}(b) & G_{m0}(c) & G_{m1}(c) & D_{mm}(c) \end{bmatrix} \begin{bmatrix} \underline{d}_0(p) \\ \vdots \\ \underline{d}_m(p) \\ \underline{d}_0(q) \\ \vdots \\ \underline{d}_m(q) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

where

$$\underline{d}_\alpha(p) = (p_{m,\alpha}, \ldots\ldots, p_{\alpha+1,\alpha}; p_{\alpha,m}, \ldots\ldots, p_{\alpha,\alpha+1}; p_{\alpha,\alpha})$$

is the component of $\underline{d}_\alpha$ (defined by (3.1)) containing the p coefficients (of $P(\underline{z})$) and $\underline{d}_\alpha(q)$ the component containing the q coefficients.

We now finally have the definitions

$$\tilde{D}'_{00} = \tilde{D}_{00} \text{ minus the } [X^{(1)}_0 \; X^{(2)}_0 \; S^{(1)}_0] \text{ column}$$

$$F = \det \tilde{D}'_{00} \prod_{\alpha=1}^{m} \det D_{\alpha\alpha}$$

and, for $0 \le \alpha \le m$,

$$\underline{z}_\alpha = \left(x^m y^\alpha, \ldots, x^{\alpha+1} y^\alpha, x^\alpha y^m, \ldots, x^\alpha y^{\alpha+1}, x^\alpha y^\alpha\right)$$

$$\underline{\underline{z}}^{(1)}_\alpha = \left(\zeta_{m\alpha}, \ldots, \zeta_{\alpha+1,\alpha}, \zeta_{\alpha m}, \ldots, \zeta_{\alpha,\alpha+1}, \zeta_{\alpha\alpha}\right)$$

$$\underline{\underline{z}}^{(2)}_\alpha = \left(\eta_{m\alpha}, \ldots, \eta_{\alpha+1,\alpha}, \eta_{\alpha m}, \ldots, \eta_{\alpha,\alpha+1}, \eta_{\alpha\alpha}\right)$$

where

$$\zeta_{\sigma,\tau} = \sum_{\mu=\sigma}^{m} \sum_{\nu=\tau}^{m} b_{\mu-\sigma,\nu-\tau}\, x^\mu y^\nu$$

and

$$\eta_{\sigma,\tau} = \sum_{\mu=\sigma}^{m} \sum_{\nu=\tau}^{m} c_{\mu-\sigma,\nu-\tau}\, x^\mu y^\nu$$

Then

$$P(x,y) = P^D(x,y)/F$$

$$Q(x,y) = Q^D(x,y)/F$$

$$R(x,y) = R^D(x,y)/F$$

and

$$f_{(\underline{m}/\underline{m}/\underline{m})}(x,y) = \frac{-Q^D(x,y) \pm \left[(Q^D(x,y))^2 - 4P^D(x,y)R^D(x,y)\right]^{\frac{1}{2}}}{2P^D(x,y)}$$

where

$$P^D(x,y) = \det \begin{vmatrix} \tilde{D}_{00}(b) & & & \tilde{D}_{00}(c) & & \\ G_{10}(b) & D_{11}(b) & 0 & G_{10}(c) & D_{11}(c) & 0 \\ \vdots & \ddots & & \vdots & \ddots & \\ G_{m0}(b) & \cdots & D_{mm}(b) & G_{m0}(c) & \cdots & D_{mm}(c) \\ \underline{z}_0 & \cdots & \underline{z}_m & 0 & \cdots & 0 \end{vmatrix}$$

$$Q^D(x,y) = \det \begin{vmatrix} \tilde{D}_{00}(b) & & & \tilde{D}_{00}(c) & & \\ G_{10}(b) & D_{11}(b) & 0 & G_{10}(c) & D_{11}(c) & 0 \\ \vdots & \ddots & & \vdots & \ddots & \\ G_{m0}(b) & \cdots & D_{mm}(b) & G_{m0}(c) & \cdots & D_{mm}(c) \\ 0 & \cdots & 0 & \underline{z}_0 & \cdots & \underline{z}_m \end{vmatrix}$$

and

$$R^{p}(x,y) = -\det \begin{vmatrix} \tilde{D}_{00}(b) & & & \tilde{D}_{00}(c) & & \\ G_{10}(b) & D_{11}(b) & 0 & G_{10}(c) & D_{11}(c) & 0 \\ \vdots & \ddots & & \vdots & \ddots & \\ G_{m0}(b) & \cdots\cdots & D_{mm}(b) & G_{m0}(c) & \cdots\cdots & D_{mm}(c) \\ Z_0^{(1)} & \cdots\cdots & Z_m^{(1)} & Z_0^{(2)} & \cdots\cdots & Z_m^{(2)} \end{vmatrix}$$

(b) DEGENERACIES IN THE APPROXIMANTS

The two variable diagonal quadratic approximants will always exist provided

$$F = \prod_{\alpha=0}^{m} \det D_{\alpha\alpha} \neq 0$$

Approximants with F=0 are termed degenerate; degeneracy occurs when

(i) $\det D'_{00} = 0$ or

(ii) $\det D_{\alpha\alpha} = 0 \quad (1 \leq \alpha \leq m)$

In both (i) and (ii) the equations can be either consistent or inconsistent. When the equations are inconsistent, the approximant does not exist. For consistent equations, degeneracy occurs when a polynomial factor is common to P(x,y), Q(x,y) and R(x,y). Such a situation occurs when f(z) is a symmetric function; to see this, consider the (2/2/2) approximant on prong 1, since it is only a degeneracy of type (ii), on the internal prongs, which occurs in this situation. On this prong we obtain the two systems (both comprising three equations)

$$F_1(p_{21}, q_{21}, p_{11}, q_{11}) = 0 \tag{3.4a}$$

and

$$F_2(p_{12}, q_{12}, p_{11}, q_{11}) = 0 \tag{3.4b}$$

Now if f(z) is symmetric, these two equations will imply that

$$p_{21} = p_{12} \text{ and } q_{21} = q_{12}$$

so that the two equations effectively reduce to the single system

$$F_1(p_{21}, q_{21}, p_{11}, q_{11}) = 0$$

However, this represents a system of three equations in four unknowns; this gives rise to degenerate but consistent equations. This example does seem to illustrate a rather undesirable feature of the diagonal approximants, for

we would not expect symmetric functions to cause any type of degeneracy. Theoretically, this presents no problem since it merely produces an arbitrary polynomial factor common to $P(\underline{z})$, $P(\underline{z})$ and $R(\underline{z})$, and hence does not alter the value of the approximant. This difficulty is encountered on all internal prongs; no such difficulty occurs on the first and last prong because these prongs involve symmetrised equations, which relate (3.4a) and (3.4b). For this same reason the above difficulty is not encountered with the two variable rational approximants.

A possible practical solution of this difficulty (12) is, for the approximant considered above, to obtain an extra equation by setting

$$r_{11}=0 \tag{3.5a}$$

and retaining the equation

$$e_{11}=0 \tag{3.5b}$$

Equation (3.5b) does not now determine r_{11} but instead provides a relation between p_{11} and q_{11}. Alternatively, we can sidestep this difficulty and, for example, instead of calculating the (2,2/2,2/2,2) diagonal approximant we could calculate the (2,2/2,1/2,2) off-diagonal approximant, with the hope that this off-diagonal approximant will not greatly violate the homographic covariance property of the diagonal approximants. In fact, in §6 we shall adopt this point of view when dealing with symmetric functions.

Although we have not yet defined the off-diagonal two variable quadratic approximants, it is convenient to illustrate here a type of degeneracy which can occur with these approximants. In (9) it was noted that the genuine two variable rational approximants behave differently from the one variable Padé approximants. The example given to illustrate this behaviour is

$$f(x,y)=1-x+x^2-x^3+c_{01}y+c_{11}xy+\cdots\cdots$$

It is easily shown that

$$f_{(1,1/2,0)}(x,y)=\frac{[1+(b_{10}-1)x][1+c_{01}y]+xy(c_{11}+c_{01})}{(1+x)[1+(b_{10}-1)x]}$$

where b_{10} is arbitrary; this means that, unless $c_{11}+c_{01}=0$, the approximant

can take almost any value. This type of behaviour cannot occur for the one variable Padé approximants.

A similar type of degeneracy can occur for the two variable quadratic approximants. If we let

$$f(x,y) = 1 - x + x^2 - x^3 + x^4 + c_{01}y + c_{11}xy + y^2$$

we find that

$$f_{(2,0/0,0/1,1)}(x,y) = \frac{-Q(x,y) \pm \left[Q^2(x,y) - 4P(x,y)R(x,y)\right]^{\frac{1}{2}}}{2P(x,y)}$$

where

$$P(x,y) = (1+x)(1+(p_{10}-1)x)$$

$$Q(x,y) = -(2 + c_{01}^2)$$

$$R(x,y) = (1 + c_{01}^2) - (p_{10} + c_{01}^2)x + c_{01}^3 y + \left[2c_{01} + c_{11}c_{01}^2 - 2c_{01}p_{10}\right]xy$$

Again the approximant can take almost any value for p_{10} arbitrary.

We can in fact overcome this difficulty by employing the hypothesis of maximum analyticity lying behind all Padé type methods; we fix the parameters, b_{10} and p_{10} above, by maximizing the distance of the nearest moveable singularity (that is, a singularity of the approximant not required to simulate a singularity of the function) from the origin. The choices

$$b_{10} = p_{10} = 1$$

both send this singularity to infinity.

4. EXTENSION TO ARBITRARY "t-POWER APPROXIMANTS" (6)

The scheme of §2 for quadratic approximants can be fairly easily extended to arbitrary t(>2)-power approximants. For diagonal two variable cubic approximants, we define the four polynomials $P(\underline{z})$, $Q(\underline{z})$, $R(\underline{z})$ and $S(\underline{z})$, assuming that

$$P(0) = p_{00}\ ,\ Q(0) = q_{00} \text{ and } R(0) = r_{00}$$

are known. We then equate to zero certain coefficients in

$$E(\underline{z}) = P(\underline{z})f^3(\underline{z}) + Q(\underline{z})f^2(\underline{z}) + R(\underline{z})f(\underline{z}) + S(\underline{z}) \qquad (4.1)$$

In Fig.5 the prong structure is indicated. Prong 0 extends to the points (4m,0) and (0,4m); we can determine q_{00} and r_{00} (assuming we choose $p_{00}=1$) by extending the prong to include the points (circled in Fig.5) (4m+1,0) and

$(0,4m+1)$. An advantage of this scheme is that these approximants will project to the (normal) Shafer approximants. On prong 1 there are $3(2m-1)$ unknown variables $\{p_{\underline{\alpha}}\}$, $\{q_{\underline{\alpha}}\}$ and $\{r_{\underline{\alpha}}\}$; these are determined by the equations

$$e_{\underline{\epsilon}} = 0$$

with

$$\underline{\epsilon} = (m+1,1), \ldots\ldots\ldots, (4m-2,1)$$

and

$$\underline{\epsilon} = (1,m+1), \ldots\ldots\ldots, (1,4m-2)$$

together with the symmetrised equation

$$e_{4m-1,1} + e_{1,4m-1} = 0$$

Prong 2 contains six fewer variables, so that each arm of the prong is three units shorter than in Prong 1, together with a symmetrised equation. This pattern continues till prong $(m-1)$ is reached. The natural scheme of determining p_{mm}, q_{mm} and r_{mm} by unsymmetrised equations at $(m+1,m)$ and $(m,m+1)$, and a symmetrised equation from $(m+2,m)$ and $(m,m+2)$, produces two linearly dependent equations corresponding to the unsymmetrised points. This situation is similar to that encountered with the quadratic approximants of §2, and we again have three possible methods of avoiding this degeneracy, whilst still preserving symmetry:

(i) Omit prong m, setting

$$P_{mm} = q_{mm} = r_{mm} = s_{mm} = 0$$

Since the rectangle rule is obeyed, the approximants will satisfy reciprocal covariance and covariance under the restricted homographic group (1.12).

(ii) Set

$$q_{mm} = r_{mm} = 0$$

and determine p_{mm} and s_{mm} as in (2.8b). Reciprocal and homographic covariance under (1.12) are both ensured.

(iii) Use (2.8b) together with

$$e_{m+2,m} + e_{m,m+2} = 0$$

As with quadratic approximants, this choice weakens the covariance properties.

By a suitable choice of weighting factors we may be able to ensure homographic covariance under (1.13).

Comparison with §2 shows that there are certain characteristic differences between the two variable approximants formed for even and odd t. We can now give the basic definitions and general properties for t-power approximants; since prong m will always produce inconsistencies if the points (m+1,m) and (m,m+1) are used (for all values of t), we must make a specific choice of the schemes available on prong m. For simplicity, we choose to omit prong m (scheme (i) above); other choices on prong m will produce the appropriate modifications in the properties listed below.

GENERAL PROPERTIES OF t-POWER APPROXIMANTS

(i) With the t constants $P(\underline{0})$, $Q(\underline{0})$,... given, $[E(\underline{0})]_{t=0}$ is determined by prong zero, which extends to the point (t+1)m along each axis.

(ii) The approximants project to modified Shafer approximants when either variable is equated to zero.

(iii) For all internal prongs (that is, all prongs excluding prongs 0 and m), the number of points on each arm of a prong decreases successively by t.

(iv) Reciprocal covariance is satisfied, provided the constants in (i) are chosen appropriately.

PROPERTIES FOR t EVEN

(v) Choosing $P(\underline{0}) = 1$ (which we can normally do), the (t-1) constants $Q(\underline{0})$, $R(\underline{0})$,... can be determined by extending prong 0 by (t-2)/2 points along each arm, and incorporating a symmetrised equation on the next two points. This final equation violates the projection property to the one variable approximants, but reciprocal covariance is preserved.

(vi) The sections of prong r ($1 \leq r < m$) lying outside the square $(\underline{0},\underline{m})$ determine the t(2m-2r-1) new variables arising; each arm thus consists of $\frac{1}{2}t(2m-2r-1)$ points. Assuming the constants $Q(\underline{0}), R(\underline{0}),...$ are given, full homographic covariance is preserved.

PROPERTIES FOR t ODD

(v) Choosing $P(\underline{0}) = 1$, the $(t-1)$ constants $Q(\underline{0})$, $R(\underline{0}),\ldots$ can be determined by extending prong 0 by $\frac{1}{2}(t-1)$ points along each axis, with no symmetrisation being required. Such an extension preserves the projection and reciprocal covariance properties. This would seem to suggest that, when the problem itself suggests no natural approximant to use, odd t-power approximants may be more useful than the corresponding even-t approximants.

(vi) The sections of prong $r\,(1\leq r<m)$ outside the square $(\underline{0},\underline{m})$ determine $t(2m-2r-1)$ new variables; since this number is odd, symmetrisation is required. The arm of each prong consists of $\frac{1}{2}(t(2m-2r-1)-1)$ points, together with the symmetrised equation obtained from the next point on each arm. With equal weights associated with these symmetrised equations, homographic covariance is restricted by $A_1=A_2$ in (1.12), but full covariance may be possible with suitable weight factors.

We finally note that, as mentioned in §2(b), the possibility of including information on more than one sheet of the function extends naturally to arbitrary t-power approximants. The method of fixing the constants $Q(\underline{0})$, $R(\underline{0}),\ldots$ by using the value of the function at the origin on $(t-1)$ sheets has the advantage of preserving the projection property. These considerations are certainly worth further study.

5. TWO VARIABLE t-POWER OFF-DIAGONAL APPROXIMANTS

The extension of t-power diagonal approximants to off-diagonal approximants has been given in (10), and is based upon the off-diagonal extension (5) of the diagonal rational approximants of §1.

We define the formal kth. power of (1.1) by

$$f^{(k)}(\underline{z}) = \sum_{\underline{\alpha}=\underline{0}}^{\infty} c_{\underline{\alpha}}^{(k)}\, \underline{z}^{\underline{\alpha}} \tag{5.1}$$

The $(t+1)$ polynomials, denoted by $P^{(k)}(\underline{z})$, required to define the off-diagonal approximant are defined by

$$P^{(k)}(\underline{z}) = \sum_{\underline{\alpha}\in S_k} P_{\underline{\alpha}}^{(k)}\, \underline{z}^{\underline{\alpha}} \qquad (k=0,1,\ldots,t) \tag{5.2}$$

where S_k denotes the rectangular lattice

$$S_K = \left\{ \underline{\alpha} \mid 0 \le \alpha_i \le m_{i;k} \ , i=1,2 \right\} \tag{5.3}$$

Here $m_{1;k}$ and $m_{2;k}$ denote the maximum powers of z_1 and z_2 respectively in $P^{(k)}(\underline{z})$. The coefficients ($P^{(k)}$) in (5.2) are determined by equating to zero certain coefficients, or linear combinations of coefficients, in

$$E(\underline{z}) = \sum_{k=1}^{t} P^{(k)}(\underline{z}) f^{(k)}(\underline{z}) + P^{(0)}(\underline{z}) \tag{5.4}$$

The coefficient of $\underline{z}^{\underline{\epsilon}}$ in (5.4) is

$$e_{\underline{\epsilon}} = \sum_{k=1}^{t} \left[\sum_{\underline{\alpha} \in S_K \cap S_{\underline{\epsilon}}} P_{\underline{\alpha}}^{(k)} c_{\underline{\epsilon}-\underline{\alpha}}^{(k)} \right] + P_{\underline{\epsilon}}^{(0)} \delta(\underline{\epsilon} \in S_0) \tag{5.5}$$

where

$$\delta(\underline{\epsilon} \in S_0) = \begin{cases} 1 \ , & \underline{\epsilon} \in S_0 \\ 0 \ , & \underline{\epsilon} \notin S_0 \end{cases}$$

and $S_{\underline{\epsilon}}$ denotes the rectangular lattice

$$S_{\underline{\epsilon}} = \left\{ \underline{\alpha} \mid 0 \le \alpha_i \le \epsilon_i \ ; i=1,2 \right\} \tag{5.6}$$

As in §2, the definition, and method of solution, of the linear system of equations defining the $P^{(k)}(\underline{z})$ is based upon the prong structure of (2); "prong σ", emanating from the point (σ,σ), is defined to be the set of lattice points

$$\pi_\sigma = \left\{ \underline{\alpha} \mid \alpha_1 = \sigma, \alpha_2 > \sigma \right\} \cup \left\{ \underline{\alpha} \mid \alpha_1 \ge \sigma, \alpha_2 = \sigma \right\} \tag{5.7}$$

where $\sigma \ge 0$.

Consider prong 0, denoted by π_0. We assume that the coefficients

$$\left\{ P_{\underline{\alpha}}^{(k)} \ ; k=0,1,\ldots,t \right\} \tag{5.8a}$$

are given, except for $P_{\underline{\alpha}}^{(0)}$, which is to be determined by π_0. The remaining coefficients occuring in

$$\left\{ e_{\underline{\epsilon}} : \underline{\epsilon} \in \pi_0 \right\}$$

are

$$\left\{ P_{\underline{\alpha}}^{(k)} \mid k=0,1,\ldots,t \ ; \underline{\alpha} \in \pi_0 \ , \underline{\alpha} \neq 0 \right\} \tag{5.8b}$$

The number, n_0, of unknown coefficients on π_0 is

$$n_0 = \sum_{i=1}^{2} \sum_{k=0}^{t} m_{i;k} + 1 \tag{5.9}$$

These n_0 unknown coefficients are determined by the equations

$$e_{\underline{\epsilon}} = 0 \tag{5.10}$$

for the set

$$\left\{ \underline{\epsilon} \;\middle|\; \underline{\epsilon} \subset S_0 , \epsilon_i \leq \sum_{k=0}^{t} m_{i;k} \right\} \quad (i=1,2) \tag{5.11}$$

Now consider a general prong π_σ with vertex (σ,σ) lying in the region

$$S = S_0 \cup S_1 \cup \cdots\cdots \cup S_t \tag{5.12}$$

This set extends the class of prongs of (5); in Figs.3 and 3.2 of (5), we allow prong vertices in both S_1 and S_5. In dealing with the off-diagonal t-power approximants we have in general (t+1) distinct regions to consider, as compared to the simpler two region structure of the off-diagonal rational approximants; use of the above extended class of prongs leads to a uniform treatment of the equations in each of the above (t+1) regions. The vertex (σ,σ) will, in general lie in a subset of S; we denote this subset by

$$\left\{ \bigcup S_k \;\middle|\; k \subset R_\sigma \right\} \tag{5.13}$$

where R_σ denotes the corresponding subset of $(0,1,\ldots,t)$. On prong σ the new coefficients introduced are

$$\left(p^{(k)}_{\sigma,\sigma} ; p^{(k)}_{\sigma+1,\sigma}, \cdots\cdots, p^{(k)}_{m_{1;k},\sigma} ; p^{(k)}_{\sigma,\sigma+1}, \cdots\cdots, p^{(k)}_{\sigma,m_{2;k}} \right) \tag{5.14}$$

where $k \subset R_\sigma$. The number, n_{π_σ}, of these new coefficients is

$$n_{\pi_\sigma} = \sum_{k \subset R_\sigma} \left(m_{1;k} + m_{2;k} - 2\sigma + 1 \right)$$

If n_σ denotes the number of integers in R_σ, that is, the number of regions S_k containing (σ,σ), then

$$n_{\pi_\sigma} = \sum_{k \subset R_\sigma} \left(m_{1;k} + m_{2;k} \right) - n_\sigma \left(2\sigma - 1 \right) \tag{5.15}$$

The expressions $e_{\underline{\epsilon}}$ corresponding to points on prong σ must therefore provide n_{π_σ} equations; below we see that this necessitates the introduction of symmetrised equations when n_σ is even, but not when n_σ is odd. Specifically:

$\underline{n_\sigma \text{ ODD}}$: The equations (5.10) have $\underline{\epsilon}$ ranging over the following points:

$$(\sigma,\sigma), \text{ the prong vertex;} \tag{5.16a}$$

$$(\sigma+1,\sigma), \cdots\cdots, \left(\sum_{k \subset R_\sigma} m_{1;k} - \left[\sigma - \tfrac{1}{2}\right]\left[n_\sigma - 1\right], \sigma \right) \tag{5.16b}$$

$$(\sigma,\sigma+1), \cdots\cdots, \left(\sigma, \sum_{k \subset R_\sigma} m_{2;k} - \left[\sigma - \tfrac{1}{2}\right]\left[n_\sigma - 1\right] \right) \tag{5.16c}$$

The number of these points is clearly n_{π_σ}.

n_σ EVEN: The equations (5.10) have $\underline{\xi}$ ranging over the following points:

$$(\sigma,\sigma), \text{ the prong vertex;} \tag{5.17a}$$

$$(\sigma+1,\sigma),\ldots\ldots,\left(\sum_{k\in R_\sigma} m_{1;k} - \tfrac{1}{2} - \left[\sigma-\tfrac{1}{2}\right]\left[n_\sigma-1\right],\ \sigma\right) \tag{5.17b}$$

$$(\sigma,\sigma+1),\ldots\ldots,\left(\sigma,\ \sum_{k\in R_\sigma} m_{2;k} - \tfrac{1}{2} - \left[\sigma-\tfrac{1}{2}\right]\left[n_\sigma-1\right]\right) \tag{5.17c}$$

The number of these points is $n_{\pi_\sigma}-1$, so that we require one more equation; this is provided by the symmetrised equation

$$e_{\alpha_1(\sigma),\sigma} + e_{\sigma,\alpha_2(\sigma)} = 0 \tag{5.18a}$$

where

$$\alpha_1(\sigma) = \sum_{k\in R_\sigma} m_{1;k} + \tfrac{1}{2} - \left(\sigma-\tfrac{1}{2}\right)\left(n_\sigma-1\right) \tag{5.18b}$$

and

$$\alpha_2(\sigma) = \sum_{k\in R_\sigma} m_{2;k} + \tfrac{1}{2} - \left(\sigma-\tfrac{1}{2}\right)\left(n_\sigma-1\right) \tag{5.18c}$$

In (5.18a) we do not include any arbitrary weighting factors; here we do not consider the problem of choosing weighting factors.

We illustrate the prong structures of this section by considering the specific cases t=1, 2 and 3.

t=1: This case corresponds to rational approximants; in this case there are only two rectangles S_o and S_1 of the form (5.3), as shown in Fig.6. Prong σ_1 has vertex in $S_o \wedge S_1$, so that $n_\sigma=2$ and a symmetrised equation (denoted by two crosses) is required on σ_1. Increasing σ by one within this region, $\alpha_1(\sigma)$ and $\alpha_2(\sigma)$ in (5.18) decrease by one, so that the symmetrised points lie on lines with unit slope. The remaining prong shown lies in S_o, but not in S_1, so that $n_\sigma=1$ and no symmetrised equation is necessary on this prong; the end points of the prong are $(m_{1;\sigma},\sigma)$ and $(\sigma,m_{2;\sigma})$, which lie on the edges of S_o, and we say that these lines have "zero slope".

t=2: For the two variable quadratic approximants there are three rectangles (Fig.7) S_o, S_1 and S_2. Successive prongs have $n_\sigma=3,3,2$ and 1, so that only in one instance is there a symmetrised equation (except for the

optional symmetrisation on prong zero, as discussed later).

t=3: The prong structure shown in Fig.8, together with Figs.6 and 7, illustrates the general prong structure for arbitrary t. The vertices of the prongs π_σ occur in groups with n_σ fixed within each group; n_σ takes the values $1,2,\ldots,(t+1)$. The prongs with $n_\sigma=1$ cover a rectangle in the corner of one of the rectangles (5.3); in this region, the edges have "zero slope" and there are no symmetrised equations. The prongs having $n_\sigma=2$ define a region with edges of unit slope, and each prong has an attached symmetrised equation. As n_σ increases to (t+1), the slope increases to t(=3 in Fig.8). Alternatively the edges do and do not correspond to sets of symmetrised equations, depending upon the parity of n_σ.

Having illustrated the general prong structure associated with t-power off-diagonal approximants, we now consider how to choose the constants (5.8a). As with the diagonal approximants of §2, two possible choices naturally suggest themselves:

(a) If the only information available about $f(\underline{z})$ is that contained in (1.1), then we can obtain the required (t-1) extra equations by extending prong 0. For t odd, $\frac{1}{2}(t-1)$ unsymmetrised equations are produced on each arm of the prong, which thus extends to the points on each axis with co-ordinates

$$\sum_{k=1}^{t} m_{i;k} + \tfrac{1}{2}(t-1) \qquad (i=1,2)$$

This set of points ensures the projection property to the Shafer approximants.

For t even, we obtain unsymmetrised equations corresponding to points on the axes up to those with co-ordinates

$$\sum_{k=1}^{t} m_{i;k} + \tfrac{1}{2}(t-2) \qquad (i=1,2)$$

together with a symmetrised equation involving the next point on each axis. This choice, as for diagonal approximants, violates the projection property.

(b) If $f(\underline{z})$ has an analytic continuation to at least (t-1) distinct Riemann sheets, we may fix the constants (5.8a) by using the function values

$f_1(\underline{0}),\ldots,f_{t-1}(\underline{0})$ at the origin of the (t-1) sheets, provided these values are distinct and not equal to $f(\underline{0})$. The extra (t-1) equations can then be written

$$\sum_{k=1}^{t} P_{\underline{0}}^{(k)} \left[f_r(\underline{0})\right]^k + P_{\underline{0}}^{(0)} = 0 \quad (r=1,\ldots,t-1)$$

and such a choice satisfies the projection property.

Whichever choice ((a) or (b)) is made, the approximant $f(\underline{z};S_k)$ defined from $f(\underline{z})$, and corresponding to the regions of (5.3), is defined to be a solution of the equation

$$\sum_{k=1}^{t} P^{(k)}(\underline{z}) \left[f(\underline{z};S_k)\right]^k + P^{(0)}(\underline{z}) = 0 \qquad (5.19)$$

In practice, the approximants of this section can be determined from matrix equations of the form (1.8)-(1.11). At the end of this chapter we give a computer program, written in Fortran, for generating off-diagonal two variable quadratic approximants. The numerical results of §6 are obtained from this program.

We conclude this section with the following remarks:

(i) The linear dependence arising on the final prong for diagonal t-power approximants does not occur for the truely off-diagonal approximants.

(ii) The reciprocal and homographic covariance properties of the off-diagonal approximants can be discussed in exactly the same as for the diagonal approximants of §2; for this reason we do not repeat the arguments of §2(b) for the off-diagonal approximants of this section.

(iii) The generalization of diagonal and off-diagonal approximants to the case of N>2 variables has been discussed in (6) and (10). Here we do not repeat these discussions and we merely remark that only for N≤3 can fully symmetrised schemes be defined for diagonal and off-diagonal approximants. In Fig.9 we illustrate the prong structure for three variable diagonal approximants.

6. NUMERICAL EXAMPLES

The previous sections of this chapter have been devoted to the

definition, method of construction and general properties of the two variable generalizations of the approximants of Chapter 1. In this section, by considering specific numerical examples, we attempt to illustrate the usefulness of these two variable approximation schemes. The examples we consider are of two types:

(a) EXAMPLES (I)

These examples consist of the following five functions with branch points, the first three of which are symmetric:

(i) $f_1(x,y) = (1-x-y)^{1/3}$

(ii) $f_2(x,y) = (1-x-y)^{1/4}$

(iii) $f_3(x,y) = 1 + \ln(1-x-y)$

(iv) $f_4(x,y) = 1 + \frac{1}{2}x + \ln(1-x-y)$

(v) $f_5(x,y) = \ln(1-x-y)/(2.5-y)$

In Tables 1-5 we tabulate the results obtained for the above functions using the two variable rational and quadratic approximants. For each table the third and fourth column provide a comparison between the two approximation schemes, both tabulated approximants requiring roughly the same number of coefficients of the series (1.1); the final column is included to indicate the degree of accuracy obtainable on the branch cut of the function. When no entry occurs in a table, this indicates that at this particular point the approximated function is complex-valued and that the single-valued (real) rational approximant cannot represent the discontinuity across the branch cut of the function. For the diagonal quadratic approximants of Tables 4 and 5 the scheme (2.8c) is used on the final prong.

From the results of Tables 1-5 one general point emerges; the quadratic approximants approximate, often very accurately, well beyond the radius of convergence of the series (1.1). More specifically:

(i) The functions f_3 and f_4 are infinitely many-valued, but their series expansions (1.1) correspond to a single Riemann sheet. The quadratic approximant is consistently more accurate than the rational approximant, in some cases four orders of magnitude more accurate. At points "on the branch

TABLE I: Diagonal rational and off-diagonal quadratic approximants to $f_1(x,y) = (1-x-y)^{1/3}$

x	y	Error in $[6,6/6,6]$ rational approximant	Error in $[4,4/4,3/4,4]$ quadratic approximant	Error in $[6,6/6,5/6,6]$ quadratic approximant
-3	-4	2×10^{-1}	5×10^{-6}	2×10^{-6}
-1		1×10^{-4}	1×10^{-6}	3×10^{-9}
1		2×10^{-4}	1×10^{-7}	$< 10^{-9}$
3		2×10^{-3}	$< 10^{-9}$	$< 10^{-9}$
-3	-2	6×10^{-5}	7×10^{-7}	2×10^{-9}
-1		6×10^{-8}	4×10^{-8}	$< 10^{-9}$
1		4×10^{-6}	$< 10^{-9}$	$< 10^{-9}$
3		4×10^{-1}	$(8 \times 10^{-2}, 8 \times 10^{-2})$	$(6 \times 10^{-2}, 5 \times 10^{-2})$
-3	0	1×10^{-6}	4×10^{-8}	$< 10^{-9}$
-1		$< 10^{-9}$	$< 10^{-9}$	$< 10^{-9}$
1		2×10^{-1}	$(8 \times 10^{-2}, 8 \times 10^{-2})$	$(5 \times 10^{-2}, 6 \times 10^{-2})$
3			$(2 \times 10^{-4}, 3 \times 10^{-4})$	$(5 \times 10^{-6}, 6 \times 10^{-7})$
-3	2	2×10^{-4}	$< 10^{-9}$	$< 10^{-9}$
-1		3×10^{-1}	$(8 \times 10^{-2}, 7 \times 10^{-2})$	$(5 \times 10^{-2}, 5 \times 10^{-2})$
1			$(2 \times 10^{-4}, 3 \times 10^{-4})$	$(6 \times 10^{-6}, 2 \times 10^{-6})$
3			$(3 \times 10^{-4}, 9 \times 10^{-4})$	$(2 \times 10^{-5}, 1 \times 10^{-5})$
-3	4	4×10^{-1}	$(6 \times 10^{-2}, 3 \times 10^{-2})$	$(4 \times 10^{-2}, 3 \times 10^{-2})$
-1			$(3 \times 10^{-4}, 4 \times 10^{-4})$	$(8 \times 10^{-6}, 2 \times 10^{-6})$
1			$(3 \times 10^{-4}, 1 \times 10^{-3})$	$(2 \times 10^{-5}, 1 \times 10^{-5})$
3			$(2 \times 10^{-3}, 7 \times 10^{-4})$	$(3 \times 10^{-5}, 4 \times 10^{-5})$

TABLE II: Diagonal rational and off-diagonal quadratic approximants to $f_2(x,y) = (1-x-y)^{\frac{1}{4}}$

x	y	Error in $[6,6/6,6]$ rational approximant	Error in $[4,4/4,3/4,4]$ quadratic approximant	Error in $[6,6/6,5/6,6]$ quadratic approximant
-3	-4	1×10^{-2}	5×10^{-6}	3×10^{-7}
-1		1×10^{-6}	1×10^{-6}	3×10^{-9}
1		2×10^{-4}	1×10^{-7}	$< 10^{-9}$
3		2×10^{-3}	$< 10^{-9}$	$< 10^{-9}$
-3	-2	5×10^{-5}	7×10^{-7}	2×10^{-9}
-1		5×10^{-8}	4×10^{-8}	$< 10^{-9}$
1		3×10^{-6}	$< 10^{-9}$	$< 10^{-9}$
3		5×10^{-1}	$(2 \times 10^{-1}, 1 \times 10^{-1})$	$(1 \times 10^{-1}, 8 \times 10^{-2})$
-3	0	1×10^{-6}	3×10^{-8}	$< 10^{-9}$
-1		$< 10^{-9}$	$< 10^{-9}$	$< 10^{-9}$
1		3×10^{-1}	$(2 \times 10^{-1}, 1 \times 10^{-1})$	$(1 \times 10^{-1}, 8 \times 10^{-2})$
3			$(3 \times 10^{-4}, 3 \times 10^{-4})$	$(5 \times 10^{-6}, 2 \times 10^{-6})$
-3	2	2×10^{-4}	$< 10^{-9}$	$< 10^{-9}$
-1		4×10^{-1}	$(2 \times 10^{-1}, 1 \times 10^{-1})$	$(1 \times 10^{-1}, 8 \times 10^{-2})$
1			$(3 \times 10^{-4}, 3 \times 10^{-4})$	$(6 \times 10^{-6}, 3 \times 10^{-6})$
3			$(8 \times 10^{-4}, 1 \times 10^{-3})$	$(2 \times 10^{-5}, 1 \times 10^{-5})$
-3	4	6×10^{-1}	$(1 \times 10^{-1}, 5 \times 10^{-2})$	$(1 \times 10^{-1}, 5 \times 10^{-2})$
-1			$(4 \times 10^{-4}, 4 \times 10^{-4})$	$(7 \times 10^{-6}, 4 \times 10^{-6})$
1			$(2 \times 10^{-5}, 1 \times 10^{-3})$	$(1 \times 10^{-5}, 2 \times 10^{-5})$
3			$(2 \times 10^{-3}, 7 \times 10^{-4})$	$(4 \times 10^{-5}, 4 \times 10^{-6})$

TABLE III: Diagonal rational and off-diagonal quadratic approximants to $f_3(x,y) = 1+\ln(1-x-y)$

x	y	Error in $\boxed{6,6/6,6}$ rational approximant	Error in $\boxed{4,4/4,3/4,4}$ quadratic approximant	Error in $\boxed{6,6/6,5/6,6}$ quadratic approximant
-3	-4	3×10^{-2}	2×10^{-5}	4×10^{-7}
-1		4×10^{-6}	4×10^{-6}	8×10^{-9}
1		6×10^{-4}	5×10^{-7}	$< 10^{-9}$
3		6×10^{-3}	2×10^{-9}	$< 10^{-9}$
-3	-2	9×10^{-5}	3×10^{-6}	6×10^{-9}
-1		1×10^{-7}	2×10^{-7}	$< 10^{-9}$
1		1×10^{-5}	$< 10^{-9}$	$< 10^{-9}$
4			$(2 \times 10^{-3}, 5 \times 10^{-4})$	$(4 \times 10^{-5}, 3 \times 10^{-5})$
-3	0	4×10^{-6}	2×10^{-7}	$< 10^{-9}$
-1		1×10^{-9}	$< 10^{-9}$	$< 10^{-9}$
2			$(1 \times 10^{-3}, 6 \times 10^{-4})$	$(2 \times 10^{-5}, 1 \times 10^{-5})$
3			$(2 \times 10^{-3}, 1 \times 10^{-4})$	$(1 \times 10^{-5}, 3 \times 10^{-5})$
-4	2	2×10^{-3}	3×10^{-8}	5×10^{-9}
-3		6×10^{-4}	$< 10^{-9}$	$< 10^{-9}$
1			$(2 \times 10^{-3}, 1 \times 10^{-4})$	$(8 \times 10^{-6}, 3 \times 10^{-5})$
3			$(3 \times 10^{-3}, 3 \times 10^{-3})$	$(7 \times 10^{-6}, 9 \times 10^{-5})$
-5	4	2×10^{-2}	1×10^{-9}	7×10^{-8}
-1			$(3 \times 10^{-3}, 1 \times 10^{-4})$	$(2 \times 10^{-5}, 5 \times 10^{-5})$
1			$(4 \times 10^{-3}, 4 \times 10^{-3})$	$(2 \times 10^{-6}, 1 \times 10^{-4})$
3			$(7 \times 10^{-3}, 3 \times 10^{-3})$	$(2 \times 10^{-4}, 7 \times 10^{-5})$

TABLE IV: Diagonal rational and quadratic approximants to $f_4(x,y) = 1+\frac{1}{2}x+\ln(1-x-y)$

x	y	Error in $[6,6/6,6]$ rational approximant	Error in $[4,4/4,4/4,4]$ quadratic approximant	Error in $[7,7/7,7/7,7]$ quadratic approximant
-4	-4	1	1×10^{-4}	3×10^{-8}
-2		5×10^{-2}	1×10^{-5}	3×10^{-9}
0		2×10^{-5}	1×10^{-6}	$< 10^{-9}$
2		1×10^{-2}	1×10^{-4}	4×10^{-6}
4		1×10^{-1}	1×10^{-1}	4×10^{-2}
-4	-2	2×10^{-1}	2×10^{-5}	1×10^{-8}
-2		3×10^{-3}	2×10^{-6}	$< 10^{-9}$
0		2×10^{-7}	1×10^{-8}	$< 10^{-9}$
2		6×10^{-4}	2×10^{-4}	1×10^{-6}
4			$(4 \times 10^{-3}, 9 \times 10^{-3})$	$(7 \times 10^{-4}, 2 \times 10^{-4})$
-4	0	5×10^{-5}	5×10^{-6}	$< 10^{-9}$
-2		4×10^{-7}	5×10^{-8}	$< 10^{-9}$
0		0	0	0
2			$(8 \times 10^{-4}, 2 \times 10^{-3})$	$(3 \times 10^{-6}, 4 \times 10^{-6})$
4			$(4 \times 10^{-3}, 1 \times 10^{-2})$	$(2 \times 10^{-5}, 2 \times 10^{-5})$
-4	2	5×10^{-2}	2×10^{-4}	4×10^{-6}
-2		1×10^{-3}	4×10^{-6}	2×10^{-7}
0			$(2 \times 10^{-3}, 6 \times 10^{-4})$	$(5 \times 10^{-7}, 5 \times 10^{-6})$
2			$(1 \times 10^{-3}, 4 \times 10^{-3})$	$(1 \times 10^{-7}, 9 \times 10^{-7})$
4			$(2 \times 10^{-2}, 6 \times 10^{-3})$	$(7 \times 10^{-5}, 3 \times 10^{-6})$
-4	4	8×10^{-2}	4×10^{-3}	6×10^{-4}
-2			$(2 \times 10^{-3}, 2 \times 10^{-4})$	$(7 \times 10^{-4}, 4 \times 10^{-4})$
0			$(1 \times 10^{-3}, 3 \times 10^{-3})$	$(4 \times 10^{-6}, 2 \times 10^{-5})$
2			$(1 \times 10^{-2}, 2 \times 10^{-3})$	$(3 \times 10^{-5}, 1 \times 10^{-5})$
4			$(1 \times 10^{-2}, 3 \times 10^{-2})$	$(2 \times 10^{-4}, 2 \times 10^{-4})$

TABLE V: Diagonal rational and quadratic approximants to $f_5(x,y) = \ln(1-x-y)/(2.5-y)$

x	y	Error in $[\underline{6},6/6,\underline{6}]$ rational approximant	Error in $[\underline{4},4/4,4/4,\underline{4}]$ quadratic approximant	Error in $[\underline{6},6/6,6/6,\underline{6}]$ quadratic approximant
-4	-4	2×10^{-2}	3×10^{-3}	5×10^{-3}
-2		2×10^{-3}	1×10^{-2}	2×10^{-3}
0		6×10^{-6}	1×10^{-6}	$< 10^{-9}$
2		1×10^{-1}	2×10^{-3}	6×10^{-3}
4		$< 10^{-9}$	$< 10^{-9}$	$< 10^{-9}$
-4	-2	1×10^{-3}	2×10^{-3}	1×10^{-3}
-2		2×10^{-4}	1×10^{-4}	8×10^{-5}
0		1×10^{-7}	8×10^{-9}	$< 10^{-9}$
2		$< 10^{-9}$	$< 10^{-9}$	$< 10^{-9}$
4			$(1 \times 10^{-2}, 7 \times 10^{-3})$	$(1 \times 10^{-2}, 6 \times 10^{-3})$
-4	0	1×10^{-5}	1×10^{-6}	$< 10^{-9}$
-2		8×10^{-8}	1×10^{-7}	$< 10^{-9}$
0		0	0	0
2			$(8 \times 10^{-4}, 2 \times 10^{-4})$	$(6 \times 10^{-6}, 4 \times 10^{-6})$
4			$(4 \times 10^{-4}, 2 \times 10^{-3})$	$(2 \times 10^{-5}, 3 \times 10^{-6})$
-4	2	6	4×10^{-1}	(1, 2)
-2		$< 10^{-9}$	$< 10^{-9}$	$< 10^{-9}$
0			$(5 \times 10^{-1}, 2 \times 10^{-1}$	$(1 \times 10^{-2}, 5 \times 10^{-3})$
2			(4, 6)	(1, 3)
4			(1, 6)	(3, 5)
-4	4	$< 10^{-9}$	$< 10^{-9}$	$< 10^{-9}$
-2			(1, 6)	$(7 \times 10^{-1}, 4)$
0			$(1 \times 10^{-1}, 2 \times 10^{-1})$	$(2 \times 10^{-3}, 2 \times 10^{-3})$
2			$(3 \times 10^{-2}, 2)$	$(6 \times 10^{-1}, 9 \times 10^{-1})$
4			$(6 \times 10^{-1}, 2)$	$(7 \times 10^{-1}, 2)$

TABLE VI: Diagonal and off-diagonal quadratic approximants to $f_4(x,y)$

x	y	Error in $[\underline{5},5/5,5/5,\underline{5}]$ quadratic approximant	Error in $[\underline{5},5/5,4/5,\underline{5}]$ quadratic approximant
-4	-4	7×10^{-6}	1×10^{-5}
-2		6×10^{-7}	3×10^{-7}
0		6×10^{-8}	7×10^{-8}
2		2×10^{-5}	2×10^{-5}
4		2×10^{-3}	4×10^{-3}
-4	-2	2×10^{-6}	4×10^{-6}
-2		2×10^{-7}	4×10^{-7}
0		1×10^{-9}	$< 10^{-9}$
2		5×10^{-7}	1×10^{-6}
4		$(7 \times 10^{-5}, 8 \times 10^{-4})$	$(3 \times 10^{-5}, 3 \times 10^{-4})$
-4	0	5×10^{-6}	4×10^{-7}
-2		$< 10^{-9}$	$< 10^{-9}$
0		0	0
2		$(4 \times 10^{-5}, 3 \times 10^{-4})$	$(1 \times 10^{-4}, 2 \times 10^{-4})$
4		$(1 \times 10^{-3}, 5 \times 10^{-5})$	$(1 \times 10^{-3}, 3 \times 10^{-4})$
-4	2	1×10^{-6}	2×10^{-6}
-2		5×10^{-8}	5×10^{-7}
0		$(1 \times 10^{-4}, 9 \times 10^{-5})$	$(2 \times 10^{-4}, 9 \times 10^{-5})$
2		$(6 \times 10^{-4}, 2 \times 10^{-4})$	$(6 \times 10^{-4}, 5 \times 10^{-4})$
4		$(5 \times 10^{-4}, 2 \times 10^{-3})$	$(5 \times 10^{-4}, 2 \times 10^{-3})$
-4	4	3×10^{-4}	3×10^{-3}
-2		$(2 \times 10^{-4}, 3 \times 10^{-5})$	$(1 \times 10^{-4}, 7 \times 10^{-5})$
0		$(2 \times 10^{-4}, 3 \times 10^{-4})$	$(2 \times 10^{-4}, 5 \times 10^{-4})$
2		$(4 \times 10^{-5}, 1 \times 10^{-3})$	$(7 \times 10^{-4}, 1 \times 10^{-3})$
4		$(4 \times 10^{-3}, 1 \times 10^{-3})$	$(2 \times 10^{-3}, 3 \times 10^{-3})$

TABLE VII: Comparison of the three schemes on the final prong for the $\overline{[5/5/5]}$ diagonal quadratic approximant to $f_4(x,y)$

x	y	Error using Scheme I	Error using Scheme II	Error using Scheme III
-4	-4	7×10^{-6}	7×10^{-6}	6×10^{-6}
-2		6×10^{-7}	6×10^{-7}	4×10^{-7}
0		6×10^{-8}	6×10^{-8}	6×10^{-8}
2		2×10^{-5}	2×10^{-5}	2×10^{-5}
4		2×10^{-3}	2×10^{-3}	9×10^{-4}
-4	-2	2×10^{-6}	2×10^{-6}	2×10^{-6}
-2		2×10^{-7}	2×10^{-7}	2×10^{-7}
0		1×10^{-9}	1×10^{-9}	1×10^{-9}
2		5×10^{-7}	4×10^{-7}	9×10^{-7}
4		$(7 \times 10^{-5}, 8 \times 10^{-4})$	$(8 \times 10^{-5}, 8 \times 10^{-4})$	$(1 \times 10^{-4}, 8 \times 10^{-4})$
-4	0	5×10^{-6}	5×10^{-6}	5×10^{-6}
-2		$< 10^{-9}$	$< 10^{-9}$	$< 10^{-9}$
0		0	0	0
2		$(4 \times 10^{-5}, 3 \times 10^{-4})$	$(4 \times 10^{-5}, 3 \times 10^{-4})$	$(4 \times 10^{-5}, 3 \times 10^{-4})$
4		$(1 \times 10^{-3}, 5 \times 10^{-5})$	$(1 \times 10^{-3}, 5 \times 10^{-5})$	$(1 \times 10^{-3}, 5 \times 10^{-5})$
-4	2	1×10^{-6}	1×10^{-6}	2×10^{-6}
-2		5×10^{-8}	1×10^{-7}	8×10^{-7}
0		$(1 \times 10^{-4}, 9 \times 10^{-5})$	$(1 \times 10^{-4}, 9 \times 10^{-5})$	$(1 \times 10^{-4}, 9 \times 10^{-5})$
2		$(6 \times 10^{-4}, 2 \times 10^{-4})$	$(6 \times 10^{-4}, 2 \times 10^{-4})$	$(6 \times 10^{-4}, 2 \times 10^{-4})$
4		$(5 \times 10^{-4}, 2 \times 10^{-3})$	$(5 \times 10^{-4}, 2 \times 10^{-3})$	$(5 \times 10^{-4}, 2 \times 10^{-3})$
-4	4	3×10^{-4}	$< 10^{-9}$	3×10^{-4}
-2		$(2 \times 10^{-4}, 3 \times 10^{-5})$	$(2 \times 10^{-4}, 3 \times 10^{-5})$	$(2 \times 10^{-4}, 2 \times 10^{-5})$
0		$(2 \times 10^{-4}, 3 \times 10^{-4})$	$(2 \times 10^{-4}, 3 \times 10^{-4})$	$(2 \times 10^{-4}, 3 \times 10^{-4})$
2		$(4 \times 10^{-5}, 1 \times 10^{-3})$	$(4 \times 10^{-5}, 1 \times 10^{-3})$	$(4 \times 10^{-5}, 1 \times 10^{-3})$
4		$(4 \times 10^{-3}, 1 \times 10^{-3})$	$(4 \times 10^{-3}, 1 \times 10^{-3})$	$(4 \times 10^{-3}, 1 \times 10^{-3})$

cut", where x+y>1, roughly six figure accuracy is obtainable with the quadratic approximant. The introduction of only one extra Riemann sheet has thus produced a dramatic improvement in the representation of these infinitely sheeted functions.

(ii) The functions f_1 and f_2 have a finite number of Riemann sheets. The quadratic approximant is again consistently more accurate than the rational approximant and, on the "cut", the quadratic approximant provides a good representation of the functions on two Riemann sheets. So, starting from the power series (1.1) defined on only one Riemann sheet, the quadratic approximants are approximating f_1, f_2, f_3 and f_4 on two Riemann sheets. In addition to analytically continuing a function on a given Riemann sheet, the quadratic approximants thus provide a practical method for the continuation of a function from one Riemann sheet to another.

(iii) The function f_5 has a pole surface in addition to a surface of infinitely-sheeted branch points, and this function illustrates the limitations of the methods of section 2. Away from the pole surface the results of Table 5 exhibit the same general features of (i) and (ii) above. However, near the pole surface, with y=2 and y=4, neither approximation scheme represents the function at all adequately.

In Table 6 we examine the conjecture, made at the end of §3(b), that we would expect the diagonal and 'slightly' off-diagonal quadratic approximants to produce similar results. For the example given in Table 6, this is seen to be the case; the errors in the two tabulated approximants are almost always of the same order of magnitude and are very often equal.

For the diagonal quadratic approximants the three suggested methods of setting up the equations on the final prong, given by (2.8a),(2.8b) and (2.8c), are considered in Table 7. Here we make a comparison of the errors arising from each of the schemes. Again we see that the errors are almost always of the same order of magnitude and very often are equal.

(b) EXAMPLES (II)

The examples considered here are the following two Feynman graphs of Chapter 3:

(i) the three point production process of §5b and

(ii) the fourth order box graph of §6,

where (i) and (ii) are considered as functions of the two invariants s and t.

(i) THREE POINT PRODUCTION PROCESS

From Appendix 2(b) of Chapter 3 we obtain the following two variable expansion for the three point production process:

$$I(s-s_0, t-t_0) = \sum_{r=0}^{\infty} (-1)^r \sum_{k=0}^{r} \binom{r}{k} (s-s_0)^{r-k} (t-t_0)^k \int_0^1 u\,du \int_0^1 dv \frac{[u^2 v(1-v)]^r [u(1-u)v]^{k-r}}{D_0^{r+1}} \quad (6.1)$$

where

$$D_0 = u^2 v(1-v)s_0 + u(1-u)vt_0 + u(1-u)(1-v)M^2 - m^2$$

and s_0, t_0 are the expansion parameters. In contrast to the one variable situation of Chapter 3, we have no working rule for assessing 'good' values of these parameters and in the following results the values chosen for s_0 and t_0 are really no more than guesses.

We note from (6.1) that the method of expansion produces the terms $(s-s_0)^{\alpha}\,(t-t_0)^{\beta}$, where (α,β) lies in a triangular region. From the preceding sections (see especially Figs.3-8) it is clear that it is in precisely this type of region that we require the power series coefficients in order to calculate the two variable approximants; the expansion (6.1) thus does not produce unwanted coefficients. In Table 8 we tabulate results comparing

(a) the (6/6) rational approximant and

(b) the (4/4/4) quadratic approximant,

with mass values M=m=1 and expansion points $s_0=3+3i$ and $t_0=1+i$. We make the following comments:

(i) For $t<4$, we can obtain reasonably good results with the rational approximants. Although these results are not as accurate (see Table 9 of

TABLE 8: TWO VARIABLE DIAGONAL RATIONAL AND QUADRATIC APPROXIMANTS TO THE THREE POINT PRODUCTION PROCESS OF CHAPTER 3,§5b.

s	t	(6/6) RATIONAL APPROXIMANT	(4/4/4) QUADRATIC APPROXIMANT
0	1	$(-0.6046001,-0.4\times10^{-7})$	$(-0.60452,-0.1\times10^{-3})$
2		$(-0.768472,-0.2\times10^{-6})$	$(-0.76832,-0.5\times10^{-4})$
4		(-1.65,-0.1)	(-1.60,-0.06)
6		(-0.6951,-0.997)	(-0.681,-0.992)
8		(-0.361,-0.8941)	(-0.365,-0.883)
10		(-0.204,-0.7888)	(-0.22,-0.785)
0	3	$(-0.822473,0.7\times10^{-7})$	$(-0.8213,0.3\times10^{-2})$
2		$(-1.117896,0.1\times10^{-4})$	$(-1.1126,0.4\times10^{-2})$
4		(-4.1,-0.983)	(-1.7,-1.9)
6		(-0.420,-1.1522)	(-0.6,-1.5)
8		(-0.1203,-1.1610)	(-0.14,-1.31)
10		(-0.0071,-0.955)	(-0.08,-1.1)
0	4.5	(-1.4,-0.44)	(-1.2,0.39)
2		(-2.1,-1.2)	(-1.2,-1.3)
4		(-0.2,-0.5)	(-0.8,-2.2)
6		(-0.2,-0.8)	(-0.65,-2.0)
8		(0.42,-1.020)	(0.84,-1.08)
10		(0.356,-0.834)	(0.65,-0.66)

Chapter 3) as those obtained with the one variable approximants (and t fixed), we do not have to compute the integrals in (6.1) anew for each value of t when using the two variable approximants. The results obtained with the quadratic approximants are disappointing, especially in view of the results of §6(a).

(ii) For $t>4$, the results are not very much in agreement. The explanation of this is provided by Appendix 2(b) of Chapter 3; for real positive $t>4m^2$ $(=4)$ there is a complex singularity on the physical sheet. This has the consequence that we cannot write single variable dispersion relations, with integrations along the real axis, for production processes - at least not in simple variables (see (11) for further details). If we fix t (>4) and form one variable approximants in $(s-s_0)$ as in Chapter 3, the resulting α-integrations become difficult to perform due to the presence of singularities in the integration region; again we cannot obtain convergent results.

(ii) FOURTH ORDER SCALAR BOX GRAPH

From Appendix 3 of Chapter 3 we obtain the following two variable expansion for the fourth order box graph:

$$I(s-s_0,t-t_0)=\sum_{j=0}^{\infty}(j+1)(-1)^j\sum_{r=0}^{j}\binom{j}{r}(s-s_0)^r(t-t_0)^{j-r}\int_0^1 du\int_0^1 dv\int_0^1 dw\,\frac{u(1-u)a^r b^{j-r}}{D_0^{j+2}}$$

where

$$a=u^2v(1-v)$$

$$b=(1-u)^2w(1-w)$$

and

$$D_0=as_0+bt_0-M^2(1-u)^2-m^2u$$

In Table 9 we compare the (6/6) rational and (4/4/4) quadratic approximants for the mass values M=m=1 and with $s_0=3+3i$ and $t_0=1+i$. Again we obtain good results using the rational approximants but disappointing results with the quadratic approximants.

7. CONCLUSIONS

In this chapter we have seen that the ideas underlying the definition

TABLE 9: TWO VARIABLE DIAGONAL RATIONAL AND QUADRATIC APPROXIMANTS TO THE FOURTH ORDER SCALAR BOX GRAPH OF CHAPTER 3, §6.

s	t	(6/6) RATIONAL APPROXIMANT	(4/4/4) QUADRATIC APPROXIMANT	EXACT VALUE
0	-1	$(0.2357114,-0.1\times10^{-6})$	$(0.23566,-0.3\times10^{-4})$	(0.2357113,0)
1		$(0.2710095,0.1\times10^{-6})$	$(0.271003,-0.3\times10^{-4})$	(0.2710096,0)
2		$(0.3253897,0.7\times10^{-7})$	$(0.32541,0.3\times10^{-5})$	(0.3253893,0)
3		$(0.428310,-0.4\times10^{-5})$	$(0.42826,0.4\times10^{-4})$	(0.428303,0)
4		(0.98,0.09)	(1.11,-0.04)	
5		(0.327,0.64921)	(0.3275,-0.6499)	
6		(0.1142,0.5455)	(0.1139,0.5450)	
7		(0.0257,0.4581)	(0.02557,0.4579)	
8		(-0.01872,0.3933)	(-0.01864,0.3930)	
9		(-0.0433,0.34366)	(-0.04367,0.34360)	
0	0	$(0.2636002,-0.2\times10^{-6})$	$(0.26361,-0.2\times10^{-4})$	(0.2636001,0)
1		$(0.3022998,0.1\times10^{-6})$	$(0.30231,0.2\times10^{-5})$	(0.3022999,0)
2		$(0.3615971,0.8\times10^{-7})$	$(0.361595,0.1\times10^{-4})$	(0.3615968,0)
3		$(0.47281,-0.4\times10^{-5})$	$(0.47277,-0.2\times10^{-4})$	(0.47280,0)
4		(1.05,0.1)	(1.25,-0.04)	(1.21,0)
5		(0.3883,0.70255)	(0.3891,0.70243)	(0.3894,0.70248)
6		(0.14949,0.6048)	(0.14965,0.6044)	(0.14962,0.6046)
7		(0.0460,0.5140)	(0.0460,0.51410)	(0.0458,0.51416)
8		(-0.00747,0.4445)	(-0.00733,0.44433)	(-0.00745,0.44429)
9		(-0.0377,0.39019)	(-0.03752,0.39034)	(-0.03758,0.39026)

and construction of the two variable rational approximants can be extended in a natural way to quadratic approximants. The approximants defined in this way have been shown to possess many of the properties of the corresponding two variable rational approximants.

The main object in defining these two variable "t-power" approximants is to try and extend the advantages of the one variable Shafer approximants to two variable series. Indeed the results of §6(a) indicate that this has been achieved; the two variable quadratic approximants do seem to provide a (sometimes considerably) better method of approximation than the currently used two variable rational approximants, especially to functions possessing branch points. However, the results of §6(b) are rather unsatisfactory and at present the explanation of this behaviour is not very clear.

If we confine our attention to real series then, on the basis of the results of §6(a), the two main conclusions we can draw are:

(i) The two variable quadratic approximants do seem to considerably increase the accuracy in two variable calculations. The functions f_1, f_2 and f_3 of §6(a) were considered because they exhibit commonly occuring types of branch points. We have not explicitly considered the function

$$F(x,y) = (1-x-y)^{1/2}$$

which has a square root branch point, since the quadratic approximants of §2 and §5 are obviously exact for this function. Since the quadratic approximants approximate these functions so well, and considerably better than the corresponding rational approximants, we expect the basic trend in the results of Tables 1-7 to be reproduced for many of the multi-valued functions which occur in practice.

(ii) Starting from a function with a power series expansion valid on only one sheet of the function, we can, using the two variable quadratic approximants, obtain information about the function on a second sheet. The approximants permit the analytic continuation of a function from one Riemann sheet to another.

On the basis of (i) and (ii), we expect the two variable quadratic

approximants discussed in this chapter to be of great value in the many areas of theoretical physics where perturbation methods are commonly employed. We also expect the "t-power" approximants (t>2) of this chapter to be of great potential use.

As a final point we note that, although the two variable quadratic approximants produce disappointing results in this connection, the methods of Chapter 3 for calculating Feynman integrals in the physical region can certainly be used in conjunction with the two variable rational approximants to produce reliable results.

FIG. 3: PRONG STRUCTURE FOR THE (m/m) DIAGONAL CHISHOLM RATIONAL APPROXIMANT

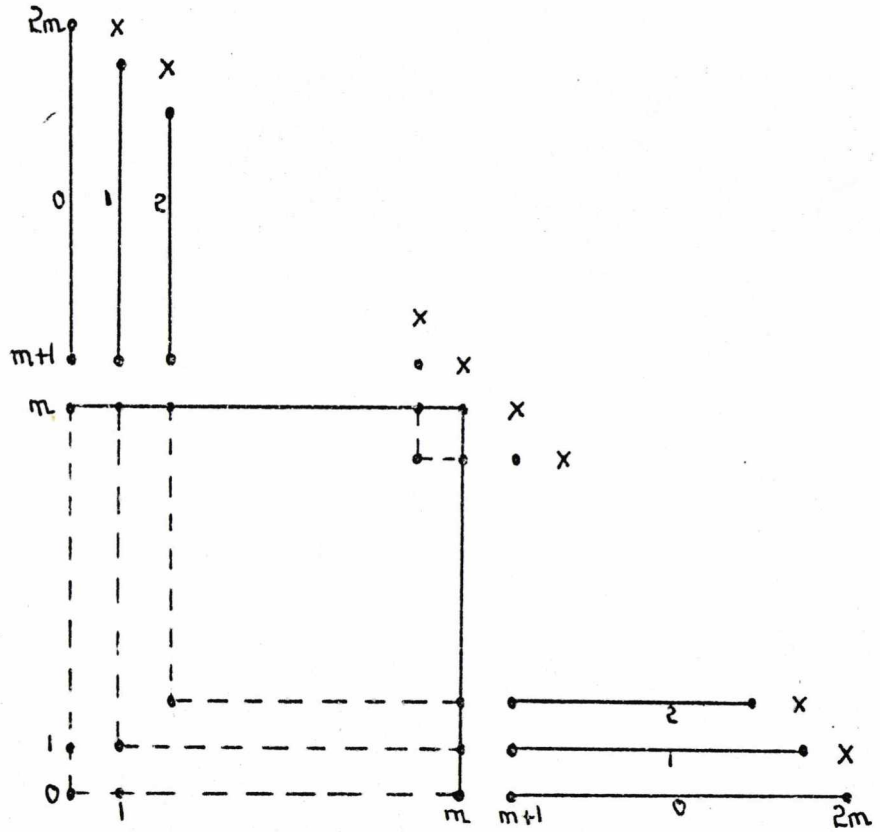


FIG. 4: PRONG STRUCTURE FOR THE (m/m/m) DIAGONAL QUADRATIC TWO VARIABLE APPROXIMANT

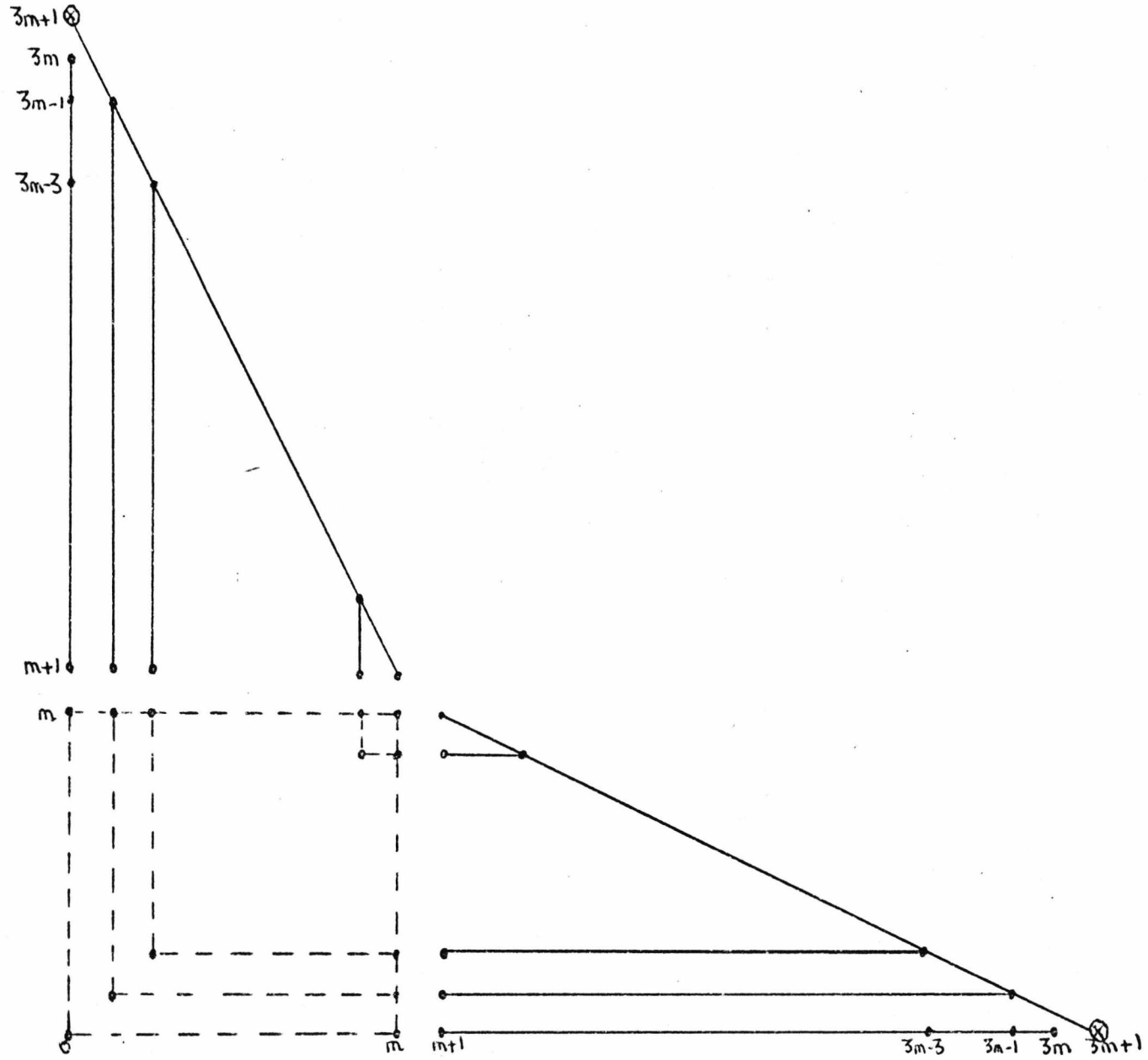


FIG. 5: PRONG STRUCTURE FOR THE (m/m/m/m) DIAGONAL TWO VARIABLE CUBIC APPROXIMANT

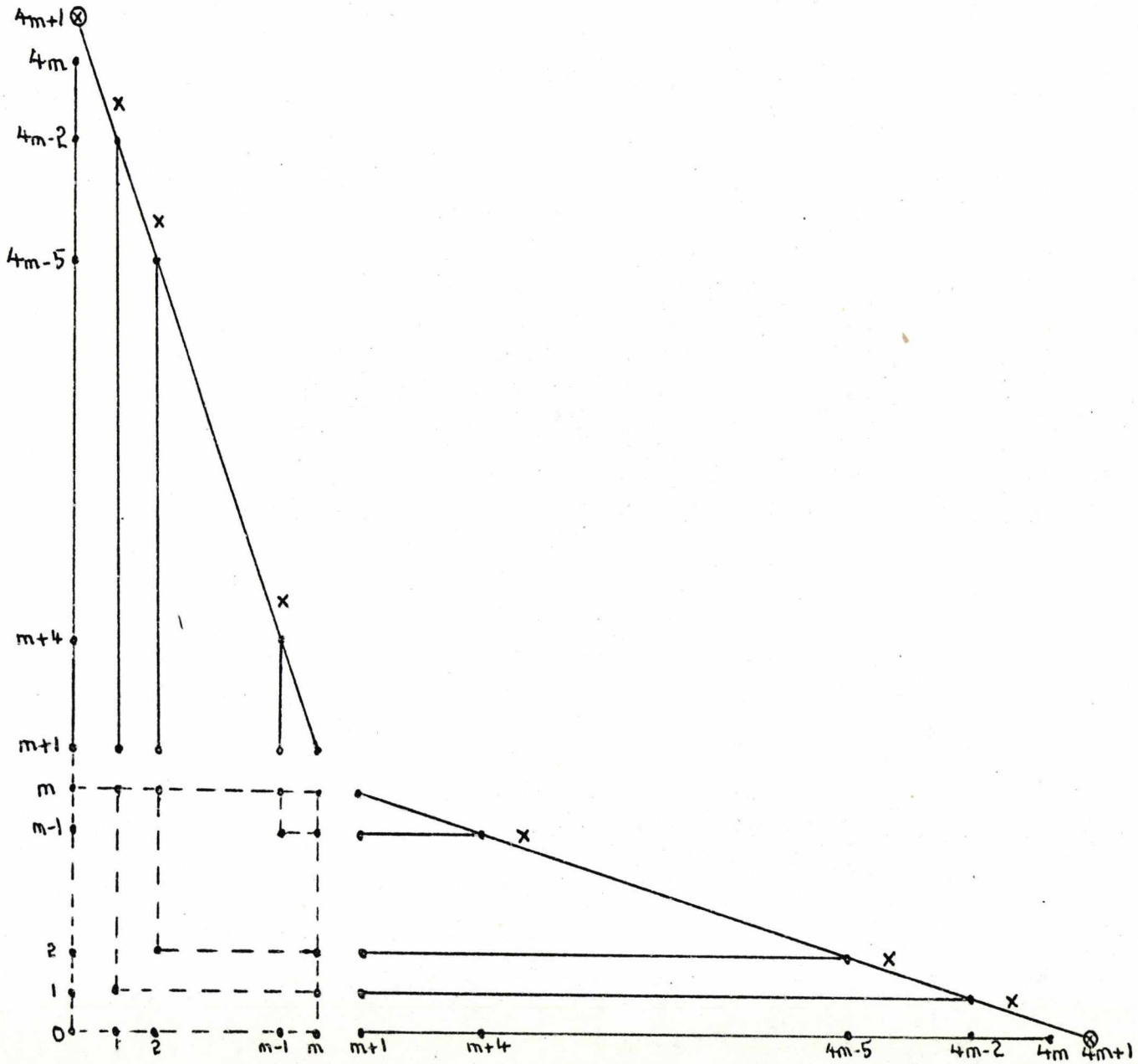


FIG. 6: PRONG STRUCTURE FOR THE $(m_{1;o}, m_{2;o}/m_{1;1}, m_{2;1})$ OFF-DIAGONAL TWO VARIABLE RATIONAL (t=1) APPROXIMANT

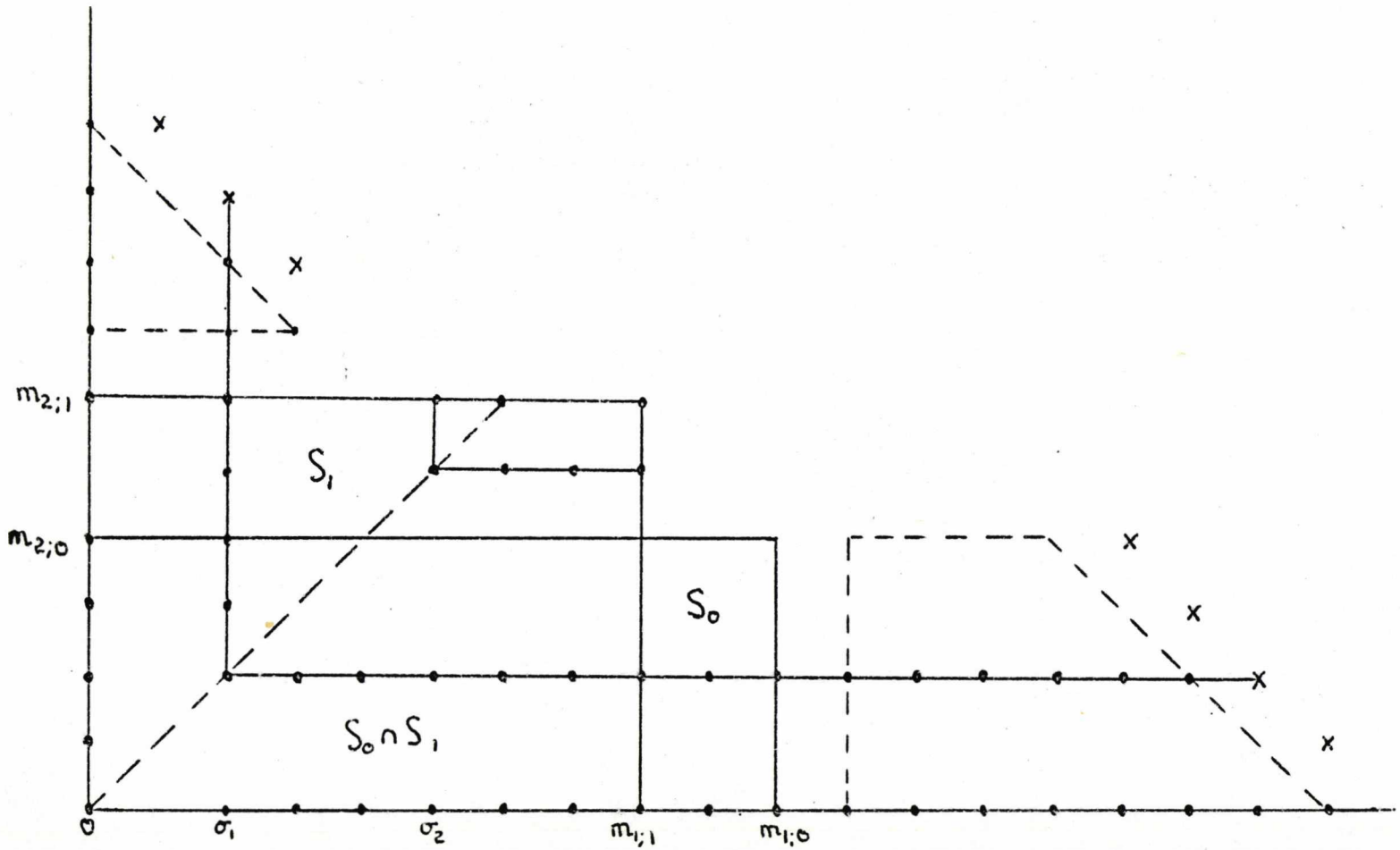


FIG. 7: PRONG STRUCTURE FOR THE (2,5/3,4/1,2) OFF-DIAGONAL TWO VARIABLE QUADRATIC (t=2) APPROXIMANT

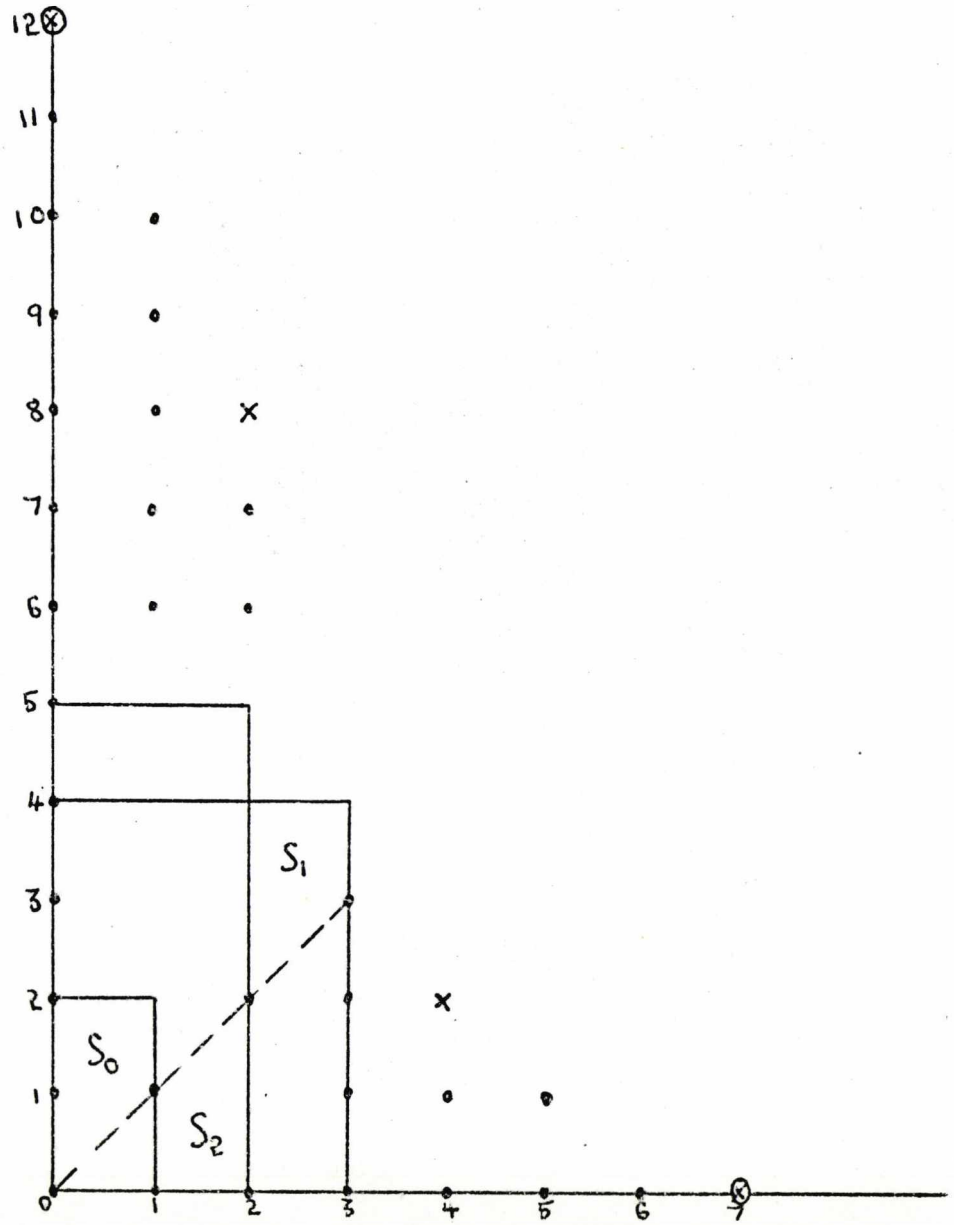


FIG. 8: PRONG STRUCTURE FOR THE (5,2/3,6/6,5/7,8) OFF-DIAGONAL TWO VARIABLE CUBIC (t=3) APPROXIMANT

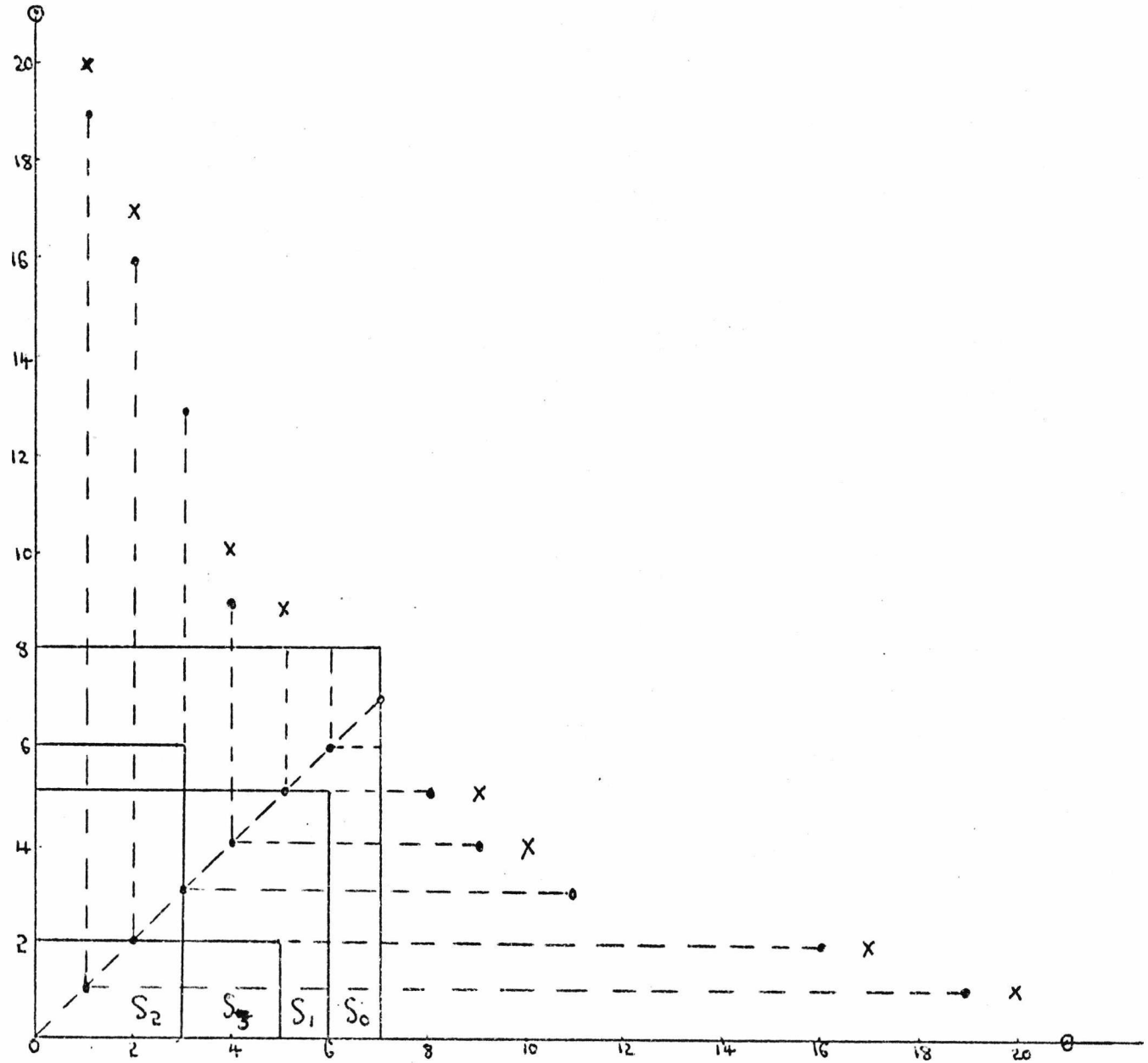
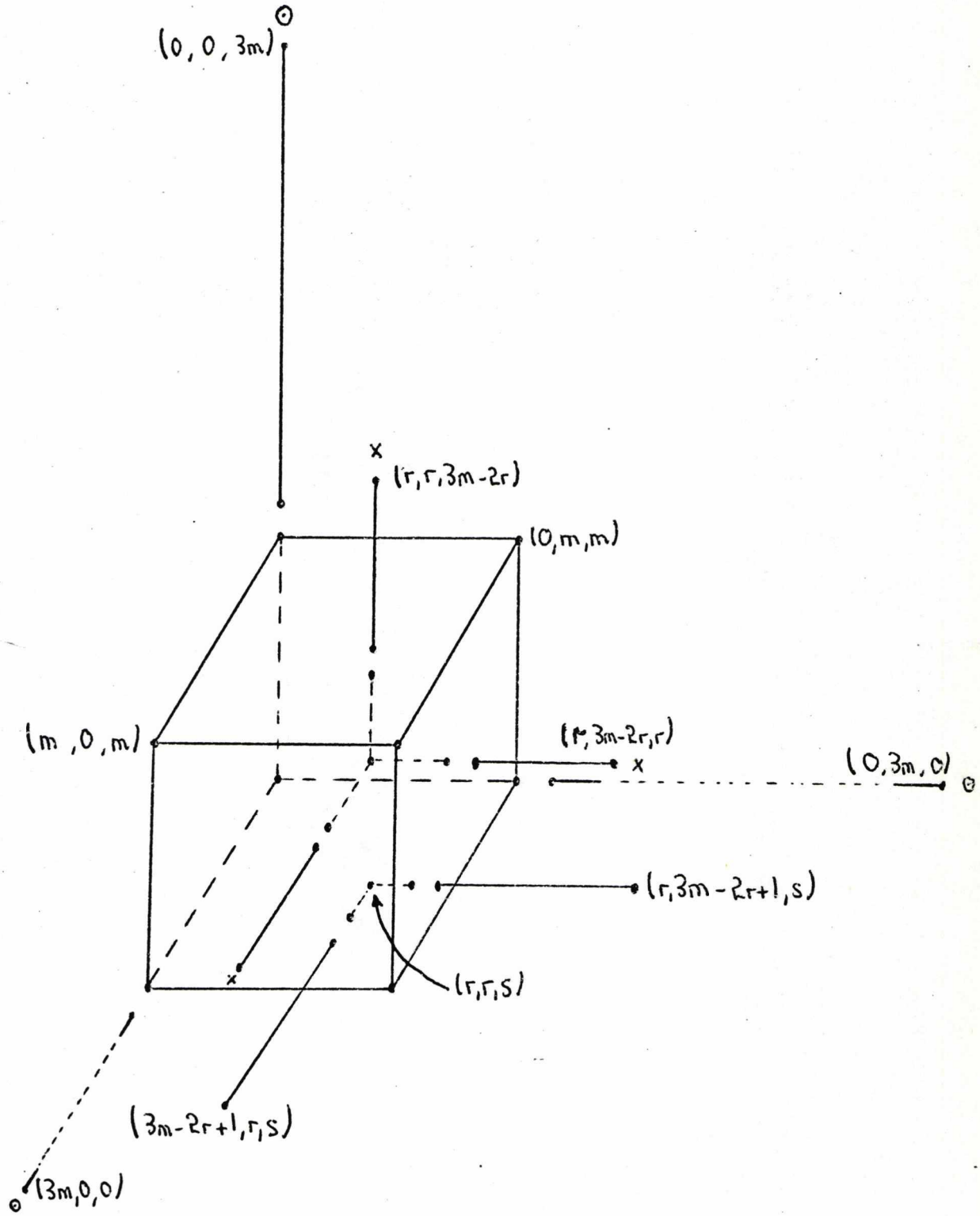


FIG.9: PRONG STRUCTURE FOR THE (m/m/m) THREE VARIABLE DIAGONAL QUADRATIC APPROXIMANT



APPENDIX I: CAUCHY-BINET THEOREM (8)

Let the determinant formed of elements taken from rows $\alpha_1, \alpha_2, \ldots, \alpha_n$ and columns $B_1, B_2, \ldots, B_n$ of a matrix M be denoted by

$$M\begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix}$$

Then if M=AB----RS, where A,B,...,R,S are of order kxm,mxn,...,rxs,sxk (so that M is of order kxk), we have

$$M = \sum\sum\sum A\begin{pmatrix} 1 & 2 & \cdots & k \\ \beta_1 & \beta_2 & \cdots & \beta_k \end{pmatrix} B\begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_k \\ \gamma_1 & \gamma_2 & \cdots & \gamma_k \end{pmatrix} \cdots\cdots R\begin{pmatrix} \rho_1 & \rho_2 & \cdots & \rho_k \\ \sigma_1 & \sigma_2 & \cdots & \sigma_k \end{pmatrix} S\begin{pmatrix} \sigma_1 & \sigma_2 & \cdots & \sigma_k \\ 1 & 2 & \cdots & k \end{pmatrix}$$

where the summations are over all sets of k columns taken independently from the columns of A,B,...,R or, alternatively, over all sets of k rows taken independently from rows B,....,R,S.

The expansion will vanish if any one of the following inequalities are true:

k>m,k>n,...., k>r,k>s

REFERENCES

1. J.S.R. Chisholm, Math. Comp.27,841 (1973).

2. R. Hughes Jones and G.J. Makinson, Jour.Inst.Math.Appl.13, 299 (1973).

3. J.S.R. Chisholm and R. Hughes Jones, Proc.Roy.Soc.A344,465 (1975).

4. J.S.R. Chisholm and J. McEwan, Proc. Roy.Soc.A336,421 (1973).

5. R. Hughes Jones, Jour.Approx.Theory 16,201 (1976).

6. J.S.R. Chisholm, "Multivariate approximants with branch points: I. Diagonal approximants", Proc.Roy.Soc., to appear.

7. J.S.R. Chisholm and P.R. Graves-Morris, Proc.Roy.Soc.A342,341 (1975).

8. A.C. Aitken, "Determinants and matrices", Chapter IV (Oliver and Boyd).

9. P.R. Graves-Morris and R. Hughes Jones, Jour.Comp.Appl.Math.2,1 (1976).

10. J.S.R. Chisholm, "Multivariate approximants with branch points: II. Off-diagonal approximants", submitted to Proc.Roy.Soc.

11. P.V. Landshoff and S.B. Trieman, Nuovo Cimento 19, 1249 (1961).

12. P.R. Graves Morris, private communication.

APPENDIX 2: FORTRAN COMPUTER PROGRAM FOR THE CALCULATION OF TWO VARIABLE OFF-DIAGONAL QUADRATIC APPROXIMANTS

The following subroutine, GQA2V, calculates the two variable off-diagonal quadratic approximants defined in Chapter 4; the routine is written in Fortran, uses double precision arithmetic throughout and has the following parameters:

C: A 25x25 matrix containing the coefficients of the power series expansion of $f(\underline{z})$.

IPX,IPY,IQX,IQY,IRX,IRY: These integers define the (IPX,IPY/IQX,IQY/IRX,IRY) quadratic approximant.

P,Q,R: These 10x10 matrices contain the calculated coefficients of the polynomials $P(\underline{z})$,$Q(\underline{z})$ and $R(\underline{z})$ respectively.

IPRONG: This integer parameter defines the prescription to be used on the final prong for diagonal approximants. The values IPRONG=-1,0,+1 correspond to the schemes (2.8c), (2.8b) and (2.8a) respectively.

For off-diagonal approximants the value of IPRONG is immaterial.

```
      SUBROUTINE GQA2V(C,IPX,IPY,IQX,IQY,LPX,LPY,P,Q,R,IPRONG)
C THIS SUBROUTINE CALCULATES 2-VARIABLE QUADRATIC APPROXIMANTS
      DOUBLE PRECISION C(25,25),B(25,25),CC(25,25),RH(25),
     1P(10,10),Q(10,10),R(10,10)
      DOUBLE PRECISION PZERO,DET,S,SEIN,ONE,OUGHT
      DATA ONE,OUGHT/0.1D 01,0.0D 00/
C DEFINITIONS AND INITIALIZATIONS
C LPX=LENGTH OF X-FORK OF PRONG ZERO(EXCLUDING SYMMETRIZED
C POINTS)
      LPX=IPX+IQX
      LPY=IPY+IQY
      LPXY=LPX+LPY
      LPXY1=LPXY+1
      IRX1=IRX+1
      IRY1=IRY+1
      IQX1=IQX+1
      IQY1=IQY+1
      IRX1=IRX+1
      IRY1=IRY+1
      P(1,1)=ONE
      PZERO=P(1,1)
      DO 5 I=1,LPXY1
      DO 5 J=1,LPXY1
    5 CC(I,J)=OUGHT
C GENERATION OF F*F
      IX=LPX+IRX1+1
      IY=LPY+IRY1+1
      CALL POWER2(C,C,B,IX,IY)
C CALCULATION OF P AND Q COEFFICIENTS
      IF(LPXY.EQ.0) GOTO 75
C ZEROTH PRONG(X-FORK)
C RH VECTOR AND CC MATRIX (X-FORK)
      IF(LPX.EQ.0) GOTO 45
      DO 40 I=1,LPX
         I1=I+IRX1
         RH(I)=-PZERO*B(I1,1)
         IF(IPX.EQ.0) GOTO 30
         DO 28 J=1,IPX
            I1=I-J+IRX1
            IF(I1.LT.1) GOTO 28
            CC(I,J)=B(I1,1)
            IF(I.EQ.LPX) CC(LPXY1,J)=B(I1+1,1)
   28    CONTINUE
   30    IF(IQX.EQ.0) GOTO 40
         DO 35 J=1,IQX
            I1=I-J+IRX1
            IF(I1.LT.1) GOTO 35
            CC(I,J+IPX)=C(I1,1)
            IF(I.EQ.LPX) CC(LPXY1,J+IPX)=C(I1+1,1)
   35    CONTINUE
   40 CC(I,LPXY1)=C(IRX1+I,1)
C RH VECTOR AND CC MATRIX (Y-FORK)
   45 IF(LPY.EQ.0) GOTO 75
      DO 70 I=1,LPY
         I1=I+IRY1
         RH(I+LPX)=-PZERO*B(1,I1)
         IF(IPY.EQ.0) GOTO 60
         DO 55 J=1,IPY
            I1=I-J+IRY1
            IF(I1.LT.1) GOTO 55
            CC(I+LPX,J+LPX)=B(1,I1)
            IF(I.EQ.LPY) CC(LPXY1,J+LPX)=B(1,I1+1)
   55    CONTINUE
   60    IF(IQY.EQ.0) GOTO 70
         DO 65 J=1,IQY
            I1=I-J+IRY1
            IF(I1.LT.1) GOTO 65
            CC(I+LPX,J+LPX+IPY)=C(1,I1)
            IF(I.EQ.LPY) CC(LPXY1,J+LPX+IPY)=C(1,I1+1)
   65    CONTINUE
   70 CC(I+LPX,LPXY1)=C(1,IRY1+I)
C LAST ENTRY OF CC-MATRIX AND RH-VECTOR
   75 CC(LPXY1,LPXY1)=C(LPX+IRX+2,1)+C(1,LPY+IRY+2)
      RH(LPXY1)=-PZERO*(B(LPX+IRX+2,1)+B(1,LPY+IRY+2))
      NEQ=LPXY1
      CALL SOLVE(CC,NEQ,RH,1,DET)
C DEFINITION OF P AND Q ON PRONG ZERO
      IF(IPX.EQ.0) GOTO 120
      DO 115 I=1,IPX
  115 P(I+1,1)=RH(I)
  120 IF(IQX.EQ.0) GOTO 130
      DO 125 I=1,IQX
  125 Q(I+1,1)=RH(I+IPX)
  130 I=(IPY.EQ.0) GOTO 140
      DO 135 I=1,IPY
```

```
  135 Q(I,IP1)=QU(I1,IQY)
  140 IF(IQY.EQ.0) GOTO 160
      DO 145 I=1,IQY
  145 Q(I,IP1)=QU(IL,IXP1QY)
  150 Q(1,1)=PH(LPXY1)
      IF(LPXY.EQ.0) GOTO 600
C NON-ZERO PRONGS
C IP DENOTES PRONG P
C LIPX=LENGTH OF X-FORK OF INTERNAL PRONG P (EXCLUDING ANY
C SYMMETRISED POINTS)
C IDIAG=+1 INDICATES THAT THE FINAL PRONG HAS BEEN REACHED
C WHEN THE REQUESTED APPROXIMANT IS DIAGONAL
      IDIAG=-1
C DEFINITION OF LIPX,LIPY
      IP=0
  200 IP=IP+1
      IJ1=IPX-IP+1
      IJ2=IPX+IQX-2*(IP-1)
      IM1=IPY-1
      IF(IPY.EQ.0) IM1=0
      IJ3=IPX+IQX+IM1-3*(IP-1)
      IP1=IP+1
      IP2=IP1+1
      IF((IPX.EQ.IPY).AND.(IQX.EQ.IQY).AND.(IRX.EQ.IRY).AND.
      (IP.EQ.IPX)) IDIAG=1
      ISIGMA=0
      IFLAGP=-1
      IFLAGQ=-1
      IFLAGR=-1
      IF((IP.GT.IPX).OR.(IP.GT.IPY)) GOTO 210
      ISIGMA=ISIGMA+1
      IFLAGP=1
  210 IF((IP.GT.IQX).OR.(IP.GT.IQY)) GOTO 220
      ISIGMA=ISIGMA+1
      IFLAGQ=1
  220 IF((IP.GT.IRX).OR.(IP.GT.IRY)) GOTO 230
      ISIGMA=ISIGMA+1
      IFLAGR=1
  230 F=0.0
      IPAR=ISIGMA-(ISIGMA/2)*2
      IF(IPAR.EQ.0) F=0.5
      INCX=0
      INCY=0
      IF(IFLAGP.LT.0) GOTO 235
      INCX=INCX+IPX
      INCY=INCY+IPY
  235 IF(IFLAGQ.LT.0) GOTO 240
      INCX=INCX+IQX
      INCY=INCY+IQY
  240 IF(IFLAGR.LT.0) GOTO 245
      INCX=INCX+IRX
      INCY=INCY+IRY
  245 IF(IFLAGR.LT.0) GOTO 250
      LIPX=IFIX(INCX-IRX-F-(IP-0.5)*(ISIGMA-1))
      LIPY=IFIX(INCY-IRY-F-(IP-0.5)*(ISIGMA-1))
      GOTO 252
  250 LIPX=IFIX(INCX-(IP-1)-F-(IP-0.5)*(ISIGMA-1))
      LIPY=IFIX(INCY-IP-F-(IP-0.5)*(ISIGMA-1))
  252 LIPX1=LIPX+1
      LIPXY=LIPX+LIPY
      LIPXY1=LIPXY+1
      IF(ISIGMA.EQ.0) GOTO 600
      IF(ISIGMA.EQ.1) GOTO 525
      IF(IDIAG.GT.0) GOTO 377
C RH VECTOR (X-FORK)
         DO 285 K=1,LIPXY1
         DO 285 J=1,LIPXY1
  285    CC(K,J)=DUGIT
      IF(LIPY.LE.0) GOTO 320
       DO 310 I=1,LIPX
         IG=IPX1+I
         IF(IFLAGR.LT.0) IG=IP+I
         S=DUGIT
         DO 280 IT=1,IP
            I1=IG-IT+1
            IF(IP.LE.IPY) S=S+B(I1,1)*P(IT,IP1)
            IF(IP.LE.IQY) S=S+C(I1,1)*Q(IT,IP1)
            LXP=IPX
            IF(IG.LT.(LXP+1)) LXP=IG-1
            LXQ=IQX
            IF(IG.LT.(LXQ+1)) LXQ=IG-1
            LXP1=LXP+1
            LXQ1=LXQ+1
            DO 270 IS=1,LXP1
               I2=IG-IS+1
               I3=IP2-IT
  270          IF((IS.LE.IPX1).AND.(IT.LE.IPY1)) S=S+B(I2,I3)*
               P(IS,IT)
```

```
                  DO 375 IS=1,LYQ1
                     I2=IG-IS+1
                     I3=ID-IT
  275          IF((IS.LE.IQX1).AND.(IT.LE.IQY1)) S=S+C(I2,I3)*
               Q(IS,IT)
  280       CONTINUE
            RH(I)=-S
C  CC-MATRIX (X-WORK)
            L1=LXQ-(IP-1)
            IF(L1.LT.1) GOTO 300
            DO 295 J=1,L1
            K=IG-IP-J+1
            IF(K.LT.1) GOTO 290
            CC(I,J)=B(K,1)
  290       CONTINUE
  295       CONTINUE
  300       L2=LXQ-(IP-1)
            IF(L2.LT.1) GOTO 310
            DO 305 J=1,L2
            K=IG-IP-J+1
            IF(K.LT.1) GOTO 305
            CC(I,J+IJ1)=C(K,1)
  305       CONTINUE
  310    CONTINUE
C  RH-VECTOR (Y-WORK)
  320    IF(LIPY.LE.0) GOTO 376
         DO 375 I=LIPX1,LIPXY
            IR=IRY
            IF(IFLAGP.LT.0) IR=IP
            ID=I+IR-LIPX+1
            S=ZUGHT
            DO 360 IS=1,IP
               I1=ID-IS+1
               IF(IP.LE.IPX) S=S+B(I,I1)*P(IP1,IS)
               IF(IP.LE.IQX) S=S+C(I,I1)*Q(IP1,IS)
               LYQ=IQY
               IF(ID.LT.(LYQ+1)) LYQ=ID-1
               LYP=IPY
               IF(ID.LT.(LYP+1)) LYP=ID-1
               LYP1=LYP+1
               LYQ1=LYQ+1
               DO 335 IT=1,LYP1
                  I2=IP2-IS
                  I3=ID-IT+1
  335          IF((IS.LE.IPX1).AND.(IT.LE.IPY1)) S=S+B(I2,I3)*
               P(IS,IT)
               DO 340 IT=1,LYQ1
                  I2=IP2-IS
                  I3=ID-IT+1
  340          IF((IS.LE.IQX1).AND.(IT.LE.IQY1)) S=S+C(I2,I3)*
               Q(IS,IT)
  350       CONTINUE
            RH(I)=-S
C  CC-MATRIX(Y-WORK)
            I1=ID-IP
            IF(IFLAGP.GT.0) CC(I,1)=B(1,I1)
            IF(IFLAGQ.GT.0) CC(I,IPX1-IP+1)=C(1,I1)
            LP=LYP-IP
            IF(LP.LT.1) GOTO 365
            DO 360 J=1,LP
               K=ID-J-IP
               IF(K.LT.1) GOTO 360
               CC(I,J+IJ2)=B(1,K)
  360       CONTINUE
  365       LQ=LYQ-IP
            IF(LQ.LT.1) GOTO 375
            DO 37) J=1,LQ
               K=ID-J-IP
               IF(K.LT.1) GOTO 370
               CC(I,J+IJ3)=C(1,K)
  370       CONTINUE
  375 CONTINUE
  376 IF(IPAR.NE.0) GOTO 465
C  LAST ENTRIES OF RH-VECTOR AND CC-MATRIX WHEN SYMMETRIZATION
C  REQUIRED
C  X-WORK
  377 IF((IDIAG.GT.0).AND.(IPROMG.EQ.1)) GOTO 710
      I=LIPXY1
      IF(IDIAG.GT.0) I=1
  380 IG=IRX1+I
      IF(IFLAGP.LT.0) IG=IP+I
      SUM=ZUGHT
      DO 390 IT=1,IP
         I1=IG-IT+1
         IF(IP.LE.IPX) SUM=SUM+B(I1,1)*P(IT,IP1)
         IF(IP.LE.IQX) SUM=SUM+C(I1,1)*Q(IT,IP1)
            LXP=IPX
            IF(IG.LT.(LXP+1)) LXP=IG-1
            LX=IQX
            IF(IG.LT.(LXQ+1)) LX=IG-1
            LX1=LXP+1
```

```
                  DO 380 IS=1,LXQ1
                     I2=IG-IS+1
                     I3=IP2-IT
  380             IF((IS.LE.IPX1).AND.(IT.LE.IPY1)) SFIN=SFIN+B(I2,I3)*
                 2(IS,IT)
                  DO 385 IS=1,LXQ1
                     I2=IG-IS+1
                     I3=IP2-IT
  385             IF((IS.LE.IQX1).AND.(IT.LE.IQY1)) SFIN=SFIN+C(I2,I3)*
                 2(IS,IT)
  390          CONTINUE
               R4(I)=-SFIN
               IF(IDIAG.GT.0) GOTO 415
               LL=LXQ-(IP-1)
               IF(LL.LT.1) GOTO 405
               DO 400 J=1,LL
                  K=IG-IP-J+1
                  IF(K.LT.1) GOTO 400
                  CC(I,J)=B(K,1)
  400          CONTINUE
               L2=LXQ-(IP-1)
  405          IF(L2.LT.1) GOTO 415
               DO 410 J=1,L2
                  K=IG-IP-J+1
                  IF(K.LT.1) GOTO 410
                  CC(I,J+IJ1)=C(K,1)
  410          CONTINUE
C Y-PORK
  415             IR=IRY
                  SFIN=DUGHT
                  IF(IFLAGR.LT.0) IR=IP
                  ID=I+IR-1-IPX+1
                  IF(IDIAG.GT.0) ID=ID+1
                  DO 440 IS=1,IP
                  I1=ID-IS+1
                  IF(IP.LE.IPX) SFIN=SFIN+B(1,I1)*P(IP1,IS)
                  IF(IP.LE.IQX) SFIN=SFIN+C(1,I1)*Q(IP1,IS)
                     LYQ=IQY
                     IF(ID.LT.(LYQ+1)) LYQ=ID-1
                     LYP=IPY
                     IF(ID.LT.(LYP+1)) LYP=ID-1
                     LYP1=LYP+1
                     LYQ1=LYQ+1
                  DO 420 IT=1,LYP1
                     I2=IP2-IS
                     I3=ID-IT+1
  420             IF((IS.LE.IPX1).AND.(IT.LE.IPY1)) SFIN=SFIN+B(I2,I3)*P(IS,IT)
                  DO 430 IT=1,LYQ1
                     I2=IP2-IS
                     I3=ID-IT+1
  430             IF((IS.LE.IQX1).AND.(IT.LE.IQY1)) SFIN=SFIN+C(I2,I3)*Q(IS,IT)
  440          CONTINUE
               R4(I)=RH(I)-SFIN
               IF(IDIAG.GT.0) GOTO 461
               I1=ID-IP
               IF(IFLAGR.GT.0) CC(I,1)=CC(I,1)+B(1,I1)
               IF(IFLAGR.GT.0) CC(I,IPX1-IP+1)=CC(I,IPX1-IP+1)+C(1,I1)
               LP=LYP-IP
               IF(LP.LT.1) GOTO 455
               DO 450 J=1,LP
                  K=ID-J-IP
                  K=K+1
                  IF(K.LT.1) GOTO 450
                  CC(I,J+IJ2)=B(1,K)
  450          CONTINUE
               LQ=LYQ-IP
  455          IF(LQ.LT.1) GOTO 465
               DO 460 J=1,LQ
                  K=ID-J-IP
                  K=K+1
                  IF(K.LT.1) GOTO 460
                  CC(I,J+IJ3)=C(1,K)
  460          CONTINUE
               GOTO 465
  461          IF(I.EQ.2) GOTO 462
               I=I+1
               GOTO 389
C CC-MATRIX ON FINAL PRONG FOR DIAGONAL APPROXIMANTS.
  462          IF(IPRONG) 700,705,710
  700          CC(1,1)=B(2,1)+B(1,2)
               CC(1,2)=C(2,1)+C(1,2)
               CC(2,1)=B(3,1)+B(1,3)
               CC(2,2)=C(3,1)+C(1,3)
               GOTO 465
  705          CC(1,1)=B(2,1)+B(1,2)
               GOTO 465
  710          B(IP1,IP1)=DUGHT
               Q(IP1,IP1)=DUGHT
               P(IP1,IP1)=DUGHT
               GOTO 600
```

```
      GOTO 600
  465 NEQ=LIPXY
      IF(IPAK.EQ.0) NEQ=LIPXY1
      IF(NEQ.EQ.0) GOTO 525
      IF((IDIAG.GT.0).AND.(IPRONG.EQ.0)) NEQ=1
      CALL SOLVE(CC,NEQ,RH,1,DET)
C DEFINITION OF P AND Q COEFFICIENTS ON PRONG P
      IM1=IPX-IP+1
      IF(IM1.LE.0) GOTO 495
      DO 490 I=1,IM1
  490 P(IP+I,IP1)=RH(I)
  495 IM2=IQX-IP+1
      IF(IM2.LE.0) GOTO 505
      DO 500 I=1,IM2
  500 Q(IP+I,IP1)=RH(I+IM1)
  505 IM3=IPY-IP
      IF(IM3.LE.0) GOTO 515
      DO 510 I=1,IM3
  510 P(IP1,IP1+I)=RH(I+IM1+IM2)
  515 IM4=IQY-IP
      IF(IM4.LE.0) GOTO 525
      DO 520 I=1,IM4
  520 Q(IP1,IP1+I)=RH(I+IM1+IM2+IM3)
      IF(IDIAG.GT.0) Q(IP1,IP1)=OUGHT
  525 IF(ISIGMA.GT.1) GOTO 200
C EVALUATION OF ANY P OR Q COEFFICIENTS REMAINING
      MPXY=IPX
      IF(IPY.LT.MPXY) MPXY=IPY
      MQXY=IQX
      IF(IQY.LT.MQXY) MQXY=IQY
      IF((IP.GT.MPXY).AND.(IP.GT.MQXY)) GOTO 600
      IF(IP.GT. IPXY) GOTO 560
C AT THIS POINT P COEFFICIENTS ARE LEFT UNDEFINED
      S=B(1,1)
      IF(DABS(S).LT.1.0D-50) GOTO 590
      IP1=IP+1
      DO 550 IG=IP1,IPX1
      DO 550 ID=IP1,IPY1
      P(IG,ID)=OUGHT
      S=OUGHT
      DO 540 IS=1,IG
      DO 540 IT=1,ID
         I1=IG-IS+1
         I2=ID-IT+1
         IF((IS.LE.IQX1).AND.(IT.LE.IQY1)) S=S+C(I1,I2)*Q(IS,IT)
  540    S=S+B(I1,I2)*P(IS,IT)
  550 P(IG,ID)=-S/B(1,1)
      GOTO 600
C AT THIS POINT Q COEFFICIENTS ARE LEFT UNDEFINED
  560 DO 580 IG=IP1,IQX1
      DO 580 ID=IP1,IQY1
      Q(IG,ID)=OUGHT
      S=OUGHT
      DO 570 IG=1,IG
      DO 570 IT=1,ID
         I1=IG-IS+1
         I2=ID-IT+1
         IF((IS.LE.IPX1).AND.(IT.LE.IPY1)) S=S+B(I1,I2)*P(IS,IT)
  570 S=S+C(I1,I2)*Q(IS,IT)
  580 Q(IG,ID)=-S/C(1,1)
      GOTO 600
  590 WRITE(2,595)
  595 FORMAT(///39H CONSTANT TERM OF TAYLOR SERIES IS ZERO///)
  600 CONTINUE
C CALCULATION OF R COEFFICIENTS
      DO 610 IG=1,IRX1
      DO 610 ID=1, IRY1
      S=OUGHT
      LX=IPX1
      IF(IG.LE.LX) LX=IG
      LY=IPY1
      IF(ID.LE.LY) LY=ID
      DO 620 IS=1,LX
      DO 620 IT=1,LY
         I1=IG-IS+1
         I2=ID-IT+1
  620 S=S+B(I1,I2)*P(IS,IT)
      LX=IQX1
      IF(IG.LE.LX) LX=IG
      LY=IQY1
      IF(ID.LE.LY) LY=ID
      DO 630 IS=1,LX
      DO 630 IT=1,LY
         I1=IG-IS+1
         I2=ID-IT+1
  630 S=S+C(I1,I2)*Q(IS,IT)
  610 R(IG,ID)=-S
      IF((IDIAG.GT.0).AND.(IPRONG.EQ.1)) R(IRX1,IRY1)=OUGHT
      RETURN
      END
```

```
      SUBROUTINE SOLVE(A,N,B,M,DET)
C     THIS SUBROUTINE IS FOR SOLVING SIMULTANEOUS LINEAR
C     EQUATIONS
      DOUBLE PRECISION A(25,25),B(25,1),PIVOT,DET,T
      DIMENSION IPVOT(25),INDEX(25,2),PIVOT(25)
      DOUBLE PRECISION ONE,ZERO
      DATA ONE,ZERO/1.0D 01,0.0D 00/
      M=1
      DET=ONE
      DO 1 J=1,N
    1 IPVOT(J)=0
      DO 24 I=1,N
      T=ZERO
      DO 6 J=1,N
      IF(IPVOT(J)-1) 2,6,2
    2 DO 5 K=1,N
      IF(IPVOT(K)-1) 3,5,24
    3 IF(DABS(T)-DABS(A(J,K)))4,4,5
    4 IROW=J
      ICOL=K
      T=A(J,K)
    5 CONTINUE
    6 CONTINUE
      IPVOT(ICOL)=IPVOT(ICOL)+1
      IF(IROW-ICOL) 7,11,7
    7 DET=-DET
      DO 8 L=1,N
      T=A(IROW,L)
      A(IROW,L)=A(ICOL,L)
    8 A(ICOL,L)=T
      IF(M) 11,11,9
    9 DO 10 L=1,M
      T=B(IROW,L)
      B(IROW,L)=B(ICOL,L)
   10 B(ICOL,L)=T
   11 INDEX(I,1)=IROW
      INDEX(I,2)=ICOL
      PIVOT(I)=A(ICOL,ICOL)
      IM1=I-1
      IF(I.EQ.1) GO TO 14
      T=DEXP(DLOG(DABS(DET))/DBLE(FLOAT(IM1)))
      IF(DABS(PIVOT(I))-T*DBLE(FLOAT(I))*0.1D-22) 12,14,14
   12 DET=ZERO
      IF(M.EQ.0) GOTO 25
      GO TO 25
   14 DET=DET*PIVOT(I)
      A(ICOL,ICOL)=ONE
      DO 15 L=1,N
   15 A(ICOL,L)=A(ICOL,L)/PIVOT(I)
      IF(M) 18,18,16
   16 DO 17 L=1,M
   17 B(ICOL,L)=B(ICOL,L)/PIVOT(I)
   18 DO 23 LI=1,N
      IF(LI-ICOL) 19,23,19
   19 T=A(LI,ICOL)
      A(LI,ICOL)=ZERO
      DO 20 L=1,N
   20 A(LI,L)=A(LI,L)-A(ICOL,L)*T
      IF(M) 23,23,21
   21 DO 22 L=1,M
   22 B(LI,L)=B(LI,L)-B(ICOL,L)*T
   23 CONTINUE
   24 CONTINUE
   25 CONTINUE
      RETURN
      END

      SUBROUTINE POWER2(A,B,C,NX,NY)
C THIS SUBROUTINE MULTIPLIES THE 2-VARIABLE POLYNOMIALS A AND
C B AND STORES THE RESULT IN C
C STORES THE RESULT IN C
C NX IS THE MAXIMUM POWER OF X IN A (AND IN B)
C NY IS THE MAXIMUM POWER OF X IN Y (AND IN B)
      DOUBLE PRECISION A(25,25),B(25,25),C(25,25)
      DO 1 I=1,NX
      DO 1 J=1,NY
    1 C(I,J)=0.0D 00
      DO 10 I=1,NX
      DO 10 K1=1,I
      L1=I+1-K1
      DO 10 J=1,NY
      DO 10 K2=1,J
      L2=J+1-K2
   10 C(I,J)=C(I,J)+A(K1,K2)*B(L1,L2)
      RETURN
      END
```

CHAPTER 5: QUADRATIC APPROXIMANTS AND LEGENDRE SERIES.

1. PADÉ APPROXIMANTS TO LEGENDRE SERIES

In many situations, especially in scattering theory (see, for example, (1)), it is more usual to expand a function f(z) as a Legendre series, rather than as a power series. In this chapter we look at the possible uses of quadratic approximants in relation to Legendre series; the methods used here can be extended in a straightforward way to the case of cubic and higher order approximants. In particular, we hope to define approximants which will converge along branch cuts of f(z).

Given the power series

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

the (M/N) Padé approximant to f(z) is given, with the notation of Chapter 1, by

$$f_{(M/N)}(z) = \frac{p_M(z)}{q_N(z)} \tag{1.1}$$

where

$$f(z)\,q_N(z) - p_M(z) = O(z^{M+N+1}) \tag{1.2}$$

Given the corresponding Legendre series

$$f(z) = \sum_{L=0}^{\infty} f_L P_L(z) \tag{1.3}$$

three types of "Legendre Padé approximant" have been defined:

(i) Fleischer (2) and Holdeman (3) have defined the (M/N) Legendre Padé approximant by the requirement that

$$f(z)\,q_N(z) - p_M(z) = O(P_{M+N+1}(z)) \tag{1.4}$$

in analogy to (1.2), where

$$q_N(z) = \sum_{K=0}^{N} b_K P_K(z) \quad \text{and} \quad p_M(z) = \sum_{L=0}^{M} a_L P_L(z)$$

are respectively the denominator and numerator of the approximant.

(ii) Fleischer (4) has given the alternative requirement

$$f(z) - \frac{p_M(z)}{q_N(z)} = O(P_{M+N+1}(z)) \tag{1.5}$$

(iii) The Legendre-Pade approximants of Common, which we discuss in §4.

(1.5) is not equivalent to (1.4), as are the analogous equations for the ordinary Padé approximants; essentially this is because of (1.8). In fact (1.5) leads to a system of non-linear equations which cannot be guaranteed to have a unique solution (or indeed any solution at all). The motivation for considering (1.5) is that the "linear Legendre Padé approximants" defined by (1.4) (so called because they lead to a linear system of equations) do not have the property that their first few expansion coefficients agree with the first few coefficients of the original series; the "non-linear" approximants have this property.

Because of the computational difficulties associated with the non-linear approximants of (1.5) (see (4) for a discussion of these), we shall only consider the linear approximants defined by (1.4). Given the Legendre series (1.3), the (N/M) Legendre Padé approximant is therefore defined by (2)

$$f_{(N/M)}(P(z)) = \frac{R_N(P(z))}{S_M(P(z))} = \frac{n_0 P_0(z) + \cdots + n_N P_N(z)}{d_0 P_0(z) + \cdots + d_M P_M(z)} \tag{1.6}$$

where

$$f(z) S_M(P(z)) - R_N(P(z)) = O(P_{M+N+1}(z)) \tag{1.7}$$

The product $f(P(z))\ S_M(P(z))$ in (1.7) is evaluated using the following expression for the product of two Legendre polynomials:

$$P_i(z) P_j(z) = \sum_{L=|i-j|}^{i+j} \alpha_L^{(i,j)} P_L(z) \tag{1.8}$$

where the $\alpha_L^{(i,j)}$ can be written in the symmetrical form (2)

$$\alpha_L^{(i,j)} = \begin{cases} \dfrac{1}{(2L+1)(i+j+L+1)} \dfrac{R\left(\frac{1}{2}(-i+j+L)\right) R\left(\frac{1}{2}(i-j+L)\right) R\left(\frac{1}{2}(i+j-L)\right)}{R\left(\frac{1}{2}(i+j+L)\right)}, & (i+j+L) \text{ even} \\ 0, & \text{otherwise} \end{cases}$$

with

$$R(k) = \binom{k-\frac{1}{2}}{k}$$

Because more than one term appears for each product on the right hand side of (1.8), the coefficients f_L $(0 \leq L \leq 2M+N)$ of (1.3) contribute in the

construction of the (N/M) approximant. In contrast, for the Padé approximants defined from the Taylor series, the maximum index contributing to (N/M) is N+M.

In analogy with the ordinary Padé approximants, we set

$$d_0 = 1$$

the remaining (N+M+1) coefficients n_j $(0 \le j \le N)$ and d_i $(1 \le i \le M)$ are determined by the requirement

$$f(z) S_M(P(z)) - R_N(P(z)) = \sum_{k=M+N+1}^{\infty} \beta_k P_k(z)$$

so that

$$\sum_{k=M+N+1}^{\infty} \beta_k P_k(z) = \left(\sum_{j=0}^{\infty} f_j P_j(z) \right) \left(\sum_{k=0}^{M} d_k P_k(z) \right) - \sum_{k=0}^{N} n_k P_k(z)$$

and

$$\sum_{k=M-N-1}^{\infty} \beta_k P_k(z) - \sum_{k=0}^{N} n_k P_k(z) = \sum_{j=0}^{\infty} f_j \sum_{k=0}^{M} d_k P_j(z) P_k(z) \qquad (1.9)$$

$$= \sum_{j=0}^{\infty} f_j \sum_{k=0}^{M} d_k \sum_{L=|j-k|}^{j+k} \alpha_L^{(j,k)} P_L(z)$$

using (1.8). The summations on the right-hand side of (1.9) are such that

(i) $$j < 2M+N+1$$

and

(ii) $$L < M+N+1$$

Multiplying (1.9) by $P_\sigma(z)$ $(0 \le \sigma < M+N+1)$ and integrating between -1 and +1, we have

$$n_L \delta_{\sigma L} = \sum_{j} f_j \sum_{k=0}^{M} d_k \alpha_\sigma^{(j,k)} \qquad (1.10)$$

provided

$$|j-k| \le \sigma \le j+k \qquad (1.11)$$

In deriving (1.10) we have used the result

$$\int_{-1}^{1} P_L(z) P_m(z) = \begin{cases} 0 & L \neq m \\ 2/(2L+1) & L = m \end{cases}$$

The summation limits on j in (1.10) are defined by (1.11). Rewriting (1.11) as

$$|\sigma-k| \le j \le \sigma+k$$

(1.10) becomes

$$\pi_L \delta_{\sigma L} = \sum_{k=0}^{M} d_k \sum_{j=|\sigma-k|}^{\sigma+k} f_j \alpha_\sigma^{(j,k)} \quad (0 \le \sigma < M+N+1) \tag{1.12}$$

If we define

$$a_{L,i} = \sum_{n=|L-i|}^{L+i} \alpha_L^{(i,n)} f_n \tag{1.13}$$

then (1.12) becomes

$$\pi_L \delta_{\sigma L} = \sum_{k=0}^{M} d_k a_{\sigma,k} \tag{1.14}$$

(since $\alpha_L^{(j,k)} = \alpha_L^{(k,j)}$). Thus, the requirement (1.4) leads to the following system of linear equations:

$$\begin{cases} \sum_{i=0}^{M} d_i a_{L,i} = 0 & (N+1 \le L \le N+M) \\ \sum_{i=0}^{M} d_i a_{L,i} = \pi_L & (0 \le L \le N) \end{cases} \tag{1.15}$$

The analogy between the linear Legendre Padē approximants, defined by (1.6) and (1.7), and the normal Padē approximants can be clearly seen from the following representation (obtained from (1.15)) of the (N/M) Legendre Padē approximant:

$$(N/M) = \frac{\det \begin{vmatrix} a_{N+1,M} & \cdots & a_{N+1,0} \\ \vdots & & \vdots \\ a_{N+M,M} & \cdots & a_{N+M,0} \\ \sum_{L=0}^{N} a_{L,M} P_L(z) & \cdots & \sum_{L=0}^{N} a_{L,0} P_L(z) \end{vmatrix}}{\det \begin{vmatrix} a_{N+1,M} & \cdots & a_{N+1,0} \\ \vdots & & \vdots \\ a_{N+M,M} & \cdots & a_{N+M,0} \\ P_M(z) & \cdots & P_0(z) \end{vmatrix}} \tag{1.16}$$

which is to be compared with the representation of the normal Padé approximants given by Theorem 3.1 of Chapter 1.

2. QUADRATIC APPROXIMANTS TO LEGENDRE SERIES

The results of §1 can be readily generalized to the case of quadratic approximants. We define polynomials Q(z), R(z) and S(z) by

$$Q(z) = \sum_{j=0}^{q} q_j P_j(z) \tag{2.1a}$$

$$R(z) = \sum_{k=0}^{r} r_k P_k(z) \tag{2.1b}$$

and

$$S(z) = \sum_{L=0}^{s} s_L P_L(z) \tag{2.1c}$$

where q,r,s are the degrees of Q,R and S respectively.

The (q/r/s) Legendre quadratic approximant to (1.3) is then defined to be the solution, g=(q/r/s), of the equation

$$Q(z)g^2(z) + R(z)g(z) + S(z) = 0 \tag{2.2a}$$

where Q,R and S are determined by

$$Q(z)f^2(z) + R(z)f(z) + S(z) = O(P_{q+r+s+2}(z)) \tag{2.2b}$$

(2.2b) leads to a system of linear equations, which we shall now determine; we do not consider the "non-linear" case arising from an equation analogous to (1.5). Defining

$$f^2(z) = \sum_{L=0}^{\infty} \tilde{f}_L P_L(z) \tag{2.3}$$

we have, from (2.2b),

$$\sum_{k=q-r-s-2}^{\infty} \beta_k P_k(z) - \sum_{j=0}^{r} s_j P_j(z) = \left(\sum_{n=0}^{\infty} \tilde{f}_n P_n(z)\right)\left(\sum_{k=0}^{q} q_k P_k(z)\right) + \left(\sum_{n=0}^{\infty} f_n P_n(z)\right)\left(\sum_{k=0}^{r} r_k P_k(z)\right)$$

$$= \sum_{n=0}^{\infty} \tilde{f}_n \sum_{k=0}^{q} q_k P_n(z) P_k(z) + \sum_{n=0}^{\infty} f_n \sum_{k=0}^{r} r_k P_n(z) P_k(z)$$

$$= \left(\sum_{n=0}^{\infty} \tilde{f}_n \sum_{k=0}^{q} q_k + \sum_{n=0}^{\infty} f_n \sum_{k=0}^{r} r_k\right) \sum_{L=|n-k|}^{n+k} \alpha_L^{(n,k)} P_L(z) \tag{2.4}$$

The summation over n in (2.4) is such that

(i) $$n < \max(2q+r+s+2,\ q+2r+s+2)$$

and

(ii) $$l < q+r+s+2$$

Multiplying (2.4) by P_σ $(0 \leq \sigma < q+r+s+2)$ and integrating between -1 and +1, we have

$$-S_j \delta_{\sigma j} = \sum_n \tilde{f}_n \sum_{k=0}^{q} q_j \alpha_\sigma^{(n,k)} + \sum_{n'} f_{n'} \sum_{m=0}^{r} r_m \alpha_\sigma^{(n',m)} \tag{2.5}$$

provided

(a) $$|n-k| \leq \sigma \leq n-k \quad (0 \leq k \leq q)$$

and

(b) $$|n'-m| \leq \sigma \leq n'-m \quad (0 \leq m \leq r)$$

Since (a) and (b) imply that

$$|\sigma-k| \leq n \leq \sigma+k$$

and

$$|\sigma-m| \leq n' \leq \sigma+m$$

(2.5) can be written

$$-S_j \delta_{\sigma j} = \sum_{k=0}^{q} q_k \sum_{n=|\sigma-k|}^{\sigma+k} \tilde{f}_n \alpha_\sigma^{(n,k)} + \sum_{k'=0}^{r} r_{k'} \sum_{n'=|\sigma-k'|}^{\sigma+k'} f_{n'} \alpha_\sigma^{(n',k')} \tag{2.6}$$

$$= \sum_{k=0}^{q} q_k \tilde{a}_{\sigma,k} + \sum_{k'=0}^{r} r_{k'} a_{\sigma,k'}$$

where

$$a_{l,i} = \sum_{n=|l-i|}^{l+i} \alpha_l^{(i,n)} f_n \tag{2.7a}$$

and

$$\tilde{a}_{l,i} = \sum_{n=|l-i|}^{l+i} \alpha_l^{(i,n)} \tilde{f}_n \tag{2.7b}$$

Hence, from (2.6) the linear system of equations defining the (linear) Legendre quadratic approximant is

$$\sum_{i=0}^{q} q_i \tilde{a}_{l,i} + \sum_{j=0}^{r} r_j a_{l,j} = -S_l \quad (0 \leq l \leq s) \tag{2.8}$$

and

$$\sum_{i=0}^{q} q_i \tilde{a}_{l,i} - \sum_{j=0}^{r} r_j a_{l,j} = 0 \quad (s+1 \leq l < q+r+s+2)$$

We note that the (q/r/s) Legendre quadratic approximant to (1.3) requires

$$q+r+s+2+\max(q,r)$$

terms of the series expansion of f(z).

From (2.8) it is possible to write down an explicit determinant representation for the Legendre quadratic approximant, in analogy to (1.16) for the Legendre Padé approximants. This will lead to a representation similar to that contained in Theorem 3.2 (of Chapter 1); we do not give the results here, since we shall make no use of them, and their form is apparent from Theorem 3.2.

3. COMPARISON OF THE LEGENDRE PADÉ AND LEGENDRE QUADRATIC APPROXIMANTS

We make a comparison of the approximation schemes of §1 and §2 by considering the following two functions, both of which are considered by Fleischer in (4):

(i)
$$f(z) = \ln\frac{1-az}{1-a} \tag{3.1}$$
$$= \left[Q_0\left(\tfrac{1}{a}\right) - Q_1\left(\tfrac{1}{a}\right)\right]P_0(z) + \sum_{L=1}^{\infty}\left[Q_{L+1}\left(\tfrac{1}{a}\right) - Q_{L-1}\left(\tfrac{1}{a}\right)\right]P_L(z)$$

where $1/a = 1.81\dot{6}$ gives the beginning of the cut, and $Q_L(x)$ is the Legendre function of the second kind.

In Table 1 we compare

(a) the linear (2/2) Padé approximant of §1, requiring 7 terms of (3.1), and

(b) the linear (1/1/1) quadratic approximant of §2, requiring 6 terms of (3.1).

Also tabulated is the partial sum of the first seven terms of (3.1). The results indicate that scheme (b) is preferable, especially near the branch point. More important, though, is the possibility of using scheme (b) for $z(>1/a)$ along the branch cut. In Table 2 we give results for the (5/5/5) quadratic approximant for $z>1/a$. Although the accuracy obtainable on the cut is not comparable to that obtained off the cut (see Table 1), it is possible to obtain roughly four significant figures along the branch cut.

TABLE 1: COMPARISON OF THE (2/2) LINEAR PADÉ AND (1/1/1) LINEAR QUADRATIC APPROXIMANTS TO THE LEGENDRE SERIES EXPANSION OF $f(z) = \ln \frac{1-az}{1-a}$, WITH $1/a = 1.816$.

z	PARTIAL SUM	LINEAR PADÉ APPROXIMANT	LINEAR QUADRATIC APPROXIMANT	f(z)
-3.5	DIVERGENT	1.865	1.8756	1.8734
-1.75	1.42	1.4735	1.4743	1.47416
-1.0	1.23797	1.23806	1.238082	1.238078
-0.5	1.04262	1.042655	1.0426543	1.0426536
0.0	0.799537	0.799521	0.7995279	0.7995276
0.5	0.47765	0.47765	0.477624	0.477628
1.0	0.0020	0.00018	-0.00006	0.0
1.5	-0.854	-0.907	-0.966	-0.947
1.75	-1.57	-1.90	-2.55	-2.51

TABLE 2: LINEAR (5/5/5) LEGENDRE QUADRATIC APPROXIMANT TO

$f(z) = \ln \frac{1-az}{1-a}$, FOR $z > \frac{1}{a} = 1.816$.

z	(5/5/5) LEGENDRE QUADRATIC APPROXIMANT	f(z)
1.75	-2.5065	-2.5057
2.0	$(-1.488, \pm 3.138)$	$(-1.494, \pm 3.14159)$
2.5	$(-0.17796, \pm 3.1432)$	$(-0.1782, \pm 3.14159)$
3.0	$(0.3719, \pm 3.1406)$	$(0.3709, \pm 3.14159)$
3.5	$(0.0722, \pm 3.1404)$	$(0.0723, \pm 3.14159)$
4.0	$(0.982, \pm 3.1418)$	$(0.983, \pm 3.14159)$
4.5	$(1.1883, \pm 3.143)$	$(1.1896, \pm 3.14159)$
5.0	$(1.36047, \pm 3.144)$	$(1.36045, \pm 3.14159)$

TABLE 3: (2/2) LEGENDRE PADÉ AND (1/1/1) LEGENDRE QUADRATIC APPROXIMANTS TO $f(z)=\frac{1}{\sqrt{1-2az+a^2}}$, WITH a=0.3.

s	PARTIAL SUM	LINEAR PADÉ APPROXIMANT	LINEAR QUADRATIC APPROXIMANT	f(z)
-3.5	DIVERGES	0.561	0.56008	0.5599
-1.75	0.77	0.6838	0.68362	0.68359
-1.0	0.76940	0.769236	0.769233	0.769231
-0.5	0.84824	0.8481884	0.84818876	0.84818893
0.0	0.95781	0.957829	0.957828	0.957826
0.5	1.12505	1.125080	1.1250865	1.1250879
1.0	1.4283	1.42849	1.42859	1.42857
1.5	2.13	2.26	2.298	2.294
1.75	2.82	3.19	5.19	5.00

TABLE 4: LINEAR (3/3/3) LEGENDRE QUADRATIC APPROXIMANT TO

$$f(z)=\frac{1}{\sqrt{1-2az+a^2}}, \text{ WITH } a=0.3.$$

z	(3/3/3) QUADRATIC LEGENDRE APPROXIMANT	f(z)
1.75	5.0000008	5
2.0	$(2\times10^{-6}, \pm 3.015116)$	$(0, \pm 3.015113)$
2.5	$(4\times10^{-6}, \pm 1.5617372)$	$(0, \pm 1.5617376)$
3.0	$(5\times10^{-7}, \pm 1.1867823)$	$(0, \pm 1.1867817)$
3.5	$(3\times10^{-8}, \pm 0.9950378)$	$(0, \pm 0.9950372)$
4.0	$(-3\times10^{-7}, \pm 0.8737046)$	$(0, \pm 0.8737041)$
4.5	$(-5\times10^{-7}, \pm 0.7881108)$	$(0, \pm 0.7881104)$
5.0	$(-8\times10^{-7}, \pm 0.7235748)$	$(0, \pm 0.7235746)$

(ii)
$$f(z) = \frac{1}{\sqrt{1-2az+a^2}} = \sum_{L=0}^{\infty} a^L P_L(z) \tag{3.2}$$

Choosing a=0.3, the branch cut starts at $x_c=(1+a^2)/2a=1.816$, as in (i) above.

In Table 3 we give the results obtained using schemes (a) and (b) above for $x<x_c$. We again find that (b) is preferable, especially near to the branch point. For $x>x_c$, we tabulate the (3/3/3) approximant in Table 4; in this case the great accuracy obtainable along the cut reflects the fact that we are essentially approximating a square root branch point by a square root branch point.

4. THE LEGENDRE PADÉ APPROXIMANTS OF COMMON

Common (5) has defined Padé approximants to Legendre series in the following way. The analytic properties of the Legendre series

$$f(z) = \sum_{n=0}^{\infty} f_n P_n(z) \tag{4.1}$$

are related (6) to those of the corresponding power series

$$g(z) = \sum_{n=0}^{\infty} f_n (-z)^n \tag{4.2}$$

If g(z) has radius of convergence r(r>1), then the Legendre series f(z) converges in an ellipse with foci at ± 1 and with semi-major axis $\frac{1}{2}(r+\frac{1}{r})$. The important relationship we require between f(z) and g(z) is (6)

$$f(z) = \frac{1}{\pi}\int_0^{\pi} g\left[-z-(z^2-1)^{1/2}\cos v\right] dv \tag{4.3}$$

The importance of (4.3) lies in the fact that it relates the problem of the analytic continuation of the Legendre series f(z) outside its ellipse of convergence to the corresponding problem of the analytic continuation of the power series g(z) outside its circle of convergence.

The most natural way of approximating f(z) is to replace g(w), where $w=-z-(z^2-1)^{\frac{1}{2}}\cos v$, in (4.3) by its Padé approximants (N/N+j). In (5) the case when g(w) is a series of Stieltjes was studied; in this case the (N/N+j) approximants have known convergence properties. The extension of this work to a general power series g(w) is given in (7). To define the Common approximants we write

$$(N/N+j) = \sum_{p=1}^{N} \frac{\alpha_{p,N}}{(1+\sigma_{p,N}w)} + \sum_{q=0}^{j} \beta_{q,N} w_q(z) \qquad (4.4)$$

where the $\alpha_{p,N}$, $\beta_{q,N}$ and $\sigma_{p,N}$ are expressible in terms of the series coefficients f_i $(0 \leq i \leq 2N+j)$. If we denote the approximant to f(z) obtained by replacing g(w) by (N/N+j) in (4.3) by $f_{N,j}(z)$, then, using (4.4), it is straightforward to show that

$$\sum_{p=1}^{N} \frac{\alpha_{p,N}}{[1-2\sigma_{p,N}z+\sigma_{p,N}^2]^{1/2}} + \sum_{q=0}^{j} \beta_{q,N} P_q(z) \qquad (4.5)$$

where the branch of $(1-2\sigma_{p,N}z+\sigma_{p,N}^2)^{\frac{1}{2}}$ is chosen to be positive for $z<0$.

An alternative derivation of (4.5), not making use of (4.3), is given in (7). Consider the (N/N+j) Padé approximant to g(z), denoted by $g_{N,j}(z)$; then, from (4.2) and (4.4),

$$g_{N,j}(z) = \sum_{p=1}^{N} \frac{\alpha_{p,N}}{1-\sigma_{p,N}z} + \sum_{q=0}^{j} \beta_{q,N} z^q$$

$$= \sum_{p=1}^{N} \alpha_{p,N} \sum_{r=0}^{\infty} \sigma_{p,N}^r z^r + \sum_{q=0}^{j} \beta_{q,N} z^q$$

Then, from (4.1),

$$f_{N,j}(z) = \sum_{p=1}^{N} \alpha_{p,N} \sum_{r=0}^{\infty} \sigma_{p,N}^r P_r(z) + \sum_{q=0}^{j} \beta_{q,N} P_q(z)$$

$$= \sum_{p=1}^{N} \frac{\alpha_{p,N}}{[1-2\sigma_{p,N}z+\sigma_{p,N}^2]^{1/2}} - \sum_{q=0}^{j} \beta_{q,N} P_q(z)$$

which is (4.5). In this derivation we have made use of the fact that $(1-2tx+t^2)$ is the generating function for the Legendre polynomials:

$$\sum_{L=0}^{\infty} t^L P_L(x) = \frac{1}{\sqrt{1-2tx+t^2}}$$

The approximants defined by (4.5) are those introduced by Common. For the case when g(w) is a series of Stieltjes, the following theorem is true (5):

THEOREM 4.1: For all z in the complex plane cut from $\frac{1}{2}(r+\frac{1}{r})$ to ∞, where r is the radius of convergence of the series of Stieltjes (4.2) and r>1,

$$\lim_{N\to\infty} f_{N,j}(z) = f(z)$$

The convergence is uniform in any finite closed region of the complex z-plane which does not include any point of the cut of f(z).

5. EXTENSION TO QUADRATIC APPROXIMANTS

An important point to note from Theorem 4.1 is that, although the approximants defined by (4.5) appear to have a square root branch point built into them, the approximation scheme (4.5) only converges at points not on the cut of f(z). We (see also (8)) propose to replace g(w) in (4.3) by its quadratic approximant, $g_{(p/q/r)}(w)$, in an attempt to obtain convergence on the branch cut; we can obviously extend this idea by replacing g(w) by any 'suitable' approximant. However, once we choose to approximate g(w) in (4.3) by a multi-valued approximant, in preference to a Padé approximant, we have no longer any equation of the form (4.4), and we are not able to give an explicit representation (of the form (4.5)) for these multi-valued approximants. In practice, the integration in the formula

$$f_{(p/q/r)}(z) = \frac{1}{\pi}\int_0^{\pi} g_{(p/q/r)}\left(z + (z^2-1)^{1/2}\cos v\right) dv \qquad (5.1)$$

is performed numerically. The following method of performing this numerical integration is due mainly to T. Stacey (8), who performed most of the numerical computations associated with this method.

To illustrate the method, we again consider the logarithmic example of §3. The power series with the same coefficients as (3.1) will have a singularity at

$$z_s = \frac{1}{a} + \left(\frac{1}{a^2} - 1\right)^{1/2} = \frac{10}{3} \qquad (5.2)$$

Let

$$f(z) = \ln\frac{1-az}{1-a} = \sum_{L=0}^{\infty} f_L P_L(z) \qquad (5.3)$$

where the f_L are given by (3.1), and consider

$$g(z) = \sum_{L=0}^{\infty} f_L z^L \qquad (5.4)$$

The branch points of the (N/N/N) quadratic approximant to g(z) are then (in the usual notation) given by the zeros of the 2Nth. degree polynomial (Q^2-4PR). In Table 5 we tabulate these zeros for N=2 and N=5, and we find the following features. Apart from the (unstable) zeros at great distances

TABLE 5: LOCATION OF THE BRANCH POINTS OF THE (N/N/N) QUADRATIC APPROXIMANT (N=2,5) TO (5.4).

ZEROS OF THE (2/2/2) QUADRATIC APPROXIMANT	ZEROS OF THE (5/5/5) QUADRATIC APPROXIMANT
-202	-2212
-2.653+0.002i	-67
-2.653-0.002i	-65
3.32	-2.88829+1.20877i
	-2.88820+1.20874i
	-2.88829-1.20877i
	-2.88820-1.20874i
	0.6275521
	0.6275517
	3.33332

from the origin and those near z_s, all the zeros occur in very close pairs. Moreover, for all the examples so far studied, this occurence of very close pairs of zeros is always found and does seem to be a characteristic feature of this quadratic method of approximation. This is a very desirable feature since we can join these adjacent zeros by cuts and produce a "maximally analytic" approximant. For, if we draw cuts which do not connect nearest neighbour branch points, we produce an approximant containing discontinuities in large regions where the original function f(z) is analytic.

For z away from the branch cut the evaluation of $f_{(N/N/N)}(z)$ from (5.1) is straightforward; for z on the branch cut the situation is not so simple. The poles of the approximant all lie far from the origin or on what we might call the "unphysical" sheet of the approximant and present no problem, except for the pole close to z_s. The presence of this pole forces us to modify (5.1) to the form

$$f_{(p/q/r)}(z) = \frac{1}{i\pi}\int_{z-\sqrt{z^2-1}}^{z+\sqrt{z^2-1}} \frac{g_{(p/q/r)}(x)}{\sqrt{1-2zx+x^2}}\,dx \qquad (5.5)$$

which is obtained from (5.1) by making the substitution

$$x = z + (z^2-1)^{1/2}\cos v$$

The path of integration in (5.5) is chosen as a path in the complex plane which avoids the pole at z_s.

For z not on the branch cut, the results obtained using (5.1) are comparable to those produced by (4.5) (see (8) for further details). In Table 6 we give results, using (5.5), for z along the branch cut. These results indicate that we can obtain very good convergence a long way from the branch point. This method is clearly far superior to that of Fleischer (4), who attempts to use the residues of the poles of the Padé Legendre approximants of §1 to obtain information about the imaginary part on the cut.

6. THE INVERSE SQUARE AND COULOMB POTENTIALS

In this section we apply the method of §5 to the following two examples from potential scattering.

(i) INVERSE SQUARE POTENTIAL

TABLE 6: ERRORS IN THE (5/5/5) QUADRATIC APPROXIMANT OF 5 TO (3.1); THE BRANCH POINT OCCURS AT z=1.82.

z	ERROR IN THE (5/5/5) QUADRATIC APPROXIMANT	
	REAL PART	IMAGINARY PART
1.8	4×10^{-5}	2×10^{-6}
2.0	2×10^{-5}	9×10^{-6}
2.2	4×10^{-10}	2×10^{-5}
2.4	2×10^{-5}	3×10^{-6}
2.6	1×10^{-5}	1×10^{-5}
2.8	4×10^{-6}	2×10^{-5}
3.0	2×10^{-5}	1×10^{-5}
3.2	3×10^{-5}	3×10^{-6}

Having a long range, the repulsive inverse square potential

$$V(r) = \lambda r^{-2} \quad (\lambda > 0) \tag{6.1}$$

has many of the features of several potentials of physical interest; most notably, a very slowly convergent partial wave expansion of the scattering amplitude. Also, the corresponding phase shifts may be obtained exactly for this potential, allowing a comparison of the various approximation schemes. For the potential (6.1) we have (9)

$$f(\cos\theta) = k^{-1} \sum_{l=0}^{\infty} f_l P_l(\cos\theta)$$

where

$$f_l = (2l+1)\, e^{i\delta_l} \sin\delta_l \tag{6.2}$$

and

$$\delta_l = \frac{\pi}{2}\left[l + \frac{1}{2} - \sqrt{\left(l+\tfrac{1}{2}\right)^2 + \frac{\lambda}{\hbar^2}} \right]$$

($\hbar k$ is the momentum of the scattered particles which are of unit mass).

In Table 7 we compare the following approximants:

(a) the linear Padé Legendre approximant of Fleischer; the values quoted are those of (9),

(b) the diagonal Padé Legendre approximants defined by (4.5); the values quoted are those of (7),

(c) the diagonal quadratic Legendre approximants defined by (5.1).

The approximants in (a), (b) and (c) are denoted in Table 7 by (N/N), $f_{(N/N)}$ and $f_{(N/N/N)}$ respectively. In each case we only quote the approximant to the real part of (6.2) (the imaginary part giving similar results) for $\lambda = \hbar^2$ and k=10. The "exact" values quoted for f(cosθ) at the end of the table are those obtained by directly summing the partial wave series, using sufficient terms to ensure stability.

The main conclusion we can draw from Table 7 is that the schemes (b) and (c) give very similar results and are both appreciably better than scheme (a).

(ii) COULOMB POTENTIAL

The scattering amplitude for the Coulomb potential

TABLE 7: APPROXIMANTS TO THE REAL PART OF THE SCATTERING AMPLITUDE FOR THE INVERSE SQUARE POTENTIAL $V(r)=\lambda r^{-2}$.

NUMBER OF COEFFICIENTS	$\theta=4^\circ$			$\theta=18^\circ$			$\theta=58^\circ$		
	Re $f_{(N/N)}$	Re $f_{(N/N/N)}$	Re (N/N)	Re $f_{(N/N)}$	Re $f_{(N/N/N)}$	Re (N/N)	Re $f_{(N/N)}$	Re $f_{(N/N/N)}$	Re (N/N)
5	-1.96			-0.330			-0.03831		
7	-2.02		-1.26	-0.3373		-0.37	-0.03823		-0.03817
11	-2.059	-2.063		-0.3368	0.3368		-0.03824	-0.03824	
13	-2.063			-0.3368					-0.03824
14		-2.064			-0.3368			-0.03824	
25	-2.064		-2.12			-0.3368			
43			-2.066						
f(cosθ)		-2.064			-0.3368			-0.03824	

$$V(r) = \frac{e^2}{r}$$

has the exact form (10)

$$f_c(z) = \frac{-2^{i/k}}{k^2}(1-z)^{-1-i/k}\,\frac{\Gamma\left(1+\frac{i}{k}\right)}{\Gamma\left(1-\frac{i}{k}\right)}$$

where k is the momentum of the particles being scattered. The partial wave expansion is

$$f_c(z) = \sum_{L\,0}^{\infty}(2L+1)\,\frac{e^{2i\delta_L}}{2ik}\,P_L(z) \qquad (6.3)$$

where

$$e^{2i\delta_L} = \frac{\Gamma\left(L+1+\frac{i}{k}\right)}{\Gamma\left(L+1-\frac{i}{k}\right)}$$

The slow convergence encountered with (6.2) is even more extreme here; in fact, the partial wave expansion for $f_c(z)$ is divergent everywhere, and $f_c(z)$ also has a branch cut from +1 to ∞.

For $z<1$, the approximants (4.5) and (5.1) give comparable results for $f_c(z)$; further details of the former approximants can be found in (11). Along the branch cut ($z>1$) the approximants (4.5) do not converge whilst the approximants defined by (5.1) produce the results of Table 8. Again we only tabulate the approximants to the real part of $f_c(z)$ with k=10. The results show that very good convergence can be obtained a long way along the branch cut.

7. CONCLUSIONS

From the results presented in this chapter, we draw the following conclusions:

(i) From the linear Padé Legendre approximants of Fleischer we can readily define linear quadratic (and higher order) Legendre approximants; in practice, these quadratic approximants can be generated as easily as the corresponding Padé approximants. Numerically, the quadratic Legendre approximants exhibit slightly better convergence properties than the Padé Legendre approximants away from any branch cut of the function; more important, though, is the convergence of the quadratic Legendre approximants

TABLE 8: THE (5/5/5) QUADRATIC LEGENDRE APPROXIMANT (OF 5) TO THE REAL PART OF THE SCATTERING AMPLITUDE FOR THE COULOMB POTENTIAL $V(r)=e^2/r$ FOR $z=\cos\theta>1$.

z	(5/5/5) QUADRATIC LEGENDRE APPROXIMANT	EXACT VALUE
1.05	$(5.314\times10^{-2},-6.212\times10^{-2})$	$(5.317\times10^{-2},-6.211\times10^{-2})$
1.48	$(6.24\times10^{-4},-6.6787\times10^{-3})$	$(6.25\times10^{-4},-6.6784\times10^{-3})$
1.90	$(-4.04\times10^{-4},-3.5494\times10^{-3})$	$(-4.05\times10^{-4},-3.5498\times10^{-3})$
2.51	$(-6.005\times10^{-4},2.112\times10^{-3})$	$(-6.008\times10^{-4},-2.108\times10^{-3})$
3.37	$(-5.79\times10^{-4},-1.3367\times10^{-3})$	$(-5.81\times10^{-4},-1.3364\times10^{-3})$
4.56	$(-5.046\times10^{-4},-8.850\times10^{-4})$	$(-5.053\times10^{-4},-8.828\times10^{-4})$
6.21	$(-4.22\times10^{-4},-6.01\times10^{-4})$	$(-4.21\times10^{-4},-5.99\times10^{-4})$
8.48	$(-3.46\times10^{-4},-4.147\times10^{-4})$	$(-3.43\times10^{-4},-4.140\times10^{-4})$

along branch cuts of the function. These quadratic approximants therefore allow the calculation of scattering amplitudes (in the form of partial wave expansions) along branch cuts of the scattering amplitude.

(ii) The quadratic Legendre approximants defined by (5.1) seem to provide a powerful method for the analytic continuation of Legendre series. Away from branch cuts they compare favourably with the Legendre-Padé approximants of Common, and also they produce very good results along branch cuts of the function. The disadvantage of the method is that the required numerical integration is a little difficult to perform, requiring integration along a path in the complex plane. Furthermore, the prescription given in §5 for joining up the branch points of the approximant by cuts is seemingly arbitrary, although the choice ensuring maximal analyticity of the approximants seems a natural one.

REFERENCES

1. M.L. Goldberger and K.M. Watson, "Collision Theory", (New York, 1964).

2. J. Fleischer, Nud.Phys.B37,59 (1972).

3. J.T. Holdeman Jr., Math. Comput. 23,275 (1969).

4. J. Fleischer, J: Math.Phys. 14,246 (1973).

5. A.K. Common, Nuovo Cimento 63A,863 (1969).

6. T. Kinoshita, J.J. Loeffel and A. Martin, Phys. Rev.135,B1464 (1964).

7. A.K. Common and T. Stacey, "Legendre Padé approximants and their application in potential scattering", U.K.C. preprint (1977).

8. T. Stacey, Ph.D. Thesis (unpublished) (1977).

9. O.D. Corbella, C.R. Garibbotti and F.F. Grinstein, "Rational Polynomial Approximants and scattering amplitudes for long range potentials", Universidad Nacional de Rosario (Argentina), preprint.

10. L.D. Landau and E.M. Lifshitz, "Quantum Mechanics" (Oxford, Pergamon Press, 1958).

11. A.K. Common and T. Stacey, "The convergence of Legendre Padé approximants to the Coulomb and other scattering amplitudes", U.K.C. preprint (1977).